

Encyclopaedia

Author and Editor

Elzein Salah Elzein

IEEE Member

- Papers in Computation

IEEE Publisher, UK



Article

Quantifying the Health–Economy Trade-Offs: Mathematical Model of COVID-19 Pandemic Dynamics

Dhika Surya Pangestu ¹, Sukono ^{2,*} , Nursanti Anggriani ²  and Najib Majdi Yaacob ³ 

¹ Mathematics Postgraduate Program, Universitas Padjadjaran, Sumedang 45363, West Java, Indonesia; dhika17001@mail.unpad.ac.id

² Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Sumedang 45363, West Java, Indonesia; nursanti.anggriani@unpad.ac.id

³ Biostatistics and Research Methodology Unit, School of Medical Sciences, Universiti Sains Malaysia Health Campus, Kubang Kerian 16150, Kelantan, Malaysia; najibmy@usm.my

* Correspondence: sukono@unpad.ac.id

Abstract: The COVID-19 pandemic has presented a complex situation that requires a balance between control measures like lockdowns and easing restrictions. Control measures can limit the spread of the virus but can also cause economic and social issues. Easing restrictions can support economic recovery but may increase the risk of virus transmission. Mathematical approaches can help address these trade-offs by modeling the interactions between factors such as virus transmission rates, public health interventions, and economic and social impacts. A study using a susceptible-infected-susceptible (SIS) model with modified discrete time was conducted to determine the cost of handling COVID-19. The results showed that, without government intervention, the number of patients rejected by health facilities and the cost of handling a pandemic increased significantly. Lockdown intervention provided the least number of rejected patients compared to social distancing, but the costs of handling the pandemic in the lockdown scenario remained higher than those of social distancing. This research demonstrates that mathematical approaches can help identify critical junctures in a pandemic, such as limited health system capacity or high transmission rates, that require rapid response and appropriate action. By using mathematical analysis, decision-makers can develop more effective and responsive strategies, considering the various factors involved in the virus's spread and its impact on society and the economy.

Keywords: pandemic; treatment costs; epidemiological models; trade-off models



Citation: Pangestu, D.S.; Sukono; Anggriani, N.; Yaacob, N.M. Quantifying the Health–Economy Trade-Offs: Mathematical Model of COVID-19 Pandemic Dynamics. *Computation* **2024**, *12*, 139. <https://doi.org/10.3390/computation12070139>

Academic Editors: Babak Shiri and Zahra Alijani

Received: 11 May 2024

Revised: 21 June 2024

Accepted: 28 June 2024

Published: 8 July 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The history of pandemics has recorded several major pandemics that affected millions of people around the world. Some of the important pandemics that occurred throughout history include the Bubonic Pandemic in 1331–1353. This pandemic spread worldwide and killed more than 75 million people, or about 30% of the world's population at the time [1]. The Spanish pandemic of 1918–1920, caused by the influenza virus, killed more than 50 million people worldwide [2]. The HIV / AIDS pandemic from 1981 to the present started in Africa and has killed millions of people worldwide [3]. The SARS pandemic began in 2002–2004. The SARS virus was first discovered in China and spread worldwide, killing more than 700 people [4]. The COVID-19 pandemic in 2019 until now was caused by the SARS-CoV-2 virus and is still happening today, killing millions of people worldwide [5]. These are just a few examples of major pandemics that have occurred throughout history, and they show how important it is to understand and prepare for future pandemics. The response to a pandemic varies depending on the pandemic that is occurring and which countries are affected. Generally, there are several steps that governments and health organizations take to cope with a pandemic, such as vaccination, isolation and quarantine, testing and contact tracing, lockdown, education, and economic assistance [6]. Responses to pandemics must be tailored to the conditions and characteristics of the current pandemic

and often require adaptation and updating over time [7]. Public health, economic, and human rights considerations must be taken into account when making decisions and actions to address the pandemic.

Financing systems during a pandemic play an important role in addressing the economic and social consequences of the pandemic [8]. There are several ways to finance the actions and programs needed to overcome the pandemic, including the government: Governments often use sources of funds such as taxes and debt to finance government programs and economic assistance for individuals and businesses affected by the pandemic. Donations: Donations from individuals, institutions, and companies can help finance the health and social programs needed during a pandemic. Loans: Loans from banks and other financial institutions can help finance government and private programs and measures to address the pandemic. International aid: Aid from other countries and international institutions can help finance government and private programs and actions to overcome the pandemic [9]. Financing systems during a pandemic should consider the balance between financing the actions needed to address the pandemic and ensuring that the financial burden does not negatively impact society and the economy in the long term [10]. Financing policies and programs during a pandemic should be tailored to the conditions and characteristics of the current pandemic and often require adaptation and updating over time.

Multiple disciplines play an important role in managing the pandemic [11]. The role of mathematical models of epidemiology in managing pandemics is to provide insights and predictions about the spread and impact of a disease [12]. These models can help decision-makers make informed choices about how to allocate resources and implement interventions to control the spread of the disease. By simulating different scenarios, these models can provide a more accurate understanding of the potential impact of different intervention strategies, including the potential advantages and disadvantages of each strategy. This information can then be used to guide public health policy and resource allocation, as well as prioritize interventions that are most likely to be effective in controlling the spread of disease [13]. Ultimately, the use of mathematical models of epidemiology can play an important role in reducing the number of infections and deaths during a pandemic and in controlling the spread of disease as quickly and effectively as possible [14].

There have been several studies that have tried to estimate the health expenditure needed to overcome COVID-19. Research conducted by Dudine et al. (2020) focuses on health spending at the country level. This health expenditure is modeled using a discrete stochastic susceptible-infected-recovered (SIR) model approach [15]. People's health expenditure is then estimated by multiplying the number of infected people who require treatment in health facilities. Other studies have focused on health expenditures in low-income countries [16]. Torres Rueda et al. (2021) estimated health expenditure in low-income countries using the susceptible-exposed-infected-recovered (SEIR) model. The SEIR model was built using a discrete approach, and health expenditure was estimated by the costs required for emergency response, testing, and tracking positive cases, as well as the costs required to treat active and deceased cases of COVID-19 [16]. Another study tried to investigate the potential impact of non-pharmaceutical interventions, such as social activity restrictions, on capital accumulation and economic development at different time scales [17]. The research conducted by Cook et al. (2020) focuses on the potential impact of non-pharmaceutical interventions on model accumulation and economic development at different time scales. This study integrates the SOLOW-type economy model with the SIR model to model the potential impact of economic interventions carried out [18]. In addition, this epidemiological model approach is also used to investigate the impact of the state's inability to implement social activity restrictions on low-income households [18]. Another study used an epidemiological approach to estimate the effectiveness of COVID-19 vaccination in reducing health expenditure [19]. Another study examined the implications of premature death and death from infection on health and the economy using the SEIR model [20].

In previous studies on modeling the cost of handling a pandemic, the effect of government intervention on the economy was not discussed. Also, models that examine the

effect of government intervention on the economy never discuss the health financing that occurs. Even though these two factors can affect the total expenditure for handling a pandemic, in this study, a model is created to estimate the health costs needed to deal with the COVID-19 pandemic in a country. The model created takes into account the possibility of changes in the severity of COVID-19 patients that have not been applied by Dudine et al. (2020) [15]. This is more realistic to do, considering the severity of COVID-19 patients can, indeed, change in the field. To do this, a discrete susceptible-infectious-susceptible (SIS) epidemiological mathematical model with interventions is used to project the number of people who need hospital care over time.

In this study, mathematical modeling is carried out to estimate the costs required for handling a pandemic, along with the trade-offs. This study comprehensively presents how much it costs, estimating the possible changes in mortality and also changes in economic costs for each policy required. The SIS model is used because of its proximity to the phenomenon that occurs, namely the possibility of patients who have been declared cured of COVID-19 and are re-infected. The model also takes into account the influence of interventions made by the government. So that each intervention step taken can have a different effect, be it lockdown, quarantine, or vaccination. Then, for each step taken, its effect on costs and potential trade-offs between reducing the number of cases and minimizing handling costs are also considered. Hence, the model designed in this study not only assesses the handling of pandemic cases in terms of the number of cases and the rate of spread but also its effect on economic conditions. This is then converted into health costs. The model provides information on the cost of handling the infected population, the cost of intervention, and the trade-off between the two. Thus, the information generated by this model is comprehensive for government policymaking. The information provided by the model ultimately allocates funds for population management as well as interventions to contain the spread of the disease. Since this study is evaluative in nature, the allocations generated by this research can then be compared with the expenditures that have been made by the government and then further evaluated. It is hoped that the information provided can be an input for the government to take a mature policy in the event of a pandemic again so that the focus of the interventions carried out is not only to reduce the number in the short term but also save the population as a whole for a longer period.

2. Overview

2.1. Epidemiological Mathematical Model

Mathematical models of epidemiology are tools for predicting and understanding the spread of disease in populations [21,22]. These models include the relationship between people who are infected, people who are susceptible to infection, and people who have recovered or are immune to the disease. Mathematical models of epidemiology can help researchers and policymakers predict the spread of disease, evaluate the effectiveness of prevention and treatment strategies, and plan appropriate public health responses [23,24]. Mathematical models of epidemiology are usually based on systems of differential equations that describe changes in the number of people in different groups over time. For example, the SIR (susceptible-infected-recovered) model is one of the most commonly used mathematical models of epidemiology. The SIR model assumes that the population is divided into three groups: susceptible people (S), infected people (I), and people who have recovered or are immune to the disease (R). The differential equation used for the SIR model is as follows,

$$N(t) = S(t) + I(t) + R(t) \quad (1)$$

Basic SIR model

1. Rate of change of susceptible to infected per unit time

$$\frac{dS}{dt} = -\beta SI \quad (2)$$

2. Rate of change of infected to recovered per unit time

$$\frac{dI}{dt} = \beta SI - \alpha I \quad (3)$$

3. Recovered cure rate per unit time

$$\frac{dR}{dt} = \alpha I \quad (4)$$

where β is the transmission rate, α is the cure or death rate, and t is time. The study by Wearing et al. (2005) provides an in-depth explanation of epidemiological mathematical models, including the SIR model and other models used in epidemiology [25]. They discuss applications of epidemiological mathematical models in predicting the spread of infectious diseases such as flu, measles, and HIV, as well as estimating the impact of vaccinations and other preventive measures. They also discuss numerical techniques for solving differential equations and the use of empirical data in validating the model.

In addition to the SIR model, there is also the SIS (susceptible-infected-susceptible) model. In the SIS model, it is assumed that the population that has been infected does not have long-term immunity, and individuals become susceptible again, so this allows reinfection. The differential equation used for the SIS model is as follows,

$$N(t) = S(t) + I(t) \quad (5)$$

Besides the SIR and SIS models, there are also other epidemiological mathematical models, such as the SEIR (susceptible-exposed-infected-recovered) model and the more complex compartmental model. However, all mathematical models of epidemiology have the same goal, which is to help in understanding and controlling the spread of disease in the population.

2.2. Discrete Epidemiological Mathematical Model

Discrete epidemiological mathematical models are used to model the spread of disease in discrete populations. These models are often used in epidemiology to predict the spread of disease among individuals in discrete populations. Some examples of commonly used discrete epidemiological mathematical models are the SIR (susceptible-infected-recovered) model and the SEIR (susceptible-exposed-infected-recovered) model. In both models, individuals in the population are categorized into classes based on their status in terms of infection. In epidemiological modeling, most models developed generally use a continuous approach [26]; this is because, mathematically, continuous models are easier to solve. However, in reality, existing epidemiological data are generally collected at discrete time intervals. Thus, this makes discrete models more appropriate to use when considering the way data are collected. In addition, discrete epidemic models are much more intuitive and easy to implement by non-mathematicians due to the similarity in the way data are recorded. This makes discrete models easier to accept and socialize for non-mathematicians. So far, there have been several studies using discrete disease transmission models [27–29]. For the simplest discrete epidemic models, Equations (6)–(8) are the models used. So, in discrete form, the equation becomes as follows,

$$S_{t+1} = S_t - \beta S_t I_t \quad (6)$$

$$I_{t+1} = I_t + \beta S_t I_t - \alpha I_t \quad (7)$$

$$R_{t+1} = R_t + \alpha I_t \quad (8)$$

where t is time, α is the cure rate, β is the infection rate, and N is the total population.

2.3. Markov Chain

Markov chains are widely used to estimate dynamic developments involving transitions between observable finite state spaces. The methodology is intuitive and easy to apply, especially for situations where different stages determine the state of observation in addition to the development of the situation. Consequently, this idea is widely recognized in studies and applications in epidemiology [30]. The discrete Markov chain asserts that the present state of the observation is totally dependent on the prior time step's state at the discrete observation interval. Mathematically, this means that the probability of the observed state i at time t transitions to state j at time $t + 1$ is denoted as follows,

$$p_{i,j}(t) = \Pr[I(t+1) = j \mid I(t) = i] \quad (9)$$

where $I(t)$ is the identification of the observed state at time t . The one-step transition probability is defined as $p_{i,j}(t)$, or it can also be referred to as the daily transition probability. For the transition probability, the t -step forward along the observations is denoted as $p_{i,j}(t)$.

2.4. Mean Absolute Percentage Error (MAPE)

In a prediction, the results of the model need to be evaluated to ensure the accuracy of the model in making predictions. So far, there have been various measures used, and the most common measures to use are absolute percentage error (APE) and mean absolute percentage error (MAPE); the formulas for calculating APE and MAPE are as follows,

$$APE = \frac{|X_t - \hat{X}_t|}{X_t}, \quad (10)$$

$$MAPE = \frac{\sum_{t=1}^n \frac{|X_t - \hat{X}_t|}{X_t}}{n} \times 100\%, \quad (11)$$

The performance of forecasting falls into the good accuracy category when the APE and MAPE values produced are less than 20% [31].

3. Materials and Methods

3.1. Materials

The object of the model in this study is a model for determining the cost of handling a pandemic. This model uniquely involves the components of the cost of handling the infected population, the cost of intervention, and the trade-off between the two. The costs of handling infected populations involved in this study are treatment costs, the cost of adding health facilities, mortality costs, and tracking costs. The costs of government intervention involved in this study are social-distancing costs and lockdown costs. Finally, trade-offs are involved to provide a better picture of the total expenditure for pandemic management. The cost of handling a pandemic is modeled using a discrete SIS model with interventions and demographics, and then the projection results from the model are combined with an ingredient-based cost model. The demographic aspect is considered because, in its journey in Indonesia, COVID-19 has been running for approximately 3 years, so to increase the accuracy of the model, the demographic aspect is considered.

The data used in this study are the number of COVID-19 cases in West Java. The cases used in this study are positive, recovered, died, and active. The data used are from the period March 2020–August 2022 in the COVID-19 pandemic. The dataset contains 3632 entries, with 908 entries for each case [32]. Then, for health facilities, the data used are the Indonesian bed occupancy rate (BOR) data. The data were obtained from the COVID-19 Task Force. This dataset is from the period March 2020–August 2022 [33]. For data on the cost of government interventions, this study used the official WHO guidelines for policies and response measures for COVID-19 to determine the set of interventions involved in a country.

Before being used, the data were checked for quality. For COVID-19 case data, the data were checked for consistency by comparing the number of cumulative cases between

t and $t + 1$. Cumulative case data are said to be consistent when the number of cases in $t + 1 \geq$ the number of cases in t . In addition, for active cases, the data are checked by comparing with existing data from the source and the data calculated by researchers. If an error or difference is found, it will be reconciled with the data source before further analysis is carried out.

3.2. Methods

In this research, two stages are carried out, namely the model development stage and the numerical illustration analysis stage, which are described as follows: First, the model development stage is carried out by examining in depth, especially the research conducted by Dudine et al. [15], which discusses the formulation of a COVID-19 costing model using the SIR model. The existing model was then modified in this study by adding the possibility of a transition in the severity of symptoms in patients infected with COVID-19. The model formulation for determining the cost of handling COVID-19 is carried out using a discrete-time transition opportunity approach. Second, the stage of numerical illustration analysis is carried out to show how the application of the model has been made in determining the cost of handling COVID-19 and to describe how the population dynamics in the COVID-19 pandemic occurred based on the model created. For the analysis of the numerical illustration here, the parameters used in the model are estimated based on the existing data. The flowchart for this research is presented in Figure 1.

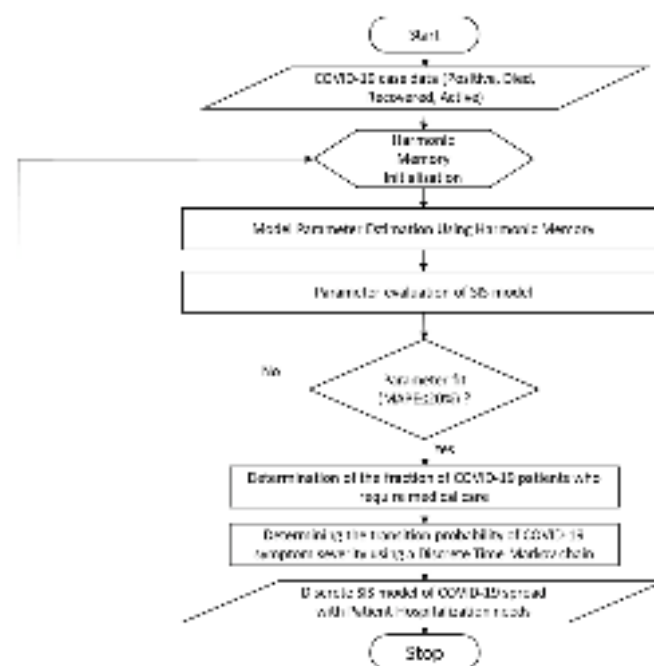


Figure 1. Modeling steps.

Modeling population dynamics of COVID-19 spread with patient hospitalization needs is conducted by first inputting the available pandemic case data. In this case, the available data are COVID-19 data. The data are then used to estimate the parameters of the SIS model formed in this study. The parameters used for this discrete SIS modeling include: The number of people assumed to have immunity at $t = 0$ is 0. The duration of infection is $\frac{1}{k}$. Because there is a transition opportunity, the duration of infection can vary so that, $\frac{1}{k_H} = \frac{1}{(1-t_{pr})}$, $\frac{1}{k_{RNH}} = \frac{1}{(1-t_{pr2})}$, and $\frac{1}{k_{NH}} = \frac{1}{(1-t_{rp})}$. Treated and non-infectable fractions are α . Parameter estimation was performed iteratively, and iteration was stopped when the parameters were fit. Parameters were considered fit when the resulting error size was minimum at a predetermined number of iterations. The error criterion used is the mean absolute percentage error (MAPE). After the parameters are fit, the fraction of COVID-19 patients who need medical care is determined, and the probability of transition

of the severity of symptoms in COVID-19 patients is estimated. The model also takes into account the impact of different interventions and control measures on the spread of the virus. By adjusting parameters, such as transmission rate and recovery rate, the model can simulate various scenarios to inform public health decision-making. Overall, this approach provides a valuable tool for predicting the trajectory of the pandemic and evaluating potential strategies to mitigate its impact.

4. Results

This section describes the results obtained in this study. Two results were obtained in this study, namely modeling results and numerical simulation results.

4.1. Modelling Result

In this study, a COVID-19 spread model is formed that considers the possibility of a transition in the severity of symptoms in infected patients, as well as the possibility of reinfection. The model formed has three compartments, namely susceptible (S), infected (I), and susceptible (S). The epidemiological mathematical model that has been created is used to model the COVID-19 pandemic in various scenarios.

In Equation (5), where $S(t)$ is the number of susceptible populations at time t and $I(t)$ is the number of infected populations at time t . The schematic diagram of the spread of COVID-19 referred to in this study is presented in Figure 2.

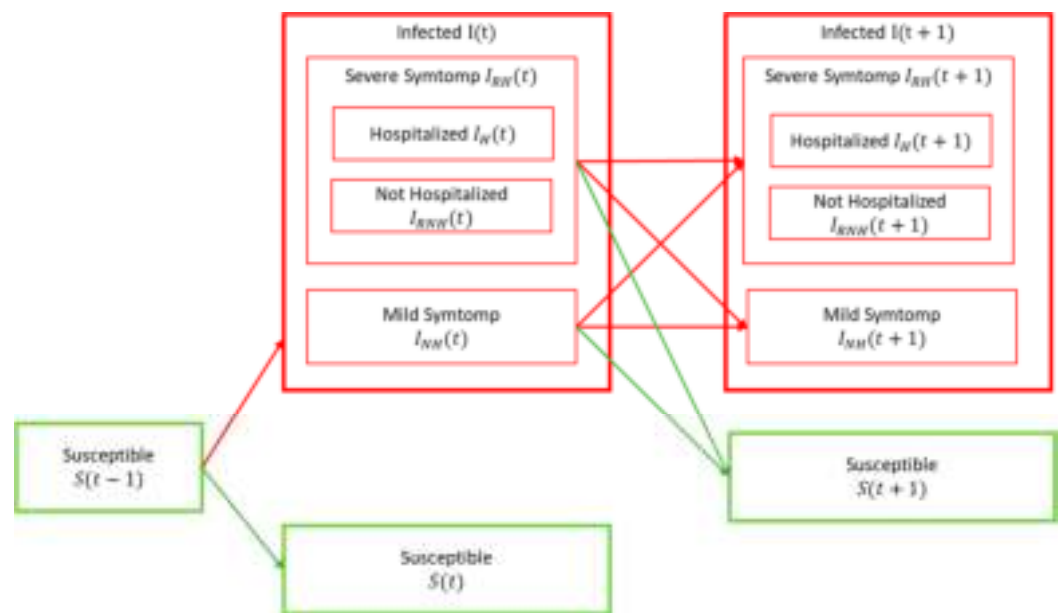


Figure 2. COVID-19 patient dynamics.

The explanation of Figure 2 is as follows:

- A person who is susceptible in period $t - 1$; $S(t - 1)$ can be infected in the next period by COVID-19 into the infected $I(t)$ group or remain susceptible $S(t)$.
- An infected person $I(t)$ can have mild symptoms $I_{NH}(t)$ or severe symptoms $I_{RH}(t)$. For $I_{NH}(t)$, the patient is assumed not to require treatment at a health facility such as a hospital. Whereas for $I_{RH}(t)$, the patient is assumed to require medical treatment at a health facility.
- Due to the limited capacity of health facilities, not all patients with severe symptoms of $I_{RH}(t)$ can receive treatment at health facilities. This results in a treated severe symptom group $I_H(t)$, and an untreated severe symptom group $I_{RNH}(t)$.
- Transitions in symptom severity may occur between $I_{RH}(t)$ and $I_{NH}(t)$. Since $I_{RH}(t)$ consists of $I_H(t)$ and $I_{RNH}(t)$, symptom severity transitions are described in Figure 3.

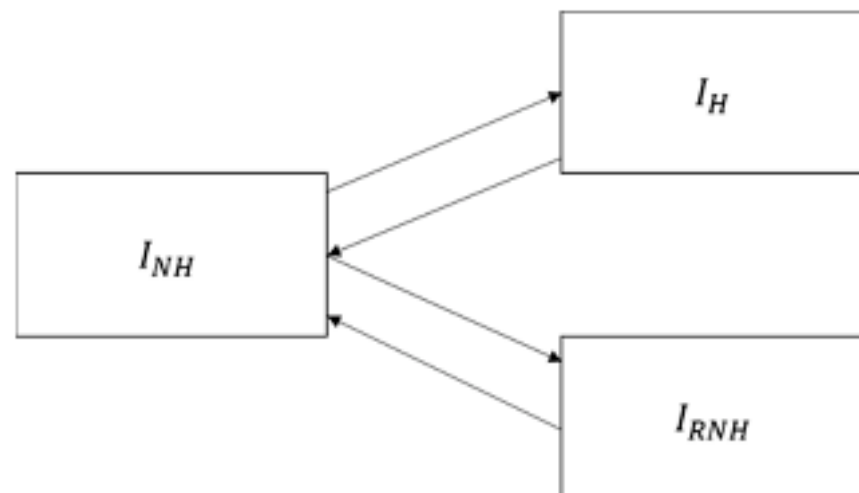


Figure 3. Symptom severity transition diagram.

Based on this explanation and Figure 3, it is known that compartment I is divided into three parts, namely:

- Does not require treatment in a health facility (I_{NH}). It is assumed that in each time period, only a portion of the members of compartment I require treatment so that the rest can self-isolate without needing to be treated in health facilities such as hospitals. For groups that are positive for COVID-19 but do not require intensive care in a hospital, the notation I_{NH} is used.
- Needed treatment at a health facility but did not get it, I_{RNH} . It is assumed that there are health facilities in each area, but their capacity is limited, resulting in the possibility of patients who need treatment but cannot be treated because of this problem. For groups that are positive for COVID-19 and require intensive care but do not receive it, the notation I_{RNH} is used. Equation (12) describes I_{RNH} .

$$I_{RNH}(t) = |\min\{0, B(t) - I_{RH}(t)\}|. \quad (12)$$

- Hospitalized I_H . This is a group that needs medical care and receives it. $I_H(t)$ is shown in Equation (13).

$$I_H(t) = \min\{I_{RH}(t), B(t)\}, \quad (13)$$

and at any time t , Equation (14) applies.

$$I(t) = I_H(t) + I_{RNH}(t) + I_{NH}(t), \quad (14)$$

Next, the capacity constraint in health facilities is modeled. The number of beds for treating COVID-19 patients at the start of the pandemic $B(0)$ is shown in Equation (15).

$$B(0) = v \cdot B_0, \quad (15)$$

where v is the fraction of COVID-19 treatment beds at the start of the pandemic, and B_0 is the total number of hospital beds. Over time, the number of COVID-19 patients will increase, which presents a situation where the number of existing hospital capacities needs to be increased. At any time t , if the number of beds $B(t-1)$ is less than the number of $I_{RH}(t-1)$, then the bed capacity will be increased based on $I_{RNH}(t-1)$. The existence of capacity constraints makes the addition of beds only possible for n fractions of $I_{RNH}(t-1)$, so the number of available bed capacities at time t is shown by Equation (16).

$$B(t) = \min\{B(t-1) + nI_{RNH}(t-1), B(0) + (b-1) \cdot B_0\}, \quad (16)$$

for each additional bed, it is assumed that the ratio between the number of beds with nurses and doctors remains the same, as in Equations (17) and (18).

$$\frac{T_{P0}}{(1-v)B_0} = \frac{T_P(t)}{B(t)}, \quad (17)$$

$$\frac{T_{D0}}{(1-v)B_0} = \frac{T_D(t)}{B(t)}, \quad (18)$$

where T_P is the number of nurses, and T_D is the number of doctors. T_{P0} and T_{D0} are the numbers of nurses and doctors in period 0 or at the beginning of the pandemic. Furthermore, Figure 3 has previously shown the dynamics of COVID-19 with the number of patients requiring medical care. It is assumed that the patient's chances of recovery and death depend on the severity of the symptoms and whether the patient receives treatment. This is shown by Equations (19) and (20).

$$d_{NH} < d_H < d_{RNH}, \quad (19)$$

$$r_{NH} > r_H > r_{RNH}, \quad (20)$$

where d is the probability of death, and r is the chance of recovery of the COVID-19 patient. After that, let k denote the probability that the patient will leave compartment I . The probability of this happening is if the patient recovers or dies. The k relations for I_{NH} , I_H , and I_{RNH} are found in Equations (21)–(23).

$$k_{NH} = d_{NH} + r_{NH}, \quad (21)$$

$$k_H = d_H + r_H, \quad (22)$$

$$k_{RNH} = d_{RNH} + r_{RNH}. \quad (23)$$

Patients can experience changes in the severity of symptoms, and this keeps the patient in compartment I , but only the severity of the symptom changes. If p denotes the probability of transition, then the probability that no transition occurs or the patient leaves compartment I is given by Equation (24).

$$1 - p = k, \quad (24)$$

and this applies to I_H , I_{RNH} , and I_{NH} , if Equation (24) is substituted with Equations (21)–(23), then Equations (25)–(27) are as follows,

$$p_b = 1 - k_{NH}, \quad (25)$$

$$p_r = 1 - k_H, \quad (26)$$

$$p_{r2} = 1 - k_{RNH}, \quad (27)$$

The relationship between the transition probability p and the probability of a patient exiting compartment I is depicted in Figure 4.

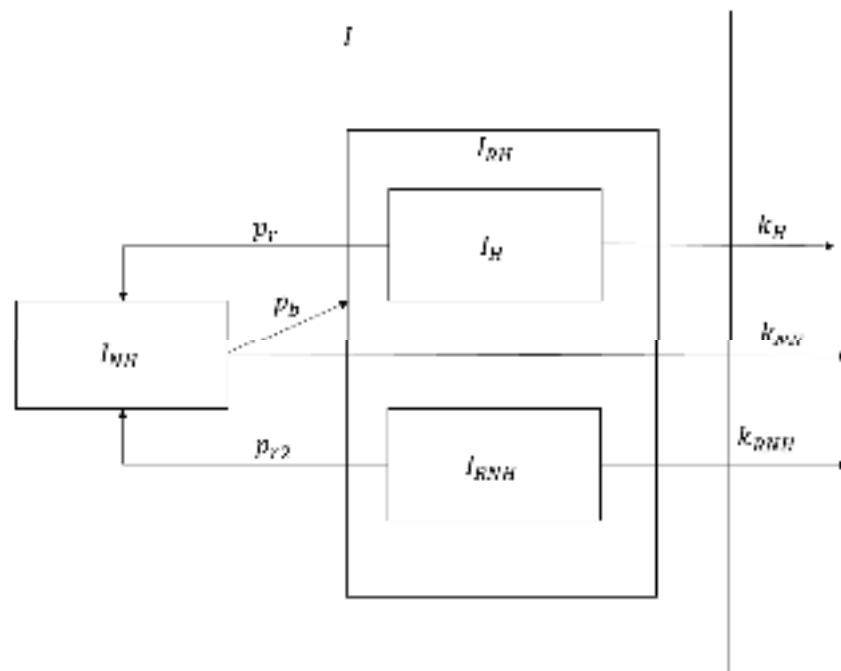


Figure 4. Diagram p and k in compartment I .

Furthermore, due to possible transitions in symptom severity, patients infected with COVID-19 may have different treatment durations. Therefore, k needs to be considered from time to time, as in Equation (28).

$$\begin{aligned} k(t) &= \frac{\Delta(S(t-1))}{I(t-1)} \\ &= \frac{r_H I_H(t-1) + r_{RNH} I_{RNH}(t-1) + r_{NH} I_{NH}(t-1) + d_H I_H(t-1) + d_{RNH} I_{RNH}(t-1) + d_{NH} I_{NH}(t-1)}{I(t-1)} \\ &= k_H \frac{I_H(t-1)}{I(t-1)} + k_{RNH} \frac{I_{RNH}(t-1)}{I(t-1)} + k_{NH} \frac{I_{NH}(t-1)}{I(t-1)}. \end{aligned} \quad (28)$$

Next, the number of patients who can no longer infect others corresponds to the number who died and recovered from COVID-19. Hence,

$$\Delta S(t) = r_H I_H(t-1) + r_{RNH} I_{RNH}(t-1) + r_{NH} I_{NH}(t-1) + d_H I_H(t-1) + d_{RNH} I_{RNH}(t-1) + d_{NH} I_{NH}(t-1) \quad (29)$$

$$\Delta S(t) = k(t) I(t-1). \quad (30)$$

Now suppose $\beta(t)$ is the average number of people infected by one person per unit time t , and $R_0(t)$ is the number of people infected by an infectious person during the course of his infection. Thus, it becomes Equation (31).

$$R_0(t) = \frac{\beta(t)}{k(t)}, \quad (31)$$

Based on the explanation that has been made, the number of new infections each time t is given in Equation (32).

$$\Delta I(t) = (I(t-1) - \alpha I_H(t-1)) \cdot \frac{S(t-1)}{N(t)} \cdot (\beta(t) - \sum_1^n i_n(t)), \quad (32)$$

$\Delta I(t)$ is always positive, $\beta(t) - \sum_1^n i_n(t)$, where i is the type of government intervention, and α is the fraction of the infected population that is properly quarantined, so that for the compartments S, I, S formed, Equations (33)–(37) apply.

$$S(t) = S(t-1) + \Delta S(t) - \Delta I(t), \quad (33)$$

$$I(t) = (1 - k(t))I(t-1) + \Delta I(t), \quad (34)$$

$$I_{RH}(t) = z \cdot I(t) + p_b(t)I_{NH}(t-1), \quad (35)$$

$$I_{RNH}(t) = |\min\{0, B(t) - I_{RH}(t)\}|, \quad (36)$$

$$I_H(t) = \min\{I_{RH}(t), B(t)\}, \quad (37)$$

After modeling the spread of COVID-19 and the number of patients who need treatment, modeling is carried out regarding the cost of handling the COVID-19 pandemic. In this study, the cost of handling the COVID-19 pandemic considers three types of financing. The first type is the cost of care for patients who experience severe symptoms of $I_H(t)$. Meanwhile, the second type is the cost of additional health facilities for patients with severe symptoms who are rejected on the grounds that the existing health facilities are full $I_{RNH}(t)$. Then, the last is the cost of interventions carried out during the time span t . In this study, the cost of treating COVID-19 patient C_p is formulated in Equation (38),

$$C_p(t) = c_p T_p(t) + c_d T_d(t) + c_m I_H(t), \quad (38)$$

where c_p is the cost of paying nurses, c_d is the cost of paying doctors, and c_m is the cost of administration and medicines needed to treat COVID-19 patients within 1 week. Meanwhile, the cost of adding health facilities C_b is formulated in Equation (39),

$$C_b = c_b \cdot \max\left\{0, \max_t\{B(t) - B_0\}\right\} + c_f \left\lfloor \frac{\max\left\{0, \max_t\{B(t) - B_0\}\right\}}{100} \right\rfloor, \quad (39)$$

where $\lfloor \cdot \rfloor$ is the floor function, c_f is the cost of building a health facility, c_b is the cost of adding one new bed, and $B(t)$ is the number of COVID-19 treatment beds available at time t . Furthermore, the cost of intervention is calculated using Equation (40).

$$CI = C_1 t_1 + C_2 t_2 + C_3 t_3 + \dots + C_n t_n$$

$$CI = \sum_{i=1}^n C_i t_i \quad (40)$$

There are various interventions that can be conducted in one time interval. Interventions can also be carried out simultaneously between intervention 1, intervention 2, intervention 3, and so on. Thus, the cost of intervention is calculated using Equation (40). Based on Equations (38)–(40), the total cost of handling COVID-19 patient C_T needs to be prepared in Equation (41),

$$C_T = t \cdot C_p(t) + C_b + CI. \quad (41)$$

4.2. Parameter Estimation Results

In this study, the parameters used were estimated using real data to provide simulation results closest to the actual phenomena in the field. There are 11 parameters used in this study. Ten are estimated parameters, and one is assumed because no data are available. In this study, the assumed parameter is parameter n , which is the bed fraction that can be added within the next week. In this study, we used the value of $n = 0.05$. We use this value by considering the time interval used; one week is a short time, meaning that even though the government has funds, there is still a time limit that can limit the number of beds that can be added. Furthermore, for other estimated parameters, we use real data

related to the definition of these parameters. All parameter estimation results in this study are summarized in Table 1.

- Parameter v is estimated using data on the number of West Java hospital beds in 2019.
- The d_{NH} parameter was estimated using daily COVID-19 mortality data. The d_H , d_{RNH} parameters are calibrated from the d_{NH} parameters using deaths that occurred during quarantine.
- The r_{NH} parameter was estimated using daily COVID-19 recovery data. For parameters r_H , r_{RNH} was calibrated from parameter r_{NH} using cures that occurred during quarantine.
- Parameter k is estimated using parameters d and r using Equations (21)–(23).
- Parameter p is estimated using parameter k and Equations (25)–(27).

Table 1. Parameter estimation results.

Parameter	Definition	Value	Details
n	Expandable bed fraction	0.05	Assumed
v	Bed fraction available at the start of the pandemic	0.1481	Estimated
d_H	Probability of death of patients with severe symptoms who are treated in health facilities	0.0011	Estimated
d_{RNH}	Probability of death of patients with severe symptoms who are not treated at a health facility	0.0017	Estimated
d_{NH}	Probability of death of a patient who does not require treatment at a health facility	0.0006	Estimated
r_H	Probability of recovery for patients with severe symptoms who are treated in health facilities	0.0654	Estimated
r_{RNH}	Probability of recovery for patients with severe symptoms who are not treated at health facilities	0.0327	Estimated
r_{NH}	Probability of recovery for patients who do not need treatment at a health facility	0.0981	Estimated
k	Probability of the patient leaving the compartment I	0.0665	Estimated
p_b	Probability of transition from mild to severe symptoms	0.0095	Estimated
p_r	Probability of transition from severe to mild symptoms	0.0221	Estimated

The parameters reported in Table 1 are then employed in this work for numerical simulations of the COVID-19 spread model. The parameter estimation findings are validated before being utilized in numerical simulations, assessing the model's correctness against the data used. The validation process ensures that the model accurately reflects real-world scenarios and can provide reliable predictions for the spread of COVID-19. This rigorous approach enhances the credibility and usefulness of the simulation results in guiding public health interventions and decision-making. This is accomplished graphically by showing a simulation of the model created using the parameters listed in Table 1. Figures 5 and 6 depict a plot of model predictions versus data from COVID-19 cases in West Java.

Figure 5 compares model predictions and real data on vulnerable or susceptible (S) populations. Based on the graph, it is known that the parameters used in this study are suitable for use in numerical simulations, as evidenced by the identical shapes of the model prediction plots and real data. The graph indicates that the model's predictions are accurate compared to the original data, and the parameters are fit for numerical simulations. The close alignment between the model predictions and actual data suggests that the model is effective in capturing the dynamics of COVID-19 cases in West Java. This validation of the parameters used enhances the reliability of the numerical simulations conducted in this study. Figure 6 compares the predictions of the model with the real data on the infected (I) population.

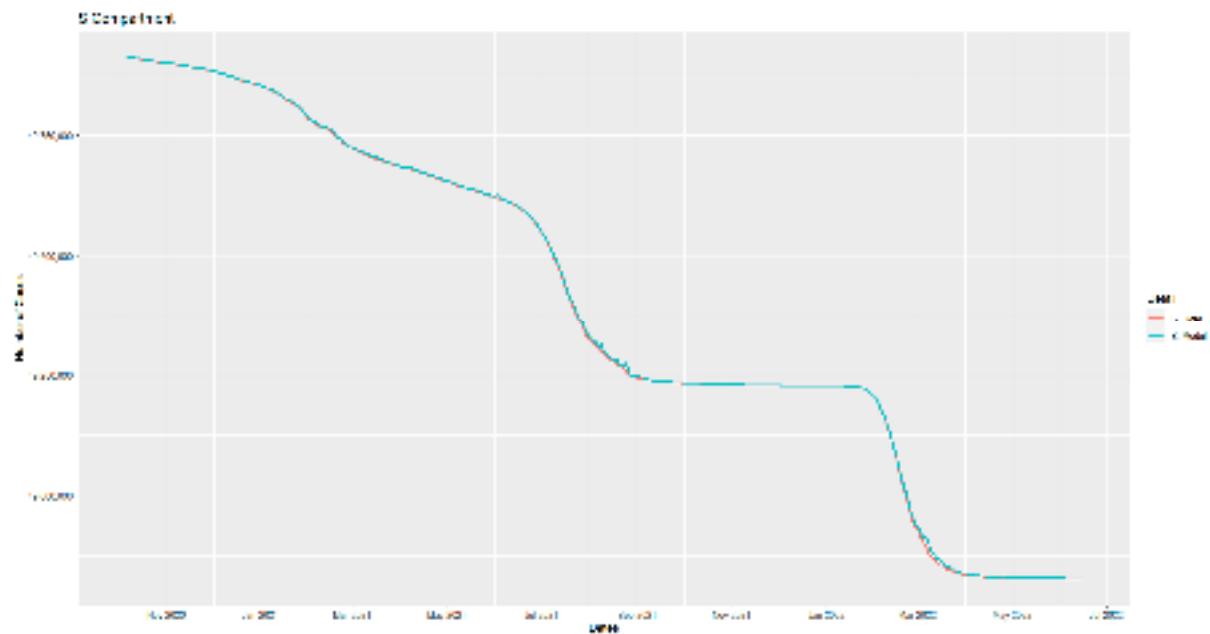


Figure 5. Comparison between model predictions and data in the S compartment.

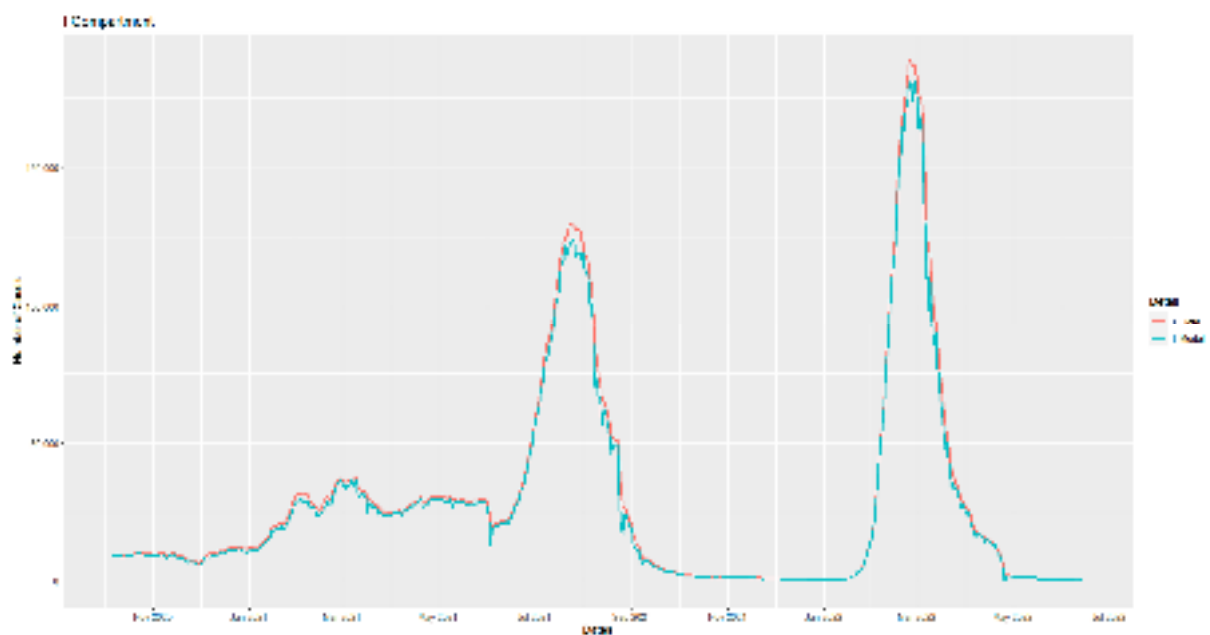


Figure 6. Comparison between model predictions and data in compartment I.

To obtain the prediction graph of the model in I, Equation (34) is used. Based on the graph, it is known that the prediction results of the model are in accordance with the real data. This is evident from the simulation graph, which is identical to the original data. This alignment between the model predictions and actual data validates the accuracy of the parameters utilized in the study. The consistency between the simulation graph and real data further supports the reliability of the numerical simulations conducted. This means that the model's predictions are accurate compared to the original data, and the parameters are fit for numerical simulations. Overall, the close match between the simulation results and real data indicates that the model is a reliable tool for predicting outcomes in similar scenarios. The successful validation of the parameters used in the study enhances the credibility of the numerical simulations performed.

4.3. Sensitivity Analysis

Before utilizing the estimated parameters in the numerical simulation, a sensitivity analysis is performed for each parameter. This procedure aids in understanding how variations in each parameter influence the overall model output, thereby facilitating necessary adjustments. By calculating the MAPE in compartment I, researchers can evaluate the model's accuracy and implement further refinements to enhance reliability. The sensitivity analysis was conducted using the one-at-a-time (OAT) method, wherein each parameter was systematically varied individually while holding the others constant. This method allows for a clear and isolated assessment of the impact of each parameter on the model's output. By adjusting one parameter at a time and observing the corresponding changes in the model's performance metrics, such as the mean absolute percentage error (MAPE), it becomes possible to identify the parameters with the most significant influence on the model. This comprehensive approach facilitates a thorough understanding of the model's sensitivity to different parameters, thereby aiding in its refinement to enhance accuracy and robustness. Furthermore, the OAT method helps identify potential interactions between parameters, ensuring that the model remains stable and reliable under varying conditions. This analysis involves varying the estimated parameters by $\pm 10\%$. Subsequently, the simulation is conducted, and the mean absolute percentage error (MAPE) is calculated for compartment I. The results of the sensitivity analysis, depicted in Figure 7, illustrate the impact of parameter variations on the MAPE baseline.

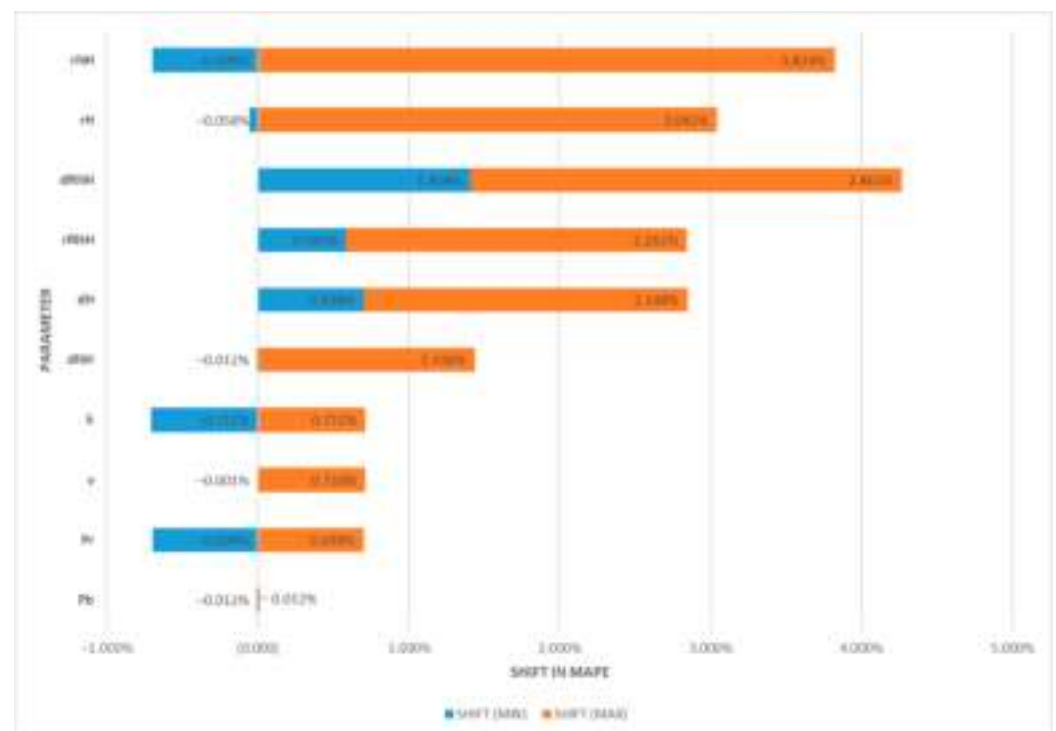


Figure 7. Change in MAPE.

Based on Figure 7, the parameter induces the most significant change in MAPE, amounting to 3.824%. This is followed by the r_H parameter, which results in a MAPE change of 3.042%. Notably, the transition parameters exhibit minimal impact on MAPE; for instance, the parameter alters MAPE by -0.012% , while the parameter causes a 0.699% change. Figure 7 further illustrates that the overall changes in MAPE are relatively modest. With a baseline MAPE of 7.12% , the largest variation still keeps the MAPE value within the 10% range. This indicates that, despite the variations in the parameters, the MAPE value remains within the acceptable threshold, as the maximum MAPE criterion for this study is set at 20% . Overall, the findings demonstrate that the transition parameters have a relatively minor effect on MAPE values, maintaining changes within permissible bounds.

This suggests that the model is robust and can effectively accommodate variations in these parameters without significantly compromising accuracy.

4.4. Numerical Simulation Results

This numerical simulation's objective is to depict the population dynamics in the COVID-19 pandemic. To obtain a simulation graph of each population compartment, apply Equations (33)–(37). The simulation graph will provide a visual representation of how each population compartment changes over time based on the equations provided. This will allow for a better understanding of the dynamics of the COVID-19 pandemic and help in making informed decisions regarding public health measures. The numerical simulation in this work makes use of the parameters presented in Table 1. Figures 8–10 depict the COVID-19 population patterns in West Java with various scenarios.

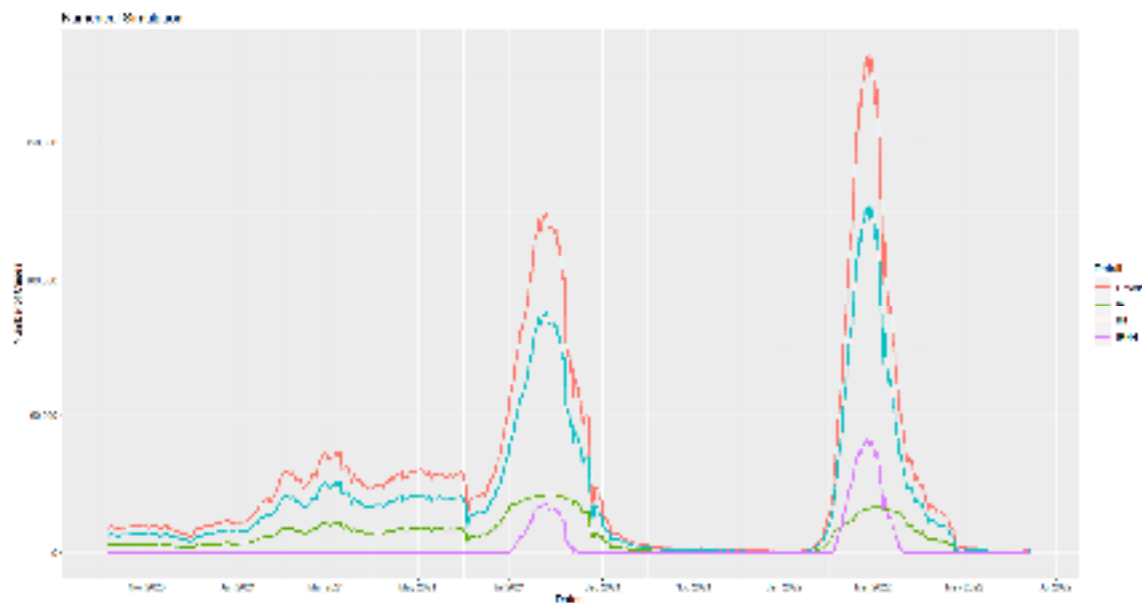


Figure 8. Numerical simulation results without intervention scenario.

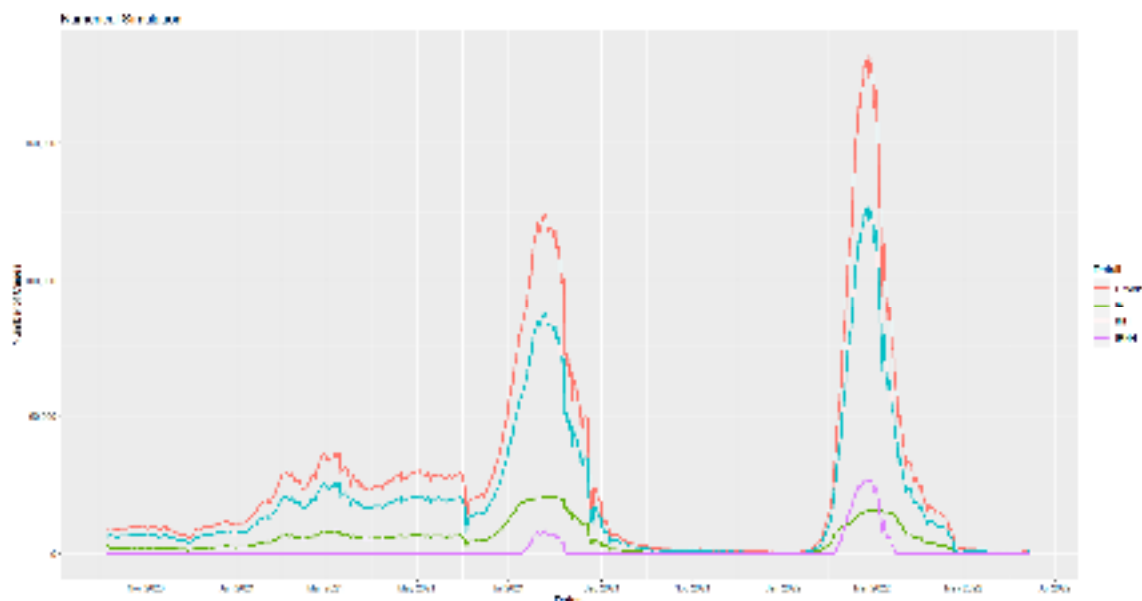


Figure 9. Numerical simulation results of social distancing scenario.

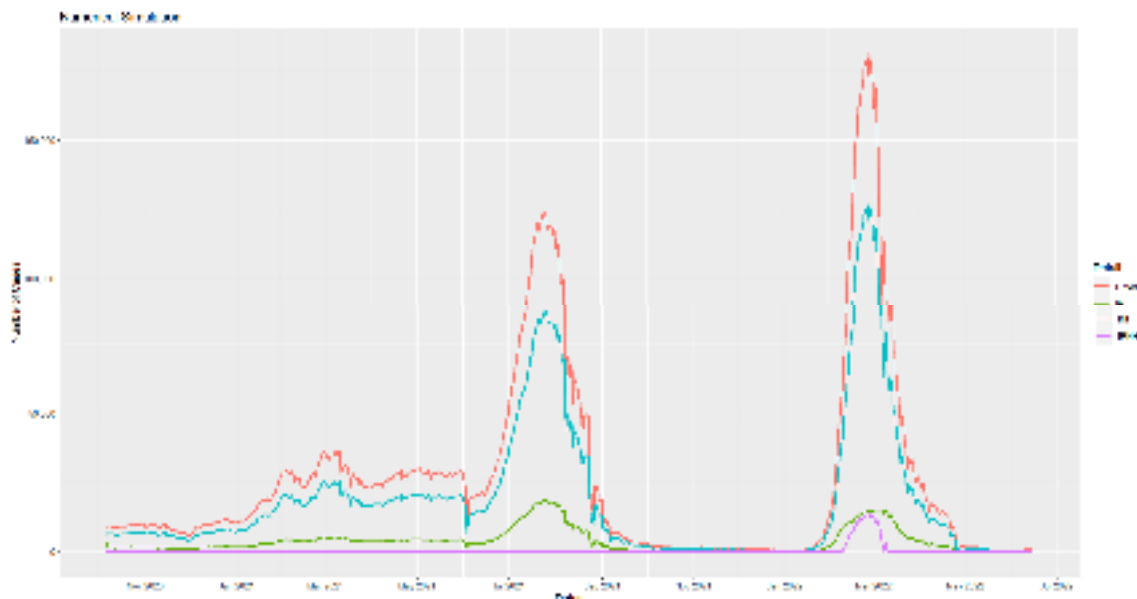


Figure 10. Numerical simulation results of lockdown scenario.

Figure 8 shows the results of numerical simulations of the spread of the COVID-19-infected population using the model that has been formed. In Figure 8, the scenario used is the scenario without intervention. The simulation illustrates a rapid increase in the number of infected individuals over time, highlighting the importance of implementing interventions to control the spread of the virus. These findings emphasize the potential impact of preventative measures on mitigating the effects of the pandemic. Figure 8 shows that without intervention, the COVID-19 cases experienced are very high. In addition, in the scenario without intervention from the government, it was found that patient rejection occurred twice, namely in August 2021 and February 2022. This indicates that without intervention, the healthcare system in West Java may become overwhelmed, leading to patient rejection. Therefore, it is crucial for authorities to enforce strict measures such as lockdowns, mask mandates, and social distancing to prevent such a scenario. Implementing these interventions can help reduce the burden on healthcare facilities and ultimately save lives. The data from Figure 8 highlights the importance of implementing effective measures to control the spread of COVID-19 in order to prevent such high numbers of cases and potential healthcare system strain.

Figure 9 shows the results of numerical simulations of the spread of the COVID-19-infected population using the model that has been formed. In Figure 9, the scenario used is the scenario with social distancing intervention. The simulation results indicate a significant reduction in the number of infected individuals compared to scenarios without social distancing measures. This highlights the effectiveness of implementing social distancing interventions in controlling the spread of COVID-19. Figure 9 shows that with social distancing intervention, the COVID-19 cases experienced have decreased when compared to without intervention. In addition, in the scenario with social distancing intervention, it was found that patient rejection occurred twice, namely in August 2021 and February 2022, with a smaller number of rejections compared to without intervention. These findings suggest that social distancing measures have played a crucial role in reducing the burden on healthcare systems by preventing a surge in cases. The data underscores the importance of continuing to adhere to these interventions to mitigate the impact of the pandemic. These findings suggest that implementing social distancing measures can effectively reduce the spread of COVID-19 and decrease the number of patient rejections. The results highlight the importance of proactive interventions in controlling the pandemic and minimizing healthcare system strain.

Figure 10 shows the results of numerical simulations of the spread of the COVID-19-infected population using the model that has been formed. In Figure 10, the scenario used is the scenario with lockdown intervention. Figure 10 shows that with lockdown intervention, the COVID-19 cases experienced have decreased when compared to no intervention and social distancing. In addition, in the scenario with lockdown intervention, it was found that patient rejection only occurred once, in February 2022, with a smaller number of rejections when compared to no intervention and social distancing. Overall, the results suggest that implementing a lockdown intervention can effectively reduce the spread of COVID-19 and minimize patient rejections. This highlights the importance of strict measures for controlling the pandemic and protecting public health.

Based on Figures 8–10, it is found that, in general, the number of patients who have severe symptoms and need treatment is 0, meaning that in most time intervals, all COVID-19 patients with severe symptoms will obtain treatment despite the increase in infected cases. This suggests that the healthcare system in West Java is able to effectively manage the influx of severe COVID-19 cases. The data from this simulation can be valuable for policymakers in determining the effectiveness of different public health measures. However, when the spike in infected cases is too high, for example, in June 2021 and January 2022, this will cause the hospital to be unable to handle the spike in cases and result in COVID-19 patients being rejected or not receiving treatment. When reviewed based on the scenario used, it was found that in the social distancing scenario, the number of patient rejections was less than without intervention, whereas in the lockdown scenario, there were only rejections in the January 2022 period. This highlights the importance of implementing timely and effective public health measures to prevent overwhelming healthcare systems during spikes in cases. It also underscores the need for continuous monitoring and adjustment of strategies based on simulation outcomes to ensure optimal outcomes. To determine how the impact on health costs needs to be prepared, Equations (40) and (41) are used. The graph of the costs that need to be prepared by the government is presented in Figures 11–13.

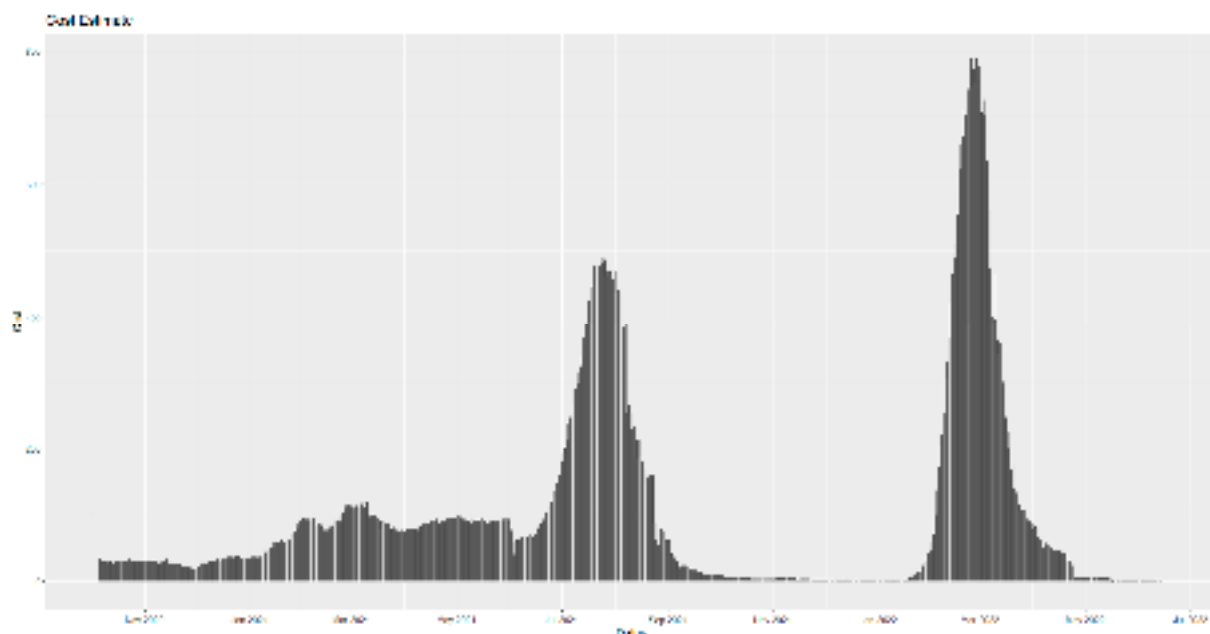


Figure 11. Cost estimation results without intervention scenario.

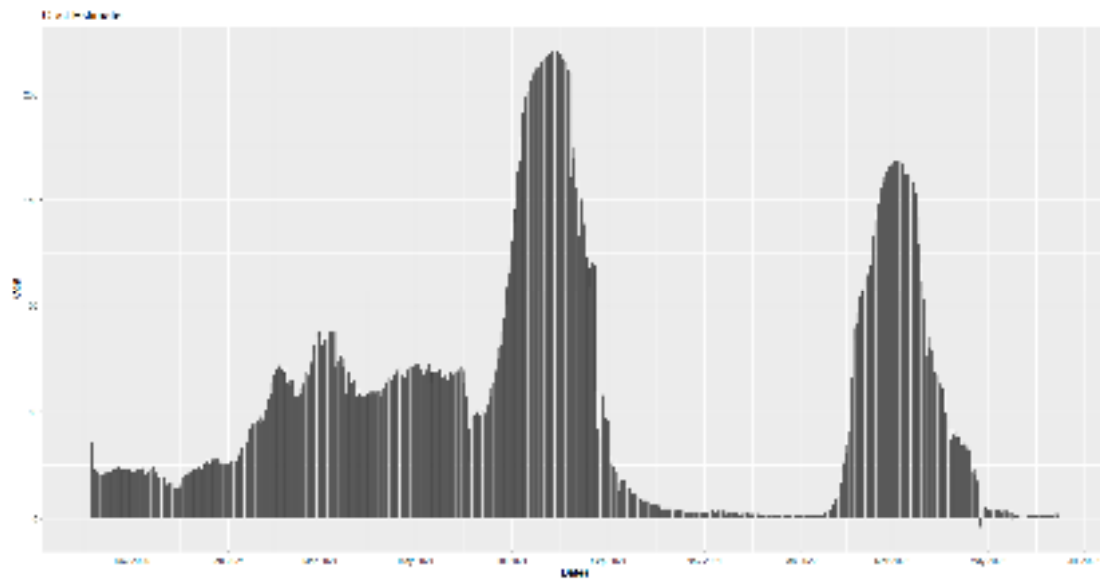


Figure 12. Cost estimation results with social distancing scenario.

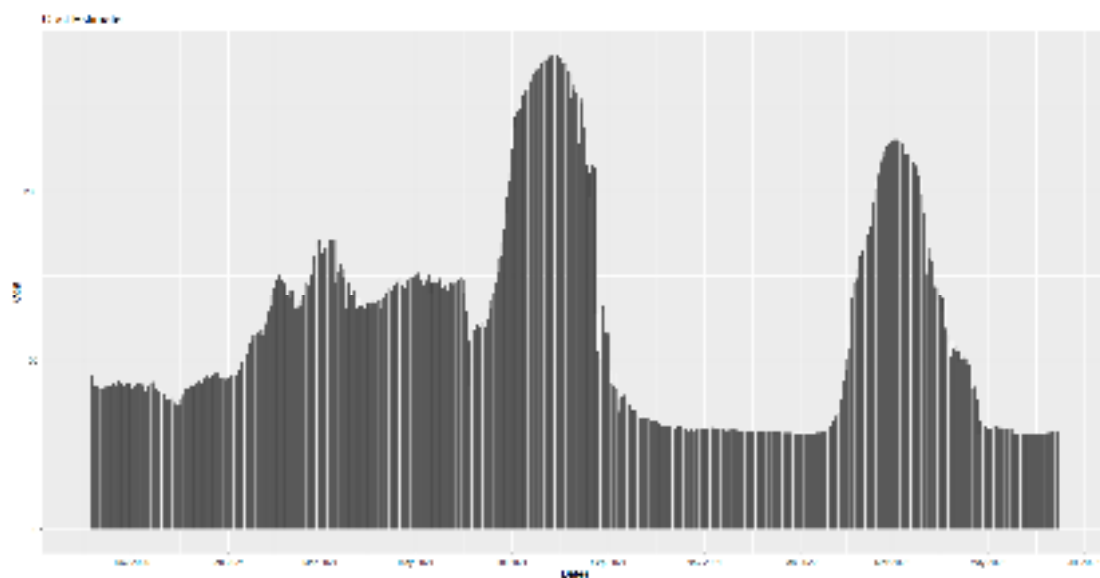


Figure 13. Cost estimation results with lockdown scenario.

Figure 11 shows the estimated cost of handling the COVID-19 pandemic based on the modeling results. The cost projections take into account various factors such as healthcare expenses, economic impacts, and government response measures. It is important to note that these estimates are subject to change based on the evolving nature of the pandemic and the effectiveness of containment efforts. Figure 11 shows that there is a spike in handling costs in the periods of August 2021 and February 2022. Figure 11 shows that there was also a surge in COVID-19 cases and patient rejection in the same period. These spikes in handling costs coincide with peaks in COVID-19 cases and patient rejection rates, indicating a direct correlation between the two. This information can help policymakers and healthcare providers better anticipate and allocate resources during future surges in the pandemic. This shows that the rejection of COVID-19 patients can increase the costs that need to be incurred by the government. This is because the government must accommodate all COVID-19 patients, and when there are cases of rejection, the bed capacity needs to be increased, and this causes a surge in the cost of handling the COVID-19 pandemic.

Figure 12 shows the estimated cost of handling the COVID-19 pandemic based on the modeling results. Based on Figure 12, the costs that need to be incurred for handling a pandemic with a social distancing scenario are less than without any intervention. This suggests that implementing social distancing measures can help reduce the financial burden associated with managing a pandemic. It is crucial for policymakers to consider these cost estimates when making decisions about public health interventions. In addition, Figure 12 shows that there is a spike in handling costs in the periods of August 2021 and February 2022, with a smaller amount when compared to the scenario without any intervention. Figure 9 shows that there was also a spike in COVID-19 cases and patient rejection during the same period. This shows that the rejection of COVID-19 patients can increase the costs that need to be incurred by the government. This is because the government must accommodate all COVID-19 patients, and when there are cases of rejection, the bed capacity needs to be increased, and this causes a surge in the cost of handling the COVID-19 pandemic. Therefore, reducing patient rejection can help mitigate the financial burden on the government and healthcare system. Implementing strategies to address patient rejection, such as improving communication and coordination between healthcare facilities, could lead to more efficient use of resources and ultimately lower costs in managing the pandemic.

Figure 13 shows the estimated cost of handling the COVID-19 pandemic based on the modeling results with the lockdown scenario. The estimated cost includes healthcare expenses, economic losses, and government stimulus packages. It is crucial for policymakers to consider these projections when making decisions regarding public health measures and economic recovery plans. Based on Figure 13, the costs that need to be incurred for handling a pandemic with a lockdown scenario are less than without any intervention. In addition, Figure 13 shows that there is a spike in handling costs in the periods of August 2021 and February 2022, with a smaller amount when compared to the scenario without intervention. This suggests that implementing a lockdown can help mitigate the overall costs associated with a pandemic. However, it is important to carefully weigh the economic impact of such measures against their potential benefits in terms of public health and long-term recovery. In Figure 10, it is shown that there was also a spike in COVID-19 cases and patient rejection during the same period. This shows that the rejection of COVID-19 patients can increase the costs incurred by the government. This is because the government must accommodate all COVID-19 patients, and when there are cases of rejection, the bed capacity needs to be increased; this causes a surge in the cost of handling the COVID-19 pandemic. This leads to a strain on resources and funding for other important public health initiatives. It is crucial for governments to address patient rejection issues promptly to ensure efficient allocation of resources and prevent unnecessary strain on the healthcare system. By implementing strategies to reduce patient rejection rates, governments can better manage the financial burden of the COVID-19 pandemic and prioritize funding for various public health initiatives.

Based on Figures 11–13, it was found that the cost of handling COVID-19 increased when the number of existing COVID-19 patients also increased. In addition, it was also found that after the spike occurred, the cost of handling COVID-19 cases decreased and tended to remain low until the next spike occurred. Therefore, it is crucial for the government to closely monitor the number of existing COVID-19 patients in order to anticipate and prepare for potential spikes in health costs. By using Equations (40) and (41) and analyzing the graph in Figures 8–10, policymakers can make informed decisions to effectively manage and allocate resources for handling COVID-19 cases. This suggests that there is increased public awareness after a spike in cases, leading to a reduction in the number of COVID-19 cases so that people begin to feel safe again and relax health protocols, which causes the number of COVID-19 cases to increase again. It is crucial for policymakers to closely follow trends and patterns in COVID-19 cases to implement timely interventions and prevent overwhelming healthcare systems. By continuously monitoring and adjusting strategies based on data analysis, the impact of potential spikes in health costs can be minimized. Based on Figures 11–13, it is found that the highest total cost is experienced by

scenario 3 with lockdown intervention, which is 73 billion IDR; the highest cost is followed by scenario 1 with no intervention scenario, which has a cost of 67.42 billion IDR; and the smallest cost is experienced by scenario 2, which costs 32.872 billion IDR. These findings suggest that implementing a lockdown intervention can lead to higher total costs compared to scenarios with no intervention. It is crucial for policymakers to carefully consider the economic implications of different strategies for managing COVID-19 outbreaks.

5. Discussion

This study aims to formulate a costing model for COVID-19 patients, including (a) patients receiving treatment until they recover and (b) additional hospital bed capacity when required. The formulation of the model to determine the cost of handling COVID-19, which involves the possibility of a transition in the severity of symptoms in this study, has been successfully carried out, the results of which are presented in Equations (33)–(37). The model takes into account various factors such as length of hospital stay, severity of symptoms, and required medical interventions to provide a comprehensive cost estimate. By incorporating these variables, healthcare providers can better understand the financial implications of treating COVID-19 patients and allocate resources effectively. The model given by the equation can be used to determine the number of costs that need to be prepared by the government to treat infected patients and increase capacity. Additionally, the model takes into account the potential fluctuation in patient conditions, ensuring a more accurate estimation of costs. By utilizing this comprehensive costing model, policymakers can better allocate resources and plan for potential surges in COVID-19 cases. This will ultimately help in optimizing healthcare resources and ensure a timely and effective response to the pandemic. Furthermore, the model can be adapted for future outbreaks or other public health emergencies to improve preparedness and response strategies.

Based on the simulations conducted, we found that over time, the susceptible population will continue to decrease. The situation arises due to the ongoing occurrence of COVID-19 infections followed by subsequent recoveries within the population over a period of time. This is what causes the vulnerable population to continue to grow over time. The opposite is true for the recovered population—over time, this population will continue to grow. An interesting thing happened in the infected population. A fluctuation was found in this population, including in June 2021 and January 2022. During these months, there was a significant increase in COVID-19 cases, leading to a spike in the infected population. However, with proper measures and interventions, the infected population gradually decreased again. There was a significant spike during that period, before further decreasing until it returned to a stable low. This increase affects both the incurred expenses and the quantity of patients who are refused admission by the hospital. It is important for healthcare facilities to monitor these fluctuations closely in order to allocate resources effectively and provide timely care to those in need. By understanding the trends in the infected population, hospitals can better prepare for potential surges in cases and adjust their capacity accordingly. Additionally, tracking these trends can also help hospitals identify any potential outbreaks in the community and implement preventive measures to contain the spread of the virus. This proactive approach can ultimately help save lives and prevent overwhelming healthcare systems. It is crucial for healthcare facilities to collaborate with public health agencies and utilize data analytics to stay ahead of the curve in managing the pandemic. By leveraging technology and predictive modeling, hospitals can make informed decisions that prioritize patient care and safety during these uncertain times.

Based on the simulations, we found that although the number of infected cases occasionally increased, these infected cases would always receive treatment in health facilities, as reflected in Figure 10. However, when there is a significant increase in cases (for example, in June 2021 and January 2022), the existing health facilities cannot handle the surge in cases, resulting in denied admission of COVID-19 patients. Strategic planning is required to address this issue. Hospitals can implement emergency response plans to expand capacity and resources during times of high demand. By proactively preparing for potential surges in

cases, hospitals can ensure that all patients receive the care they need during the pandemic. Additionally, hospitals can collaborate with other healthcare facilities to coordinate patient transfers and optimize bed availability. This coordinated effort can help alleviate strain on individual hospitals and ensure that patients are able to receive timely and appropriate care. This increased the number of health expenditures, as shown in Figures 11–13, and the number of deaths. Based on these findings, there are two things that the government needs to focus on in handling COVID-19. The first is that the government needs to take a policy that can stabilize the number of newly infected cases every week. Of course, reducing the number of daily new infections to 0 is impossible considering the current status of COVID-19, which has become endemic. The second focus should be on strengthening and expanding the existing health facilities to accommodate the increasing number of cases. This emphasizes that proper investment in additional beds in healthcare facilities will help reduce the burden on hospitals and prevent the rejection of COVID-19 patients, which will ultimately lead to lower healthcare costs and mortality rates.

Additionally, investing in public health education and awareness campaigns can help in promoting preventive measures and responsible behavior among the population. By emphasizing the importance of vaccination, mask-wearing, and social distancing, we can collectively work towards effectively managing the spread of the virus. However, it is possible to suppress the number of new cases at a stable low rate. If this can be done, there will be no cases of patients being denied admission by the hospital. The second focus is the addition of health facilities; the greater the government's capacity to add beds per period, the less likely it is that healthcare providers will be unable to admit patients. In this study, we used the parameter $n = 0.05$, meaning that if the government's capacity to add beds is greater than that, the number of denied admission patients will decrease when a spike occurs, or maybe it can be 0. However, if the government's capacity is smaller, then the number of denied admission cases will be even greater. Patients who are rejected from medical care are often those with severe symptoms that necessitate appropriate treatment at a healthcare facility. These individuals require urgent attention to manage their conditions effectively. When a large number of these patients are turned away, it can lead to a critical situation where their health deteriorates rapidly. Without timely medical intervention, their chances of recovery diminish significantly. This increase in untreated severe cases can substantially raise mortality rates. Therefore, the healthcare system must ensure that patients with severe symptoms receive the necessary care. Failing to do so not only impacts individual health outcomes but also contributes to a higher overall mortality rate within the community.

This highlights the importance of strategic planning and investment in healthcare infrastructure to ensure adequate capacity during times of increased demand. It also underscores the potential consequences of inadequate government intervention in healthcare system management.

6. Conclusions

This study developed a discrete susceptible-infected-susceptible (SIS) model to project the number of patients requiring hospitalization. The model considers the possibility of transitioning the severity of the patient's symptoms, either from mild to severe or vice versa. The results showed that early intervention and effective containment measures could significantly reduce the burden on healthcare systems. This dynamic approach to modeling patient outcomes can help inform resource allocation and public health strategies during disease outbreaks. By accurately predicting hospitalization needs, healthcare facilities can better prepare for surges in patients and allocate resources effectively. This proactive approach can ultimately save lives by ensuring that patients receive the appropriate level of care in a timely manner. Additionally, it highlights the importance of implementing preventative measures to control the spread of infectious diseases. This proactive approach can ultimately save lives and prevent overwhelming the healthcare system. The results of the projected number of patients are then used to determine the cost of treatment and additional beds following the provisions set by the Government of Indonesia. This allows

for efficient planning and budgeting to ensure that hospitals have the necessary resources to accommodate the predicted influx of patients.

By utilizing this model, healthcare facilities can optimize their response to potential increases in patient volume and provide better care for those in need. The model that has been created is simulated on COVID-19 data from West Java Province. This data-driven approach helps hospitals anticipate and prepare for surges in patient numbers, ultimately improving the overall quality of care provided. Additionally, by analyzing trends and patterns in the data, healthcare facilities can make informed decisions to enhance their capacity and response strategies. This study found that the social distancing scenario gave the best result in COVID-19 intervention. Social distancing reduces the number of hospitalized persons while still maintaining a reasonable cost. The resulting estimated cost of treating COVID-19 patients can be used as input for policies and as an alert to the government about expenses that may occur in various scenarios. This information can help healthcare facilities, and policymakers make informed decisions about resource allocation and planning for future outbreaks. By analyzing the data from this study, authorities can implement effective measures to control the spread of the virus and minimize the economic impact on the healthcare system. Additionally, understanding the cost implications of hospitalizations can aid in developing strategies to ensure adequate funding and support for healthcare systems during times of crisis. This data can also be valuable for predicting potential financial burdens on individuals and families affected by COVID-19.

This research underscores the critical importance of effectively managing infection rates and adequately preparing for patient care in the event of a pandemic. The COVID-19 pandemic has profoundly impacted both the population and the economy, illustrating the necessity for robust contingency plans to ensure better preparedness for future pandemics. Initial governmental responses played a pivotal role in containing the spread of the virus. Swift interventions, such as lockdowns, travel restrictions, and social distancing mandates, were implemented to prevent the healthcare system from being overwhelmed by new infections. Simultaneously, the government undertook significant efforts to increase the bed capacity in healthcare facilities, including the rapid construction of temporary hospitals, the conversion of existing facilities to COVID-19 treatment centers, and the procurement of additional medical equipment and supplies. These measures were designed to address potential worst-case scenarios involving sudden and substantial surges in infection rates. Proactive strategies are essential for minimizing mortality rates and preserving human lives. By expanding healthcare capacity, governments can ensure that adequate medical care is available for all patients, thus reducing the strain on existing resources and maintaining the functionality of the healthcare system for essential services beyond COVID-19 treatment. This study illustrates that a comprehensive approach, encompassing both infection control and healthcare capacity expansion, is paramount in mitigating the detrimental effects of pandemics. The effective management of infection rates through targeted interventions, combined with strategic enhancements to healthcare infrastructure, can significantly reduce the impact on public health and the economy. This research advocates for the continuous development and refinement of pandemic preparedness plans to safeguard against future outbreaks and ensure a resilient healthcare system capable of responding to crises efficiently and effectively.

In this study, we used the parameter $n = 0.05$, meaning that if the government's capacity to add beds is greater than that, then the number of denied admissions will decrease the government's capacity to add beds is greater than that, then the number of patients denied admission will decrease when there is a surge, or it could possibly be 0. However, if the capacity of the government is smaller, then the number of cases denied admission will be even greater. This highlights the importance of strategic planning and investment in health infrastructure to ensure adequate capacity during times of increased demand. It also underscores the potential consequences of inadequate government intervention in healthcare system management. The limitation of this study is that it only examines the dynamics between the financing components of pandemic intervention without paying attention to its impact on the macroeconomy as a whole. The link between pandemic

financing and macroeconomics is also important, considering that the sustainability of society in a pandemic situation depends on both.

In future research, accurately estimating this parameter will yield more precise results regarding the costs associated with managing the COVID-19 pandemic. With a well-estimated value of the parameter n , hospitals can reduce the number of patients they must turn away and diminish the expenses associated with expanding hospital capacity. This is because an accurately estimated n will enable the determination of the optimal number of additional beds required by the government. Consequently, this minimizes the disparity between the current number of available beds and the number of new infections, ensuring a more efficient allocation of resources and better preparedness for future waves of the pandemic.

Author Contributions: Conceptualization, D.S.P. and S.; methodology, N.A.; software, D.S.P.; validation, D.S.P., S. and N.M.Y.; formal analysis, S.; investigation, N.A. and N.M.Y.; resources, N.A.; data curation, D.S.P.; writing—original draft preparation, D.S.P.; writing—review and editing, S. and N.M.Y.; visualization, D.S.P.; supervision, N.A.; project administration, S.; funding acquisition, S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Padjadjaran Postgraduate Excellence Scholarship (BUPP) grant number 2203/UN6.3.1/PT.00/2022.

Data Availability Statement: The original data presented in the study are openly available in Github at <https://github.com/DhikaSuryaP/SIS-DATASET>.

Acknowledgments: The authors would like to thank the Dean of the Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, and the Directorate of Research and Community Service (DRPM), who has given a grant: Pendidikan Magister Menuju Doktor Bagi Sarjana Unggul (PMDSU/BUPP).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Walløe, L. 3 Medieval and modern bubonic plague: Some clinical continuities. *Med. Hist.* **2008**, *52*, 59–73. [CrossRef]
2. Tumpey, T.M.; Basler, C.F.; Aguilar, P.V.; Zeng, H.; Solórzano, A.; Swayne, D.E.; Cox, N.J.; Katz, J.M.; Taubenberger, J.K.; Palese, P.; et al. Characterization of the reconstructed 1918 Spanish influenza pandemic virus. *Science* **2005**, *310*, 77–80. [CrossRef] [PubMed]
3. Sabin, C.A.; Lundgren, J.D. The natural history of HIV infection. *Curr. Opin. HIV AIDS* **2013**, *8*, 311. [CrossRef]
4. Lam, W.K.; Zhong, N.S.; Tan, W.C. Overview on SARS in Asia and the world. *Respirology* **2003**, *8*, S2–S5. [CrossRef] [PubMed]
5. Ciotti, M.; Ciccozzi, M.; Terrinoni, A.; Jiang, W.C.; Wang, C.B.; Bernardini, S. The COVID-19 pandemic. *Crit. Rev. Clin. Lab. Sci.* **2020**, *57*, 365–388. [CrossRef] [PubMed]
6. Hale, T.; Angrist, N.; Goldszmidt, R.; Kira, B.; Petherick, A.; Phillips, T.; Webster, S.; Cameron-Blake, E.; Hallas, L.; Majumdar, S.; et al. A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). *Nat. Hum. Behav.* **2021**, *5*, 529–538. [CrossRef] [PubMed]
7. Izumi, T.; Sukhwani, V.; Surjan, A.; Shaw, R. Managing and responding to pandemics in higher educational institutions: Initial learning from COVID-19. *Int. J. Disaster Resil. Built Environ.* **2021**, *12*, 51–66. [CrossRef]
8. Bai, H.M.; Zaid, A.; Catrin, S.; Ahmed, K.; Ahmed, A. The socio-economic implications of the coronavirus pandemic (COVID-19): A review. *Int. J. Surg.* **2020**, *8*, 8–17.
9. Rahim, F.; Allen, R.; Barroy, H.; Gores, L.; Kutzin, J. *COVID-19 Funds in Response to the Pandemic*; International Monetary Fund: Washington, DC, USA, 2020.
10. Sen-Crowe, B.; McKenney, M.; Elkbulli, A. Public health prevention and emergency preparedness funding in the United States: Are we ready for the next pandemic? *Ann. Med. Surg.* **2020**, *59*, 242. [CrossRef]
11. Rudolph, C.W.; Allan, B.; Clark, M.; Hertel, G.; Hirschi, A.; Kunze, F.; Shockley, K.; Shoss, M.; Sonnentag, S.; Zacher, H. Pandemics: Implications for research and practice in industrial and organizational psychology. *Ind. Organ. Psychol.* **2021**, *14*, 1–35. [CrossRef]
12. Bakker, A.; Cai, J.; Zenger, L. Future themes of mathematics education research: An international survey before and during the pandemic. *Educ. Stud. Math.* **2021**, *107*, 1–24. [CrossRef] [PubMed]
13. Adiga, A.; Dubhashi, D.; Lewis, B.; Marathe, M.; Venkatramanan, S.; Vullikanti, A. Mathematical models for covid-19 pandemic: A comparative analysis. *J. Indian Inst. Sci.* **2020**, *100*, 793–807. [CrossRef] [PubMed]
14. Jewell, N.P.; Lewnard, J.A.; Jewell, B.L. Predictive mathematical models of the COVID-19 pandemic: Underlying principles and value of projections. *JAMA* **2020**, *323*, 1893–1894. [CrossRef] [PubMed]
15. Dudine, P.; Hellwig, K.P.; Jahan, S. *A Framework for Estimating Health Spending in Response to COVID-19*; International Monetary Fund: Washington, DC, USA, 2020.

16. Torres-Rueda, S.; Sweeney, S.; Bozzani, F.; Naylor, N.R.; Baker, T.; Pearson, C.; Eggo, R.; Procter, S.R.; Davies, N.; Quaife, M.; et al. Stark choices: Exploring health sector costs of policy responses to COVID-19 in low-income and middle-income countries. *BMJ Glob. Health* **2021**, *6*, e005759. [CrossRef] [PubMed]
17. Cook, M.J.; Dri, G.G.; Logan, P.; Tan, J.B.; Flahault, A. COVID-19 Down Under: Australia's initial pandemic experience. *Int. J. Environ. Res. Public Health* **2020**, *17*, 8939. [CrossRef]
18. McQuade, S.T.; Weightman, R.; Merrill, N.J.; Yadav, A.; Trélat, E.; Allred, S.R.; Piccoli, B. Control of COVID-19 outbreak using an extended SEIR model. *Math. Models Methods Appl. Sci.* **2021**, *31*, 2399–2424. [CrossRef]
19. Samuel, J.; Sinha, S. Optimal control in pandemics. *Phys. Rev. E* **2021**, *103*, L010301. [CrossRef] [PubMed]
20. De Lara-Tuprio, E.P.; Estuar, M.R.; Sescon, J.T.; Lubangco, C.K.; Castillo, R.C.; Teng, T.R.; Tamayo, L.P.; Macalalag, J.M.; Vedeja, G.M. Economic losses from COVID-19 cases in the Philippines: A dynamic model of health and economic policy trade-offs. *Humanit. Soc. Sci. Commun.* **2022**, *9*, 111. [CrossRef]
21. Huppert, A.; Katriel, G. Mathematical modelling and prediction in infectious disease epidemiology. *Clin. Microbiol. Infect.* **2013**, *19*, 999–1005. [CrossRef]
22. Alamo, T.; Reina, D.G.; Millán, P. Data-driven methods to monitor, model, forecast and control covid-19 pandemic: Leveraging data science, epidemiology and control theory. *arXiv* **2020**, arXiv:2006.01731.
23. Myall, A.; Price, J.; Peach, R.; Abbas, M.; Mookerjee, S.; Ahmad, I.; Ming, D.; Zhu, N.J.; Ramzan, F.; Weisse, A.; et al. Prediction of hospital-onset COVID-19 using networks of patient contact: An observational study. *Int. J. Infect. Dis.* **2022**, *116*, S109–S110. [CrossRef]
24. Jordan, E.; Shin, D.E.; Leekha, S.; Azarm, S. Optimization in the context of COVID-19 prediction and control: A literature review. *IEEE Access* **2021**, *9*, 130072–130093. [CrossRef] [PubMed]
25. Wearing, H.J.; Rohani, P.; Keeling, M.J. Appropriate models for the management of infectious diseases. *PLoS Med.* **2005**, *2*, e174. [CrossRef] [PubMed]
26. Brauer, F.; Fenga, Z.; Castillo-Chaveza, C. Discrete Epidemic Models. *Math. Biosci. Eng.* **2010**, *7*, 1–15. [PubMed]
27. Allen, L.J.; Burgin, A.M. Comparison of deterministic and stochastic SIS and SIR. *Math. Biosci.* **2000**, *163*, 1–33. [CrossRef] [PubMed]
28. Allen, L.J.S. Some Discrete-Time SI, SIR, and SIS Epidemic Models. *Math. Biosci.* **1994**, *124*, 83–105. [CrossRef] [PubMed]
29. Zhou, Y.; Ma, Z. A discrete epidemic model for SARS transmission and control in China. *Math. Comput. Model.* **2004**, *40*, 1491–1506. [CrossRef] [PubMed]
30. Škulj, D. Discrete time Markov chains with interval probabilities. *Int. J. Approx. Reason.* **2009**, *50*, 1314–1329. [CrossRef]
31. Goodwin, P.; Lawton, R. On the asymmetry of the symmetric MAPE. *Int. J. Forecast.* **1999**, *15*, 405–408. [CrossRef]
32. Jawa Barat, P.P. Dashboard Statistik Kasus COVID-19 Provinsi Jawa Barat. Pemerintah Daerah Provinsi Jawa Barat; 2024. Available online: <https://dashboard.jabarprov.go.id/id/dashboard-pikobar/trace/statistik> (accessed on 21 June 2024).
33. Jawa Barat, P.P. Keterisian Tempat Tidur (BOR). Pemerintah Daerah Provinsi Jawa Barat; 2024. Available online: <https://dashboard.jabarprov.go.id/id/dashboard-pikobar/treatment/fasyankes> (accessed on 21 June 2024).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Design and Evaluation of a Hypersonic Waverider Vehicle Using DSMC

Angelos Klothakis  and Ioannis K. Nikolos * 

Turbomachines & Fluid Dynamics Laboratory, School of Production Engineering and Management, Technical University of Crete, 73100 Chania, Greece; aklothakis@gmail.com

* Correspondence: inikolos@tuc.gr

Abstract: This work investigates the aerodynamic performance of a hypersonic waverider designed to operate at Mach 7, focusing on optimizing its design through advanced computational methods. Utilizing the Direct Simulation Monte Carlo (DSMC) method, the three-dimensional flow field around the specifically designed waverider was simulated to understand the shock wave interactions and thermal dynamics at an altitude of 90 km. The computational approach included detailed meshing around the vehicle's critical leading edges and the use of three-dimensional iso-surfaces of the Q-criterion to map out the shock and vortex structures accurately. Additional simulation results demonstrate that the waverider achieved a lift–drag ratio of 2.18, confirming efficient aerodynamic performance at a zero-degree angle of attack. The study's findings contribute to the broader understanding of hypersonic flight dynamics, highlighting the importance of precise computational modeling in developing vehicles capable of operating effectively in near-space environments.

Keywords: hypersonic vehicles; direct simulation Monte Carlo; aerodynamic performance; computational fluid dynamics; hypersonic waverider



Citation: Klothakis, A.; Nikolos, I.K. Design and Evaluation of a Hypersonic Waverider Vehicle Using DSMC. *Computation* **2024**, *12*, 140. <https://doi.org/10.3390/computation12070140>

Academic Editor: Jeffrey S. Marshall

Received: 17 May 2024

Revised: 24 June 2024

Accepted: 3 July 2024

Published: 9 July 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The waverider is a distinct configuration of a hypersonic vehicle engineered to maintain the shock wave generated by the vehicle attached to its leading edge, offering numerous advantages [1]: (a) this shock wave effectively separates the flow fields on the upper and lower surfaces of the vehicle, thus preventing high-pressure flow from spilling from the lower to the upper surface at the leading edge; (b) it ensures a uniform flow on the lower surface, which is optimal for scramjet engine intake; (c) the design of the waverider is derived through reverse engineering from a known flow field around a conical shape. This design approach will be elaborated on in later sections of this work.

The concept of the waverider was first put forward by Nonweiler in 1963 [2]. At that time, the available technology was not advanced enough to produce a hypersonic air-breathing vehicle. Sandlin and Pessin offered an in-depth explanation of the design methodology for waveriders [1]. A broad analysis of the significant design challenges associated with hypersonic flight vehicles is discussed in the work of Sziroczak and Smith [3], where two types of vehicles, hypersonic transports and space launchers, are examined. Although these are distinct from waveriders, they share many design and construction challenges. Zhang et al. employed the three-dimensional Class–Shape Function Transformation (CST) method to determine the geometric shapes of Hypersonic Gliding Vehicles (HGVs) [4]. Inviscid modeling methods, relying on modified Newton theory and the Prandtl–Meyer equation, were used to assess the aerodynamics of these vehicles. The use of the NSGA-II multi-objective optimization algorithm helped achieve designs with high lift–drag ratios and substantial volumetric efficiency. In his Ph.D. thesis, K. Kontogiannis considered the parameterization and manipulation of waverider forebody geometries, proposing a new three-dimensional leading-edge waverider design method, shifting from inverse to direct

design methodologies [5]. Additionally, Son et al. described an optimization framework for the inverse design of hypersonic waveriders [6].

Since its introduction, the waverider has been a focal point of research. After conceptualizing the design, the next phase involves its analytical assessment. Various methods such as engineering correlations, flight tests, wind tunnel tests, and numerical simulations are employed not only for analysis but also for validating concepts and accumulating knowledge. While effective in the subsonic domain, these methods have shown limitations in reliability and implementation at hypersonic speeds. Ongoing research aims to refine these methods to master the hypersonic domain.

During the early quest for hypersonic flight, NASA designed, constructed, and tested a hypersonic vehicle [7]. The vehicle demonstrated that it was capable of achieving speeds up to Mach 8. However, a closer look at wind tunnel data revealed an interesting phenomenon. It was observed that the vehicle's aerodynamic performance parameter lift-drag ratio (L/D) decreased as Mach number increased. Kuchemann documented a decreasing trend in maximum L/D with increasing Mach numbers through extensive empirical research [8], highlighting a significant challenge in traditional aerodynamic design and methodology in the hypersonic regime. Moreover, Bowcutt introduced a multidisciplinary approach to optimize waverider configurations, achieving relatively higher L/D ratios at higher Mach numbers [9], a finding later corroborated by other researchers [10,11]. Nonweiler originally developed the waverider design technique in the 1950s to create a hypersonic wing for re-entry purposes [12]. Despite early challenges due to the inherent sharpness of the waverider's leading edge, which led to significant aerothermodynamic loads, and the limited volumetric efficiencies of initial thin configurations [7] that did not account for viscous effects, advancements in numerical methods and computational resources in the 1980s enabled Bowcutt to optimize waveriders by predicting the viscous effects on their performance [9]. This has spurred continued interest and research into waverider designs, although practical application in producing operational hypersonic vehicles remains a challenge.

Since Nonweiler [2] introduced the waverider design concept, leveraging wedge flow dynamics to create the first waverider in aerospace science, numerous design studies have aimed to refine and optimize his initial concept. The majority of these studies model the flow as two-dimensional and inviscid, either planar or axisymmetric [13–22]. A smaller subset has considered three-dimensional inviscid flow, both without and with chemical reactions, as in references [23–28] and [29,30], respectively. Only a limited number of investigations have addressed three-dimensional viscous flow, incorporating turbulence [31]. This study focuses on analyzing the full three-dimensional viscous flow field around the designed waverider.

2. Waverider Design Methodology

Waveriders are crafted using an inverse design methodology, a process that leverages the characteristics of a hypersonic flow environment. This approach utilizes the streamlines of the inviscid flow field post-shock as the basis for designing the compression stream surface, typically the lower surface of the waverider geometries. The end result is a streamlined geometry optimized for hypersonic flows. The geometry of the lower surface, defined by a set of adjacent streamlines, ensures the leading edge remains attached to the shock wave. Below is an outline of the algorithm used to generate waverider geometries:

- Determining and designing the initial flow field in the direction of the flow.
- Solving for the flow field using computational fluid dynamics methods, such as the Taylor–Maccoll equation [32], or the Euler, or Navier–Stokes equations.
- Mapping out the streamlines.
- Applying osculating cone theory along the spanwise direction.
- Refining and finalizing the model design.

To specify the initial flow conditions from which the waverider is derived, a cone with a predetermined angle δ is first designed. Then, the inlet boundary conditions like Mach

number, pressure, and temperature are set, and the resultant flow field is computed. After calculating the flow field, the streamlines can be computed by numerically solving the following set of ordinary differential equations (ODEs):

$$\begin{aligned}\frac{dx}{dt} &= u(x, y, z) \\ \frac{dy}{dt} &= v(x, y, z) \\ \frac{dz}{dt} &= w(x, y, z)\end{aligned}\quad (1)$$

Using the Runge–Kutta method, the above set of equations can be integrated in time to obtain each streamline coordinate.

A schematic overview of the streamlines traced from a conical shock is depicted in Figure 1. To establish a waverider geometry using the cone-derived method, one must first define a base curve. This curve is then projected onto the shock, aligned parallel to the freestream, to form the leading edge. Subsequently, the streamlines are traced backward from the leading-edge positions on the shock to the base plane, creating the waverider's compression (lower) surface. Beyond the analytical Equation (1), an alternative, simpler method can be employed to trace the streamlines by following tangential paths to the base cone surface, thus facilitating the generation of the compression surface geometry. The base curve equation, essential for designing the waverider, is illustrated around the symmetry plane in Figure 2. The total width of the waverider is denoted by l , with l_u representing the flat section of the curve. The curved section of the curve is determined by the equation depicted in Figure 2. The value of b is calculable from l , l_u , and h , values that are defined following the resolution of the initial conical flow field.

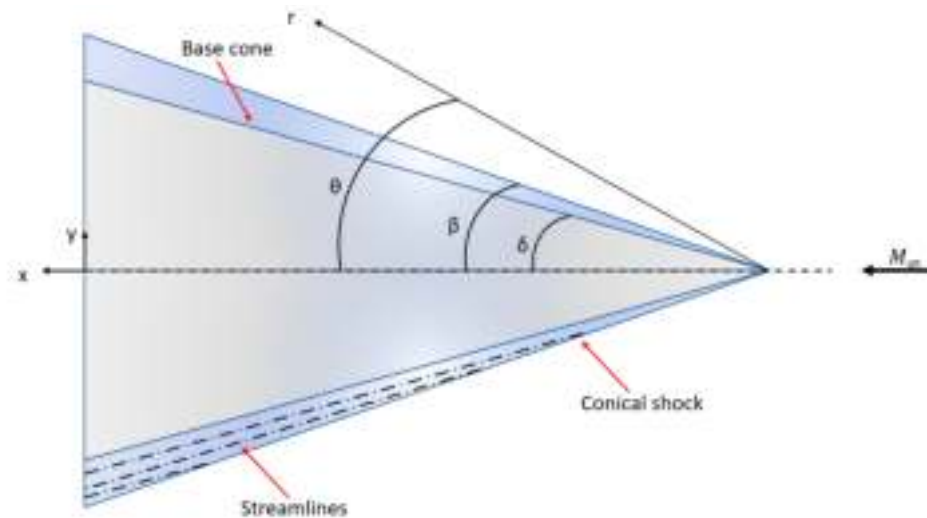


Figure 1. Schematic representation of the waverider design methodology.

For a practical waverider design, addressing the sharpness of the leading edge is crucial. Thus, a pragmatic design strategy incorporates blunting techniques to manage this characteristic. Blunting the leading edge and the junctions where two stream surfaces converge mitigates heating effects and results in a geometry more feasible for manufacturing. However, blunting does deviate from the ideal waverider design, allowing some pressure leakage from the lower to the upper surface, which in turn increases drag and diminishes the aerodynamic efficiency of the waverider. Studies [33,34] have demonstrated that while blunting reduces the heat flux experienced by the vehicle, it adversely impacts the lift–drag ratio (L/D). Despite this, a blunted waverider design can still offer superior aerodynamic performance compared to other design alternatives. The extent and nature of the blunting are crucial and require a balance between reducing thermal stresses and optimizing aerodynamic efficiency. With various possible configurations for a waverider, there are numerous approaches to designing a blunt leading edge.

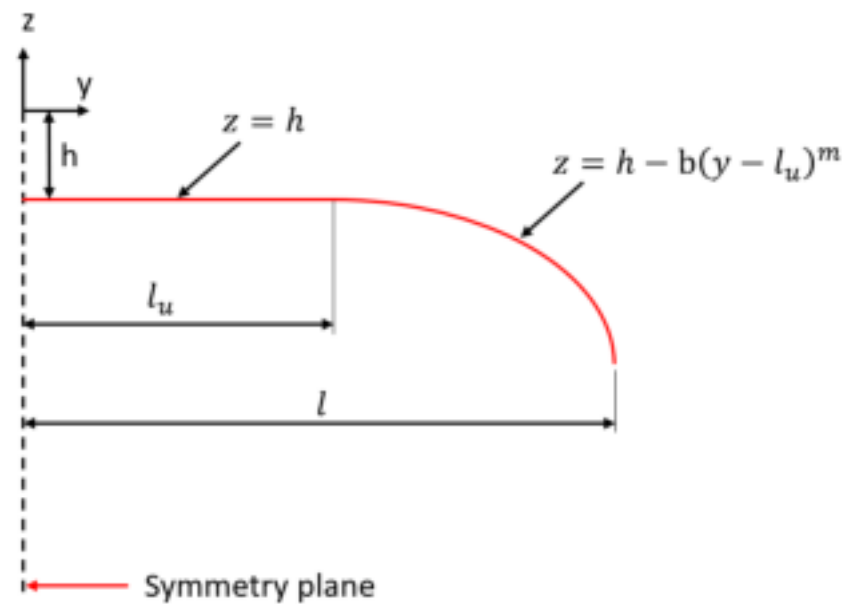


Figure 2. Schematic representation of the waverider design methodology.

There are two primary methods for blunting the leading edge of a waverider: one involves removing material from the edge, while the other entails adding material to it [35]. Tinch and Burnett have proposed that the addition of material to the leading edge incurs a lower aerodynamic penalty compared to material removal [35]. This approach is favored because it allows for maintaining more of the original aerodynamic profile, potentially minimizing the negative impacts on performance that are associated with leading edge modifications.

3. Direct Simulation Monte Carlo (DSMC) Methodology

The Direct Simulation Monte Carlo (DSMC) method, initially developed by Graeme Bird in the late 1960s [36], employs a particle-based statistical approach to study rarefied gas flows. Unlike traditional methods that solve equations to predict flow behavior, DSMC simulates the motion and interactions of a large number of simulated particles, each representing a fraction of the actual molecules in the flow field. This method decouples the motion of the particles and computes intermolecular collisions over brief time intervals. It has been mathematically demonstrated that with a sufficient number of particles per cell, DSMC converges to the Boltzmann equation [37]. The DSMC technique comprises four main steps:

1. Advancing the position of the particles.
2. Sorting the particles into respective cells.
3. Calculating the intermolecular collisions.
4. Gathering and analyzing the particle properties.

In practice, the computational domain is segmented into numerous cells, potentially further divided into sub-cells based on a predefined grid. Initially, the particles are randomly assigned within each cell, with specific initial positions, velocities, and energy states. These steps are cyclically repeated for each timestep to model the time evolution of the flow. Macroscopic flow properties like velocity, pressure, density, and temperature are derived from the weighted averages of these microscopic particle properties. For all the flow simulations discussed in this study, the DSMC method as detailed by Graeme Bird [36] and as implemented in the Stochastic Parallel Rarefied-gas Time-accurate Analyzer (SPARTA) at Sandia National Laboratories [38], was utilized to examine the flow characteristics.

4. Application to a Specific Waverider Design

In the development of a waverider designed for Mach 7, the initial step involved crafting a 7-degree cone, 2 m in length with a 1 cm blunt leading edge, as depicted in Figure 3. The three-dimensional flow field around this cone was then resolved using the Direct Simulation Monte Carlo (DSMC) method. The simulation conditions were derived from the US standard atmosphere model, representing the flow characteristics at an altitude of 90 km. The specific flow conditions and the parameters used in the DSMC simulation are detailed in Tables 1 and 2, respectively. In the present simulations, the gas comprises 79% N_2 and 21% O_2 ; rotational and vibrational relaxation numbers are treated as constants. The rotational relaxation number was set to $Z_{rot} = 5$ for both N_2 and O_2 , whereas a vibrational relaxation number $Z_{vib} = 52,600$ for N_2 and $Z_{vib} = 17,900$ for O_2 have been used. Elastic collisions are performed using the Variable Soft Sphere (VSS) model [39] while inelastic collisions are performed using the Borgnakke–Larsen model [40]. For the selection of collision partners, we utilized the near-collision partner algorithm. This algorithm selects collision partners within a sphere whose radius is the distance a particle travels in one timestep. This approach enables the use of a grid cell size larger than one mean free path while maintaining collision accuracy [41] although our cell size is close to 1/3 of the mean free path. Due to the low temperatures occurring in these flight conditions, chemical reactions are not enabled; thus, the flow is assumed to be in thermal equilibrium. All the simulations were conducted in the ARCHER2 UK supercomputing facility.

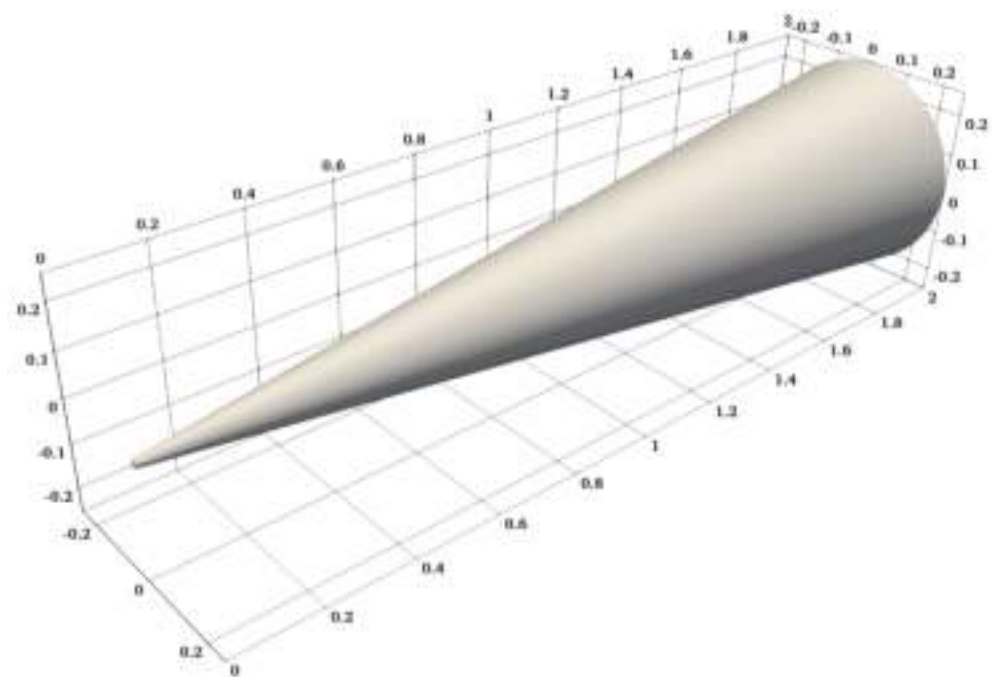


Figure 3. The geometry of the 7-degree half cone used for the calculation of the initial flow field.

As it was previously mentioned, all the simulations were carried out using the open-source parallel DSMC code SPARTA (Stochastic Parallel Rarefied-gas Time-accurate Analyzer), which was developed at Sandia National Laboratories [38]. After computing the three-dimensional flow field around the 7-degree cone, the flow streamlines were extracted at specific parallel planes to delineate the shockwave boundaries. Figure 4 offers a visualization of these calculated flow field streamlines. The shock boundaries are discernible by the bending of the flow streamlines in both the side and top planes, and by a change in the color of the streamwise velocity flow field at the rear plane behind the cone.

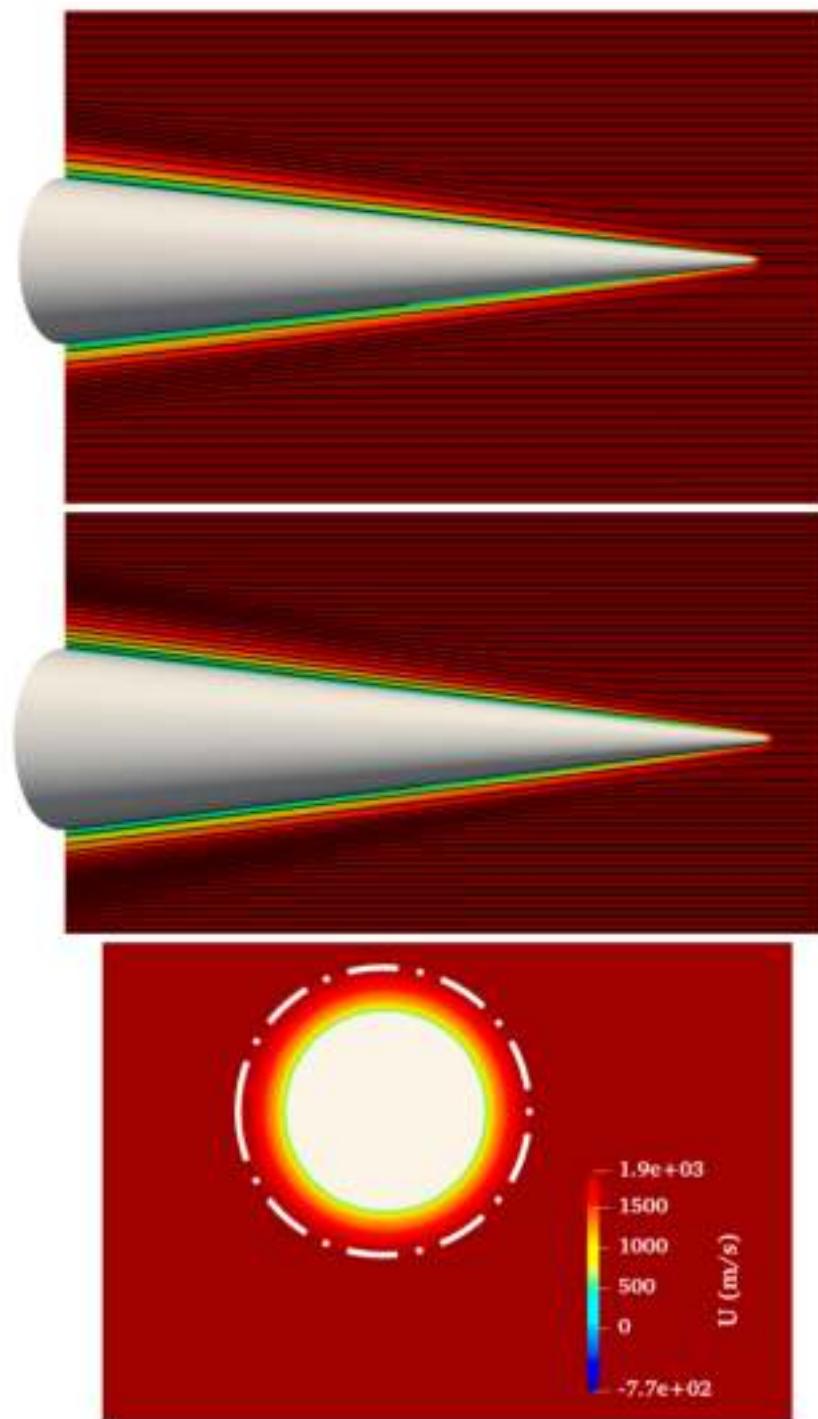


Figure 4. Streamlines of the three-dimensional flow field around the 7-degree cone. Side view (**top**), top view (**middle**), rear view (**bottom**).

After obtaining the flow planes from the DSMC simulation, these are imported into a Computer-Aided Design (CAD) software (SolidWorks 2021) to precisely define the geometry of the waverider. The design process begins at the rear of the vehicle, crafting curves that are tangent to the observed circular shock shape, with the first characteristic cross-section displayed in Figure 5 (bottom) and Figure 6. As seen in Table 1, due to the low Reynolds number the flow falls within the laminar regime.

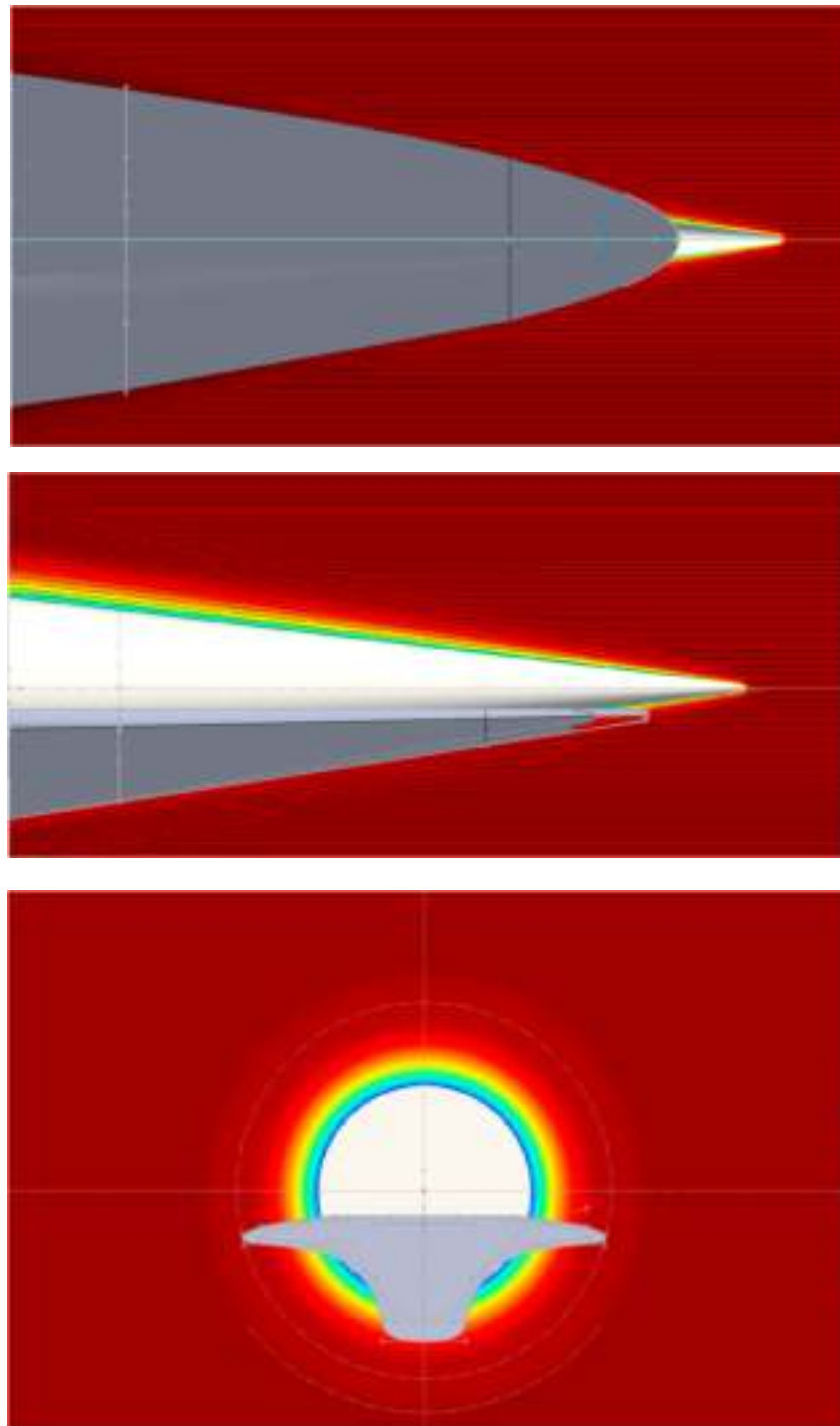


Figure 5. Waverider surface along the flow streamlines, in comparison with the initial cone. Top view (**top**), side view (**middle**), and rear view (**bottom**).

To ensure the design remains practical for real-world applications, the leading edges of the waverider are fashioned to be blunt rather than sharp. To accurately control the bluntness of the leading edges, a four-point Bezier curve is utilized to sculpt the leading-edge profile. This method provides the flexibility to finely adjust the curvature and thus tailor the aerodynamic and thermal properties of the edge. Similarly, additional sections along the waverider's longitudinal axis are designed at intervals of every 30 cm. Both the upper and lower surfaces of the waverider are defined by the shockwave boundaries generated from the initially prescribed 7-degree cone geometry. This approach ensures

that the final waverider design is optimized within the constraints set by the initial flow conditions and the aerodynamic needs of hypersonic travel.

Table 1. Parameters for a Mach 7 flow at 90 km altitude.

Mach number, Ma_∞	7
Reynolds, Re_∞	2873
Freestream velocity, U_∞ (m/s)	1965
Knudsen number, Kn	0.6
Angle of attack, α (degrees)	0
Freestream density, ρ_∞ (kg/m ³)	9.62×10^{-6}
Freestream pressure, p_∞ (Pa)	0.55
Freestream temperature, T_∞ (K)	196
Wall temperature, T_w (K)	300
Gas	Air

Table 2. DSMC computational parameters for Mach 7 flow at 90 km altitude.

Number density, Nd_∞ (#/m ³)	2.00×10^{20}
Timestep (s)	3.0×10^{-7}
Transient period	40,000
Sample period	100,000
Number of particles	10.5×10^9
Number of cells	183,664,000
Wall-clock time (CPU hours)	195,840

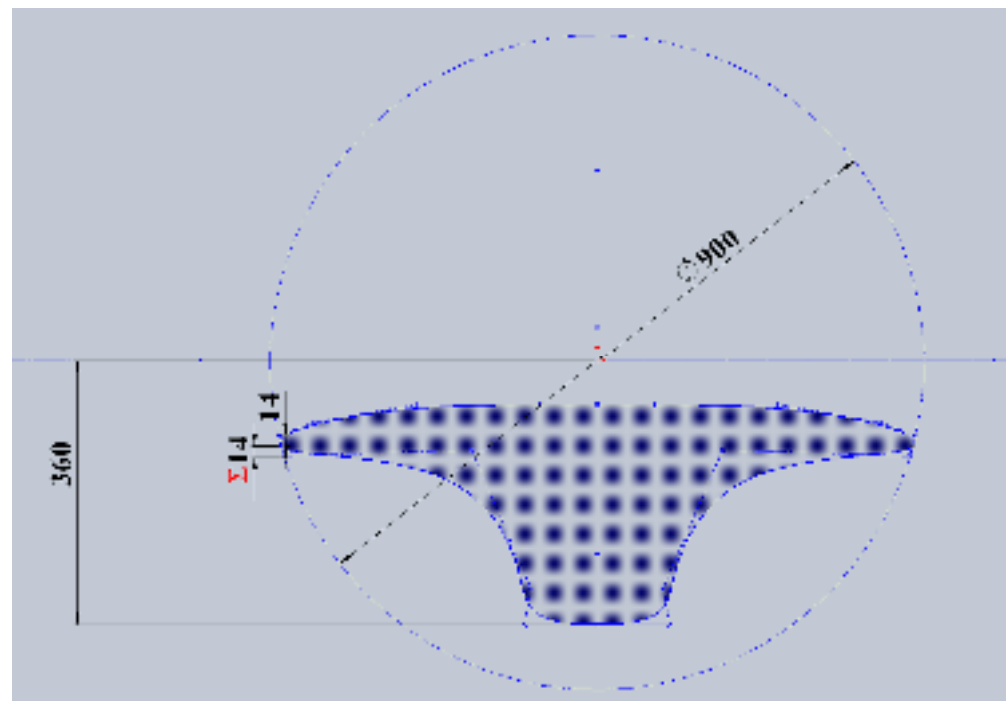


Figure 6. Back section of the waverider (units in mm).

Figure 7 provides a visual representation of the various sections of the waverider geometry, although not all the sections are displayed to maintain clarity. Upon completing the design of all the sections, a loft operation with tangency constraints is employed to generate the surfaces of the waverider. To simplify the design process and leverage the symmetry of the flow field, only half of the final geometry is initially designed. This half is then mirrored around the symmetry plane to produce the complete vehicle structure.

The methodology of the osculating cones, as described earlier, serves as a robust computational framework for designing hypersonic waveriders. Hypersonic aerodynamics involves dealing with extreme temperatures, high pressures, and intricate shock wave interactions, making the design process inherently complex. By breaking down the methodology into discrete steps, this approach allows for a detailed focus on both the theoretical and practical aspects, guiding the creation of these advanced hypersonic vehicles. This structured design process ensures that each aspect of the vehicle's performance and structural integrity is considered, leading to the development of effective and efficient hypersonic waveriders.

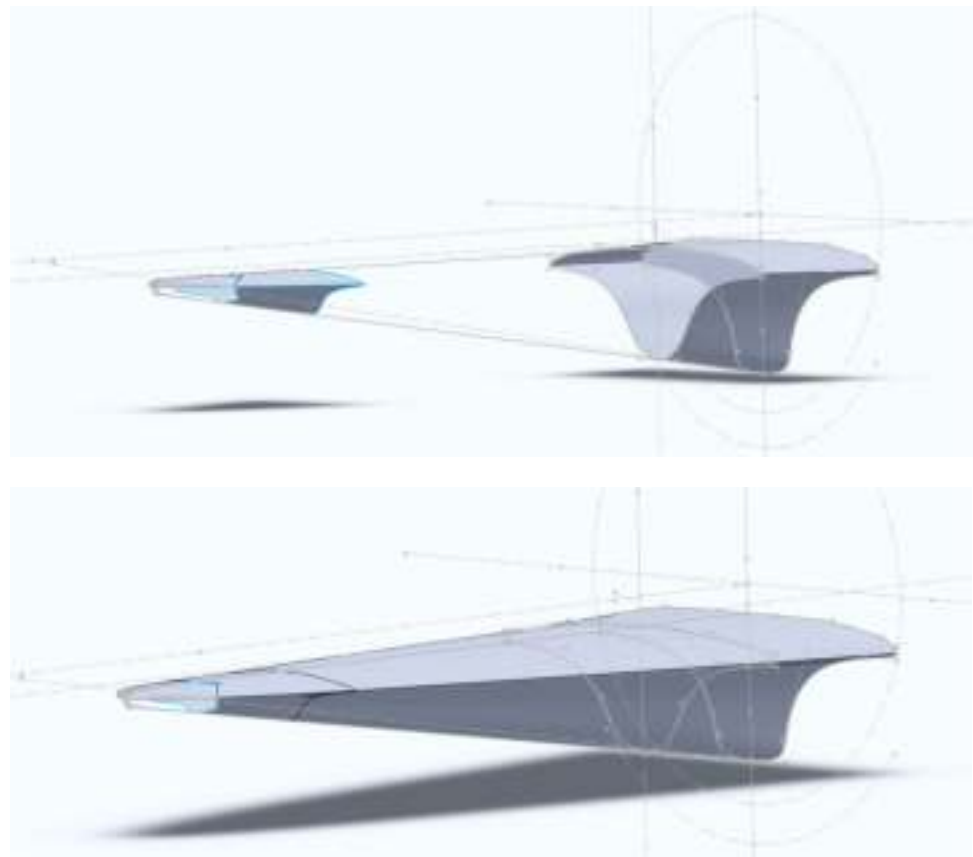


Figure 7. Surface lofts along the vehicle profiles (**top**). Vehicle overview without the nose section (**bottom**).

The design process for a hypersonic waverider involves a meticulous and structured approach, crucial for achieving aerodynamic efficiency and structural integrity under extreme conditions. Here is a detailed breakdown of each step in the design procedure:

Step 1: Initiating at the Aft End of the Cone

The design process begins at the trailing end of a cone, which is pivotal for setting the initial conditions of the flow field. Here, the shock wave boundaries are meticulously defined, shown as dashed white lines in Figure 4 (bottom). This foundational step is critical as it establishes the first profile of the waverider. The design ensures that both the leading edge and the lower surface of this profile are tangentially aligned with the shock wave boundaries, minimizing drag and enhancing aerodynamic efficiency.

Step 2: Sequential Cross-section Profiling

Moving forward, the cross-sectional planes are taken at intervals of about 30 to 40 cm along the cone's length. Each plane provides a template for shaping a segment of the vehicle's body, tailored to the specific aerodynamic conditions of that segment. This step

is akin to sculpting the individual slices of the vehicle, each optimized for its particular flow dynamics.

Step 3: Lofting Operation

The individual profiles generated in Step 2 are then interconnected through a lofting operation. This technique merges the profiles to form the continuous outer skin of the waverider, as seen in Figure 7. The operation uses interpolation methods to ensure a smooth contour transition between sections. However, the profiles narrow as they approach the cone's apex, which presents challenges in maintaining the structural and aerodynamic continuity.

Step 4: Addressing the Nose Region

In the nose region, where profiles become exceedingly slender, traditional lofting techniques are less effective. To address this, boundary surfaces that conform to the curvature of the original cone are introduced. As depicted in Figure 8, these surfaces form the nose section's outer skin, integrating seamlessly with the lofted sections to maintain aerodynamic integrity.

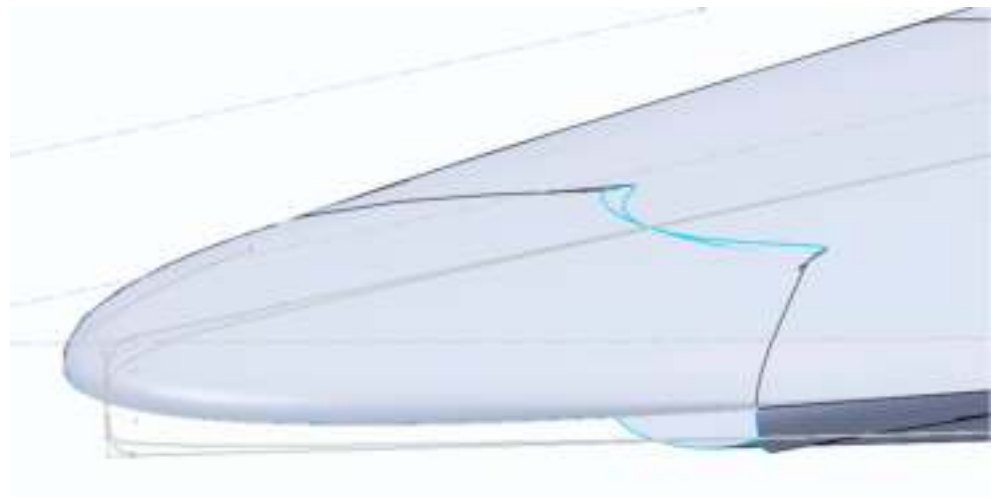


Figure 8. Nose section with upper boundary surface.

Step 5: Knitting Surfaces for Fluid Simulation

Following the design of the nose and main body sections, a knitting operation is employed to merge all the surface segments into a unified, watertight structure. This integration is crucial for the accuracy of subsequent fluid dynamics simulations, ensuring no gaps or inconsistencies exist that could compromise the model's stability or precision. The completion of this step results in a complete vehicle geometry, illustrated in Figure 9.

Step 6: Mirroring Operation

To enhance design efficiency and reduce computational demands, only half of the vehicle is modeled initially. The other half is created through a mirroring operation along a symmetry plane. This strategy not only simplifies the construction process but also facilitates faster iterations and refinements during the design phase, effectively halving the computational resources required.

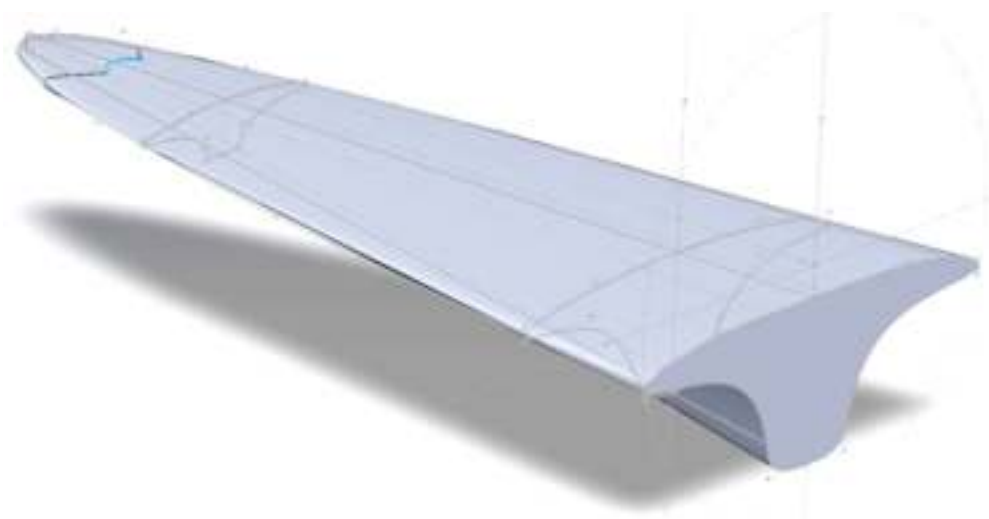


Figure 9. (Top): Waverider sections and the complete geometry. (Bottom): Flowchart of the design methodology.

Step 7: Refined Leading-Edge Design

This step in the waverider design methodology underscores the importance of optimizing the leading edges, which play a critical role in the vehicle's overall performance and survivability in hypersonic conditions. Unlike the typical hypersonic designs that feature sharp leading edges to minimize drag and maximize speed, this approach recognizes the trade-offs such sharp edges bring, including high stress concentrations and reduced thermal resilience. These factors can lead to structural failures when subjected to the intense heat and pressure of hypersonic flight.

To address these challenges, the design incorporates three-point Bezier curves to blunt the leading edges. This technique allows for precise control over the curvature and tangency of the edges, making it possible to tailor them specifically to the demands of high-speed aerodynamics and thermal stress management. The Bezier curves not only help in smoothing the transition of the flow over the vehicle but also enhance the thermal robustness by distributing the stresses and heat more evenly across the surface. The use of Bezier curves in this context provides a balance between aerodynamic efficiency and structural integrity, ensuring the leading edges are not only effective in minimizing aerodynamic drag but are also capable of withstanding the harsh conditions of hypersonic flight. This nuanced approach results in a leading-edge design that is both geometrically optimized and thermodynamically durable, setting a new standard for hypersonic vehicle design. Through these steps, the waverider design is developed with precision, addressing both the high demands of hypersonic travel and the practical aspects of vehicle manufacture and simulation.

The osculating cones methodology is a sophisticated framework that draws from the fields of aerodynamics, computational fluid dynamics (CFD), and applied mathematics. This approach forms a holistic and in-depth pathway for designing hypersonic waveriders, effectively addressing the complex challenges associated with hypersonic flows and structural dynamics. The integration of these multi-disciplinary domains ensures that the resulting waverider design achieves not only aerodynamic efficiency but also exhibits enhanced thermal and structural robustness.

In the first phase of this computational exploration, the creation of an accurate computational mesh is crucial. This mesh is instrumental in precisely capturing the intricate geometric details of the waverider, particularly in the areas of high sensitivity, such as the leading edges. The precision in modeling these edges is paramount because even slight inaccuracies can lead to significant aerodynamic repercussions. This meticulousness in the mesh design is essential for replicating the intended geometrical features and ensuring that the simulations reflect the true aerodynamic behavior of the waverider. The computational surface mesh used in the Direct Simulation Monte Carlo (DSMC) method is specifically tailored to accommodate the unique characteristics of the waverider's design, as illustrated in Figure 10. This mesh allows for the detailed and accurate simulation of the flow dynamics around the waverider, capturing the essential interactions between the vehicle's surface and the hypersonic flow. The DSMC method, known for its ability to model rarefied gas dynamics effectively, is particularly well-suited for analyzing the aerodynamic properties of vehicles operating at the upper edges of the atmosphere, where conventional CFD methods might struggle. For the simulation of the final waverider design, we used the SPARTA DSMC code, consistent with the simulation of the initial shock-generating cone. This consistency was maintained because the flight conditions did not change. As shown in Figure 11 (bottom) the Knudsen number falls within the upper end of the slip regime or within the transition regime.

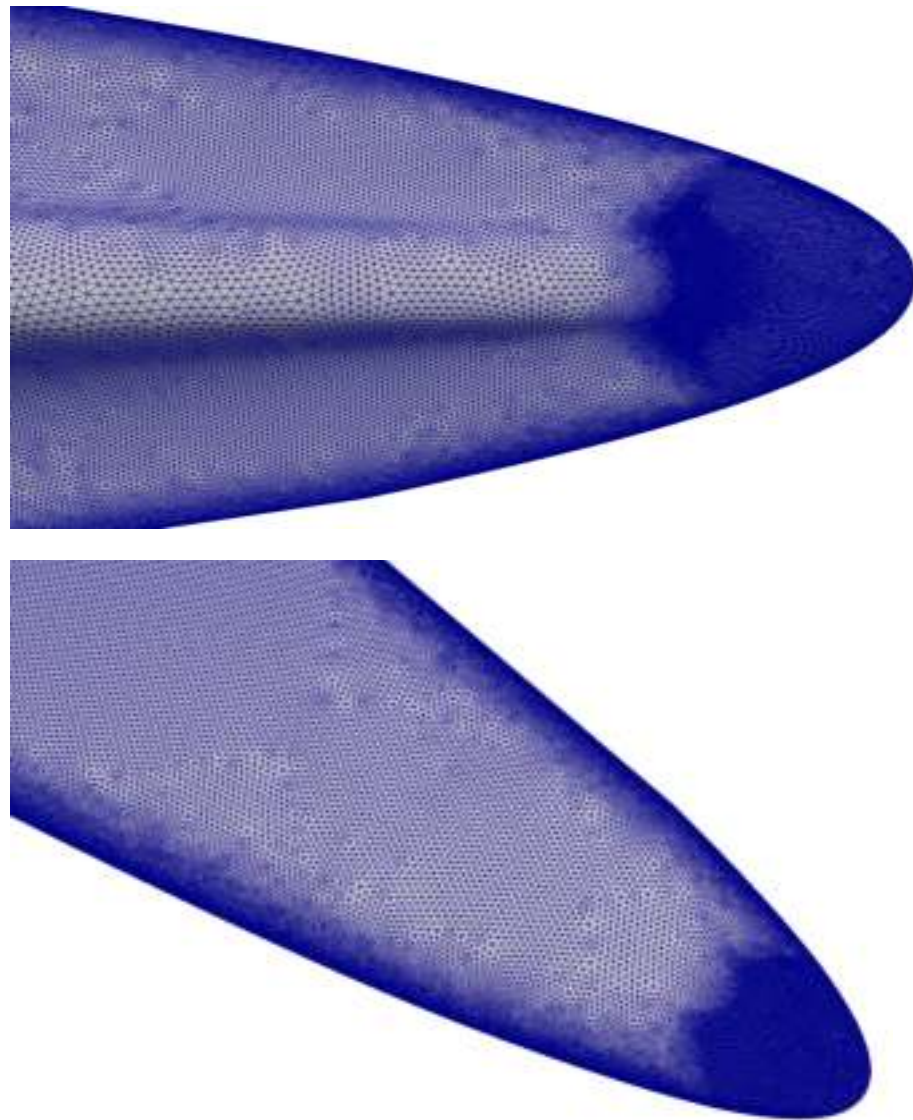


Figure 10. The utilized surface mesh. Lower surface (**top**) and isometric view (**bottom**).

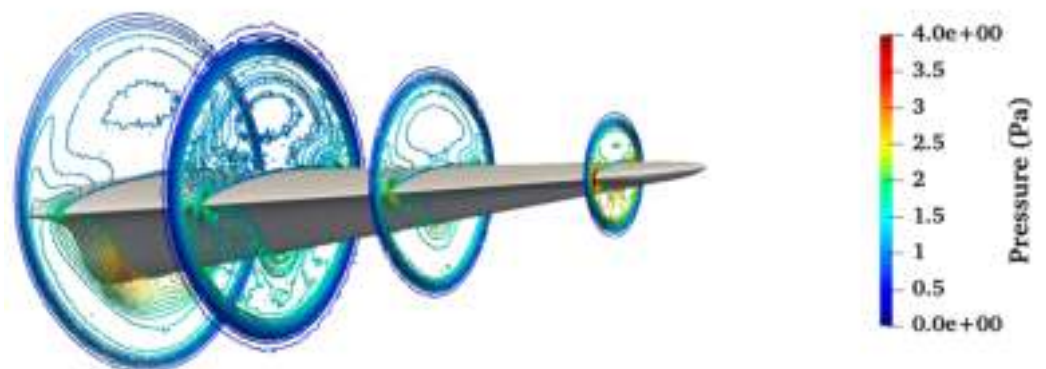


Figure 11. *Cont.*

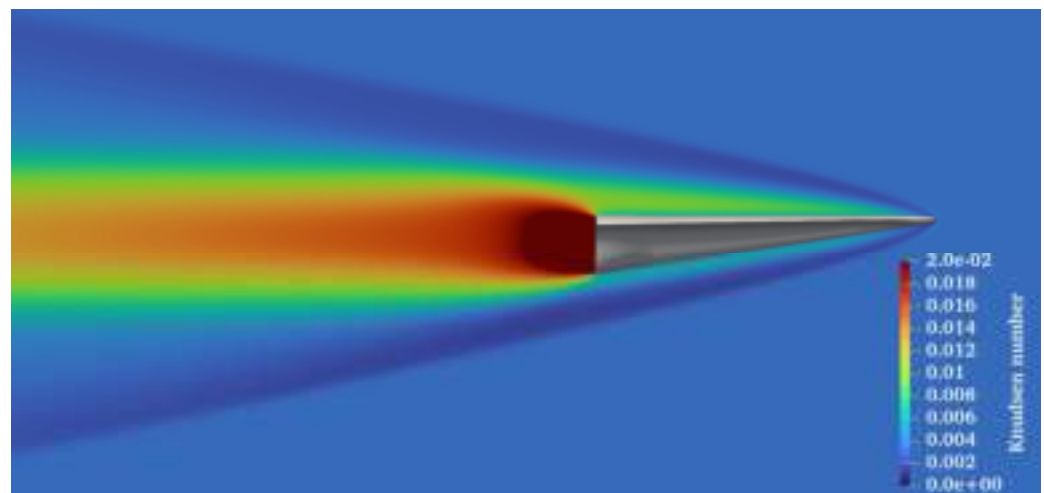


Figure 11. Pressure contours around the vehicle (**top**), Knudsen number of the flow field based on the waverider length (**bottom**).

5. Simulation Results

Due to the extensive computational demands associated with simulating hypersonic flows, several strategies are implemented to optimize the use of computational resources. One effective technique is the simulation of only half of the waverider's geometry, with the resulting flow field then being symmetrically mirrored across a predefined plane. This approach significantly reduces the computational load without sacrificing the accuracy of the results. A key objective of these simulations is to confirm the correct positioning of the shock wave relative to the vehicle's leading edges. This verification is conducted by analyzing the pressure contours and Q-criterion plots, which are displayed in Figure 11 and Figure 12, respectively.

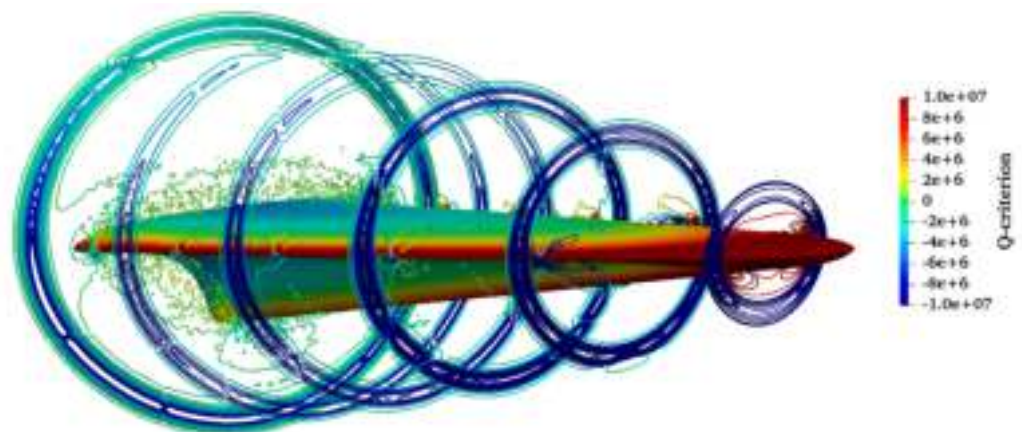


Figure 12. Q-criterion contours around the vehicle.

The Q-criterion, calculated using Equation (2), involves Ω the antisymmetric component of the velocity gradient representing vorticity, and S , the symmetric component representing the rate of strain.

$$Q = \frac{1}{2} (\|\Omega\|^2 - \|S\|^2) \quad (2)$$

According to the established criteria, positive Q-criterion values signify regions within the flow field where vorticity predominates, suggesting the areas of rotational flow. Conversely, negative values indicate the dominance of strain rate or viscous stresses [42,43]. This criterion is a critical diagnostic tool for analyzing the flow characteristics around

the waverider. For instance, in Figure 13, the negative Q-criterion values observed below the leading edge signal the presence of a shock in that area. The associated pressure contour plots, illustrated in Figure 11, reveal a distinct pressure buildup around the leading edges of the waverider, confirming the effective attachment of the shock wave to the vehicle's geometry.

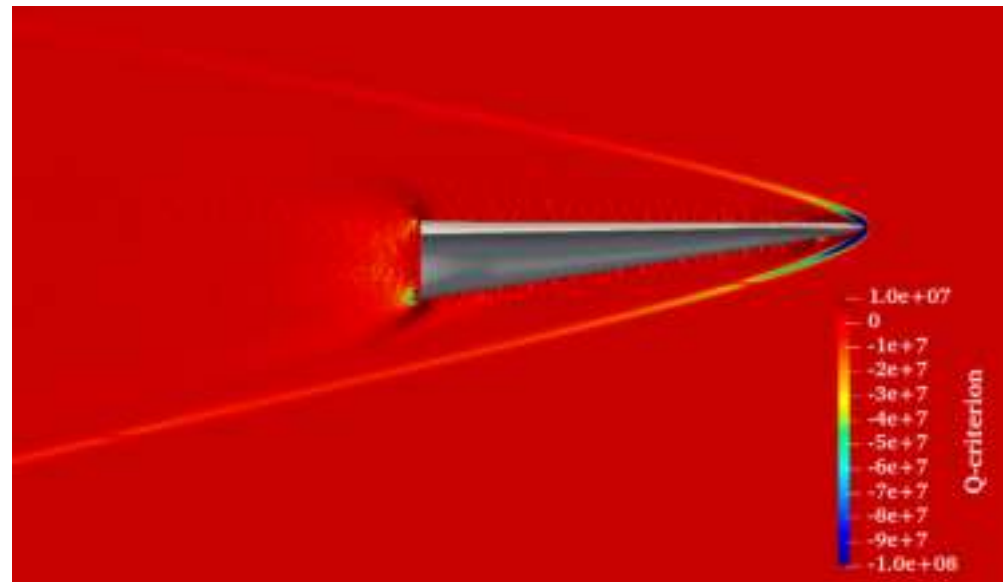


Figure 13. Q-criterion contours on the plane of symmetry.

Moreover, Figure 11 graphically depicts the pressure distribution on the surface of the waverider. As anticipated, increased pressure levels are prominent around the leading edges and nose section, validating the presence of an attached shock wave. This integration of various computational diagnostics—from mesh optimization to detailed pressure and Q-criterion analyses—offers a thorough assessment of the waverider's aerodynamic performance. These results underscore the computational rigor and physical validity of the waverider design, particularly regarding the attachment of shock waves to the leading edges. Figure 13 depicts the Q-criterion contours on the plane of symmetry, while Figure 14 shows the corresponding contours on a vertical plane at the rear of the vehicle, where the concentrated blue regions at the wing tips and under the fuselage indicate attached shocks, further demonstrating the aerodynamic efficacy of the design.

Figures 15 and 16 present the streamwise velocity contours and the temperature field around the vehicle. These figures are crucial for understanding how the vehicle interacts with the flow at high speeds. In particular, the simulations reveal important thermal dynamics and fluid flow characteristics. For the specific flow conditions applied in this study, the temperature field around the vehicle shows significant thermal variations. A peak temperature of 1400 Kelvin is observed just in front of the nose section of the waverider, indicating areas where aerodynamic heating is most intense due to the rapid flow deceleration. This information is vital for assessing the thermal resistance and structural integrity of the waverider under operational conditions.

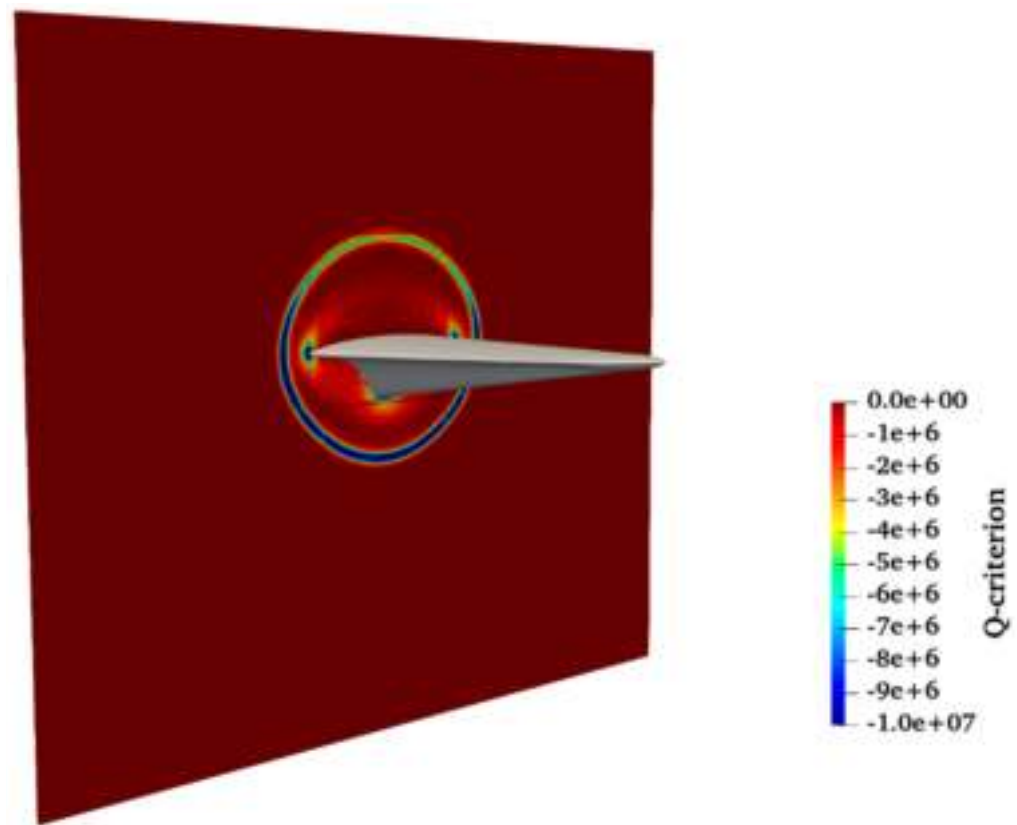


Figure 14. Q-criterion contours on a vertical plane.

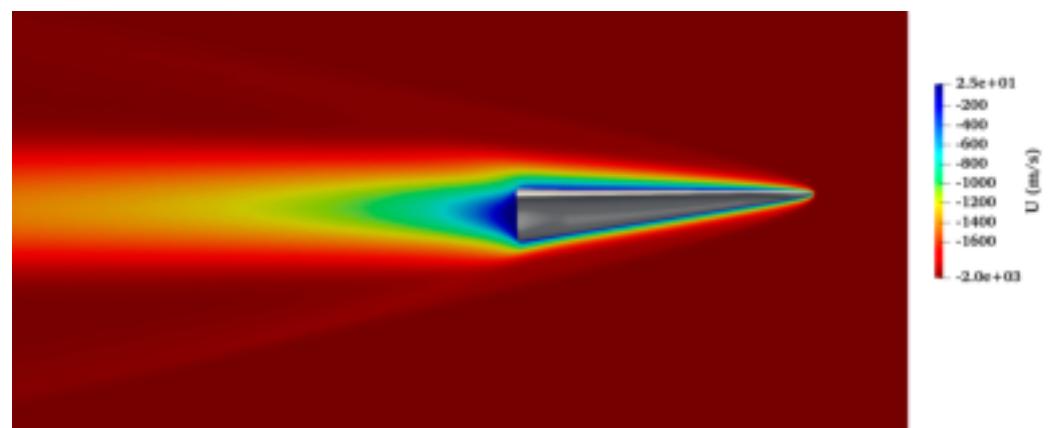


Figure 15. Streamwise velocity contours on the symmetry plane of the waverider.

In Figures 17–20, various flow properties are illustrated on a plane parallel to the waverider, with the surface of the vehicle distinctly highlighted by the variations in the surface pressure. Specifically, Figure 17 displays the pressure contours on this plane, providing a clear visualization of how the pressure varies in response to the shape and aerodynamics of the vehicle. Notably, considering that the free-stream pressure is established at 0.55 Pa, there is a significant pressure increase observed around the nose of the waverider, rising to an order of magnitude higher than the ambient conditions. Additionally, the pressure around the leading edges of the waverider is approximately four times higher than that of the free-stream pressure. This detailed mapping of pressure distribution is critical for evaluating the aerodynamic performance and structural stresses experienced by the waverider during operation.

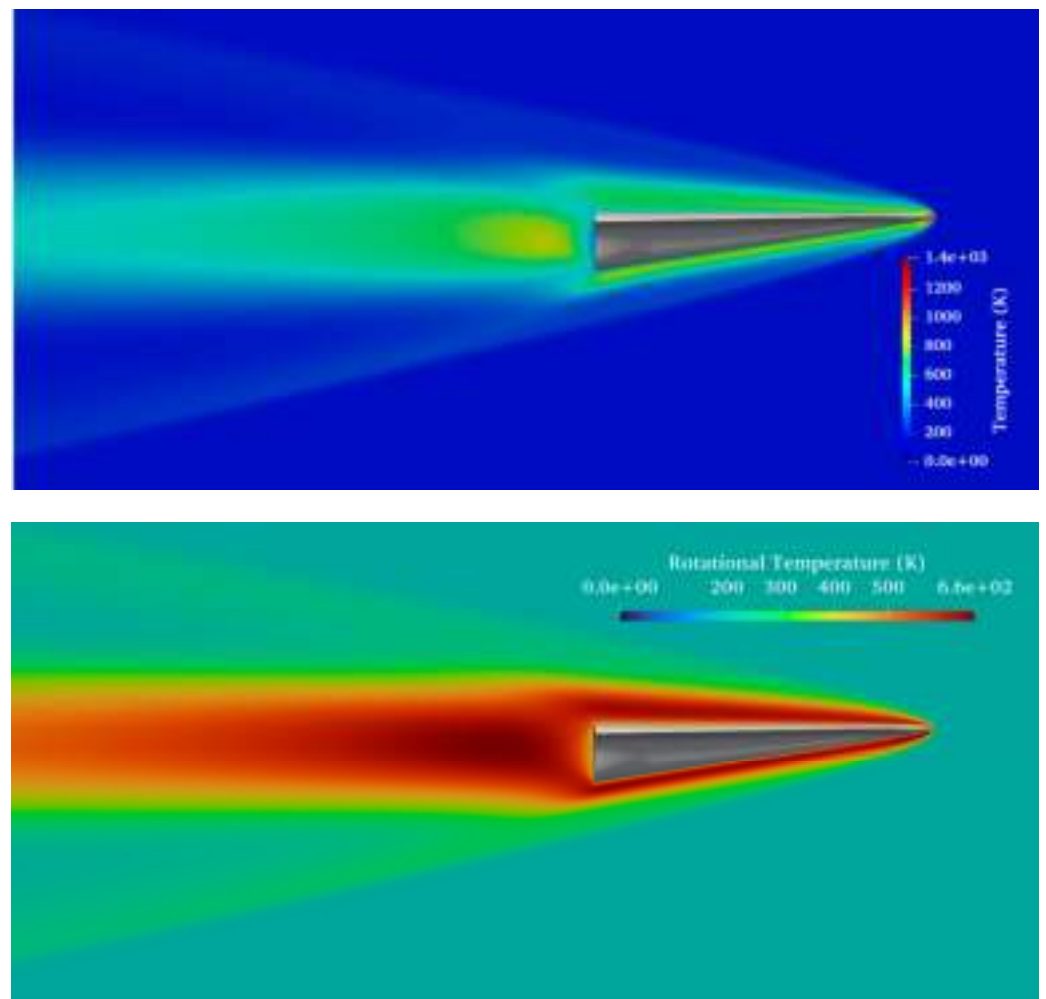


Figure 16. (Top): total temperature field around the vehicle. (Bottom): rotational temperature field (on the symmetry plane of the waverider).

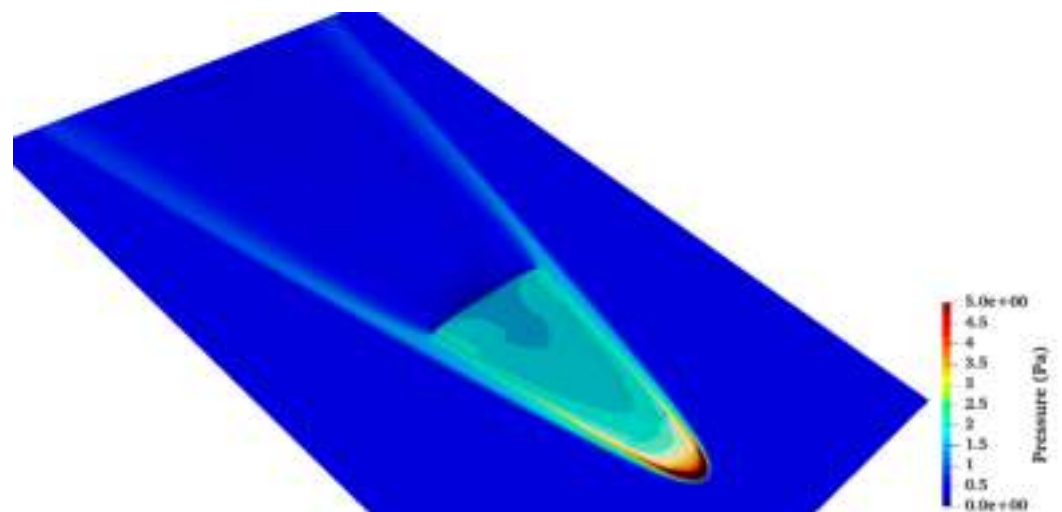


Figure 17. Pressure contours on a horizontal plane parallel to the vehicle.

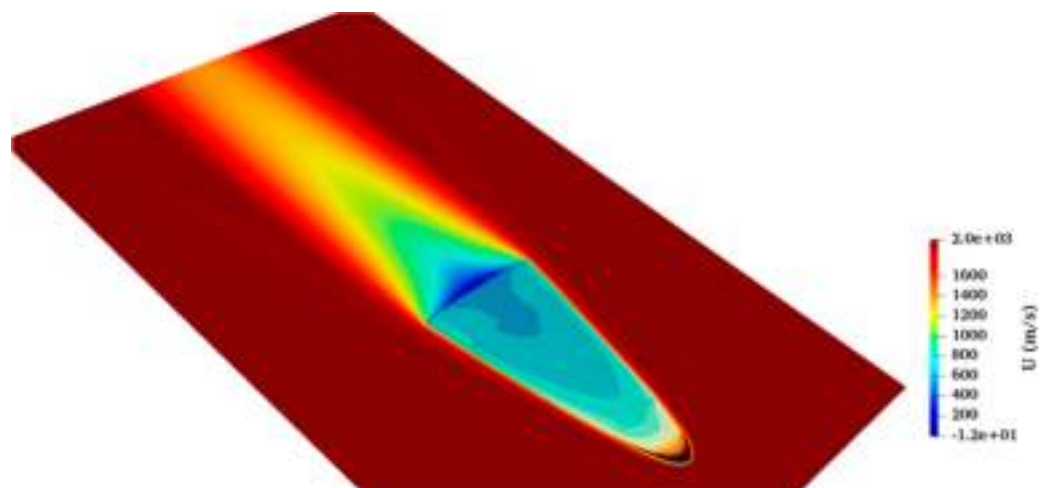


Figure 18. Streamwise velocity contours on a horizontal plane parallel to the vehicle.

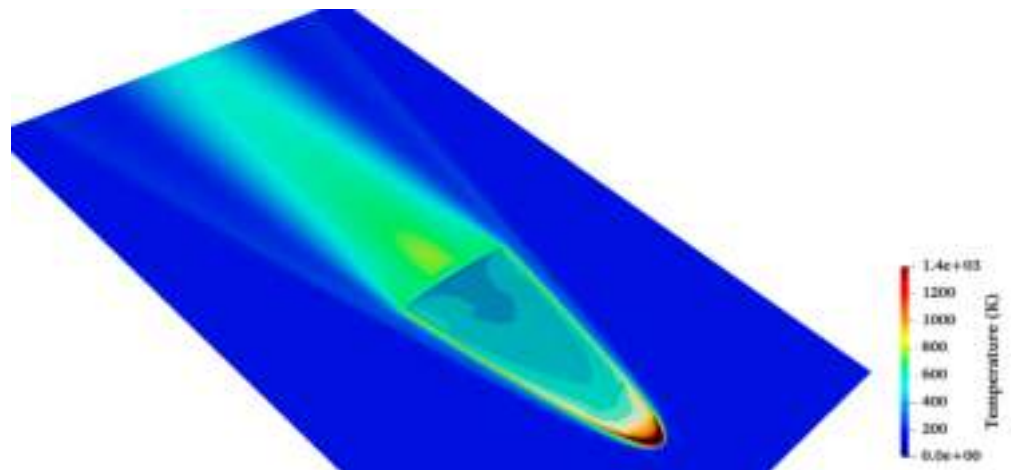


Figure 19. Temperature contours on a horizontal plane parallel to the vehicle.

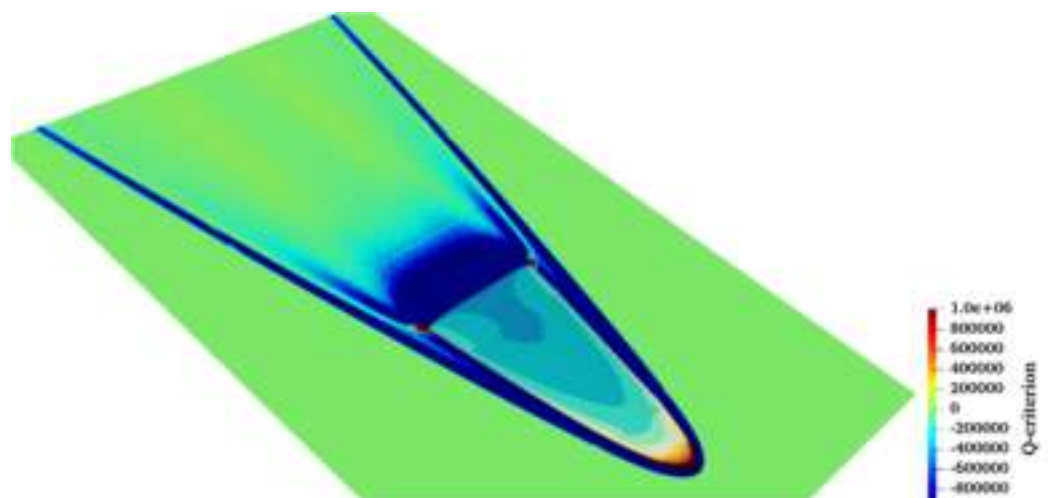


Figure 20. Q-criterion contours on a horizontal plane parallel to the vehicle.

Figure 18 contains the velocity contours on the same plane. In this figure, the attachment of the shock to the leading edge and the nose of the vehicle can be clearly observed. As can be observed in Figure 19, the temperature around the leading edges is 800 K, which is about four times higher than the free-stream temperature of 196 K (with a peak of around 1400 K at the nose of the vehicle).

In Figure 20, the Q-criterion is meticulously illustrated on the same plane parallel to the waverider, effectively capturing the entire field of the attached shock wave. This visualization underscores the precise areas where the shock remains attached to the surface of the waverider, a critical aspect of its aerodynamic design. Additionally, the positive values of the Q-criterion in this figure indicate the formation of two distinct vortices, located symmetrically, one on each side, at the tip of the vehicle's rear section. These vortices are pivotal in understanding the flow dynamics at the rear of the waverider, particularly how they influence the overall stability and control of the vehicle at high speeds. For a more detailed observation of these vortices, Figure 21 provides an enhanced view where the vorticity at the rear of the waverider is plotted. In this figure, the two vortices are distinctly highlighted in red, offering a stark contrast against the rest of the flow field. This color differentiation not only emphasizes the location and intensity of the vortices but also aids in analyzing their structure and interaction with the surrounding flow. The depiction of these vortices in such vivid detail is crucial for engineers and designers to evaluate the aerodynamic interactions at play and to refine the waverider's design for optimized performance.

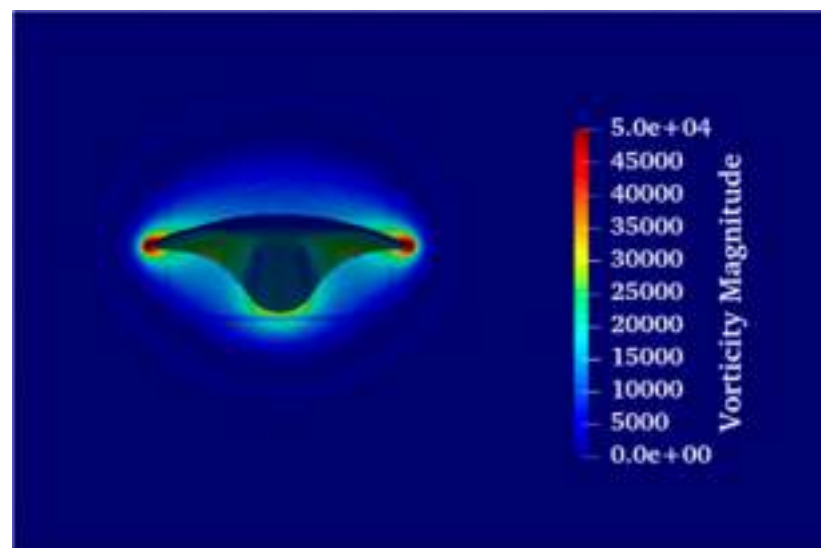


Figure 21. Vorticity magnitude at the back of the vehicle.

Given the intricate design of the waverider, it was deemed necessary to employ the three-dimensional iso-surfaces of the Q-criterion to fully capture the complex spatial characteristics of the shocks and the structured vortices surrounding the vehicle. Figure 22 illustrates these three-dimensional Q-criterion contours enveloping the vehicle, with the contours colored by vorticity magnitude to emphasize the variations in the rotational flow behavior. The transparent grey iso-surfaces distinctly outline the external shape of the shock waves, providing a comprehensive view of their three-dimensional form. Notably, parts of the shock wave close to the compression surface are also visible in the same figure, offering a detailed look at how the shock interacts with the vehicle's structure.

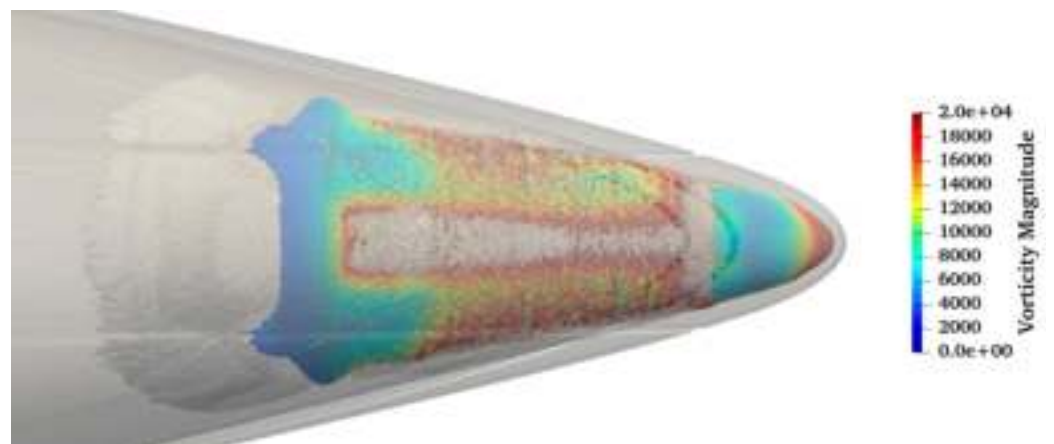


Figure 22. Three-dimensional Q-criterion contours, colored by vorticity magnitude.

In Figure 23, these contours are similarly plotted but are instead colored by velocity magnitude to highlight the differences in flow speed. As depicted in this figure, approximately 30 cm downstream from the nose of the vehicle—where the compression surface becomes steeper—the vortices detach from the main flow and continue downstream towards the higher pressure regions at the back of the vehicle. This separation is critical for understanding the aerodynamic behavior of the waverider. Additionally, as the angle of the compression surface increases towards the vehicle's rear, a noticeable pressure rise occurs, and the flow velocity significantly reduces, slowing down by about 1000 m/s within just a few centimeters. These detailed visualizations and analyses of flow dynamics are essential for assessing the performance implications of the vehicle's design under operational conditions.

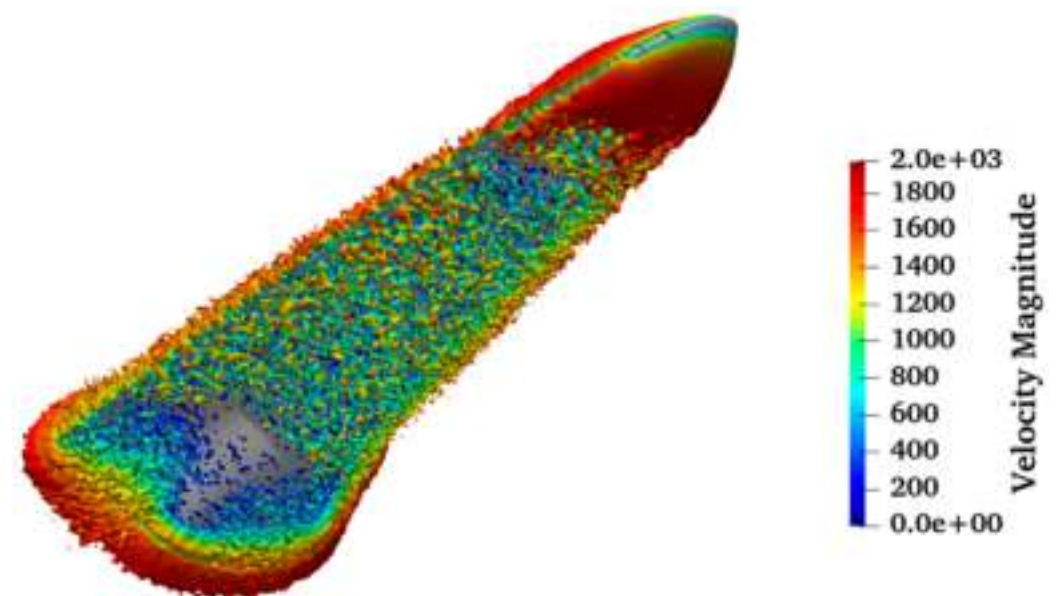


Figure 23. Overview of the three-dimensional Q-criterion contours around the vehicle, colored by velocity magnitude.

To determine the aerodynamic forces acting on the waverider, the shear stresses and pressure components across its three-dimensional surface were meticulously calculated. These calculations facilitate the precise measurement of lift and drag forces per unit surface area, with the results displayed in Figure 24. The analysis of this figure reveals that the maximum drag primarily occurs at the tip of the nose, while a significant amount of lift is generated in the area just below the nose. By integrating the forces measured across all the

surface cells, the total lift and drag for the waverider were computed, yielding a lift–drag ratio (L/D) of 2.18. According to existing studies, this L/D ratio is typical for a waverider operating at a 0-degree angle of attack [44], indicating that a value above 2.0 is indicative of effective aerodynamic performance under the chosen design conditions.

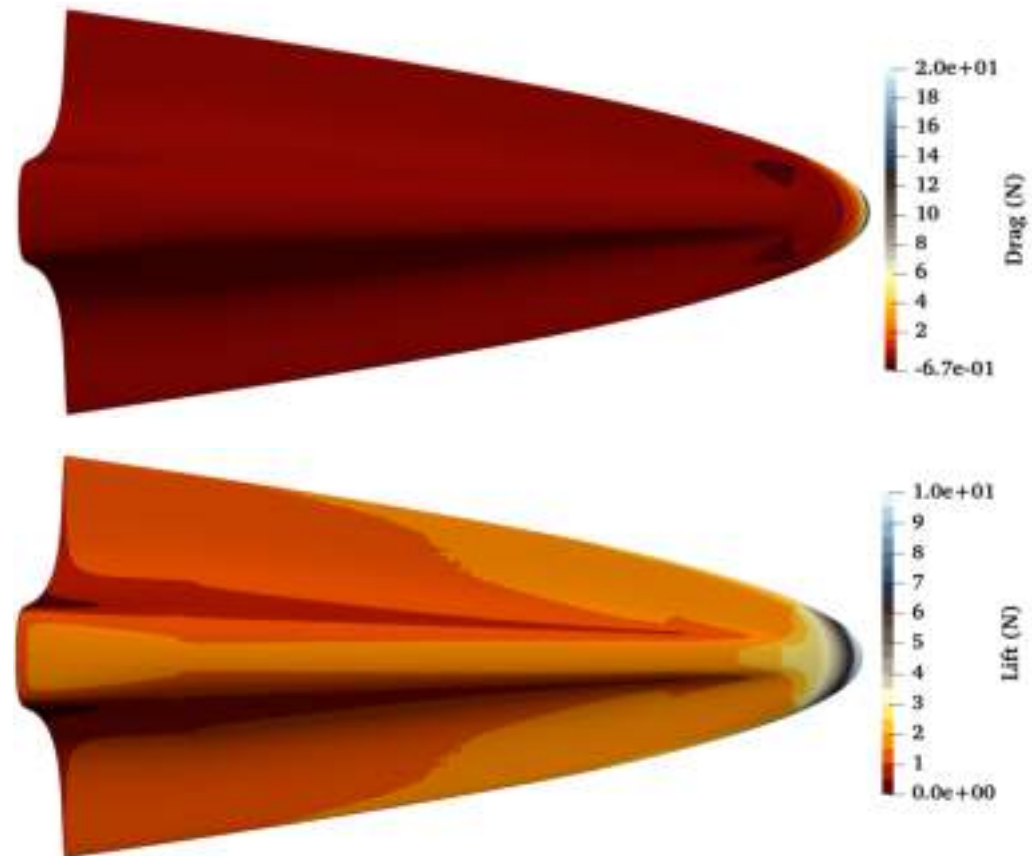


Figure 24. Overview of the lift (**bottom**) and drag (**top**) per unit surface, exerted on the waverider's lower surface.

The Knudsen number contours, based on the waverider length, are depicted in Figure 11 (bottom), while the Mach number contours at the plane of symmetry are depicted in Figure 25. As demonstrated the Knudsen number around the vehicle is within the transitional regime, which justifies the use of the DSMC method.

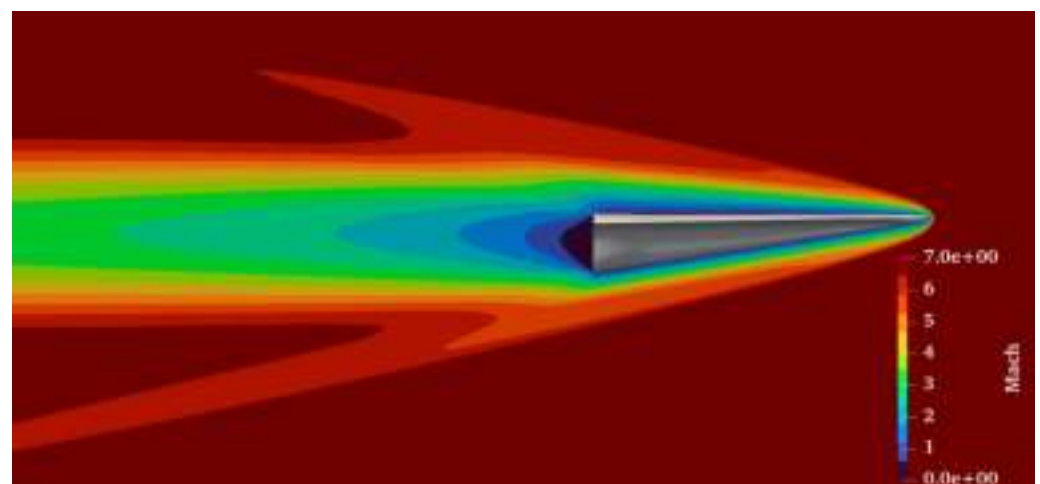


Figure 25. Mach number contours around the vehicle at the plane of symmetry.

For this study, only a single angle of attack was investigated; however, the authors in [44] suggest that a slight increase in the angle of attack could potentially enhance the L/D ratio by two to three times, thereby significantly improving performance. It is also noteworthy that waveriders are commonly analyzed at altitudes between 25 km and 40 km in the open literature, unlike the 90 km altitude considered in this study. The elevated altitude impacts the generation of lift due to different fluid dynamics, necessitating a larger compression surface than what is typically seen in waveriders operating at lower altitudes. This adaptation is a critical design modification that enhances the vehicle's performance in the rarified atmospheric conditions encountered at high altitudes.

6. Conclusions

In the comprehensive design and analysis of a Mach 7 waverider, intricate computational methodologies were employed to accurately simulate the vehicle's aerodynamic properties. After applying the osculating cone methodology, for the definition of the waverider's geometry, a precise computational mesh was constructed to capture the complex surface geometry, particularly around the waverider's critical leading edges. An advanced simulation technique, the Direct Simulation Monte Carlo (DSMC) method, was used to resolve the three-dimensional flow field around the vehicle. This simulation not only detailed the pressure and thermal conditions surrounding the waverider but also highlighted the structural interactions of the shock waves, validated through the visualizations of streamwise velocity and temperature fields.

Further refinement of the simulation model involved creating the three-dimensional iso-surfaces of the Q-criterion to intricately map out the shock and vortex structures around the waverider. This allowed for an in-depth examination of the flow dynamics, particularly the attachment of shock waves and the formation of vortices at the critical sections of the vehicle. Such detailed analyses were essential for confirming the aerodynamic efficiency of the design. Additionally, the use of layered meshing techniques facilitated the accurate capture of boundary layer phenomena, crucial for assessing the vehicle's performance under the extreme conditions of hypersonic flight.

The culmination of these extensive simulations was the calculation of lift and drag forces, where the lift–drag ratio was determined to be 2.18, indicative of efficient aerodynamic performance for a waverider at a zero-degree angle of attack. These findings are significant, as they not only demonstrate the waverider's capability under the stipulated design conditions but also suggest potential improvements for future designs. For instance, adjustments in the angle of attack could further optimize the vehicle's performance. This study's insights into high-altitude aerodynamics, where different fluid dynamics necessitate modifications such as larger compression surfaces, pave the way for the enhanced designs of future hypersonic vehicles operating in near-space environments.

Author Contributions: Conceptualization, A.K. and I.K.N.; methodology, A.K. and I.K.N.; software, A.K.; validation, A.K. and I.K.N.; resources, A.K.; writing—original draft preparation, A.K.; writing—review and editing, I.K.N.; visualization, A.K.; supervision, I.K.N.; project administration, I.K.N.; funding acquisition, I.K.N. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the Integrated Air and Missile Defence Centre of Excellence (IAMDCOE) and any intellectual property resulting from the work covered by this will be property of the IAMDCOE. This paper reflects only the IAMDCOE policies and its author(s)' positions and it is not intended to create any legal obligations nor does it reflect NATO's policies or positions, or engage NATO in any way.

Data Availability Statement: The data presented in this study are available on request from the corresponding author due to the fact that this project was funded by NATO IAMDCOE, and sharing the data requires their explicit permission.

Acknowledgments: The authors would like to thank the UKTC and EPSRC for computational time on the UK supercomputing facility ARCHER2 via projects EP/R029326/1 and EP/X035484/1.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Sandlin, D.R.; Pessin, D.N. *Aerodynamic Analysis of Hypersonic Waverider Aircraft*; NASA-CR-192981; NASA: Washington, DC, USA, 1993.
2. Nonweiler, T.R.F. Delta wings of shape amenable to exact shock wave theory. *J. R. Aerosp. Soc.* **1963**, *67*, 39. [\[CrossRef\]](#)
3. Sziroczak, D.; Smith, H. A review of design issues specific to hypersonic flight vehicles. *Prog. Aerosp. Sci.* **2016**, *84*, 1–28. [\[CrossRef\]](#)
4. Zhang, T.T.; Wang, Z.G.; Huang, W.; Yan, L. Parameterization and optimization of hypersonic-gliding vehicle configurations during conceptual design. *Aerosp. Sci. Technol.* **2016**, *58*, 225–234. [\[CrossRef\]](#)
5. Kontogiannis, K. On Developing Efficient Parametric Geometry Models for Waverider-Based Hypersonic Aircraft. Ph.D. Thesis, University of Southampton, Southampton, UK, 2017.
6. Son, J.; Son, C.; Yee, K. A Novel Direct Optimization Framework for Hypersonic Waverider Inverse Design Methods. *Aerospace* **2022**, *9*, 348. [\[CrossRef\]](#)
7. Flower, J.W. Configurations for high supersonic speeds derived from simple shockwaves and expansions. *R. Aeronaut. Soc. J.* **1963**, *67*, 287–290. [\[CrossRef\]](#)
8. Kuchemann, D. *The Aerodynamic Design of Aircraft: A Detailed Introduction to the Current Aerodynamic Knowledge and Practical Guide to the Solution of Aircraft Design Problems*; Pergamon Press: Oxford, UK, 1978.
9. Bowcutt, K.G. Multidisciplinary optimization of airbreathing hypersonic vehicles. *J. Propuls. Power* **2001**, *17*, 1184–1190. [\[CrossRef\]](#)
10. Cockrell, C.E., Jr. Interpretation of waverider performance data using computational fluid dynamics. *J. Aircr.* **1994**, *31*, 1095–1100. [\[CrossRef\]](#)
11. Corda, S.; Anderson, J., Jr. Viscous optimized hypersonic waveriders designed from axisymmetric flow fields. In Proceedings of the 26th Aerospace Sciences Meeting, Reno, NV, USA, 14 January 1988.
12. Nonweiler, T.R.F. Aerodynamic problems of manned space vehicles. *R. Aeronaut. Soc. J.* **1959**, *63*, 521–528. [\[CrossRef\]](#)
13. Zubin, M.A.; Ostapenko, N.A. Experimental investigation of some singularities of the supersonic flow around V-shaped wings. *Fluid Dyn.* **1975**, *10*, 647–652. [\[CrossRef\]](#)
14. Ostapenko, N.A. Aerodynamic characteristics of V-shaped wings with shock waves detached from leading edges at hypersonic speeds. *Fluid Dyn.* **1993**, *28*, 545–552. [\[CrossRef\]](#)
15. Maksimov, F.A.; Ostapenko, N.A. V-shaped wings with an opening angle exceeding π at super- and hypersonic flow velocities. *Dokl. Phys.* **2016**, *61*, 412–417. [\[CrossRef\]](#)
16. Mazhul, I.I.; Rakhimov, R.D. Numerical investigation of off-design regimes of flow past power-law waveriders based on the flows behind plane shocks. *Fluid Dyn.* **2003**, *38*, 806–814. [\[CrossRef\]](#)
17. Mazhul, I.I.; Rakhimov, R.D. Hypersonic power-law shaped waveriders. *J. Aircr.* **2004**, *41*, 839–845. [\[CrossRef\]](#)
18. Mazhul, I.I. Off-design regimes of flow past waveriders based on isentropic compression flows. *Fluid Dyn.* **2003**, *45*, 271–280. [\[CrossRef\]](#)
19. Jones, J.G.; Moore, K.C.; Pike, J.; Roe, P.L. A method for designing lifting configurations for high supersonic speeds, using axisymmetric flow fields. *Ing.-Arch.* **1968**, *37*, 56–72. [\[CrossRef\]](#)
20. Mangin, B.; Benay, R.; Chanetz, B.; Chpoun, A. Optimization of viscous waveriders derived from axisymmetric power-law blunt body flows. *J. Spacecr. Rocket.* **2006**, *43*, 990–998. [\[CrossRef\]](#)
21. Corda, S. Viscous Optimized Hypersonic Waveriders Designed from Flows over Cones and Minimum Drag Bodies. Ph.D. Thesis, University of Maryland, College Park, MD, USA, 1988.
22. Ding, F.; Liu, J.; Shen, C.-B.; Huang, W. Novel approach for design of a waverider vehicle generated from axisymmetric supersonic flows past a pointed von Karman ogive. *Aerosp. Sci. Technol.* **2015**, *42*, 297–308. [\[CrossRef\]](#)
23. Rasmussen, M.L. Lifting-body Configurations Derived from Supersonic Flows Past Inclined Circular and Elliptic Cones. In Proceedings of the 5th Atmospheric Flight Mechanics Conference for Future Space Systems, AIAA Paper, Boulder, CO, USA, 6–8 August 1979; p. 1665.
24. Rasmussen, M.L. Waverider configurations derived from inclined circular and elliptic cones. *J. Spacecr. Rocket.* **1980**, *17*, 537–545. [\[CrossRef\]](#)
25. Rasmussen, M.L.; Jischke, M.C.; Daniel, D.C. Experimental forces and moments on cone-derived waveriders for $m = 3$ to 5. *J. Spacecr. Rocket.* **1982**, *19*, 592–598. [\[CrossRef\]](#)
26. Rasmussen, M.L.; Jischke, M.C.; Daniel, D.C. Experimental surface pressures on cone-derived waveriders for $M = 3$ –5. *J. Spacecr. Rocket.* **1983**, *20*, 539–545.
27. Lin, S.C.; Rasmussen, M.L. Cone-derived Waveriders with Combined Transverse and Longitudinal Curvature. In Proceedings of the 26th Aerospace Sciences Meeting, AIAA Paper 88–03714, Reno, NV, USA, 11–14 January 1988.
28. Liu, C.Z.; Bai, P.; Chen, B.Y.; Ji, C.Q. Rapid Design and Optimization of Waverider from 3D Flow. In Proceedings of the 16th AIAA Aviation Technology, Integration, and Operations Conference, AIAA Paper 2016–3288, Washington, DC, USA, 13–17 June 2016.
29. McLaughlin, T.A. Viscous Optimized Hypersonic Waveriders for Chemical Equilibrium Flow. Master’s Thesis, University of Maryland, College Park, MD, USA, 1990.

30. Anderson, J.D.; Chang, J.; McLaughlin, T.A. Hypersonic Waveriders: Effects of Chemically Reacting Flow and Viscous Interaction. In Proceedings of the 30th Aerospace Sciences Meeting and Exhibit, AIAA Paper 92-0302, Reno, NV, USA, 6–9 January 1992.
31. Li, M.Y.; Zhou, H. Research on the design method of waverider aerodynamic configuration based on commercial CFD software. *Chin. Q. Mech.* **2014**, *35*, 293–299.
32. Taylor, G.I.; Macoll, J.W. The air pressure on a cone moving at high speeds I. *Proc. R. Soc. A* **1933**, *139*, 278–297.
33. Santos, W.F. Bluntness impact on lift-to-drag ratio of hypersonic wedge flow. *J. Spacecr. Rocket.* **2009**, *46*, 329–339. [[CrossRef](#)]
34. Chen, X.Q.; Hou, Z.X.; Liu, J.X.; Gao, X.Z. Bluntness impact on performance of waverider. *Comput. Fluids* **2011**, *48*, 30–43. [[CrossRef](#)]
35. Tinch, D.J.; Burnett, D.W. Hypersonic waverider test vehicle—a logical next step. *J. Spacecr. Rocket.* **1994**, *31*, 392–399. [[CrossRef](#)]
36. Bird, G.A. *Molecular Gas Dynamics and the Direct Simulation of Gas Flows*; Clarendon Press: Oxford, UK, 1994.
37. Wagner, W. A convergence proof for Bird’s direct simulation Monte Carlo method for the Boltzmann equation. *J. Stat. Phys.* **1992**, *66*, 1011–1044. [[CrossRef](#)]
38. Gallis, M.A.; Torczynski, J.R.; Plimpton, S.J.; Rader, D.J.; Koehler, T. Direct Simulation Monte Carlo: The quest for speed. In Proceedings of the 29th Rarefied Gas Dynamics (RGD) Symposium, Xi’an, China, 13–18 July 2014.
39. Koura, K.; Matsumoto, H. Variable soft sphere molecular model for air species. *Phys. Fluids A Fluid. Dyn.* **1992**, *4*, 1083–1085. [[CrossRef](#)]
40. Borgnakke, C.; Larsen, P.S. Statistical Collision Model for Monte Carlo Simulation of Polyatomic Gas Mixture. *J. Comput. Phys.* **1975**, *18*, 405–420. [[CrossRef](#)]
41. Gallis, M.A.; Torczynski, J.R. Effect of Collision-Partner Selection Schemes on the Accuracy and Efficiency of the Direct Simulation Monte Carlo Method. *Int. J. Numer. Methods Fluids* **2011**, *67*, 1057–1072. [[CrossRef](#)]
42. Zhan, J.M.; Li, Y.T.; Wai, W.H.O.; Hu, W.Q. Comparison between the Q criterion and Rortex in the application of an in-stream structure. *Phys. Fluids* **2019**, *31*, 121701. [[CrossRef](#)]
43. Jeong, J.; Hussain, F. On the identification of a vortex. *J. Fluid Mech.* **1995**, *285*, 69–94. [[CrossRef](#)]
44. Huang, W.; Ma, I.; Wang, Z.; Pourkashanian, D.B.; Luo, S.; Lei, J.A. parametric study on the aerodynamic characteristics of a hypersonic waverider vehicle. *Acta Aeronaut.* **2011**, *69*, 135–140. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Hybrid Nanofluid Flow over a Shrinking Rotating Disk: Response Surface Methodology

Rusya Iryanti Yahaya ¹, Norihan Md Arifin ^{1,2,*}, Ioan Pop ³, Fadzilah Md Ali ^{1,2}
and Siti Suzilliana Putri Mohamed Isa ^{1,4}

- ¹ Institute for Mathematical Research, Universiti Putra Malaysia, Serdang 43400, Malaysia; rusyairyanti@gmail.com (R.I.Y.); fadzilahma@upm.edu.my (F.M.A.); ctsuzilliana@upm.edu.my (S.S.P.M.I.)
² Department of Mathematics and Statistics, Universiti Putra Malaysia, Serdang 43400, Malaysia
³ Department of Mathematics, Babeş-Bolyai University, 400084 Cluj-Napoca, Romania; popm.ioan@yahoo.co.uk
⁴ Centre of Foundation Studies for Agricultural Science, Universiti Putra Malaysia, Serdang 43400, Malaysia
* Correspondence: norihana@upm.edu.my

Abstract: For efficient heating and cooling applications, minimum wall shear stress and maximum heat transfer rate are desired. The current study optimized the local skin friction coefficient and Nusselt number in Al_2O_3 -Cu/water hybrid nanofluid flow over a permeable shrinking rotating disk. First, the governing equations and boundary conditions are solved numerically using the bvp4c solver in MATLAB. Von Kármán's transformations are used to reduce the partial differential equations into solvable non-linear ordinary differential equations. The augmentation of the mass transfer parameter is found to reduce the local skin friction coefficient and Nusselt number. Higher values of these physical quantities of interest are observed in the injection case than in the suction case. Meanwhile, the increase in the magnitude of the shrinking parameter improved and reduced the local skin friction coefficient and Nusselt number, respectively. Then, response surface methodology (RSM) is conducted to understand the interactive impacts of the controlling parameters in optimizing the physical quantities of interest. With a desirability of 66%, the local skin friction coefficient and Nusselt number are optimized at 1.528780016 and 0.888353037 when the shrinking parameter (λ) and mass transfer parameter (S) are -0.8 and -0.6 , respectively.

Keywords: disk; hybrid nanofluid; RSM; shrinking; suction/injection



Citation: Yahaya, R.I.; Arifin, N.M.; Pop, I.; Ali, F.M.; Isa, S.S.P.M. Hybrid Nanofluid Flow over a Shrinking Rotating Disk: Response Surface Methodology. *Computation* **2024**, *12*, 141. <https://doi.org/10.3390/computation12070141>

Academic Editors: Ali Cemal Benim, Abdulmajeed A. Mohamad, Sang-Ho Suh, Rachid Bennacer, Paweł Oclon and Jan Taler

Received: 28 June 2024
Revised: 6 July 2024
Accepted: 8 July 2024
Published: 10 July 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Rotating machinery, gas turbine rotators, and thermal power generating systems are among the numerous practical applications of the classical rotating-disk flow [1]. The Navier–Stokes equations for rotating-disk flow were first presented by Von Kármán [2]. In this study, Von Kármán's transformations were introduced to reduce the partial differential equations into a system of ordinary differential equations. Later, these equations were employed in various rotating-disk flow problems. Ayano et al. [3] did a numerical computation on magnetohydrodynamics (MHD) nanofluid flow over a stretching rotating disk with Brownian motion, thermophoresis, and gyrotactic microorganisms. Von Kármán's transformations were employed in this study, and the increased stretching parameter was found to reduce the temperature profile. Meanwhile, the augmentation of the thermophoresis and Brownian motion enhanced and reduced the concentration profile. Then, the unsteady flow of nanofluid over a stretching rotating disk with variable viscosity, Brownian motion, and thermophoresis was analyzed by Sharma et al. [4]. It was found that the increased stretching rate promoted fluid flow in the radial direction and enhanced horizontal velocity. Puneeth et al. [5] then examined the unsteady flow over a permeable stretching rotating disk. It was observed that a high nanoparticle volume fraction improved and reduced the temperature and velocity profiles of the nanofluid.

Researchers have also examined the rotating-disk flow of hybrid nanofluids containing the dispersion of two different nanoparticles. Rashid et al. [6] analyzed the MHD hybrid nanofluid flow over a rotating disk with Brownian motion and thermophoresis. It was noted that the hybrid nanofluid possessed better thermal conductivity than the nanofluid. Moreover, the augmentation of the magnetic field reduced axial and tangential velocities while increasing the surface drag. Meanwhile, thermophoresis and nanoparticle volume fraction enhanced the heat transfer performance. Alkuhayli [7] extended this study with thermal radiation and heat source/sink. The enhancement of the thermal radiation and variable thermal conductivity parameters was found to have boosted the temperature profile and local Nusselt number. Further extension of this study was made by Kumar et al. [8], who considered the electromagnetohydrodynamic flow over a stretching rotating disk with an irregular heat source, Arrhenius energy, and non-linear radiation. The increment in the electric parameter was noted to reduce the skin friction coefficient and enhance the heat transfer rate. Meanwhile, Kulkarni [9] found that the heat transfer rate increased with the sphericity of the nanoparticles in hybrid nanofluid flow over a permeable rotating disk with mixed convection, magnetic field, and slip. Lone et al. [10] then analyzed the MHD flow of $\text{Al}_2\text{O}_3\text{-Cu}$ /water hybrid nanofluid over a rotating disk with velocity slip, zero mass flux, and convective boundary conditions. Next, Khashi'ie et al. [11] conducted a stability analysis on multiple solutions obtained in MHD hybrid nanofluid flow over a permeable stretching/shrinking rotating disk with heat generation. The findings revealed that only the first and second solutions were stable and physically significant, while the third solution was unstable. In addition, only the suction parameter enhanced the heat transfer performance of the hybrid nanofluid. Vijay and Sharma [12] then studied the Soret and Dufour effects on MHD hybrid nanofluid flow over a permeable decelerating rotating disk. Later, Abu Bakar et al. [13] analyzed the unsteady flow of $\text{Al}_2\text{O}_3\text{-Cu}$ /water hybrid nanofluid over an inclined, shrinking, rotating disk with thermal radiation and velocity slip. Similarly, only the first solution was found to be stable, while the second solution was unstable. The nanoparticle volume fraction, radiation, and shrinking parameters increased the heat transfer rate. Recently, the rotating-disk flow of ternary and tetra hybrid nanofluids, consisting of three and four different nanoparticles dispersed in a base fluid, was also analyzed by researchers (see Ali et al. [14], Singh et al. [15], and Patgiri and Paul [16]). In addition, researchers have also incorporated machine learning approaches to study fluid flow over a rotating disk (see Ali et al. [17] and Kumar et al. [18]).

Moreover, researchers have also included statistical analysis to improve theoretical investigations of fluid flow over a rotating disk. A sophisticated statistical tool called the response surface methodology (RSM) is commonly employed to analyze the effects and significance of various controlling parameters on specified responses. The optimal conditions at which to achieve the desired response can also be estimated using the RSM. Mehmood et al. [19] utilized the RSM to optimize the local skin friction coefficients and Nusselt number in MHD nanofluid flow past a stretched rotating disk with entropy generation and autocatalytic chemical reaction. Both physical quantities of interest were found to be optimized by the higher value of the magnetic parameter. Then, Shafiq et al. [20] considered the Darcy–Forchheimer flow of nanofluid past a stretched rotating disk with thermal radiation. Based on the RSM and sensitivity analysis, the local Nusselt number was most sensitive towards the solid volume fraction, followed by the radiation parameter and Biot number. Later, Mathew et al. [21] examined the effects of magnetic field and stratification in bioconvective stagnation-point flow over a stretched rotating disk. This study employed the RSM to investigate the relationship between the magnetic field, thermal stratification, the nanoparticle volume fraction of magnetite, and the velocity ratio parameter and the heat transfer rate. The thermal stratification and magnetic field parameters were found to have a negative correlation with the heat transfer rate. Recently, the RSM was utilized by Hussain et al. [22] to optimize the MHD nanofluid flow past a rotating disk with thermal radiation and Joule heating. Meanwhile, Khashi'ie et al. [23] found through the RSM that the magnetic and suction parameters significantly affected the skin friction coefficient in

unsteady hybrid nanofluid flow over a rotating disk with heat generation. In the meantime, the heat transfer rate was affected by the magnetic, suction, and heat generation parameters.

Due to the scarcity of rotating-disk flow studies that combine numerical and statistical investigations, the current research will focus on these approaches to analyze the flow and heat transfer performance in Al_2O_3 -Cu/water hybrid nanofluid flow past a shrinking rotating disk. The surface is assumed to be permeable to allow for suction/injection. Partial differential equations and boundary conditions governing this flow problem will be solved numerically using the bvp4c solver in MATLAB. Then, RSM will be executed to analyze the interaction of mass transfer with shrinking parameters and their significance in optimizing the physical quantities of interest. The current study aims to fill the existing literature gap and offer a better understanding of the impacts of mass transfer and shrinking parameters on the considered flow problem.

2. Mathematical Formulation

The working fluid contains alumina (Al_2O_3) and copper (Cu) nanoparticles dispersed in water. This combination is chosen because Cu nanoparticles offer high thermal conductivity, Al_2O_3 nanoparticles exhibit good inertness, and water is widely available. Moreover, an experimental study by Siddiqui et al. [24] demonstrated that an Al_2O_3 -Cu/water hybrid nanofluid with a 50:50 mixing ratio resulted in an overall enhanced hydrothermal property with improved thermal conductivity and better stability. Meanwhile, the correlations and values of thermophysical properties for the Al_2O_3 -Cu/water hybrid nanofluid are based on Takabi et al. [25] and Oztop and Abu-Nada [26]. At a low-volume concentration of this hybrid nanofluid, the correlations proposed by Takabi et al. [25] were found to be in good agreement with the experiment results presented by Suresh et al. [27]. Following this, a low-volume concentration of Cu and Al_2O_3 nanoparticles is used in the current study (i.e., $\phi_{\text{Al}_2\text{O}_3} = \phi_{\text{Cu}} = 0.02$).

Meanwhile, the flow configuration is described using cylindrical coordinates (r, φ, z) and velocity components (u, v, w) , respectively. The schematic representation is depicted in Figure 1.

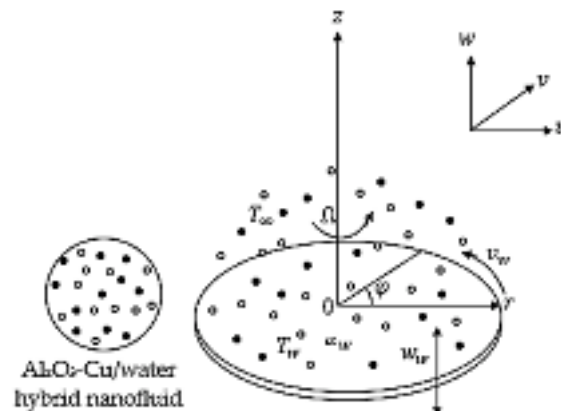


Figure 1. Schematic diagram of the flow problem.

The fluid is assumed to flow in the region $z \geq 0$ over an infinite rotating disk with a uniform shrinking rate a in the radial direction r ; angular velocity Ω ; mass flux velocity $w_w = \sqrt{\Omega \nu_{bf}} S$ with mass transfer parameter S and kinematic viscosity of base fluid ν_{bf} ; and constant temperature T_w . Meanwhile, the surrounding hybrid nanofluid is assumed to have a temperature of T_∞ , with $T_w > T_\infty$ for a hot disk. In the current study, the flow is considered to be steady and axisymmetric (i.e., $\partial/\partial\varphi = 0$ for all variables).

Following the stated assumptions, the governing equations and boundary conditions for the laminar, incompressible Newtonian hybrid nanofluid flow are as follows [28,29]:

$$\frac{\partial u}{\partial r} + \frac{u}{r} + \frac{\partial w}{\partial z} = 0, \quad (1)$$

$$u \frac{\partial u}{\partial r} + w \frac{\partial u}{\partial z} - \frac{v^2}{r} = \frac{\mu_{hnf}}{\rho_{hnf}} \left[\frac{1}{r} \frac{\partial u}{\partial r} - \frac{u}{r^2} + \frac{\partial^2 u}{\partial r^2} + \frac{\partial^2 u}{\partial z^2} \right], \quad (2)$$

$$u \frac{\partial v}{\partial r} + w \frac{\partial v}{\partial z} + \frac{u v}{r} = \frac{\mu_{hnf}}{\rho_{hnf}} \left[\frac{1}{r} \frac{\partial v}{\partial r} - \frac{v}{r^2} + \frac{\partial^2 v}{\partial r^2} + \frac{\partial^2 v}{\partial z^2} \right], \quad (3)$$

$$u \frac{\partial v}{\partial r} + w \frac{\partial v}{\partial z} + \frac{u v}{r} = \frac{\mu_{hnf}}{\rho_{hnf}} \left[\frac{1}{r} \frac{\partial v}{\partial r} - \frac{v}{r^2} + \frac{\partial^2 v}{\partial r^2} + \frac{\partial^2 v}{\partial z^2} \right], \quad (4)$$

$$u \frac{\partial T}{\partial r} + w \frac{\partial T}{\partial z} = \frac{k_{hnf}}{(\rho C_p)_{hnf}} \left[\frac{1}{r} \frac{\partial T}{\partial r} + \frac{\partial^2 T}{\partial r^2} + \frac{\partial^2 T}{\partial z^2} \right], \quad (5)$$

$$\left. \begin{aligned} u = u_w = ar, \quad v = v_w = \Omega r, \quad w = w_w, \quad T = T_w \quad \text{at } z = 0 \\ u \rightarrow 0, \quad v \rightarrow 0, \quad T \rightarrow T_\infty \quad \text{as } z \rightarrow \infty \end{aligned} \right\}. \quad (6)$$

In these equations, the subscript *hnf* denotes hybrid nanofluid, μ is the dynamic viscosity, ρ is the density, k is the thermal conductivity, C_p is the specific heat, T is the temperature, and p is the pressure. Then, these partial differential equations are subjected to the following Von Kármán's transformations [2]:

$$\left. \begin{aligned} u = \Omega r f(\eta), \quad v = \Omega r g(\eta), \quad w = \sqrt{\Omega \nu_{bf}} h(\eta), \quad \theta(\eta) = \frac{T - T_\infty}{T_w - T_\infty} \\ p - p_\infty = 2 \mu_{bf} \Omega P(\eta), \quad \eta = z \sqrt{\frac{\Omega}{\nu_{bf}}} \end{aligned} \right\} \quad (7)$$

to form the following non-linear ordinary differential equations and boundary conditions:

$$2f + h' = 0, \quad (8)$$

$$\frac{\mu_{hnf}/\mu_{bf}}{\rho_{hnf}/\rho_{bf}} f'' - h f' - f^2 + g^2 = 0, \quad (9)$$

$$\frac{\mu_{hnf}/\mu_{bf}}{\rho_{hnf}/\rho_{bf}} g'' - h g' - 2fg = 0, \quad (10)$$

$$\frac{k_{hnf}/k_{bf}}{(\rho C_p)_{hnf}/(\rho C_p)_{bf}} \theta'' - Pr h \theta' = 0, \quad (11)$$

$$\left. \begin{aligned} g(0) = 1, \quad f(0) = \lambda, \quad h(0) = S, \quad \theta(0) = 1 \\ g(\eta) \rightarrow 0, \quad f(\eta) \rightarrow 0, \quad \theta(\eta) \rightarrow 0 \quad \text{as } \eta \rightarrow \infty \end{aligned} \right\}, \quad (12)$$

where ' $'$ ' denotes differentiation with respect to η ; *bf* refers to the base fluid; f , g , and h are the radial, tangential, and axial velocities; $\lambda = a/\Omega$ is the shrinking parameter; S is the mass transfer parameter with $S < 0$ for injection and $S > 0$ for suction; and $Pr = (\mu C_p)_{bf}/k_{bf}$ is the Prandtl number. According to Turkylmazoglu [28], the fluid pressure can be determined by integrating Equation (4), and the resulting pressure equation is decoupled from the flow field. Since the interest of the current study is to analyze the flow and thermal fields of the hybrid nanofluid, the pressure function can be neglected.

Then, the physical quantities of interest are the skin friction coefficient (C_f) and Nusselt number (Nu_r) [28]:

$$C_f = \frac{\mu_{hmf} \sqrt{\left[\left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial \varphi} \right)_{z=0} \right]^2 + \left[\left(\frac{\partial v}{\partial z} + \frac{1}{r} \frac{\partial w}{\partial \varphi} \right)_{z=0} \right]^2}}{\rho_{bf} (\Omega r)^2},$$

$$Nu_r = \frac{r k_{nf}}{k_{bf} (T_w - T_\infty)} \left(-\frac{\partial T}{\partial z} \right)_{z=0},$$

which result in the following:

$$Re_r^{1/2} C_f = \frac{\mu_{hmf}}{\mu_{bf}} \sqrt{[f'(0)]^2 + [g'(0)]^2}, \quad Re_r^{-1/2} Nu_r = -\frac{k_{hmf}}{k_{bf}} \theta'(0), \quad (13)$$

where $Re_r^{1/2} C_f$ is the local skin friction coefficient; $Re_r^{-1/2} Nu_r$ is the local Nusselt number; and $Re_r = \frac{\Omega r^2}{\nu_{bf}}$ is the local Reynolds number. Here, an increase in the angular velocity can raise the local Reynolds number.

3. Response Surface Methodology (RSM)

Box and Wilson [30] introduced an experimental design-based approach, incorporating statistical and mathematical techniques, called the response surface methodology (RSM). The RSM can be used to model, analyze, and optimize a system containing dependent (response) and independent parameters. The relationship between the response and independent parameters can be described as follows [31]:

$$\text{Response} = F(X_1, X_2, X_3, \dots, X_N) \pm E, \quad (14)$$

where F is the unknown function of response with $X_1, X_2, X_3, \dots, X_N$ as independent parameters; and E is the random experimental error assumed to have a zero mean. Meanwhile, the general quadratic model is

$$\text{Response} = c_0 + \sum_{i=1}^N c_i x_i + \sum_{i=1}^N c_{ii} x_i^2 + \sum_{i=1}^{N-1} \sum_{j=i+1}^N c_{ij} x_i x_j, \quad (15)$$

with c_0 , c_i , c_{ii} , and c_{ij} representing the intercept, linear, quadratic, and bilinear two-factor terms, respectively.

The current study utilizes the RSM and face-centered central composite design to investigate the significance of the mass transfer and shrinking parameters on the local skin friction coefficient and Nusselt number, as well as the optimal settings of these parameters to minimize and maximize the responses. Therefore, the mass transfer and shrinking parameters are denoted as the independent parameters, and the local skin friction coefficient and Nusselt number are the responses. The independent parameters are labeled as A and B and are assigned with actual and coded values, as shown in Table 1. The ranges for the actual values are taken based on the numerical investigation conducted in the current study.

Table 1. The actual and coded values of the independent parameters.

Independent Parameter	Symbol	Level		
		Low (−1)	Medium (0)	High (1)
λ	A	−1.0	−0.9	−0.8
S	B	−0.6	−0.1	0.4

The central composite design is the most commonly used response surface design to determine the coefficients of the quadratic model. According to Bhattacharya [32], the

central composite design has the benefit of estimating the non-linear relationship between the independent parameters and the response, as well as the ability to provide maximum information with minimal experimental data, and fewer experiments are needed to predict the quadratic term in a quadratic model. The central composite design is comprised of several types, including the face-centered type. The face-centered central composite design requires three levels of each factor or independent parameter (i.e., $-1, 0, 1$) with $2^N + 2N + C$ experiments. Here, N is the number of independent parameters; 2^N denotes the factorial points; $2N$ denotes the axial points; and C denotes the center points. As stated by Sahoo and Barman [33], the center points may vary from three to six. In this current study, with $N = 2$ and $C = 5$, a total of 13 runs are carried out. The respective responses, calculated using MATLAB for each run, are then tabulated in Table 2. Meanwhile, Figure 2 presents the flowchart of the numerical and statistical procedures involved in the current study.

Table 2. Experimental design and response values.

Runs	Coded Values		Actual Values		Responses	
	A	B	λ	S	$Re_r^{1/2}Cf$ (Response 1)	$Re_r^{-1/2}Nu_r$ (Response 2)
1	−1	−1	−1.0	−0.6	1.803245471	0.232133118
2	1	−1	−0.8	−0.6	1.529299905	0.975650061
3	−1	1	−1.0	0.4	1.475361873	0.000000014
4	1	1	−0.8	0.4	1.229514429	0.000001138
5	−1	0	−1.0	−0.1	1.642240690	0.000228587
6	1	0	−0.8	−0.1	1.380007580	0.004928016
7	0	−1	−0.9	−0.6	1.659170147	0.525530320
8	0	1	−0.9	0.4	1.346068736	0.000000126
9	0	0	−0.9	−0.1	1.504436107	0.001106468
10	0	0	−0.9	−0.1	1.504436107	0.001106468
11	0	0	−0.9	−0.1	1.504436107	0.001106468
12	0	0	−0.9	−0.1	1.504436107	0.001106468
13	0	0	−0.9	−0.1	1.504436107	0.001106468

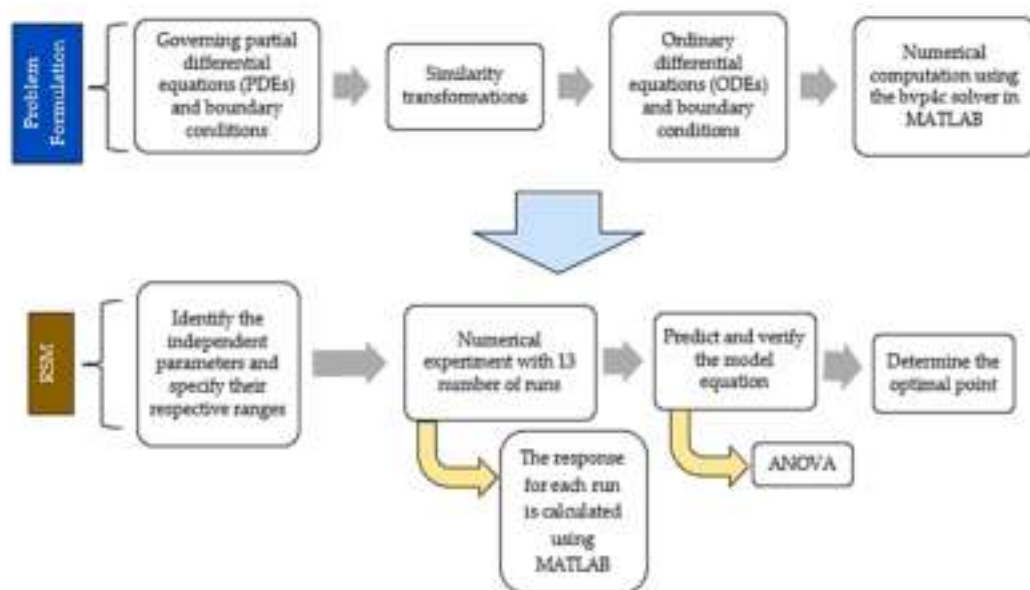


Figure 2. Flowchart of the numerical and statistical procedures.

4. Results and Discussion

The boundary value problem (8)–(12) is numerically solved using the bvp4c solver in MATLAB, and the algorithm is presented in Supplementary Materials. A comparison of

results is made to validate the correctness of the mathematical approach employed in this current study. For $\phi_{hmf} = 0$, the boundary value problem reduces into a classical viscous fluid flow problem, and the numerical results can be compared with those computed by Rashidi et al. [34] and Turkyilmazoglu [28]. These previous studies utilized the shooting and spectral Chebyshev collocation methods, respectively. Despite the difference in computation techniques, the results generated by the bvp4c solver are in close agreement with the previous studies (see Table 3). Furthermore, by assuming the finite value of the free stream boundary condition (i.e., $\eta \rightarrow \infty$) to be $\eta_{\max} = 10$, all profiles acquired in this investigation approach the free stream boundary conditions (12) asymptotically. Therefore, it enhances confidence in the reliability of the numerical findings.

Table 3. Comparison of results for $\phi_{Al_2O_3} = \phi_{Cu} = S = \lambda = 0$ and $Pr = 6.2$.

	Rashidi et al. [34]	Turkyilmazoglu [28]	Present Study
$f'(0)$	0.510186	0.51023262	0.510232394
$-g'(0)$	0.61589	0.61592201	0.615921866
$-\theta'(0)$	-	0.93387794	0.933877577

The effects of shrinking and mass transfer parameters on the velocity and temperature profiles are presented in Figure 3. The radial, tangential, and axial velocities are denoted by $f(\eta)$, $g(\eta)$, and $h(\eta)$, respectively. As observed in Figure 3a, the radial velocity profile of the hybrid nanofluid decreases with the augmentation of the mass transfer parameter and the magnitude of the shrinking parameter. However, the opposite behavior is observed at some distance from the disk. As the disk is assumed to have a no-slip condition and to shrink in the radial direction, the radial velocity of the hybrid nanofluid near the disk decreases as the magnitude of the shrinking parameter increases. The shrinking effects towards the fluid velocity are shown to be diminished at some distance away from the disk (see Figure 3a). Meanwhile, the shrinking of the disk does not affect the tangential and axial velocities of the hybrid nanofluid. The tangential velocity of the hybrid nanofluid continues to increase throughout the boundary layer when the mass transfer parameter and magnitude of the shrinking parameter increase (see Figure 3b). Figure 3c demonstrates that the axial velocity profile near the disk rises with both the mass transfer parameter and the magnitude of the shrinking parameter. As the mass transfer occurs in the axial direction, the fluid velocity increases with the mass transfer parameter. Meanwhile, the radial velocity near the disk, shown in Figure 3a, decreases due to fluid removal from the boundary by suction ($S > 0$). The increment in the mass transfer parameter and magnitude of the shrinking parameter then enhances the temperature profile, as shown in Figure 3d.

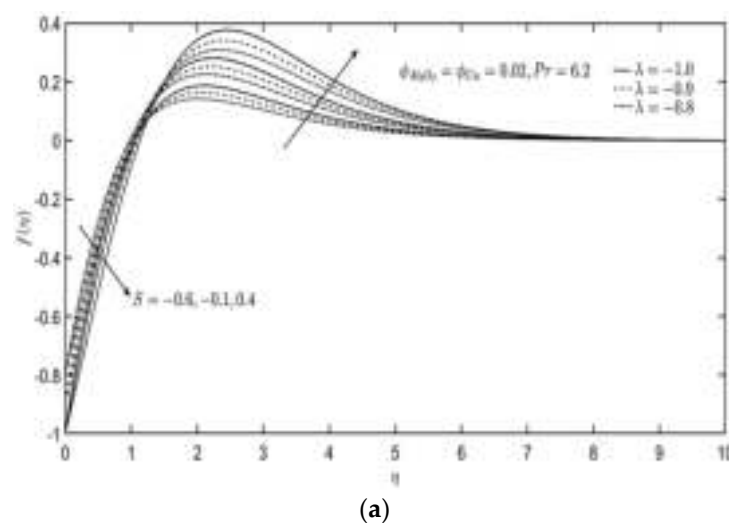


Figure 3. Cont.

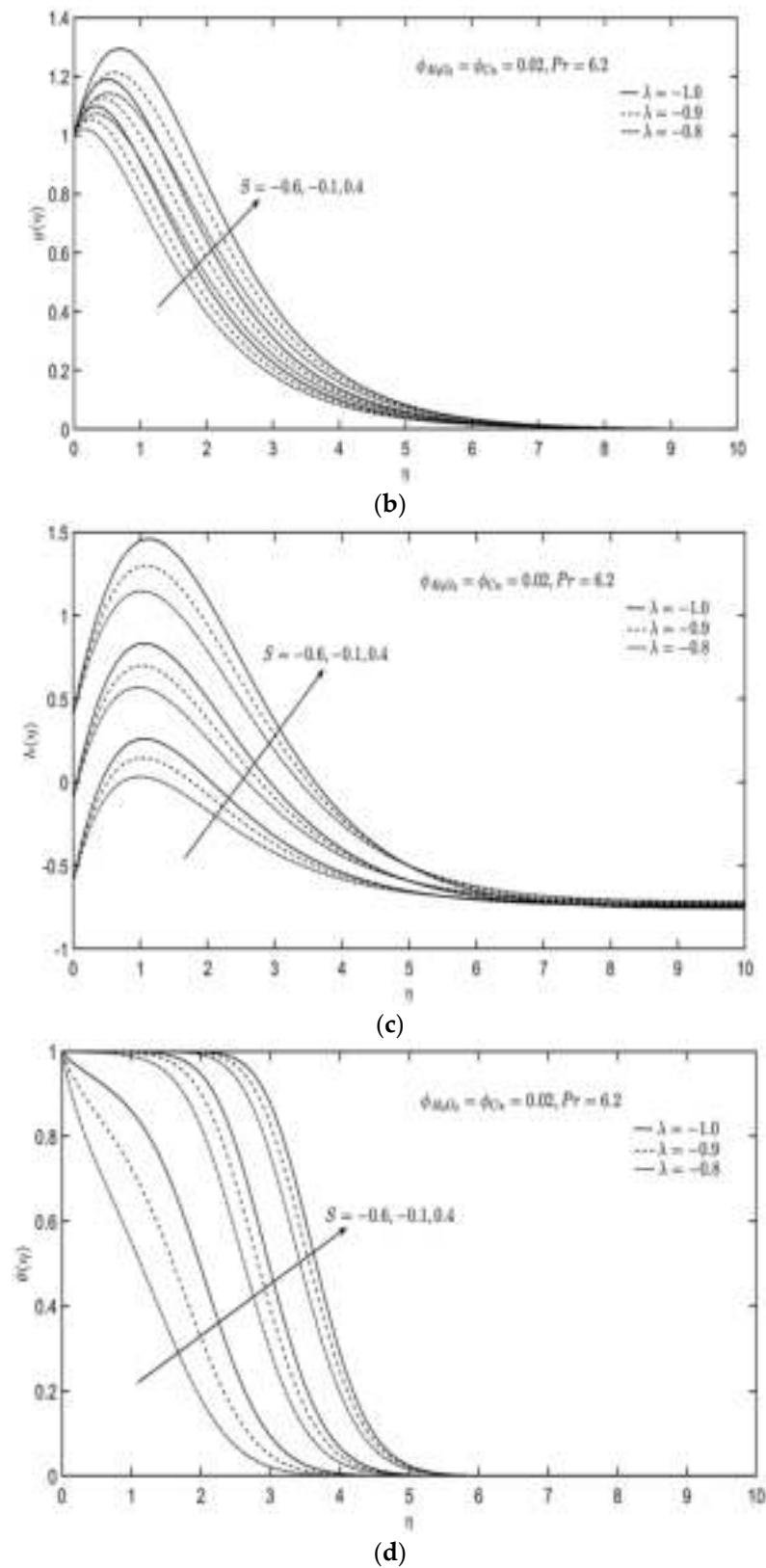


Figure 3. The profiles of (a) radial velocity, (b) tangential velocity, (c) axial velocity, and (d) temperature with various values of λ and S .

Next, the effects of shrinking and mass transfer parameters on the local skin friction coefficient and Nusselt number are illustrated in Figure 4. The enhancement of the mass transfer parameter is found to reduce the local skin friction coefficient and Nusselt number

in Figure 4a,b, respectively. Meanwhile, the increment in the magnitude of the shrinking parameter raises the local skin friction coefficient, as shown in Figure 4a. Thus, the wall shear stress rises with the increase in the magnitude of the shrinking parameter. In contrast, the local Nusselt number diminishes as $|\lambda|$ increases. The observed behavior may be due to the reduction of the surface area available for heat transfer. Hence, it causes the heat transfer rate between the disk and hybrid nanofluid to decrease. Moreover, it can be noted that the values of the local skin friction coefficient and Nusselt number for the injection case ($S < 0$) are higher than the suction case ($S > 0$).

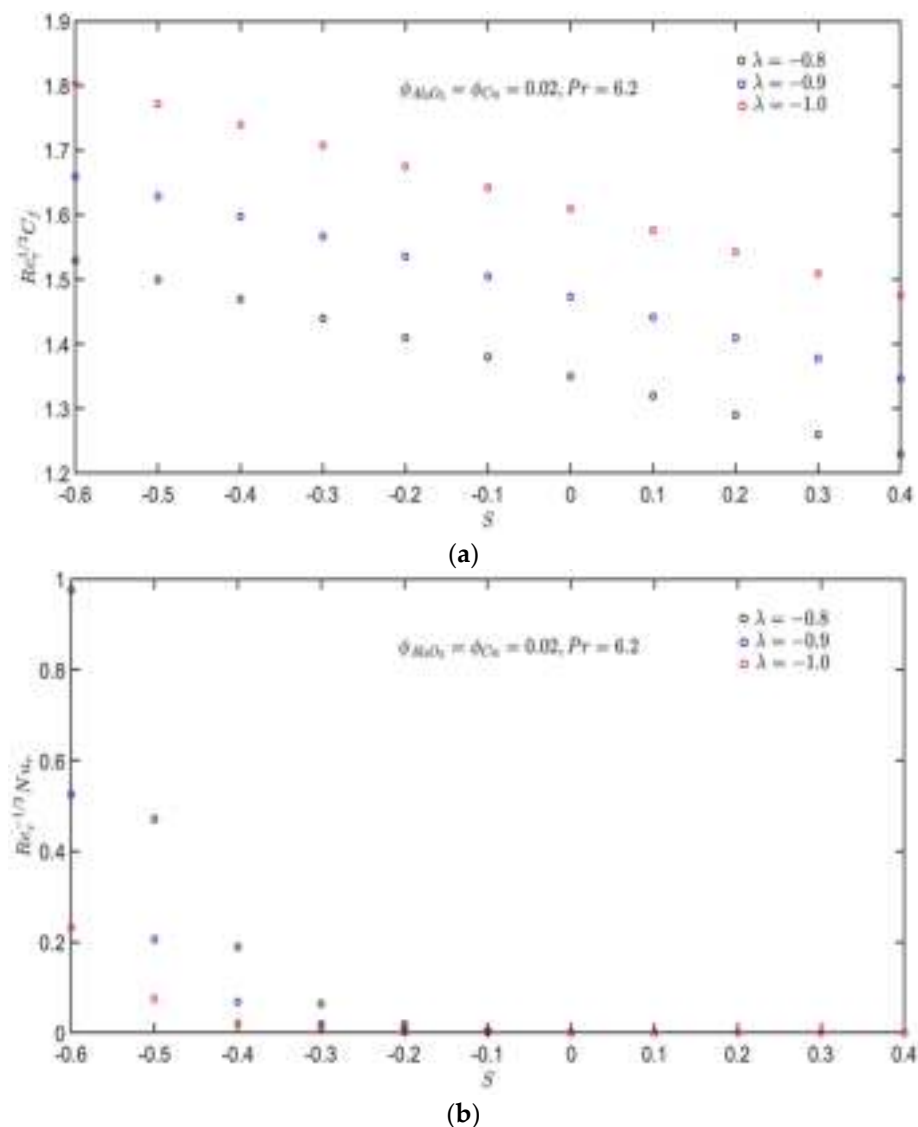


Figure 4. The profiles of (a) local skin friction coefficient and (b) local Nusselt number with various values of λ and S .

Meanwhile, the interacting impacts of the shrinking and mass transfer parameters on the physical quantities of interest are analyzed using the RSM in Minitab. The correlations between the independent parameters and the responses are given by the following:

$$\text{Response 1} = c_0 + c_1A + c_2B + c_{11}AA + c_{22}BB + c_{12}AB, \quad (16)$$

$$\text{Response 2} = c_0 + c_1A + c_2B + c_{11}AA + c_{22}BB + c_{12}AB. \quad (17)$$

The significance of each term in Equations (16) and (17) are identified through the analysis of variance (ANOVA). The results are presented in Tables 4 and 5.

Table 4. Analysis of variance for Response 1.

Source	DF	Adj SS	Adj MS	F-Value	p-Value
Model	5	0.249759313	0.049951863	174774.611709837	0.000000000
Linear	2	0.249435660	0.124717830	436370.320105496	0.000000000
A	1	0.101927475	0.101927475	356630.043436460	0.000000000
B	1	0.147508184	0.147508184	516110.596774533	0.000000000
Square	2	0.000126277	0.000063139	220.913397825	0.000000474
AA	1	0.000124396	0.000124396	435.243856915	0.000000146
BB	1	0.000008884	0.000008884	31.084684334	0.000837291
2-Way Interaction	1	0.000197376	0.000197376	690.591542539	0.000000030
AB	1	0.000197376	0.000197376	690.591542539	0.000000030
Error	7	0.000002001	0.000000286		
Lack-of-Fit	3	0.000002001	0.000000667	*	*
Pure Error	4	0.000000000	0.000000000		
Total	12	0.249761314			

The symbol * denotes the value is too small.

Table 5. Analysis of variance for Response 2.

Source	DF	Adj SS	Adj MS	F-Value	p-Value
Model	5	1.000087703	0.200017541	29.206710770	0.000150454
Linear	2	0.594033446	0.297016723	43.370603881	0.000113785
A	1	0.093304904	0.093304904	13.624451749	0.007745032
B	1	0.500728543	0.500728543	73.116756012	0.000059439
Square	2	0.267850314	0.133925157	19.555851501	0.001363026
AA	1	0.001069278	0.001069278	0.156136818	0.704502971
BB	1	0.216321666	0.216321666	31.587451304	0.000798846
2-Way Interaction	1	0.138203943	0.138203943	20.180643085	0.002824322
AB	1	0.138203943	0.138203943	20.180643085	0.002824322
Error	7	0.047938393	0.006848342		
Lack-of-Fit	3	0.047938393	0.015979464	*	*
Pure Error	4	0.000000000	0.000000000		
Total	12	1.048026096			

The symbol * denotes the value is too small.

The symbol * denotes the value is too small.

The symbol * denotes the value is too small.

According to Alhadri et al. [35], a term is statistically significant if its *p*-value is less than 0.05. Based on Tables 4 and 5, all terms are significant for correlation (16), while the term AA can be removed from correlation (17) as its *p*-value is greater than 0.05. After eliminating the insignificant term, a second ANOVA is carried out for Response 2. Then, the following new correlations, containing only the significant terms, are obtained:

$$\text{Response 1} = 1.50443 - 0.130338A - 0.156795B + 0.006711AA - 0.001794BB + 0.007025AB, \quad (18)$$

$$\text{Response 2} = 0.0015 + 0.1247A - 0.2889B + 0.2874BB - 0.1859AB. \quad (19)$$

Based on Equations (18) and (19), the term B is seen to have a negative correlation with Response 1 and Response 2. Hence, the increment in the mass transfer parameter has an unfavorable impact on the local skin friction coefficient and Nusselt number. These findings agree with the results obtained in Figure 4.

Then, the interactive impacts between the shrinking and mass transfer parameters can be analyzed through the surface and contour plots displayed in Figures 5 and 6. Based on Figure 5, the increase in A and B reduces Response 1. The higher and medium levels of shrinking parameter (A) and mass transfer parameter (B) are observed to produce low values of $Re_r^{1/2}C_f$ (Response 1). The maximum $Re_r^{1/2}C_f$ (Response 1) occurs at low levels

of these independent parameters. Meanwhile, Figure 6 shows that lower levels of the mass transfer parameter (B) and higher levels of the shrinking parameter (A) produce high values of $Re_r^{-1/2}Nu_r$ (Response 2). However, at the moderate level of mass transfer parameter (B), the increment in the shrinking parameter (A) has no significant effect on $Re_r^{-1/2}Nu_r$ (Response 2).

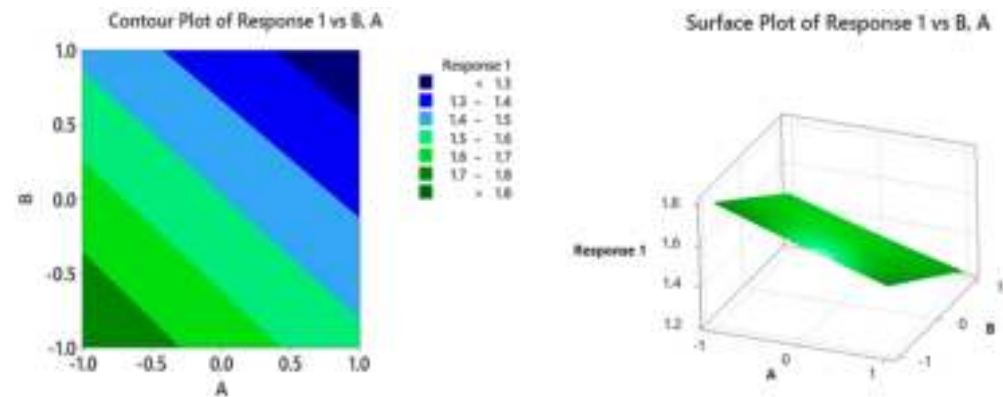


Figure 5. Contour and surface plots for Response 1.

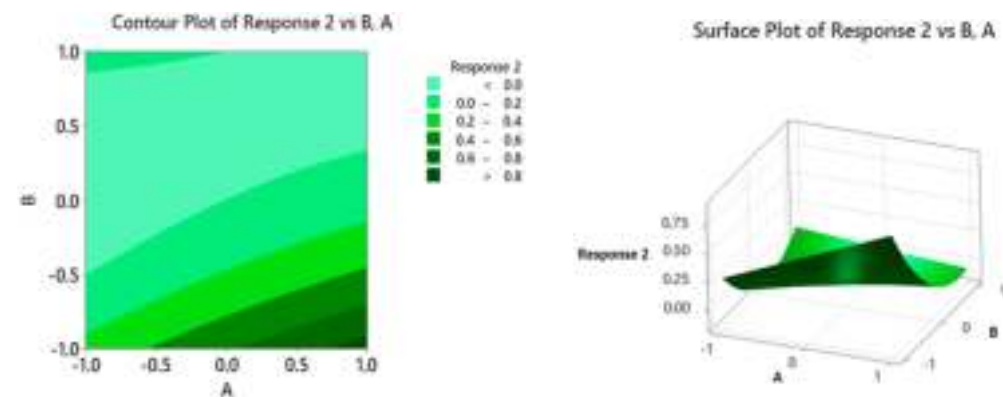


Figure 6. Contour and surface plots for Response 2.

Next, optimization is conducted to minimize and maximize the local skin friction coefficient and Nusselt number, respectively. The local skin friction coefficient is related to the wall shear stress. A low skin friction coefficient is sought to reduce energy consumption in most applications. Meanwhile, a high heat transfer rate, represented by the local Nusselt number, is desired for better heat transfer performance in cooling or heating processes. The composite desirability assesses how the configurations of the independent parameters optimize a given set of responses. The desirability is measured on a scale from zero to one. A value of zero signifies that one or more responses exceed their permissible limits, whereas a value of one represents the optimal scenario. Therefore, in the present investigation, the optimization can be observed by examining the composite desirability.

Table 6 presents the optimization result for the local skin friction coefficient and Nusselt number. The minimum local skin friction coefficient (Response 1) and maximum local Nusselt number (Response 2) can be achieved by setting the shrinking parameter (A) and mass transfer parameter (B) at high and low levels, respectively. With a desirability of 66% at $\lambda = -0.8$ and $S = -0.6$, the local skin friction coefficient is minimized at 1.528780016, while the local Nusselt number is maximized at 0.888353037. Hence, the lowest magnitude of the shrinking parameter and lowest mass transfer parameter optimized the physical quantities of interest in the current flow problem. Different behavior may be observed for other controlling parameters and flow conditions ranges.

Table 6. Optimization of responses.

Solution	A	B	Response 2 Fit	Response 1 Fit	Composite Desirability
1	1	−1	0.888353037	1.528780016	0.659986551

5. Conclusions

The current study analyzes the rotating-disk flow of a hybrid nanofluid. The disk is considered to shrink in the radial direction and is permeable to allow for suction/injection. The local skin friction coefficient and Nusselt number for this flow problem are optimized using the RSM. First, the governing equations for the flow problem, consisting of differential equations and boundary conditions, are presented and solved using the bvp4c solver in MATLAB. Then, the Minitab 21 software is utilized to perform the RSM. The results are outlined as follows:

1. As the mass transfer parameter and the magnitude of the shrinking parameter ($|\lambda|$) increase, the radial velocity profile of the hybrid nanofluid near the disk decreases.
2. The increment in the mass transfer parameter and the magnitude of the shrinking parameter improves the profiles of tangential velocity, axial velocity, and temperature.
3. The local skin friction coefficient and Nusselt number diminish with the increase in the mass transfer parameter. The injection case demonstrates higher values of these physical quantities of interest than the suction case.
4. The increment in the magnitude of the shrinking parameter enhances the local skin friction coefficient and reduces the local Nusselt number.
5. Based on the RSM, low values of $Re_r^{1/2}C_f$ are produced at higher and medium levels of shrinking parameter (λ) and mass transfer parameter. Meanwhile, the highest value of $Re_r^{1/2}C_f$ is observed at low levels of these independent parameters.
6. High values of $Re_r^{-1/2}Nu_r$ occur at lower levels of the mass transfer parameter and higher levels of the shrinking parameter.
The local Nusselt number is not significantly impacted by the increase in the shrinking parameter at the moderate level of the mass transfer parameter.
7. With a desirability of 66%, the local skin friction coefficient is minimized at 1.528780016, while the local Nusselt number is maximized at 0.888353037 when $\lambda = -0.8$ and $S = -0.6$.

The current study can be extended to other types of hybrid nanofluids, such as non-Newtonian hybrid nanofluids, carbon nanotubes hybrid nanofluids, and ternary hybrid nanofluids. In addition, the complex non-linear interaction of various controlling parameters in this flow problem can be further analyzed using a sophisticated machine learning technique known as the artificial neural network (ANN).

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/computation12070141/s1>.

Author Contributions: Conceptualization, I.P.; methodology, I.P. and R.I.Y.; validation, I.P. and N.M.A.; formal analysis, R.I.Y.; writing—original draft preparation, R.I.Y. and I.P.; writing—review and editing, N.M.A. and I.P.; supervision, N.M.A., F.M.A. and S.S.P.M.I.; funding acquisition, N.M.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Universiti Putra Malaysia, grant number GP-GPB 9711400.

Data Availability Statement: The original contributions presented in the study are included in the article/Supplementary Materials, further inquiries can be directed to the corresponding author/s.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Gamachu, D.; Ibrahim, W. Mixed Convection Flow of Viscoelastic Ag-Al₂O₃/Water Hybrid Nanofluid Past a Rotating Disk. *Phys. Scr.* **2021**, *96*, 125205. [\[CrossRef\]](#)
2. Kármán, T.V. Über Laminare Und Turbulente Reibung. *ZAMM J. Appl. Math. Mech. Z. Angew. Math. Mech.* **1921**, *1*, 233–252. [\[CrossRef\]](#)
3. Ayano, M.S.; Otegbeye, O.; Mathunjwa, J.S. Numerical Simulation Of Nanofluid Flow Due To A Stretchable Rotating Disk. *Theor. Appl. Mech.* **2023**, *50*, 55–72. [\[CrossRef\]](#)
4. Sharma, K.; Vijay, N.; Bhardwaj, D.; Jindal, R. Flow of Water Conveying Fe₃O₄ and Mn–ZnFe₂O₄ Nanoparticles over a Rotating Disk: Significance of Thermophoresis and Brownian Motion. *J. Magn. Magn. Mater.* **2023**, *574*, 170710. [\[CrossRef\]](#)
5. Puneeth, V.; Aly, E.H.; Pop, I. Nanofluid Flowing over a Rotating Disk That Is Stretching and Permeable: An Unsteady Model. *Int. J. Mod. Phys. B* **2023**, *37*, 23502491. [\[CrossRef\]](#)
6. Rashid, A.; Dawar, A.; Ayaz, M.; Islam, S.; Galal, A.M.; Gul, H. Homotopic Solution of the Chemically Reactive Magnetohydrodynamic Flow of a Hybrid Nanofluid over a Rotating Disk with Brownian Motion and Thermophoresis Effects. *ZAMM Z. Angew. Math. Mech.* **2023**, *103*, e202200262. [\[CrossRef\]](#)
7. Alkuhayli, N.A.M. Heat Transfer Analysis of a Hybrid Nanofluid Flow on a Rotating Disk Considering Thermal Radiation Effects. *Case Stud. Therm. Eng.* **2023**, *49*, 103131. [\[CrossRef\]](#)
8. Kumar, P.; Poonia, H.; Areekara, S.; Sabu, A.S.; Mathew, A.; Ali, L. Significance of Irregular Heat Source and Arrhenius Energy on Electro-Magnetohydrodynamic Hybrid Nanofluid Flow over a Rotating Stretchable Disk with Nonlinear Radiation. *Numer. Heat. Transf. A Appl.* **2024**, *85*, 1866–1888. [\[CrossRef\]](#)
9. Kulkarni, M. Mixed Convective Magnetized GO–MoS₂/H₂O Hybrid Nanofluid Flow about a Permeable Rotating Disk. *Asia Pac. J. Chem. Eng.* **2023**, *18*, e2923. [\[CrossRef\]](#)
10. Lone, S.A.; Anwar, S.; Raizah, Z.; Almusawa, M.Y.; Saeed, A. A Theoretical Investigation of the MHD Water-Based Hybrid Nanofluid Flow over a Convectively Heated Rotating Disk Surface with Velocity Slips and Zero Mass Flux Conditions. *Eur. Phys. J. Plus* **2023**, *138*, 1–14. [\[CrossRef\]](#)
11. Khashi'ie, N.S.; Waini, I.; Zainal, N.A.; Hamzah, K.; Arifin, N.M.; Pop, I. Multiple Solutions and Stability Analysis of Magnetic Hybrid Nanofluid Flow Over a Rotating Disk with Heat Generation. *J. Adv. Res. Fluid. Mech. Therm. Sci.* **2023**, *102*, 59–72. [\[CrossRef\]](#)
12. Vijay, N.; Sharma, K. Magnetohydrodynamic Hybrid Nanofluid Flow over a Decelerating Rotating Disk with Soret and Dufour Effects. *Multidiscip. Model. Mater. Struct.* **2023**, *19*, 253–276. [\[CrossRef\]](#)
13. Abu Bakar, S.; Pop, I.; Md Arifin, N. Unsteady Flow of Hybrid Nanofluid over a Permeable Shrinking Inclined Rotating Disk with Radiation and Velocity Slip Effects. *Neural Comput. Appl.* **2024**, *36*, 11525–11544. [\[CrossRef\]](#)
14. Ali, F.; Zaib, A.; Khan, M.I.; Alzahrani, F.; Eldin, S.M. Irreversibility Analysis in Stagnation Point Flow of Tri-Hybrid Nanofluid over a Rotating Disk; Application of Kinetic Energy. *J. Indian. Chem. Soc.* **2023**, *100*, 100873. [\[CrossRef\]](#)
15. Singh, S.P.; Kumar, M.; Yaseen, M.; Rawat, S.K. Ternary Hybrid Nanofluid (TiO₂–SiO₂–MoS₂/Kerosene Oil) Flow over a Rotating Disk with Quadratic Thermal Radiation and Cattaneo-Christov Model. *J. Cent. S. Univ.* **2023**, *30*, 1262–1278. [\[CrossRef\]](#)
16. Patgiri, B.; Paul, A. Inspection of Viscoelastic Ag+Cu+Fe₃O₄+Al₂O₃/Kerosene Oil Tetra-Hybrid Nanofluid Flow across a Stretchable Rotating Disk with Exponentially Varying Viscosity. *J. Taibah Univ. Sci.* **2024**, *18*, 2336327.
17. Ali, I.; Gul, T.; Khan, A. Unsteady Hydromagnetic Flow over an Inclined Rotating Disk through Neural Networking Approach. *Mathematics* **2023**, *11*, 1893. [\[CrossRef\]](#)
18. Varun Kumar, R.S.; Kumar, R.N.; Ben Ahmed, S.; Madhu, J.; Verma, A.; Punith Gowda, R.J. Unsteady Flow of a Ternary Nanofluid over a Slow-Rotating Disk Subject to Uniform Suction and Backpropagated Neural Network. *Numer. Heat Transf. Part B Fundam.* **2023**, 1–21. [\[CrossRef\]](#)
19. Mehmood, T.; Ramzan, M.; Howari, F.; Kadry, S.; Chu, Y.M. Application of Response Surface Methodology on the Nanofluid Flow over a Rotating Disk with Autocatalytic Chemical Reaction and Entropy Generation Optimization. *Sci. Rep.* **2021**, *11*, 4021. [\[CrossRef\]](#)
20. Shafiq, A.; Sindhu, T.N.; Al-Mdallal, Q.M. A Sensitivity Study on Carbon Nanotubes Significance in Darcy–Forchheimer Flow towards a Rotating Disk by Response Surface Methodology. *Sci. Rep.* **2021**, *11*, 8812. [\[CrossRef\]](#)
21. Mathew, A.; Areekara, S.; Sabu, A.S. Significance of Magnetic Field and Stratification Effects on the Bioconvective Stagnation-Point Flow of Ferro-Nanofluid over a Rotating Stretchable Disk: Four-Factor Response Surface Methodology. *J. Indian Chem. Soc.* **2022**, *99*, 100615. [\[CrossRef\]](#)
22. Hussain, S.; Ali, A.; Rasheed, K.; Pasha, A.A.; Algarni, S.; Alqahtani, T.; Irshad, K. Application of Response Surface Methodology to Optimize MHD Nanofluid Flow over a Rotating Disk with Thermal Radiation and Joule Heating. *Case Stud. Therm. Eng.* **2023**, *52*, 103715. [\[CrossRef\]](#)
23. Khashi'ie, N.S.; Waini, I.; Hamzah, K.B.; Mukhtar, M.F.; Kasim, A.R.M.; Arifin, N.M.; Pop, I. Numerical Solution and Statistical Analysis of the Unsteady Hybrid Ferrofluid Flow with Heat Generation Subject to a Rotating Disk. *ZAMM J. Appl. Math. Mech. Z. Angew. Math. Mech.* **2023**, *103*, e202200384. [\[CrossRef\]](#)
24. Siddiqui, F.R.; Tso, C.Y.; Chan, K.C.; Fu, S.C.; Chao, C.Y.H. On Trade-off for Dispersion Stability and Thermal Transport of Cu–Al₂O₃ Hybrid Nanofluid for Various Mixing Ratios. *Int. J. Heat Mass Transf.* **2019**, *132*, 1200–1216. [\[CrossRef\]](#)

25. Takabi, B.; Gheithaghy, A.M.; Tazraei, P. Hybrid Water-Based Suspension of Al_2O_3 and Cu Nanoparticles on Laminar Convection Effectiveness. *J. Thermophys. Heat Trans.* **2016**, *30*, 523–532. [[CrossRef](#)]
26. Oztop, H.F.; Abu-Nada, E. Numerical Study of Natural Convection in Partially Heated Rectangular Enclosures Filled with Nanofluids. *Int. J. Heat Fluid Flow* **2008**, *29*, 1326–1336. [[CrossRef](#)]
27. Suresh, S.; Venkitaraj, K.P.; Selvakumar, P.; Chandrasekar, M. Synthesis of Al_2O_3 -Cu/Water Hybrid Nanofluids Using Two Step Method and Its Thermo Physical Properties. *Colloids Surf. A Physicochem. Eng. Asp.* **2011**, *388*, 41–48. [[CrossRef](#)]
28. Turkyilmazoglu, M. Nanofluid Flow and Heat Transfer Due to a Rotating Disk. *Comput. Fluids* **2014**, *94*, 139–146. [[CrossRef](#)]
29. Yin, C.; Zheng, L.; Zhang, C.; Zhang, X. Flow and Heat Transfer of Nanofluids over a Rotating Disk with Uniform Stretching Rate in the Radial Direction. *Propuls. Power Res.* **2017**, *6*, 25–30. [[CrossRef](#)]
30. Box, G.E.P.; Wilson, K.B. On the Experimental Attainment of Optimum Conditions. *J. R. Stat. Soc. Ser. B (Methodol.)* **1951**, *13*, 1–38. [[CrossRef](#)]
31. Mahanthesh, B.; Thriveni, K.; Rana, P.; Muhammad, T. Radiative Heat Transfer of Nanomaterial on a Convectively Heated Circular Tube with Activation Energy and Nanoparticle Aggregation Kinematic Effects. *Int. Commun. Heat Mass Transf.* **2021**, *127*, 105568. [[CrossRef](#)]
32. Bhattacharya, S. Central Composite Design for Response Surface Methodology and Its Application in Pharmacy. In *Response Surface Methodology in Engineering Science*; IntechOpen: London, UK, 2021.
33. Sahoo, P.; Barman, T.K. ANN Modelling of Fractal Dimension in Machining. In *Mechatronics and Manufacturing Engineering*; Woodhead Publishing: Cambridge, UK, 2012.
34. Rashidi, M.M.; Abelman, S.; Mehr, N.F. Entropy Generation in Steady MHD Flow Due to a Rotating Porous Disk in a Nanofluid. *Int. J. Heat Mass Transf.* **2013**, *62*, 515–525. [[CrossRef](#)]
35. Alhadri, M.; Raza, J.; Yashkun, U.; Lund, L.A.; Maatki, C.; Khan, S.U.; Kolsi, L. Response Surface Methodology (RSM) and Artificial Neural Network (ANN) Simulations for Thermal Flow Hybrid Nanofluid Flow with Darcy-Forchheimer Effects. *J. Indian Chem. Soc.* **2022**, *99*, 100607. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Review

Computational Fluid Dynamics-Based Systems Engineering for Ground-Based Astronomy

Konstantinos Vogiatzis ^{1,2,*}, George Angeli ¹, Gelys Trancho ² and Rod Conan ¹

¹ Giant Magellan Telescope Organization, Pasadena, CA 91101, USA; gyangeli@gmail.com (G.A.); rconan@gmto.org (R.C.)

² TMT International Observatory, Pasadena, CA 91124, USA; gtranco@tmt.org

* Correspondence: konvoyages@gmail.com

Abstract: This paper presents the state-of-the-art techniques employed in aerothermal modeling to respond to the current observatory design challenges, particularly those of the next generation of extremely large telescopes (ELTs), such as the European ELT, the Thirty Meter Telescope International Observatory (TIO), and the Giant Magellan Telescope (GMT). It reviews the various aerothermal simulation techniques, the synergy between modeling outputs and observatory integrating modeling, and recent applications. The suite of aerothermal modeling presented includes thermal network models, Computational Fluid Dynamics (CFD) models, solid thermal and deformation models, and conjugate heat transfer models (concurrent fluid/solid simulations). The aerothermal suite is part of the overall observatory integrated modeling (IM) framework, which also includes optics, dynamics, and controls. The outputs of the IM framework, nominally image quality (IQ) metrics for a specific telescope state, are fed into a stochastic framework in the form of a multidimensional array that covers the range of influencing operational parameters, thus providing a statistical representation of observatory performance. The applications of the framework range from site selection, ground layer characterization, and site development to observatory performance current best estimate and optimization, active thermal control design, structural analysis, and an assortment of cost–performance trade studies. Finally, this paper addresses planned improvements, the development of new ideas, attacking new challenges, and how it all ties to the “Computational Fluid Dynamics Vision 2030” initiative.

Keywords: computational fluid dynamics; aerothermal modeling; systems engineering; ground-based telescopes



Citation: Vogiatzis, K.; Angeli, G.; Trancho, G.; Conan, R. Computational Fluid Dynamics-Based Systems Engineering for Ground-Based Astronomy. *Computation* **2024**, *12*, 143. <https://doi.org/10.3390/computation12070143>

Academic Editor: Dimitrios Koubogiannis

Received: 10 June 2024

Revised: 29 June 2024

Accepted: 3 July 2024

Published: 11 July 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

A significant fraction (approximately one-third) of the total image aberration of ground-based optical observatories has aerothermal causes, most of it attributed to thermal seeing (~80%), wind jitter, and thermal deformations of the telescope structure and optics. These aberrations also carry the largest uncertainty due to their random nature. Through the years, various empirical aerodynamic and thermal management design and operating strategies have been developed and implemented. In the last two decades, however, the optimization of the configuration and operation of these tools is performed by modeling the mountain–observatory interaction and the enclosure interior–structure–optics interaction.

In this paper, we present the state-of-the-art techniques employed in aerothermal modeling to address the current observatory design challenges, particularly those of the next generation of ELTs (extremely large telescopes). This also happens to be the acronym of the first one, currently under construction in Chile by the European Southern Observatory (ESO), with a primary mirror diameter of ~39 m [1]. To avoid confusion, it is often called E-ELT (European ELT or ESO ELT). Across the Atlantic, two such projects exist, the Giant Magellan Telescope (GMT) [2] and the Thirty Meter Telescope International Observatory (TIO) [3]. The two telescopes of ~25 m and ~30 m primary mirror diameter and currently

planned for the southern and northern hemispheres, respectively, can provide diverse research opportunities and enable pivotal breakthroughs in nearly all areas of astrophysics from our Solar System to the most distant stars and galaxies, from fundamental physics and cosmology to the search for evidence of life on planets around other stars.

Recent aerothermal modeling applications range from site selection, ground layer characterization, and site development to observatory aerothermal design optimization, thermal control design, structural analysis, and an assortment of cost–performance trade studies. A review of the aerothermal model types employed by the TIO and GMT will be presented. These models are extendable to all optical observatories, including solar, and aspects of them even to radio telescopes. The synergy between modeling outputs (such as surface pressures, integrated forces, moments, temperature fields, and heat transfer coefficients) and observatory integrating modeling (IM) will follow. Finally, planned improvements, the development of new ideas, attacking new challenges, and how it all ties to the “Computational Fluid Dynamics Vision 2030” will be visited.

2. Integrated Modeling Synergy

2.1. Performance Metrics and Requirement Validation

The ELT projects established Key Performance Parameters (KPPs) for measuring, tracking, and managing the evolution of expected observatory performance through construction and commissioning.

The general foundation documents start at the Operations Requirements Document (for TIO) or Concept of Operations (for GMT) and the Science Requirements Document (SRD) (all Level 1) which are derived from predetermined science cases. The SRD requirements are then flowed down to the Observatory Requirements Document (ORD; Level 2) which is the project’s response to the SRD and contains technical requirements that establish the baseline for the construction project. Requirements in the ORD are given in engineering units that can be measured and verified during the Assembly, Integration, Verification, and Commissioning (AIVC) phase of the project.

The KPPs are defined in the ORD and are the engineering requirements that would have the most significant impact to science if they were not met. Many of the Level 2 KPPs link to several of the science requirements. Every Level 2 requirement also has an associated performance budget in the Level 3 Observatory Architecture Document (OAD). Figure 1 indicates the current best estimate (CBE) analysis: allocations from the error budgets, including the KPP budgets, are flowed down in engineering units to the Level 4 subsystem requirements documents.



Figure 1. Cycle of compliance and CBE analysis.

2.2. Integrated Modeling Framework

The integrated model is a single computing framework that combines subsets of specialized models, which are detailed numerical representations of specific properties and functionalities of an ELT [4,5]. It is a composition of several physical models, part of which are also the aerothermal models, and a distributed control system that binds together the sensors and actuators emulated by the different models. An IM instance is a time domain simulation configured with a set of parameters representing a telescope design specification. In addition,

models of environmental disturbances provide external boundary conditions to the physical models. Examples of external disturbances are atmosphere and local seeing, wind forces, heat sources, etc. Environmental disturbances are stochastic by nature, and one random occurrence of the disturbances is drawn for each model instance. It follows that optical performance is stochastic too, and the performance of a telescope design is established based on the statistical moments of the optical performance metric (in this case, KPP). Thus, the goal of the IM is to estimate the Level 3 error budget terms and the Level 2 KPPs.

Each simulation of the IM generates an output (KPP value) that is valid for a particular set of disturbance values, called the System State. Each disturbance (or influencing parameter) is discretized with several values that cover its expected range. To span the entire operating envelope, the simulation must be repeated for a different combination of disturbance values. Thus, a multidimensional array of KPP values or look-up tables (LUTs) is generated.

2.3. Stochastic Framework

The observatory system can be considered a stochastic process; its behavior is non-deterministic, in that the system's subsequent state is determined both by predictable actions and by random elements. Consequently, any metric describing the system performance is a statistical variable. While some performance aspects can be predicted deterministically given a set of inputs, these inputs themselves are usually statistical in nature [6,7].

A wide range of imperfections and external disturbances are truly random in time or through realizations. As an example, the optical effects of thermal deformations of optical elements and their support structure can be deterministically calculated. The inputs to thermal deformation calculations are random environmental and operational parameters. The usual response to this complexity is to intuitively define "representative" and "worst case" scenarios and then to calculate the "mean" and "limiting" performance.

To gain a better insight, we developed a stochastic framework for defining and assessing performance. It is based on site testing data collected by the projects at the sites of the observatories, Maunakea and Las Campanas, as well as on the history of operation sequences of the Gemini North and Magellan observatories [8,9]. The performance of an ELT is expressed for and evaluated against a real observing program carried out in the same environment where it is to be built.

It is not practical for a single, unified model to run through all time frames relevant to an ELT's performance, from milliseconds to years. Therefore, our approach is to separate physical processes into distinct mathematical models that can be explored somewhat independently. Using the Point Source Sensitivity-Normalized (PSSn) metric [10–12] as an example KPP, individual results can be combined to system-level estimates. There are noted exceptions to this approach. In some cases, the time frames of coupled processes overlap sufficiently to justify the development of integrated models. A particular case involves wind, telescope structural dynamics, and control dynamics. For ELTs, this is thoroughly investigated, including the wind buffeting of the primary mirror and the close interaction between the control and structure. Similarly, image jitter due to telescope vibrations is directly addressed in adaptive optics simulations. Over longer time frames, solid (glass and steel) and aerothermal processes are closely coupled. Addressing this coupling requires iterations between the CFD and FEA models or the development of high-resolution conjugate heat transfer models.

The essence of the performance framework is depicted in Figure 2. Group A (red) denotes the available inputs. Group B (blue) is what can be controlled by design and operation. Group C (brown) shows the developed integrated modeling framework output with the pre-calculated aberrations in LUT form (the optics/controls/dynamics/CFD/FEA box is actually the integrated modeling framework). Finally, functions in Group D (green) correspond to mathematical/statistical processing and are in fact the core of the time domain simulation.

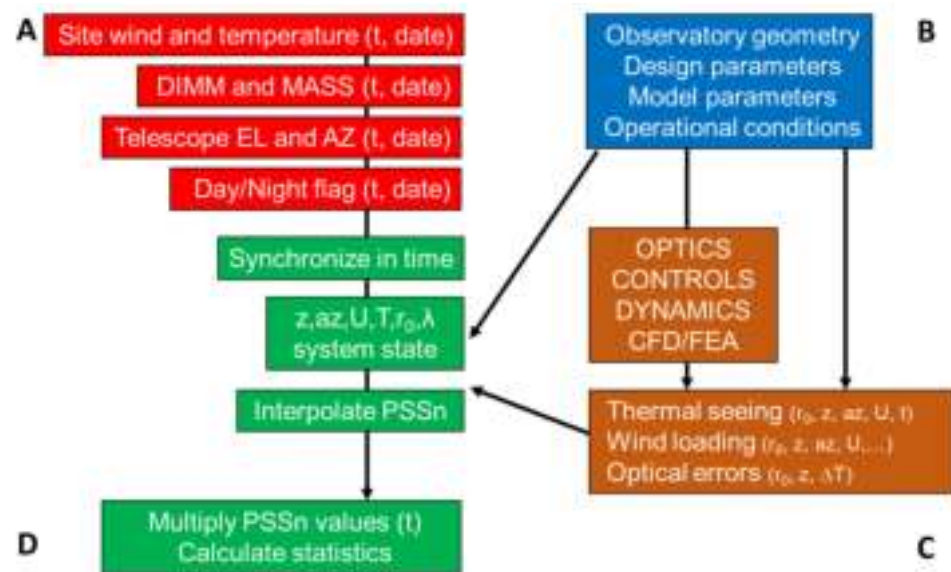


Figure 2. Stochastic framework flow-chart. (A) environmental inputs, (B) controlled parameters, (C) integrated modeling outputs, (D) data fusion and statistics.

The advantages of this framework are summarized below:

- Allows the statistical representation of errors that are randomized by environmental and operating conditions;
- Allows the statistical representation of errors that depend on previous temporal states of the system;
- Allows the combined representation of errors that are correlated and cannot be assessed independently;
- Allows the combined representation of errors that exhibit disparate time and spatial scales;
- Provides probability distributions of errors to denote the range of expected performance under various seeing conditions, knowledge particularly useful to queue observing;
- Allows the investigation of observatory behavior on different candidate sites;
- Allows the optimization of certain operating strategies, such as venting, daytime thermal treatment, and calibration measurement frequency;
- Unveils the true impact of a particular design decision on the expected observatory performance and subsequent science productivity, enabling cost-effective trade-offs.

3. Aerothermal Modeling Synergy

Thermal management is key to controlling observatory image quality. As tools to minimize thermal effects, ELTs use air conditioning during the day and passive ventilation and heat release control in various forms (subsystem cooling, proper location of heat sources, coatings of desired emissivity) at night. Enclosure surface emissivity and reflectivity are optimized to reduce overheating during the day and overcooling at night.

It is prudent to minimize the difference between the temperature of the telescope (structure and optics) and the average temperature expected at night. Especially for the optics, given their time constant, this is important both for mitigating thermal deformation and mirror seeing. In addition, spatial temperature gradients will be minimized since they are a major contributor to telescope misalignment.

At sunrise, the enclosure will close, and the air trapped inside will start warming up due to the daytime heat release of equipment operating inside the enclosure, solar radiation, and the infiltration of external air. Air Handling Units (AHUs) are connected to chillers/dry coolers in the facilities area. The AHU nozzles supply air to the enclosure interior at several locations, while the returns are located near the observing floor. The goal is to minimize vertical temperature gradients at least up to the height of the telescope and maintain the

environment at the expected nighttime temperature by removing all internally generated and externally imposed heat.

At night, radiation loss to the sky through the enclosure opening is balanced by convective heat transfer through the enclosure opening and vents, as well as the heat released into the interior air. The non-uniform spatial distribution of released heat can be particularly detrimental to dome and mirror seeing. Its effect is mitigated by flushing the enclosure with ambient air through the vents. However, at higher wind speeds, enclosure flushing needs to be restricted to limit wind buffeting of the telescope structure and the mirrors. The venting strategy is optimized to balance the thermal seeing and wind buffeting effects. As a baseline, the enclosure vents are expected to be fully open at low external wind speeds, fully closed at high external wind speeds, and partially open at median external wind speeds.

The waste heat removed from the observatory is concentrated into exhaust plumes vented from the farthest point of the facility building which is in the predominantly downwind direction from telescope. During nighttime, the plume's maximum temperature deviation from ambient is restricted to ensure that any effects on seeing in this direction are negligible.

The models used to simulate the above physical processes are grouped into three main categories based on scope, geometric fidelity, and turnaround time:

- Thermal network models;
- CFD models;
- Conjugate heat transfer models.

Well-known commercial off-the-shelf general-purpose software suites are employed to develop the modeling tools. STAR-CCM+ (2210 or later) [13] is employed to develop three-dimensional representative observatory and subsystem models, while one-dimensional thermal network models, performance frameworks, and post-processors are developed primarily in MATLAB (2020a or later) [14] and Python (3.7 or later).

3.1. Thermal Network Models

These are lumped-mass or one-dimensional transient thermal models. They are computationally inexpensive with short turnaround times. Their main purpose is to estimate surface temperatures, such as for the enclosure skin, as well as other components. These temperatures are used to support the design, to validate requirements, and as boundary conditions to high-fidelity models. Since they are thermal network models, they can accept air as another node so they can also report the mean temperature behavior of a confined volume of air. The true usefulness of these models lies in the fact that the environmental and operational input conditions along with the virtual simulation time span years, thus allowing for the statistical analysis of temperature differentials.

3.2. Computational Fluid Dynamics Models

These are high-fidelity, three-dimensional steady or transient models. They are computationally expensive with long turnaround times. As the name suggests, their purpose is to estimate fluid (air) velocity, temperature, and pressure behavior in space and time, at the ELT sites, around and inside the enclosures, and inside components.

Steady simulations above topography provide ground layer (GL) estimates for site characterization studies. Unsteady daytime-long simulations inside the enclosure provide HVAC performance. Finally, unsteady nighttime simulations provide aerothermal observatory performance, through dome seeing and telescope wind loading. The simulation times for the latter are in the order of minutes.

3.3. Solid and Conjugate Heat Transfer Models

These are three-dimensional transient models. In general, they are computationally inexpensive with short turnaround times, with a few exceptions. Their purpose is not only to estimate component thermal inertia behavior in time but also temperature spatial variability and the resulting deformation. In cases where the effect of a component on

a confined volume of air is needed or convective heat transfer variability is important, solid/fluid interaction (conjugate) models are developed. The simulation time for these models is normally a few diurnal cycles.

3.4. Post-Processing Tools

A series of post-processing tools also exist to handle the various intermediate outputs and provide results in the final variable format. Some are incorporated in the software simulation files, and some are standalone.

Figure 3 depicts the current typical ELT aerothermal modeling synergy. The suite consists of a set of thermal network, CFD, and conjugate heat transfer models. It also incorporates a special structural model that has been developed just for thermal deformations. The purpose of these models is not only to provide specific outputs for IQ estimates but also to complement and support each other. Blue corresponds to the high-fidelity CFD models, gold to the conjugate heat transfer, thermal network, and thermal deformation models, green are IQ tools that belong to the IM framework, and yellow are the input/output variables. The CFD models provide heat transfer coefficients to the mirror and telescope structure Finite Element Analysis (FEA) solid thermal models, which in turn provide temperature and heat fluxes to the CFD models as boundary conditions for the telescope and optical surfaces. The network thermal models provide heat fluxes to the CFD models for the enclosure surfaces and other components. They can also provide the FEA models with correct radiation properties. Finally, they can provide fast qualitative information about optimum daytime thermal control strategies to minimize the computationally intensive daytime CFD simulations. The convective heat transfer coefficients needed by the thermal network and FEA models are provided by CFD simulations.

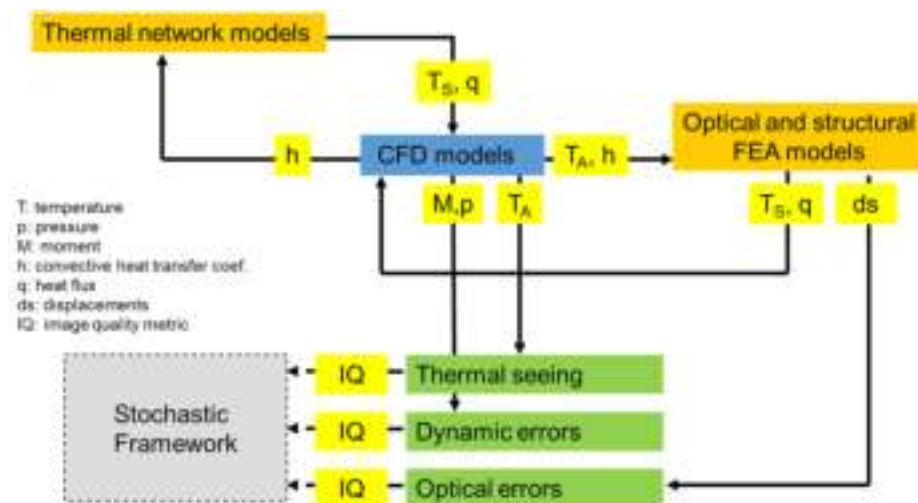


Figure 3. Aerothermal modeling information flow-chart.

The aerothermal modeling suite can provide a wide range of outputs to validate and/or verify subsystem design requirements. The thermal CFD simulations yield force/moment and temperature fields used, respectively, by the dynamic and thermal seeing blocks, while the thermal FEA simulations yield displacements used directly by the optical block. Thus, it can produce estimates of subsystem aerothermal-related error budget terms and system-level performance budgets.

The available inputs, parameters that can be controlled by design and operation, and pre-calculated aberrations from the IM blocks in multidimensional matrix form (look-up table, LUT) are combined inside the stochastic framework to define the System State in time. Finally, functions corresponding to mathematical/statistical processing produce system-level performance parameters, such as the PSSn.

3.5. Subsystem-Level Inquiry Processing

ELT IM teams have an internal plan for aerothermal model development. The scheduling is tied to system-level reviews, while the specific models are prioritized according to the corresponding error budget term significance. The need for initial estimates of local seeing and telescope wind loading dictated that the first model to be developed and the first subsystem design to be supported and optimized should be the enclosure. A conceptual design-level telescope would be placed inside. Thermal boundary conditions would come from an initial thermal network model that considered large surface areas that directly affect the optical path. The rest of the heat sources would come from the corresponding thermal requirements. This constituted the observatory CFD model. Note here that there exists a nighttime and a daytime version of this model. The model would be updated annually with increasing geometric fidelity. Subsystem models would be developed either to increase thermal boundary condition fidelity for the observatory model, thus increasing performance estimate accuracy, or to support subsystem reviews, as their design matured. This latter process generates what is referred to herein as a subsystem aerothermal inquiry.

The subsystem team contacts the IM team with an inquiry (Figure 4). Typical inquiries can be design trade-off studies (geometric considerations, equipment selection, flow rate and heat dissipation options), input data for survival analysis, and the performance impact of irreversible design choices. These can often be review action items. A meeting or a series of meetings between the teams are held to trace the subsystem requirements directly involved (Level 4) and the corresponding performance requirements (typically Level 3). This is important to define the scope of the action expected by the IM team: what models are required, with what output, and by when. Once the scope is defined, all prerequisites can be identified and required resources estimated.

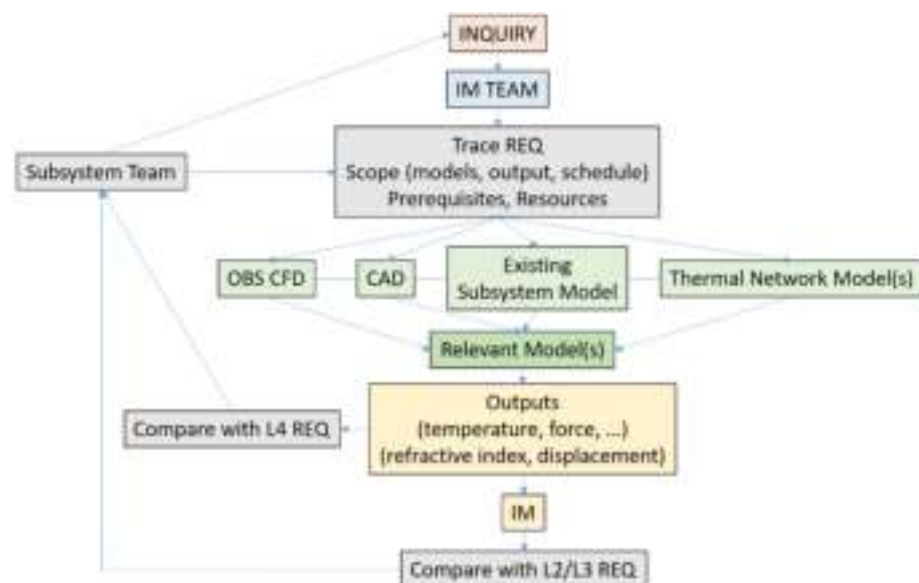


Figure 4. Representative ELT aerothermal requirement validation process.

Prerequisites include generic inputs (wind speed, telescope orientation, component power/flow rates, etc.), an existing subsystem model, which must be used with different inputs, a CAD model for a new geometry or to update an existing model representation, thermal network models for additional temperature calculations, and system- or subsystem-level heat dissipation budgets. Note that the overwhelming majority of requests to date require information provided by the observatory CFD model.

Often, either the urgency or time scale associated with the inquiry is not compatible with the estimated resources, leading to a de-scoping exercise (minimum amount of effort to produce acceptable answers).

Once all the relevant models and inputs have been identified, the simulations are performed, and outputs are collected. Integrated modeling may or may not be required based on the level of requirements to be validated (or verified). It is common for the first round of results to generate additional inquiries and modifications, along with possibly new model development.

Over the next several sections, the various types and major applications of the aerothermal models will be visited in more detail. To close this section, a summary of the impact aerothermal modeling had on ELT design is presented in Table 1, mostly in timeline order (depending also on the maturity of subsystem design). The summary is not exhaustive. Most models are periodically updated, and their fidelity increases as a result of design maturity and modeling lessons learned. Also, there have been several instantiations of overall performance CBE prior to the development of many of these models. Note, however, that not all observatories followed strictly that order, especially in the site and enclosure selection and design stages.

Table 1. Aerothermal modeling impact on ELTs.

ELT Development Aspect	Modeling Action	Systems Engineering Function
Site selection	GL modeling. Comparison with site measurements.	Trade-off studies. Risk mitigation (model validation).
Site development	GL modeling: minimize impact on site quality, select specific location within site, and investigate interactions with existing observatories where applicable. Comparison with site measurements.	Trade-off studies. Risk mitigation (model validation).
Enclosure and facilities	Observatory modeling: enclosure type selection, aerodynamic shape optimization, survival wind loads, vent area, location, and number, cooling sizing and performance, thermal control of drives, utility equipment and interstitial spaces, surface temperatures/paint emissivity, and insulation thickness.	Trade-off studies. Requirements validation.
Telescope mount (structure)	Solid thermal modeling: thermal deformation, telescope pointing, surface temperatures/paint emissivity, and temperature sensor location/alignment look-up table. Observatory modeling: aerodynamic optimization of top end and secondary support structure.	Requirements validation. Trade-off studies.
Optical assemblies	Conjugate heat transfer modeling: surface and component temperatures, deformation wavefront error, thermal control, and effects on metrology sensors.	Requirements validation.
Instruments and Laser Guide Star Systems	Conjugate heat transfer modeling: surface and component temperatures, focus errors, internal optical path turbulence, and design optimization.	Requirements validation. Trade-off studies.
Performance CBE	Observatory modeling: thermal seeing/effects on metrology, wind loading on structure and optics, thermal deformation of structure and optics, and telescope pointing. Modeling existing observatories for comparison and mesh resolution studies (for multiple models).	Requirements verification. Risk mitigation (model validation).
AIVC/operations	Observatory modeling: aerothermal environment during construction, internal thermal gradients in absence of air conditioning, external wind profiles, and effects on metrology. Reduced Order Modeling (ROM) and Artificial Neural Network (ANN) modeling for first light characterization and operations optimization/queue observing.	Requirements validation. Requirements verification.

4. Thermal Network Modeling

The stochastic framework is the basis for any thermal network model to statistically cover a wide range of environmental conditions. Even though the main variable of interest is the ambient temperature, wind and telescope pointing variability provide the required diversity in convective heat transfer coefficients and radiative view factors for both external and internal enclosure surfaces, as well as telescope components. Diurnal environmental records are vital inputs for any thermal network model, since due to thermal inertia, component behavior exhibits “memory”; its current thermal state depends on previous thermal states.

The legacy version of this framework incorporated a lumped-mass thermal model for the TIO primary mirror, since mirror seeing could be calculated using a semi-empirical relation [15]. This relation has been since found not to be valid for giant mirrors and all flow patterns inside ELT enclosures and is not used anymore.

The framework has since been coupled with several thermal network models, which will be presented below. The output parameters of this framework that are used as inputs for other models or for design purposes (operating ranges) include the following:

- Ambient temperature drift since last telescope alignment, used by the telescope/mirror thermal deformation models;
- Temperature differential between sunrise and a particular time of night or mean of night, used for enclosure AHU sizing and to determine supply temperature;
- Temperature differential between daytime external ambient and interior temperatures, using the air conditioning model results, used in combination with external wind speed to determine the enclosure infiltration load.

4.1. Enclosure

Proper thermal boundary conditions on the enclosure surfaces are important for correctly capturing the temperature field inside the optical path. In places where the wind speed is low, diffusion and buoyancy dominate, and the temperature gradients on the surfaces transform into plumes that affect the optical path. Also, as the wind moves over the area around the enclosure aperture, the exterior surface temperature introduces a gradient which then passes through the optical path. In order to estimate the correct heat flux on the enclosure surfaces, a model that can take into account the constant movement of the enclosure components, conduction, convection, and, most importantly, radiation was implemented. It incorporates nodes for the enclosure skin, insulation, enclosure structure, and interior air. It can also provide information on temperature temporal gradients for the interstitial space air, observing floor, and even bulk telescope structure. It can estimate the limits on heat dissipation for specific components. Finally, it can estimate the daytime air conditioning power requirement due to solar radiation, infiltration, and interior heat sources.

A full 3D version of this model for the TIO was first presented in 2011, and the current version is included in [16]. The GMT equivalent can be found in [17].

The impact is as follows:

- Provides Cumulative Distribution Functions of component temperatures and fluxes: exterior skin, enclosure structure, interior insulation, observing floor, enclosure interior air, and interstitial space air;
- Requirements validation: exterior paint and insulation emissivity;
- Provides maximum expected temperature in interstitial space;
- Provides heat dissipation statistics through insulation and observatory floor;
- Provides input boundary conditions to observatory performance CFD simulations.

4.2. Optics

Even though 3D solid thermal models also exist for certain optical components, it is often useful to develop fast lumped-mass counterparts (ex. [18]). The main goal of these models is the requirement verification of allowed residual dissipation after cooling (if re-

quired) coupled with the statistical representation of compliance to a surface temperature requirement (typically for thermal seeing purposes).

The impact is as follows:

- Provides glass temperature differential from ambient under various environmental conditions;
- Validation/modification of thermal specifications for primary and secondary mirror systems.

5. Astronomical Site CFD Modeling

5.1. Site Selection

Very early, in 2002, a campaign to identify possible sites for the next generation of a U.S. observatory began. An integral part of this campaign was the use of CFD simulations for validation, site characterization, and short-listing, since equipment could not be deployed on all candidate sites.

Digital Elevation Maps (DEMs) of sites of interest had to be obtained and processed to generate computational domains and meshes. The DEMs covered existing observatory sites, such as Maunakea (Hawaii, USA), Las Campanas (Chile), Paranal/Armazones (Chile), Tololo/Pachon (Chile), San Pedro Martir (Baja peninsula, Mexico), and La Palma (Canary Islands, Spain), and some new ones in the Atacama Desert in Chile (Negro, Honar, Chajnantor, Tolonchar, Quimal) and even in Antarctica (Dome C).

Even though the simulations were initially steady-state and isothermal, focusing on topographically induced mechanical turbulence [19,20], the eventual scope included the resolution and characterization of the GL (optical turbulence), which is more cumbersome to measure experimentally. By 2007, significant progress, including initial model validation efforts, had been achieved on this front as well [21–24]. A functional dependence on thermal ε_T and mechanical ε_M turbulent dissipation was used to estimate the refractive index structure function parameter, $C_n^2 \sim A \varepsilon_T \varepsilon_M^{-1/3}$, from steady-state simulations, while the weak scattering theory was used during processing unsteady fields. The CFD simulations produced turbulent kinetic energy and dissipation rates that were subsequently used to calculate the turbulent length scales.

The required inputs include the following:

- Upwind wind and temperature conditions, which, along with topography temperature, result in a realistic GL profile between 7 m and 400 m from the ground. The only variables are wind speed and direction. The ground temperature was chosen to be a calculated quantity rather than input, either by imposing the expected net heat flux or by invoking a gray thermal radiation model with an emissivity of 0.8 and appropriate effective sky temperature. The reference location was that of the corresponding site monitoring system.
- Single reference pressure, reflecting the average site altitude. A user-defined function was used for the corresponding density vertical profile, which resulted in the expected reference density at the average site altitude.

The impact is as follows:

- Primary candidate ELT site selection;
- Risk reduction: validation of GL modeling (comparison with measurements where available);
- Given the appropriate inputs, modeled GL profiles are realistic, while measurement and modeling uncertainties for optical turbulence proved comparable, making the use of thermal modeling an economic and flexible alternative to experiments and in some cases the only option.

5.2. Ground Layer

5.2.1. Maunakea, 13N

The 13N site on Maunakea (HI, USA) is the selected site for the TIO. To that end, a $20 \times 13 \text{ km}^2$ DEM of Maunakea at 10 m resolution was obtained. The topographic boundary

of the computational domain is centered at UTM-5Q [X, Y] coordinates of [241000, 2192500]. Due to the smooth slopes, the height of the computational domain was chosen to be 8000 m for full-scale simulations and 5000 m for simulations confined to $3.5 \times 3.0 \text{ km}^2$.

During the TIO site selection campaign, temporal records of wind speed, direction, and temperature were collected above the site, along with optical turbulence data. Based on these and the fact that 13N is on the north side, below the summit, steady-state simulations were performed for three wind directions (E, SE, W) under median wind speeds [25].

5.2.2. Las Campanas

Las Campanas Peak (Chile) is the selected site for the GMT. To that end, a $22 \times 24 \text{ km}^2$ DEM at 20m resolution was fused with a $2 \times 2 \text{ km}^2$ high-resolution DEM (5 m) of the summit. The topographic boundary of the computational domain is centered at Las Campanas, with UTM-19J [X, Y] coordinates of [336216, 6785453]. The height of the computational domain was chosen to be 5000 m.

Temporal records of wind speed, direction, and temperature, along with turbulent inner-scale and C_n^2 values, were collected from the dual-tower environmental monitoring station. Both steady and unsteady CFD simulations were performed for the predominant site conditions, NNE median winds. A procedure was developed to minimize the uncertainty of topographic boundary conditions [26].

5.2.3. La Palma, ORM

The Observatorio del Roque de Los Muchachos (ORM) on the Canary Island of La Palma (Spain) has been selected as the alternate site for the TIO. Several potential locations on the summit needed to be investigated in terms of GL strength. Moreover, the presence of existing observatories necessitated a study of the interaction between these observatories and the TIO. The lack of localized site testing and the nature of the terrain led to the use of CFD simulations combined with a seeing model for GL optical turbulence estimates in 2016 [27].

The location of the ORM on the island of La Palma in combination with rugged terrain and the unavailability of realistic input profiles close to the summit dictated a large computational domain that can encompass almost the entire island with an additional buffer zone at sea level for realistic GL development. At the boundaries of such a domain, velocity and temperature profiles are known with higher certainty. To that effect, a $40 \times 40 \text{ km}^2$ DEM of the island at 50 m resolution was fused with a $4 \times 3 \text{ km}^2$ high-resolution DEM (10 m) of the ORM summit. The topographic boundary of the computational domain is centered at the ORM, with UTM-28R [X, Y] coordinates of [217900, 3185029]. The height of the computational domain was chosen to be 7500 m.

Steady-state simulations were performed for four different wind directions (NE, N, W, SW). Two separate studies were conducted. The first one focused on GL comparison between the four candidate sites. Then, for the preferred site based on the results and for two of these wind directions, the presence of nearby observatories such as the GTC and TNG was also considered, to determine mutual interference that could affect the local GL [27]. A more detailed report on the ORM GL variability is given in [28].

The additional impact is as follows:

- Alternate candidate TIO site location selection;
- Evaluation of interaction between the TIO and existing observatories, required by Instituto de Astrofísica de Canarias.

6. Observatory CFD Modeling

Appropriate fidelity CFD models for the observatories have been developed from 3D CAD models of the telescope mount and enclosure. They are used to simulate and analyze the aerothermal environment around the observatories. The developed multimillion-element models account for the major observatory components such as the primary (M1) and secondary (M2) assemblies, the secondary supporting truss-work, other subcomponents of the telescope mount, and the enclosure along with the auxiliary building(s) on the summits.

Upon the conception of the idea of an ELT, the immediate objection that arose was that excessive wind buffeting would render the project unfeasible. The telescope would be too exposed and not stiff enough to perform within the desired specifications. Years of extensive studies had to be performed to resolve the issue [29–32]. Wind pressure and velocity measurements inside the dome of the still-under-construction Gemini South, wind tunnel studies of generic ELT shapes at Caltech, at the NRC Institute for Aerospace Research in Canada, and in Europe, and CFD simulations and integrated modeling of the structure, controls, and optics finally suggested that an ELT, incorporating the right wind loading mitigation strategies, would not be a futile cause [33]. Armed with additional confidence, legacy studies were performed to converge to acceptable enclosure configurations by simulating conventional dome-shutter, carousel, and Calotte designs and multiple vent and wind mitigation configurations, as well as fixed-base and facility building options. The same techniques developed for the TIO and GMT were also employed for the design optimization of the Vera Rubin Observatory (Large Synoptic Survey Telescope at the time). CFD simulations literally shaped the facilities building architecture, enclosure, light baffles, and camera requirements [34–38].

Across the Atlantic, E-ELT has also performed some nighttime observatory-scale simulations, but they are not tied directly to optical performance, they are focused on velocity levels around the telescope and survival conditions [39].

Detailed reports of the simulation setup can be found in [40,41]. Frequent updates/modifications are necessary due to design evolution and lessons learned from studies, such as the ones reported in [28,42–45]. The most important ones are summarized in the following:

- The geometries are updated to reflect the latest designs and increased fidelity.
- The heat dissipation budgets are similarly updated, and additional boundaries are created to accommodate the resolved heat sources.
- Advances in meshing methodologies, solver development, and computational power are considered both for simulation setup and output post-processing.
- Upwind domain inputs have been revised to reflect more realistic GL conditions.
- The models retain the relative motion between tessellated parts and the corresponding local coordinates systems to support it. This way, the boundary conditions can be only applied once, and the model templates can be used for multiple orientations and wind speeds. However, for volume meshing, the parts are combined to minimize numerical errors generated at the volume interfaces, to which thermal seeing PSSn calculation is very sensitive. The procedure can still be automated through scripting, if desired.

6.1. Daytime (Air Conditioning)

Daytime CFD simulations were reported in [16,46] between 2010 and 2024 for the TIO and in [47] for the GMT. The simulations performed were selected to demonstrate that the enclosure Heating, Ventilation and Cooling (HVAC) system can perform as expected at least 90% of the time. This is based on the expected daytime heat budget and temperature swing between sunrise temperature and target temperature. The geometry corresponds to a stationary telescope and enclosure at their nominal parking positions.

The impact is as follows:

- Time records of enclosure air temperatures at various heights from the observatory floor and temperature vertical profiles at specified times during the day;
- Validation/verification of HVAC design (max flow rate capacity, split between primary supply and recirculation rates, nozzle location, operating strategy);
- Optimum HVAC nozzle temperature as a function of target temperature to meet performance requirements;
- Provides heat transfer coefficients for convection calculations to component thermal models.

6.2. Nighttime

The current aerothermal performance LUTs consist of results for 3 different telescope zenith angles (0° , 30° , 60° – 65°), 5 telescope relative-to-wind azimuth angles (0° , 45° , 90° ,

135°, 180°), 5–6 wind speeds (2–3 for open vents, 2–3 for closed), and up to 3 different sets of thermal boundary conditions, corresponding to different parts of the night (nominally ~1 h after sunset, middle of night, and before sunrise). An effective use of the model involves positioning the mesh at a particular orientation and then running one simulation varying the inlet wind speed between zero and some threshold with open vents. The transition from open to closed vents occurs for an inlet wind speed of ~5–7 m/s. Then, a separate simulation corresponding to wind speeds up to 17–18 m/s with closed vents is performed for that orientation. This way, most of the entire performance wind speed range is covered. The inlet wind speed function is defined so that enough time is spent at the LUT reference wind speeds to eliminate transient effects and accumulate enough time for statistics. This is particularly important for dome seeing due to the longer time scales associated with the diffusion of thermal boundary conditions to the enclosure interior air. Even though this approach has been used to estimate TIO aerothermal performance in the past [45], an alternative will be discussed in a subsequent section.

The additional required inputs are as follows:

- Current thermal boundary conditions on observatory surfaces. The values are grouped into three categories: standard (such as adiabatic, symmetry, etc.), estimated, and assumed. Estimated values are a result of modeling. Assumed values are for components that are either temperature-controlled and have a requirement specification or have insufficient information for accurate thermal modeling. Very few boundaries fall in this category since at least a thermal network model can provide some estimate for most components. For many subsystems, there already exist conjugate heat transfer models that provide surface temperatures. These include optical assemblies, adaptive optics facilities, and some “first light” instruments, with mature enough design. Most boundary conditions are applied as temperature differentials from the ambient reference temperature. In some cases, equilibrium is assumed between convection and radiation, ignoring conduction. These correspond mainly to low-thermal-inertia components, such as the secondary mirror assembly support structure and top end, enclosure and facilities building exterior, and interior surface cladding. Higher thermal inertia component temperatures, such as the enclosure concrete components and telescope lower structure, are given by the corresponding models. Note that proper boundary condition implementation ultimately requires knowledge of the heat balance of the observatory [48].

The outputs include the following:

- Three-dimensional temperature fields inside the optical volume from the M1 vertex until the end of the enclosure boundary layer at a specified temporal sampling rate. These are used as inputs in the thermal seeing post-processor [49,50]. Hence, a byproduct of these simulations is also a time record of Optical Path Difference (OPD) maps. For a historical overview of early thermal seeing modeling, the reader may consult [51,52].
- Time records of pressure maps on optical surfaces. These are used as inputs for wind-induced image blur calculation [53,54]. Pressure maps are also available upon request for any other resolved component.
- Time records of moment components around the telescope coordinate system origin from wind forces on the entire telescope. Forces and moments on isolated resolved telescope components are available upon request. These are used as inputs for image jitter calculation [9,42].
- Convective heat transfer coefficients on selected surfaces for use as inputs in solid and conjugate heat transfer models of components.

Figure 5 presents representative contour snapshots of the refractive index spatial gradient, a measure of optical turbulence, along the normal to the telescope elevation axis plane for the TIO and GMT. Similarly, representative instantaneous OPD maps are shown in Figure 6.

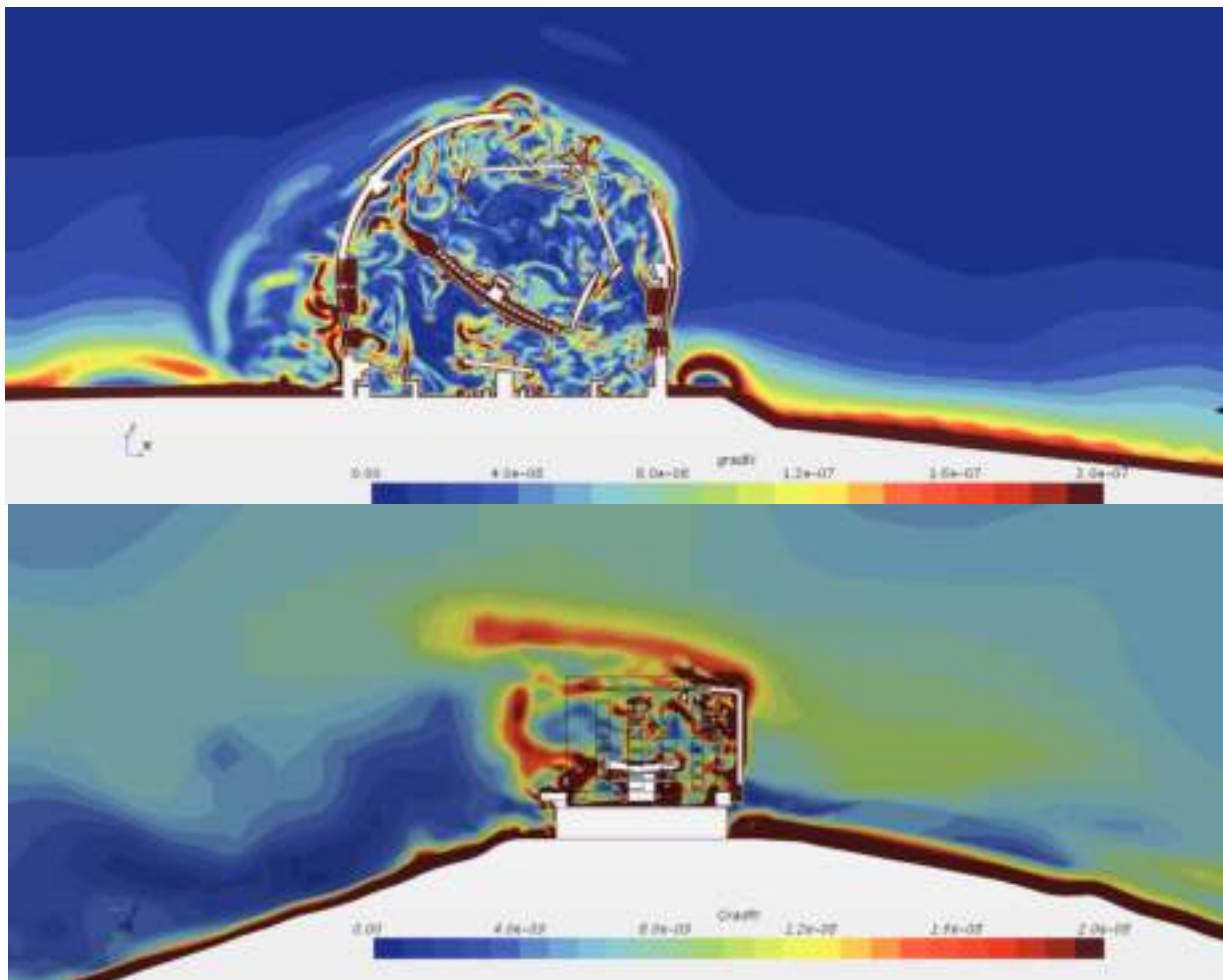


Figure 5. Representative contour snapshots of the refractive index spatial gradient along the normal to the telescope elevation axis plane: (**top**) TIO, 30° zenith and 45° azimuth relative to wind, and (**bottom**) GMT, 0° zenith and 180° azimuth relative to wind. Wind speeds are median (6–7 m/s), and venting configuration is “open”.

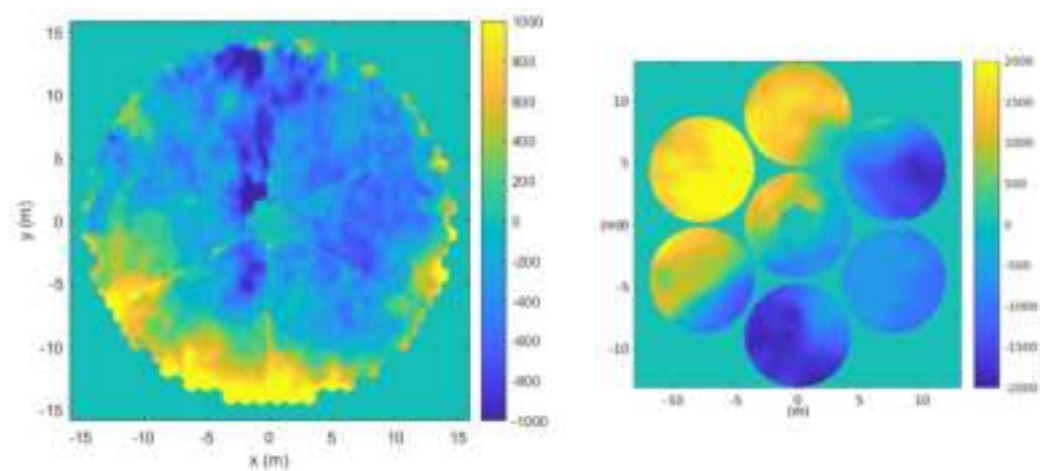


Figure 6. Representative OPD maps (in nm) for TIO (left) and GMT (right).

The impact is as follows:

- Observatory-wide aerodynamic optimization and enclosure operating strategy development through a trade-off between wind jitter and thermal seeing;

- Observatory performance and error budget terms estimate: thermal seeing, image jitter, and wind-induced image blur;
- Provides thermal seeing sensitivity to heat sources: heat dissipation budget update and design choices;
- Provides metrology system sensitivity to enclosure environment: expected measurement error;
- Provides heat transfer coefficients for convection and component view factors for radiation calculations to component thermal models.

6.3. Daytime (Solar)

It would be remiss not to mention the special category of observatories where the observations occur during the day: solar observatories. These performance estimate techniques have been recently extended to include enclosure and telescope platform optimization and the validation of requirements on heat shields and active thermal control for the European Solar Observatory (EST) [55].

6.4. Validation

By 2010, significant progress, including initial model validation efforts, had been achieved towards resolving and characterizing the GL [21,25]. In 2018, an additional validation effort was presented, using unsteady simulations this time [26]. With the proper knowledge of environmental conditions, current modeling methodologies can produce realistic and useful results. An illustrative example is depicted in Figure 7 for Las Campanas [28]. The GL is compared to that measured at Paranal, ESO's Very Large Telescope (VLT) site. Paranal is also very close to E-ELT's site. These profiles are used by both the GMT and E-ELT to train their Ground Layer Adaptive Optics (GLAO) modes of operation. Similar studies have been being performed for Maunakea and the ORM, as mentioned above.

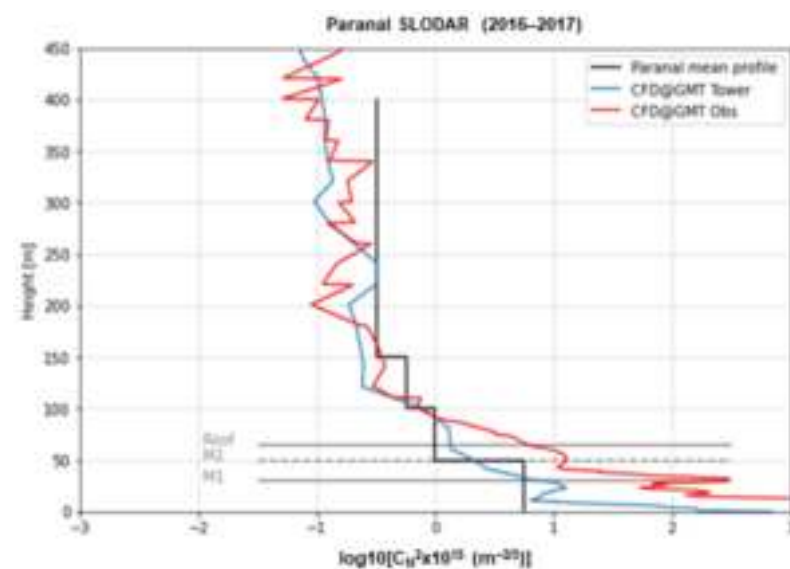


Figure 7. C_n^2 profile at Las Campanas at the monitoring tower and inside the GMT enclosure. The coarse SLODAR-measured profile from the nearby Paranal site is superimposed.

The procedure for nighttime observatory performance simulations with focus on thermal seeing has been validated by studies performed in 2011 and reported in [56,57]. They involved measurements and simulations at two observatories, Keck II and the CFHT (Canada–France–Hawaii Telescope), which share the TIO baseline site and thus computational domain and upwind conditions.

The goal of the studies from the TIO perspective was to validate the modeling procedure necessary to capture local seeing: the numerical method (solver settings and models),

required domain size, required observatory geometric fidelity, spatial and temporal resolution, boundary conditions, and representation of heat sources on observatory surfaces.

The Keck observatories were interested in optimizing the use of their active ventilation system to improve performance and minimize operating costs.

The CFHT was interested in the differential performance of a passive ventilation system before deciding to move forward with its approval.

A laser scintillometer was installed at Keck II that provided integrated C_n^2 and average optical turbulence inner-scale values on a 63 m path between M1 and M2, along with sonic anemometers at the top-end support ring to capture mechanical turbulence. The active ventilation system was operated at different speeds. An infrared camera provided temperature differentials from ambient for all observatory surfaces. Similar temperatures were collected at the CFHT. A combination of data from CFHT MegaCam delivered IQ; a DIMM installed at the base of the aperture and one installed at the CFHT monitoring tower provided estimates of the apparent dome seeing.

For all measured subsets, the telescope orientations, external wind speed, direction, and temperature were extracted from the CFHT tower data and the Maunakea Weather Center. Four distinct orientations were simulated for a long period of time for Keck II and three for the CFHT, with and without passive ventilation. In the case of Keck II, 240 combinations of orientation, external wind speed, and fan speed were sampled.

The Keck II outputs are as follows:

- Refractive index variance along the scintillometer path, integrated C_n^2 : reasonable agreement with measurements;
- Turbulent energy dissipation rate and inner scale: reasonable agreement with measurements;
- Integral length scale: matched aperture width as expected;
- Matched behavior of C_n^2 vs. external wind speed;
- Matched behavior of C_n^2 vs. fan speed;
- Identified optimal fan speed; half of what Keck was previously using.

The CFHT outputs are as follows:

- Point Spread Function (PSF) of dome seeing along optical path, PSF Full Width Half Maximum (FWHM): reasonable agreement with measurements;
- Significant reduction in dome seeing FWHM with passive ventilation: at least 25% without including surface temperature differential improvement (it proved to be closer to 50%).

The impact is as follows:

- Risk reduction: validation of dome seeing modeling;
- Optimal Keck fan speed selection: operational cost reduction;
- Support of recommendation to ventilate the CFHT dome: differential IQ improvement;
- Overall comment as in the case of site CFD: modeling and measured uncertainties were comparable.

As a first approximation, one could envision the thermal seeing optical volume exhibiting OPD behavior similar to atmospheric with a von Karman spectrum of prescribed outer scale and spectral slope. Simulations have shown that the spectral slope quickly converges to Kolmogorov as resolution increases. The optical path length of current and under design/construction large and extremely large telescopes is several tens of meters, long enough to be considered a “good case” of the GL. Under these conditions, the PSSn could be approximated as follows:

$$\frac{FWHM_{ATM}^p}{FWHM_{TOTAL}^p}$$

In general, this is a quick and dirty way (with some associated error margin) to convert the FWHM to PSSn for most aberrations for purposes of error budgeting. Note that the PSSn is the integral of the square of the PSF, normalized with that of the reference atmosphere,

incorporating information from the PSF wings as well, while the FWHM only refers to the width of the PSF where its strength is half its maximum. Thus, the approximation may not be valid for all types of optical errors. More on this can be found in [11–13]. Normally, $p = 2$ is used (equivalent to a root sum square operation between the error and the reference atmosphere). For the specific thermal seeing term, however, since the thermal seeing optical volume is part of the total seeing column, the value of $5/3$ should be used:

$$\frac{FWHM_{ATM}^{5/3}}{FWHM_{ATM}^{5/3} + FWHM_{TS}^{5/3}}$$

With the Keck and CFHT validation campaign and the fact that the authors were involved in systems engineering for three large and extremely large observatories under design, PSF maps for five observatories have become available. Figure 8 summarizes the thermal seeing results (multiple data sets, wind speeds, and orientations per point). The results have been rescaled to correspond to the same reference site atmosphere with Fried parameter $r_0 = 18\text{cm}$ and a wavelength of $0.5\text{ }\mu\text{m}$. The simulated PSFs confirm that the above expression for the conversion between the FWHM and PSSn is valid. Moreover, the values for Keck and the CFHT have been independently estimated to be realistic, which renders the predicted performance of the observatories under design also realistic, while the Vera Rubin observatory (LSST) is at its final stages of construction with first light expected in 2025. Note that the nominal total error budget for an ELT in the PSSn is ~ 0.85 .

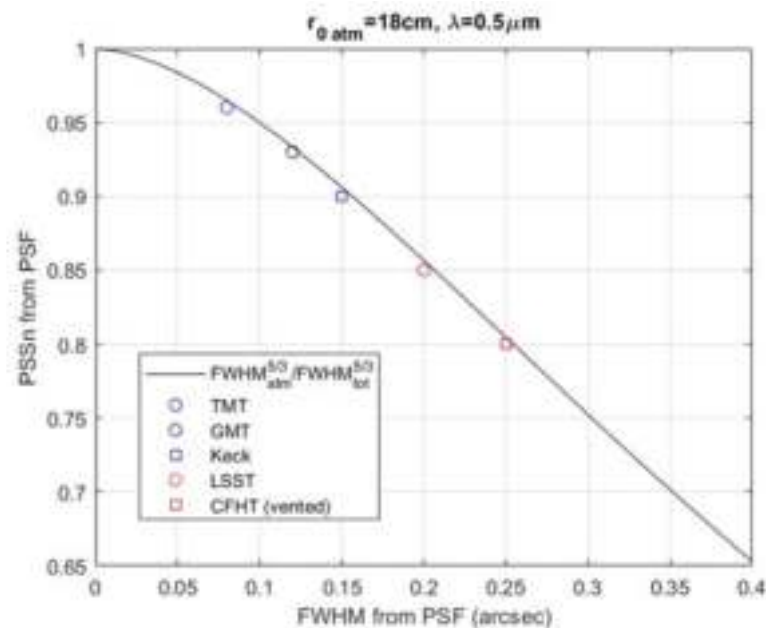


Figure 8. Observatory simulations: PSSn vs. FWHM.

7. Conjugate Heat Transfer and Solid Thermal Modeling

As explained, the purpose of these types of models is not only to estimate component thermal inertia behavior in time but also temperature spatial variability and the resulting deformation. Thus, the typical required inputs include the following:

- Material and thermal properties of components, including emissivity;
- Ambient temperature profile (either user-defined typical diurnal or subsets of the stochastic framework);
- Active heat loads (component-generated, temporal records if possible);
- Convective heat transfer coefficients (from CFD) for surfaces in contact with air volumes not modeled;
- Thermal resistance between conducting components.

The outputs include the following:

- Time (and space if desired) records of bulk and/or surface temperature differential from ambient for each assembly component;
- Time and space records of node displacements from a reference geometry.

7.1. Mirror Assemblies

The first and most characteristic example of this model category was the TIO Segment Support Assembly (SSA) model. It is a transient model that incorporates conductive, convective, and radiative heat transfer between the segment, the assembly frame, the twelve edge sensors, the three actuators, the corresponding electronics boxes, and the environment.

An earlier version of this model did not incorporate a fluid component. The drawback was that since there was no actual flow pattern away from the segment, convective heat transfer was only treated as a parameter. Since then, the model evolved into its intermediate version, presented in [58]. Here, the fundamental difference is that both the solid and the air temperatures are being solved behind the segment, which eliminates the need for heat transfer coefficient information or “feedback loop simulations” between CFD and FEA. Therefore, the plumes from heat sources, such as actuators and electronics boxes, are accurately represented. For the top segment, surface heat transfer coefficients from the observatory CFD model simulations are used. The correct material properties are used for the segment itself, but for some assembly components that are not 100% solid, density is adjusted so that, given the volume they occupy, the resulting mass is correct. Figure 9 depicts the resolved geometry along with end-of-night surface temperature deviation from ambient, for the final version of the model with increased geometric fidelity.

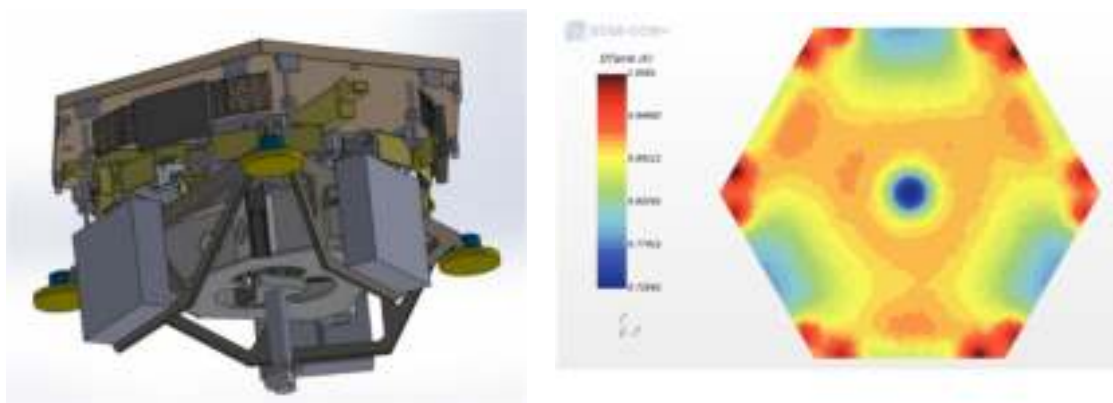


Figure 9. TIO Segment Support Assembly resolved components (**left**) and representative front sheet temperature distribution (**right**).

The GMT M1 mirror segment is made of borosilicate glass with a flat back surface, a parabolic top surface, and mostly hexagonal cores (air pockets) connecting the two sides. The M1 thermal control system (M1TCS) is required to flush the cores with (slightly below) ambient air temperature to track the ambient nighttime air as close as possible. Single-core conjugate heat transfer models were developed for various core heights and nozzle lengths and were simulated for various nozzle flow rates. The goal was to produce a convective heat transfer coefficient (HTC) on the core surfaces that is uniform and comparable to that of the mirror front surface, dictated by the external wind speed, telescope orientation, and venting strategy and obtained from the nighttime observatory CFD simulations.

This way, the deformation of the segment’s front sheet should also be minimal. The air is conditioned by Air Handling Units (AHUs) in the lower plenum of the M1 cell and is guided to the cores by a forest of nozzles. The core flow exits to an upper plenum (UP) volume, separated from the lower plenum (LP) by an insulated steel plate. The upper plenum air can be directed back to the AHUs through intakes on the plate (closed cycle)

or exit radially, in which case enclosure air is used by the AHUs (open cycle). The upper plenum can also be sealed from or allowed to communicate with the enclosure air through a gap at the edge of the segment.

Single-core conjugate heat transfer models were developed for various core heights and nozzle lengths and were simulated for various nozzle flow rates, to generate a look-up table of front-sheet HTC and core flushing efficiency. The HTC on the back sheet, however, depends on the upper plenum flow pattern. So, next, the UP CFD model was developed, resolving all air path inlets and exits, along with the forest of mirror load spreaders (LSs) and core nozzles.

At the same time, it became apparent that, to obtain deformations corresponding to the various design options, a 3D solid model of an entire segment, which takes the expected environment air temperature and surface HTC maps and reports the resulting glass temperature and consequent deformation fields, was required. Note that the on-axis segment varies from the off-axis segments, which means that even though the strategy is identical, all the manual and computational effort is duplicated.

Finally, to confirm that the pressure level and flow pattern created by the AHUs in the lower plenum ensures the desired velocity and temperature uniformity through the nozzles, the LP model was developed, which incorporated all structural components of influence, the AHUs and their ductwork, electrical cabinets and cooling pipework, the actuators and their heat sources, and, of course, the nozzles.

Details of the GMT framework can be found in [18]. Figure 10 shows the M1TCS modeling flow chart with representative model depictions and resulting maps.

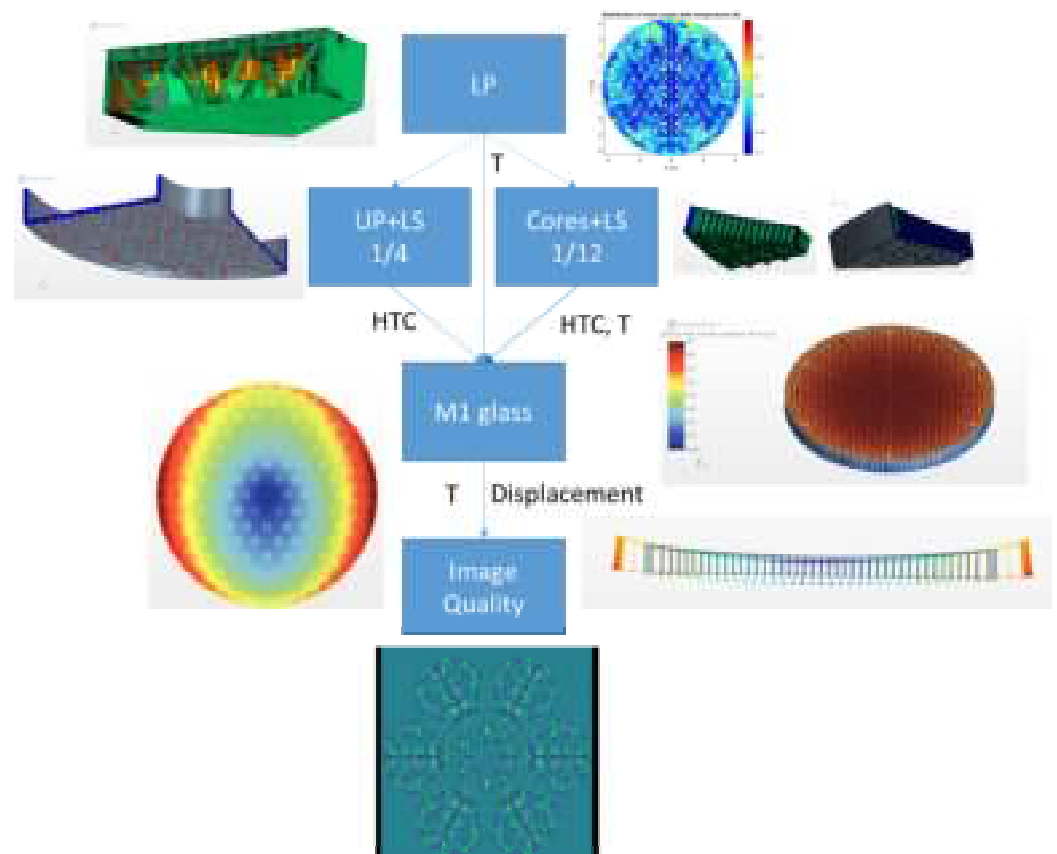


Figure 10. The GMT M1TCS modeling flow chart.

The GMT Adaptive Secondary Mirror System (ASMS) comprises seven deformable 1 m class mirror segments that are designed to function as a single mirror and seven hexapod positioners. Each mirror is deformed by 675 actuators whose dissipated heat is thermally controlled by a cooling plate. The seven crates of the required electronic boards are also

actively cooled, while the forty-two hexapod actuators are not. There are also power switch hubs and batteries occupying the top-end space between the support trusses.

A high-resolution solid thermal model of a segment was built, which included the hexapod, frame, electronics crate, cooling plate, reference body, wind skirt, and mirror [59]. The model was attached as the on-axis segment to a reduced version of the mount model (top-end structure, M1 segments, “sky”) in order to properly account for conduction and radiation effects. The remaining six segments and thirty-six actuators were represented as the coarser resolution shapes to be included in the observatory model, and their densities were adjusted in order for their surface temperature to behave similar to the on-axis model over the course of a night.

The need for additional information, such as the refractive index variability at the segment edge sensor locations and the expected differential pressure on the deformable mirror, resulted in modifying the already fine model in order to properly resolve the 100 nm gap between the mirror and the reference body, the actuators themselves as cylinders, and the gaps between them and the reference body as well.

After inspecting the 60 observatory CFD cases in stock, simulations were performed for several wind speeds at the expected impinging angle between the wind and segment, and the differential pressure maps were obtained. Then, the simulations were performed with the coarser resolution segments included in the observatory model, which only provide front pressure maps. A conversion function was formulated so that differential pressure maps from all 60 cases could be generated, which cover the entire range of wind speeds and orientations of interest.

As for the edge sensor environment, 120 s long records of pressure and temperature at 20 Hz were extracted at the sensor locations from the 60 cases, and refractive index spatiotemporal statistics were calculated. These outputs are valuable inputs to the segment co-phasing model.

Figure 11 shows a representative pressure coefficient distribution under median winds on the surfaces of an ASMS segment (half section view).

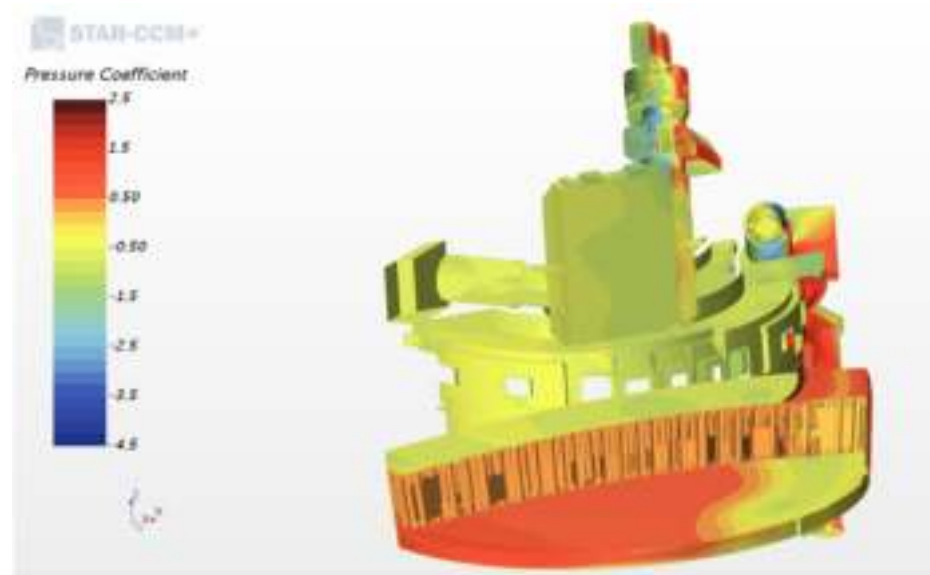


Figure 11. GMT ASMS model and representative pressure coefficient distribution.

The impact is as follows:

- Validation of thermal requirements on optical assemblies;
- Error budget terms estimate: mirror seeing, segment thermal deformation, and dome seeing;
- Provides realistic expected mirror shapes (thermal, pressure-induced deformation) for actuator correction and phasing.

7.2. Enclosure Systems

The TIO enclosure design consists of a dual-shell geometry, which creates an interstitial space between the external cladding and the interior insulation. Most of the enclosure's structural thermal mass and heat sources associated with its operation reside in this volume. The required level of insulation might not be achievable everywhere (e.g., vent doors), and seals and gaps might result in undesired infiltration. The telescope azimuth drives reside above the observing floor. In contrast, the GMT enclosure design employs a single-shell geometry, with all the structural thermal mass residing inside the enclosure volume. Most of the enclosure heat sources are behind insulation, and the telescope azimuth drives are below the observing floor.

Transient models have been developed to assess the thermal environment in these interstitial spaces and investigate potential mitigation strategies. The models include all major components and expected heat sources involved, such as the thermal inertia of steel and concrete structures, thermal loads from electrical panels, and interactions with the exterior environment and ground. Particular emphasis is given to the enclosure azimuth corridors, due to the large drive power output and bogie thermal mass in a confined environment with limited and/or expensive ventilation options. The scope is to maintain the temperature below a specified limit given the insulation specifications and the pressure at a level that will deny leaks to the observing chamber either through the seals or through the enclosure vents, when open.

More specifically, the TIO model [16] consists of the following: exterior enclosure cladding, interstitial air in isolated volumes (cap, azimuth corridor, fixed base, vent space), and insulation. Approximately 100 major electrical panels are resolved as blocks. The important heat sources are resolved as solid with internal heat dissipation and the rest as boundaries with prescribed surface heat flux. Even though in the assembly model the enclosure drive heat dissipation is treated as a volumetric source, the magnitude of that source as convective heat transfer is obtained from individual azimuth and cap bogie models, such as the one depicted in Figure 12.

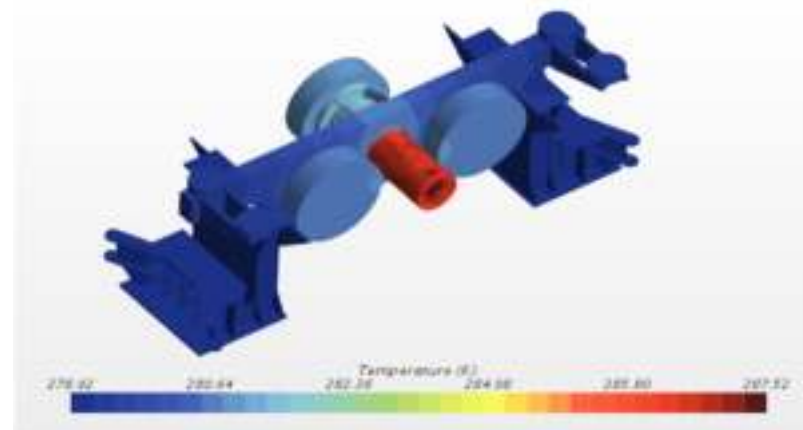


Figure 12. TIO enclosure azimuth bogie with drive motor: typical nighttime surface temperature distribution.

The GMT model consists of two parts, the Pier Ventilation System (PVS) model (outer and inner fixed base rings, pier, cable wrap, telescope azimuth track, telescope azimuth disk, insulation, interstitial air) and the Mechanical Corridor model (insulation, azimuth track, bogies, break masses, four drive assemblies, and electrical panels). In Figure 13, the PVS model and a steady-state representative air temperature distribution are presented.

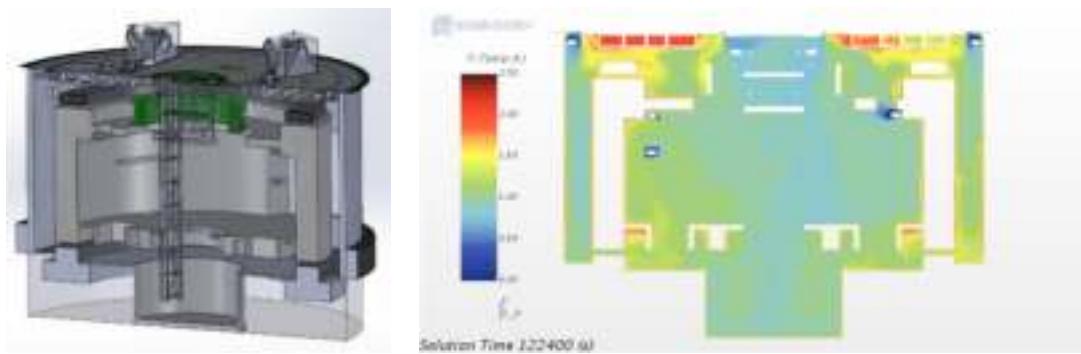


Figure 13. GMT PVS model (left) and steady-state mid-plane air temperature contours (right).

The impact is as follows:

- Validation of enclosure heat dissipation requirements and insulation performance;
- Design choice: enclosure drives' and electrical cabinets' active/passive cooling option and flow rate.

7.3. Telescope Structure

The thermal behavior of an ELT telescope structure (STR) and the STR mounted subsystems depends on the heat load of the system and the thermal properties of component materials and the environment as well as their interactions through convection, conduction, and radiation.

A legacy TIO model from 2010 [60] was based on one-dimensional FEA elements. The resulting displacements were fed into the TIO Merit Function Routine (MFR), which converted them into translations and rotations of the optical surfaces. They, in turn, were multiplied by the TIO optical sensitivity matrix that delivers the corresponding pointing error. Thus, the thermal performance of the structure could be assessed for requirement compliance, and design guidance could be provided. In addition, thermal drift correction strategies and LUTs could be developed.

The latest developed model from 2014 tracks the diurnal temperature variation in the STR and the corresponding deformations [61]. This model is three-dimensional, resolving the thickness of the structural members. It accommodates diurnal simulations and variable elevation angles. It also incorporates all subsystems and components that are thermally coupled to the structure and calculates their temperature as well. Finally, it introduces a “sky” body the size of the aperture, which enables the correct calculation of radiation coupling to the STR. The model consists of the M1 cell, the M3 tower and the lower tube as a single body, the elevation journals, and the upper tube and the spider as a single body. As a later addition, it also incorporates the azimuth structure below the Nasmyth platform, elevation drives and magnets, and hydrostatic pads. Contact interfaces between the distinct bodies are equivalent to welded surfaces. The segment handling and cleaning systems have been omitted, as well as several thin, grated platforms. Additional components, which are only coupled to the structure through radiation, are also included: the M1 control system node boxes, the SSA as a single block of equivalent volume, the M1 as a single body (no segment gaps), the M3 and M2 assemblies, the LGSF, M3, and top-end electronics cabinets, the beam transfer optics duct section above the spider, and the laser launch telescope. The appropriate/desired level of contact thermal resistance between these components and the structure is a parameter. Figure 14 illustrates the end-of-night telescope structure thermal deformation under median expected conditions.

This model will be revisited in the future, and the study will be updated to reflect the latest design.

A modeling framework of similar fidelity and functionality was recently developed for the GMT mount [62]. The telescope model consists of the M1 segment cells, the elevation journals (C-rings) and hydrostatic pads, the Gregorian Instrument Rotator (GIR), the

three secondary support trusses, and the telescope top end with the adaptive secondary mirror. It also incorporates electronics cabinets, actuators, optical baffles, and other equipment that may be thermally coupled with the structure through radiation or conduction. The framework, however, consists of two separate models: one thermal for unsteady temperature calculation and one FEA for steady-state deformation calculation. Three-dimensional snapshot temperature fields from the thermal model, corresponding to various parts of the night, are mapped onto the FEA model to estimate the resulting deformation relative to an initial zero-strain condition, nominally at the time of telescope alignment around sunset. Deformation of the entire structure is available, but the displacements at the hardpoints, the interface between the structure and the optics, are further processed by the GMT optical model to provide the resulting optical misalignment. The goal is to eventually correlate the structure temperature sensor information through a spatial gradient to the optical misalignment and develop an LUT of corrective adjustments as a function of external environmental conditions. Figure 15 depicts the thermal and FEA models.

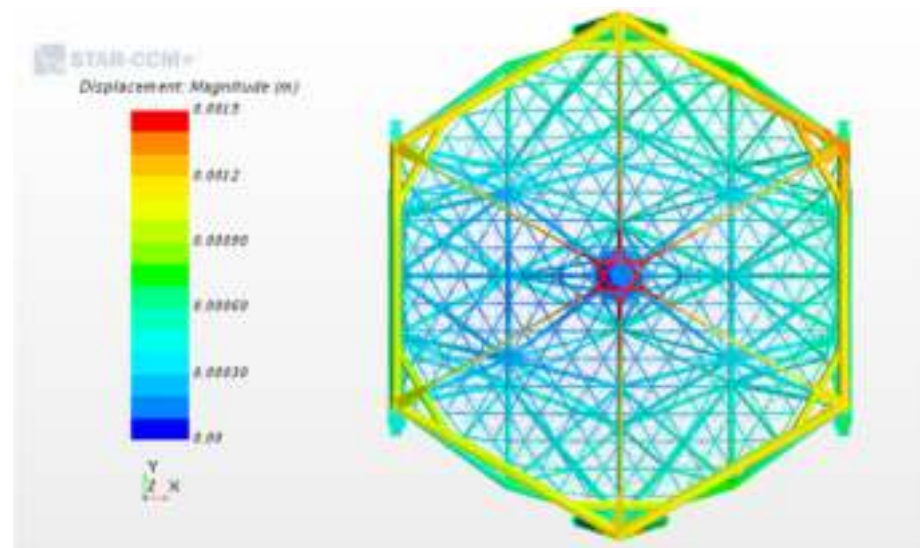


Figure 14. TIO telescope structure thermal deformation magnitude (m) 12h after sunset.

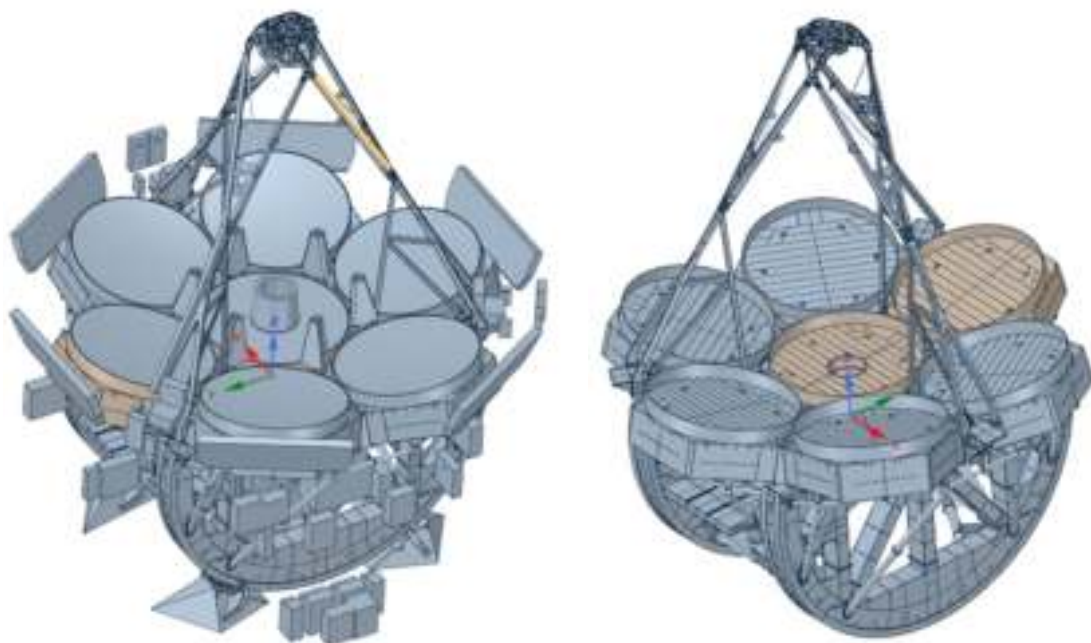


Figure 15. GMT telescope structure thermal (right) and deformation (left) models.

The impact is as follows:

- Requirement verification: minimum segment gap distance, residual on-sky tracking error, telescope pointing error, maximum actuator stroke, and maximum back focal distance shift;
- Provides realistic expected thermal misalignments for correction strategy testing;
- Provides input boundary conditions to observatory performance CFD simulations.

7.4. Subsystems/Components

7.4.1. Elevation Drives

The modeling of elevation drives started in 2014 with a standalone model to estimate the required cooling flow rate so that the drive surfaces do not exceed a temperature detrimental to dome seeing and the drive coils do not exceed the maximum allowable temperature.

It consists of four major components: the support structure, coils, forcers, and magnets. It is a transient model that incorporates conductive, convective, and radiative heat transfer between the components and the environment. The model was later incorporated into the TIO Telescope Structure Model, as shown in Figure 16.

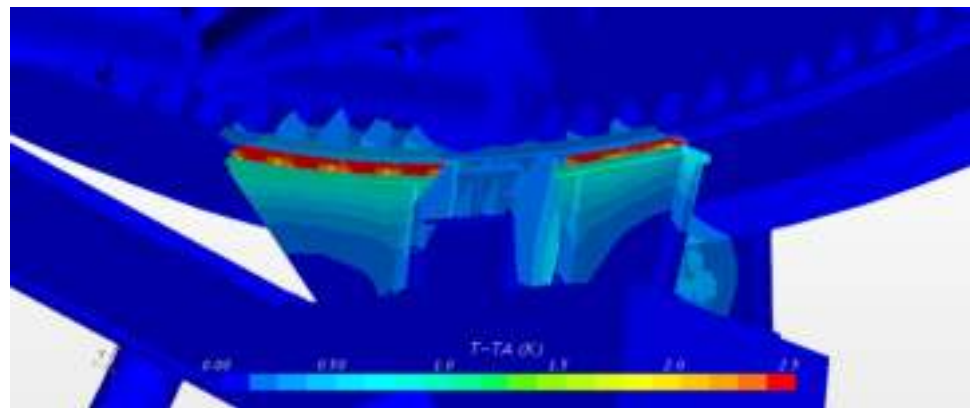


Figure 16. TIO elevation drive model detail and surface temperature deviation from ambient.

The impact is as follows:

- Validation of EL drives' cooling flow rate and max coil temperature requirements;
- Provides input boundary conditions to observatory performance CFD simulations.

7.4.2. TIO Laser Guide Star Facility (LGSF)

A model of the LGSF beam transfer optics duct (BTD) was developed in 2013 [63]. It resolved the duct thickness and laser beam transfer mirrors and their support framework for most of the laser beam path that is subject to significant temperature gradients and/or large vertical change. It also resolved the air inside the duct and its thermal interaction with the above components through conjugate heat transfer. The thermal interaction of the laser beam with the optics was also captured. It did not include the path from the eight lasers to the truss pointing array (TPA) or the path inside the laser launch telescope (LLT). The model provided guidance to the LGSF design team and a first estimate of the laser beam stability performance and requirement compliance.

As the telescope structure design evolved through the years, a new optical path was proposed for the LGSF. Both the original and the new optical paths were compared against optical, mechanical, and other telescope performance-related criteria. The optical performance criteria included a first-order analysis of the optical turbulence generated within the ducts. The new proposed path exhibited similar thermal behavior but significantly reduced optical turbulence and was adopted. Design modifications as a result of the study included increasing distances between certain optics and aerodynamic reshaping of the envelopes around the transfer optics. The initial beam jitter estimate was well within the budgeted values.

Between 2021 and 2023, the model was updated to include the laser bench (LB) array, the top-end (TE) assembly, and the LLT [64]. Higher geometric fidelity was also achieved with additional equipment and a support framework. The model was split into five sub-models (LB array, TPA, BTM, TE, LLT), each one feeding the next in terms of flow rate and air temperature. The optical path was split into multiple segments to identify the major contributors to beam jitter and focus error from optics, especially lens, thermal deformation. Beam jitter was calculated as a function of time over the course of the night at both low and high temporal frequency. Additional standalone models of the laser head (LH), standard laser electronics (LE), and top-end electronics (TE.EL) cabinets were developed to verify the expected cooling efficiency and exterior surface temperature. Figure 17 illustrates the modeling flow of information with representative model depictions and resulting optical turbulence contours.

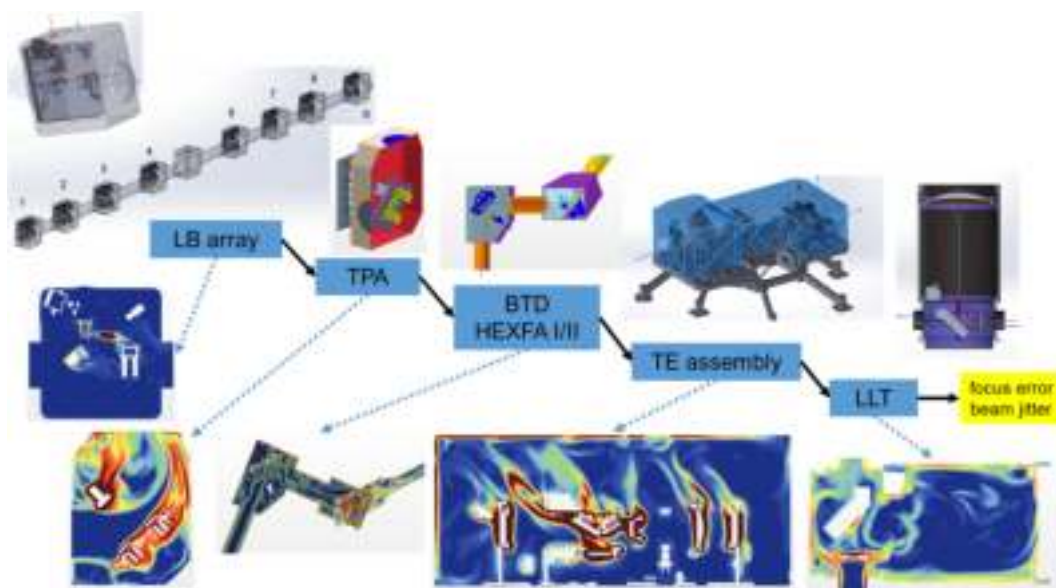


Figure 17. TIO LGSF modeling flow chart.

The impact is as follows:

- Design modifications to minimize path turbulence and enhance lens cooling;
- Estimates of beam jitter and focus error budget terms.

7.4.3. Electronic/Electrical Cabinets

These are transient models that focus on the shape and volume of a particular component, which typically encompasses an internal heat source. The volume may include multiple parts of different materials, as well as air. Therefore, only the total mass and an average heat capacitance are required to capture their thermal inertia. Their main goal is to provide the distribution of resulting surface temperature and the time it takes the internally dissipated heat to reach the surface given a variable heat source and external environment.

Initially, they were developed as standalone models for a few given components, but recently, they have been routinely included in all assembly models; it is the typical way to treat the average heat source component.

8. Support CFD Modeling

Support CFD models are models of limited scope developed to answer specific design and/or modeling questions. They can be steady or unsteady, and they typically correspond to isolated components.

8.1. TIO Enclosure Aperture Deflectors

The TIO aperture deflectors are a unique feature developed to shield the top end of the telescope from wind as a result of reducing the enclosure diameter to the minimum necessary, but they also act as a seal between the enclosure and the aperture shutter. Figure 18 illustrates the effect of the deflectors on the velocity levels around the top end of the telescope through this early steady-state simulation during the conceptual design phase in 2006. This was equivalent to a ~ 2.5 m increase in the enclosure diameter. The deflectors have a radius of ~ 5 m, a chord of ~ 5 m (tip to hinge), and a projected length normal to the aperture plane of ~ 3.5 m. When deployed, the angle between the chord and the aperture plane is 45° . During the detailed design phase, the lowest structural mode was estimated approximately at 5 Hz. For both fatigue and survival considerations, the natural vortex shedding frequency and its proximity to the lowest structural modes were evaluated in 2011.



Figure 18. Impact of the TIO deflector design on velocity levels around the telescope top end (velocity scale normalized with upwind value).

A detailed CFD model of an isolated, single deflector at a 30° angle of attack (the angle expected to determine the maximum shedding frequency) was developed, and unsteady simulations were performed for both median and survival wind conditions [58]. The Strouhal number was calculated, along with the mean and dynamic force components. Under survival conditions, the shedding frequency approaches the first structural mode. The results could be used as inputs to the deflector FEA model to determine the effects of wind loading and investigate design modifications if necessary.

8.2. Spider Beams

The term “spider” refers to the thin multimember truss structure that supports the secondary mirror assembly of a telescope and is present in many existing observatories.

In 2012, an isolated, single spider beam model, at the time $0.25 \text{ m} \times 0.6 \text{ m} \times 20 \text{ m}$, was developed to investigate several modeling assumptions.

The original parametric model used to estimate wind jitter based on steady CFD inputs had assumptions on the beam drag coefficient, the slope of the force spectrum, and the force decorrelation length scale along the beam. The simulations would provide accurate estimates for these assumed values as well as the spatial and temporal resolution required to properly resolve the physics.

Unsteady simulations were performed using various surface resolutions, wind speeds, upwind turbulent intensity, and integral length scale and sampling rate, for flows perpendicular to both the wide and narrow beam sides, as well as angled flow at 30° (combination). The results collected included pressure records on selected locations and integrated force records, as well as mean and RMS drag and lift coefficients.

The results indicated that the parametric model was not valid for local wind speeds above 1 m/s. It also confirmed the assumption that the integrated force spectrum around a solid object has a slope steeper than that of point pressure due to the aerodynamic attenuation effect. The appropriate surface resolution that captures vortex shedding at the expected frequencies was determined and was adopted for the entire TIO telescope

structure above the hex ring, the M2 assembly, and the laser launch telescope (LLT). Lower tube resolution was also scaled accordingly. Unsteady simulations would directly provide force records for image jitter estimates going forward. The GMT upper telescope structure is similarly modeled.

8.3. Flow-Induced Pipe Vibration

The phenomenon of flow-induced vibration is well known and potentially problematic for ELTs, especially in AO mode, as the cooling system comprises numerous pipes and bends.

In support of vibration error budgeting, several simple pipe and valve flow models were developed between 2013 and 2021 in a joint TIO-GMT effort to investigate the magnitude of the dynamic force the turbulent flow imposes on the pipe due to friction or separation, as well as its spectrum. A turbulent fully developed straight pipe flow case was first considered and provided the inlet conditions for the remainder of the simulations. These included a 90° bend and an S-shaped pipe, two consecutive 90° bends. Different pipe diameters, ball valve angles, bend radii, and flow rates were also considered.

The results indicated that a single valve could cause larger vibrations than many feet of piping. The piping network should be designed to ensure that valve angles are kept within a threshold limit ($<30^\circ$). Even though separation frequencies of turbulent flow past a 90° bend can be as high as 200 Hz, the absolute magnitude of the resulting dynamic force is much smaller than the expected pump pressure fluctuations.

8.4. Summit Facilities' Exhaust

Observatories employ several refrigeration systems used for enclosure daytime air conditioning and the cooling of instrumentation, optics, telescope hydrostatic bearings, etc. The management of the refrigeration cycle heat rejection can be important for IQ. Often, heat rejection occurs at a facility building adjacent to the telescope enclosure. The location and orientation of exhaust vents is selected considering the prevailing wind direction. This approach is based on statistics and might not be very effective for bimodal or multimodal wind environments. A more expensive solution involves building a tunnel to direct the warm air away from the observatory and the optical path. Both the TIO and GMT sites are bimodal, with roughly NE/SW wind orientations.

The chiller of the TIO summit facilities building is therefore located at the NW corner but is expected to discharge warm air in close proximity to the observatory. Under low wind speeds and for certain wind orientations, it is expected that this plume might rise high enough and interfere with the optical path before it has time to diffuse into the air. However, it is expected that the plume can only be a problem under westerly winds and for low elevation angles. Under east winds, the plume dissipates before it crosses the TIO optical path, even under low wind speeds (although the complete absence of wind has never been simulated).

In 2010, a study was performed to assess the plume's impact on the reference atmospheric seeing. Three CFD steady cases for low (2.5 m/s), median (6.5 m/s), and high wind speeds (10 m/s) from the west (270°) were processed. It is assumed that the results are also valid for west winds between 225° (more common) and 315° (less common). The telescope zenith angle was fixed at 33° . The nominal flow rate of 150,000 CFM resulted in an exhaust velocity of 2.5 m/s and a temperature of ~ 10.5 K above ambient. The values correspond to an air-cooled unit.

Various optical paths were processed to estimate the equivalent r_0 of the first 50 m past the enclosure. Three azimuth (315° , 0° , 45°) and three zenith (65° , 55° , 45°) telescope angles were considered. For each specific telescope orientation, three paths were considered, at $X = 0$ m and $Y = -15$ m, $Y = 0$, and $Y = +15$ m in the telescope coordinate system. The initial results indicated that considerable impact is detected only for the 65° zenith angle and only for the $Y = -15$ m path.

As the design evolved, the total power required to be exhausted increased dramatically, and a trade-off between temperature and flow rate was necessary. A new study was per-

formed in 2020 [16] considering various combinations of exhaust temperature differentials and flow rates. Two major changes were implemented for this study. The simulations were unsteady, since advances in GL modeling, as explained in previous sections, allowed for more accurate estimates of the expected FWHM along a path. The exhaust is expected to have louvers that can be optimally arranged to minimize the plume's effect. Therefore, a 45° down angle was selected. Based on previous experience, the study focused on west 2 m/s winds and the processing of the 315°–45° azimuth sector.

The new study verified that only the lowest part of the optical path is significantly affected at a zenith angle of 65°. It also verified that the plume is not expected to enter the observatory through the vents. An acceptable temperature differential of 15 K above ambient was selected. However, for that temperature, seeing degrades as flow rate increases, which suggests possible exhaust angle and other geometric adjustments. On average, under west low winds and for low elevations, the worst seeing will occur when pointing NE.

Overall, it appears that the plume crosses in some way the telescope optical path when (a) the wind direction is between 225° and 315°, (b) the wind speed is less than a nominal 4.5 m/s, and (c) the telescope azimuth angle is between 300° and 60° and the telescope zenith angle is higher than a nominal 50°. From the stochastic framework, we conclude that this occurs in ~0.4% of the observing time.

The GMT dry coolers are located E-SE of the enclosure, further away and downhill. Their exhaust flow has also been incorporated in the observatory CFD model. Even though they reject heat vertically and at a higher velocity (5 m/s), it is even less likely to affect the optical path. None of the 60 baseline simulations used in the performance assessment suggest any optical path–plume interaction.

The impact is as follows:

- Requirements validation: chiller exhaust temperature differential from ambient, flow rate, orientation, and potential geometric modifications near the facilities' exhaust area.

8.5. TIO Mechanical Room and Utility Tunnel

In 2015, an extension of the observatory CFD model was developed that included the cable wrap opening on the observing floor and the utility tunnel, as well as the utility and mechanical rooms.

The scope of this study was to investigate the direction of the flow at the cable wrap opening and inside the utility tunnel, the expected temperatures inside the tunnel and utility and mechanical rooms, and the adequacy of the design flow rate inside the tunnel for hazardous gas removal.

Simulations were performed for a configuration that would produce pressure levels inside the enclosure low enough to cause potential flow reversal in the tunnel or locally at the cable wrap opening (median to high wind speed, closed vent, zenith pointing).

Under normal observing conditions, there is only thermal diffusion through the cable wrap that does not really affect the optical path. Under high wind speeds, however, there is flow from the cable wrap opening to the enclosure interior. The tunnel gas removal pipe should not discharge in the utility room but should be connected to the utility room exhaust duct. In order to prevent thermal leakage through the rooms and into the tunnel, the mechanical room exhaust flow rate should be 15% higher than the utility room intake flow rate.

In 2021, the model was revisited and significantly extended to reflect design changes and incorporate higher fidelity [16]. Additions included a bulkhead door that divides the tunnel volume during a fire emergency, the pintle and cable wrap obstructions, and pier and foundation concrete. Additional goals were to assess the impact of closing the bulkhead door on the tunnel flow pattern and the impact of directing the exhausted tunnel air into the mechanical room (as opposed to diverting it directly into the fluid cooling system; we had already determined not to discharge in the utility room).

The study verified that closing the bulkhead door will not have an impact (tunnel flow does not reach the utility room). It also showed that directing the tunnel air into the mechanical room can reduce the exhaust temperature by 2–4 K. Finally, it verified the

expected nighttime exhaust temperature level, even though in the first two hours after sunset this level will be significantly higher due to thermal inertia and volume flushing turnaround time (daytime load is much larger).

The impact is as follows:

- Requirements validation: flow sign at cable wrap opening, tunnel fan and utility exhaust flow rates, tunnel gas discharge location, mechanical room exhaust temperature, and temperature environment around equipment.

8.6. Segment Ion Beam Figuring

Several segmented mirror telescopes (e.g., E-ELT, TIO) use Ion Beam Figuring (IBF) to bring the M1 segments to their final figure specifications. During this process, which can last many hours, the segment temperature rises considerably. The bonded joints in the back of the segment can weaken if their temperature exceeds a certain value.

A transient thermal model of the IBF process was developed for the TIO in 2020 with the goal of being able to cheaply test various scanning patterns in order to optimize the process and minimize the temperature rise [65]. Temperature measurements during the IBF experiments for the TIO were used to validate the model. The measurements consisted of a calibration test on a small sample disk (simulated) and three segment runs (two of which were simulated) corresponding to different scan patterns and speeds.

One of the key aspects of the model is the implementation of the Ion Beam flux arriving at the segment surface. This is accomplished by a Gaussian distribution function either on the surface or in a volume of specified thickness. Other important physical phenomena to be considered are diffuse irradiation from the Ion Beam gun as an infrared source, transmissivity through the glass, and potential radiation reflection on surrounding support surfaces. Despite several uncertainties, an initial set of inputs has been identified that could satisfactorily predict the maximum expected temperature at the back of the segment.

8.7. Hexagonal Segment Profilometer

Several existing and future observatories employ segmented primary mirrors consisting of hexagonal segments. In order to verify the final figure specifications of these segments, optical metrology measurements must be conducted. For a particular type of measurement, a laser profilometer, a specific chamber had been proposed by a TIO vendor. In 2012, a model was developed to investigate the relative magnitude of thermal effects such as segment deformation due to non-uniform temperature distribution in the test chamber and variable refractive index along the laser path.

It consists of the following components: laser probe, outer test chamber, glass barriers, segment, support structure, and air inside the test chamber. It is a transient model that incorporates conductive, convective, and radiative heat transfer between the components and the environment. Two configurations of active test chamber ventilation were considered (parallel and counter-flow), results were obtained, and possible error mitigation strategies were suggested.

8.8. Factory Acceptance Tests

More often than not the environment inside a factory or laboratory, where final acceptance tests will be carried out for an observatory component, deviates from that at the actual observatory site. Sometimes a separate set of requirements must be established: if performance during a factory test is X then performance on the mountain should be Y. In such cases a model is used to connect the two. Let us take as an example the “air knife” assembly of the Daniel K. Inouye Solar Telescope (DKIST). It is a key thermal systems component that serves the dual purpose of ventilating the ceiling of the Coude laboratory and as the interface between laboratory and ambient environments using an air curtain. During factory acceptance testing, flow visualization of the air curtain revealed that flow was deflected upwards, suggesting an unexpected pressure gradient, which was confirmed by pressure sensors between the bottom of the assembly and the factory floor. It was

hypothesized that large-scale eddies, generated by air knife outflow reversing due to floor proximity, may have induced the increased pressure at the measured locations. The actual CFD model of the assembly in observatory installation configuration suggested that the air knife would meet or exceed all performance requirements [66].

9. CFD Vision 2030

It is evident from the previous sections that the ability to simulate aerothermal flows using CFD has progressed rapidly during the last two decades and has fundamentally changed the aerospace design process. Advanced simulation capabilities not only enable reductions in ground-based (and flight) testing requirements but also provide added physical insight and enable superior designs at a reduced cost and risk. Despite considerable success, the reliable use of CFD has remained confined to a small region of the operating envelope due, partly, to the inability of current methods to accurately and/or efficiently predict turbulent, separated flows. As a relevant example, let us mention the effort that began in 2015 to predict the high-resolution asymptotic value of the PSSn in the context of thermal seeing [43]. Even though an uncertainty quantification study was finally performed with satisfactory results [28] and recent E-ELT efforts for localized effects, without resolving the full observatory, also yield realistic results [67], the status can still be summarized in the following lines: *“The estimate of the local thermal seeing as expressed in the PSSn metric does not appear to converge with spatial resolution at the limit of feasible, or rather, reasonable computing power. The extent to which this phenomenon is numerical and to which fine-scale contributions can be parameterized and accounted for is still under investigation. Methodologies have been developed to predict the asymptotic value past the simulation’s target resolution cutoff. However, the error margin within which this might be achieved is probably larger than the desired precision for conducting individual moderate heat source sensitivity analysis. Moreover, the time variability of the metric is also larger than the desired precision”*.

Computational scientists are summoned to answer whether CFD has reached its limits or should further research investments make CFD accurate, efficient, and reliable enough to enable future revolutionary gains in aerospace system (in our case observatory) performance and ensure verification by analysis. Under NASA sponsorship, Boeing, Pratt & Whitney, Stanford University, Massachusetts Institute of Technology, University of Wyoming, and the National Center for Supercomputing Applications recently conducted the CFD Vision 2030 Study [68].

The study constitutes a roadmap for the development of enabling technologies such as High-Performance Computing (HPC), physical modeling, algorithms and numerical methods, geometry and mesh generation, knowledge extraction, Multidisciplinary Analysis and Optimization (MDAO) with quantified uncertainties, and experimental testing for computational validation campaigns. A CFD Vision 2030 Integration Committee was established within the American Institute of Aeronautics and Astronautics (AIAA) [69], but its scope of promoting revolutionary advances in computational simulation technologies for the analysis, design, certification, and qualification of aerospace systems is very relevant to ELTs as well. The roadmap consists of various milestones to address and overcome certain challenges with the ultimate goal of simulating a complete operating system within a time window long enough to assess mission success. The simulation realization and post-processing, along with estimating performance parameters and their uncertainties, should be one of many, and the turnaround timescale should be weeks, if not days. In the aerospace community, this means, for example, a maneuvering aircraft with a fully resolved and operating engine(s), as well as critical subsystems such as fuel supply. A database of 1000 simulations will create a real-time multi-fidelity feedback tool during system operation. In the next subsections, we will review how this may translate to the case of a ground-based observatory.

9.1. Geometry and Mesh Generation

The first step in a modeling process is establishing the desired geometric fidelity of the simulation. This of course depends on the physical phenomena involved and the simulation scope (outputs). In the case of trade studies between multiple design options, the geometric differences must be well resolved. Unfortunately, the geometric definition of a component, the digital representation of it (CAD model), and the actual manufactured product are not all the same due to tolerance and translation errors between software formats. This may sometimes have implications for performance predictions. The main current challenge, however, is the incompatibility of CAD models with simulation requirements. There are several reasons for this, the first being the aforementioned tolerance errors. Subsystem components put together to create a complete system model do not always combine to create a closed surface that can be meshed either for air (externally) or solid (internally). Aerothermal simulation requirements might also exclude small features that are not expected to affect the aerothermal environment, but their removal is cumbersome, while leaving gaps and holes that again disrupt surface continuity. Modern software developers, both CAD and simulation, are actively working on implementing automated repair techniques, but dividing the effort between them is an active source of contention. . . On the simulation side, few software suites are able to “shroud” the geometry, within some tolerance, and eliminate unwanted gaps and features. This technique, however, can significantly alter other important details that affect the flow, and manual user intervention is tedious and time consuming. As a result, both the TIO and GMT currently maintain separate simulation CAD models for subsystems. Generating and updating these models to match design progress has been a labor-intensive task. To accommodate an increasing number of models and configurations, best practices from both systems engineering teams and vendors should be implemented. CAD modelers should understand simulation requirements and create assemblies that can be converted to simulation models with minimal effort, maximizing model interoperability. Simulation modelers should continue to follow repair methodology advances, especially on the surface meshing field.

With a geometry representation in place, the next step is the generation of the computational mesh. There are two main mesh categories, surface/normal-to-surface and volume. The former dictates the accuracy of calculated forces on structures and heat dissipation. The latter is responsible for properly resolving vortices in shear layers and wakes. The aerospace community is actively working towards automatic mesh adaptation in high-velocity-gradient regions, either through re-alignment, morphing, or refinement. Complex geometries or moving parts are resolved with multiple, overlapping meshes (overset) that exchange information through interpolation. Minimizing the interpolation error is another active field of research. At this point, let us define the goal for our operating observatory simulation: the telescope is tracking, then slews to another target at a different orientation and continues tracking, while operating the vents and windscreen (where applicable) under variable wind conditions.

In 2014 [70], an attempt was made to develop an observatory CFD model that would accommodate all configurations, orientations, and modes of operation with a single mesh, based on sliding regions. It had the ability to slew, open, and close the vents and convert to daytime configuration for air conditioning simulations. This, however, was achieved with certain geometric and physical compromises. Vent operation relied on porosity, while parts of the telescope azimuth and elevation structure were removed to accommodate their relative motion. Moreover, it has been proven [71] that uniform hexahedral meshes aligned with the optical volume minimize errors in estimating thermal seeing. Therefore, overset is the only way to go for a multi-target, variable-configuration, transient simulation. Vent operation can be achieved with morphing, since motion is one-dimensional. Mesh adaptation around the enclosure aperture and to the wake might also be needed, especially when the telescope points away from the wind at low elevation angles.

9.2. Physical Modeling, Algorithms, and Numerical Methods

The focus of the decade is the accurate representation of unsteady, turbulent separated flows, a common condition inside an enclosure as well. It has been shown [45] that Reynolds-Averaged Navier–Stokes (RANS) models are not able to resolve optical turbulence and that Detached Eddy Simulations (DES) or Large Eddy Simulations (LES) are needed. Currently, great effort is made toward Wall-Modeled and Wall-Resolved LES. Vortices, such as the ones we observe in Figure 5, are best resolved in conjunction with high-order spatial and temporal discretization schemes. Second-order schemes are being replaced by third-, fourth-, or fifth-order ones. Computational scientists should monitor developments in the implementation of these models in the available software and employ them. ELTs currently employ third-order models.

Efforts are continuing toward the prediction of the high-resolution asymptotic value of the PSSn for thermal seeing. In the case of the TIO, with a nominal OPD resolution of ~ 16 mm, resolving the optical volume alone requires 10 billion elements. As a result, novel approaches are now considered to characterize local seeing PSF behavior from coarser OPD map spectral analysis [31,47,69]. To this end, the help of validation studies and artificial intelligence (see following sections) will be employed.

9.3. HPC

Significant financial resources are invested annually by industry and government organizations in the purchase of petascale HPC systems and the development of exascale HPC systems, capable of performing 10^{15} and 10^{18} Floating Point Operations per Second (FLOPS), respectively. These efforts are also transitioning to quantum and bio-computing. Similarly, significant funds and efforts are spent to allow CFD codes to efficiently utilize the new architectures. There are several reasons that dictate the use of HPC systems for observatory CFD simulations:

- Large volume of simulations. The LUT resolution required to characterize the aerothermal performance of the observatory (zenith and azimuth angles, wind speeds and directions, heat release scenarios) can result in hundreds of simulations and dozens of meshes. These performance LUTs need to be updated every time a major design change occurs and subsets of them on a regular basis to accommodate trade studies; thus, the required turnaround timescale is in months.
- Large computational meshes. Accurate wind jitter and thermal seeing require high enough resolution, resulting in a minimum of several tens of millions of elements. Also, to accurately estimate the net heat transfer on enclosure and telescope surfaces, conjugate heat transfer analysis should be enabled, with the enclosure structure, telescope structure, and other important components to be meshed as well in the same simulation. This will double the required mesh size. Finally, proper GL characterization and interaction with the enclosure requires relatively high-resolution mesh to extend in the first 100 m above the topography.
- Large number of timesteps. Thermal seeing PSSn calculation requires simulations in excess of 10 min of virtual time at a sampling rate of 20 Hz or more, especially if the response of an adaptive secondary mirror is to be properly assessed. Moreover, conjugate heat transfer simulations require much longer integration times, dramatically increasing the number of timesteps, where a high-fidelity solution of the air flow is also desired.

ELT integrated modeling teams are working to automate the mesh generation and simulation setup through scripting, a practice also employed by major aerospace manufacturers. Production CFD runs are deployed on an Amazon Web Services (AWSs) cloud system, a cluster of nodes. The configuration of the AWS cluster is driven by the amount of memory and processors of each node, the processor performance, and the data bandwidth required for inter-node communications. After performing scalability analysis, we can settle on the number of nodes of the specific type of instance that provides the required memory, cores, and network bandwidth. The automated process will configure and create the cluster, upload the simulation file, and launch the CFD run. A sweep of the CFD parameter space that generally requires

performing approximately 100 CFD runs could easily lead to the need for a 500-node HPC cluster, if we wished to launch all the runs at once.

Let us examine the example of a GMT AWS CFD simulation [59]. Following the launch of a simulation, a local process is started that synchronizes the data produced locally with a remote S3 repository. Some files, once uploaded on S3, trigger Lambda functions that transform the data in the files into new data products that are compatible with the IM input files. Each CFD simulation generates (currently) ~1 TB of data requiring the local hard drive to be configured with the highest read/write IO throughput available. At the end of each simulation, a post-processing script is executed that extracts and consumes data that are only available once a simulation is completed.

The steps involved in a CFD parameter space sweep are as follows:

- Configure and create the HPC cluster with ParallelCluster (the nodes are requested as Spot instances);
- Upload the CFD configuration files to the cluster;
- Launch the CFD simulations on the cluster;
- For each simulation, start the synchronization process;
- For each completed simulation, post-process the output data;
- For each completed and post-processed simulation, terminate the synchronization process;
- For each terminated simulation, delete the local data.

All these steps used to be performed manually by an operator that is responsible for monitoring and managing the entire simulation cycle. The GMT IM is in the process of replacing them with a streamlined, event-driven operation flow that relies on AWS services to monitor and manage the simulation cycle. As an example, a CFD configuration file is uploaded to an S3 repository that, in turn, will trigger an event that transfers the file to the HPC cluster where the CFD case runs, and associated processes are sequenced and launched automatically, providing significant savings in manual labor.

9.4. Knowledge Extraction

Each HPC simulation generates files of 3D temperature fields for thermal seeing, 2D pressure maps on the optical surfaces for wind blur, and integrated forces on various components at each timestep. In addition, various “screenshots” of variable contours are saved for animation purposes. With the volume of simulations, the mesh size, and the number of timesteps increasing, data manipulation and post-processing becomes a major challenge both for purposes of communications and simulation slow-down. This process also requires automation. Once the simulation is on the way, a synchronization process with a remote AWS repository with enough disk space to receive the output data files of the simulation is required. Additional scripts post-process the simulation data, the data are saved locally, the synchronization process is terminated, and the remote data are deleted.

9.5. Validation

The CFD Vision 2030 called for efforts to develop and execute integrated experimental testing to support a computational validation campaign that will benchmark the new models and methodologies. The astronomical community has the test beds already set: existing observatories. Given the limited body of data regarding dome seeing for ground-based observatories [56,57,72,73], its significant contribution to image degradation, and the modeling limitations already mentioned, it is important to continue validating the CFD methodology and results using data from operating telescopes. There exist current efforts to build dome seeing monitoring instruments [73,74].

Similarly, the estimation of the wind jitter of the telescope structure is based on CFD simulations that have not been fully validated experimentally. One of the largest uncertainties lies in the connection between surface pressure fluctuations and resulting telescope mount torques. Wind tunnel tests with physical models could be conducted to validate the analysis of the forces and torques imparted to the mount by wind buffeting. More specifically, validation is required for dynamic wind loading modeling results for

both M1 piston and mount top-end forces, both in terms of predicting differential pressure fluctuations and the relationship between velocity fluctuations, unsteady forces, and the resulting mount torque.

9.6. MDAO

With a volume of CFD simulation results and observatory and wind tunnel measurements, a multi-fidelity database can be created. This extends the current stochastic frameworks to be able to perform uncertainty quantification for the various properties and data sources. It will also enable the observatory to utilize this database as a real-time performance estimator and operation optimizer. This will be particularly useful when quick engineering assessments will need to be made to resolve problems during observatory commissioning.

Several of the error mitigation strategies already defined, such as daytime air conditioning and nighttime ventilation, rely on environmental inputs. Current modeling depends on certain reference upwind conditions, which are moreover coarsely discretized in parameter space. However, the information used by observatory operators will be based on available data from engineering sensors and a site or environmental conditions monitoring system, describing temperature, wind speed, wind direction, atmospheric seeing, etc. These measurements are expected to deviate from the modeling reference conditions, and they could be influenced by summit facilities. As such, an operational plan should be formulated to re-reference the LUTs to these readings. Sensitivity analysis techniques and Artificial Neural Network (ANN) methodologies, using the simulation results at the monitoring locations as training data, can be employed to predict changes in system response to varying parameter inputs in near real time (order of minutes, even seconds). The applicability of this real-time tool will include thermal seeing and wind loading mitigation through vent operation and knowledge of dominant turbulent length scales, telescope pointing error, and thermal misalignment correction strategies based on temperature sensors. Thus, combined with the deterministic optical errors and the current measured atmospheric seeing, a snapshot of the system performance in the delivered IQ metric of choice will be available at any given time. Finally, performance forecasting might even be possible through the coupling of the results to computational boundary conditions which are outputs of an atmospheric mesoscale model or any other weather forecasting service. To the extent that the levels and even temporal behavior of atmospheric seeing, wind speed, and temperature can be predicted for the upcoming night a few hours in advance, delivered IQ scenarios can be generated that can enable queue observing, that is, selecting the order of targets, operating mode (seeing limited or adaptive optics), and instrument throughout the night, so as to optimize the IQ and science output.

The status of observatory aerothermal modeling, the existing challenges, and possible mitigations are summarized in Table 2. As expected, due to the non-profit and purely scientific nature of the observatories, the main challenge is funding. A future concern can also be climate/weather change, since ELTs are expected to operate until the end of the century. Note that smaller support studies are not an issue and are routinely performed.

Table 2. Observatory modeling status.

Category	Status	Challenges	Mitigations
Stochastic and thermal network models using multiyear environmental records	Developed	Climate change.	On-site environmental monitoring systems and periodic record updates.
Subsystem conjugate heat transfer models	Partially developed; most instruments are still under development, and CADs are not mature	CAD preparation/update.	CAD configuration control; mesh by parts to minimize update time during design development.

Table 2. Cont.

Category	Status	Challenges	Mitigations
Observatory CFD models	Developed, but overset mesh motion and mesh adaptivity are not yet implemented	Insufficient funding for computational resources and/or time to completion. Concurrent simulation with many subsystem models is computationally expensive. Simulation at desired OPD resolution is computationally prohibitive. Data handling. Validation.	Await future increased computing power at same cost. Develop ROM and ANN models. In situ processing. Measurements on existing observatories.
ROM and ANN models for observatory operation (AIVC phase)	Early development	Insufficient funding for required FTEs and computational resources. Uncertainty for the number of parameters ANNs can stably accommodate and the number of simulations required for training.	If insufficient funding persists, roll to operations and investigate alternative funding avenues. Break down to simpler tasks, few inputs/outputs per task. Test software on existing observatories.

10. Concluding Remarks

This paper attempted to provide a summary on the role aerothermal modeling has played in ground-based astronomy over the last two decades and to suggest a possible roadmap for the coming decade.

From GL characterization to enclosure aerodynamic optimization, observatory day-time air conditioning, nighttime heat release control, and subsystem design support, CFD and conjugate heat transfer modeling have become indispensable tools in ELT systems engineering. Extensive system modeling and simulations have been conducted to predict system performance and evaluate against technical budgets (e.g., image quality, heat dissipation) and requirements (e.g., pointing accuracy, environmental constraints). Such performance prediction models support four important systems engineering functions: requirement validation, verification, risk mitigation, and trade-off studies. Simulations have significantly influenced ELT design choices. In the trade space today, perhaps there exist fewer opportunities to modify design than to instead support integration, verification, commissioning, and troubleshooting, as well as optimize performance via operational strategies. Hidden challenges and new opportunities, however, lie abundant.

As the projects continue determined on their path of exploration and discovery, so will the efforts to overcome the new challenges. Using novel techniques and embracing modern technologies, aerothermal modeling will make new discoveries to ensure the success of this groundbreaking engineering endeavor.

Author Contributions: Conceptualization, K.V. and G.A.; methodology, K.V.; software, K.V. and R.C.; validation, K.V.; formal analysis, K.V. and R.C.; investigation, K.V.; resources, K.V. and R.C.; data curation, K.V. and R.C.; writing—original draft preparation, K.V.; writing—review and editing, K.V.; visualization, K.V. and R.C.; supervision, G.A., G.T. and R.C.; project administration, G.A., G.T. and R.C.; funding acquisition, G.A., G.T. and R.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by TIO, contract number TIO.BUS.CON.24.063 and GMTO, contract number GMT-CON-01543.

Data Availability Statement: The data presented in this study are not readily available because they reside in access-controlled repositories. Some are no longer retrievable. Most of these studies are ongoing and the design not finalized. Requests to access the datasets should be directed to granchot@tmt.org and rconan@gmto.org.

Acknowledgments: The TMT International Observatory (TIO) gratefully acknowledges the support of the TIO Members and the Gordon and Betty Moore Foundation. The Members and their respective collaborating institutions are as follows: California Institute of Technology (Caltech); the National

Research Council of Canada (NRC-CNRC); the Association of Canadian Universities for Research in Astronomy (ACURA); the Department of Science and Technology, Government of India (DST); the Department of Atomic Energy of India (DAE); the National Institute of Natural Sciences of Japan (NINS); the National Astronomical Observatory of Japan (NAOJ); and the Regents of the University of California (UC). This work was supported as well by the Association of Universities for Research in Astronomy (AURA) and U.S. National Science Foundation (NSF).

This work was supported by the GMTO Corporation, a non-profit organization operated on behalf of an international consortium of universities and institutions: Arizona State University, Astronomy Australia Ltd., the Australian National University, the Carnegie Institution for Science, Harvard University, the Korea Astronomy and Space Science Institute, the São Paulo Research Foundation, the Smithsonian Institution, the University of Texas at Austin, Texas A&M University, the University of Arizona, the University of Chicago, and the Weizmann Institute of Science.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. European Extremely Large Telescope. Available online: <https://elt.eso.org> (accessed on 10 June 2024).
2. Giant Magellan Telescope. Available online: <https://giantmagellan.org> (accessed on 10 June 2024).
3. Thirty Meter Telescope International Observatory. Available online: <https://www.tmt.org> (accessed on 10 June 2024).
4. Angeli, G.Z.; Roberts, S.; MacMynowski, D.; Fitzsimmons, J.; Dunn, J.; Vogiatzis, K.; Sheinis, A. *Modeling Tools to Estimate the Performance of the Thirty Meter Telescope: An Integrated Approach*; SPIE 5497(24); SPIE: Bellingham, WA, USA, 2004.
5. Conan, R.; Angeli, G.; Sitarski, B.; Dribusch, C.; Fitzpatrick, H.; Romano, H.; Vogiatzis, K.; Das, K. *GMTO Approach to Integrated Modeling-Based Systems Engineering*; SPIE 11450(26); SPIE: Bellingham, WA, USA, 2020.
6. Angeli, G.Z.; Vogiatzis, K. *Statistical Approach to Systems Engineering for the Thirty Meter Telescope*; SPIE 7738(45); SPIE: Bellingham, WA, USA, 2010.
7. Vogiatzis, K.; Das, K.; Sitarski, B.; Schwartz, D.; Conan, R.; Angeli, G.Z. *Using a Stochastic Approach to Estimate the Aerothermal Performance of the Giant Magellan Telescope*; AIAA 2019-2319; American Institute of Aeronautics and Astronautics: Reston, VA, USA, 2019.
8. Vogiatzis, K.; Angeli, G.Z. *Monte Carlo Simulation Framework for TMT*; SPIE 7017(29); SPIE: Bellingham, WA, USA, 2008.
9. Sitarski, B.; Vogiatzis, K.; Conan, R.; Angeli, G.Z.; Goodrich, R. Performance Estimates for the Giant Magellan Telescope via Statistical Analysis. In Proceedings of the AO4ELT6, Quebec City, QC, Canada, 9–14 June 2019.
10. Seo, B.-J.; Angeli, G.Z.; Nelson, J. Analysis of normalized point source sensitivity as a performance metric for large telescopes. *J. Appl. Opt.* **2009**, *48*, 5997–6007. [CrossRef] [PubMed]
11. Angeli, G.Z.; Seo, B.-J.; Nissly, C.; Troy, M. *A Convenient Telescope Performance Metric for Imaging through Turbulence*; SPIE 8127(4); SPIE: Bellingham, WA, USA, 2011.
12. Conan, R. *PSSn as a Function of Field Aberration and M1 Segment Polishing Residuals*; GMT-DOC-01940; GMTO Corporation: Pasadena, CA, USA, 2017.
13. Siemens Simcenter STAR-CCM+. Available online: <https://plm.sw.siemens.com/en-US/simcenter/fluids-thermal-simulation/star-ccm/> (accessed on 10 June 2024).
14. Mathworks MATLAB. Available online: <https://www.mathworks.com/products/matlab.html> (accessed on 10 June 2024).
15. Zago, L. The Effect of the Local Atmospheric Environment on Astronomical Observations. Ph.D. Thesis, Ecole Polytechnique Fédérale de Lausanne, Vaud, Switzerland, 1995. Thèse No. 1394.
16. Vogiatzis, K.; Thompson, H.; Tranco, G. *Aerothermal Support for TMT International Observatory Enclosure and Summit Facilities Subsystems*; SPIE 13094(27); SPIE: Bellingham, WA, USA, 2024.
17. Sitarski, B.; Vogiatzis, K. *HVAC Sizing Thermal Simulations*; GMT-DOC-03957; GMTO Corporation: Pasadena, CA, USA, 2019.
18. Vogiatzis, K.; Fitzpatrick, H.; Conan, R.; Ranka, T.; Muller, G.; Xin, B.; Ashby, D. *GMT Primary Mirror Thermal Control System and Thermal Deformation Modeling Framework*; SPIE 13099(102); SPIE: Bellingham, WA, USA, 2024.
19. Vogiatzis, K.; De Young, D.S. Site Characterization through the Use of Computational Fluid Dynamics. In Proceedings of the 2nd Baekaskog Workshop on Extremely Large Telescopes, Baekaskog Castle, Sweden, 9–11 September 2003. SPIE 5382(74).
20. Vogiatzis, K.; Hiriart, D. Numerical Wind Modeling for the San Pedro Martir Sierra in Baja California. *Rev. Mex. Astron. Astrophys.* **2004**, *40*, 81–98.
21. Els, S.; Vogiatzis, K. *TMT Studies on the Influence of Short-Term Changes of Local Thermodynamic Conditions on the Seeing: CFD vs. Observations*; SPIE 6267(55); SPIE: Bellingham, WA, USA, 2006.
22. Els, S.; Vogiatzis, K. Revealing the Onset of Convection in Terrestrial Planet Atmospheres. In Proceedings of the IAU S239 Convention of Astrophysics, Prague, Czech Republic, 21–25 August 2006.
23. Els, S.; Schöck, M.; Skidmore, W.; Travouillon, T.; Riddle, R.; Vogiatzis, K.; Wang, L.; Bustos, E.; Seguel, J.; Vasquez, J.; et al. Observations of the Turbulent Atmospheric Ground Layer during the Site Testing for the Thirty Meter Telescope (TMT). In Proceedings of the EMS 2007-A-00109, Annual Meeting of the European Meteorological Society and 8th European Conference on Applications of Meteorology, San Lorenzo de El Escorial, Spain, 1–5 October 2007.

24. Els, S.; Schöck, M.; Skidmore, W.; Travouillon, T.; Riddle, R.; Bustos, E.; Seguel, J.; Vasquez, J.; Vogiatzis, K.; Walker, D.; et al. The Site Testing for the Thirty Meter Telescope and its Potential Role in Developing the Short Term Forecasting of Observing Conditions. *Astron. Nachrichten* **2007**, *328*, 7.
25. Els, S.G.; Vogiatzis, K.; Otarola, A.; Schöck, M.; Travouillon, T.; Riddle, R.L.; Skidmore, W.A. *Optical Turbulence above Mountains Seen in 3D*; SPIE 7733(58); SPIE: Bellingham, WA, USA, 2010.
26. Das, K.; Vogiatzis, K.; Angeli, G.Z.; Rosenthal, W.; Bouchez, A.; Goodrich, W. *GMT Aerothermal Modeling Validation through Site Measurement*; SPIE 10705(2)s; SPIE: Bellingham, WA, USA, 2018.
27. Vogiatzis, K.; Thompson, H.; Roberts, S.; Dumas, C. *Ground Layer Studies for the Alternate TMT Site*; SPIE 10705(61); SPIE: Bellingham, WA, USA, 2018.
28. Vogiatzis, K.; Thompson, H.; Rogers, J.; Trancho, G. *On the Relationship between Thermal Seeing and Observatory Design*; SPIE 12187(74); SPIE: Bellingham, WA, USA, 2022.
29. De Young, D.S.; Vogiatzis, K. Numerical Simulations of Flow in Extremely Large Telescope Enclosures. In Proceedings of the SPIE 5382-44, 2nd Baekaskog Workshop on Extremely Large Telescopes, Baekaskog Castle, Sweden, 9–11 September 2003.
30. MacMynowski, D.G.; Angeli, G.Z.; Vogiatzis, K.; Fitzsimmons, J.; Padin, S. *Parametric Modeling and Control of Telescope Wind-Induced Vibration*; SPIE 5497(26); SPIE: Bellingham, WA, USA, 2004.
31. Vogiatzis, K.; Segurson, A.; Angeli, G.Z. *Estimating the Effect of Wind Loading on Extremely Large Telescope Performance Using Computational Fluid Dynamics*; SPIE 5497(30); SPIE: Bellingham, WA, USA, 2004.
32. MacMynowski, D.G.; Blaurock, C.A.; Angeli, G.Z.; Vogiatzis, K. *Modeling Wind-Buffering of the Thirty Meter Telescope*; SPIE 6271(25); SPIE: Bellingham, WA, USA, 2006.
33. MacMynowski, D.G.; Vogiatzis, K.; Angeli, G.Z.; Fitzsimmons, J.; Nelson, E.J. Wind Loads on Ground-Based Telescopes. *Appl. Opt.* **2006**, *45*, 7912–7923. [[CrossRef](#)]
34. Sebag, J.; Vogiatzis, K. *LSST Camera Heat Requirements Using CFD and Thermal Seeing Modeling*; SPIE 7738(13); SPIE: Bellingham, WA, USA, 2010.
35. Sebag, J.; Vogiatzis, K.; Barr, J.; Neill, D. *LSST Summit Enclosure: Facility Design Optimization Using Aero-Thermal Modeling*; SPIE 8449(3); SPIE: Bellingham, WA, USA, 2012.
36. Cho, M.; Vogiatzis, K.; Sebag, J.; Neill, D. *Wind Responses of the LSST Secondary Mirror*; SPIE 8449(4); SPIE: Bellingham, WA, USA, 2012.
37. Sebag, J.; Vogiatzis, K. *Estimating Dome Seeing for LSST*; SPIE 9150(26); SPIE: Bellingham, WA, USA, 2014.
38. Vogiatzis, K.; Xin, B.; Claver, C.; Angeli, G.Z.; Selvy, B. *Dome Seeing Sensitivity Analysis for LSST*; SPIE 10705(3); SPIE: Bellingham, WA, USA, 2018.
39. Marchiori, G.; Rampini, F.; Mian, S.; Ghedin, L.; Marcuzzi, E.; Bressan, R. *ELT Dome and Telescope: Performance Analysis Overview*; SPIE 10705(13); SPIE: Bellingham, WA, USA, 2018.
40. Vogiatzis, K.; MacMartin, D.; Wang, L.; Trancho, G. *The TMT International Observatory Aerothermal Performance Estimation Procedure*; SPIE 13099(100); SPIE: Bellingham, WA, USA, 2024.
41. Vogiatzis, K.; Das, K.; Angeli, G.Z.; Bigelow, B.; Burgett, W. *Computational Fluid Dynamics Modeling of GMT*; SPIE 10705(28); SPIE: Bellingham, WA, USA, 2018.
42. MacMartin, D.G.; Vogiatzis, K. *Unsteady Wind Loads for TMT: Replacing Parametric Models with CFD*; SPIE 9150(25); SPIE: Bellingham, WA, USA, 2014.
43. Vogiatzis, K.; Thompson, H. *On the Precision of Aero-Thermal Modeling for TMT*; SPIE 9911(40); SPIE: Bellingham, WA, USA, 2016.
44. Vogiatzis, K. *Effects of Meshing, Solver Parameters and Post-Processing on Aero-Optical Image Quality Metric Modeling*; AIAA 2018-1502; American Institute of Aeronautics and Astronautics: Reston, VA, USA, 2018.
45. Das, K.; Vogiatzis, K.; Angeli, G.Z. *Assessment of Turbulence Models and Mesh Refinement for Aero-Optical Simulations of Large Ground Based Telescopes*; AIAA 2018-4156; American Institute of Aeronautics and Astronautics: Reston, VA, USA, 2018.
46. Vogiatzis, K. *Thermal Modeling Environment for TMT*; SPIE 7738(11); SPIE: Bellingham, WA, USA, 2010.
47. Vogiatzis, K.; Conan, R. *Preliminary Evaluation of the GMT HVAC System Using CFD*; GMT-DOC-03938; GMTO Corporation: Pasadena, CA, USA, 2019.
48. Thompson, H.A.; Vogiatzis, K. *Heat Balance and Thermal Management of the TMT Observatory*; SPIE 9150(83); SPIE: Bellingham, WA, USA, 2014.
49. Pazder, J.; Vogiatzis, K. *Dome and Mirror Seeing Estimates for the Thirty Meter Telescope*; SPIE 7017(26); SPIE: Bellingham, WA, USA, 2008.
50. Conan, R.; van Dam, M.; Vogiatzis, K.; Das, K.; Bouchez, A. Modeling the Giant Magellan Telescope dome seeing using fluid dynamics simulations. In Proceedings of the AO4ELT6, Quebec City, QC, Canada, 9–14 June 2019.
51. Vogiatzis, K.; Angeli, G.Z. *Strategies for Estimating Mirror and Dome Seeing for TMT*; SPIE 6271(27); SPIE: Bellingham, WA, USA, 2006.
52. Vogiatzis, K.; Upton, R. *TMT Studies on Thermal Seeing Modeling: Mirror and Dome Seeing Model Validation*; SPIE 6271(48); SPIE: Bellingham, WA, USA, 2006.
53. McMynowski, D.G.; Colavita, M.M.; Skidmore, W.A.; Vogiatzis, K. *Primary Mirror Dynamic Disturbance Models for TMT: Vibration and Wind*; SPIE 7738(14); SPIE: Bellingham, WA, USA, 2010.
54. Vogiatzis, K.; McMynowski, D.G. *Unsteady Aerodynamic Simulations for TMT Primary Mirror Segment Wind Loading*; SPIE 7738(61); SPIE: Bellingham, WA, USA, 2010.
55. Soler, M.; Vogiatzis, K.; Cozar-Castellano, J.; Bonaque, S.; Bello-Asin, M.; Nunez Cagigal, M.; Baretto Cabrera, M. *Local Seeing Analysis for the European Solar Telescope*; SPIE 13094(153); SPIE: Bellingham, WA, USA, 2024.

56. Vogiatzis, K.; Otarola, A.; Skidmore, W.; Travouillon, T.; Angeli, G.Z. *Local Thermal Seeing Modeling Validation through Observatory Measurements*; SPIE 8449(1); SPIE: Bellingham, WA, USA, 2012.
57. Salmon, D.; Benedict, T.; Baril, M.; Wang, S.-Y.; Thanjavur, K.; Vogiatzis, K.; Racine, R. *A Fluid Dynamic Study of the CFHT Dome Using CFD and Water Tunnel Tests*; SPIE 8449(2); SPIE: Bellingham, WA, USA, 2012.
58. Vogiatzis, K. Aerothermal modeling framework for TMT. In *Proceedings of the Integrated Modeling of Complex Optomechanical Systems*, Kiruna, Sweden, 15–17 August 2011. SPIE 8336(10).
59. Vogiatzis, K.; Fitzpatrick, H.; Das, K.; Conan, R.; Angeli, G.; Sitarski, B.; Romano, R.; Dribusch, C.; Walls, B. *Aerothermal Modeling for Design Support, Requirement Validation and Performance Assessment of GMT Subsystems*; SPIE 12187(22); SPIE: Bellingham, WA, USA, 2022.
60. Cho, M.; Corredor, A.; Vogiatzis, K.; Angeli, G.Z. *Thermal Analysis of the TMT Telescope Structure*; SPIE 7738(12); SPIE: Bellingham, WA, USA, 2010.
61. Vogiatzis, K.; Sadjadpour, A.; Roberts, S. *TMT Telescope Structure Thermal Model*; SPIE 9150(72); SPIE: Bellingham, WA, USA, 2014.
62. Fitzpatrick, H.; Conan, R.; Vogiatzis, K.; Das, K.; Dribusch, C. *A Computational Framework for the Study of the Diurnal Thermomechanical Behavior of the Giant Magellan Telescope*; SPIE 13099(98); SPIE: Bellingham, WA, USA, 2024.
63. Vogiatzis, K.; Boyer, C.; Wei, K.; Tang, J.; Ellerbroek, B. *Aero-Thermal Simulations of the TMT Laser Guide Star Facility*; SPIE 9148(243); SPIE: Bellingham, WA, USA, 2014.
64. Vogiatzis, K.; Wang, L.; Trubey, M.; Irarrazaval, B.; Boyer, C. *Aerothermal Modeling of the TMT International Observatory Laser Guide Star Facility*; SPIE 13099(101); SPIE: Bellingham, WA, USA, 2024.
65. Vogiatzis, K. *Ion Beam Figuring Modeling*; Document TMT.SEN.TEC.20.070; TMT International Observatory: Pasadena, CA, USA, 2020.
66. McQuillen, I.; Phelps, L. *Factory Acceptance Testing and Model Refinement for the Daniel K. Inouye Solar Telescope air Knife Assembly*; SPIE 10705(25); SPIE: Bellingham, WA, USA, 2018.
67. Martins, D.; Holzlohner, R.; Verinaud, C.; Kleinclaus, C. *Transient Wavefront Error from Cooled Air Downwind of Telescope Spiders*; SPIE 12187(12); SPIE: Bellingham, WA, USA, 2022.
68. CFD Vision 2030 Study. Available online: <http://ntrs.nasa.gov/search.jsp?R=20140003093> (accessed on 10 June 2024).
69. AIAA CFD Vision 2030 Integration Committee Website. Available online: <https://cfd2030.com> (accessed on 10 June 2024).
70. Vogiatzis, K. *Transient Aero-Thermal Simulations for TMT*; SPIE 9150(24); SPIE: Bellingham, WA, USA, 2014.
71. Vogiatzis, K.; Thompson, H.; Roberts, S. *On the Precision of Aero-Thermal Simulations for TMT: Revisited*; SPIE 10705(4); SPIE: Bellingham, WA, USA, 2018.
72. Racine, R.; Salmon, D.; Cowley, D.; Sovka, J. Mirror, dome, and natural seeing at CFHT. *Publ. Astron. Soc. Pac.* **1991**, *103*, 1020–1032. [[CrossRef](#)]
73. Bustos, E.; Tokovinin, A. *Dome Seeing Monitor and Its Results for the 4m Blanco Telescope*; SPIE 10700; SPIE: Bellingham, WA, USA, 2018.
74. Lai, O.; Withington, K.; Chun, M.R.; Laugier, R. Investigating ground layer and dome turbulence in astronomical observatories using a localized optical turbulence sensor. In *Proceedings of the Imaging and Applied Optics Congress*, Washington, DC, USA, 22–26 June 2020.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Fractional-Order Degn–Harrison Reaction–Diffusion Model: Finite-Time Dynamics of Stability and Synchronization

Ma'mon Abu Hammad ¹, Issam Bendib ², Waseem Ghazi Alshanti ³, Ahmad Alshanty ⁴, Adel Ouannas ⁵, Amel Hioual ^{5,*}  and Shaher Momani ^{6,7}

¹ Department of Mathematics, Al-Zaytoonah University of Jordan, Amman 11733, Jordan; m.abuhammad@zuj.edu.jo

² Applied Mathematics & Modeling Laboratory, Department of Mathematics, Faculty of Exact Sciences, University of Brothers Mentouri, Constantine 25000, Algeria; bendib.issam@doc.umc.edu.dz

³ Department of Mathematics, Al Zaytoonah University of Jordan, Amman 11733, Jordan; w.alshanti@zuj.edu.jo

⁴ Cyber Security Department, Al Zaytoonah University of Jordan, Amman 11733, Jordan; a.alshanty@zuj.edu.jo

⁵ Department of Mathematics and Computer Science, University of Oum EL-Bouaghi, Oum El Bouaghi 04000, Algeria; ouannas.adel@univ-oeb.dz

⁶ Nonlinear Dynamics Research Center (NDRC), Ajman University, Ajman 346, United Arab Emirates; s.momani@ajman.ac.ae

⁷ Department of Mathematics, The University of Jordan, Amman 11942, Jordan

* Correspondence: amel.hioual@univ-oeb.dz

Abstract: This study aims to address the topic of finite-time synchronization within a specific subset of fractional-order Degn–Harrison reaction–diffusion systems. To achieve this goal, we begin with the introduction of a novel lemma specific for finite-time stability analysis. Diverging from existing criteria, this lemma represents a significant extension of prior findings, laying the groundwork for subsequent investigations. Building upon this foundation, we proceed to develop efficient dependent linear controllers designed to orchestrate finite-time synchronization. Leveraging the power of a Lyapunov function, we derive new, robust conditions that ensure the attainment of synchronization within a predefined time frame. This innovative approach not only enhances our understanding of finite-time synchronization, but also offers practical solutions for its realization in complex systems. To validate the efficacy and applicability of our proposed methodology, extensive numerical simulations are conducted. Through this comprehensive analysis, we aim to contribute valuable insights to the field of fractional-order reaction–diffusion systems while paving the way for practical implementations in real-world applications.

Keywords: finite-time synchronization; Degn–Harrison reaction–diffusion systems; finite-time stability; Lyapunov function



Citation: Hammad, M.A.; Bendib, I.; Alshanti, W.G.; Alshanty, A.; Ouannas, A.; Hioual, A.; Momani, S. Fractional-Order Degn–Harrison Reaction–Diffusion Model: Finite-Time Dynamics of Stability and Synchronization. *Computation* **2023**, *12*, 144. <https://doi.org/10.3390/computation12070144>

Academic Editors: Endre Kovács, Denis Butusov and Valerii Ostrovskii

Received: 3 June 2024

Revised: 7 July 2024

Accepted: 10 July 2024

Published: 12 July 2024



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Developing appropriate controllers is essential for dynamic system analysis, particularly for achieving synchronization. Researchers have explored various mathematical models to understand the unique synchronization patterns of these systems. Mathematical models, such as reaction–diffusion equations, are employed to describe continuous-space systems like oscillating biological or chemical media, revealing complex dynamical structures such as bifurcations, spatial patterns, and turning instability. Studies suggest that reaction–diffusion systems, resembling low-dimensional oscillators, can exhibit synchronization. For instance, reference [1] examined synchronization in the Degn–Harrison reaction–diffusion system, reference [2] focused on synchronization control in the Lengyel–Epstein system, and reference [3] discussed synchronization in the FitzHugh–Nagumo model using a specific control approach. Over the past decade, there has been a significant increase in interest in fractional reaction–diffusion systems, especially regarding synchronization [4]. For example, in [5], synchronization in a specific type of fractional-order

spatiotemporal partial differential systems was investigated using the fractional Lyapunov technique. Furthermore, a hybrid approach was developed to synchronize between two integer and fractional-order reaction–diffusion systems, with applications to specific chemical models discussed in [6].

Within this framework, various methods and strategies have emerged to tackle the challenge of synchronizing partial differential equations. These approaches span a wide spectrum, encompassing active control techniques alongside both linear and nonlinear control schemes, as outlined in [7,8]. Diverse forms of synchronization have been delineated, each with its unique characteristics and applications. These include finite-time synchronization, phase synchronization, anticipated synchronization, complete synchronization, lag synchronization, function Q-S synchronization, projective synchronization, and generalized synchronization, as elaborated upon in references [9–11]. Through an extensive and thorough review of various methodologies, our objective is to delve deeply into the complexities of synchronization in partial differential equations. We aim to elucidate the diverse array of approaches that researchers and practitioners have developed and utilized within this field. By integrating insights from a broad spectrum of existing literature and synthesizing the collective findings of numerous past studies, we aspire to make a meaningful contribution to the ongoing scholarly conversation surrounding synchronization theory. Our work not only highlights the current state of research, but also identifies potential pathways for future advancements, thereby fostering further progress in this critical area of study.

Finite-time synchronization refers to a particular type of synchronization where coordinated behavior among systems is attained within a predetermined and finite time interval. Unlike traditional synchronization, which might require an indefinite period to converge, finite-time synchronization ensures that systems achieve synchrony within a defined time frame. This concept finds applications across various fields, such as control systems, communication networks, and robotics. For example, in the field of robotics, finite-time synchronization can enhance the efficiency and collaboration among multiple robotic agents by guaranteeing their coordinated movement within a designated time period. In communication networks, rapid synchronization is crucial for ensuring effective and reliable data transmission, as highlighted in references [12–15]. Despite its importance, research on finite-time synchronization of spatially extended systems described by reaction–diffusion systems (RDSs) remains limited.

In recent years, various numerical methods have been developed to solve reaction–diffusion equations effectively. Standard methods such as finite difference, finite element, and spectral methods have been widely used due to their robustness and computational efficiency [16–18]. Advancements in numerical analysis have led to more techniques, such as adaptive mesh refinement and parallel computing, enhancing the accuracy and efficiency of simulations. For fractional-order systems like the Degrn–Harrison reaction–diffusion model, specialized methods such as fractional finite difference and fractional spectral methods have been proposed. These methods effectively address the unique challenges of fractional derivatives and provide deeper insights into the system’s dynamics [19,20].

This research presents a significant and original contribution to the study of finite-time synchronization within a particular subset of fractional-order spatiotemporal partial differential systems. Our main goal is to conduct a comprehensive examination of both finite-time stability and finite-time synchronization in interconnected fractional reaction–diffusion systems. These systems are characterized by their complex dynamic behaviors, which are influenced by spatial and temporal variables, making them challenging yet crucial to study. To achieve finite-time stability, we employ a fractional Lyapunov method. This approach involves constructing a suitable Lyapunov function that helps in deriving the conditions under which the system’s equilibrium point can be reached in a finite amount of time. This method is pivotal, as it ensures that the system’s state will converge to the desired equilibrium within a predetermined timeframe, thereby enhancing the predictability and control of such systems. For attaining finite-time synchronization, we suggest a robust linear

control technique. This control strategy is designed to synchronize the states of fractional reaction–diffusion systems, ensuring that they exhibit coordinated behavior within a finite time interval. Our approach is general enough to be applicable to a wide range of fractional reaction–diffusion systems, making it a versatile tool for researchers and practitioners in the field. To illustrate the efficacy and practicality of our proposed control methods, we conduct detailed analyses of the synchronization behaviors of interconnected fractional-order Degn–Harrison systems. These systems serve as a representative example due to their complexity and relevance in various applications. Through simulations and theoretical analysis, we demonstrate how our control techniques can effectively achieve finite-time synchronization, thereby validating the robustness and applicability of our approach.

The paper is structured as follows: In Section 2, we delve into the foundational principles of fractional calculus, providing readers with a comprehensive understanding of the mathematical underpinnings essential for grasping the subsequent discussions. Section 3 provides a description of the Degn–Harrison models. Section 4 aims to investigate the finite-time stability of equilibrium points in reaction–diffusion models. This analysis is crucial for understanding the transient behavior of the system and determining whether it can achieve a stable state within a finite time. The practical implications of this analysis are significant, including applications in control system design, pattern formation in chemical and biological processes, and the synchronization of dynamic systems. Section 5 focuses on establishing a finite-time synchronization scheme for reaction–diffusion systems, specifically through the study of the Degn–Harrison reaction–diffusion system. Section 6 presents a diverse set of examples to illustrate the principles of finite-time stability and finite-time synchronization within the Degn–Harrison system.

2. Basic Tools

The aim of this section is to introduce the fundamental definitions and key theorems associated with fractional calculus, which are essential for understanding the subsequent analysis.

Definition 1 ([21,22]). *The Riemann–Liouville fractional-order integral operator of an arbitrary integrable function $\mathcal{E}$ can be defined as:*

$$\mathcal{I}_{\mathfrak{z}}^{\aleph} \mathcal{E}(\mathfrak{z}) = \frac{1}{\Gamma(\aleph)} \int_0^{\mathfrak{z}} (\mathfrak{z} - \varsigma)^{\aleph-1} \mathcal{E}(\varsigma) d\varsigma, \quad (1)$$

where $\mathfrak{z} \geq 0, 0 < \aleph < 1$, and $\Gamma(\cdot)$ is the gamma function.

Definition 2 ([21,22]). *The Caputo fractional derivative of $\mathcal{E}$ is defined as:*

$${}_0^C \mathcal{D}_{\mathfrak{z}}^{\aleph} \mathcal{E}(\mathfrak{z}) = \frac{1}{\Gamma(1-\aleph)} \int_0^{\mathfrak{z}} (\mathfrak{z} - \varsigma)^{-\aleph} \frac{d}{d\varsigma} \mathcal{E}(\varsigma) d\varsigma, \quad (2)$$

where $0 < \aleph < 1$.

We consider the following fractional-order nonlinear system:

$$\begin{cases} {}_0^C \mathcal{D}_{\mathfrak{z}}^{\aleph} \mathcal{E}(\mathfrak{z}) = \chi(\mathfrak{z}, \mathcal{E}(\mathfrak{z})), & \aleph \in (0, 1), \mathfrak{z} > 0, \\ \mathcal{E}(0) = \mathcal{E}_0. \end{cases} \quad (3)$$

where $\chi : \mathbb{R}^n \rightarrow \mathbb{R}^n, \mathcal{E}(\mathfrak{z}) \in \mathbb{R}^n$, and ${}_0^C \mathcal{D}_{\mathfrak{z}}^{\aleph}$ is the Caputo fractional derivative of order $\aleph$, and a point $\mathcal{E}^* \in \mathbb{R}^n$ is an equilibrium point of fractional-order nonlinear system (3) if $\chi(\mathcal{E}^*) \equiv 0$.

Definition 3 ([23]). $\mathcal{E}^*$ is called as the initial equilibrium point of fractional-order nonlinear system (3) if $\mathcal{E}_0 = \mathcal{E}^*$, or equivalently, $\mathcal{E}(\mathfrak{z}) \equiv \mathcal{E}^*, \forall \mathfrak{z} \in [\mathfrak{z}^*, +\infty)$.

Definition 4 ([23]). $\mathcal{E}^*$ is called as the finite-time equilibrium point of fractional-order nonlinear system (3) if there is a $z^* > 0$, such that $\mathcal{E}(z) \neq \mathcal{E}^*, z \in [0, z^*)$, and $\mathcal{E}(z) \equiv \mathcal{E}^*, \forall z \in [z^*, +\infty)$.

Definition 5 ([23]). Suppose $D \subseteq \Omega$ is an open neighborhood of 0, if there is a function $z^*(\mathcal{E}_0) : D \setminus \{0\} \rightarrow \mathbb{R}_+^*$ called the setting-time satisfying the upcoming conditions, then $\mathcal{E}^*$ is called a finite-time stable equilibrium point of the system (3).

1. $\forall \mathcal{E}_0 \in D \setminus \{0\}, \mathcal{E}(z) = \mathcal{E}(z, \mathcal{E}_0) \in D \setminus \{0\}, \forall z \in [0, z^*(\mathcal{E}_0)), \lim_{z \rightarrow z^*(\mathcal{E}_0)} \mathcal{E}(z) = 0$, and $\mathcal{E}(z) \equiv 0$ for $z \in [z^*(\mathcal{E}_0), +\infty)$.
2. For every open neighborhood U_ϵ of 0, there is an open subset U_δ containing 0 in D , such that for $\forall \mathcal{E}_0 \in U_\delta \setminus \{0\}$ and $\forall z \in [0, z^*(\mathcal{E}_0)), \mathcal{E}(z) = \mathcal{E}(z, \mathcal{E}_0) \in U_\epsilon$.

Theorem 1 ([23]). $\mathcal{E}^*$ is an finite-time stability equilibrium point of the fractional-order nonlinear systems (3) if there exists a differentiable Lyapunov function $\mathcal{L} : [0, +\infty) \times \Omega \rightarrow \mathbb{R}_+$. three class $\mathcal{M}$ functions $\beta_1, \beta_2, \beta_3$, and $\delta > 0$ such that

1. $\beta_1 \|\mathcal{E}(z)\| \leq \mathcal{L}(z, \mathcal{E}(z)) \leq \beta_2 \|\mathcal{E}(z)\|$,
2. $\mathcal{L}'(z, \mathcal{E}(z)) < -\beta_3 \mathcal{E}(z, \mathcal{E}(z))$,
3. $\int_0^\epsilon \frac{d\mathcal{E}}{\beta_3(\mathcal{E})} < +\infty, (\forall \epsilon : 0 < \epsilon \leq \delta)$.

Lemma 1 ([24,25]). For all $\mathcal{E}(z) \in \mathbb{R}^n$ and $\aleph \in (0, 1)$, the following inequality holds

$${}_0^C \mathcal{D}_z^\aleph (\mathcal{E}^T(z) \mathcal{E}(z)) \leq 2\mathcal{E}^T(z) {}_0^C \mathcal{D}_z^\aleph \mathcal{E}(z), \quad z \geq 0. \quad (4)$$

Definition 6 ([26]). The Mittag–Leffler function with $\aleph > 0$ noted by $E_\aleph$ and defined by

$$E_\aleph(z) = \sum_{j=0}^{+\infty} \frac{z^\aleph}{\Gamma(\aleph j + 1)}, \quad (5)$$

Theorem 2 ([27]). Let $\mathcal{L}(z)$ be a continuous and positive definite function which satisfies

$${}_0^C \mathcal{D}_z^\aleph \mathcal{L}(z) \leq -\Omega \mathcal{L}(z), \quad \forall z \geq 0, \quad (6)$$

where $0 < \aleph < 1$ and Ω is a positive constant. Then,

$$\mathcal{L}(z) \leq \mathcal{L}(0) E_\aleph(-\Omega z^\aleph), \quad \forall z \in [0, z^*), \quad (7)$$

where $E_\aleph$ is a Mittag–Leffler function, with the setting time.

$$z^* = \left(\frac{\Gamma(\aleph + 1)}{\Omega} \right)^{\aleph^{-1}}, \quad (8)$$

Moreover, $\mathcal{L}(0)$ is positive and $\mathcal{L}(z) = 0$ for $z \geq z^*$.

3. Description of the Models

The Degn–Harrison system illustrates the reaction process involving two substances: oxygen, denoted as $\mathcal{E}_1$, and a nutrient, denoted as $\mathcal{E}_2$. A key aspect of this system is that an excess of oxygen inhibits respiration, which is represented by a nonlinear rate function $\frac{\mathcal{E}_1}{1 + q\mathcal{E}_1^2}$. This function captures the saturation effect, where the inhibitory influence of oxygen increases with its concentration, but levels off due to the presence of the quadratic term in the denominator.

Considering that similar inhibitory phenomena can be described using comparable reaction schemes, we extend the inhibitory law using the same nonlinear function $\frac{\mathcal{E}_1}{1 + q\mathcal{E}_1^2}$.

This extension allows for the modeling of a broader range of reaction processes where inhibition plays a crucial role.

The reaction–diffusion model established based on this extended inhibitory law is known as the Degn–Harrison reaction–diffusion model. This model is characterized by partial differential equations that describe the spatial and temporal evolution of the concentrations of $\mathcal{E}_1$, and $\mathcal{E}_2$. The governing equations for the Degn–Harrison reaction–diffusion model are given by:

$$\begin{cases} {}_0^C \mathcal{D}_z^\alpha \mathcal{E}_1(\tau, z) = \mathfrak{k}_1 \Delta \mathcal{E}_1 + \mathfrak{a} - \mathcal{E}_1 - \frac{\mathcal{E}_1 \mathcal{E}_2}{1+q\mathcal{E}_1^2}, & \tau \in \Omega, z > 0, \\ {}_0^C \mathcal{D}_z^\alpha \mathcal{E}_2(\tau, z) = \mathfrak{k}_2 \Delta \mathcal{E}_2 + \mathfrak{b} - \frac{\mathcal{E}_1 \mathcal{E}_2}{1+q\mathcal{E}_1^2}, & \tau \in \Omega, z > 0, \\ \frac{\partial \mathcal{E}_1}{\partial \tau} = \frac{\partial \mathcal{E}_2}{\partial \tau} = 0, & \tau \in \partial\Omega, z > 0, \\ \mathcal{E}_1(\tau, 0) = \mathcal{E}_{1,0}(\tau) > 0, \quad \mathcal{E}_2(\tau, 0) = \mathcal{E}_{2,0}(\tau) > 0, & \tau \in \Omega. \end{cases} \quad (9)$$

In this context, Ω represents a bounded domain in $\mathbb{R}^n$ with a sufficiently smooth boundary, denoted as $\partial\Omega$. The coefficients $\mathfrak{k}_1$ and $\mathfrak{k}_2$ indicate the diffusion rates of the reactants $\mathcal{E}_1$ and $\mathcal{E}_2$, respectively. The parameter q acts as a scaling factor influencing the strength of interaction between the variables $\mathcal{E}_1$ and $\mathcal{E}_2$. These diffusion coefficients are crucial parameters that determine how quickly substances spread throughout the domain Ω .

The constants $\mathfrak{a}$, $\mathfrak{b}$, q , $\mathfrak{k}_1$, and $\mathfrak{k}_2$ are assumed to be positive throughout the reaction process. These constants typically represent reaction rate parameters and other specific characteristics of the system's kinetics.

The Laplace operator, which accounts for the spatial diffusion of the reactants, is denoted by Δ . In an n -dimensional space, the Laplace operator is defined as:

$$\Delta = \sum_{i=1}^n \frac{\partial^2}{\partial \tau_i^2},$$

where $\frac{\partial^2}{\partial \tau_i^2}$ represents the second partial derivative with respect to the i -th spatial coordinate τ_i . References [28,29] provide further mathematical background and applications of the Laplace operator in reaction–diffusion systems, emphasizing its importance in describing the spatial aspects of the process.

The reaction–diffusion model, incorporating these definitions and assumptions, is crucial for analyzing the behavior of the Degn–Harrison system under the influence of diffusion and nonlinear reactions. The precise formulation of the model allows for a detailed study of the dynamics and synchronization properties of the system within the domain Ω .

4. Stability Analysis

The aim of investigating finite-time stability equilibrium points in reaction–diffusion models is to understand the transient behavior of the system and determine whether it can settle into a stable state within a finite time frame. This type of analysis is crucial for predicting the system's behavior over limited time periods, which has significant implications for various practical applications. These applications include the design of control systems, the formation of patterns in chemical and biological processes, and the synchronization of dynamic systems.

In our study, we focus on a unique equilibrium point, which serves as the solution to the following system of Equation (10). By analyzing this equilibrium point, we can gain insights into the stability and synchronization properties of the reaction–diffusion

system. This analysis helps in ensuring that the system reaches a desired state quickly and maintains that state despite the presence of disturbances or variations in initial conditions.

$$\begin{cases} \mathfrak{k}_1 \Delta \mathcal{E}_1^* + \mathfrak{a} - \mathcal{E}_1^* - \frac{\mathcal{E}_1^* \mathcal{E}_2^*}{1 + \mathfrak{q} \mathcal{E}_1^{*2}} = 0, \\ \mathfrak{k}_2 \Delta \mathcal{E}_2^* + \mathfrak{b} - \frac{\mathcal{E}_1^* \mathcal{E}_2^*}{1 + \mathfrak{q} \mathcal{E}_1^{*2}} = 0, \end{cases} \quad (10)$$

The unique equilibrium point of systems (10) is given by

$$\mathcal{E}^* = (\mathcal{E}_1^*, \mathcal{E}_2^*) = \left(\mathfrak{a} - \mathfrak{b}, \mathfrak{b} \frac{1 + \mathfrak{q}(\mathfrak{a} - \mathfrak{b})^2}{\mathfrak{a} - \mathfrak{b}} \right), \quad (11)$$

Lemma 2 ([30]). *The following inequality holds*

$$\left| \frac{\bar{\mathcal{E}}_1 \bar{\mathcal{E}}_2}{1 + \mathfrak{q} \bar{\mathcal{E}}_1^2} - \frac{\mathcal{E}_1 \mathcal{E}_2}{1 + \mathfrak{q} \mathcal{E}_1^2} \right| \leq \mathfrak{P}(|\bar{\mathcal{E}}_1 - \mathcal{E}_1| + |\bar{\mathcal{E}}_2 - \mathcal{E}_2|), \quad (12)$$

where

$$\mathfrak{P} \geq \max \left\{ \frac{5}{4} \mathfrak{K}, \frac{1}{2\sqrt{\mathfrak{q}}} \right\}, \quad |\bar{\mathcal{E}}_2| < \mathfrak{K}. \quad (13)$$

Lemma 3 ([31]). Ω is a bounded region in $\mathbb{R}^m$ with a smooth boundary denoted by $\partial\Omega$, which belongs to class $\mathcal{C}^2$. $\Xi(\mathfrak{r}) \in H_0^1(\Omega)$ represents a real-valued function, and $\frac{\partial \Xi(\mathfrak{r})}{\partial \eta}|_{\partial\Omega} = 0$. Then,

$$\zeta_1 \int_{\Omega} |\Xi(\mathfrak{r})|^2 d\mathfrak{r} \leq \int_{\Omega} |\nabla \Xi(\mathfrak{r})|^2 d\mathfrak{r}, \quad (14)$$

where ζ_1 is determined as the positive eigenvalue of the following problem:

$$\begin{cases} -\Delta \Xi(\mathfrak{r}) = \zeta \Xi(\mathfrak{r}), & \mathfrak{r} \in \Omega, \\ \frac{\partial \Xi(\mathfrak{r})}{\partial \eta} = 0, & \mathfrak{r} \in \partial\Omega. \end{cases} \quad (15)$$

Theorem 3. *The equilibrium point $\mathcal{E}^* = \left(\mathfrak{a} - \mathfrak{b}, \mathfrak{b} \frac{1 + \mathfrak{q}(\mathfrak{a} - \mathfrak{b})^2}{\mathfrak{a} - \mathfrak{b}} \right)$ represents a finite-time stable equilibrium point of the fractional-order nonlinear systems (9), provided that the following conditions are satisfied:*

$$|\mathcal{E}_2| < \mathfrak{K}, \quad \mathfrak{P} \geq \max \left\{ \frac{5}{4} \mathfrak{K}, \frac{1}{2\sqrt{\mathfrak{q}}} \right\}, \quad \min \{ \mathfrak{k}_1 \zeta_1 - 2\mathfrak{P} + 1, \mathfrak{k}_2 \zeta_2 - 2\mathfrak{P} \} > 0. \quad (16)$$

The time of finite-time stability of equilibrium point is estimated as

$$\mathfrak{z}_1^* = \frac{\mathcal{L}_2(0)}{2 \min \{ \mathfrak{k}_1 \zeta_1 - 2\mathfrak{P} + 1, \mathfrak{k}_2 \zeta_2 - 2\mathfrak{P} \} \mathcal{L}_2(\mathfrak{z}_{\max})}. \quad (17)$$

Proof. We construct the positive Lyapunov function as follows:

$$\mathcal{L}_2(\mathfrak{z}) = \frac{1}{2} \int_{\Omega} (\mathcal{E}_1 - \mathcal{E}_1^*)^2 + (\mathcal{E}_2 - \mathcal{E}_2^*)^2 d\mathfrak{r}, \quad (18)$$

where ${}_0\mathcal{D}_{\mathfrak{z}}$ denotes the normal derivative. According to Lemma 2, thus, we obtain:

$$\begin{aligned}
{}_0\mathcal{D}_3\mathcal{L}_2(\mathfrak{z}) &= \frac{1}{2} \int_{\Omega} {}_0\mathcal{D}_3(\mathcal{E}_1 - \mathcal{E}_1^*)^2 + {}_0\mathcal{D}_3(\mathcal{E}_2 - \mathcal{E}_2^*)^2 d\mathfrak{r} \\
&= \int_{\Omega} (\mathcal{E}_1 - \mathcal{E}_1^*) {}_0\mathcal{D}_3(\mathcal{E}_1 - \mathcal{E}_1^*) d\mathfrak{r} + \int_{\Omega} (\mathcal{E}_2 - \mathcal{E}_2^*) {}_0\mathcal{D}_3(\mathcal{E}_2 - \mathcal{E}_2^*) d\mathfrak{r} \\
&= \int_{\Omega} (\mathcal{E}_1 - \mathcal{E}_1^*) \left[\mathfrak{k}_1 \Delta(\mathcal{E}_1 - \mathcal{E}_1^*) - (\mathcal{E}_1 - \mathcal{E}_1^*) - \left(\frac{\mathcal{E}_1 \mathcal{E}_2}{1 + \mathfrak{q} \mathcal{E}_1^2} - \frac{\mathcal{E}_1^* \mathcal{E}_2^*}{1 + \mathfrak{q} \mathcal{E}_1^{*2}} \right) \right] d\mathfrak{r} \\
&\quad + \int_{\Omega} (\mathcal{E}_2 - \mathcal{E}_2^*) \left[\mathfrak{k}_2 \Delta(\mathcal{E}_2 - \mathcal{E}_2^*) - \left(\frac{\mathcal{E}_1 \mathcal{E}_2}{1 + \mathfrak{q} \mathcal{E}_1^2} - \frac{\mathcal{E}_1^* \mathcal{E}_2^*}{1 + \mathfrak{q} \mathcal{E}_1^{*2}} \right) \right] d\mathfrak{r} \\
&\leq \int_{\Omega} (\mathcal{E}_1 - \mathcal{E}_1^*) (\mathfrak{k}_1 \Delta(\mathcal{E}_1 - \mathcal{E}_1^*) - (\mathcal{E}_1 - \mathcal{E}_1^*)) d\mathfrak{r} + \int_{\Omega} |\mathcal{E}_1 - \mathcal{E}_1^*| \left| \frac{\mathcal{E}_1 \mathcal{E}_2}{1 + \mathfrak{q} \mathcal{E}_1^2} - \frac{\mathcal{E}_1^* \mathcal{E}_2^*}{1 + \mathfrak{q} \mathcal{E}_1^{*2}} \right| d\mathfrak{r} \\
&\quad + \int_{\Omega} (\mathcal{E}_2 - \mathcal{E}_2^*) \mathfrak{k}_2 \Delta(\mathcal{E}_2 - \mathcal{E}_2^*) d\mathfrak{r} + \int_{\Omega} |\mathcal{E}_2 - \mathcal{E}_2^*| \left| \frac{\mathcal{E}_1 \mathcal{E}_2}{1 + \mathfrak{q} \mathcal{E}_1^2} - \frac{\mathcal{E}_1^* \mathcal{E}_2^*}{1 + \mathfrak{q} \mathcal{E}_1^{*2}} \right| d\mathfrak{r} \\
&\leq \mathfrak{k}_1 \int_{\Omega} (\mathcal{E}_1 - \mathcal{E}_1^*) \Delta(\mathcal{E}_1 - \mathcal{E}_1^*) d\mathfrak{r} - \int_{\Omega} (\mathcal{E}_1 - \mathcal{E}_1^*)^2 d\mathfrak{r} \\
&\quad + \mathfrak{P} \int_{\Omega} |\mathcal{E}_1 - \mathcal{E}_1^*| (|\mathcal{E}_1 - \mathcal{E}_1^*| + |\mathcal{E}_2 - \mathcal{E}_2^*|) d\mathfrak{r} \\
&\quad + \mathfrak{k}_2 \int_{\Omega} (\mathcal{E}_2 - \mathcal{E}_2^*) \Delta(\mathcal{E}_2 - \mathcal{E}_2^*) d\mathfrak{r} + \mathfrak{P} \int_{\Omega} |\mathcal{E}_2 - \mathcal{E}_2^*| (|\mathcal{E}_1 - \mathcal{E}_1^*| + |\mathcal{E}_2 - \mathcal{E}_2^*|) d\mathfrak{r}.
\end{aligned} \tag{19}$$

Using Green's formula, we achieve:

$$\begin{aligned}
{}_0\mathcal{D}_3\mathcal{L}_2(\mathfrak{z}) &\leq -\mathfrak{k}_1 \int_{\Omega} |\nabla(\mathcal{E}_1 - \mathcal{E}_1^*)|^2 d\mathfrak{r} - \mathfrak{k}_2 \int_{\Omega} |\nabla(\mathcal{E}_2 - \mathcal{E}_2^*)|^2 d\mathfrak{r} - (1 - \mathfrak{P}) \int_{\Omega} (\mathcal{E}_1 - \mathcal{E}_1^*)^2 d\mathfrak{r} \\
&\quad + \mathfrak{P} \int_{\Omega} (\mathcal{E}_2 - \mathcal{E}_2^*)^2 d\mathfrak{r} + 2\mathfrak{P} \int_{\Omega} |\mathcal{E}_1 - \mathcal{E}_1^*| |\mathcal{E}_2 - \mathcal{E}_2^*| d\mathfrak{r} \\
&\leq -\mathfrak{k}_1 \int_{\Omega} |\nabla(\mathcal{E}_1 - \mathcal{E}_1^*)|^2 d\mathfrak{r} - \mathfrak{k}_2 \int_{\Omega} |\nabla(\mathcal{E}_2 - \mathcal{E}_2^*)|^2 d\mathfrak{r} - (1 - \mathfrak{P}) \int_{\Omega} (\mathcal{E}_1 - \mathcal{E}_1^*)^2 d\mathfrak{r} \\
&\quad + \mathfrak{P} \int_{\Omega} (\mathcal{E}_2 - \mathcal{E}_2^*)^2 d\mathfrak{r} + \mathfrak{P} \int_{\Omega} (|\mathcal{E}_1 - \mathcal{E}_1^*|^2 + |\mathcal{E}_2 - \mathcal{E}_2^*|^2) d\mathfrak{r} \\
&\leq -\mathfrak{k}_1 \int_{\Omega} |\nabla(\mathcal{E}_1 - \mathcal{E}_1^*)|^2 d\mathfrak{r} - \mathfrak{k}_2 \int_{\Omega} |\nabla(\mathcal{E}_2 - \mathcal{E}_2^*)|^2 d\mathfrak{r} - (1 - 2\mathfrak{P}) \int_{\Omega} (\mathcal{E}_1 - \mathcal{E}_1^*)^2 d\mathfrak{r} \\
&\quad + 2\mathfrak{P} \int_{\Omega} (\mathcal{E}_2 - \mathcal{E}_2^*)^2 d\mathfrak{r}.
\end{aligned} \tag{20}$$

Using Lemma 3, we obtain:

$$\begin{aligned}
{}_0\mathcal{D}_3\mathcal{L}_2(\mathfrak{z}) &\leq -(\mathfrak{k}_1 \zeta_1 - 2\mathfrak{P} + 1) \int_{\Omega} (\mathcal{E}_1 - \mathcal{E}_1^*)^2 d\mathfrak{r} - (\mathfrak{k}_2 \zeta_2 - 2\mathfrak{P}) \int_{\Omega} (\mathcal{E}_2 - \mathcal{E}_2^*)^2 d\mathfrak{r} \\
&\leq -2 \min\{\mathfrak{k}_1 \zeta_1 - 2\mathfrak{P} + 1, \mathfrak{k}_2 \zeta_2 - 2\mathfrak{P}\} \mathcal{L}_2(\mathfrak{z}).
\end{aligned} \tag{21}$$

Let $\beta_{3,1} = 2 \min\{\mathfrak{k}_1 \zeta_1 - 2\mathfrak{P} + 1, \mathfrak{k}_2 \zeta_2 - 2\mathfrak{P}\}$ be a strictly positive constant. Thus,

$$\int_0^\epsilon \frac{1}{\beta_{3,1}} d\varsigma = \frac{\epsilon}{2 \min\{\mathfrak{k}_1 \zeta_1 - 2\mathfrak{P} + 1, \mathfrak{k}_2 \zeta_2 - 2\mathfrak{P}\}} < +\infty, \tag{22}$$

The Lyapunov function $\mathcal{L}_2$ is decreasing over the interval $0 \leq \mathfrak{z} < \mathfrak{z}_1^* \leq \mathfrak{z}_{\max}$. This implies that $\mathcal{L}_2(\mathfrak{z}) > \mathcal{L}_2(\mathfrak{z}_1^*) > \mathcal{L}_2(\mathfrak{z}_{\max})$. Furthermore, we have:

$$\begin{aligned}
\mathcal{L}_2(\mathfrak{z}) &\leq \mathcal{L}_2(0) - 2 \min\{\mathfrak{k}_1 \zeta_1 - 2\mathfrak{P} + 1, \mathfrak{k}_2 \zeta_2 - 2\mathfrak{P}\} \int_0^{\mathfrak{z}} \mathcal{L}_2(s) d\varsigma \\
&\leq \mathcal{L}_2(0) - 2 \min\{\mathfrak{k}_1 \zeta_1 - 2\mathfrak{P} + 1, \mathfrak{k}_2 \zeta_2 - 2\mathfrak{P}\} \int_0^{\mathfrak{z}} \mathcal{L}_2(\mathfrak{z}_{\max}) d\varsigma \\
&= \mathcal{L}_2(0) - 2 \min\{\mathfrak{k}_1 \zeta_1 - 2\mathfrak{P} + 1, \mathfrak{k}_2 \zeta_2 - 2\mathfrak{P}\} \mathcal{L}_2(\mathfrak{z}_{\max}) \mathfrak{z}.
\end{aligned}$$

By applying Definitions 4 and 5, as $\mathfrak{z}$ approaches $\mathfrak{z}_1^*$, the function $\mathcal{L}_2(\mathfrak{z})$ tends to zero:

$$\lim_{\mathfrak{z} \rightarrow \mathfrak{z}_1^*} \mathcal{L}_2(\mathfrak{z}) \leq \mathcal{L}_2(0) - 2 \min\{\mathfrak{k}_1 \zeta_1 - 2\mathfrak{P} + 1, \mathfrak{k}_2 \zeta_2 - 2\mathfrak{P}\} \mathcal{L}_2(\mathfrak{z}_{\max}) \mathfrak{z}_1^* = 0. \quad (23)$$

The settling time can be defined as:

$$\mathfrak{z}_1^* = \frac{\mathcal{L}_2(0)}{2 \min\{\mathfrak{k}_1 \zeta_1 - 2\mathfrak{P} + 1, \mathfrak{k}_2 \zeta_2 - 2\mathfrak{P}\} \mathcal{L}_2(\mathfrak{z}_{\max})}.$$

Thus, $\mathcal{E}^* = \left(\mathfrak{a} - \mathfrak{b}, \mathfrak{b} \frac{1+q(\mathfrak{a}-\mathfrak{b})^2}{\mathfrak{a}-\mathfrak{b}} \right)$ represents a finite-time stable equilibrium point of the nonlinear systems (9), as stated in Theorem 1. $\square$

5. Finite-Time Synchronization Scheme

The objective of finite-time synchronization in reaction–diffusion systems is to achieve coordinated behavior among the elements of the system within a specified and limited time duration. This ensures that the systems synchronize within a defined timeframe, which is essential for various applications such as control system design, pattern generation, and synchronization in complex networks.

In this study, we explore the response of the Degrn–Harrison reaction–diffusion system, which is governed by a set of partial differential equations incorporating nonlinear reaction terms and diffusion processes. The system’s behavior is described by:

$$\begin{cases} {}_0^C \mathcal{D}_{\mathfrak{z}}^{\aleph} \bar{\mathcal{E}}_1(\mathfrak{r}, \mathfrak{z}) = \mathfrak{k}_1 \Delta \bar{\mathcal{E}}_1 + \mathfrak{a} - \bar{\mathcal{E}}_1 - \frac{\bar{\mathcal{E}}_1 \bar{\mathcal{E}}_2}{1+q\bar{\mathcal{E}}_1^2} + \mathcal{C}_1, & \mathfrak{r} \in \Omega, \mathfrak{z} > 0, \\ {}_0^C \mathcal{D}_{\mathfrak{z}}^{\aleph} \bar{\mathcal{E}}_2(\mathfrak{r}, \mathfrak{z}) = \mathfrak{k}_2 \Delta \bar{\mathcal{E}}_2 + \mathfrak{b} - \frac{\bar{\mathcal{E}}_1 \bar{\mathcal{E}}_2}{1+q\bar{\mathcal{E}}_1^2} + \mathcal{C}_2, & \mathfrak{r} \in \Omega, \mathfrak{z} > 0, \\ \partial_{\bar{\mathcal{E}}_1} = \partial_{\bar{\mathcal{E}}_2} = 0, & \mathfrak{r} \in \partial\Omega, \mathfrak{z} > 0, \\ \bar{\mathcal{E}}_1(\mathfrak{r}, 0) = \bar{\mathcal{E}}_{1,0}(\mathfrak{r}) > 0, \quad \bar{\mathcal{E}}_2(\mathfrak{r}, 0) = \bar{\mathcal{E}}_{2,0}(\mathfrak{r}) > 0, & \mathfrak{r} \in \Omega. \end{cases} \quad (24)$$

The synchronization discrepancies in Equations (9)–(24) are considered.

$$\mathfrak{e}(\mathfrak{r}, \mathfrak{z}) = \begin{pmatrix} \mathfrak{e}_1 \\ \mathfrak{e}_2 \end{pmatrix} = \begin{pmatrix} \bar{\mathcal{E}}_1 - \mathcal{E}_1 \\ \bar{\mathcal{E}}_2 - \mathcal{E}_2 \end{pmatrix}, \quad (25)$$

In the subsequent discussion, we will determine the linear controllers $\mathcal{C}_1$ and $\mathcal{C}_2$ that drive the error system solution towards zero as $\mathfrak{z}$ approaches $\mathfrak{z}^*$.

Definition 7 ([32,33]). If there is a specific time parameter $\mathfrak{z}^* > 0$ such that

$$\lim_{\mathfrak{z} \rightarrow \mathfrak{z}^*} \|\mathfrak{e}_1(\mathfrak{z})\| + \|\mathfrak{e}_2(\mathfrak{z})\| = 0, \quad (26)$$

and

$$\|\mathfrak{e}_1(\mathfrak{z})\| + \|\mathfrak{e}_2(\mathfrak{z})\| \equiv 0, \quad \forall \mathfrak{z} \geq \mathfrak{z}^*, \quad (27)$$

then the derived-response systems (9)–(24) are synchronized in finite time.

Theorem 4. The master–slave systems, denoted by Equations (9) and (24), attain stable and synchronized states within a finite duration by utilizing a two-dimensional linear control law

$$\begin{cases} \mathcal{C}_1 = (1 - 2\mathfrak{P})\mathfrak{e}_1, \\ \mathcal{C}_2 = -2\mathfrak{P}\mathfrak{e}_2, \end{cases} \quad (28)$$

and the time of synchronization is estimated as

$$\mathfrak{z}_2^* = \left(\frac{\Gamma(\aleph + 1)}{2 \min\{\mathfrak{k}_1 \zeta_1, \mathfrak{k}_2 \zeta_2\}} \right)^{\aleph-1}. \quad (29)$$

Proof. By integrating the control described in the theorem into the error system, we obtain the following dynamics:

$$\begin{cases} {}^C_0\mathcal{D}_3^\aleph \mathfrak{e}_1(\mathfrak{r}, \mathfrak{z}) = \mathfrak{k}_1 \Delta \mathfrak{e}_1 - 2\mathfrak{P} \mathfrak{e}_1 - \left(\frac{\bar{\mathcal{E}}_1 \bar{\mathcal{E}}_2}{1+q\bar{\mathcal{E}}_1^2} - \frac{\mathcal{E}_1 \mathcal{E}_2}{1+q\mathcal{E}_1^2} \right), & \mathfrak{r} \in \Omega, \mathfrak{z} > 0, \\ {}^C_0\mathcal{D}_3^\aleph \mathfrak{e}_2(\mathfrak{r}, \mathfrak{z}) = \mathfrak{k}_2 \Delta \mathfrak{e}_2 - 2\mathfrak{P} \mathfrak{e}_2 - \left(\frac{\bar{\mathcal{E}}_1 \bar{\mathcal{E}}_2}{1+q\bar{\mathcal{E}}_1^2} - \frac{\mathcal{E}_1 \mathcal{E}_2}{1+q\mathcal{E}_1^2} \right), & \mathfrak{r} \in \Omega, \mathfrak{z} > 0, \\ \partial_{\mathfrak{e}_1} = \partial_{\mathfrak{e}_2} = 0, & \mathfrak{r} \in \partial\Omega, \mathfrak{z} > 0, \\ \mathfrak{e}_1(\mathfrak{r}, 0) = \bar{\mathcal{E}}_{1,0}(\mathfrak{r}) - \mathcal{E}_{1,0}(\mathfrak{r}), \quad \mathfrak{e}_2(\mathfrak{r}, 0) = \bar{\mathcal{E}}_{2,0}(\mathfrak{r}) - \mathcal{E}_{2,0}(\mathfrak{r}), & \mathfrak{r} \in \Omega. \end{cases} \quad (30)$$

We employ Lyapunov's direct method in conjunction with a positive definite Lyapunov function.

$$\mathcal{L}_3(\mathfrak{z}) = \frac{1}{2} \int_{\Omega} (\mathfrak{e}_1^2 + \mathfrak{e}_2^2) d\mathfrak{r}, \quad (31)$$

Using Lemmas 1–3, we have:

$$\begin{aligned} {}^C_0\mathcal{D}_3^\aleph \mathcal{L}_3(\mathfrak{z}) &= \frac{1}{2} \int_{\Omega} {}^C_0\mathcal{D}_3^\aleph \mathfrak{e}_1^2 d\mathfrak{r} + \frac{1}{2} \int_{\Omega} {}^C_0\mathcal{D}_3^\aleph \mathfrak{e}_2^2 d\mathfrak{r} \\ &\leq \int_{\Omega} \mathfrak{e}_{10} \mathcal{D}_3^\aleph \mathfrak{e}_1 d\mathfrak{r} + \int_{\Omega} \mathfrak{e}_{20} \mathcal{D}_3^\aleph \mathfrak{e}_2 d\mathfrak{r} \\ &= \int_{\Omega} \mathfrak{e}_1 \left[\mathfrak{k}_1 \Delta \mathfrak{e}_1 - 2\mathfrak{P} \mathfrak{e}_1 - \left(\frac{\bar{\mathcal{E}}_1 \bar{\mathcal{E}}_2}{1+q\bar{\mathcal{E}}_1^2} - \frac{\mathcal{E}_1 \mathcal{E}_2}{1+q\mathcal{E}_1^2} \right) \right] d\mathfrak{r} \\ &\quad + \int_{\Omega} \mathfrak{e}_2 \left[\mathfrak{k}_2 \Delta \mathfrak{e}_2 - 2\mathfrak{P} \mathfrak{e}_2 - \left(\frac{\bar{\mathcal{E}}_1 \bar{\mathcal{E}}_2}{1+q\bar{\mathcal{E}}_1^2} - \frac{\mathcal{E}_1 \mathcal{E}_2}{1+q\mathcal{E}_1^2} \right) \right] d\mathfrak{r} \\ &\leq \mathfrak{k}_1 \int_{\Omega} \mathfrak{e}_1 \Delta \mathfrak{e}_1 d\mathfrak{r} + \mathfrak{k}_2 \int_{\Omega} \mathfrak{e}_2 \Delta \mathfrak{e}_2 d\mathfrak{r} - 2\mathfrak{P} \int_{\Omega} \mathfrak{e}_1^2 d\mathfrak{r} - 2\mathfrak{P} \int_{\Omega} \mathfrak{e}_2^2 d\mathfrak{r} \\ &\quad + \int_{\Omega} (|\mathfrak{e}_1| + |\mathfrak{e}_2|) \left| \frac{\bar{\mathcal{E}}_1 \bar{\mathcal{E}}_2}{1+q\bar{\mathcal{E}}_1^2} - \frac{\mathcal{E}_1 \mathcal{E}_2}{1+q\mathcal{E}_1^2} \right| d\mathfrak{r} \\ &\leq -\mathfrak{k}_1 \int_{\Omega} |\nabla \mathfrak{e}_1|^2 d\mathfrak{r} - \mathfrak{k}_2 \int_{\Omega} |\nabla \mathfrak{e}_2|^2 d\mathfrak{r} - 2\mathfrak{P} \int_{\Omega} (\mathfrak{e}_1^2 + \mathfrak{e}_2^2) d\mathfrak{r} \\ &\quad + \mathfrak{P} \int_{\Omega} (|\mathfrak{e}_1| + |\mathfrak{e}_2|) (|\bar{\mathcal{E}}_1 - \mathcal{E}_1| + |\bar{\mathcal{E}}_2 - \mathcal{E}_2|) d\mathfrak{r} \\ &\leq -\mathfrak{k}_1 \int_{\Omega} |\nabla \mathfrak{e}_1|^2 d\mathfrak{r} - \mathfrak{k}_2 \int_{\Omega} |\nabla \mathfrak{e}_2|^2 d\mathfrak{r} - 2\mathfrak{P} \int_{\Omega} (\mathfrak{e}_1^2 + \mathfrak{e}_2^2) d\mathfrak{r} + 2\mathfrak{P} \int_{\Omega} (|\mathfrak{e}_1|^2 + |\mathfrak{e}_2|^2) d\mathfrak{r} \\ &\leq -\mathfrak{k}_1 \zeta_1 \int_{\Omega} \mathfrak{e}_1^2 d\mathfrak{r} - \mathfrak{k}_2 \zeta_2 \int_{\Omega} \mathfrak{e}_2^2 d\mathfrak{r} \\ &= -2 \min\{\mathfrak{k}_1 \zeta_1, \mathfrak{k}_2 \zeta_2\} \mathcal{L}_3(\mathfrak{z}). \end{aligned} \quad (32)$$

As the parameter $\aleph$ approaches 1 from the left, the expression ${}^C_0\mathcal{D}_3^\aleph \mathcal{L}_3(\mathfrak{z})$ converges to ${}_0\mathcal{D}_3 \mathcal{L}_3(\mathfrak{z})$ [34]. This convergence implies

$${}_0\mathcal{D}_3 \mathcal{L}_3(\mathfrak{z}) \leq -2 \min\{\mathfrak{k}_1 \zeta_1, \mathfrak{k}_2 \zeta_2\} \mathcal{L}_3(\mathfrak{z}), \quad (33)$$

where $\beta_{3,2} = 2 \min\{\mathfrak{k}_1 \zeta_1, \mathfrak{k}_2 \zeta_2\}$ is a strictly positive constant. Consequently, we can derive

$$\int_0^\epsilon \frac{1}{\beta_{3,2}} d\varsigma = \frac{\epsilon}{2 \min\{\mathfrak{k}_1 \zeta_1, \mathfrak{k}_2 \zeta_2\}} < +\infty, \quad (34)$$

Using Theorem 1, we establish that the zero solution of the error system (30) indicates finite-time stability of the equilibrium point $(\mathfrak{e}_1, \mathfrak{e}_2) = (0, 0)$. Moreover, employing Theorem 2, we infer:

$$\mathcal{L}_3(\mathfrak{z}) \leq \mathcal{L}_3(0) E_{\aleph} \left(-2 \min\{\mathfrak{k}_1 \zeta_1, \mathfrak{k}_2 \zeta_2\} \mathfrak{z}^{\aleph} \right), \quad \forall \mathfrak{z} \in [0, \mathfrak{z}_2^*], \quad (35)$$

with the settling-time

$$\mathfrak{z}_2^* = \left(\frac{\Gamma(\aleph + 1)}{2 \min\{\mathfrak{k}_1 \zeta_1, \mathfrak{k}_2 \zeta_2\}} \right)^{\aleph^{-1}},$$

when $\mathfrak{z} \geq \mathfrak{z}_2^* = \left(\frac{\Gamma(\aleph + 1)}{2 \min\{\mathfrak{k}_1 \zeta_1, \mathfrak{k}_2 \zeta_2\}} \right)^{\aleph^{-1}}$, we can conclude that

$$\mathcal{L}_3(\mathfrak{z}) \equiv 0, \quad (36)$$

Consequently, the master–slave systems, denoted by Equations (9) and (24), attain stable and synchronized states within a finite duration according to Definition 7. $\square$

6. Numerical Examples with Simulations

In this section, we present a diverse set of examples to illustrate the principles of finite-time stability and finite-time synchronization within the Degn–Harrison system. These examples explore various scenarios, including different initial conditions, parameter values, and control strategies, offering insights into the system's behavior. Through detailed analysis and numerical simulations, we aim to provide a comprehensive understanding of how finite-time stability and synchronization manifest in practical contexts, shedding light on the dynamics of the Degn–Harrison system.

The numerical simulations in these examples are conducted using the Finite Difference method, a crucial tool for analyzing fractional-order Degn–Harrison reaction–diffusion systems. This method discretizes partial differential equations into algebraic equations, enabling precise numerical approximation and detailed modeling of system dynamics. By facilitating simulations of complex phenomena, stability analysis, and parameter optimization, the Finite Difference method enhances our understanding of chemical and biological systems.

Example 1. We define the domain $\Omega = [0, 25]$ and the parameter $\mathfrak{z}$ ranging from $[0, 100]$. The parameter values are selected as follows:

Variable	Value
$\mathfrak{k}_1$	1
$\mathfrak{k}_2$	1
$\mathfrak{a}$	0.78
$\mathfrak{b}$	0.01275
$\mathfrak{q}$	0.75
$\mathfrak{P}$	0.5774
ζ_1	10^4
ζ_2	10^4
$\aleph$	0.99
$\aleph$	150

with initial conditions:

$$\begin{cases} \mathcal{E}_{1,0}(\mathfrak{r}) = 0.5(1 - \cos(5\mathfrak{r})), \\ \mathcal{E}_{2,0}(\mathfrak{r}) = 0.15(1 - \sin(4\mathfrak{r})). \end{cases} \quad (37)$$

The conditions of Theorem 3 are met:

$$\begin{aligned} |\mathcal{E}_2| < \aleph = 0.3, \quad \mathfrak{P} \geq \max \left\{ \frac{5}{4} \aleph, \frac{1}{2\sqrt{\mathfrak{q}}} \right\} &= 0.5774, \\ \min\{\mathfrak{k}_1 \zeta_1 - 2\mathfrak{P} + 1, \mathfrak{k}_2 \zeta_2 - 2\mathfrak{P}\} &= 9.9988 \times 10^3. \end{aligned} \quad (38)$$

Upon calculation, we obtain:

$$\begin{cases} \mathcal{L}_2(0) = 16.8056, \\ \mathcal{L}_2(100) = 8.4449 \times 10^{-6}, \end{cases} \quad (39)$$

The time of finite-time stability of the equilibrium point is estimated as:

$$\mathfrak{z}_1^* = \frac{\mathcal{L}_2(0)}{2 \min\{\mathfrak{k}_1 \zeta_1 - 2\mathfrak{P} + 1, \mathfrak{k}_2 \zeta_2 - 2\mathfrak{P}\} \mathcal{L}_2(100)} = 99.5134s, \quad (40)$$

where

$$\mathcal{L}_2(99.5082) = 8.6830 \times 10^{-6}. \quad (41)$$

The figures presented in Figures 1 and 2 depict the spatiotemporal solutions of system (9) with homogeneous Neumann boundary conditions. According to Theorem 3, these solutions exemplify finite-time stability at the equilibrium point $\mathcal{E}^* = (0.7672, 0.0240)$. To validate this assertion numerically, Figure 3 showcases the system's spatiotemporal solutions alongside the corresponding errors, focusing specifically on one-dimensional space. This visual representation compellingly illustrates that the Lyapunov function converge to zero as $\mathfrak{z}$ approaches $\mathfrak{z}_1^* = 99.5134s$ (as depicted in Figure 4).

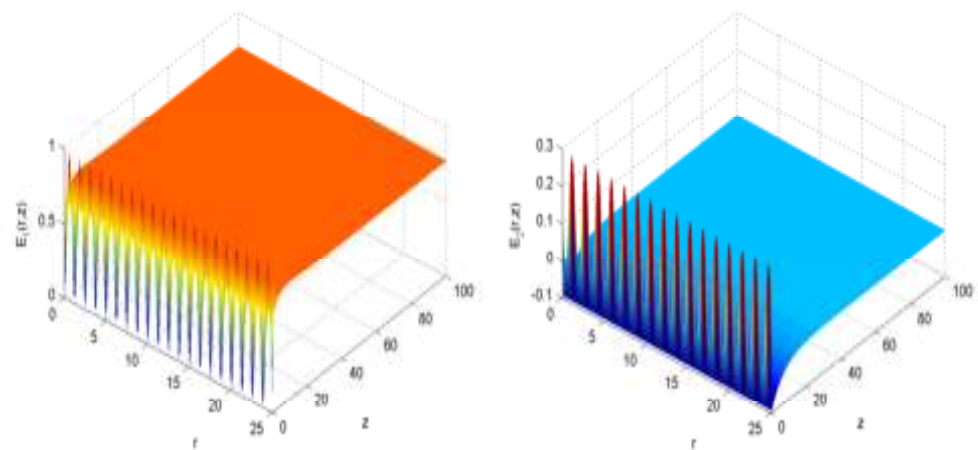


Figure 1. Dynamic behavior of the system (9): $\mathcal{E}_1(\tau, \mathfrak{z})$ and $\mathcal{E}_2(\tau, \mathfrak{z})$.

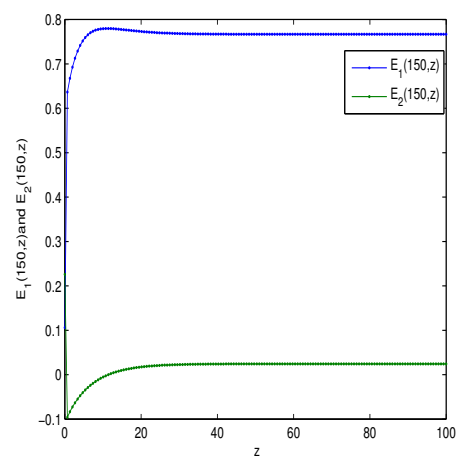


Figure 2. Solutions of the system (9): $\mathcal{E}_1(150, \mathfrak{z})$, $\mathcal{E}_2(150, \mathfrak{z})$.

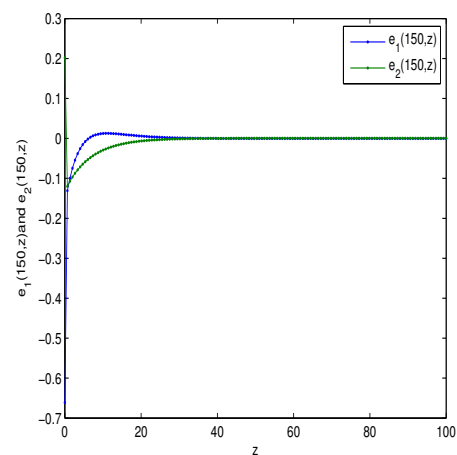


Figure 3. Solutions of the errors system (9): $e_1(150, z)$, $e_2(150, z)$.

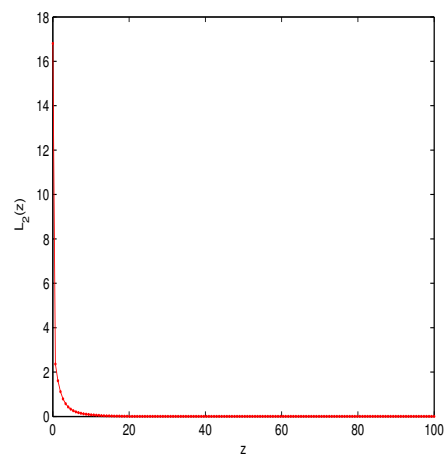


Figure 4. Estimation of the Lyapunov function $\mathcal{L}_2(z)$.

Example 2. The parameters are specified as follows:

Parameter	Value
t_1	0.1885
t_2	0.1885
a	3
b	1.25
q	1
N	0.9
$\mathfrak{P}$	3.7
ζ_1	0.6
ζ_2	0.6
$\mathbb{N}$	50

with the domain $\Omega = [0, 10]$ and $z \in [0, 5]$. The initial conditions for the master–slave systems (9) and (24) are established as follows:

$$\begin{cases} \mathcal{E}_{1,0}(\tau) = 1.5, & \bar{\mathcal{E}}_{1,0}(\tau) = 1, \\ \mathcal{E}_{2,0}(\tau) = 1.5, & \bar{\mathcal{E}}_{2,0}(\tau) = 1, \end{cases} \quad (42)$$

We aim to formulate two controllers, denoted as $\mathcal{C}_1$ and $\mathcal{C}_2$, as described below:

$$\begin{cases} \mathcal{C}_1 = -6.4e_1, \\ \mathcal{C}_2 = -7.4e_2, \end{cases} \quad (43)$$

we have $\mathcal{L}_3(0) = 12.5$ and

$$\mathcal{L}_3(\mathfrak{z}) \leq M(\mathfrak{z}) = 12.5E_{0.9}\left(-0.2262\mathfrak{z}^{0.9}\right), \quad \forall \mathfrak{z} \in [0, \mathfrak{z}_2^*), \quad (44)$$

The time of synchronization is estimated as:

$$\mathfrak{z}_2^* = \left(\frac{\Gamma(\aleph + 1)}{2 \min\{\mathfrak{k}_1\zeta_1, \mathfrak{k}_2\zeta_2\}} \right)^{\aleph^{-1}} = 4.9937 \text{ s}, \quad (45)$$

and

$$\mathcal{L}_3(5) \approx \mathcal{L}_3(4.9937) = 4.6605 \times 10^{-6}, \quad (46)$$

where $4.9937 \leq \mathfrak{z} \leq 5$. We can conclude that

$$\mathcal{L}_3(\mathfrak{z}) \equiv 0. \quad (47)$$

Figures 5–7 showcase the spatiotemporal dynamics of systems (9) and (24), providing insights into both two-dimensional and one-dimensional spaces. To furnish numerical evidence, we present the spatiotemporal solutions of the error synchronization system (30) in Figures 7 and 8. This evolution clearly demonstrates that the errors converge to zero as $\mathfrak{z}$ approaches $\mathfrak{z}_2^* = 4.9937 \text{ s}$, thereby illustrating the finite-time behavior of the system. Furthermore, Figure 9 exemplifies this finite-time convergence within the specified duration $\mathfrak{z}_2^* = 4.9937 \text{ s}$.

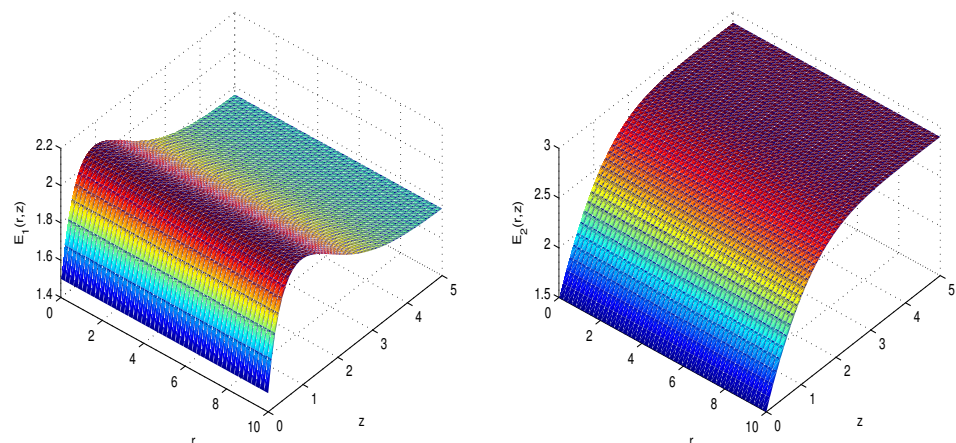


Figure 5. Dynamic behavior of the system (9): $\mathcal{E}_1(\tau, \mathfrak{z})$ and $\mathcal{E}_2(\tau, \mathfrak{z})$.

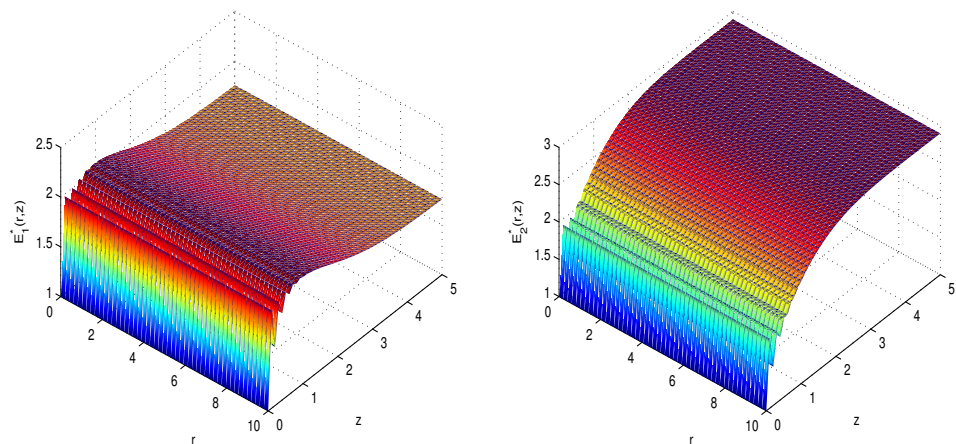


Figure 6. Dynamic behavior of the system (24): $\bar{\mathcal{E}}_1(\tau, \mathfrak{z})$ and $\bar{\mathcal{E}}_2(\tau, \mathfrak{z})$.

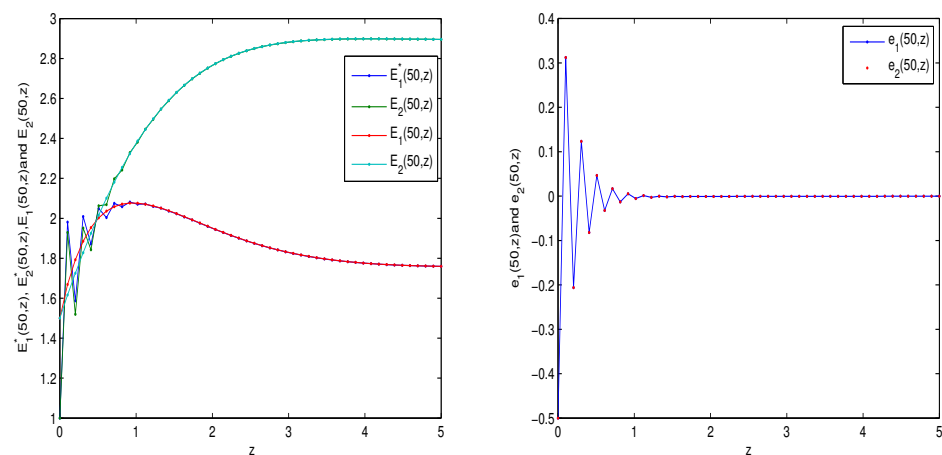


Figure 7. Solutions of the master–slave systems (9), (24), and the error system (30) at $\tau = 50$.

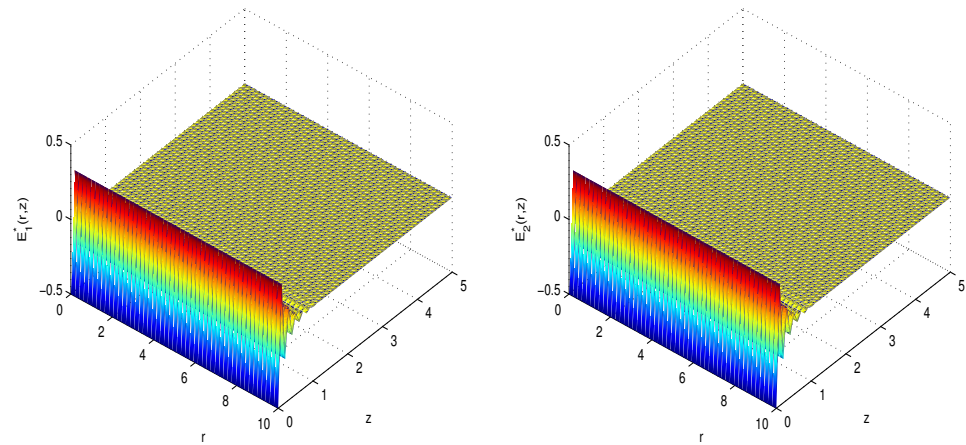


Figure 8. Dynamic behavior of the error system (30): $e_1(\tau, \mathfrak{z})$ and $e_2(\tau, \mathfrak{z})$.

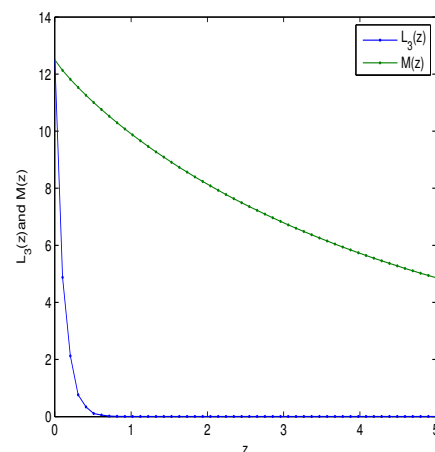


Figure 9. Estimation of the Lyapunov function $\mathcal{L}_3(\mathfrak{z})$ and $M(\mathfrak{z})$.

7. Conclusions

This study marks the beginning of an innovative exploration into the Degrn–Harrison reaction–diffusion systems, introducing novel variations that promise to reveal new layers of complexity within spatiotemporal models. Our research aims to pioneer methodologies for investigating both finite-time stability and finite-time synchronization within these models, providing a comprehensive understanding of their dynamics, particularly in the context of nonlinear bacterial colonies. To initiate our investigation, we utilize the

foundational principles of Lyapunov theory to establish robust frameworks for analyzing the finite-time stability of equilibrium points and synchronization phenomena. These theoretical foundations serve as the basis for our subsequent experimental endeavors. To empirically validate the effectiveness and reliability of our proposed methodologies, we conduct extensive computational experiments. These experiments involve thorough simulations of fractional-order Degrn–Harrison systems, designed to capture the dynamics and behaviors of bacterial colonies within reaction–diffusion frameworks.

As we move forward, our next phase of research will focus on extending and refining our theoretical frameworks, exploring new applications in biological systems beyond bacterial colonies, and enhancing our computational models for greater efficiency and accuracy. We plan to validate our findings through collaborative experimental work, fostering interdisciplinary collaborations to integrate diverse perspectives and expertise. Additionally, we will actively disseminate our results through various channels. Through these efforts, we aim to advance the field of reaction–diffusion systems and contribute to a deeper understanding of complex spatiotemporal phenomena. Our commitment to pushing the boundaries of knowledge in this dynamic area of research promises exciting discoveries and innovations in the near future.

Author Contributions: Conceptualization, M.A.H.; Formal analysis, I.B.; Investigation, A.A.; Methodology, W.G.A.; Resources, A.O. and A.H.; Software, S.M.; Validation, A.O.; Visualization, I.B. and M.A.H.; Writing—review and editing, A.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in this study are included in the article; further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Mesdoui, F.; Ouannas, A.; Shawagfeh, N.; Grassi, G.; Pham, V.T. Synchronization methods for the Degrn–Harrison reaction–diffusion systems. *IEEE Access* **2020**, *8*, 91829–91836. [\[CrossRef\]](#)
2. Ouannas, A.; Abdelli, M.; Odibat, Z.; Wang, X.; Pham, V.T.; Grassi, G.; Alsaedi, A. Synchronization control in reaction–diffusion systems: Application to Lengyel–Epstein system. *Complexity* **2019**, *2019*, 2832781. [\[CrossRef\]](#)
3. Ambrosio, B.; Aziz-Alaoui, M.A. Synchronization and control of coupled reaction–diffusion systems of the FitzHugh–Nagumo type. *Comput. Math. Appl.* **2012**, *64*, 934–943. [\[CrossRef\]](#)
4. Turkyilmazoglu, M.; Alofi, A.S. Liquid Vortex Formation in a Swirling Container Considering Fractional Time Derivative of Caputo. *Fractal Fract.* **2024**, *8*, 231. [\[CrossRef\]](#)
5. Ouannas, A.; Wang, X.; Pham, V.T.; Grassi, G.; Huynh, V.V. Synchronization results for a class of fractional-order spatiotemporal partial differential systems based on fractional Lyapunov approach. *Bound. Value Probl.* **2019**, *2019*, 74. [\[CrossRef\]](#)
6. Wang, B.; Ouannas, A.; Karaca, Y.; Xia, W.F.; Jahanshahi, H.; Alkhateeb, A.F.; Nour, M. A Hybrid Approach For Synchronizing Between Two Reaction Diffusion Systems of Integer-And Fractional-Order Applied On Certain Chemical Models. *Fractals* **2022**, *30*, 2240145. [\[CrossRef\]](#)
7. Srivastava, M.; Ansari, S.P.; Agrawal, S.K.; Das, S.; Leung, A.Y.T. Anti-synchronization between identical and non identical fractional-order chaotic systems using active control method. *Nonlinear Dyn.* **2014**, *76*, 905–914. [\[CrossRef\]](#)
8. Jahanshahi, H.; Sajjadi, S.S.; Bekiros, S.; Aly, A.A. On the development of variable-order fractional hyper chaotic economic system with a nonlinear model predictive controller. *Chaos Solitons Fractals* **2021**, *144*, 110698. [\[CrossRef\]](#)
9. Khennaoui, A.A.; Pham, V.T.; Thoai, V.P.; Ouannas, A.; Grassi, G.; Momani, S. From Lozi map to fractional memristive Lozi map. *Eur. Phys. J. Spec. Top.* **2023**, *232*, 2385–2393. [\[CrossRef\]](#)
10. Khennaoui, A.A.; Ouannas, A.; Boulaaras, S.; Pham, V.T.; Taher Azar, A. A fractional map with hidden attractors: Chaos and control. *Eur. Phys. J. Spec. Top.* **2020**, *229*, 1083–1093. [\[CrossRef\]](#)
11. Ouannas, A.; Khennaoui, A.A.; Wang, X.; Pham, V.T.; Boulaaras, S.; Momani, S. Bifurcation and chaos in the fractional form of Hénon–Lozi type map. *Eur. Phys. J. Spec. Top.* **2020**, *229*, 2261–2273. [\[CrossRef\]](#)
12. Wu, Y.; Huang, T. Finite-time synchronization of chaotic systems with time-varying delays and parameter mismatches. *Nonlinear Dyn.* **2024**, *77*, 729–736.
13. Hu, J.; Wu, X. Finite-time synchronization of fractional-order chaotic systems. *Commun. Nonlinear Sci. Numer. Simul.* **2018**, *59*, 440–452.

14. Hamadneh, T.; Hioual, A.; Saadeh, R.; Abdoon, M.A.; Almutairi, D.K.; Khalid, T.A.; Ouannas, A. General methods to synchronize fractional discrete reaction–diffusion systems applied to the glycolysis model. *Fractal Fract.* **2023**, *7*, 828. [\[CrossRef\]](#)
15. Hamadneh, T.; Hioual, A.; Alsayyed, O.; Al-Khassawneh, Y.A.; Al-Husban, A.; Ouannas, A. The FitzHugh–Nagumo Model Described by Fractional Difference Equations: Stability and Numerical Simulation. *Axioms* **2023**, *12*, 806. [\[CrossRef\]](#)
16. Garvie, M.R. Finite-difference schemes for reaction–diffusion equations modeling predator–prey interactions in M ATLAB. *Bull. Math. Biol.* **2007**, *69*, 931–956. [\[CrossRef\]](#)
17. Houston, P.; Schwab, C.; Süli, E. Discontinuous hp-finite element methods for advection–diffusion–reaction problems. *SIAM J. Numer. Anal.* **2002**, *39*, 2133–2163. [\[CrossRef\]](#)
18. Bueno-Orovio, A.; Kay, D.; Burrage, K. Fourier spectral methods for fractional-in-space reaction–diffusion equations. *BIT Numer. Math.* **2014**, *54*, 937–954. [\[CrossRef\]](#)
19. Mesdoui, F.; Shawagfeh, N.; Ouannas, A. Global synchronization of fractional-order and integer-order N component reaction diffusion systems: Application to biochemical models. *Math. Methods Appl. Sci.* **2021**, *44*, 1003–1012. [\[CrossRef\]](#)
20. Song, G.; Zhang, X.; Wang, Z.; Chen, Y.; Chen, P. The asymptotic local finite-difference method of the fractional wave equation and its viscous seismic wavefield simulation. *Geophysics* **2020**, *85*, T179–T189. [\[CrossRef\]](#)
21. Al-Husban, A.; Djenina, N.; Saadeh, R.; Ouannas, A.; Grassi, G. A New Incommensurate Fractional-Order COVID 19: Modelling and Dynamical Analysis. *Mathematics* **2023**, *11*, 555. [\[CrossRef\]](#)
22. Samko, S.G.; Kilbas, A.A.; Marichev, O.I. *Fractional Integrals and Derivatives: Theory and Applications*; Gordon and Breach Science Publishers: Basel, Switzerland, 1993.
23. Feng, Z.; Xiang, Z. Finite-time stability of fractional-order nonlinear systems. *Chaos* **2024**, *34*, 023105. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Ouannas, A.; Abdelmalek, S.; Bendoukha, S. *Coexistence of Some Chaos Synchronization Types in Fractional-Order Differential Equations*; Texas State University, Department of Mathematics: San Marcos, TX, USA, 2017.
25. Aguila-Camacho, N.; Duarte-Mermoud, M.A.; Gallegos, J.A. Lyapunov functions for fractional order systems. *Commun. Nonlinear Sci. Numer. Simul.* **2014**, *19*, 2951–2957. [\[CrossRef\]](#)
26. Yaagoub, Z.; Danane, J.; Hammouch, Z.; Allali, K. Mathematical analysis of a fractional order two strain SEIR epidemic model. *Results Nonlinear Anal.* **2024**, *7*, 156–175.
27. Jiang, J.; Li, H.; Zhao, K.; Cao, D.; Guirao, L.G. Finite time stability and sliding mode control for uncertain variable fractional order nonlinear systems. *Adv. Differ. Equ.* **2021**, *2021*, 127. [\[CrossRef\]](#)
28. Fairén V.; Velarde, M.G. Time-periodic oscillations in a model for the respiratory process of a bacterial culture. *J. Math. Biol.* **1979**, *8*, 147–157. [\[CrossRef\]](#)
29. Lisena, B. Some global results for the Degrn–Harrison system with diffusion. *Medit. J. Math.* **2017**, *14*, 91. [\[CrossRef\]](#)
30. Abu Falahah, I.; Hioual, A.; Omar Al-Qadri, M.; AL-Khassawneh, Y.A.; Al-Husban, A.; Hamadneh, T.; Ouannas, A. Synchronization of Fractional Partial Difference Equations via Linear Methods. *Axioms* **2023**, *12*, 728. [\[CrossRef\]](#)
31. Pan, J.; Liu, X.; Zhong, S. Stability criteria for impulsive reaction–diffusion Cohen–Grossberg neural networks with time-varying delays. *Math. Comput. Model.* **2010**, *51*, 1037–1050. [\[CrossRef\]](#)
32. Wang, L.; Yang, X.; Liu, H.; Chen, X. Synchronization in Finite Time of Fractional-Order Complex-Valued Delayed Gene Regulatory Networks. *Fractal Fract.* **2023**, *7*, 347. [\[CrossRef\]](#)
33. Li, Y.; Yang, X.; Shi, L. Finite-time synchronization for competitive neural networks with mixed delays and non-identical perturbations. *Neurocomputing* **2016**, *185*, 242–253. [\[CrossRef\]](#)
34. Podlubny, I. *Fractional Differential Equations: An Introduction to Fractional Derivatives, Fractional Differential Equations, to Methods of Their Solution and Some of Their Applications*; Elsevier: Amsterdam, The Netherlands, 1998.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Molecular Dynamics Simulation of Melting of the DNA Duplex with Silver-Mediated Cytosine–Cytosine Base Pair

Elena B. Gusarova and Natalya A. Kovaleva *

N.N. Semenov Federal Research Center for Chemical Physics, Russian Academy of Sciences, 4 Kosygin Street, Moscow 119991, Russia

* Correspondence: natykov@gmail.com

Abstract: Metal-mediated base pairs in DNA double helix molecules open up broad opportunities for biosensors based on DNA clusters with silver due to their low toxicity and applicability in drug design. Despite intensive experimental and computational research, molecular mechanisms of stabilization of a double helix by silver-mediated base pairs are mainly unknown. We conducted all-atom molecular dynamics simulations of a dodecameric DNA double helix (sequence 5'-TAGGTCAATACT-3'-3'ATCCACTTATGA-5') with either cytosine–cytosine or cytosine–Ag⁺–cytosine mismatch in the center of the duplex. We extended the previously proposed set of interaction parameters for a silver ion in the silver-mediated pair in order to allow for its dissociation. With this new potential, we studied how the addition of a silver ion could stabilize a DNA double helix containing a single cytosine–cytosine mismatch. In particular, we found out that the helix with cytosine–Ag⁺–cytosine mismatch has a greater melting temperature than the helix with cytosine–cytosine one. This stabilization effect of the silver ion is in qualitative agreement with experimental data. The central region of the duplex with cytosine–Ag⁺–cytosine mismatch (unlike with cytosine–cytosine mismatch) is stable enough to prevent bubble formation at moderate temperatures during melting. The results of this simulation can be used to devise novel metal-mediated DNA structures.

Keywords: metal-base pair; silver-mediated DNA; melting temperature; molecular dynamics

Citation: Gusarova, E.B.; Kovaleva, N.A. Molecular Dynamics Simulation of Melting of the DNA Duplex with Silver-Mediated Cytosine–Cytosine Base Pair. *Computation* **2024**, *12*, 145. <https://doi.org/10.3390/computation12070145>

Academic Editor: Claudio Amovilli

Received: 2 May 2024

Revised: 4 July 2024

Accepted: 10 July 2024

Published: 12 July 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

DNA double helix is often used in nanotechnology for a wide range of applications. One promising example is the extensive search for alternatives for hydrogen bonds (H-bonds) between complementary nitrogenous bases. An attractive approach is to create a base pair with metal cations that have a high affinity for nitrogenous bases. This approach is promising for the creation of DNA-based nanomaterials with improved thermal, mechanical and chemical properties, as well as with new functionalities [1–4]. A DNA double helix in which nucleotide bases are paired by a metal ion instead of hydrogen bonds in Watson–Crick base pairs is called metalized DNA (M-DNA) [5]. Such systems are of interest due to their potential use as sensitive and specific metal sensors and biosensors [6,7]. As a rule, the inclusion of non-canonical base pairs in homochiral or heterochiral DNA leads to a decrease in the double helix stability [8]. Such decrease depends to a large extent on the specific structure of the non-canonical base pairs, their nature, and the sequence in which they are included. In 2009, Ono and Miyake [9] described two structures with metal-mediated pairs, in particular, cytosine–Ag⁺–cytosine (C:Ag:C). Their research has opened wide opportunities for DNA nanotechnology, nanoelectronics and DNA-origami due to the high specificity of silver binding [10,11]. Silver ions have attracted special attention due to their low toxicity and applicability in drug design [12,13]. As a result, numerous studies on artificial DNA containing non-canonical base pairs with silver ions were reported [14–17]. In natural DNA, silver-mediated pairing has been explored in canonical double helices with base pair mismatches [9] and in strands without canonical pairs [18–21]. High chemical

and thermal stability of the C:Ag:C base pair was demonstrated [22–26], and the role of pH in the stability of the C:Ag:C pair was studied [27]. The effect of silver on the melting temperature of DNA double helix with cytosine–cytosine (C:C) inclusions as a complementary pair has been experimentally observed [28,29]. The melting point of an oligonucleotide with C:C mismatch is increased in the presence of silver ions in solution.

To date, a number of theoretical models of base-pairing dynamics in DNA of various degrees of detailization have been developed. These models stimulate significant progress in the field of DNA melting [30,31]. The proposed models vary from the mesoscopic level (models based on the Peyrard–Bishop (PB) model [32,33]) to the coarse-grained model (CG) with several beads per nucleotide [34]. Mesoscopic models such as the PB model and its variations use a set of effective potentials to describe the main interactions in the DNA double helix, namely the hydrogen bonds and stacking interactions [35–38]. Using the mesoscopic models, the influence of single, double, and triple mismatches on duplex stability was investigated. It was found that many combinations of multiple mismatches are surprisingly stable [39]. Melting points of sequences containing metal-mediated C:C and thymine–thymine pairs were calculated [40]. The use of a CG model, where only monovalent and divalent ions are included, allows one to reproduce experimental melting temperatures for oligonucleotides [41]. In many mesoscopic and CG models, the effective non-linear potential, Lennard–Jones potential [42–44] or Morse potential [45–47] are used to mimic hydrogen bonding to describe the melting processes.

DNA melting has also been studied by means of all-atom molecular dynamics (MD) simulation [48–50]. However, the number of all-atom MD simulation studies of DNA with metal-mediated pairs is very limited. The shortest possible Ag^+ -mediated cytosine homobase duplex, a tetramer formed by Ag^+ bridges between two DNA strands, has previously been studied [16]. The Lopez-Acevedo group reported a study of Ag^+ -paired cytosine homobase DNA strands using a combination of quantum and classical MD calculations [51]. They showed that a parallel double helix has high stability, whereas an antiparallel double helix untwists and loses its helical structure. Cytosines are paired in a transoid conformation in a parallel double helix, while in an antiparallel double helix, they are in a cisoid conformation (see Figure 1a,b). The bond in the C:Ag:C pair in the cisoid conformation is weaker than in the transoid one due to an additional hydrogen bond formed between the oxygen of one cytosine and the amino group of the complementary cytosine [51–53]. The approach developed in the works of Lopez-Acevedo et al. [51] was successfully used in the MD simulation of silver-mediated base pairing in parallel-stranded DNA [29]. However, the MD simulation of the destruction of the C:Ag:C pair is challenging to date. The potentials are usually used to describe the bond of a silver ion with the nitrogen of cytosine as a covalent bond. MD's description of a covalent bond is based on a harmonic potential that does not provide a possibility for bond breaking. Therefore, a different potential has to be used to describe the destruction of the C:Ag:C pair. One possibility is to use the Morse potential, which explicitly includes the effects of bond breaking and the existence of unbound states.

In this article, we propose to use the Morse potential to describe the bond of a silver ion with cytosine by analogy with a hydrogen bond in some mesoscopic models. We simulated an antiparallel double-stranded DNA helix with C:Ag:C mismatch. We fitted parameters of the Morse potential that reproduced the process of melting a pair mediated by silver. We obtained melting curves and estimated melting temperatures. We believe that melting temperature studies of M-DNA will certainly enhance our knowledge about DNA double helix properties associated with metal-mediated base pairs.

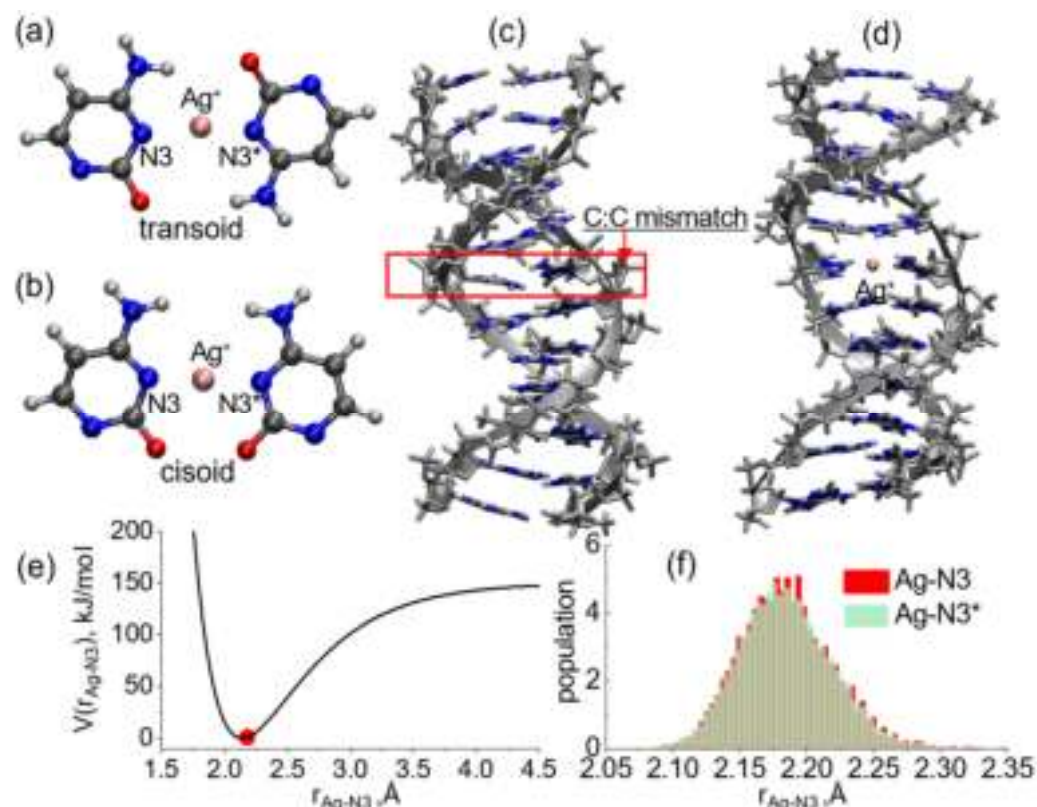


Figure 1. Structural variations for the C:Ag:C base pair: (a) transoid and (b) cisoid conformations. N3* defines the N3 atom in complementary base. DNA double helix: (c) with pure C:C mismatch and (d) with C:Ag:C formed during the classical MD. A snapshot of the double helix at $t = 100$ ns is shown. (e) The fitted Morse potential curve. The red point corresponds to the equilibrium value r_{Ag-N3} in the MD simulation at a temperature of 300 K. (f) The histogram (population in percentage) of the distance between silver ion and cytosine nitrogen r_{Ag-N3} at temperature 300 K. Histogram was constructed using 75 bins between 2.05 and 2.35.

2. Materials and Methods

Double-stranded DNA helix (12 base pairs) with the sequence 5'-TAGGTCAAATACT-3'-3'-ATCCACTTATGA-5' with a mismatch of complementarity in the sixth position, as C:C was considered. The initial structure was generated using the Avogadro program [54].

Initially, the DNA double helix was placed in a cubic box so that the minimum distance between the helix and the edges of the box exceeded 1 nm. Then, water molecules were added to the system. We used the TIP3P water model. To neutralize the charge of the system, 22 Na⁺ ions were added to the calculation cell. Calculations were carried out in a periodic cell using stochastic dynamics as a thermostat. Equations of motion were integrated using a leap-frog version of the Verlet algorithm. The integration step was 2 fs. Constant temperature was maintained using a stochastic thermostat, and constant pressure was maintained using a Parrinello–Rahman barostat with a time constant of 2 ps. Bonds with hydrogen atoms were constrained using the LINCS algorithm. The cutoff radii for Coulomb and van der Waals interactions were taken to be 15 Å. PME method was used to consider long-range Coulomb interactions. For MD simulation, the Gromacs 2023.1 software package [55] was used.

The parameters of potentials are a combination of the Amber99-parmbsc1 force field [56] and original parameters for Ag⁺ and cytosine interactions. Parameters of interaction between cytosine and a silver ion were taken from the literature, with some modifications as follows. Partial charges of atoms on cytosine interacting with silver ions have been corrected according to the paper [51]. The 12-6 Lennard-Jones (LJ) model was used for van der Waals interactions. The values of non-bonded interaction parameters for

the LJ potential for Ag^+ ($\sigma = 2.731 \text{ \AA}$ and $\varepsilon = 0.16316922 \text{ kJ/mol}$) and the partial charge of a silver ion ($q_{\text{Ag}^+} = 0.38272$) were taken from [57], which previously were validated for silver-mediated base pairing in double-stranded DNA in [29,51].

To simulate the bond dissociation of the silver ion from the N3 atom of cytosine, we chose the Morse potential. The Morse potential explicitly includes bond-breaking effects such as the existence of unbound states:

$$V(r) = D_e(1 - e^{-\alpha(r-r_e)})^2, \quad (1)$$

where r is the distance between N3 and Ag atoms, r_e is the equilibrium bond distance, D_e is the depth of the well, and α controls the “width” of the potential. The equilibrium bond distance $r_e = 2.14 \text{ \AA}$ is known from quantum mechanical calculations and has been successfully used in MD simulations of both transoid and cisoid pairs C:Ag:C [29,51]. The parameters of the Morse potential α and D_e were obtained as a result of an optimization procedure described below.

The optimization procedure of the Morse potential parameters consisted of two steps. At first, we determined the value of the parameter α . The idea is to keep the geometry of the C:Ag:C pair determined from quantum mechanical calculations [51], in particular, by fitting the equilibrium value of the angle N3-Ag⁺-N3* to one derived by the DFT method. To carry out this procedure, we set the parameter $D_e = 200 \text{ kJ/mol}$, which exceeds the energy of the hydrogen bond. Such a large value of D_e is required to prevent the C:Ag:C pair from spontaneous dissociation upon thermal oscillations. The initial value of parameter α was chosen so that the “width” of the Morse potential was similar to the “width” of the harmonic potential from [51].

Authors of ref. [51] introduced harmonic bonded potential as well as angular and dihedral ones in order to keep the angle N3-Ag⁺-N3* close to quantum chemical values. We used Morse potential rather than these ones and, therefore, had to optimize the “width” of the potential well.

We selected a range of values of parameter α . For each value α , we conducted an MD simulation at a temperature of 300 K with a length of 10 ns and determined the equilibrium value of the valent angle N3-Ag⁺-N3*. It turned out that $\alpha = 2.0 \text{ \AA}^{-1}$ gives the closest to quantum chemical values [51].

The next step was the parameter D_e optimization. The parameter D_e is connected with dissociation energy. We conducted a series of MD simulations heating the system as described above. We reduced the parameter D_e for each simulation. We calculated the melting curves at each value of the parameter D_e . The optimized value is $D_e = 150 \text{ kJ/mol}$. Further decrease was accompanied by the melting of the double helix with the silver ion at the same temperatures as without a silver ion. Figure 1e shows the fitted curve of the potential Morse with optimized parameters.

We conducted two types of MD simulations of a DNA double helix: equilibrium dynamics at 300 K and simulation of melting. Both simulations began with a relaxation stage followed by a production run. The relaxation stage consisted of energy minimization using the conjugate gradient method and MD simulations under conditions of constant volume (NVT ensemble) and constant pressure (NPT ensemble) at a constant pressure of 1.0 bar and temperature of 300 K for 10 ns. After relaxation, the dynamics of the system were calculated for 100 ns at a temperature of 300 K. The simulation of melting started from the configuration of the system taken from the equilibrium dynamics at 300 K. The system was heated every 20 ns with a temperature step of 10 K. MD simulations were run at temperatures ranging between 300 K and 460 K.

Based on the obtained MD trajectories, the hydrogen bonds (H-bonds) were calculated. In detecting H-bonds, the well-known geometric criterion was used: H-bond between base pairs exists if the distance between two electronegative atoms is less than $3.5 \text{ \AA}$ and the H-bond angle is less than 30° [58].

3. Results

3.1. Dynamics and Stability of the DNA Double Helix with Either C:C or C:Ag:C Mismatch

Since we offer a new potential, we examined the stability of the dodecamer double helix in the presence of a silver ion. Therefore, after the equilibration process was completed, an MD simulation was carried out under constant temperature (300 K) and pressure (NPT) for 100 ns, of which the last 50 ns were used for analysis. The original potential was used to describe the interaction of the silver ion with cytosine atoms. These changes do not affect any other interaction parameters and should not affect the structure of the double helix outside the C:C pair mismatch. The double helix in the presence of the silver ion at 300 K was stable and had an average number of hydrogen bonds even greater than in the simulation with a C:C mismatch in the absence of a silver ion. The presence of silver ions in a system with a C:C mismatch leads to an increase in the stability of the double helix. Figure 1c,d shows typical snapshots at the end of the simulation in the absence and presence of the silver ion.

Figure 1f shows a histogram of the $r_{\text{Ag-N3}}$ distance between the silver ion and cytosine nitrogen. One can see that the MD simulation equilibrium distance $r_{\text{Ag-N3}} = 2.18 \pm 0.003 \text{ \AA}$ is scattered in close proximity around the equilibrium distance between the silver ion and cytosine nitrogen, in agreement with the literature data [51]. We also compared this equilibrium distance with the potential minimum (Figure 1e). The equilibrium value is close to the potential minimum. The displacement is due to the presence of other highly charged cytosine atoms in the immediate vicinity of the nitrogen–silver bond.

The increase in the stability of the C:C pair in the presence of the silver ion is well characterized by the $r_{\text{N3-N3}}$ distance between the N3 nitrogen atoms of cytosines in the silver-mediated pair. Figure 2a shows the $r_{\text{N3-N3}}$ distance. The average distance between N3 atoms of cytosines in the absence of silver ions at temperature 300 K the N3-N3 distance is $r_{\text{N3-N3}} = 3.97 \pm 0.27 \text{ \AA}$. In the presence of the silver ion, the N3-N3 distance is $r_{\text{N3-N3}} = 4.31 \pm 0.07 \text{ \AA}$, which is higher than in the C:C pair, but the C:Ag:C pair is more stable, as indicated by the standard deviation from the average. In this case, we observed large deformations of the C:C pair and the temporary formation of a hydrogen bond between the NH_2 group of one cytosine and the N3 atom of another cytosine. However, such a hydrogen bond turns out to be unstable.

We also calculated the local base pair parameters of the DNA double helix in the absence and presence of the silver ion. The local base pair parameters for the C:C pair are significantly different compared to adjacent pairs based on the absence of a silver ion. Local base pair parameters such as opening, shear and stretch turned out to be the most noticeable (see Figure 3a–c).

In the presence of the silver ion, the profile of shear and stretch parameters along the helix is almost flat (see Figure 3d,e). This indicates greater stability in the silver-mediated C:C pair. The non-typical value of opening is evidence of the cisoid conformation C:C pair and reflects the particular features of this non-canonical pair.

The structure of the silver-mediated pair of cytosines in the transoid and cisoid conformations can be different, as was demonstrated in [51] by the DFT method. In our case, the C:C pair is in the cisoid conformation, as in antiparallel helices. The initial cisoid conformation does not transform into a transoid one during simulation. The angles formed between the silver ion and cytosines in the coordination bond are shown in Table 1. A slight shift in the equilibrium angles is observed, indicating a displacement of the silver ion towards the cytosine oxygens.

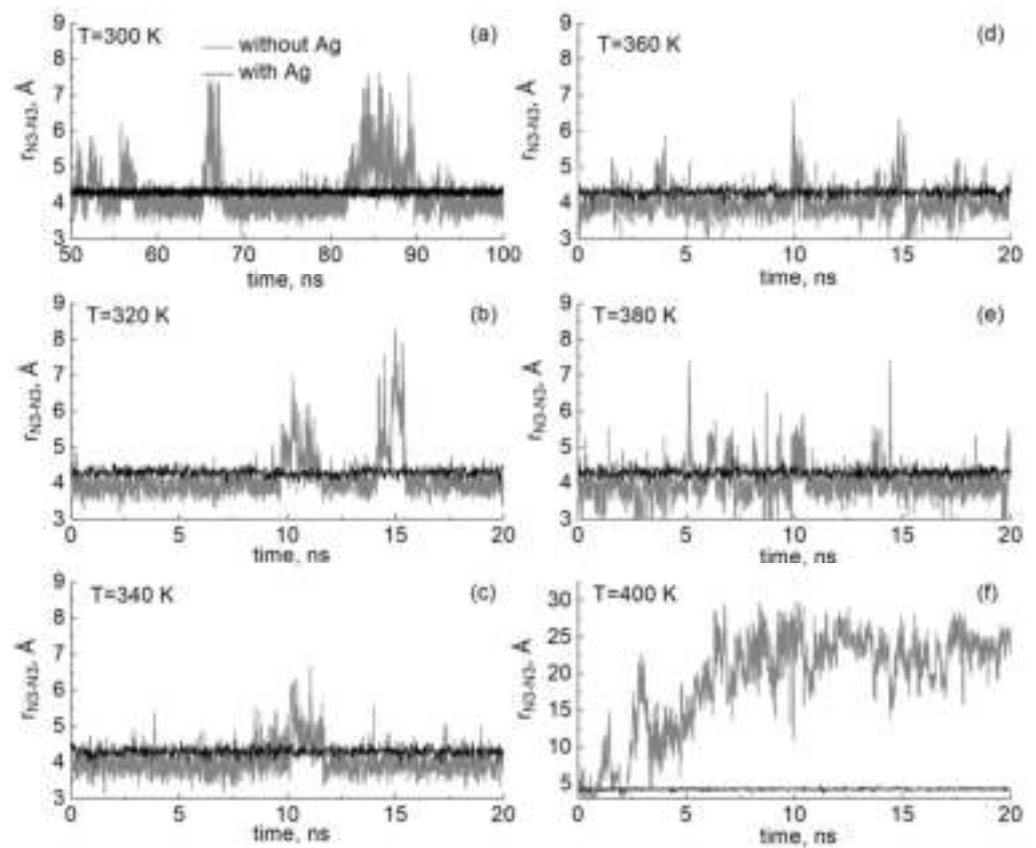


Figure 2. The distance r_{N3-N3} between nitrogen atoms N3 of cytosines in the absence of a silver ion (thick gray line) and in the presence of the silver ion (thin black line) reflects the stability of the C:C and C:Ag:C pairs: (a) A 50 ns interval at the end of the analyzed time interval for the temperature 300 K. (b–f) A 20 ns time interval for the temperatures 320 K (b), 340 K (c), 360 K (d), 380 K (e), 400 K (f).

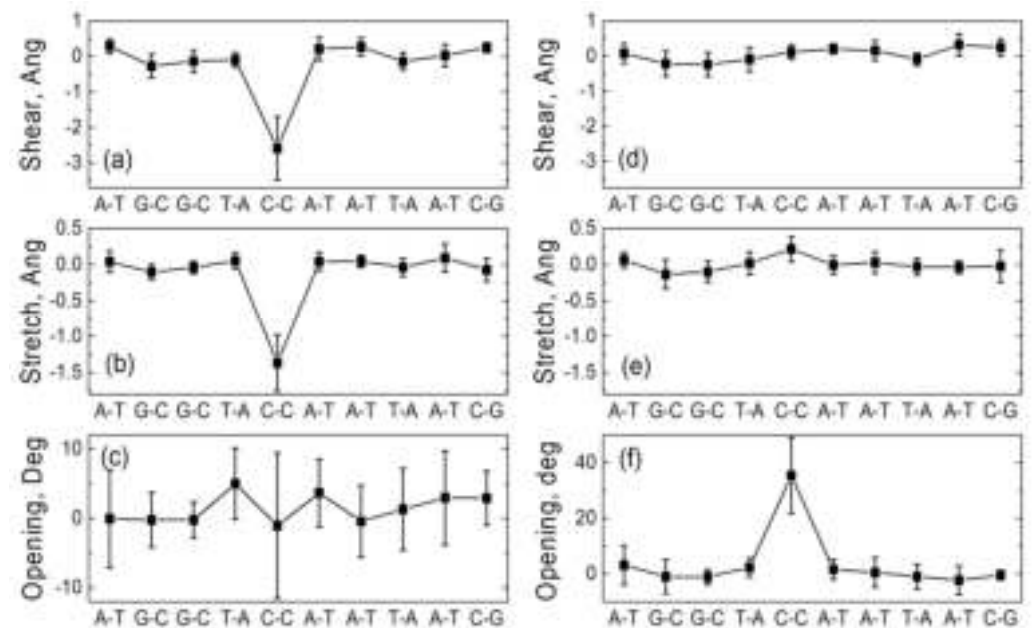


Figure 3. Local base pair parameters: shear, stretch and opening for case of absence (a–c) and presence (d–f) of the silver ion.

Table 1. Angles formed by the coordination bond between the silver ion and nitrogen N3 atom of cytosine. Standard atom notations for nucleotide bases were used.

Angle	θ , (deg)
N3-Ag ⁺ -N3*	163 ± 8
Ag ⁺ -N3-C2	103 ± 7
Ag ⁺ -N3-C4	138 ± 7

3.2. Melting of the DNA Double Helix with Either C:C or C:Ag:C Mismatch

We performed an MD simulation of DNA heating in the range from 300 to 400 K with a temperature step of 10 K. With increasing temperature, the C:C mismatch distance fluctuates strongly, which is often accompanied by the opening of a C:C pair. Figure 2 shows the r_{N3-N3} for some temperatures in the range from 300 to 400 K. At a temperature of 400 K, the C:C pair is already completely destroyed (Figure 2f). In the presence of the silver ion at a temperature of 400 K, the C:Ag:C pair is stable. The destruction of this pair coincides with the melting of the double helix at 410 K.

The maximum number of hydrogen bonds in the complementary base pairs in the sequence under consideration is 25. However, during the course of MD simulation, a smaller number of hydrogen pairs is observed as a result of base pairs fraying, a known process in MD simulations. When a DNA double helix with a C:C pair without a silver ion is heated, melting begins from the terminal base pairs of the double helix and from the C:C pair. Without a stable connection between these cytosines, the double helix forms a bubble, and three melting hot spots form in this case: two terminal pairs and the C:C mismatch. With the silver ion between two cytosines, we observe a more stable double helix, and melting occurs only from the terminal base pairs. The melting curves of the DNA double helix in the presence and absence of a silver ion are presented in Figure 4. It can be seen that in the presence of the silver ion, the melting curve is shifted towards higher temperatures, just as observed in the experiment [28].

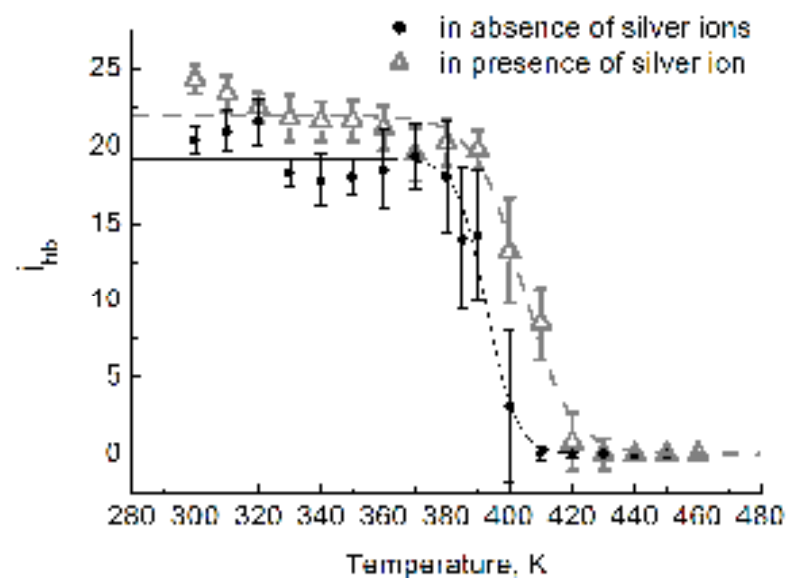
**Figure 4.** The presence of the silver ion shifts the melting curve towards higher temperatures. The temperature dependence of the number of hydrogen bonds i_{hb} without (circles) and with (triangles) silver ion is presented.

Table 2 shows the melting temperatures of the DNA double helix in the absence and presence of the silver ion. As usual for MD simulations, we obtain the higher absolute values of melting temperatures in both cases (due to the superheating, see Discussion); however, we are interested in the relative effect—the melting temperature shift due to the

silver ion presence. This shift, in our case, is 12 K, which is close to the experimental value of 7.5 K. When the DNA melts, the metal-mediated pair breaks down, and the silver ion is released into the solution. The metal-mediated pair C:Ag:C is thermally more stable than the hydrogen bonds of complementary bases, but in this case, it has one silver-mediated base pair, which is also subject to thermal destruction, as observed in the experiment [28].

Table 2. The melting temperatures for double helix in the absence and the presence of the silver ion in comparison with experimental data. The bottom row shows the shift in melting temperature due to the presence of silver ions.

Silver Ion	T _m Experiment ¹ , K	T _m MD, K
absence	299.5	392
presence	307	404
temperature shift ΔT_m	7.5	12

¹ Experimental results were taken from [28] for the same sequence as studied in this work.

4. Discussion

We took this sequence as a robust object for potential validation since its experimentally found melting point is well known in both the absence and the presence of silver ions [28]. We chose the standard antiparallel double helix since, for such objects, the MD approach is well-validated.

By introducing bond potentials to describe the coordination bond between the silver ion and cytosine, we were able to directly model the dynamics of coordinated bond breaking, which was previously impossible while treating these bonds as covalent. The main feature of our potential in comparison with early works is that the structure of the C:Ag:C pair is not fixed near the equilibrium position but can fluctuate significantly near it up to the destruction of the pair with increasing temperature. Our calculations show that the structure of the pair corresponds to the early work [51].

The discrepancies in the absolute values of melting temperatures between experiments [28] and MD simulations are common and attributed to the methodology used to simulate the nanoscale samples [59]. In MD simulations, the melting point is determined by “slow” heating; thus, the rate of temperature change compared to a real experiment is very high and can exceed the experimental value by several orders of magnitude. Since the transition does not occur instantly but requires a small (on an experimental scale) time, a shift is observed. This effect can be corrected by using Monte Carlo simulations instead of molecular dynamics, but this requires more expensive calculations. In addition, the melting temperature discrepancy is affected by the size of the simulation cell and the thermostats used. In molecular dynamics, we are dealing with a small cell, nanometers in size. In this case, the temperature change occurs instantly throughout the entire volume. In the experiment, the measured cell is macroscopic, and temperature changes inside it at the molecular level occur non-uniformly. Thus, we believe that the discrepancy between the melting temperatures in the real experiment and our results is not a disadvantage. We obtained a qualitative agreement of the melting temperature shift in the presence of a silver ion compared to the case when a silver ion is absent. In this case, we observe a satisfactory shift in the melting temperature, comparable with the experimental results.

5. Conclusions

We proposed and verified the Morse potential for the silver-mediated base pair bond that can describe the destruction of the coordination bond between cytosine and a silver ion. The verified potential parameters allowed us to reproduce the process of melting of the DNA double helix with a silver-mediated pair and the destruction of the coordination bond between the silver ion and the cytosine base with increasing temperature. The results of the simulations are in good agreement with experimental data on the melting temperature shift of the DNA double helix with a C:C mismatch in the absence and the presence of a silver ion. Since the B-form double helix is a basic building block not only in the genetic code of

living organisms but also in DNA nanotechnology, metal-mediated double helices may be used for self-assembly dynamic DNA nanomachines. Our results open new possibilities for the development of new metal-mediated 3D DNA materials with increased resistance to thermal, chemical, and possibly enzymatic degradation due to the additional stability provided by silver-mediated base pairs.

Author Contributions: Conceptualization, N.A.K.; methodology, N.A.K.; validation, E.B.G. and N.A.K.; investigation, N.A.K.; writing—original draft preparation, N.A.K.; writing—review and editing, N.A.K.; visualization, E.B.G.; supervision, N.A.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Program of Fundamental Research of the Russian Academy of Sciences (project No. FFZE-2022-0009).

Data Availability Statement: Data are contained within the article.

Acknowledgments: The calculations were carried out in the Joint Supercomputer Center of the Russian Academy of Sciences.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Yang, H.; Mettera, K.L.; Sleiman, H.F. DNA Modified with Metal Complexes: Applications in the Construction of Higher Order metal–DNA Nanostructures. *Coord. Chem. Rev.* **2010**, *254*, 2403–2415. [[CrossRef](#)]
2. Park, K.S.; Park, H.G. Technological Applications Arising from the Interactions of DNA Bases with Metal Ions. *Curr. Opin. Biotechnol.* **2014**, *28*, 17–24. [[CrossRef](#)] [[PubMed](#)]
3. Scharf, P.; Müller, J. Nucleic Acids with Metal-Mediated Base Pairs and Their Applications. *ChemPlusChem* **2013**, *78*, 20–34. [[CrossRef](#)]
4. Swasey, S.M.; Gwinn, E.G. Silver-Mediated Base Pairings: Towards Dynamic DNA Nanostructures with Enhanced Chemical and Thermal Stability. *New J. Phys.* **2016**, *18*, 045008. [[CrossRef](#)]
5. Albarqouni, Y.; Ali, G.A.; Hairul, A.R.M.; Chong, K.F. On the Way of DNA Metallization: Principle, Methods, and Recent Applications. *Chem. Adv. Mater.* **2020**, *5*, 1–14.
6. Chatterjee, S.; Lou, X.Y.; Liang, F.; Yang, Y.W. Surface-functionalized gold and silver nanoparticles for colorimetric and fluorescent sensing of metal ions and biomolecules. *Coord. Chem. Rev.* **2022**, *459*, 214461. [[CrossRef](#)]
7. Ibrahim, N.; Jamaluddin, N.D.; Tan, L.L.; Mohd Yusof, N.Y. A review on the development of gold and silver nanoparticles-based biosensor as a detection strategy of emerging and pathogenic RNA virus. *Sensors* **2021**, *21*, 5114. [[CrossRef](#)]
8. Funai, T.; Adachi, N.; Aotani, M.; Wada, S.I.; Urata, H. Effects of metal ions on thermal stabilities of DNA duplexes containing homo- and heterochiral mismatched base pairs: Comparison of internal and terminal substitutions. *Nucleosides Nucleotides Nucleic Acids* **2020**, *39*, 310–321. [[CrossRef](#)] [[PubMed](#)]
9. Ono, A.; Cao, S.; Togashi, H.; Tashiro, M.; Fujimoto, T.; Machinami, T.; Oda, S.; Miyake, Y.; Okamoto, I.; Tanaka, Y. Specific Interactions between Silver(I) Ions and Cytosine–cytosine Pairs in DNA Duplexes. *Chem. Commun.* **2008**, *39*, 4825–4827. [[CrossRef](#)]
10. Chai, Y.; Guo, X.; Leonard, P.; Seela, F. Heterochiral DNA with Complementary Strands with α -d and β -d Configurations: Hydrogen-Bonded and Silver-Mediated Base Pairs with Impact of 7-Deazapurines Replacing Purines. *Chem. Eur. J.* **2020**, *26*, 13973–13989. [[CrossRef](#)]
11. Jash, B.; Müller, J. Metal-Mediated Base Pairs: From Characterization to Application. *Chem.–Eur. J.* **2017**, *23*, 17166–17178. [[CrossRef](#)] [[PubMed](#)]
12. Sun, Q.; Xie, X.; Song, Y.; Sun, L. A review on silver-mediated DNA base pairs: Methodology and application. *Biomater. Res.* **2022**, *26*, 9. [[CrossRef](#)] [[PubMed](#)]
13. Raju, S.K.; Karunakaran, A.; Kumar, S.; Sekar, P.; Murugesan, M.; Karthikeyan, M. Silver complexes as anticancer agents: A perspective review. *Ger. J. Pharm. Biomater.* **2022**, *1*, 6–28. [[CrossRef](#)]
14. Sinha, I.; Fonseca Guerra, C.; Müller, J. A Highly Stabilizing Silver(I)-Mediated Base Pair in Parallel-Stranded DNA. *Angew. Chem. Int. Ed.* **2015**, *54*, 3603–3606. [[CrossRef](#)] [[PubMed](#)]
15. Johannsen, S.; Megger, N.; Böhme, D.; Sigel, R.K.O.; Müller, J. Solution Structure of a DNA Double Helix with Consecutive Metal-Mediated Base Pairs. *Nat. Chem.* **2010**, *2*, 229–234. [[CrossRef](#)]
16. Mandal, S.; Hepp, A.; Müller, J. Unprecedented Dinuclear Silver(I)-Mediated Base Pair Involving the DNA Lesion 1,N6-Ethenoadenine. *Dalton Trans.* **2015**, *44*, 3540–3543. [[CrossRef](#)] [[PubMed](#)]
17. Polonius, F.-A.; Müller, J. An Artificial Base Pair, Mediated by Hydrogen Bonding and Metal-Ion Binding. *Angew. Chem. Int. Ed.* **2007**, *46*, 5602–5604. [[CrossRef](#)]

18. Swasey, S.M.; Leal, L.E.; Lopez-Acevedo, O.; Pavlovich, J.; Gwinn, E.G. Silver(I) as DNA Glue: Ag⁺-Mediated Guanine Pairing Revealed by Removing Watson-Crick Constraints. *Sci. Rep.* **2015**, *5*, 10163. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Berdakin, M.; Taccone, M.I.; Pino, G.A.; Sánchez, C.G. DNA-Protected Silver Emitters: Charge Dependent Switching of Fluorescence. *Phys. Chem. Chem. Phys.* **2017**, *19*, 5721–5726. [\[CrossRef\]](#)
20. Megger, D.A.; Müller, J. Silver(I)-Mediated Cytosine Self-Pairing Is Preferred over Hoogsteen-Type Base Pairs with the Artificial Nucleobase 1,3-Dideaza-6-Nitropurine. *Nucleosides Nucleotides Nucleic Acids* **2010**, *29*, 27–38. [\[CrossRef\]](#)
21. Ramazanov, R.R.; Sych, T.S.; Reveguk, Z.V.; Maksimov, D.A.; Vdovichev, A.A.; Kononov, A.I. Ag–DNA Emitter: Metal Nanorod or Supramolecular Complex? *J. Phys. Chem. Lett.* **2016**, *7*, 3560–3566. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Tanaka, Y.; Kondo, J.; Sychrovský, V.; Šebera, J.; Dairaku, T.; Saneyoshi, H.; Urata, H.; Torigoe, H.; Ono, A. Structures, physicochemical properties, and applications of T–Hg II–T, C–Ag I–C, and other metallo-base-pairs. *Chem. Commun.* **2015**, *51*, 17343–17360. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Mistry, L.; El-Zubir, O.; Dura, G.; Clegg, W.; Waddell, P.G.; Pope, T.; Hofer, W.A.; Wright, N.G.; Horrocks, B.R.; Houlton, A. Addressing the properties of “Metallo-DNA” with a Ag (i)-mediated supramolecular duplex. *Chem. Sci.* **2019**, *10*, 3186–3195. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Marzilli, L.G.; Kistenmacher, T.J.; Rossi, M. An extension of the role of O (2) of cytosine residues in the binding of metal ions. Synthesis and structure of an unusual polymeric silver (I) complex of 1-methylcytosine. *J. Am. Chem. Soc.* **1977**, *99*, 2797–2798. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Nakagawa, O.; Aoyama, H.; Fujii, A.; Kishimoto, Y.; Obika, S. Crystallographic Structure of Novel Types of AgI-Mediated Base Pairs in Non-canonical DNA Duplex Containing 2′-O, 4′-C-Methylene Bridged. *Nucleic Acids. Chem. Eur. J.* **2021**, *27*, 3842–3848. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Hossain, M.N.; Ahmad, S.; Kraatz, H.-B. Consecutive Silver (I) Ion Incorporation into Oligonucleotides containing Cytosine-Cytosine Mispairs. *ChemPlusChem* **2021**, *86*, 224–231. [\[CrossRef\]](#)
27. Bhai, S.; Ganguly, B. Role of pH in the stability of cytosine-cytosine mismatch and canonical AT and GC base pairs mediated with silver ion: A DFT study. *Struct. Chem.* **2022**, *33*, 35–47. [\[CrossRef\]](#)
28. Zhang, A.; Budow-Busse, S.; Leonard, P.; Seela, F. Anomeric and Enantiomeric 2′-Deoxycytidines: Base Pair Stability in the Absence and Presence of Silver Ions. *Chem.–Eur. J.* **2021**, *27*, 10574–10577. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Schönrrath, I.; Tsvetkov, V.B.; Zatsepin, T.S.; Aralov, A.V.; Müller, J. Silver (I)-mediated base pairing in parallel-stranded DNA involving the luminescent cytosine analog 1, 3-diaza-2-oxophenoxazine. *J. Biol. Inorg. Chem.* **2019**, *24*, 693–702. [\[CrossRef\]](#)
30. Manghi, M.; Destainville, N. Physics of base-pairing dynamics in DNA. *Phys. Rep.* **2016**, *631*, 1–41. [\[CrossRef\]](#)
31. Singh, A.; Maity, A.; Singh, N. Structure and Dynamics of dsDNA in Cell-like Environments. *Entropy* **2022**, *24*, 1587. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Peyrard, M.; Bishop, A.R. Statistical mechanics of a nonlinear model for DNA denaturation. *Phys. Rev. Lett.* **1989**, *62*, 2755–2758. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Singh, A.; Singh, N. Effect of salt concentration on the stability of heterogeneous DNA. *Phys. A* **2015**, *419*, 328–334. [\[CrossRef\]](#)
34. Knotts, T.A.; Rathore, N.; Schwartz, D.C.; De Pablo, J.J. A coarse grain model for DNA. *J. Chem. Phys.* **2007**, *126*, 084901. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Rodrigues Leal, M.; Weber, G. Sharp DNA denaturation in a helicoidal mesoscopic model. *Chem. Phys. Lett.* **2020**, *755*, 137781. [\[CrossRef\]](#)
36. Zoli, M. Base pair fluctuations in helical models for nucleic acids. *J. Chem. Phys.* **2021**, *154*, 194102. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Valle-Orero, J.; Wildes, A.R.; Theodorakopoulos, N.; Cuesta-López, S.; Garden, J.L.; Danilkin, S.; Peyrard, M. Thermal denaturation of A-DNA. *New J. Phys.* **2014**, *16*, 113017. [\[CrossRef\]](#)
38. Singh, A.; Singh, N. Pulling short DNA molecules having defects on different locations. *Phys. Rev. E* **2015**, *92*, 032703. [\[CrossRef\]](#) [\[PubMed\]](#)
39. Oliveira, L.M.; Long, A.S.; Brown, T.; Fox, K.R.; Weber, G. Melting temperature measurement and mesoscopic evaluation of single, double and triple DNA mismatches. *Chem. Sci.* **2020**, *11*, 8273–8287. [\[CrossRef\]](#)
40. Silva, L.G.; Weber, G. Mesoscopic model confirms strong base pair metal mediated bonding for T–Hg²⁺–T and weaker for C–Ag⁺–C. *Chem. Phys. Lett.* **2022**, *803*, 139847. [\[CrossRef\]](#)
41. Freeman, G.S.; Hinckley, D.M.; de Pablo, J.J. A coarse-grain three-site-per-nucleotide model for DNA with explicit ions. *J. Chem. Phys.* **2011**, *135*, 165104. [\[CrossRef\]](#)
42. Florescu, A.; Joyeux, M. Thermal and mechanical denaturation properties of a DNA model with three sites per nucleotide. *J. Chem. Phys.* **2011**, *135*, 085105. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Sayar, M.; Avşaroğlu, B.; Kabakçoğlu, A. Twist-writhe partitioning in a coarse-grained DNA minicircle model. *Phys. Rev. E* **2010**, *81*, 041916. [\[CrossRef\]](#) [\[PubMed\]](#)
44. Sambriski, E.J.; Ortiz, V.; de Pablo, J.J. Sequence effects in the melting and renaturation of short DNA oligonucleotides: Structure and mechanistic pathways. *J. Phys. Condens. Matter* **2009**, *21*, 034105. [\[CrossRef\]](#)
45. Zhang, F.; Collins, M.A. Model simulations of DNA dynamics. *Phys. Rev. E* **1995**, *52*, 4217–4224. [\[CrossRef\]](#)
46. Drukker, K.; Wu, G.; Schatz, G.C. Model simulations of DNA denaturation dynamics. *J. Chem. Phys.* **2001**, *114*, 579–590. [\[CrossRef\]](#)
47. Ouldridge, T.E.; Louis, A.A.; Doye, J.P.K. Structural, mechanical, and thermodynamic properties of a coarse-grained DNA model. *J. Chem. Phys.* **2011**, *134*, 085101. [\[CrossRef\]](#)

48. Wong, K.Y.; Pettitt, B.M. The pathway of oligomeric DNA melting investigated by molecular dynamics simulations. *Biophys. J.* **2008**, *95*, 5618–5626. [[CrossRef](#)] [[PubMed](#)]
49. Lomzov, A.A.; Vorobjev, Y.N.; Pyshnyi, D.V. Evaluation of the Gibbs free energy changes and melting temperatures of DNA/DNA duplexes using hybridization enthalpy calculated by molecular dynamics simulation. *J. Phys. Chem. B* **2015**, *119*, 15221–15234. [[CrossRef](#)]
50. Kundu, S.; Mukherjee, S.; Bhattacharyya, D. Melting of polymeric DNA double helix at elevated temperature: A molecular dynamics approach. *J. Mol. Model.* **2017**, *23*, 1–10. [[CrossRef](#)]
51. Chen, X.; Karpenko, A.; Lopez-Acevedo, O. Silver-mediated double helix: Structural parameters for a robust DNA building block. *ACS Omega* **2017**, *2*, 7343–7348. [[CrossRef](#)] [[PubMed](#)]
52. Mistry, L.; Waddell, P.G.; Wright, N.G.; Horrocks, B.R.; Houlton, A. Transoid and cisoid conformations in silver-mediated cytosine base pairs: Hydrogen bonding dictates argentophilic interactions in the solid state. *Inorg. Chem.* **2019**, *58*, 13346–13352. [[CrossRef](#)] [[PubMed](#)]
53. Espinosa Leal, L.A.; Karpenko, A.; Swasey, S.; Gwinn, E.G.; Rojas-Cervellera, V.; Rovira, C.; Lopez-Acevedo, O. The role of hydrogen bonds in the stabilization of silver-mediated cytosine tetramers. *J. Phys. Chem. Lett.* **2015**, *6*, 4061–4066. [[CrossRef](#)] [[PubMed](#)]
54. Hanwell, M.D.; Curtis, D.E.; Lonie, D.C.; Vandermeersch, T.; Zurek, E.; Hutchison, G.R. Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminform.* **2012**, *4*, 4–17. [[CrossRef](#)] [[PubMed](#)]
55. Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A.E.; Berendsen, H.J. GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **2005**, *26*, 1701–1718. [[CrossRef](#)] [[PubMed](#)]
56. Ivani, I.; Dans, P.D.; Noy, A.; Pérez, A.; Faustino, I.; Hospital, A.; Walther, J.; Andrio, P.; Goñi, R.; Balaceanu, A.; et al. Parmbsc1: A refined force field for DNA simulations. *Nat. Methods* **2015**, *13*, 55. [[CrossRef](#)]
57. Li, P.; Song, L.F.; Merz, K.M. Systematic Parameterization of Monovalent Ions Employing the Nonbonded Model. *J. Chem. Theory Comput.* **2015**, *11*, 1645–1657. [[CrossRef](#)]
58. Filho, E.D.; Ruggiero, J.R. H-bond simulation in DNA using a Harmonic Oscillator Isospectral Potential. *Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.* **1997**, *56*, 4486–4488. [[CrossRef](#)]
59. Zou, Y.; Xiang, S.; Dai, C. Investigation on the efficiency and accuracy of methods for calculating melting temperature by molecular dynamics simulation. *Comput. Mater. Sci.* **2020**, *171*, 109156. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Natural Convection Fluid Flow and Heat Transfer in a Valley-Shaped Cavity

Sidhartha Bhowmick ¹, Laxmi Rani Roy ¹, Feng Xu ² and Suvash C. Saha ^{3,*} 

¹ Department of Mathematics, Jagannath University, Dhaka 1100, Bangladesh; dr.sidhartha@math.jnu.ac.bd (S.B.); lrroy.jnu@gmail.com (L.R.R.)

² School of Civil Engineering, Beijing Jiaotong University, Beijing 100044, China; fxu@bjtu.edu.cn

³ School of Mechanical and Mechatronic Engineering, University of Technology Sydney, Sydney, NSW 2007, Australia

* Correspondence: suvash.saha@uts.edu.au

Abstract: The phenomenon of natural convection is the subject of significant research interest due to its widespread occurrence in both natural and industrial contexts. This study focuses on investigating natural convection phenomena within triangular enclosures, specifically emphasizing a valley-shaped configuration. Our research comprehensively analyses unsteady, non-dimensional time-varying convection resulting from natural fluid flow within a valley-shaped cavity, where the inclined walls serve as hot surfaces and the top wall functions as a cold surface. We explore unsteady natural convection flows in this cavity, utilizing air as the operating fluid, considering a range of Rayleigh numbers from $Ra = 10^0$ to 10^8 . Additionally, various non-dimensional times τ , spanning from 0 to 5000, are examined, with a fixed Prandtl number ($Pr = 0.71$) and aspect ratio ($A = 0.5$). Employing a two-dimensional framework for numerical analysis, our study focuses on identifying unstable flow mechanisms characterized by different non-dimensional times, including symmetric, asymmetric, and unsteady flow patterns. The numerical results reveal that natural convection flows remain steady in the symmetric state for Rayleigh values ranging from 10^0 to 7×10^3 . Asymmetric flow occurs when the Ra surpasses 7×10^3 . Under the asymmetric condition, flow arrives in an unsteady stage before stabilizing at the fully formed stage for $7 \times 10^3 < Ra < 10^7$. This study demonstrates that periodic unsteady flows shift into chaotic situations during the transitional stage before transferring to periodic behavior in the developed stage, but the chaotic flow remains predominant in the unsteady regime with larger Rayleigh numbers. Furthermore, we present an analysis of heat transfer within the cavity, discussing and quantifying its dependence on the Rayleigh number.

Keywords: unsteady flow; natural convection; valley-shaped cavity; bifurcations; heat transfer



Citation: Bhowmick, S.; Roy, L.R.; Xu, F.; Saha, S.C. Natural Convection Fluid Flow and Heat Transfer in a Valley-Shaped Cavity. *Computation* **2024**, *12*, 146. <https://doi.org/10.3390/computation12070146>

Academic Editors: Ali Cemal Benim, Abdulmajeed A. Mohamad, Sang-Ho Suh, Rachid Bennacer, Paweł Oclon and Jan Taler

Received: 4 June 2024

Revised: 11 July 2024

Accepted: 12 July 2024

Published: 14 July 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The study of natural convection within confined spaces has attracted considerable attention from scientists due to its widespread applicability across various disciplines [1,2]. Researchers have utilized different types of enclosures with diverse boundary conditions as experimental setups to investigate natural convection, aiming to deepen our understanding of heat transfer mechanisms and fluid dynamics. This interest stems from the multitude of applications where natural convection phenomena are integral, spanning fields such as geophysics, building insulation, geothermal reservoir management, and industrial separation processes. However, the complexity of the Earth's surface, characterized by irregular geometries highly influenced by slanted terrains, poses challenges for traditional geometric configurations. Despite the prevalence of irregular shapes in both natural and industrial settings, natural convection remains a subject of ongoing study owing to its significance in understanding fluid flow behavior. While much research has focused on natural convection within conventional square or rectangular enclosures due to their simplicity [3,4], the investigation of natural convection within triangular-shaped cavities holds particular

importance as it contributes significantly to our understanding of phenomena such as Rayleigh–Bénard convection [5,6]. Recently, Cui et al. [7] focused on the study of mixed convection and heat transfer in an arc-shaped cavity with inner heat sources under the conditions of bottom heating and top wall cooling. Chuhan et al. [8] examined the thermal behavior of a power law fluid in a plus-shaped cavity as a result of natural convection, taking into account the Darcy number and magnetohydrodynamics. Sarangi et al. [9] explored the heat loss induced by radiation and persistently laminar natural convection in a solar cooker cavity with a rectangular or trapezoidal cavity.

Enhancing our understanding of natural convection flows holds paramount importance in improving predictions of heat transfer and facilitating flow control mechanisms. Researchers intrigued by this phenomenon can investigate detailed explanations of buoyancy-induced natural convection flows [10]. Natural convection near a thermal boundary is often idealized as flow adjacent to a thermally well-conducted infinite or semi-infinite flat plate, leading to the formation of Rayleigh–Bénard convection. Studies by Manneville and Bodenschatz et al. [11,12] provide comprehensive insights into Rayleigh–Bénard instability findings, whereas Sparrow and Husar [13] focus on exploring Rayleigh–Bénard convection on inclined flat plates. Investigating cavities with vertical and horizontal temperature gradients reveals two distinct natural or free convection flow scenarios. While natural convection is predominantly studied in rectangular or square cavities due to their simplicity, Batchelor’s [14] research offers a significant understanding of natural convection flows within differentially heated cavities, particularly emphasizing the influence of conduction on heat transfer, especially at low Rayleigh numbers. As the Rayleigh number exceeds a threshold, convection controls the flow dynamics. Early studies have primarily examined stable natural convection flows [15–17]. The intriguing phenomenon of baroclinity, engendered by the thermal vertical wall, instigates spontaneous convection flows within the interior cavity through viscous shear under symmetric conditions [18]. At lower Rayleigh numbers, the thermal boundary layer proximal to the wall remains steady, while convective instability induces discrete traveling waves as the Rayleigh numbers increase [18]. Recent studies have delved into the dynamics and transient nature of transient natural convection flows, building upon earlier investigations [19,20]. These findings contribute to our understanding of the complex interplay between thermal gradients and fluid dynamics within confined spaces, shedding light on the transient behaviors inherent to natural convection phenomena. Recent research indicates the occurrence of baroclinity-induced spontaneous convection flows and turbulent Rayleigh–Bénard convection [21]. Rahaman et al. [22,23] explore transitional natural convection flows in a trapezoidal enclosure heated from below, supported by numerical investigations, further enriching our comprehension of this complex phenomenon.

Natural convection flows within cavities featuring inclined walls have attracted considerable attention due to their prevalence and ease of observation [24]. Specifically, investigations into the natural convection flows within triangular cavities have revealed their significant enhancement of Rayleigh–Bénard convection, as documented in prior studies [5,6]. Extensive research explores the characteristics and heat transfer mechanisms of natural convection flows within attic-shaped cavities characterized by temperature gradients between the bottom surface and the inclined borders [25–31]. Notably, studies by Asan and Namli [25,26] and Salman [27] have examined the instability and bifurcations of natural convection flow solutions in triangular cavities, revealing empirical evidence showcasing the profound influence of the aspect ratio and Rayleigh number on both the temperature and flow fields. Analytical findings indicate that as the Rayleigh numbers increase, the aspect ratios decrease, leading to the occurrence of multiple vortex flow patterns. Poulikakos and Bejan [29] investigated heat transmission and natural convection flows within attic spaces during night or winter conditions, establishing scaling relationships and engaging in discussions regarding natural convection flow dynamics [30]. Flack [31] elucidated heat transport dynamics using the Nusselt number–Rayleigh number relationship, while Holtzman et al. [32] observed the transition from symmetric to asymmetric

flow patterns with increasing Grashof numbers, identifying a Pitchfork bifurcation based on experimental data. Furthermore, Lei et al. [33] provided evidence of the existence of transient natural convection flows within attic cavities, further enriching our understanding of this complex phenomenon.

The evolution of flow under sudden heating and cooling undergoes three distinct stages: early, transitional, and steady or quasi-steady [33]. It has been observed that lower Rayleigh numbers result in heating across the entire flow zone rather than the splitting of the thermal boundary layer, with layer separation occurring as the Rayleigh numbers increase. Particularly in scenarios of rapid cooling, the thickness of the thermal boundary layer exceeds the vertical distance from the center of the inclined surface to the horizontal bottom, especially at lower Rayleigh numbers. The thermal boundary layer stabilizes prior to cavity cooling as the Rayleigh numbers increase, with fluid passing the sloping surface and contacting the bottom tip before entering the interior. Attic heat transmission has been studied under thermal stimuli by Saha et al. [34–38]. Additionally, researchers have investigated the behavior and heat transfer dynamics of natural convection flows in wedge-shaped cavities, serving as models for various shallow-water bodies with inclined bottom surfaces, such as reservoir sidearms and seashores. As described by Lei and Patterson [39–41], convective flow within wedges undergoes three stages from isothermal and stationary states. Mao et al. [42–44] argued that the horizontal location of the wedge determines the primary heat transfer mode and flow status for varying Rayleigh numbers, with different flow regimes observed in the shallow littoral zone. Bednarz et al. [45] explored natural convection flows within reservoir-shaped cavities, observing the formation of two heating boundary layers: inflow at the bottom and unstable reflux below the water for different Grashof values. Understanding natural convection flows in V-shaped cavities is crucial, given their prevalence in both natural and industrial systems [46–57]. Kenjeres [58] quantified the heat transfer and turbulent natural convection fluxes in V-shaped cavities, noting that upslope flow increases with rising Rayleigh numbers. Kimura et al. [59] utilized angled V-shaped channel open chambers to calculate heat transfer dynamics.

The flow pattern and heat transfer quantities have been meticulously measured in previous studies [30,31]. Despite the detailed examinations conducted, natural convection flows within an attic or a wedge-shaped triangular cavity may not accurately depict the flow dynamics in a V-shaped cavity over time. Given their natural occurrence, natural convection flows within V-shaped triangular enclosures with opposing boundary conditions have been deemed worthy of thorough investigation. Bhowmick et al. [60–62] investigated the transition from symmetric steady flow to asymmetric unsteady flow in a V-shaped triangular enclosure, where heating occurred from below and cooling from above. It was observed that different fluids exhibit distinct movement behaviors. Despite witnessing the transition from regular to chaotic flow in the valley-shaped cavity heated from below, the mechanisms driving unsteady flow in such cavities remain unclear. To the best of the authors' knowledge, no study has examined the unsteady flow structures present at various times in the valley-shaped cavity. Moreover, there is a need for further quantitative analysis of the dynamics of heat transfer in such configurations.

This literature review highlights the significance of investigating unsteady natural convection in V-shaped cavities for air, providing valuable insights into the flow patterns, transition processes, and heat transfer phenomena in such systems. However, a comprehensive understanding of the intricate physics governing unsteady flow mechanisms within V-shaped cavities remains imperative. Analysis of the existing literature reveals a dearth of studies focusing on characterizing unsteady flow structures at different non-dimensional times within V-shaped cavities. To address this gap, the present study aims to examine unsteady flow in a valley-shaped cavity through two-dimensional numerical simulations, considering air with Rayleigh numbers ranging from 10^0 to 10^8 , a Prandtl number of 0.71, and an aspect ratio of 0.5. The impact of various Ra and τ values on the flow structure, heat transfer, and transient flow characteristics in the fully formed stage of the valley-shaped cavity are covered in this study. This research contributes to existing knowledge

by enhancing our understanding of how fluid characteristics affect flow behavior and transition processes. The evolution of flow mechanisms and changes within the cavity are elucidated using the temperature and velocity differences over time. Such insights are crucial for enhancing the energy efficiency of engineering applications and optimizing the design of thermal systems like heat exchangers and cooling systems. Consequently, this study is poised to make a novel contribution to the field, offering distinct and valuable conclusions that will benefit specialists engaged in modeling and experimentation on flow over complex geometries.

2. Numerical Model Formulations

The primary objective of this study is to analyze the characteristics of natural convection flows in a triangular cavity with a valley-shaped configuration. To achieve this, a two-dimensional numerical simulation approach is employed. The physical model and its corresponding boundary conditions are visually represented in Figure 1. In order to rectify the singularity present at the connection between the top and inclined walls, a strategic adjustment is made by removing an appropriate amount of substance from both the top corners. Specifically, 4% of the length is carefully extracted in the form of minuscule points. It should be noted that this minor adjustment does not exhibit any discernible impact on the mechanics of fluid flow and heat transfer, as indicated by previous studies [18,20–24,29–31]. The dimensions of the cavity are defined as follows: the horizontal length is $2L$, and the height is H , where $L = 2H$ and the ratio $A = H/L = 0.5$. At time $T = T_0$, the fluid inside the cavity is initially at a uniform temperature and not in motion. At a given temperature of $T_c = T_0 - \Delta T/2$ and $T_h = T_0 + \Delta T/2$, respectively, the top and inclined walls undergo instantaneous cooling and heating processes. In every given circumstance, all the boundaries are motionless.

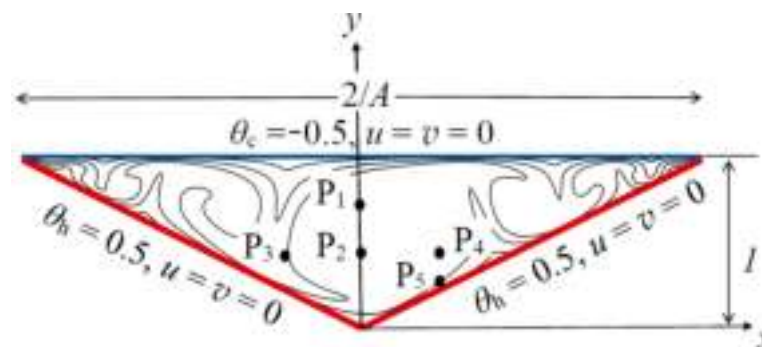


Figure 1. Physical domain and non-dimensional boundary conditions with the figuring points P_1 (0, 0.825), P_2 (0, 0.46), P_3 (−0.5, 0.5), P_4 (0.5, 0.5), and P_5 (0.5, 0.255), which are applied in the following figures.

This study analyses two-dimensional natural convection flows within a valley-shaped enclosure. The governing equations employed in this investigation are presented below, utilizing the Boussinesq approximation as a simplifying assumption [62].

$$\frac{\partial U}{\partial X} + \frac{\partial V}{\partial Y} = 0, \quad (1)$$

$$\frac{\partial U}{\partial t} + U \frac{\partial U}{\partial X} + V \frac{\partial U}{\partial Y} = -\frac{1}{\rho} \frac{\partial P}{\partial X} + \nu \left(\frac{\partial^2 U}{\partial X^2} + \frac{\partial^2 U}{\partial Y^2} \right), \quad (2)$$

$$\frac{\partial V}{\partial t} + U \frac{\partial V}{\partial X} + V \frac{\partial V}{\partial Y} = -\frac{1}{\rho} \frac{\partial P}{\partial Y} + \nu \left(\frac{\partial^2 V}{\partial X^2} + \frac{\partial^2 V}{\partial Y^2} \right) + g\beta(T - T_0), \quad (3)$$

$$\frac{\partial T}{\partial t} + U \frac{\partial T}{\partial X} + V \frac{\partial T}{\partial Y} = \kappa \left(\frac{\partial^2 T}{\partial X^2} + \frac{\partial^2 T}{\partial Y^2} \right). \quad (4)$$

U and V in the equations, respectively, represent the horizontal and the vertical flow rates. For a two-dimensional coordinate system:

- P is the pressure;
- ρ is the density;
- t is the time;
- T is the temperature, at time T_0 , where $T_0 = (T_c + T_h)/2$, the fluid medium in the triangular cavity is isothermally stationary.

The dimensionless variables utilized are as follows:

$$x = \frac{X}{H}, y = \frac{Y}{H}, u = \frac{UH}{\kappa Ra^{\frac{1}{2}}}, v = \frac{VH}{\kappa Ra^{\frac{1}{2}}}, p = \frac{PH^2}{\rho \kappa^2 Ra}, \theta = \frac{T - T_\infty}{T_h - T_c}, \tau = \frac{t \kappa Ra^2}{H^2}. \quad (5)$$

In the aforementioned equations, the variables v, x, y, p, τ , and θ represent the normalized counterparts of U, V, X, Y, P, t , and T , respectively. The three governing variables, the aspect ratio (A), Prandtl number (Pr), and Rayleigh number (Ra), have an impact on the enclosure's natural convective flows (view [6] for further details), which are best described as follows:

$$A = \frac{H}{L}, Pr = \frac{\nu}{\kappa}, Ra = \frac{g\beta(T_h - T_c)H^3}{\nu \kappa}. \quad (6)$$

The above dimensionless variables are added, and then Equations (1)–(4) become (for more information, see [61]):

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0, \quad (7)$$

$$\frac{\partial u}{\partial \tau} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} = -\frac{\partial p}{\partial x} + \frac{Pr}{Ra^{\frac{1}{2}}} \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right), \quad (8)$$

$$\frac{\partial v}{\partial \tau} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} = -\frac{\partial p}{\partial y} + \frac{Pr}{Ra^{\frac{1}{2}}} \left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} \right) + Pr\theta, \quad (9)$$

$$\frac{\partial \theta}{\partial \tau} + u \frac{\partial \theta}{\partial x} + v \frac{\partial \theta}{\partial y} = \frac{Pr}{Ra^{\frac{1}{2}}} \left(\frac{\partial^2 \theta}{\partial x^2} + \frac{\partial^2 \theta}{\partial y^2} \right). \quad (10)$$

The present study used a finite-volume Navier–Stokes solver [62], Fluent 15.0 in ANSYS, to model natural convection within a valley-shaped cavity. The principles of mass conservation, momentum conservation, and energy conservation are the foundations of the computational model that is used to simulate unsteady natural convection flow in an air cavity. The continuity equation, the Navier–Stokes equations, and the energy equations have all been solved using finite volume methods after the appropriate numerical techniques have been applied. The non-uniform 2D mesh system is created with the help of the commercial software ICEM 15.0, and the numerical results have been displayed graphically with the help of the post-processing program TECPLOT 360.

The governing Equations (7)–(10) are solved by means of a finite volume technique using the SIMPLE method, which is thoroughly explained in reference [63] and Saha [64]. The numerical procedure for the SIMPLE method is shown in Figure 2.

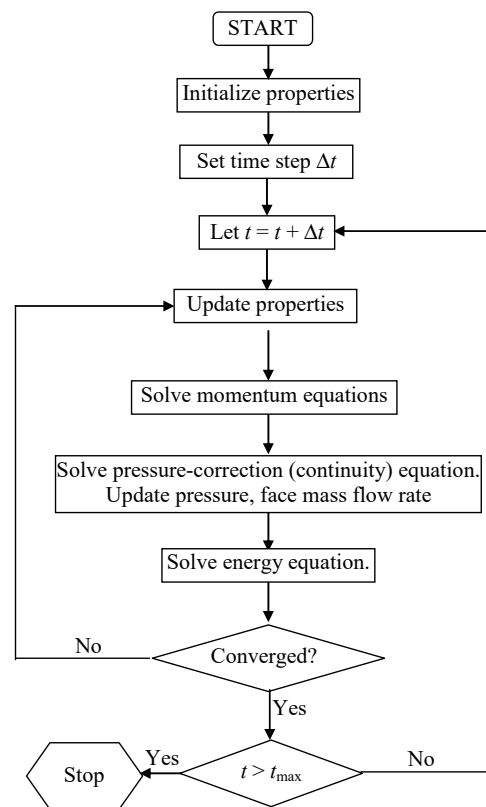


Figure 2. Flowchart of the SIMPLE method for transient flow.

3. Grid Dependency Test

The mesh and time step dependence of the largest Rayleigh number ($Ra = 10^8$) for the Prandtl number ($Pr = 0.71$) and aspect ratio ($A = 0.5$) was investigated in the present inquiry. Three symmetrical, non-uniform meshes were employed in the test, with dimensions of 600×100 , 800×150 , and 1200×200 . These meshes were designed to have finer grids near the boundaries and coarser grids in the inner zone. The 800×150 mesh underwent a 3% expansion, ranging from a minimum width of 0.00025 near the wall to a maximum width of 0.02 in the inside. Figure 3 illustrates the time series of the Nusselt number at the right wall of the cavity. The data were obtained using different meshes and time steps for a Rayleigh number of 10^8 . The Nusselt numbers obtained from the different meshes and time steps show uniformity during the initial stage, along with some divergence during the mature stage. In addition, Table 1 provides a calculation and list of the average Nusselt numbers during the fully developed stage. Based on the analysis conducted, it has been determined that the variations in outcomes observed across the different meshes and time steps are within the acceptable threshold of 2%. Taking this into consideration, the numerical simulation in this study used a mesh size of 800×150 and a time step of 0.0025.

Table 1. Nusselt numbers (Nu) for different grids and time steps.

Mesh and Time Step	Average Nu	Difference
600×100 and $\Delta\tau = 0.0025$	121.63	1.94%
800×150 and $\Delta\tau = 0.00125$	122.47	1.27%
800×150 and $\Delta\tau = 0.0025$	124.04	-
1200×200 and $\Delta\tau = 0.0025$	122.95	0.87%

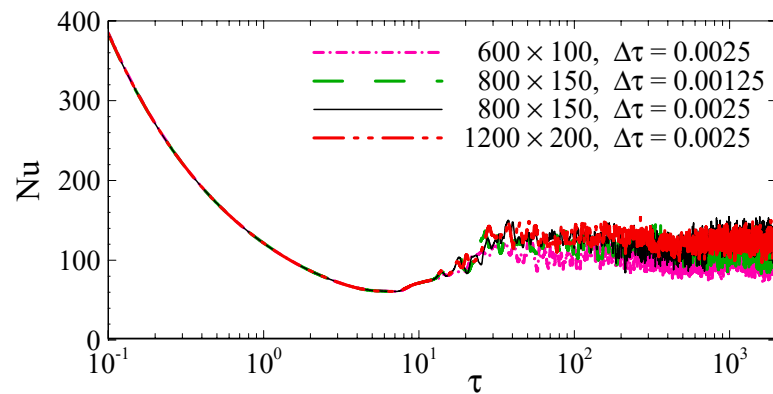


Figure 3. Nusselt numbers time series at the right inclined wall in the valley-shaped cavity for $Ra = 10^8$ with definite grids and time steps.

4. Validation

Figure 4 compares the laboratory experiment conducted by Holtzman et al. [32] with the current numerical results for additional confirmation. In contrast to the triangular cavity in this study, it is vertically inverted and the Rayleigh numbers have been employed in place of the corresponding Grashof numbers. In comparison to the experimental results in Figure 4b for the Rayleigh numbers $Ra = 3.5 \times 10^3$, Figure 4d for $Ra = 7 \times 10^3$, and Figure 4f for $Ra = 7 \times 10^4$, the numerical results depicting the symmetric flow for $Ra = 7 \times 10^3$ in Figure 4a and the asymmetric flow for $Ra = 1.2 \times 10^4$, and $Ra = 10^5$ in Figure 4c,e are accurate. As indicated by Xu et al. [18] and Patterson and Armfield [20], the presence of inconsistencies between numerical Ra values and experimental Ra findings gives rise to a notable contradiction: if the numerical findings and experiment results are accurate, the experimental Rayleigh number is approximately 1.5 to 2.5 times the numerical Rayleigh number. The numerical methods utilized in this research can be used to depict a transitional flow in a triangular cavity, as the experimental and computational results are identical.

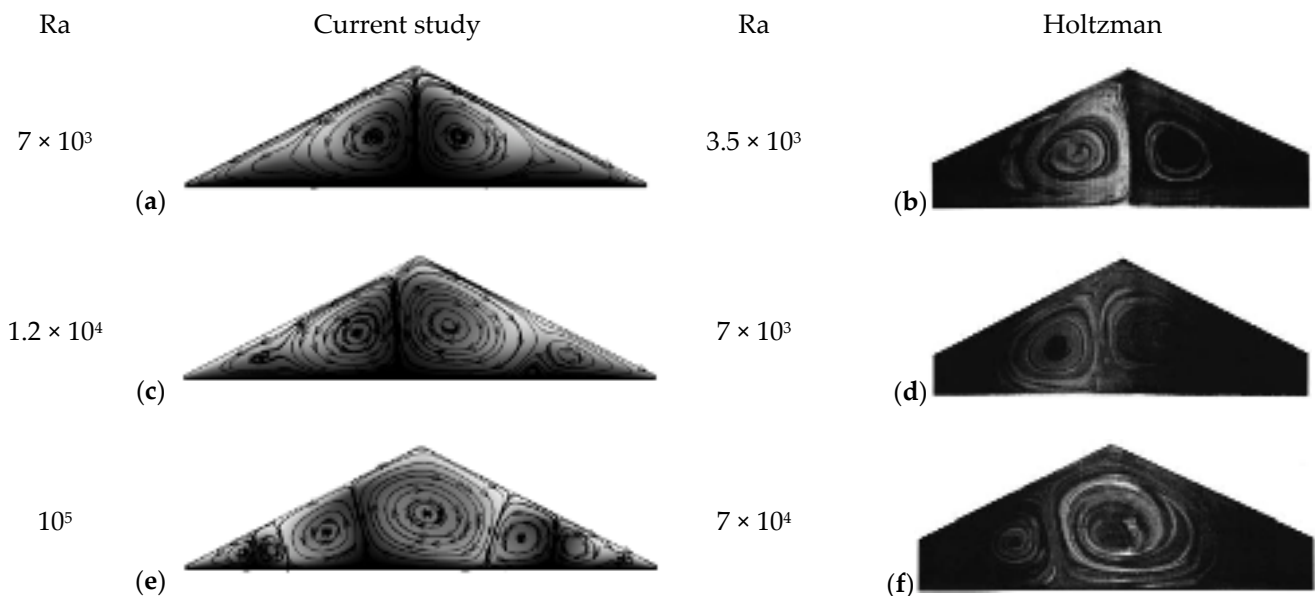


Figure 4. Comparison of the experimental results of Holtzman (b,d,f) [32] for different Rayleigh numbers with the current study (a,c,e).

5. Numerical Results and Discussion

The subsequent section provides an explanation of the primary characteristics of the fluid flows within a V-shaped cavity, wherein the flow is stimulated by heat from the

inclined walls and dissipated by cooling from the top wall. The analysis focuses on a range of Rayleigh numbers, spanning from $Ra = 10^0$ to 10^8 , a Prandtl number, $Pr = 0.71$, and an aspect ratio, $A = 0.5$. Various non-dimensional times are considered to comprehensively examine the flow growth phenomenon. In this work, a 2D numerical simulation has been conducted. According to the numerical simulations, the evolution of the flow for these Rayleigh numbers in different times from $\tau = 0$ to 5000 may be split into symmetric flow, asymmetric flow and unsteady flow.

To produce the simulation results, we used high-performance computing facilities with 64 processors, with which it took 15 days for a small $Ra = 10^5$ and 20 days for a large $Ra = 10^8$ for non-dimensional time, $\tau = 5000$ with time step size, $\Delta\tau = 0.0025$.

5.1. Symmetric Flow

It was discovered that for $Ra = 10^0, 10^1$ and 10^2 , the transitional flow development with non-dimensional time did not exhibit an increasing or decreasing plume, which indicates that for those Rayleigh numbers, the flow exhibited an ongoing level of stability due to the prevailing influence of conduction dominance. In this study, the results for $Ra = 10^0, 10^1$ and 10^2 at $\tau = 0, 0.1, 0.5, 1$ and 2000 are included in Figure 5. Initially, at $\tau = 0$, the temperature is constant, and at this time, there is no fluid flow in the cavity. Fluid flows develop when the inclined walls are suddenly heated and the top wall is suddenly cooled. In Figure 5, the temperature rapidly changes for a small $Ra = 10^0$ and 10^1 at $\tau = 0.1$ and becomes symmetric, but for $Ra = 10^2$, it becomes symmetric at $\tau = 0.5$ and this symmetric tendency is observed from $\tau = 0.5$ to 2000. For all the Ra and τ values, the cavity contains a pair of symmetrical cells. Convective flows exhibit relatively low magnitudes, with the cells having a prevailing conduction dominance that provides stability.

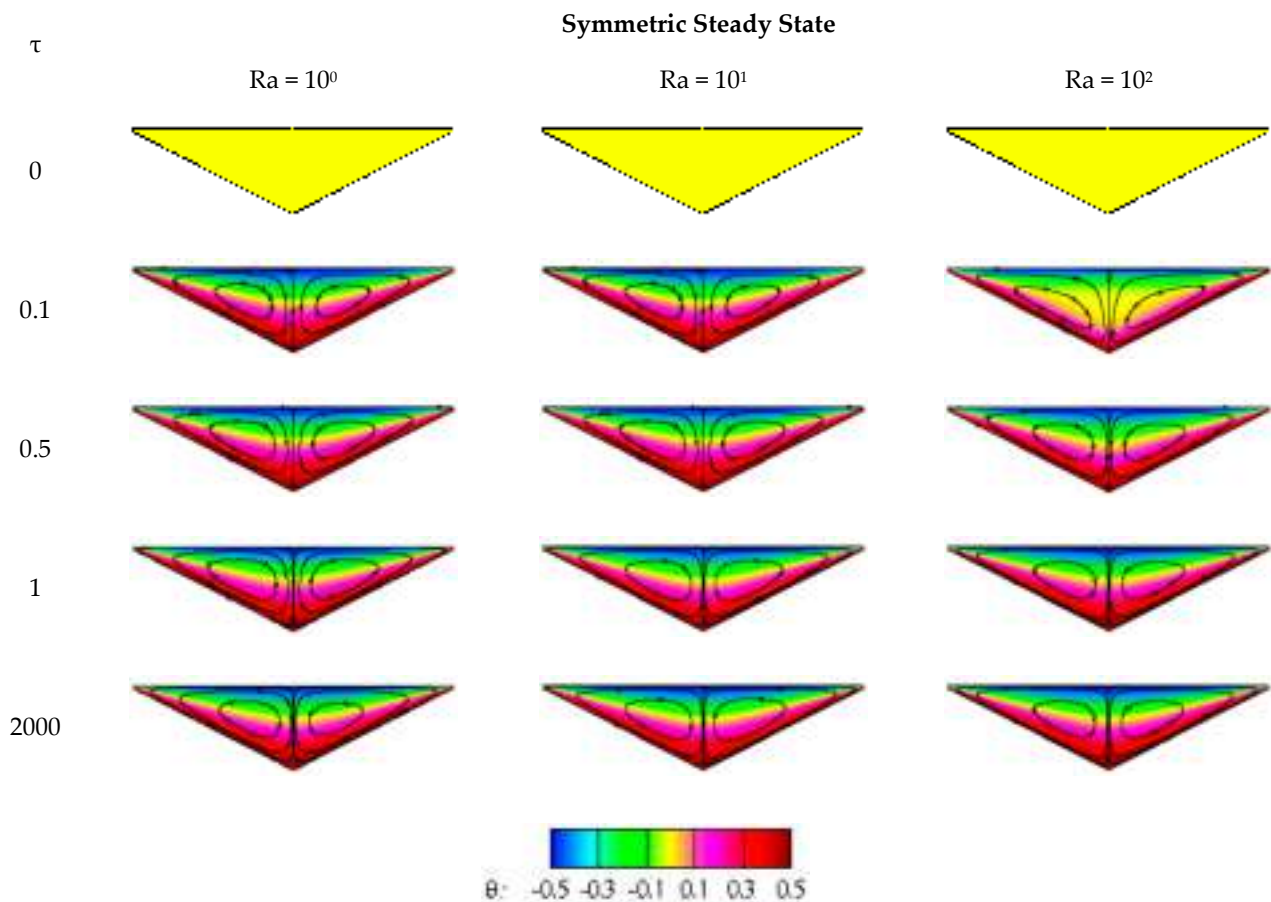


Figure 5. Streamlines and isotherms at various non-dimensional time intervals, τ , and different small Rayleigh numbers, Ra , for the symmetric steady state.

5.2. Asymmetric Flow

The flow of different Rayleigh numbers of $Ra = 10^3$, 7×10^3 and 2×10^4 is shown in Figure 6 on the basis of the time with the displayed streamlines and isotherms. Initially, at $\tau = 0$, the temperature is constant and no fluid moves in the cavity. Fluid flows start when the inclined walls are suddenly heated and the top wall is suddenly cooled. In Figure 6, the temperature gradually changes for $Ra = 10^3$ and 7×10^3 at $\tau = 0.1$ and becomes symmetric, whereas for $Ra = 2 \times 10^4$, it becomes symmetric at $\tau = 1$. In the initial stage, when $\tau = 1$, the steady flows for all the Rayleigh numbers are symmetric. The cavity has two symmetric cells. Since the cells have a conduction dominance that keeps them eternally stable, convective flows are really extremely low. The baroclinity near the inclined walls produces viscous shear, which induces natural convection flows in the fundamental symmetric state, but as time passes, the flow for $Ra = 7 \times 10^3$ and 2×10^4 starts becoming gradually asymmetric. As the figure shows, in the initial transitional stage at $\tau = 1$, the convection in $Ra = 7 \times 10^3$ and 2×10^4 is increased, although it is not sufficiently strong to disrupt the symmetric flow structure (see Figure 7), while the convection in $Ra = 10^3$ remains static. Though there are only two cells when $Ra = 10^3$ and $Ra = 7 \times 10^3$, four cells occur in the cavity for $Ra = 2 \times 10^4$ at $\tau = 10$. That is, aside from symmetric break, the Rayleigh number causes an increase in the number of cells in the cavity. It should be noted that depending on the initial perturbations, any one of the two cells may be bigger and advance into the cavity. The convection of $Ra = 10^3$ is found to be constant and weak in all the subsequent stages of time, like $\tau = 100$, 1000, 2000 and 4000, but the convection in $Ra = 7 \times 10^3$ becomes stronger slowly in $\tau = 100$, 1000, 2000 and 4000, though it is not enough to break the symmetric state. On the other hand, for $Ra = 2 \times 10^4$, the convection grows so strongly that it breaks the symmetrical structure and creates more than two new cells, which is evident at $\tau = 100$, and the flow becomes asymmetric (in Figure 7). Six cells are present for $Ra = 2 \times 10^4$ at $\tau = 100$ and 1000. Depending on the initial perturbations, one of the two cells at $\tau = 1000$ becomes larger and dominates the other cell in the cavity, moving toward the cavity's center. At $\tau = 2000$, the number of cells decreases to five, and this trend persists until $\tau = 4000$. Upon careful examination, it is evident that during the further development stage, the asymmetric flow has attained a state of equilibrium commonly referred to as a steady-state situation. The transition from a symmetric to an asymmetric state, occurring within the range of $Ra = 7 \times 10^3$ to $Ra = 2 \times 10^4$, can be understood as a supercritical Pitchfork bifurcation. This transformation is driven by the onset of Rayleigh–Bénard instability. It is seen that there is more circulation of cells and that the flow pattern is becoming continually asymmetric.

Figure 7 depicts the time series of the temperature and velocity for $Ra = 10^3$, $Ra = 7 \times 10^3$, and $Ra = 2 \times 10^4$ at different points P_1 , P_2 , and P_5 . The fluid within the valley-shaped cavity initially exhibits isothermal and stationary behavior. The cooling and heating of the top wall and the inclined walls occur simultaneously, with a non-dimensional temperature denoted as θ_c for the top wall and θ_h for the inclined walls. Figure 7a demonstrates that there is no difference in temperature between any of the points P_1 and P_2 , and temperature is different at point P_5 in the early stage for a small $Ra = 10^3$. As a result of the passage of time, the temperature increases and decreases at the three different points in the transitional period. Finally, there is no temperature fluctuation in the completely developed stage, and then the flow becomes steady. For more accuracy of the flow steadiness, here, we have used non-dimensional times from 0 to 5000. In Figure 7d, initially, the velocity is zero; as the time increases, the velocity remains the same at points P_1 and P_2 ; and at point P_5 , the velocity increases and finally becomes constant. Similar consequences have been seen in Figure 7b,e. The figures clearly indicate that the flow within the cavity is in a state of steady symmetry. From this, it would seem that the flow is always steady at symmetric states. Now, in Figure 7c, the temperature decreases suddenly and then increases at points P_1 and P_2 during the transitional period and finally stabilizes at the completely developed stage. At point P_5 , there are no changes because of its position. However, the velocity (Figure 7f)

clearly overshoots suddenly and then becomes constant in the completely formed stage, and it indicates that the flow becomes an asymmetric steady state for $Ra = 2 \times 10^4$.

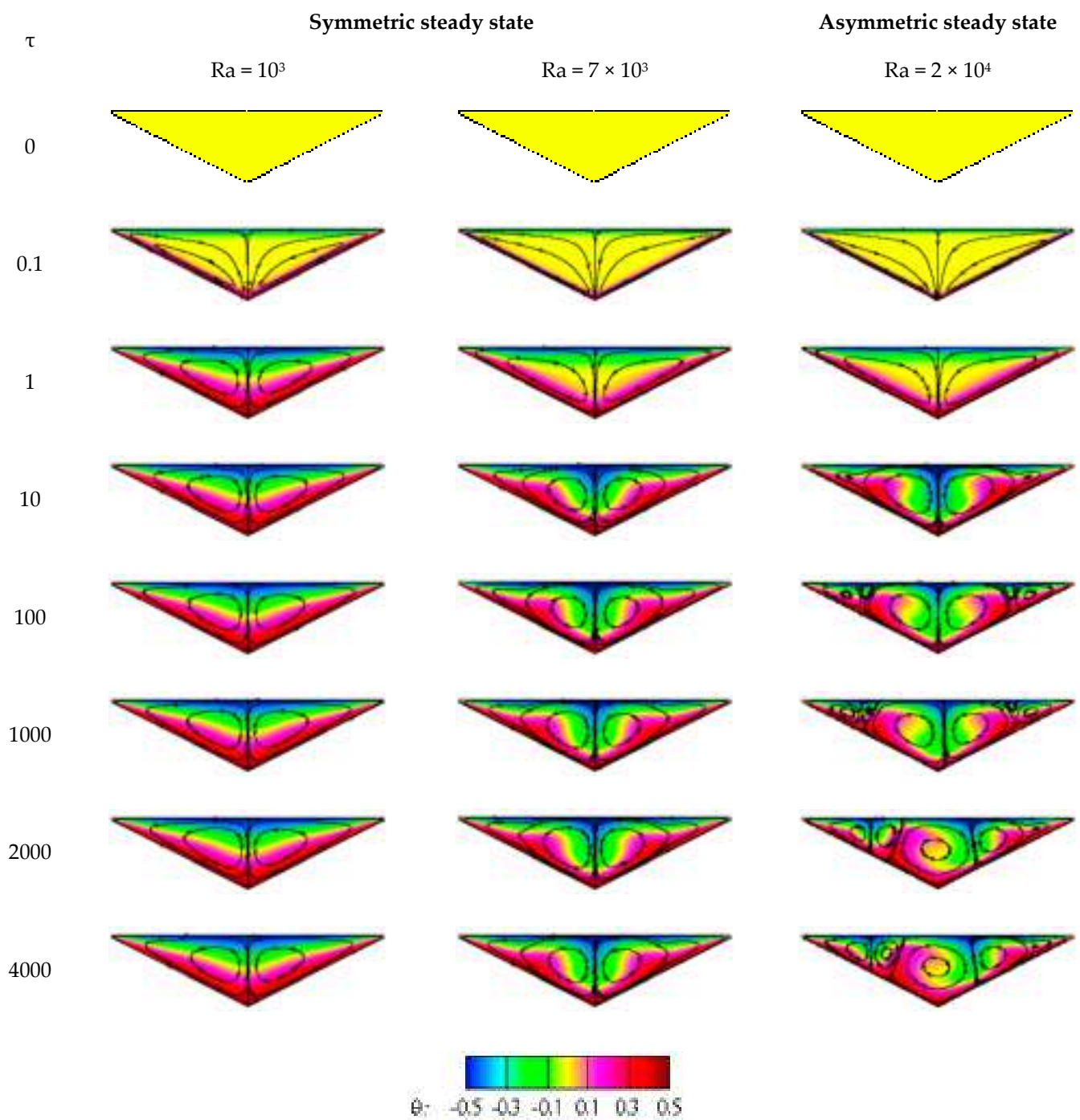


Figure 6. Streamlines and isotherms at various non-dimensional time intervals for the steady-state Rayleigh numbers, $Ra = 10^3$, $Ra = 7 \times 10^3$ and $Ra = 2 \times 10^4$.

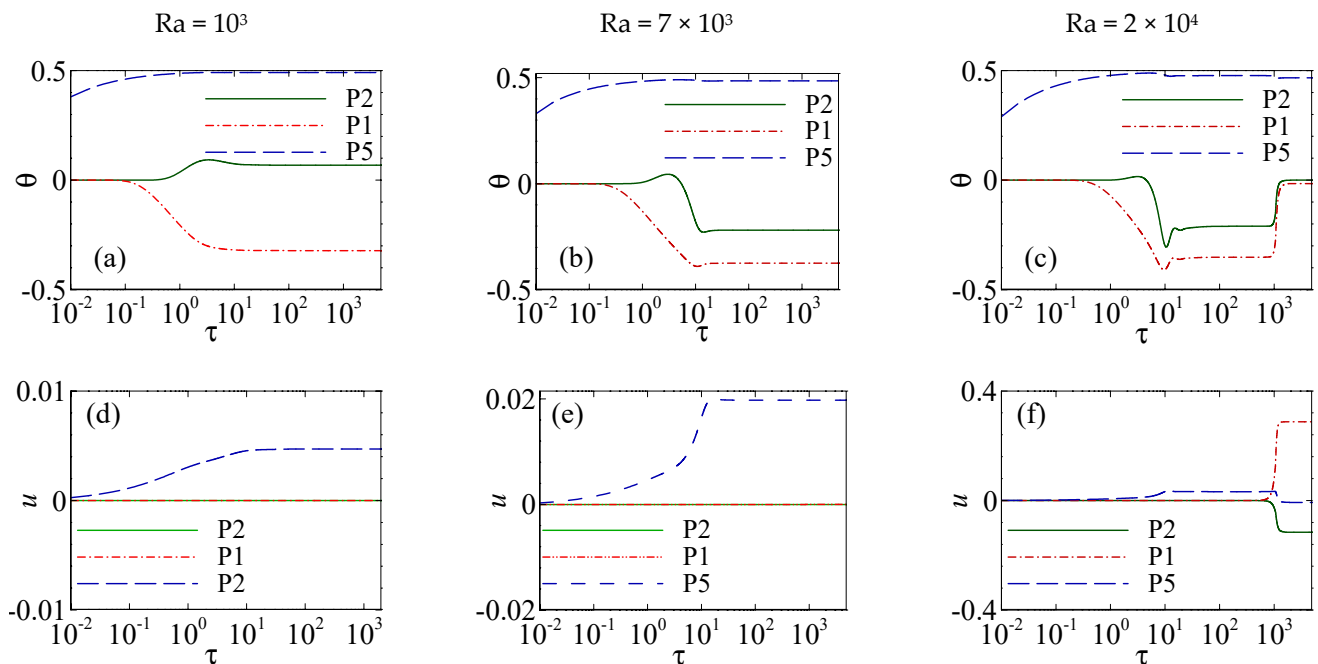


Figure 7. Times series of the temperature and velocity at three distinct points, P₁ (0, 0.825), P₂ (0, 0.46) and P₅ (0.5, 0.255), for (a,d) when $Ra = 10^3$, for (b,e) when $Ra = 7 \times 10^3$, and for (c,f) when $Ra = 2 \times 10^4$.

5.2.1. Pitchfork Bifurcation

Figure 8 illustrates the isotherms and streamlines corresponding to a Rayleigh number of $Ra = 10^4$. Figure 8a demonstrates the continued presence of clear symmetry in the flow at $Ra = 7 \times 10^3$. At $Ra = 10^4$ and $Ra = 1.3 \times 10^4$, the flow shows asymmetry, as depicted in Figure 8b,c. It is to be noted that depending on the initial perturbations, any one of the two cells might grow bigger and move toward the cavity. When the Rayleigh number reaches a value of $Ra = 1.3 \times 10^4$, as depicted in Figure 8c, an additional cell emerges in the top-right region of the cavity (depending on the initial perturbations), and such an asymmetric flow configuration becomes more clearly evident as being the value of the Rayleigh number increases. That is, aside from the symmetric break, the Rayleigh number is responsible for the growing number of cells within the cavity. For instance, Figure 8b has four cells for $Ra = 10^4$, whereas Figure 8c has five cells for $Ra = 1.3 \times 10^4$. The transition observed around $Ra = 10^4$, where a symmetric state transforms to an asymmetric state, can be characterized as a supercritical Pitchfork bifurcation, which happens when Rayleigh–Bénard instability begins to occur.

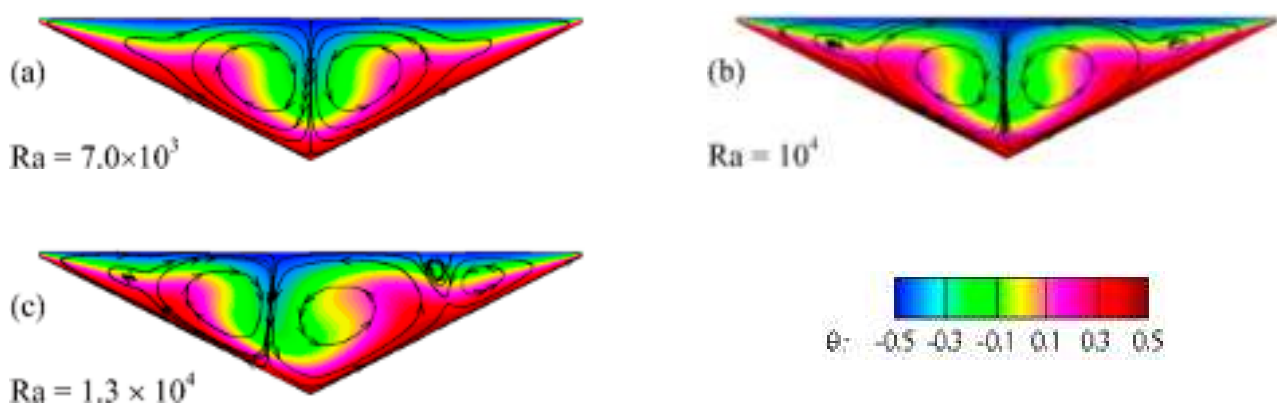


Figure 8. Streamlines and isotherms at $Ra = 7 \times 10^3$, $Ra = 10^4$, and $Ra = 1.3 \times 10^4$.

Table 2 presents the x -velocity at point P_2 (0, 0.46) for various Rayleigh numbers ($Ra = 10^1$ to 10^6) to comprehend the transition from a symmetric to an asymmetric state during the fully developed stage ($\tau = 2000$) of the Pitchfork bifurcation. For Ra values less than or equal to 7.5×10^3 , the x -velocity is nearly zero due to the symmetry of the flow. Once the Rayleigh number reaches or exceeds 7.6×10^3 , the cell's polarity shifts to positive when the x -velocity increases and negative when the x -velocity decreases.

Table 2. Velocities in the x -direction for different Rayleigh numbers at point P_2 (0, 0.46).

Rayleigh Number	Velocity
10^1	0
10^2	0
10^3	0
7.5×10^3	0
7.6×10^3	$\pm 9.47 \times 10^{-5}$
10^4	± 0.00669
1.2×10^4	± 0.0557
1.5×10^4	± 0.09
10^5	± 0.096
10^6	± 0.0993

At the initial stage at $\tau = 0$, the temperature is constant, as depicted in Figure 9. The figure indicates that the flows are symmetric and continuous in the early stage at $\tau = 1$ for $Ra = 5 \times 10^4$, $Ra = 10^5$ and $Ra = 10^6$. The Pitchfork bifurcation causes the flow to intensify over time and eventually become asymmetric. The graphs show the asymmetric streamlines and isotherms at Rayleigh numbers $Ra = 5 \times 10^4$, $Ra = 10^5$ and $Ra = 10^6$ at the transitional stages ($\tau = 100$ and 800). It is undoubtedly interesting that the flow oscillates for a considerable amount of time with a higher Rayleigh number. In this figure, for all the Ra values, the flows in the valley-shaped cavity are asymmetric and steady with the passing of time. For a higher $Ra = 10^6$, the flow becomes asymmetric and steady at $\tau = 800$, whereas $Ra = 5 \times 10^4$ and $Ra = 10^5$ take more time to reach the asymmetric steady-state situation.

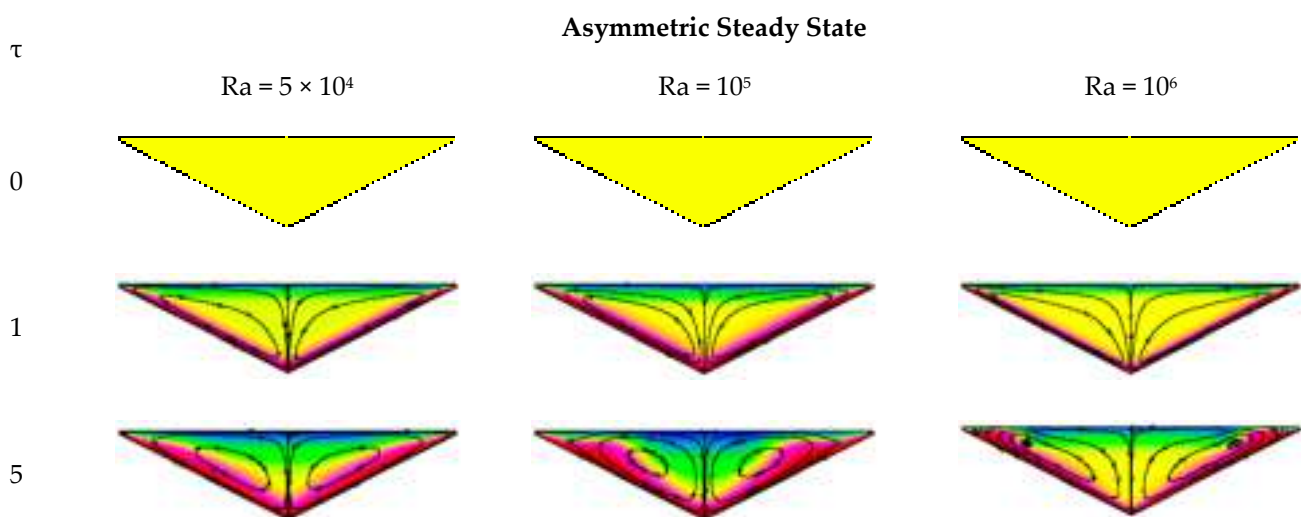


Figure 9. Cont.

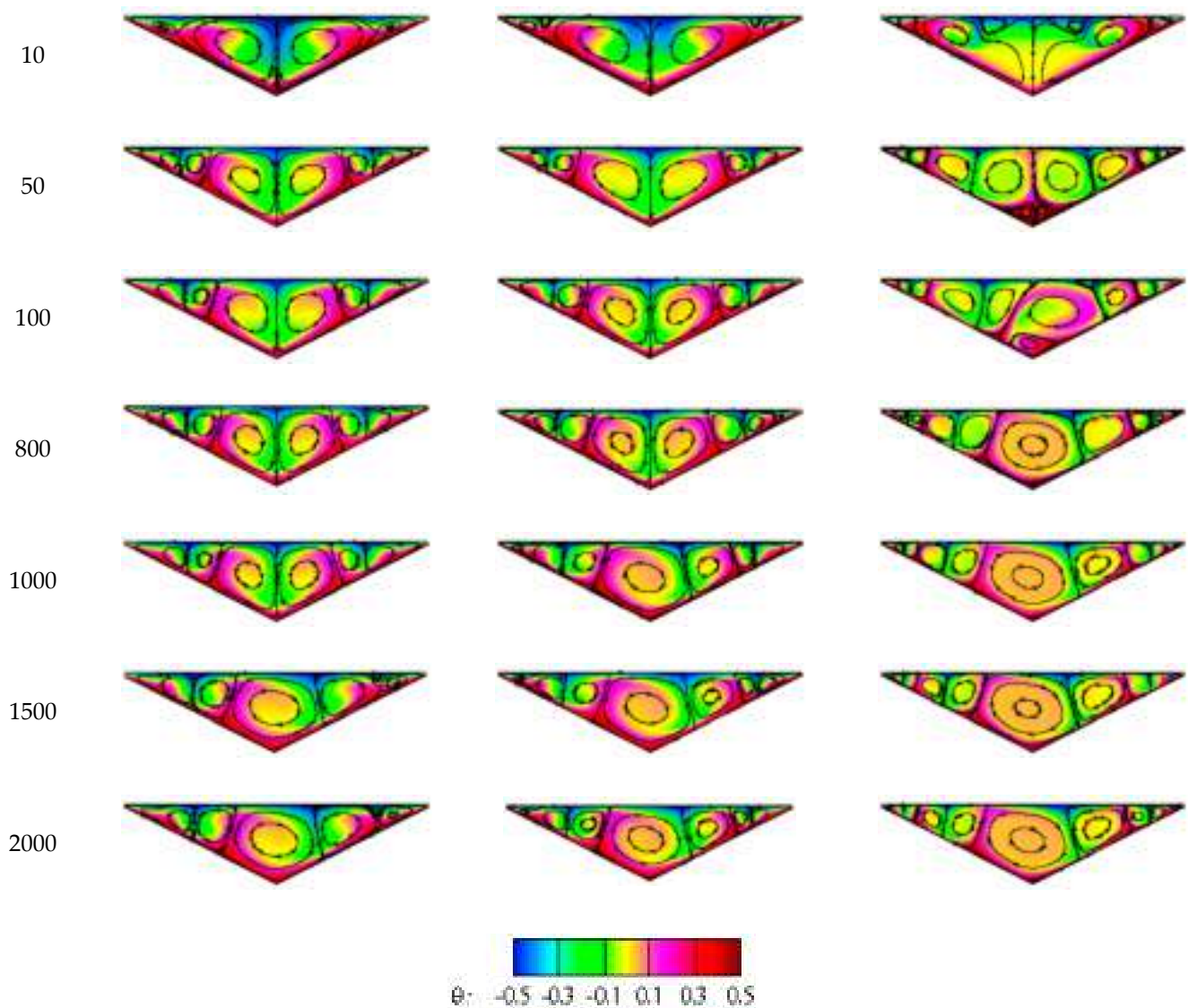


Figure 9. Streamlines and isotherms at various non-dimensional time intervals for $Ra = 5 \times 10^4$, $Ra = 10^5$ and $Ra = 10^6$.

Clearly, the temperature and velocity effect on the flow in the cavity with time, as has been calculated in Figure 10 for more precision at a range of Rayleigh numbers and various points. Figure 10a,b demonstrate that, at first, the temperatures are uniform at points P_1 and P_2 , but point P_5 is located close to the heated wall (see Figure 1), which causes it to heat up more quickly than the other points in the early stage. As time goes on, the temperature in the transitional stage first drops and then rises, ultimately achieving its constant point in the fully developed stage, when there are no longer any temperature variations. Also, from Figure 10d,e, it is clearly seen that the velocity increases with the passing of time. The velocity overshoots in the transitional period before becoming stable in the fully developed stage. Figure 10c shows that the temperature increases in the transitional stage and becomes stable in the fully developed stage. In contrast, the velocity in Figure 10f fluctuates during the transitional period before becoming constant during the last stage. The preceding figures and analyses indicate that in an asymmetric state, the flow becomes unsteady in the transitional period and then steady in the fully developed stages.

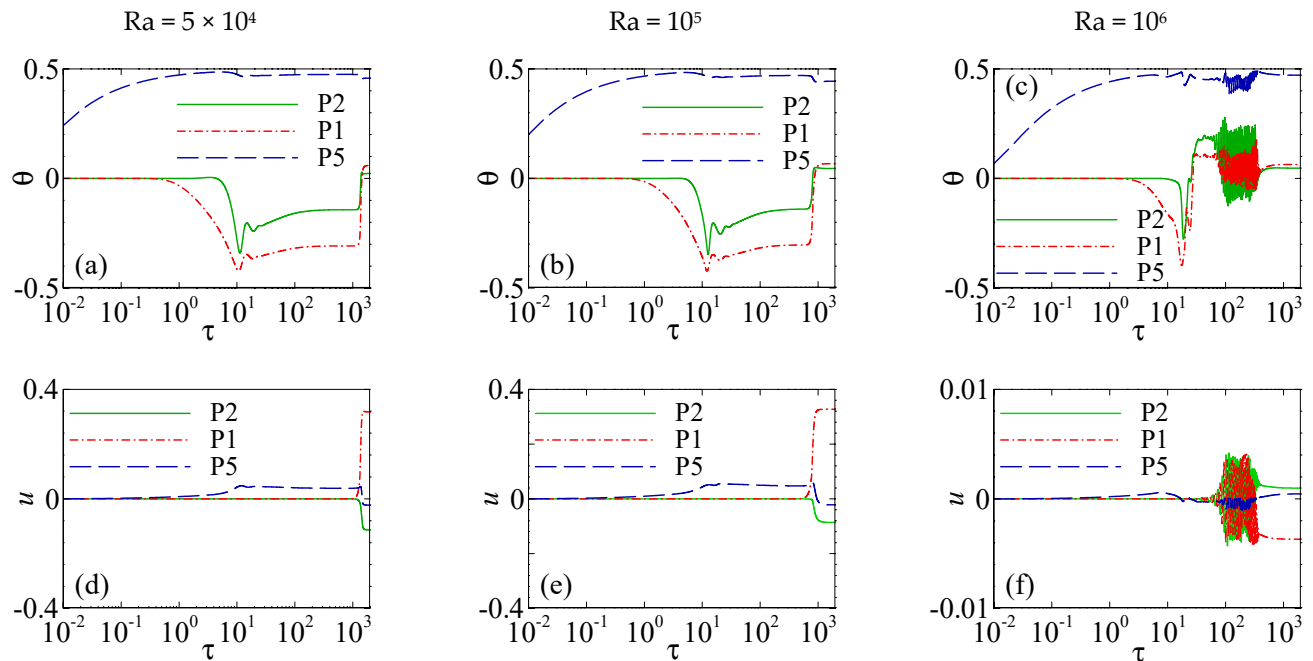


Figure 10. Times series of the temperature and velocity at three distinct points, P_1 (0, 0.825), P_2 (0, 0.46), and P_5 (0.5, 0.255), for (a,d) when $Ra = 5 \times 10^4$, for (b,e) when $Ra = 10^5$, and for (c,f) when $Ra = 10^6$.

5.2.2. Other Bifurcations

For higher Rayleigh numbers, the asymmetric isotherms and streamlines are depicted in Figure 11. It is evident that when the Rayleigh number increases, several types of further bifurcations occur, leading to an increase in the number of cells, as seen in Figure 11. In Figure 11a, there are two cells in the cavity. With the increase in the Ra , we see in Figure 11b that the cell number becomes four; in Figure 11c, the cell number is five. For instance, the cell number increases from six at a Rayleigh number of 5×10^4 in Figure 11d to more than six, as shown in Table 3. Figure 11 demonstrates that when the Rayleigh number rises, more new tiny cells develop in the left or right side of the cavity. However, a single large cell consistently persists at the center of the cavity. This indicates that the flow configuration within the cavity becomes increasingly complex for an asymmetric steady state when the Rayleigh number rises.

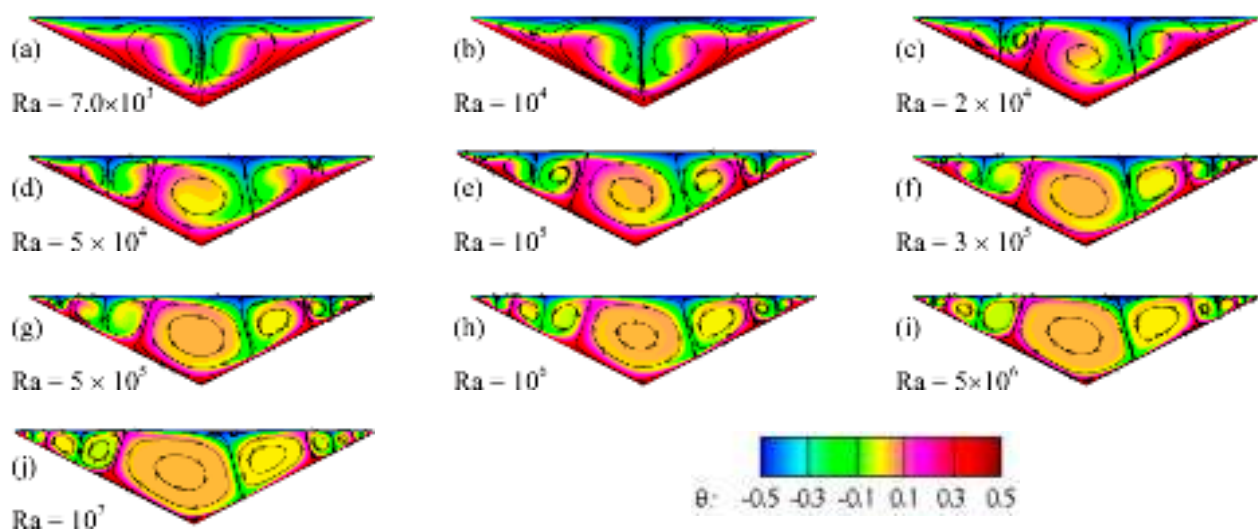


Figure 11. At the fully developed stage, streamlines and isotherms for various Rayleigh numbers.

Table 3. Number of cells with corresponding Rayleigh numbers.

Rayleigh Number	Number of Cells
7×10^3	2
10^4	4
2×10^4	5
5×10^4	6
10^5	7
3×10^5	9
5×10^5	10
10^6	11
5×10^6	13
10^7	15

Table 3 shows that the Rayleigh number and the number of cells are linked. The Rayleigh number increased from 7×10^3 to 10^7 as the number of bifurcations increased from two to fifteen cells. For Ra values between 10^4 and 10^7 , it is clear that there is a nearly linear link between the number of cells and the Ra.

5.3. Unsteady Flow

An asymmetric flow structure is created as a consequence of a Pitchfork bifurcation that arises during the later stages of the transitional period. As already noted, the numerical simulation experiences a Pitchfork bifurcation early on. Figure 12 displays the streamlines and isotherms for $Ra = 10^7$, $Ra = 5 \times 10^7$, and $Ra = 10^8$ to better study the flow at greater Rayleigh numbers for different times. In Figure 12, initially at $\tau = 0$, the temperature is constant for all the Ra values. The figure depicts that the fluids are symmetric for all the Rayleigh numbers when $\tau = 1$, but as time passes, when $\tau = 5$, they become asymmetric and create two additional tiny cells in the cavity at the right and left upper corners. With increasing cell numbers, all the Ra values at $\tau = 10$ in Figure 12 exhibit the same flow pattern. When $\tau = 50$ for $Ra = 10^7$, a tiny cell appears in the top center of the two largest cells in the cavity, but for $Ra = 5 \times 10^7$ and $Ra = 10^8$ at $\tau = 50$, the steady flow breaks and becomes unsteady. In Figure 11, the flow is more convoluted for $Ra = 5 \times 10^7$ and $Ra = 10^8$ at $\tau = 50$, even though it is still constant, and the biggest cell has a few smaller cells at its upper right and left side. This indicates that between $Ra = 10^7$ and 5×10^7 , there is a Hopf bifurcation (see [19] for details on the bifurcation). When $Ra = 10^7$, the asymmetric steady state remains similar at all the subsequent transitional, developed transitional and fully developed stages when $\tau = 100, 1000, 1500$ and 2000 , respectively. The observed trends remain consistent during both the developed transitional stage ($\tau = 1000$ and 1500) and the completely developed stage ($\tau = 2000$) for two distinct Rayleigh numbers, namely $Ra = 5 \times 10^7$ and $Ra = 10^8$. But as τ rises, as shown in Figure 12 for $Ra = 10^8$ at $\tau = 2000$, both cells grow in the middle of the two largest cells. The largest central cell exhibits a right-to-left movement. The unsteady flow becomes more and more complex due to its lack of stability.

Over time, the temperature series has been followed and analyzed spectrally to gain an understanding of the unsteady flow at higher Rayleigh numbers. Figure 13 displays the temperature variation over time and the power spectral densities according to various Ra values. Figure 13a demonstrates a steady flow throughout the fully established stage for $Ra = 10^7$, while Figure 13b exhibits periodic flow for $Ra = 5 \times 10^7$. Additionally, Figure 13c displays the temperature power spectral density located in Figure 13b. The periodic flow's fundamental frequency, with harmonic modes, is $f_p = 0.397$. The periodic flow changes as the Rayleigh number rises. This indicates the occurrence of one more bifurcation for which a periodic solution transforms into another. As the Rayleigh number continues to rise, the unstable flow becomes chaotic, as seen in Figure 13d, which depicts the fully developed stage for $Ra = 10^8$. According to Figure 13e, the unique frequency with harmonic modes vanishes for $Ra = 10^8$, while the flow is chaotic but unstable at the fully developed stage.

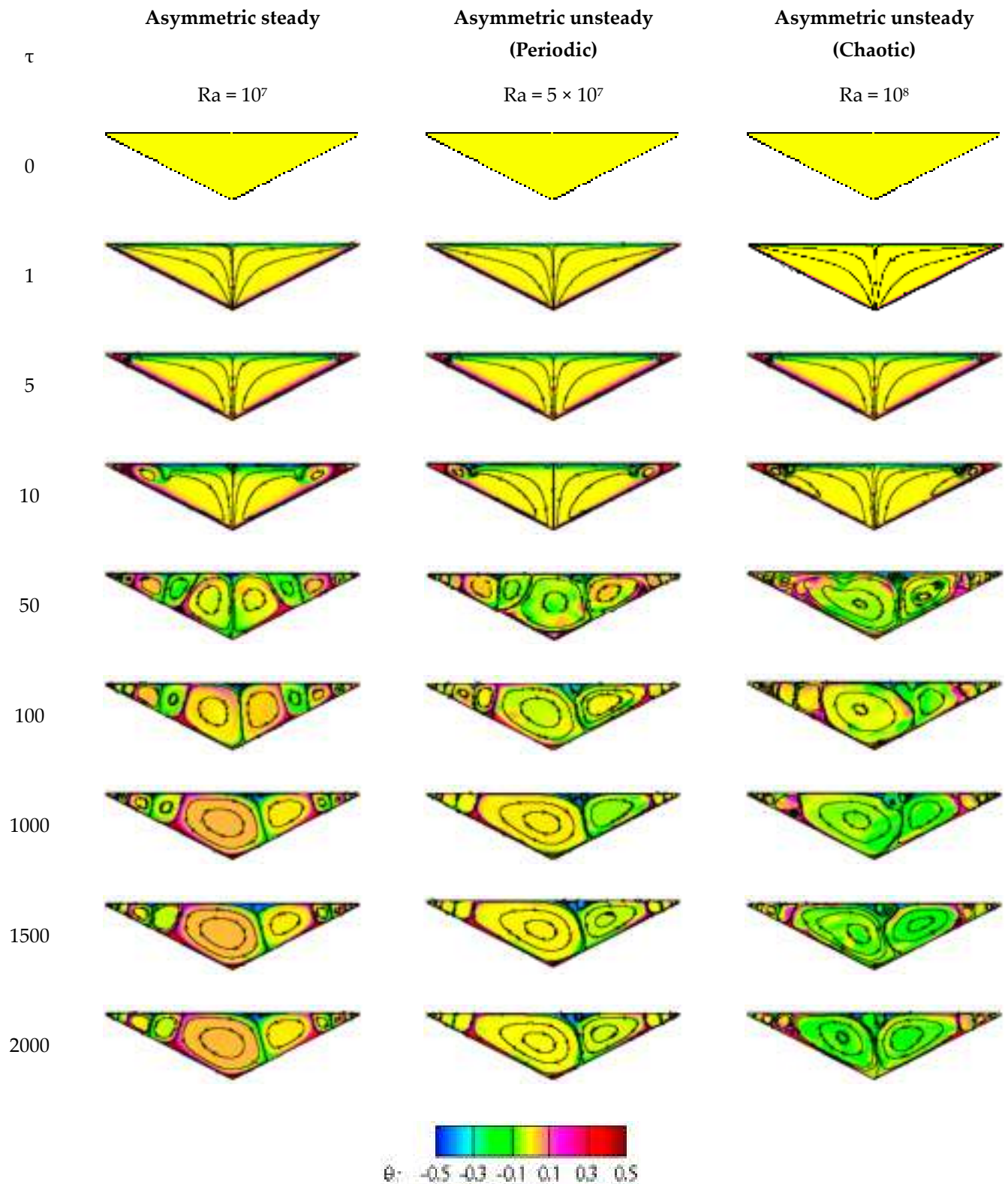


Figure 12. Streamlines and isotherms at various non-dimensional time intervals for distinct higher Rayleigh numbers: $Ra = 10^7$, $Ra = 5 \times 10^7$ and $Ra = 10^8$.

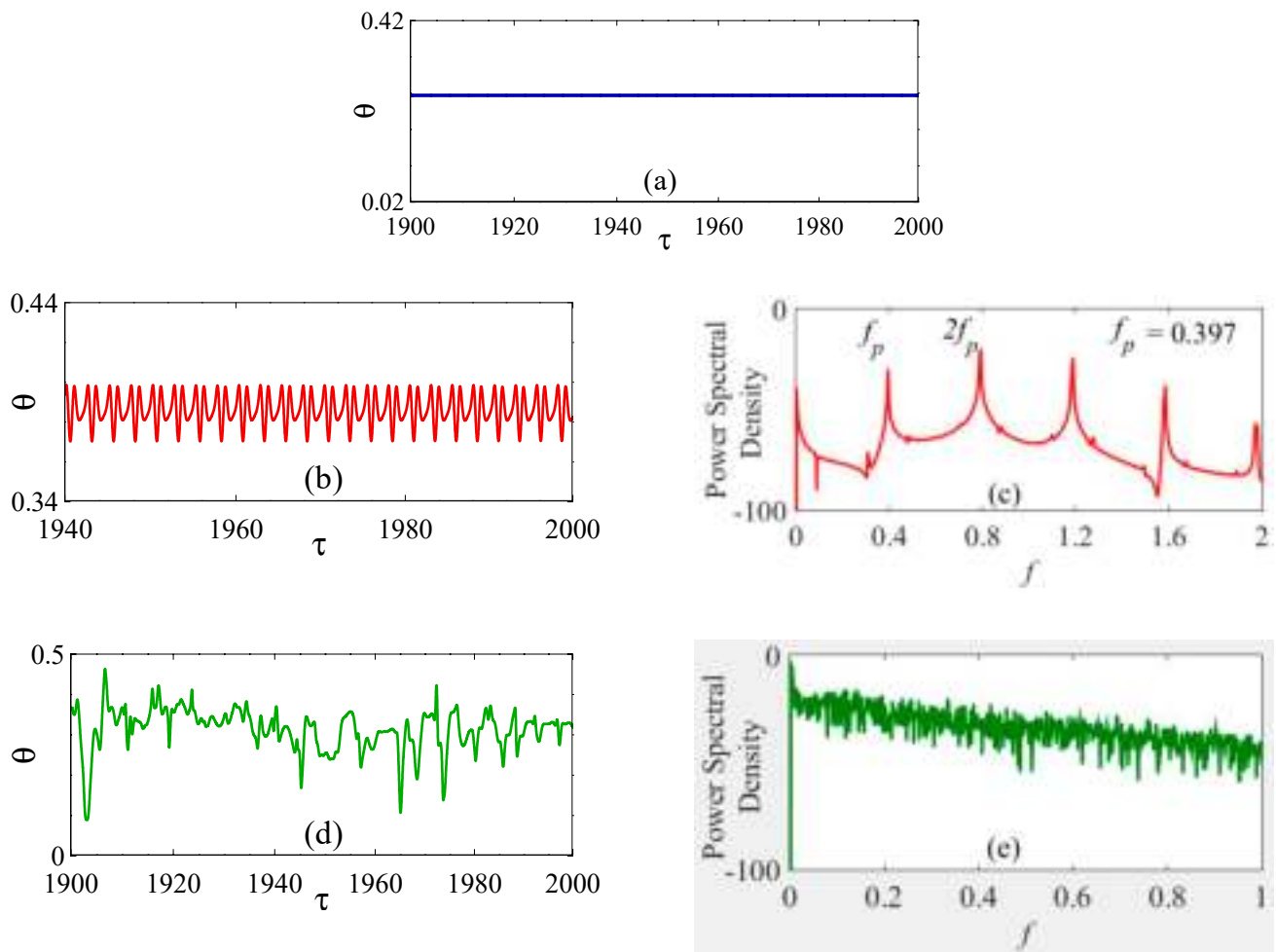


Figure 13. A time series of the temperature at the completely developed stage and the power spectral density at point P_5 (0.5, 0.255) (a) for $Ra = 10^7$, (b,c) for $Ra = 5 \times 10^7$, and (d,e) for $Ra = 10^8$.

Figure 14 depicts a time series analysis of the temperature as well as the velocity at three distinct points, P_1 , P_2 , and P_5 , with higher Rayleigh numbers, specifically $Ra = 10^7$, 5×10^7 and 10^8 . Because the temperature is initially isothermal and constant in Figure 14a–c, it is the same at various stages in the early stage for greater Rayleigh numbers. In the transitional period, the temperature fluctuates with the time for each Ra in Figure 14. In the advanced stage of development, as Figure 14a shows, the temperature becomes constant; that is, there are no changes in the temperature, but in Figure 14b, we see that it becomes periodic. In Figure 14c, with the time increases, the flow becomes chaotic and more complex in the fully developed stage. However, the velocity (Figure 14d–f) starts off at zero and changes as time moves on during the transitional stage. Figure 14d–f all show that the flow is stable in the developed stage, periodic in Figure 14e, and chaotic in Figure 14f. In addition, most importantly, it seems that the flow of a periodic state changes from a periodic one to a chaotic one at a stage of transition. Furthermore, in a chaotic state, the flow is constantly chaotic at the transitional stage and at the developed stages.

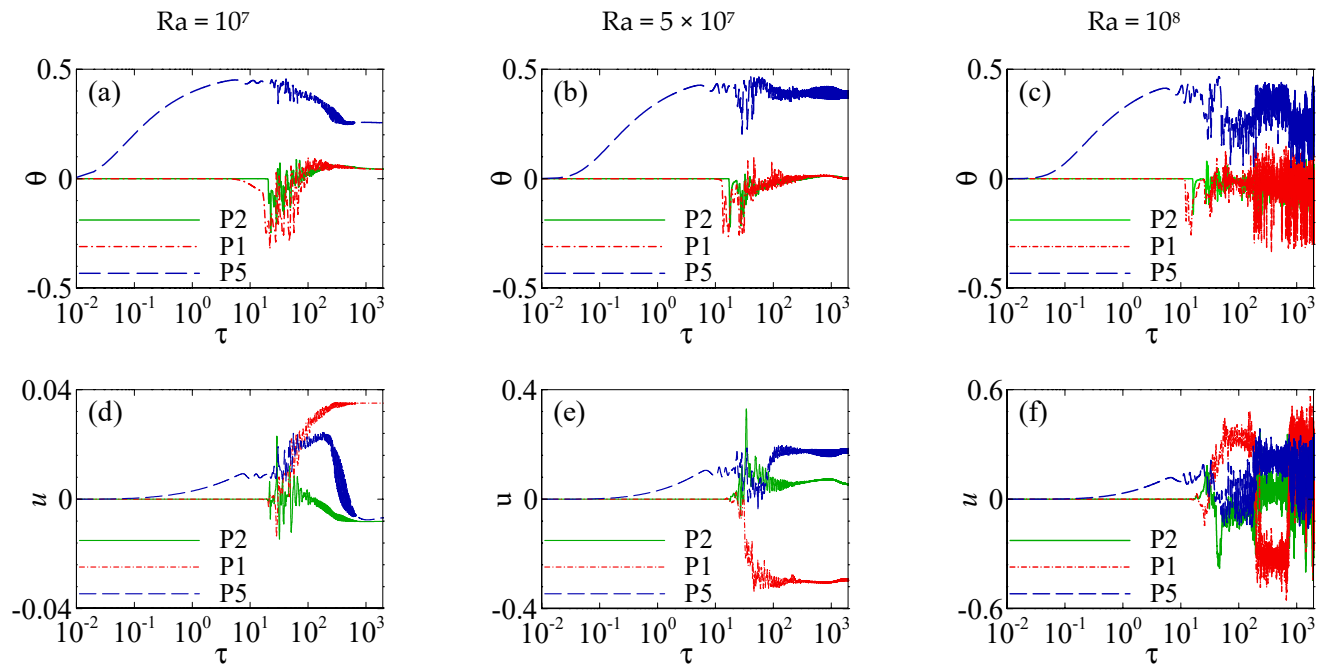


Figure 14. Times series of the temperature and velocity at three distinct points, P_1 (0, 0.825), P_2 (0, 0.46) and P_3 (0.5, 0.255), for (a,d) when $Ra = 10^7$, for (b,e) when $Ra = 5 \times 10^7$, and for (c,f) when $Ra = 10^8$.

5.3.1. Hopf Bifurcation

Figure 15 displays the streamlines and isotherms for $Ra = 10^7$ and $Ra = 2 \times 10^7$ to allow for a more in-depth analysis of the flow in the case of greater Rayleigh numbers. Figure 15a illustrates the more complicated flows throughout the fully developed stage for $Ra = 10^7$, even though it is still steady. However, a deeper inspection of the numerical results reveals that more than two tiny cells alternately occur; as a result, the flow becomes unsteady for $Ra = 2 \times 10^7$ in the fully developed stage. This indicates that there is a Hopf bifurcation between $Ra = 10^7$ and 2×10^7 .

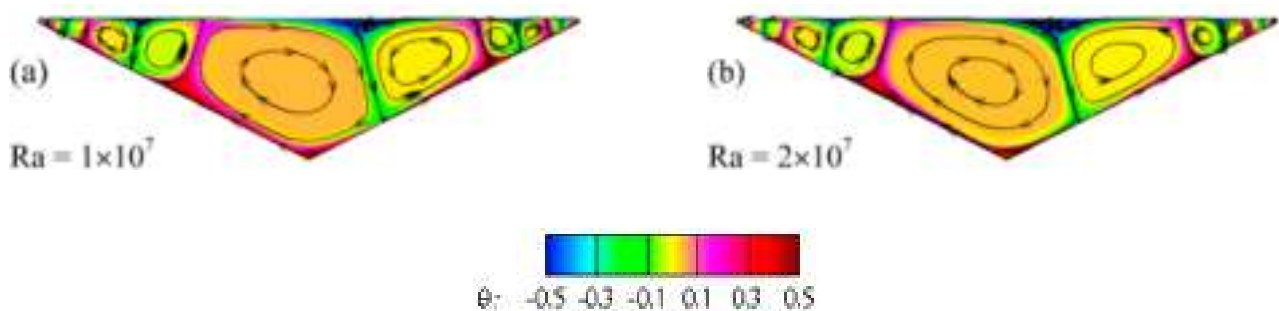


Figure 15. Streamlines and isotherms at $Ra = 10^7$ and $Ra = 2 \times 10^7$.

Figure 16 shows the attractors with values ranging from $\tau = 300$ to 2000 for $Ra = 10^7$ and $\tau = 1000$ to 1500 for $Ra = 5 \times 10^7$ at the defining point P_1 (0, 0.825) in order to facilitate an understanding of the Hopf bifurcation, which takes place during the transition from the steady state to the periodic stage. Figure 16a illustrates that the u - θ plane curve reaches a specific value when $Ra = 10^7$. Figure 16b displays a limit cycle for $Ra = 5 \times 10^7$. As a result, when $Ra = 5 \times 10^7$, a Hopf bifurcation takes place (for further information on the Hopf bifurcation, see [65]).

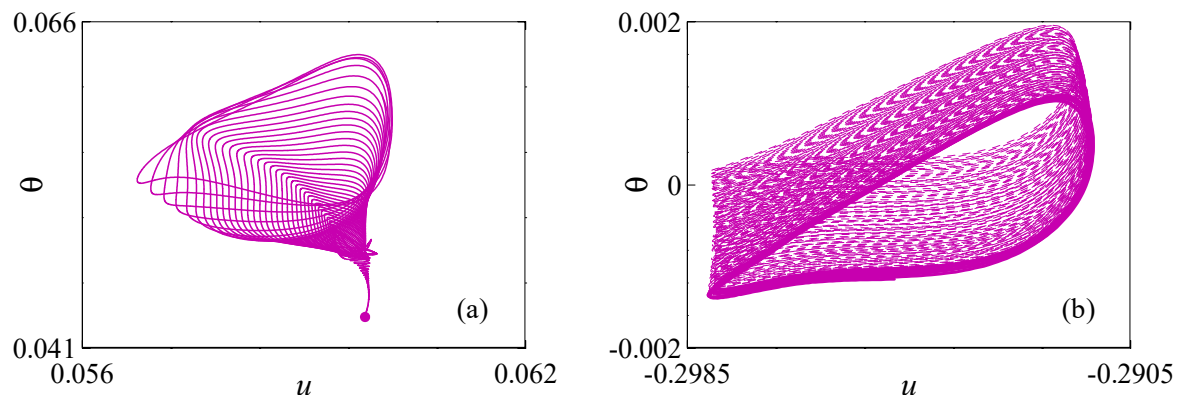


Figure 16. The limit point and limit cycle for (a) $Ra = 10^7$, and for (b) $Ra = 5 \times 10^7$ on the plane of u - θ at point P_1 (0, 0.825).

5.3.2. Chaotic

Figure 17a illustrates that the two cells in the upper right portion of the biggest cell become bigger precisely as the Rayleigh number increases. The biggest cell in the center likewise travels across the right as well as the left sides when $Ra = 10^8$, as seen in Figure 17b. The unsteady flow becomes increasingly complicated, which is known as chaotic.

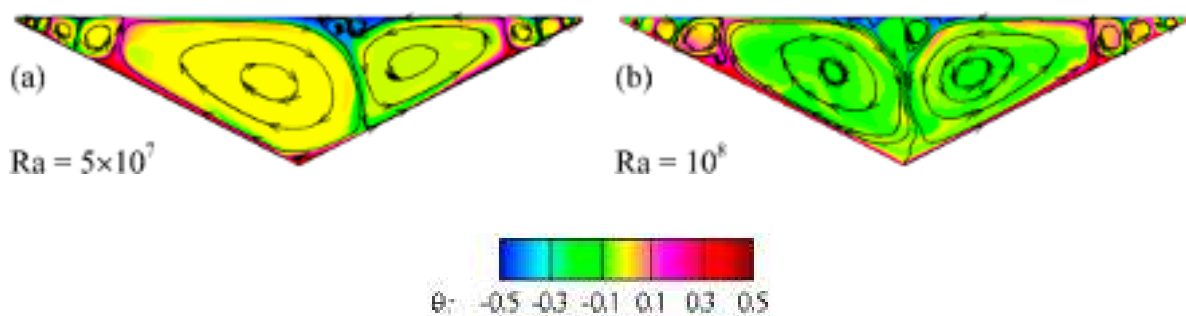


Figure 17. Streamlines and isotherms at $Ra = 5 \times 10^7$ and $Ra = 10^8$.

Figure 18 shows the trajectories through the stage space in the u - θ plane at point P_5 (0.5, 0.255) for $Ra = 5 \times 10^7$ and 10^8 in order to more clearly illustrate how the periodic condition changes to a chaotic state. The observed limit cycle depicted in Figure 18a demonstrates the periodic nature of the unsteady flow at a Rayleigh number of $Ra = 5 \times 10^7$. This finding aligns with the information presented in Figure 16 and further supports the existence of a limit cycle. In Figure 18b, the trajectory for $Ra = 10^8$ shows that the periodic flow changes to chaotic, which takes place between $Ra = 5 \times 10^7$ and 10^8 . For a detailed explanation of the stage-space trajectories, see [66].

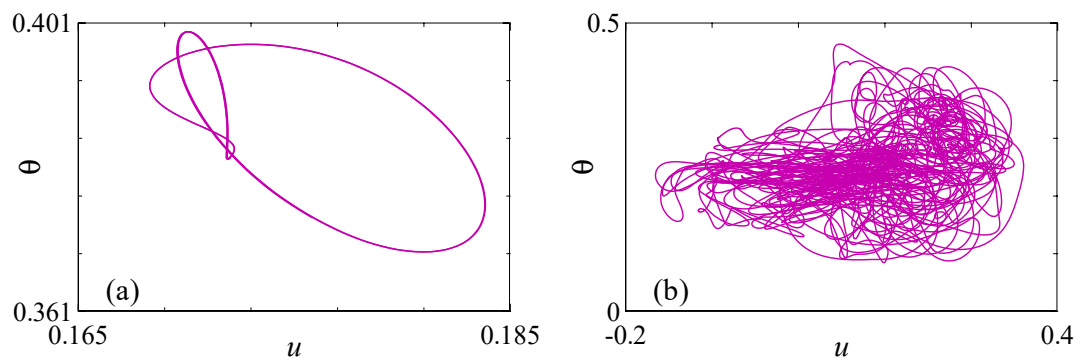


Figure 18. Temperature and x-velocity trajectories in the stage space for the values of $Ra = 5 \times 10^7$ and $Ra = 10^8$ at the point P_5 (0.5, 0.255).

5.4. Temperature and Velocity

Figure 19 displays the temperature and velocity at designated points P_1 (0, 0.825) over time across various Rayleigh numbers. This information is provided to help understand the formation of natural convection flow patterns within the cavity in response to sudden heating from the inclined walls and cooling from the top wall. The simulations were conducted using various Rayleigh numbers, ranging from $Ra = 10^0$ to 10^8 . An observation was made about the various fluctuating flow properties throughout a range of Rayleigh numbers. Figures 5, 6, 9 and 12 depict the isotherms and corresponding streamlines for various Rayleigh numbers, specifically focusing on the case where $A = 0.5$. The observed numerical outcomes for the various Ra values, as depicted in Figure 19, exhibit discernible changes. At the lowest Rayleigh number, convective flow instabilities may first be seen. However, the number of waves and the unsteadiness increase with increasing Rayleigh numbers. Based on the symmetry and continuous flow, it is expected that the flow is weaker and displays symmetric behavior at $Ra = 10^3$. During the transitional stage of fluid flow, it is observed that the flow develops asymmetrically for the Rayleigh numbers of $Ra = 10^4$, 10^5 , and 10^6 . However, as the flow progresses toward the fully developed stage, it stabilizes. Finally, the flow changes into periodic and chaotic states at $Ra = 10^7$ and $Ra = 10^8$, respectively, as shown in Figure 19a. As illustrated in Figure 19b, a similar characteristic can also be seen in the x -velocity at the point P_1 .

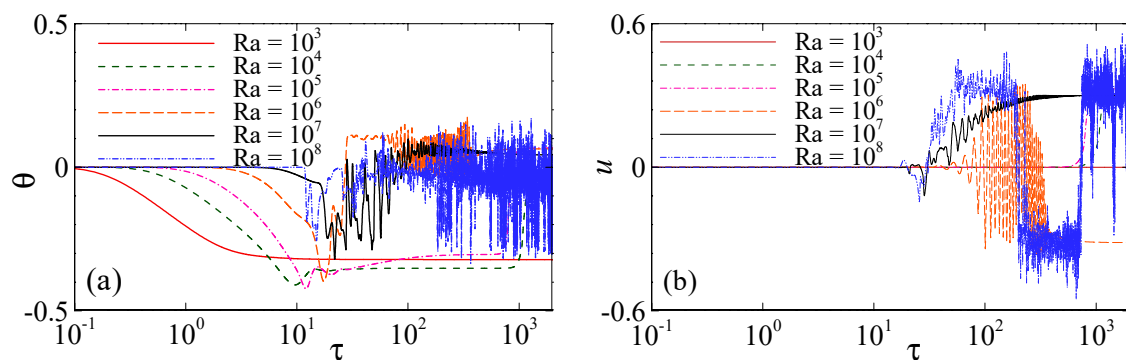


Figure 19. (a) Temperature time series and (b) x -velocity time series for different Rayleigh numbers at point P_1 (0, 0.825).

5.5. Heat Transfer

During the transitional stage, convective heat transfer dominates in the V-shaped cavity. Convective heat transfer is enhanced by the irregular fluctuations and vortices that encourage fluid mixing. For air, the Nusselt number that expresses the proportion of convective to conductive heat transfer is examined. The Nusselt number Nu [60–62] is defined as:

$$Nu = \frac{1}{\ln} \int_{\ln}^{\frac{\partial \theta}{\partial n}} ds. \quad (11)$$

In accordance with Figure 20a, a time series representation has been provided to illustrate the average Nusselt number associated with the inclined wall. This particular measurement has been employed to quantify the amount of heat transfer occurring across the cavity wall. It is important to highlight that Figure 20b additionally presents the representation of the Nu after being normalized by $Ra^{1/4}$. The presence of a significant temperature difference between the fluid and the wall may contribute to the substantial heat transfer phenomenon. Consequently, it is expected that a large value of the Nusselt number (Nu) is observed. The observed phenomenon can be attributed to the simultaneous application of heating and cooling to both the inclined walls and the top wall. Figure 20b illustrates how, as time moves on, throughout the beginning stage, the Nu dramatically decreases, and throughout the fully devolved stage, it gradually attains a value that is either constant or oscillatory, and this value is determined by the Rayleigh number. This

is similar to those in Figure 13, where the Nusselt number oscillates for $Ra \geq 5 \times 10^7$ but remains constant for $Ra = 10^7$. In contrast to those in Figure 20a, it is obvious that the Nu versus the curves in Figure 20b collapse together. As a result, $Nu \sim Ra^{1/4}$ scaling works effectively, considering the current set of Ra values. Figure 20b demonstrates that $NuRa^{-1/4}$ almost maintains a constant value of 1.473 as the Rayleigh number rises, although somewhat decreasingly.

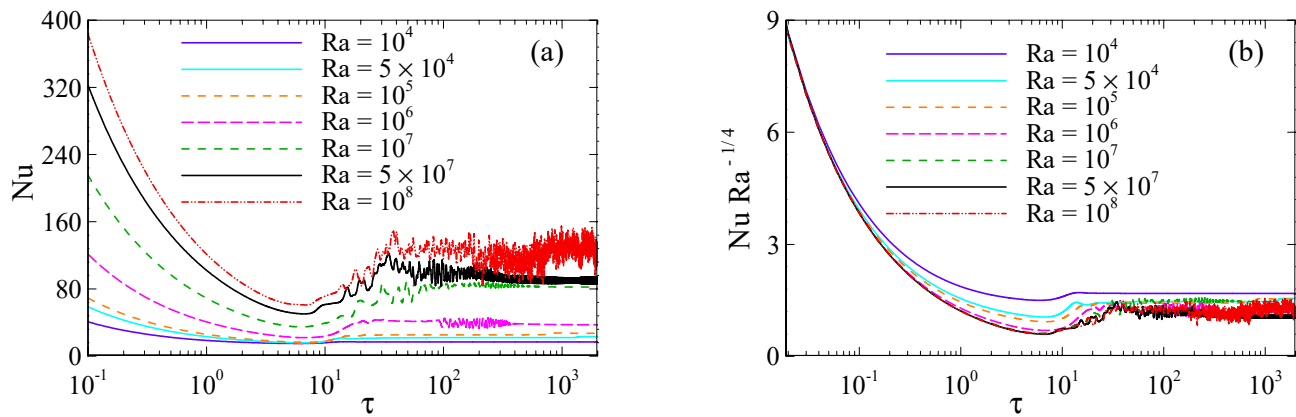


Figure 20. (a) The Nusselt number and (b) the normalized Nusselt number time series for several kinds of Rayleigh numbers.

6. Conclusions

The current investigation delves into the numerical exploration of unsteady 2D natural convection flows within a valley-shaped cavity filled with water, where heating occurs along the inclined walls while cooling is facilitated through the top wall. The study encompasses a broad spectrum of Rayleigh numbers ranging from $Ra = 10^0$ to 10^8 , under the Prandtl number $Pr = 0.71$ and the aspect ratio $A = 0.5$, spanning various non-dimensional times from $\tau = 0$ to $\tau = 5000$. The research delineates the diverse flow structures evolving within the cavity over time and scrutinizes the relationship between heat transfer dynamics and Rayleigh numbers.

Analysis reveals that the progression of natural convection flows manifests in three discernible stages: an initial stage, a transitional stage, and a fully developed stage, following the sudden application of heating through inclined walls and cooling through the top wall. This study examines the symmetric, asymmetric, and unsteady flow patterns characterizing these stages, supported by numerical findings. Specifically, the investigation primarily focuses on elucidating the flow mechanisms across all the stages. It is observed that natural convection flows remain steady at the symmetric state for Rayleigh numbers ranging from $Ra = 10^0$ to 7×10^3 . Beyond $Ra > 7 \times 10^3$, flows exhibit asymmetry. Furthermore, during the asymmetric state, the flow transitions into an unsteady regime in the transitional stage before stabilizing at the fully developed stage for $7 \times 10^3 < Ra < 10^7$. This study highlights that periodic unsteady flows evolve into chaotic states during the transitional stage, reverting to periodic behavior in the developed stage at $Ra = 5 \times 10^7$, while at higher Rayleigh numbers like $Ra = 10^8$, the chaotic flow remains predominant in the unsteady regime.

Additionally, the investigation discusses the notable bifurcations observed in the fully developed states. Detailed analyses, including the power spectral density and stage space trajectories for $Pr = 0.71$, are provided. Numerical studies elucidate the intricacies of heat transfer and demonstrate the influence of the Rayleigh number on both the Nusselt number and flow rate dynamics.

Author Contributions: Conceptualization, S.B. and S.C.S.; investigation, L.R.R.; methodology, L.R.R.; project administration, S.C.S.; software, L.R.R.; supervision, S.B. and S.C.S.; validation, L.R.R.; writing—original draft, S.B. and L.R.R.; writing—review and editing, F.X. and S.C.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: No data is used in this study. Use of Artificial Intelligence: During the preparation of this work, we used ChatGPT in order to improve the language and readability.

Conflicts of Interest: The authors declare no conflict of interest.

Nomenclature

κ	thermal diffusivity
τ	non-dimensional time
$\Delta\tau$	non-dimensional time step
θ	non-dimensional temperature
ν	kinematic viscosity
ρ	Density
β	thermal expansion coefficient
A	aspect ratio
g	gravitational acceleration
Gr	Grashof number
L	half length of the cavity
H	height of the cavity
k	thermal conductivity
l	inclined wall length
n	non-dimensional coordinate normal to the inclined wall
ln	non-dimensional length of the inclined wall
Nu	Nusselt number
P	Pressure
p	non-dimensional pressure
Pr	Prandtl number
Ra	Rayleigh number
Ra_c	critical Rayleigh number for the break of the symmetric state
T	dimensional temperature
T_0	initial temperature
T_c	cold top wall temperature
T_h	hot inclined wall temperature
t	Time
ΔT	temperature difference
U, V	velocity components
u, v	the non-dimensional velocity components
X, Y	coordinates in the horizontal and vertical
x, y	the non-dimensional coordinates in the horizontal and vertical

References

1. Gill, E. The boundary-layer regime for convection in a rectangular cavity. *J. Fluid Mech.* **1966**, *26*, 515–536. [\[CrossRef\]](#)
2. Prandtl, L. *Essentials of Fluid Dynamics*; Blackie: London, UK, 1952; pp. 422–425.
3. Ravi, M.R.R.; Henkes, R.A.W.M.; Hoogendoorn, C.J. On the high-Rayleigh-number structure of steady laminar natural-convection flow in a square enclosure. *J. Fluid Mech.* **1994**, *262*, 325–351. [\[CrossRef\]](#)
4. Xu, F.; Patterson, J.C.; Lei, C. An experimental study of the unsteady thermal flow around a thin fin on a sidewall of a differentially heated cavity. *Int. J. Heat Fluid Flow* **2008**, *29*, 1139–1153. [\[CrossRef\]](#)
5. Cui, H.; Xu, F.; Saha, S.C. Transition to unsteady natural convection flow in a prismatic enclosure of triangular section. *Int. J. Therm. Sci.* **2017**, *111*, 330–339. [\[CrossRef\]](#)
6. Saha, S.C.; Khan, M.M.K. A review of natural convection and heat transfer in attic-shaped space. *Energy Build.* **2011**, *43*, 2564–2571. [\[CrossRef\]](#)
7. Cui, H.; An, H.; Wang, W.; Han, Z.; Hu, B.; Xu, F.; Liu, Q.; Saha, S.C. Numerical Study of Mixed Convection and Heat Transfer in Arc-Shaped Cavity with Inner Heat Sources. *Appl. Sci.* **2023**, *13*, 1029. [\[CrossRef\]](#)

8. Chuhan, I.S.; Li, J.; Guo, Z.; Yaqub, M.; Manan, M.A. Heat Transfer Investigation in Plus-Shaped Enclosure Using Power Law Fluid: A Finite Element Approach. *Appl. Sci.* **2023**, *13*, 11042. [\[CrossRef\]](#)
9. Sarangi, A.; Sarangi, A.; Sahoo, S.S.; Mallik, R.K.; Awad, M.M. Conjugate Radiation and Convection Heat Transfer Analysis in Solar Cooker Cavity Using a Computational Approach. *Energies* **2023**, *16*, 3868. [\[CrossRef\]](#)
10. Gebhart, B.; Jaluria, Y.; Mahajan, R.L.; Sammakia, B. *Buoyancy-Induced Flows and Transport*; Hemisphere: Washington, DC, USA, 1988.
11. Manneville, P. Rayleigh-Bénard convection, thirty years of experimental, theoretical, and modeling work, Dynamics of spatio-temporal cellular structures. In *Dynamics of Spatio-Temporal Cellular Structures: Henri Bénard Centenary Review*; Springer: New York, NY, USA, 2006; Volume 207, pp. 41–65.
12. Bodenschatz, E.; Pesch, W.; Ahlers, G. Recent developments in Rayleigh Benard convection. *Annu. Rev. Fluid Mech.* **2000**, *32*, 709–778. [\[CrossRef\]](#)
13. Sparrow, E.M.; Husar, R.B. Longitudinal vortices in natural convection flow on inclined plates. *J. Fluid Mech.* **1969**, *37*, 251–255. [\[CrossRef\]](#)
14. Batchelor, G.K. Heat transfer by free convection across a closed cavity between vertical boundaries at different temperatures. *Q. Appl. Math.* **1954**, *12*, 209–233. [\[CrossRef\]](#)
15. Eckert, E.R.G.; Carlson, W.O. Natural convection in an air layer enclosed between two vertical plates at different temperatures. *Int. J. Heat Mass Transf.* **1961**, *2*, 106–120. [\[CrossRef\]](#)
16. Elder, J.W. Laminar free convection in a vertical slot. *J. Fluid Mech.* **1965**, *23*, 77–98. [\[CrossRef\]](#)
17. Xu, F.; Patterson, J.C.; Lei, C. An experimental study of the coupled thermal boundary layers adjacent to a partition in a differentially heated cavity. *Exp. Therm. Fluid Sci.* **2014**, *54*, 12–21. [\[CrossRef\]](#)
18. Xu, F.; Patterson, J.C.; Lei, C. Transient natural convection flows around a thin fin on the sidewall of a differentially heated cavity. *J. Fluid Mech.* **2009**, *639*, 261–290. [\[CrossRef\]](#)
19. Patterson, J.C.; Imberger, J. Unsteady natural convection in a rectangular cavity. *J. Fluid Mech.* **1980**, *100*, 65–86. [\[CrossRef\]](#)
20. Patterson, J.C.; Armfield, S.W. Transient features of natural convection in a cavity. *J. Fluid Mech.* **1990**, *219*, 469–497. [\[CrossRef\]](#)
21. De, A.K.; Eswaran, V.; Mishra, P.K. Scaling of heat transport and energy spectra of turbulent Rayleigh-Bénard convection in a large-aspect-ratio box. *Int. J. Heat Fluid Flow* **2017**, *67*, 111–124. [\[CrossRef\]](#)
22. Rahaman, M.M.; Bhowmick, S.; Mondal, R.N.; Saha, S.C. Unsteady natural convection in an initially stratified air-filled trapezoidal enclosure heated from below. *Processes* **2022**, *10*, 1383. [\[CrossRef\]](#)
23. Rahaman, M.M.; Bhowmick, S.; Mondal, R.N.; Saha, S.C. A Computational Study of Chaotic Flow and Heat Transfer within a Trapezoidal Cavity. *Energies* **2023**, *16*, 5031. [\[CrossRef\]](#)
24. Princevac, M.; Fernando, H.J.S. Morning breakup of cold pools in complex terrain. *J. Fluid Mech.* **2008**, *616*, 99–109. [\[CrossRef\]](#)
25. Asan, H.; Namli, L. Laminar natural convection in a pitched roof of triangular cross-section: Summer day boundary conditions. *Energy Build.* **2000**, *33*, 69–73. [\[CrossRef\]](#)
26. Asan, H.; Namli, L. Numerical simulation of buoyant flow in a roof of triangular cross-section under winter day boundary conditions. *Energy Build.* **2001**, *33*, 753–757. [\[CrossRef\]](#)
27. Salmun, H. Convection patterns in a triangular domain. *Int. J. Heat Mass Transf.* **1995**, *38*, 351–362. [\[CrossRef\]](#)
28. Salmun, H. The stability of a single-cell steady-state solution in a triangular enclosure. *Int. J. Heat Mass Transf.* **1995**, *18*, 363–369. [\[CrossRef\]](#)
29. Poulikakos, D.; Bejan, A. The fluid dynamics of an attic space. *J. Fluid Mech.* **1983**, *131*, 251–269. [\[CrossRef\]](#)
30. Poulikakos, D.; Bejan, A. Natural convection experiments in a triangular enclosure. *J. Heat Transf.* **1983**, *105*, 652–655. [\[CrossRef\]](#)
31. Flack, R.D. The experimental measurement of natural convection heat transfer in triangular enclosures heated or cooled from below. *J. Heat Transf.* **1980**, *102*, 770–772. [\[CrossRef\]](#)
32. Holtzman, G.A.; Hill, R.W.; Ball, K.S. Laminar natural convection in isosceles triangular enclosures heated from below and symmetrically cooled from above. *J. Heat Transf.* **2000**, *122*, 485–491. [\[CrossRef\]](#)
33. Lei, C.; Armfield, S.W.; Patterson, J.C. Unsteady natural convection in a water-filled isosceles triangular enclosure heated from below. *Int. J. Heat Mass Transf.* **2008**, *51*, 2637–2650. [\[CrossRef\]](#)
34. Saha, S.C.; Lei, C.; Patterson, J.C. Natural convection in a triangular enclosure subject to periodic thermal forcing. In Proceedings of the 13th International Heat Transfer Conference, Sydney, Australia, 13–18 August 2006.
35. Saha, S.C.; Lei, C.; Patterson, J.C. Effect of aspect ratio on natural convection in attics subject to periodic thermal forcing. *Anziam J.* **2007**, *48*, C618–C632. [\[CrossRef\]](#)
36. Saha, S.C.; Patterson, J.C.; Lei, C. Natural convection in attic-shaped spaces subject to sudden and ramp heating boundary conditions. *Heat Mass Transf.* **2010**, *46*, 621–638. [\[CrossRef\]](#)
37. Saha, S.C.; Patterson, J.C.; Lei, C. Natural convection in attics subject to instantaneous and ramp cooling boundary conditions. *Energy Build.* **2010**, *42*, 1192–1204. [\[CrossRef\]](#)
38. Saha, S.C.; Patterson, J.C.; Lei, C. Natural convection and heat transfer in attics subject to periodic thermal forcing. *International J. Therm. Sci.* **2010**, *49*, 1899–1910. [\[CrossRef\]](#)
39. Lei, C.; Patterson, J.C. Natural convection in a reservoir sidearm subject to solar radiation: Experimental observations. *Exp. Fluids* **2002**, *32*, 590–599. [\[CrossRef\]](#)

40. Lei, C.; Patterson, J.C. Unsteady natural convection in a triangular enclosure induced by absorption of radiation. *J. Fluid Mech.* **2002**, *460*, 181–209. [\[CrossRef\]](#)
41. Lei, C.; Patterson, J.C. Natural convection in a reservoir sidearm subject to solar radiation: A two-dimensional simulation. *Numer. Heat Transf.* **2002**, *42*, 13–32. [\[CrossRef\]](#)
42. Mao, Y.; Lei, C.; Patterson, J.C. Unsteady natural convection in a triangular enclosure induced by absorption of radiation—A revisit by improved scaling analysis. *J. Fluid Mech.* **2009**, *622*, 75–102. [\[CrossRef\]](#)
43. Mao, Y.; Lei, C.; Patterson, J.C. Unsteady near-shore natural convection induced by surface cooling. *J. Fluid Mech.* **2010**, *642*, 213–233. [\[CrossRef\]](#)
44. Mao, Y.; Lei, C.; Patterson, J.C. Characteristics of instability of radiation-induced natural convection in shallow littoral waters. *Int. J. Therm. Sci.* **2010**, *49*, 170–181. [\[CrossRef\]](#)
45. Bednarz, T.P.; Lei, C.; Patterson, J.C. A numerical study of unsteady natural convection induced by iso-flux surface cooling in a reservoir model. *Int. J. Heat Mass Transf.* **2009**, *52*, 56–66. [\[CrossRef\]](#)
46. Whiteman, C.D. Breakup of temperature inversions in deep mountain valleys: Part I. Observations. *J. Appl. Meteorol.* **1982**, *21*, 270–289. [\[CrossRef\]](#)
47. Roach, W.T.; Brown, R.; Caughey, S.J.; Garland, J. The physics of radiation fog: Part I: A field study. *Q. J. R. Meteorol. Soc.* **1976**, *102*, 313–333.
48. Pilie, R.J.; Mack, E.J.; Kocmond, W.C.; Rogers, C.W.; Eadie, W.J. The life cycle of valley fog. Part I: Micrometeorological characteristics. *J. Appl. Meteorol.* **1975**, *14*, 347–363. [\[CrossRef\]](#)
49. Fuzzi, S.; Facchini, M.C. The Po Valley fog experiment 1989, an overview. *Tellus* **1992**, *44B*, 448–468. [\[CrossRef\]](#)
50. Jiusto, J.E.; Lala Grieser, J.G.G. *Radiation Fog Field Programs Recent Studies*; Publication no. 869; Atmospheric Sciences Research Center, State University of New York (ASRC, SUNY): New York, NY, USA, 1983.
51. Zihua, L.; Limin, Z.; Qinghong, Z. The physical structure of the winter fog in Chongqing metropolitan area and its formation process. *Acta Meteorol. Sin.* **1994**, *8*, 316–328.
52. Haiden, T.; Whiteman, C.D. Katabatic flow mechanisms on a low-angle slope. *J. Appl. Meteorol.* **2005**, *44*, 113–126. [\[CrossRef\]](#)
53. Catalano, F.; Cenedese, A. High-resolution numerical modeling of thermally driven slope winds in a valley with strong capping. *Am. Meteorol. Soc.* **2010**, *49*, 1859–1880. [\[CrossRef\]](#)
54. Clark, P.A.; Hopwood, W.P. One-dimensional site-specific forecasting of radiation fog. Part 1: Model formulation and idealized sensitivity studies. *Meteorol. Appl.* **2001**, *8*, 279–286. [\[CrossRef\]](#)
55. Baumbach, G.; Vogt, U. Experimental determination of the effect of mountain-valley breeze circulation on air pollution in the vicinity of Freiburg. *Atmos. Environ.* **1999**, *33*, 4019–4027. [\[CrossRef\]](#)
56. Lehner, M.; Gohm, A. Idealized simulations of daytime pollution transport in a steep valley and its sensitivity to thermal stratification and surface albedo. *Bound.-Layer Meteorol.* **2010**, *134*, 327–351. [\[CrossRef\]](#)
57. Liu, H.; Liang, B.; Zhu, F.; Zhang, B.; Sang, J. Water-tank experiment on the thermal circulation induced by the bottom heating in an asymmetric valley. *Adv. Atmos. Sci.* **2004**, *21*, 536–546. [\[CrossRef\]](#)
58. Kenjeres, S. Heat transfer enhancement induced by wall inclination in turbulent thermal convection. *Phys. Rev. E* **2015**, *92*, 053006. [\[CrossRef\]](#)
59. Kimura, F.; Noguchi, T.; Kitamura, K. Fluid flow and heat transfer of natural convection around heated V-shaped cavity. *Mech. Eng. J.* **2016**, *3*, 16–00330. [\[CrossRef\]](#)
60. Bhowmick, S.; Saha, S.C.; Qiao, M.; Xu, F. Transition to a chaotic flow in a V-shaped triangular cavity heated from below. *Int. J. Heat Mass Transf.* **2019**, *128*, 76–86. [\[CrossRef\]](#)
61. Bhowmick, S.; Xu, F.; Molla, M.M.; Saha, S.C. Chaotic phenomena of natural convection for water in a V-shaped enclosure. *Int. J. Therm. Sci.* **2022**, *176*, 107526. [\[CrossRef\]](#)
62. Bhowmick, S.; Xu, F.; Zhang, X.; Saha, S.C. Natural convection and heat transfer in a valley-shaped cavity filled with initially stratified water. *Int. J. Therm. Sci.* **2018**, *128*, 59–69. [\[CrossRef\]](#)
63. Patankar, S.V. *Numerical Heat Transfer and Fluid Flow*; Hemisphere Publishing Corporation: New York, NY, USA, 1980.
64. Saha, S.C. Natural Convection Adjacent to an Inclined Flat Plate and in an Attic Space under Various Thermal Forcing Conditions. Ph.D. Thesis, School of Engineering and Physical Sciences, Townsville, Australia, 2009.
65. Drazin, P.G. *Introduction to Hydrodynamic Stability*; Cambridge University Press: Cambridge, UK, 2002.
66. McLaughlin, J.B.; Orszag, S.A. Transition from periodic to chaotic thermal convection. *J. Fluid Mech.* **1982**, *122*, 123–142. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Modeling the Properties of Magnetostrictive Elements Using Quantum Emulators

Edvard Karpukhin ¹, Alexey Bormotov ^{1,*} and Luiza Manukyan ²

¹ Faculty of Industrial Technologies, Penza State Technological University, pr. Baidukova/ul. Gagarina, 1a, 11, 440039 Penza, Russia; edvar1@rambler.ru

² Institute of Aerospace Engineering, Samara National Research University, ul. Moskovskoe Highway 34, 443086 Samara, Russia; lu-iza03@mail.ru

* Correspondence: aleks21618@yandex.ru

Abstract: The article discusses mathematical and numerical methods for modeling magnetostrictive multielectronic systems based on a combination of quantum and classical methods. The algorithm development suitable for the investigation of magnetostrictive phenomena at the micro level using the classical-quantum method implemented on a modern classical computer is justified. The algorithms and structure of the software package are given. The adequacy of the quantum-classical method is verified by comparing the calculated results of the properties of known magnetostrictive materials with the real properties of magnetostrictive alloys.

Keywords: quantum-classical numerical methods; investigation of magnetostrictive phenomena at the micro level

1. Introduction

When modeling magnetostrictive phenomena at the micro level, the problem of solving the Schrodinger equation (SE) arises. Currently, many numerical methods for solving it are known [1–10]. Depending on the type of computer designed for their implementation, all these methods can be divided into classical and quantum. Classical methods are designed to be implemented on classical computers and are based on the classical approach to the numerical solution of differential equations. They allow you to effectively solve a wide range of tasks in an acceptable time and with high enough accuracy. However, in many cases, for example, when modeling polyatomic systems, the use of such an approach leads either to an excessively large error or to an exponentially increasing solution time, unacceptable with the modern power of classical computers [11–19]. In scientific papers [20–27], modifications of classical methods are proposed that can significantly increase accuracy and reduce calculation time, but still this problem remains relevant.

Another fundamentally new and actively developing way of numerically solving the SE is the use of quantum methods [11,12,28–31]. Such methods are based on algorithms designed to be performed by quantum computers, which, as is known [23,32–36], are more effective for solving such problems. The main problem with the use of quantum computing is the insufficient level of modern technological development in this direction: the lack of available fully functional quantum computers, and the small number of known effective quantum algorithms and scientific works in this direction. However, existing quantum computer emulators make it possible to develop and test new quantum numerical methods for solving particular problems of modeling polyatomic systems and to substantiate the theoretical possibility of using these methods in the general case in the future, when full-fledged quantum computers appear [37–41].

It is also possible to apply a combined approach to modeling quantum systems using quantum-classical methods using the resources of conventional computers and quantum emulators. The algorithms of quantum-classical numerical methods (QCNM) are designed



Citation: Karpukhin, E.; Bormotov, A.; Manukyan, L. Modeling the Properties of Magnetostrictive Elements Using Quantum Emulators. *Computation* **2024**, *12*, 147. <https://doi.org/10.3390/computation12070147>

Academic Editor: Rafael Lahoz-Beltra

Received: 7 May 2024

Revised: 24 June 2024

Accepted: 26 June 2024

Published: 15 July 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

to be executed on a classical computer in a software environment equipped with a classical computing module and a module for emulating quantum computing, with the possibility of exchanging data between them (Figure 1).

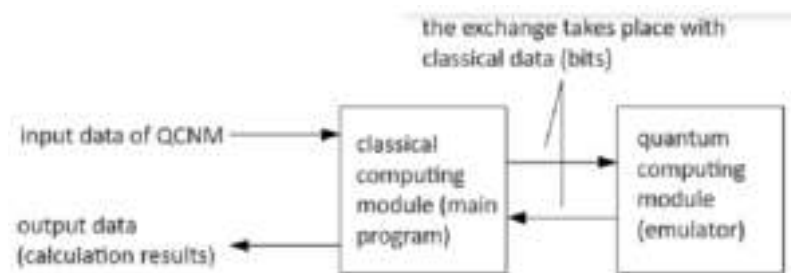


Figure 1. Simplified structure of the software package for the implementation of the QCNM.

This method of modeling will make it possible to use the most useful properties of quantum algorithms to speed up the work and increase the efficiency of classical ones.

The main task considered in this article is the development of a QCNM algorithm suitable for the investigation of magnetostrictive phenomena by their mathematical modeling at the micro level using a classical computer. At the same time, it should be borne in mind that the capabilities of a classical computer for the implementation of quantum computing emulators are currently severely limited by the amount of quantum memory, which is a maximum of about 400 qubits (when using cloud technologies).

The proposed QCNM algorithm is implemented as part of a software package (Figure 1), and with its help, a study of magnetostrictive phenomena in *Ni-Mn-Ga* Geisler alloys has been carried out [42–47].

2. Materials and Methods

Atoms with N electrons differ from the hydrogen atom and hydrogen-like systems. Firstly, the Hamiltonian takes into account the contributions to the kinetic and potential energy of all N electrons, and a summand describing the interelectronic interaction arises in the potential energy. Secondly, the wave function must obey the Pauli principle—it must be antisymmetric when rearranging the coordinates of any pair of electrons.

This leads to the fact that the solution of the Schrodinger equation for multielectronic atoms becomes much more complicated. It cannot be obtained analytically and requires approximate numerical solution methods [13,20,48,49].

However, the properties of multielectronic systems can be described qualitatively quite simply if we take advantage of the fact that certain relations related to the angular momentum must be fulfilled for the atom.

1. Each state of the atom must have a certain value of the total angular momentum.
2. The total moment of all electrons in an atom must obey the rules of quantization and spatial quantization.
3. Since the spin of electrons is also the moment of momentum, it should also contribute to the total moment of the atom.
4. The moments and spins of individual electrons should be combined so that the total moment remains an observable quantity and obeys the postulates and equations of quantum mechanics.

The total moment is associated with the moments of individual electrons, but since the moment of each electron is ambiguously oriented in space (spatial quantization does not allow you to uniquely define all projections of the vector), the total moment, according to the superposition principle, can become a combination of any possible orientations of the vectors of individual moments. These principles were used in the modeling of multielectronic Geisler alloys of the *Ni-Mn-Ga* type.

Consider the Schrodinger time equation:

$$\hat{H}\Psi = \left(-\frac{\hbar^2}{2m}\nabla^2 + U(\bar{r}, t) \right) \Psi(\bar{r}, t) = j\hbar \frac{\partial \Psi(\bar{r}, t)}{\partial t}, \quad (1)$$

where $\hat{H}$ —the Hamiltonian; $\Psi(\bar{r}, t)$, $U(\bar{r}, t)$ —the wave function and potential energy of a particle located at a point in space with coordinates (x, y, z) and a radius-vector $\bar{r} = x\bar{i} + y\bar{j} + z\bar{k}$ at a time t ; $\nabla = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ —the Laplace operator; m —particle mass; and $\hbar$ —reduced Planck constant.

The solution of Equation (1) on a quantum computer can be performed by the Zalka–Wiesner method [15], which is applicable, including for polyatomic systems. The general scheme of the numerical solution of the SE by this method is as follows:

1. Write down the formally exact solution of the SE (1) Ψ in the form of a vector $|\psi(t)\rangle$:

$$|\psi(t)\rangle = e^{-j\hat{H}t} |\psi(0)\rangle. \quad (2)$$

2. Set the initial state $|\psi(0)\rangle$, the time point t_f , or whichever is required to determine the solution and discretization step Δt , such that $i\Delta t = t_f$, for some integer i . To accept $|\psi_0\rangle = |\psi(0)\rangle$, $i = 0$.
3. Construct a quantum circuit of $U_{\Delta t}$ in the form:

$$U_{\Delta t} = \text{IQFT} \exp\left(-\frac{j}{\hbar} \frac{p^2}{2m} \Delta t\right) \text{QFT} \exp\left(-\frac{j}{\hbar} V(\bar{r}) \Delta t\right), \quad (3)$$

where $V(\bar{r})$ —potential energy of a quantum particle system, and QFT, IQFT—forward and reverse quantum Fourier transform.

4. The beginning of the iterative process.
5. At one iteration step, the value $|\psi_{i+1}\rangle = U_{\Delta t} |\psi_i\rangle$ is calculated, and the counter i is incremented by one.
6. The iterative process stops when the condition $i\Delta t \geq t_f$ is met. The result is returned to $|\psi(t_f)\rangle = |\psi_i\rangle$.

This algorithm is performed on a quantum computer in polynomial time, is the most versatile, and is suitable for modeling a wide range of quantum systems, including magnetostrictive phenomena. The result is obtained with an error, the value of which can be obtained from the estimate:

$$\left| \langle \psi(t_f) | e^{-j\hat{H}t_f} | \psi(0) \rangle \right|^2 \geq 1 - \varepsilon. \quad (4)$$

The Zalka–Wesner method without modifications can be implemented on a classical computer using the software package in Figure 1, but it is suitable only for modeling a quantum system with a small number of particles. To increase its efficiency, it is proposed to use the QCNM based on the Trotter–Suzuki transformation [24,50–52], which can be expressed by the following recursive formulas:

$$\left(\prod_{i=1}^m e^{-jH_i t/2n} \prod_{i=m}^1 e^{-jH_i t/2n} \right)^n = e^{-jHt} + O\left(\frac{m^3 t^3}{n^2}\right) = U_2(t), \quad (5)$$

$$U_{2k}(t) = [U_{2k-2}(s_k t)]^2 U_{2k-2}([1 - 4s_k]t) [U_{2k-2}(s_k t)]^2 = e^{-jHt} + O\left(\frac{(mt)^{2k+1}}{n^2}\right), \quad s_k = \left(4 - 4^{1/(2k-1)}\right)^{-1}.$$

The use of this transformation makes it possible to increase the accuracy of the approximation using the same computing resources. The effectiveness of the Trotter–Suzuki method can be clearly and simplistically demonstrated using the illustration in Figure 2. If we conditionally assume that the quantum algorithm must move from state 1 to state 2, and the lengths of the sides of the triangle A , B , and $U_{2k}(t)$ —the total amount of computing re-

sources needed for this transition, then the shortest path is $U_{2k}(t)$, which is a Trotter–Suzuki transform of order $2k$.

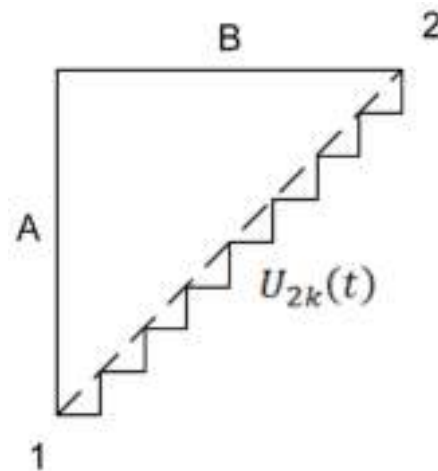


Figure 2. The principle of operation of the Trotter–Suzuki transformation.

An important point in the implementation of the proposed QCNM is the use of the quantum Fourier transform (QFT). The peculiarity of QFT in comparison with the usual Fourier transform is that it does not speed up the transformation itself, but makes it possible to more efficiently search for initial approximations for the eigenvalues of unitary operators, which are needed for other numerical methods.

In an orthonormal basis $|0\rangle, \dots, |N-1\rangle$ QFT is a linear operator acting as follows [14]:

$$|x\rangle \rightarrow \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} e^{-2\pi j k x / N} |k\rangle. \quad (6)$$

Or if you accept $N = 2^n$, where n —an integer, and represent the state $|x\rangle$ as a binary number $x = x_1 x_2 \dots x_n$, (6), then you can rewrite it in a more convenient form for practical purposes:

$$\frac{|x\rangle = |x_1 x_2 \dots x_n\rangle \rightarrow \left(|0\rangle + e^{-2\pi j 0.x_n} |1\rangle \right) \left(|0\rangle + e^{-2\pi j 0.x_{n-1} x_n} |1\rangle \right) \dots \left(|0\rangle + e^{-2\pi j 0.x_1 x_2 \dots x_n} |1\rangle \right)}{2^{n/2}}, \quad (7)$$

where the notation of type $0.x_1 x_2 \dots x_n$ denotes a binary fraction $\frac{x_1}{2} + \frac{x_2}{4} + \dots + \frac{x_n}{2^n}$.

Representation (7) allows to construct a simple quantum scheme for calculating QFT based on quantum gates of Hadamard H and controlled phase shifts $R_k = R_\phi$ at $\phi = \pi 2^{1-k}$, shown in Figure 3.

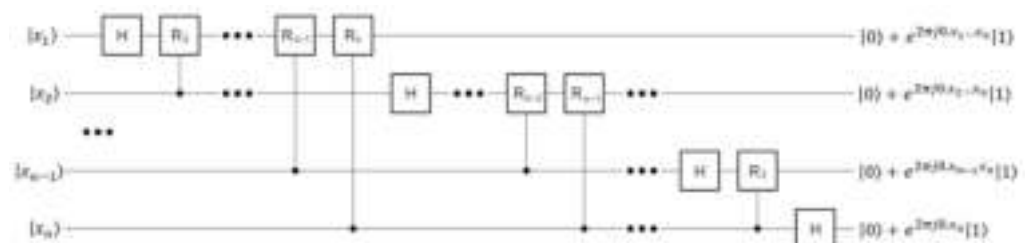


Figure 3. A quantum scheme for calculating the QFT.

Thus, to conduct studies of magnetostrictive phenomena at the microlevel, the proposed QCNM was used, based on the modified Zalka–Wesner method, which can be implemented on a classical computer due to quantum simulation of the Trotter–Suzuki transformation and QFT.

The implementation of the software package for conducting research in accordance with the proposed algorithm was performed in the MS Visual Studio environment using the C#—programming languages for the classical computing module and Q#—for the quantum emulator module (see Figure 1). The scheme of the algorithm of the classical computing module is shown in Figure 4.

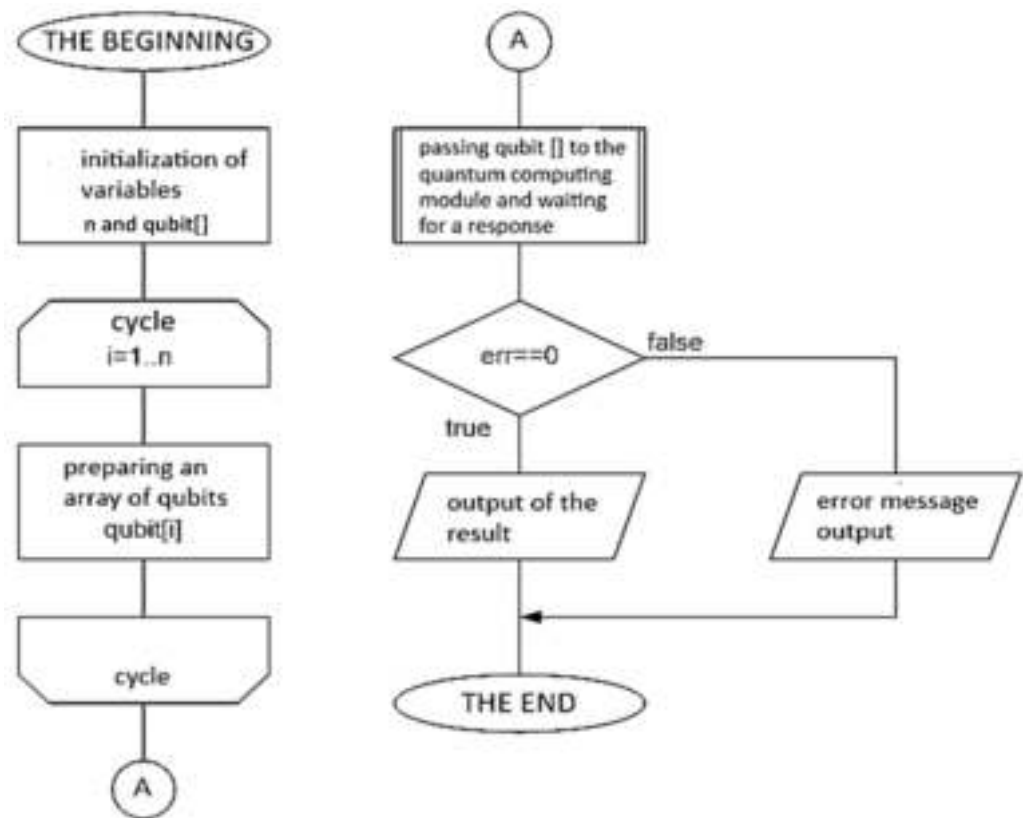


Figure 4. The scheme of the algorithm of the classical computing module.

The classical computing module is the main module of the program, and its task is to manage the QCNM, namely, to prepare the initial data, transfer them to the quantum emulator module and process the results obtained. At the initial stage of the module algorithm, an array of qubits $[i]$ of length n is initialized, which is necessary for modeling the quantum system under consideration. Next, the elements of this array are prepared, depending on the type of Hamiltonian $\hat{H}$ of SE (1), and transferred to the quantum computing module for simulation on a quantum emulator. After receiving a response from this, in the absence of an error flag, output data are generated based on the obtained simulation results.

The scheme of the algorithm of the quantum computing module is shown in Figure 5.

At the initial stage, this module converts the resulting qubit array into the corresponding Hamiltonian matrix and simulates its quantum evolution using the proposed modified Zalka–Wiesner method. At the same time, to emulate quantum operations, it is possible to use built-in Q# language operations, such as *Exp*—to calculate the matrix exponent, *DcomposedIntoTimeStepsCA*—to perform a single step in the evolution of the Hamiltonian, etc. [12,53,54]. The main part of the quantum module algorithm is the iterative process of calling the Trotter–Suzuki transformation procedure, in accordance with expression (5). In this case, the QFT algorithm is used, based on formulas (6) and (7), implemented according to the scheme in Figure 3. Upon completion of the iterative process, an array of “result” is formed containing the numerical data of the simulation and the variable “err”, equal to 1 in case of errors in the operation of the module and 0 in their absence. These data are passed back to the classical computing module for processing.

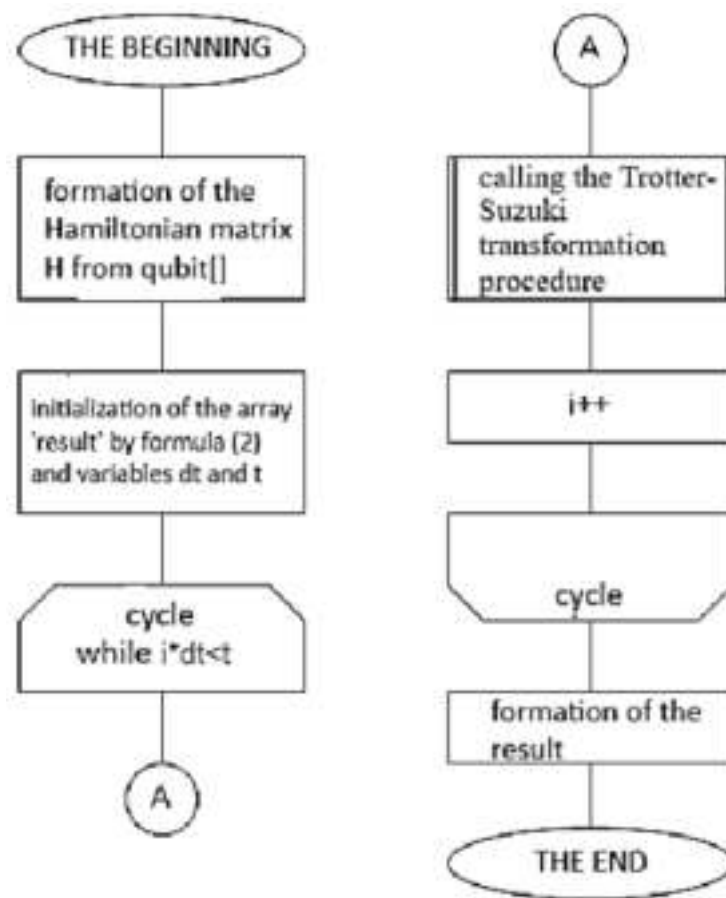


Figure 5. The scheme of the algorithm of the quantum computing module.

3. Results

Using the developed software package implementing the algorithm of the proposed QCNM, we conducted a study of a magnetostrictive Geisler alloy of the *Ni-Mn-Ga* type, the crystal lattice structure of which is shown in Figure 6. In the case when the atom of matter in question contains electrons, it is necessary to take into account the influence of each of them on the total kinetic and potential energy, as well as the interelectronic interaction between them. The Hamiltonian of the SE (1) for such an alloy, as is known [42,53–57], can be written in the Born–Oppenheimer approximation as:

$$\hat{H} = \sum_{i=1}^N \left[-\frac{\hbar^2}{2m_i} \left(\frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2} \right) - k \frac{Ze^2}{|\bar{r}_i|} + k \sum_{j=1}^{i-1} \frac{e^2}{|\bar{r}_i - \bar{r}_j|} \right], \quad (8)$$

where m_i —the mass of the i -particle, $\hbar$ —the reduced Planck constant, Z —the total charge of the atom, e —the charge of the electron, k —the proportionality coefficient in Coulomb's law, and $\bar{r}_i$ —the radius-vectors of electrons.

As a result of modeling the microparticle system described by the Hamiltonian (8) using the QCNM algorithm on a classical computer, numerical data of its energy characteristics were obtained. When studying magnetostrictive phenomena, the values of possible magnetostrictive deformations of a material, depending on its structure, temperature, the presence of an external magnetic field, etc., are of the greatest interest. Based on the numerical data obtained, using a well-known technique, given for example in [42], graphs of some of these dependencies were constructed. Figure 7 shows graphs of the dependence of magnetostrictive deformations e_i on the modulus of elasticity a (Figure 7a) and temperature t (Figure 7b) when the alloy is in magnetic fields of different intensity.

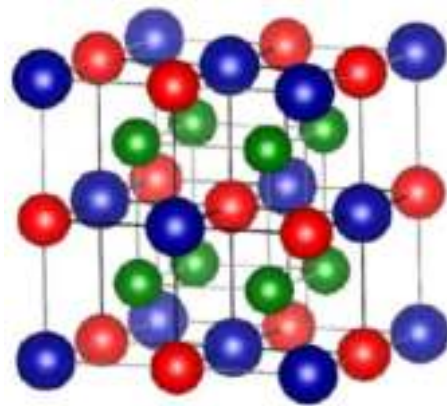


Figure 6. The crystal lattice cell of the *Ni-Mn-Ga* alloy [42].

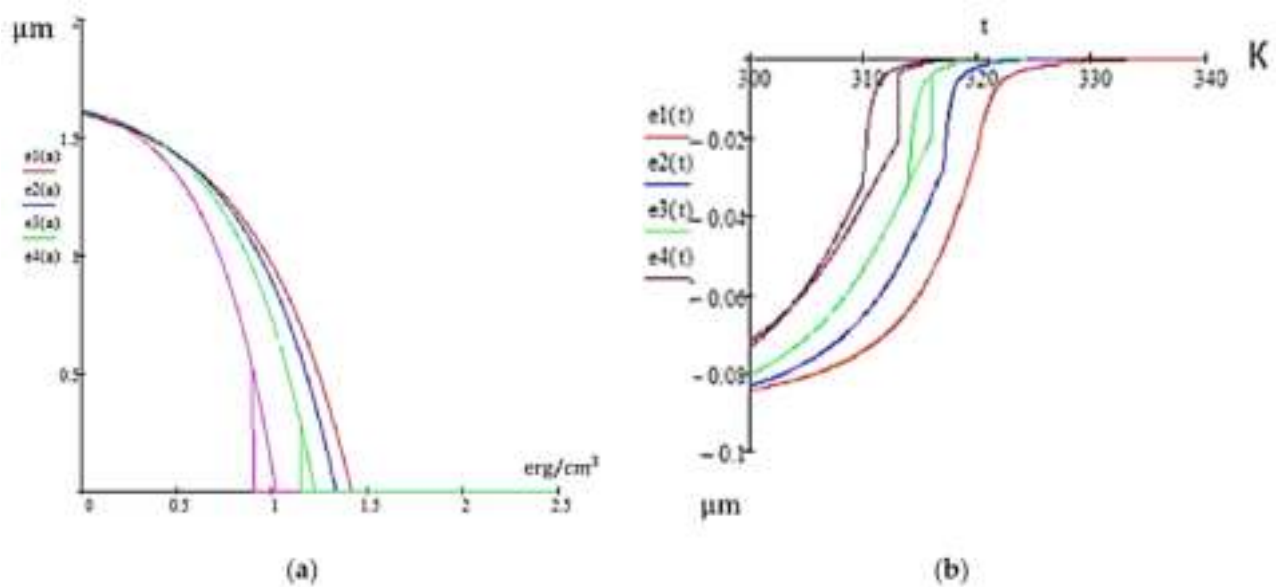


Figure 7. Dependency graphs of magnetostrictive deformations of *Ni-Mn-Ga* alloy at constant magnetic fields of 5, 14, 25, and 30 T. (a) graphs of the dependence of magnetostrictive deformations on the modulus of elasticity; (b) graphs of the dependence of magnetostrictive deformations on temperature.

The temperature and elastic dependences of deformations at constant magnetic fields of 5, 14, 25, and 30 T, shown in Figure 7, show that the hysteresis width decreases with increasing value of the external field, and in a field with an induction value of more than 25 T, the hysteresis disappears. A comparison of the graphs in Figure 7 with the known results of modeling *Ni-Mn-Ga* alloys and experimental data of similar characteristics shown in Figure 8 shows deviations of no more than 5% when modeling the dependence of deformation on the modulus of elasticity and no more than 7% when constructing its dependence on temperature, at the boundary points of the hysteresis loop or at the inflection points of graphs, in its absence. This indicates the correctness of the QCNM algorithm in the investigation of magnetostrictive phenomena.

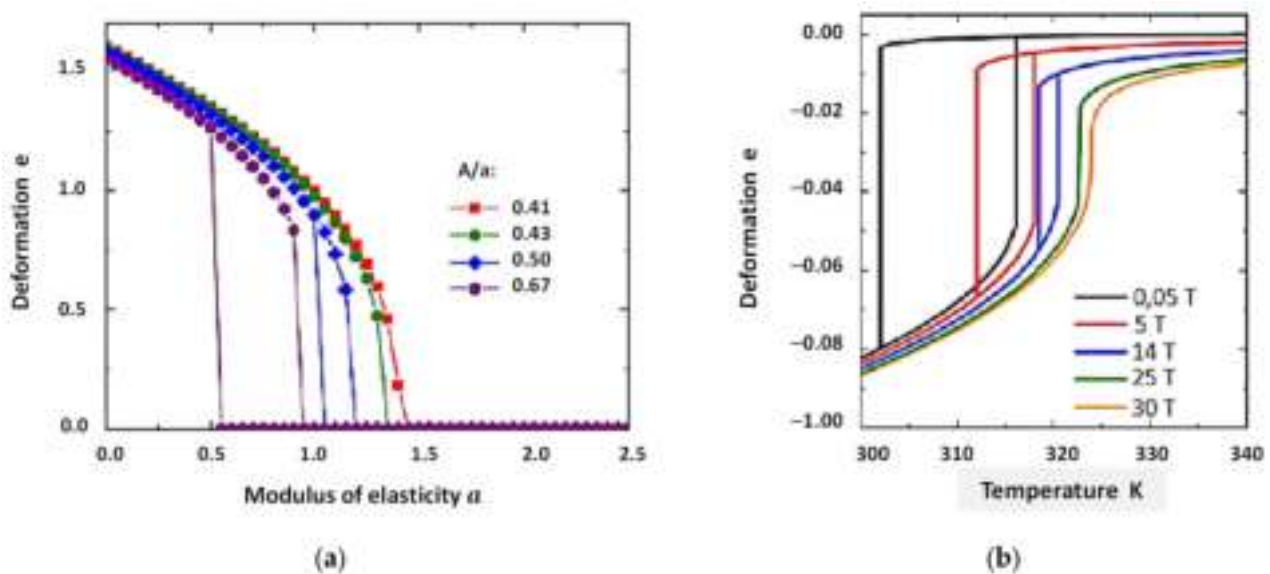


Figure 8. Experimental characteristics of magnetostrictive deformations of *Ni-Mn-Ga* alloy [42–47]. (a) graphs of the dependence of magnetostrictive deformations on the modulus of elasticity; (b) graphs of the dependence of magnetostrictive deformations on temperature.

4. Conclusions

Within the framework of the study, the development of a QCNM algorithm was carried out, suitable for modeling polyatomic systems, including magnetostrictive phenomena, at the micro level, suitable for modeling the magnetostrictive properties [58–61]. The approach used to solve the problem is the most effective because it allows one to combine the capabilities of modern numerical methods, complementing them with the advantages of quantum methods.

The proposed QCNM algorithm is based on a modification of the Zalka–Wesner method for solving the Schrodinger equation by using the Trotter–Suzuki transform and the quantum Fourier transform implemented on a quantum emulator.

During the research of magnetostrictive phenomena in *Ni-Mn-Ga* Geisler alloys using the developed software package, similar results were obtained to the known experimental ones, which confirms the correctness of the algorithm of the proposed QCNM.

Author Contributions: Conceptualization, A.B. and E.K.; methodology, A.B. and E.K.; software, A.B. and E.K.; validation, A.B. and E.K.; formal analysis, A.B. and E.K.; investigation, A.B. and E.K.; resources, A.B., E.K., and L.M.; data curation, A.B. and E.K.; writing—original draft preparation, A.B., E.K., and L.M.; writing—review and editing, A.B., E.K., and L.M.; visualization, A.B. and E.K.; supervision, A.B.; project administration, A.B. and L.M.; funding acquisition, A.B. All authors have read and agreed to the published version of the manuscript.

Funding: The research was carried out at the expense of a grant from the Russian Science Foundation (project № 23-29-00207), <https://rscf.ru/project/23-29-00207/> (accessed on 13 February 2023).

Data Availability Statement: The raw data cannot be shared at this time as the data also form part of an ongoing study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Barabanov, A.L. *Quantum Mechanics*; MIPT: Dolgoprudny, Russia, 2015; 97p.
2. Kashurnikov, V.A.; Krasavin, A.V. *Computational Methods in Quantum Physics*; MEPHI: Moscow, Russia, 2005; 412p.
3. Linev, A.V. *Technique for Solving the Nonstationary Schrodinger Equation Using a Parallel Version of the Runge-Kutta Method*; Linev, A.V., Satanin, A.M., Eds.; Nizhny Novgorod State University: Nizhny Novgorod, Russia, 2014; 89p.

4. Serov, V.V. *Numerical Methods for Solving Quantum Mechanical Problems*; Saratov State University: Saratov, Russia, 2008; 39p.
5. Chuluunbaatar, O. Multilayer schemes for the numerical solution of the nonstationary Schrödinger equation by the finite element method. *Bull. RUDN Math. Ser. Comput. Sci. Phys.* **2008**, 69–84. Available online: <https://cyberleninka.ru/article/n/mnogosloynnye-shemy-dlya-chislennogo-resheniya-nestatsionarnogo-uravneniya-shryodinger-a-metodom-konechnyh-elementov/viewer> (accessed on 6 May 2024).
6. Liu, Y.; Ran, M.; Zhang, L. Hamiltonian-preserving schemes for the two-dimensional fractional nonlinear Schrödinger wave equations. *Comput. Math. Appl.* **2023**, 150, 54–69. [[CrossRef](#)]
7. Zhang, Q.; Duan, J. Constraint minimization problem of the nonlinear Schrödinger equation with the Anderson Hamiltonian. *J. Math. Anal. Appl.* **2024**, 538, 128360. [[CrossRef](#)]
8. Chen, Y.P.; Wei, E.I.; Jiang, L.; Meng, M.; Wu, Y.M.; Chew, W.C. A unified Hamiltonian solution to Maxwell–Schrödinger equations for modeling electromagnetic field–particle interaction. *Comput. Phys. Commun.* **2017**, 215, 63–70. [[CrossRef](#)]
9. Lorin, E.; Yang, X. Quasi-optimal domain decomposition method for neural network-based computation of the time-dependent Schrödinger equation. *Comput. Phys. Commun.* **2024**, 299, 109129. [[CrossRef](#)]
10. Gabrick, E.C.; Sayari, E.; de Castro, A.S.M.; Trobia, J.; Batista, A.M.; Lenzi, E.K. Fractional Schrödinger equation and time dependent potentials. *Commun. Nonlinear Sci. Numer. Simul.* **2023**, 123, 107275. [[CrossRef](#)]
11. Ignatov, S.K. *Quantum Chemistry. Quantum Mechanics and the Structure of the Atom*; Nizhny Novgorod State University: Nizhny Novgorod, Russia, 2015; 113p.
12. Kaiser, S.; Granad, K. *Study Quantum Computing in Python and Q#*; DMK Press: Moscow, Russia, 2021; 430p.
13. Karpukhin, E.V.; Bormotov, A.N. Numerical methods to calculate magnetic fields of magnetostriction level converter EDP Sciences. In Proceedings of the International Scientific and Practical Conference “Environmental Risks and Safety in Mechanical Engineering” (ERSME-2023), Rostov-on-Don, Russia, 1–3 March 2023; Volume 376, p. 01101. [[CrossRef](#)]
14. Bormotov, A.N.; Orlov, D.I.; Khnyreva, E.S.; Bratkova, M.E. Experience in using the algorithm for unloading a flywheel engine using magnetic actuators to ensure stabilization of the Aist-2D remote sensing satellite. *Mod. Probl. Remote Sens. Earth Space* **2023**, 20, 157–164.
15. Li, Z.; Ye, J.; Wang, Y.; Liu, L.; Shi, Y.; Li, Y.; Ye, J. A novel micromechanical method for nonlinear analysis of magnetostrictive composites. *Compos. Struct.* **2023**, 311, 116633. [[CrossRef](#)]
16. Chen, L.; Zhu, Y.; Ling, J.; Zhang, M. Temperature dependence modeling and experimental evaluation of a multidimensional discrete magnetostrictive actuator. *Appl. Therm. Eng.* **2023**, 230, 120736. [[CrossRef](#)]
17. Mizukawa, Y.; Ahmed, U.; Blažević, D.; Rasilo, P. Modeling and efficiency maximization of magnetostrictive energy harvester under free vibration. *J. Sound Vib.* **2023**, 558, 117759. [[CrossRef](#)]
18. Xie, J.; Li, M. A fast BDF2 Galerkin finite element method for the one-dimensional time-dependent Schrödinger equation with artificial boundary conditions. *Appl. Numer. Math.* **2023**, 187, 89–106. [[CrossRef](#)]
19. Navia, D.; Sanz, Á.S. Exploring the nonclassical dynamics of the “classical” Schrödinger equation. *Ann. Phys.* **2024**, 463, 169637. [[CrossRef](#)]
20. Bormotov, A.N.; Proshin, I.A.; Vasilkov, A.V. *Investigation of Mathematical Models of Structure Formation of Composite Materials by Analytical Methods*; Bulletin of the Bryansk State Technical University: Bryansk, Russian, 2011; pp. 62–70. Available online: <https://www.elibrary.ru/item.asp?id=16388945> (accessed on 10 June 2024).
21. Bormotov, A.N.; Proshin, I.A.; Kuznetsova, M.V. Quality management system in mathematical modeling and multicriteria synthesis of nanomodified composites for special purposes. *Proc. Inst. Syst. Anal. Russ. Acad. Sci.* **2014**, 64, 110–118.
22. Konovalova, I.I.; Selezneva, S.V.; Konovalov, A.N.; Bormotov, A.N. Algorithm for setting adaptive regulators. *World Sci. Discov.* **2013**, 22–33. Available online: <https://www.elibrary.ru/item.asp?id=19049611> (accessed on 10 June 2024).
23. Nielsen, M.; Chang, I. *Quantum Computing and Quantum Information*; Publishing House “Mir”: Moscow, Russian, 2006; 824p.
24. Ozhigov, Y.I. Modified Zalka–Wiesner scheme for modeling quantum systems. *Math. Model.* **2012**, 24, 109–119.
25. Satanin, A.M. *Additional Chapters of Quantum Mechanics: Methods for Solving the Scattering Problem Using High-Performance Computing*; Nizhny Novgorod State University: Nizhny Novgorod, Russia, 2014; 56p.
26. Li, A.; Goto, K.; Kobayashi, Y.; Hara, Y.; Jia, Y.; Shi, Y.; Soutis, C.; Kurita, H.; Narita, F.; Otsuka, K.; et al. Energy harvesting using a magnetostrictive transducer based on switching control. *Sens. Actuators A Phys.* **2023**, 355, 114303. [[CrossRef](#)]
27. Kurita, H.; Keino, T.; Senzaki, T.; Narita, F. Direct and inverse magnetostrictive properties of Fe–Co–V alloy particle-dispersed polyurethane matrix soft composite sheets. *Sens. Actuators A Phys.* **2022**, 337, 113427. [[CrossRef](#)]
28. Bogdanov, Y.I.; Bogdanova, N.A.; Fastovets, D.V.; Lukichev, V.F. The solution of the Schrodinger equation on a quantum computer by the Zalka–Wiesner method taking into account quantum noise. *Lett. JETP* **2021**, 114, 391–399.
29. De Raedt, H.; Hams, A.H.; Michielsen, K.; De Raedt, K. Quantum Computer Emulator. *Comput. Phys. Commun.* **2000**, 132, 1–20. [[CrossRef](#)]
30. Pfister, R.; Schubert, G.; Kröll, M. Transfer of Logistics Optimizations to Material Flow Resource Optimizations using Quantum Computing. *Procedia Comput. Sci.* **2024**, 232, 32–42. [[CrossRef](#)]
31. Lee, K.F.; Reilly, D.R.; Moraw, P.; Kanter, G.S. Emulation of up-conversion based quantum key distribution scheme using active pump-controlled basis selection and adaptive thresholding. *Opt. Commun.* **2020**, 475, 126258. [[CrossRef](#)]
32. Osca, J.; Vala, J. Implementation of photon partial distinguishability in a quantum optical circuit simulation. *Comput. Phys. Commun.* **2023**, 289, 108773. [[CrossRef](#)]

33. Rakyta, P.; Morse, G.; Nádori, J.; Majnay-Takács, Z.; Mencer, O.; Zimborás, Z. Highly optimized quantum circuits synthesized via data-flow engines. *J. Comput. Phys.* **2024**, *500*, 112756. [\[CrossRef\]](#)
34. Islam, M.T.; Nandwana, D.; Healy, J.; Jaklich, J.K.; Dong, B.; Yu, A.; Moore, E.E.; Ijiri, Y.; McCall, S.K.; Willard, M.A. Structural and magnetic properties of magnetostrictive Fe-Ga-Zr nanocrystalline alloy. *J. Alloys Compd.* **2023**, *958*, 170541. [\[CrossRef\]](#)
35. Zhan, Y.-S.; Lin, C.-H. Micromechanics-based constitutive modeling of magnetostrictive 1–3 and 0–3 composites. *Compos. Struct.* **2021**, *260*, 113264. [\[CrossRef\]](#)
36. Hubert, O.; Daniel, L.; Bernard, L. Multiscale Modeling of Magnetostrictive Materials. *Encycl. Smart Mater.* **2022**, *5*, 337–354. [\[CrossRef\]](#)
37. Ajagekar, A.; You, F. Variational quantum circuit based demand response in buildings leveraging a hybrid quantum-classical strategy. *Appl. Energy* **2024**, *364*, 123244. [\[CrossRef\]](#)
38. Cranganore, S.S.; De Maio, V.; Brandic, I.; Deelman, E. Paving the way to hybrid quantum–classical scientific workflows. *Future Gener. Comput. Syst.* **2024**, *158*, 346–366. [\[CrossRef\]](#)
39. Hevia, J.L.; Peterssen, G.; Piattini, M. qSOA[®]: Dynamic integration for hybrid quantum/Classical software systems. *J. Syst. Softw.* **2024**, *214*, 112061. [\[CrossRef\]](#)
40. Shukla, A.; Vedula, P. A hybrid classical-quantum algorithm for solution of nonlinear ordinary differential equations. *Appl. Math. Comput.* **2023**, *442*, 127708. [\[CrossRef\]](#)
41. Gay-Balmaz, F.; Tronci, C. Evolution of hybrid quantum–classical wavefunctions. *Phys. D Nonlinear Phenom.* **2022**, *440*, 133450. [\[CrossRef\]](#)
42. Zagrebin, M.A. Crystal Structure, Phase Diagrams, Electronic and Magnetic Properties of Three-, Four- and Five-Component Geisler Alloys: Dissertation for the degree of Doctor of Physical and Mathematical Sciences. Ph.D. Thesis, Chelyabinsk State University Publishing House, Chelyabinsk, Russia, 2021; 326p.
43. Chernenko, V. Magnetostrictive Ni-Mn-Based Heusler Alloys. *Encycl. Smart Mater.* **2022**, *4*, 160–176. [\[CrossRef\]](#)
44. Chen, Z.X.; Sakuraba, Y.; Masuda, K.; Xu, X.D.; Kumara, L.S.R.; Tajiri, H.; Kotani, Y.; Kushwaha, V.; Perumal, A.; Hono, K. Microstructure and atomic order analyses in Co-Fe-Cr-Al Heusler alloy thin films: Interpretation of spin gapless semiconductor-like transport properties. *Acta Mater.* **2022**, *232*, 117958. [\[CrossRef\]](#)
45. Pfeuffer, L.; Gràcia-Condal, A.; Gottschall, T.; Koch, D.; Faske, T.; Bruder, E.; Lemke, J.; Taubel, A.; Ener, S.; Scheibel, F.; et al. Influence of microstructure on the application of Ni-Mn-In Heusler compounds for multicaloric cooling using magnetic field and uniaxial stress. *Acta Mater.* **2021**, *217*, 117157. [\[CrossRef\]](#)
46. Budruk, A.; Phatak, C.; Petford-Long, A.K.; De Graef, M. In situ lorentz TEM magnetization study of a Ni–Mn–Ga ferromagnetic shape memory alloy. *Acta Mater.* **2011**, *59*, 4895–4906. [\[CrossRef\]](#)
47. Hao, Z.; Liu, R.; Fan, Y.; Wang, L. First-principles calculations of a new half-metallic Heusler alloy Fe-Cr-As. *J. Alloys Compd.* **2020**, *820*, 153118. [\[CrossRef\]](#)
48. Mori, K.; Horibe, T.; Maejima, K. Evaluation of the axial force in an FeCo bolt using the inverse magnetostrictive effect. *Measurement* **2020**, *165*, 108131. [\[CrossRef\]](#)
49. Hao, X.; Yang, B.; Li, J.; Wang, D.; Li, Z.; Yan, H.; Zhang, Y.; Esling, C.; Zhao, X.; Zuo, L. Achieving a linear magnetostrictive effect in textured MnCoSiGe alloys. *Acta Mater.* **2023**, *242*, 118486. [\[CrossRef\]](#)
50. Wittek, P.; Cucchiatti, F.M. A second-order distributed Trotter–Suzuki solver with a hybrid CPU–GPU kernel. *Comput. Phys. Commun.* **2013**, *184*, 1165–1171. [\[CrossRef\]](#)
51. Ariasoca, T.A.; Sholihun; Santoso, I. Trotter–Suzuki-time propagation method for calculating the density of states of disordered graphene. *Comput. Mater. Sci.* **2019**, *156*, 434–440. [\[CrossRef\]](#)
52. Wittek, P.; Calderaro, L. Extended computational kernels in a massively parallel implementation of the Trotter–Suzuki approximation. *Comput. Phys. Commun.* **2015**, *197*, 339–340. [\[CrossRef\]](#)
53. Semenova, A.; Dyachenko, S.A.; Korotkevich, A.O.; Lushnikov, P.M. Comparison of split-step and Hamiltonian integration methods for simulation of the nonlinear Schrödinger type equations. *J. Comput. Phys.* **2021**, *427*, 110061. [\[CrossRef\]](#)
54. Zhang, X.; Yang, B.; Wei, C.; Luo, M. Quantization method and Schrödinger equation of fractional time and their weak effects on Hamiltonian: Phase transitions of energy and wave functions. *Commun. Nonlinear Sci. Numer. Simul.* **2021**, *93*, 105531. [\[CrossRef\]](#)
55. Graf, T.; Felser, C.; Parkin, S.S.P. Simple rules for the understanding of Heusler compounds. *Prog. Solid State Chem.* **2011**, *39*, 1–50. [\[CrossRef\]](#)
56. Pirge, G.; Hyatt, C.V.; Altıntaş, S. Characterization of Ni-Mn-Ga magnetic shape memory alloys. *J. Mater. Process. Technol.* **2004**, *155–156*, 1266–1272. [\[CrossRef\]](#)
57. Webster, P.J.; Ziebeck, K.R.A. Dynamic magnetic properties of ferromagnetic and antiferromagnetic Heusler alloys. *J. Magn. Magn. Mater.* **1985**, *50*, 7–18. [\[CrossRef\]](#)
58. Scherban, A.B.; Meshcheryakova, E.N. Approach to the formalization of models of identification and structural management of complex systems. *J. Adv. Res. Dyn. Control Syst.* **2019**, *11*, 124–131.
59. Vorontsov, A.A.; Slesarev, Y.N. Mathematical Modeling and Experimental Check of Output Signals of Magnetostrictive Converters of Movement. In Proceedings of the 2019 International Russian Automation Conference, RusAutoCon, Sochi, Russia, 8–14 September 2019.

60. Pashchenko, D.; Martyshkin, A.; Trokoz, D. Modeling of Process Control Algorithms for Parallel Computing Systems Using Nondeterministic Automata. *Stud. Syst. Decis. Control* **2021**, 337, 235–249.
61. Sedelnikov, A.; Kurkin, E.; Quijada-Pioquinto, J.G.; Lukyanov, O.; Nazarov, D.; Chertykov-tseva, V.; Kurkina, E.; Hoang, V.H. Algorithm for Propeller Optimization Based on Differential Evolution. *Computation* **2024**, 12, 52. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Mathematical Modeling of the Heat Transfer Process in Spherical Objects with Flat, Cylindrical and Spherical Defects

Pavel Balabanov, Andrey Egorov , Alexander Divin , Sergey Ponomarev, Victor Yudaev, Sergey Baranov and Huthefa Abu Zetoonh

Department of Mechatronics and Technological Measurements, Tambov State Technical University, 392000 Tambov, Russia; balabanov.pv@mail.tstu.ru (P.B.); agdv@yandex.ru (A.D.); svponom@yahoo.com (S.P.); pt-arta@yandex.ru (V.Y.); baranov.sergey127@gmail.com (S.B.); hodefax2020x@yandex.ru (H.A.Z.)

* Correspondence: egorov.andrey@list.ru; Tel.: +7-(910)-6522100

Abstract: This paper proposes a method for determining the optimal parameters for the thermal testing of plant tissues of fruits and vegetables containing surface and subsurface defects in the form of areas of plant tissues with different thermophysical characteristics. Based on well-known mathematical models for objects of predominantly flat, cylindrical and spherical shapes containing flat, spherical and cylindrical regions of defects, numerical solutions of three-dimensional, non-stationary temperature fields were found, making it possible to measure the power and time of the thermal exposure of the sample surface to the radiation from infrared lamps using the finite element method. This made it possible to ensure the reliable detection of a temperature contrast of up to 4 °C between the defect and defect-free regions of the test object using modern thermal imaging cameras. In this case, subsurface defects can be detected at a depth of up to 3 mm from the surface. To determine the parameters of mathematical models of temperature fields, such as thermal conductivity and a coefficient of the thermal diffusivity of plant tissues, a new method of a pulsed heat flux from a flat heater is proposed; this differs in the method of processing experimental data and makes it possible to determine the required characteristics with high accuracy during the active stage of the experiment in a period not exceeding 1–3 min.

Keywords: thermophysical control; plant diseases; modeling of temperature fields



Citation: Balabanov, P.; Egorov, A.; Divin, A.; Ponomarev, S.; Yudaev, V.; Baranov, S.; Abu Zetoonh, H. Mathematical Modeling of the Heat Transfer Process in Spherical Objects with Flat, Cylindrical and Spherical Defects. *Computation* **2024**, *12*, 148. <https://doi.org/10.3390/computation12070148>

Academic Editor: Ali Cemal Benim

Received: 10 June 2024

Revised: 15 July 2024

Accepted: 15 July 2024

Published: 17 July 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Computer vision systems can effectively detect defects in plant tissue through the analysis of images of fruits and vegetables obtained in the range of 8–13 microns [1]. In this case, the main conditions are significant differences in the thermophysical characteristics (TPCs) of defect and defect-free zones of the test object, as well as the correctly selected operating parameters of the non-contact, non-destructive thermal testing process, which provide the required temperature contrast between the different quality zones of plant tissue [2]. To determine the operating parameters of thermal testing, it is necessary to solve the problem of the mathematical modeling of heat transfer in heterogeneous spherical bodies containing inclusions (defects) with thermal characteristics that differ from the characteristics of the base material. In this regard, the problem of determining the parameters of mathematical models—the coefficient of thermal diffusivity and the thermal conductivity of substances included in the defect and defect-free zones of the test object—arises.

The analysis of literary sources makes it possible to conclude that, in most cases, mathematical models of heat transfer in bodies made of heterogeneous materials are developed for objects in the form of a plate, cylinder or sphere. The formulation of the problem of heat transfer in a heterogeneous material in the form of an unlimited plate is described in [3,4], where the problem of the freezing of wet soil in a thawed state and having a certain specified distribution of temperature and moisture can be considered classical.

At the same time, it should be borne in mind that the process of heat transfer in plant tissues of fruits and vegetables differs significantly from similar processes in the soil due to the different shapes of the objects. In [3,5], the coupled problems of stationary thermal conductivity in a non-uniform body consisting of several homogeneous parts, in each of which the thermal conductivity coefficient of the material depends on the temperature distribution in the body, are considered. However, this approach is not applicable when modeling the process of unsteady heat transfer. From this perspective, Reference [6] is of particular interest, as it considers a mathematical model of the temperature field in a multilayer inhomogeneous ball consisting of several layers. Each layer, at each moment of time, is characterized by internal and external radii, temperature, specific heat capacity, density, thermal conductivity, humidity and heat release density. In this study, dependence was proposed for calculating the layer temperature as a function of the Fourier number. The proposed model can be used to calculate the temperature field in objects of plant origin if their shape is close to spherical, such as apples. However, it does not take into account the influence of defects localized in a limited volume of the object. In [7], a mathematical model of a multilayer cylinder of finite dimensions with a defect in the inner layer is considered. The solution can be obtained using a numerical method. At the same time, the proposed mathematical model makes it possible to determine the temperature fields in a multilayer cylinder during its cooling and the presence of defects in the inner layer. The presented model can be used to describe the temperature field in cylindrical objects of plant origin (for example, carrots and onions). In [8], a model of the temperature field in a flat object with cylindrical defects was proposed. In [9], heat transfer in layered rod structures is considered [9]. In [10], based on a statistical model, an artificial neural network model and a finite element model, the temperature fields of flat objects with various inclusions are studied.

In [11–14], mathematical models of temperature fields of crack-shaped defects are presented. It is noted that the spread of heat in an object leads to the appearance of a complex temperature field distribution on its surface. By formally fixing this and solving the inverse problem of non-stationary thermal conductivity, the approximate position and dimensions of the defect can be found. Thus, the known methods of the analytical modeling of temperature fields in heterogeneous media can be used for certain geometric shapes and inclusions. At the same time, due to the need to calculate three-dimensional temperature fields, the models contain a number of simplifications, which can significantly reduce their accuracy. The most promising seems to be the numerical modeling of the parameters of heterogeneous objects [15], including temperature fields [16], using the finite element method [17]. In [18,19], this modeling method was used for objects containing a solid phase with pores and cracks [20] filled with another substance, which made it possible to calculate changes in temperature, mechanical stress and phase deformation in heterogeneous objects of an arbitrary shape, including nanoobjects. Modeling heat transfer processes in plant tissues is complicated by the need to take into account the diffusion of moisture and its subsequent evaporation. Similar studies were carried out to simulate the processes of drying corn grains [21] and cocoa beans [22], potatoes [23], etc. The mathematical models used to study the drying process make it possible to determine the optimal operating parameters that ensure minimal energy consumption to remove moisture from the product. However, their complexity inevitably entails costs for computing power and time resources to calculate the three-dimensional temperature fields in control objects. When determining the optimal operating parameters of active thermal testing, one should take into account the need to limit the power of thermal exposure in order to ensure the safety of plant tissues. This also makes it possible to heat the plant tissues so that the thermophysical properties can be considered constant. In [24,25], models were developed that make it possible to determine the operating parameters of laser radiation (power, heating spot area, etc.) acting on the surface of a sample according to a non-stationary temperature distribution, providing the necessary temperature contrast between the defect and defect-free zones of

the test object, and also the determination of their geometric parameters, calculated using regression models.

Reference [26], focusing on the development of mathematical models in problems of thermal diagnostics, suggests that forecasting the durability of composite structural elements (CSEs) with defects can also be useful in developing algorithms to determine the shelf life of plant products. The task of predicting durability is considered to be calculating the damage parameters of CSEs with diagnosed defects under the influence of external mechanical factors for a given period, as well as calculating the limiting time to failure. The study also contains a mathematical formulation of the geometric inverse problem of thermal conductivity to recognize three-dimensional shapes and types of defects in CSEs and a numerical method for solving it.

In [27–29], the process of thermal conductivity in a heterogeneous medium is modeled by a homogeneous matrix with cylindrical inclusions separated by a contact layer. Also, in these papers, a study was carried out of the influence of the thickness and material of the layer on the effective properties of the matrix–layer–inclusion system and temperature profiles were obtained for configurations with different concentrations of inclusions and the surface layer, as well as for the different thermal conductivity coefficients of the components.

To determine the thermophysical characteristics of materials, one can use experimental data on the thermophysical characteristics of plant tissues taken from authoritative scientific publications [30,31], or calculated dependencies [30], including those obtained using machine learning technologies and artificial neural networks (ANNs) [32]. However, it should be borne in mind that thermophysical characteristics may differ depending on the variety, as well as the place of growth. Therefore, the most reliable method for obtaining information about the values of thermophysical properties is still a thermophysical experiment.

Thus, the literature review showed, on the one hand, that existing mathematical models of heat transfer in heterogeneous media, as a rule, do not take into account the finite dimensions of the test object and are designed to determine the presence of inclusions of a certain geometric shape. On the other hand, their numerical implementation requires significant computing resources due to their complexity and, as a rule, their nonlinearity. The objectives of the research described in this paper are to adapt known mathematical models of temperature fields for such objects as fruit trees and vegetable crops containing surface and subsurface defects; to obtain a numerical solution of the presented mathematical models in order to determine the optimal power and time of external thermal exposure, ensuring an acceptable temperature contrast between the defect and defect-free zones; and to develop a method and configure a measuring device for the express goal of determining the thermophysical characteristics of plant tissues of control objects, the values of which are used as parameters of mathematical models.

2. Materials and Methods

2.1. A Mathematical Model of the Temperature Field of a Spherical Object with Defects

When writing mathematical models in the form of boundary value problems of thermal conductivity, which make it possible to calculate the temperature fields of test objects (TO) (or their local areas), the following assumptions were made. Test objects of plant origin, such as apples, potatoes, oranges and lemons, have a spherical shape (Figure 1). The defects they contain, including rot, bruises, hail holes and scabs, are close in shape to the shape of the surface area of the object on which they are located. Let us assume that V_1 , V_2 , V_3 and V_4 are a set of points x , y and z , belonging to healthy 1 (Figure 1) tissues of the test object, and tissues containing defects of flat shape 2 (thickness H), cylindrical shape 3 or spherical shape 4 (radius R), respectively. Defects are located at a distance h from the surface of the object. Similarly, S_1 , S_2 , S_3 and S_4 are a set of points x , y and z belonging to the outer surface of healthy tissues and tissues damaged by defects in the corresponding shape. The initial temperature T_0 of plant tissues will be considered constant and uniformly distributed throughout the volume of the test object. When the test object is exposed to

a heat flux from a heat source, the change in temperature T on the surface and inside the object does not exceed 3–5 °C. With such a change in the surface temperature of the test object, the thermophysical properties of plant tissues can be considered independent of temperature and constant over time.

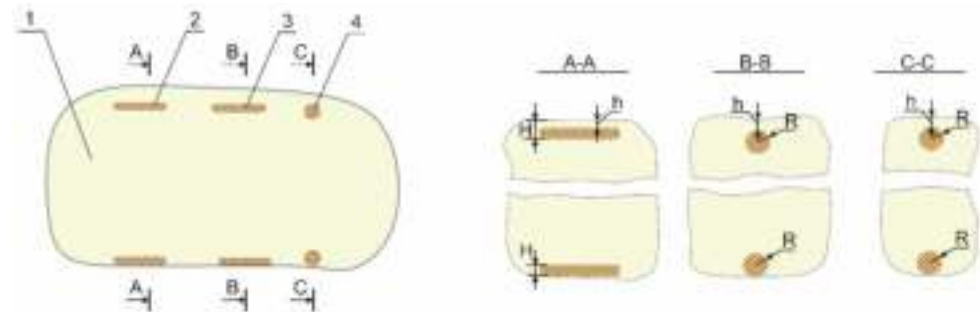


Figure 1. Physical model of spherical objects with defects. A, B, C are the cutting planes for tissues containing defects of flat, cylindrical and spherical shapes respectively.

Taking the temperature field model of a spherical body [33,34] as the basis, we proposed a mathematical model of the temperature field of a spherical body containing all three types of defects under consideration. In the general case, this can be written in the form of a boundary value problem of thermal conductivity [35]:

$$c_i \rho_i \frac{\partial T_i(x, y, z, \tau)}{\partial \tau} = \text{div}[\lambda_i \cdot \text{grad } T_i(x, y, z, \tau)], \quad (1)$$

$$i = \overline{1, 4}, \tau > 0, x, y, z \in V_i$$

$$T_i(x, y, z, 0) = T_0 = \text{const}, \quad (2)$$

$$\lambda_1 \frac{\partial T_1(x, y, z, \tau)}{\partial n} \Big|_{x, y, z \in S_1} = q(\tau) = \begin{cases} q = \text{const}, & 0 < \tau < \tau_1, \\ 0, & \tau > \tau_1 \end{cases} \quad (3)$$

$$\begin{cases} T_1(x, y, z, \tau) \Big|_{x, y, z \in S_{i+1}} = T_{i+1}(x, y, z, \tau) \Big|_{x, y, z \in S_{i+1}}, \\ \lambda_1 \frac{\partial T_1(x, y, z, \tau)}{\partial n_{i+1}} \Big|_{x, y, z \in S_{i+1}} = \lambda_{i+1} \frac{\partial T_{i+1}(x, y, z, \tau)}{\partial n_{i+1}} \Big|_{x, y, z \in S_{i+1}}, \end{cases} i = \overline{1, 3}, \quad (4)$$

$$\frac{\partial T_1(x, y, z, \tau)}{\partial x} \Big|_{x=0} = \frac{\partial T_1(x, y, z, \tau)}{\partial y} \Big|_{y=0} = \frac{\partial T_1(x, y, z, \tau)}{\partial z} \Big|_{z=0} = 0, \quad (5)$$

where λ , c , ρ and a are the thermal conductivity, specific heat capacity, density and thermal diffusivity of plant tissues, respectively; τ is time; n is a single vector normal to the surface S of the corresponding region of the test object; q is specific heat flux density from an external heat source. Since, during the experiment, the test object was exposed to a thermal effect lasting about 20–30 s, for the specified time, in the center of the test object ($x = y = z = 0$), the temperature changes so insignificantly that we neglected this change. Therefore, at the center of the sample, the boundary conditions specified by Equation (5) can be used. Also, due to the insignificant heating of the test object, we neglected the change in thermophysical properties.

Note that taking into account the insignificant, relative to the initial values, temperature difference in plant tissues during the thermal testing, instead of the boundary condition specified by the Stefan–Boltzmann law, we used the boundary condition of the second kind (3), where τ_1 is the duration of heat exposure on the test object when diagnosing defects.

When mathematically describing the temperature field of a spherical body with defects 1', 2' and 3', located on its surface, we supplemented boundary Condition (3) with the following equations:

$$\lambda_i \frac{\partial T_i(x, y, z, \tau)}{\partial n_i} \Big|_{x, y, z \in S_i} = q(\tau), \quad i = \overline{2, 4}.$$

When modeling the temperature field of objects of plant origin that have a spherical shape and contain internal and external defects, it is necessary to know the thermophysical characteristics of the corresponding plant tissues. To study TPCs, the method described below is proposed.

2.2. A Mathematical Model of the Temperature Field in a Flat Sample When Exposed to a Pulsed Heat Source

To determine the TPCs of plant tissues, samples 1–3 were cut out from the spherical body under study (Figure 2a), having the shape of plates with a thickness l . A low-inertia flat heater 4 was placed between samples 2 and 3. The electric heater (Figure 2b) was made of permalloy foil 4, fixed in dielectric holders 6. Slots 7 were made in the foil to ensure the specified electrical resistance of the heater: 1.57 Ohms. The outermost elements of the heater were equipped with electrical contacts 8, to which wires are connected to supply electrical voltage. Since the heater was made of thin foil, the duration of the power pulse supplied to it was 20–30 s; then, during mathematical modeling, we neglected the inertia during heating. The main junction of a differential thermocouple 5 was installed between samples 1 and 2, and an additional junction of the same thermocouple was installed on the lower surface of sample 3.

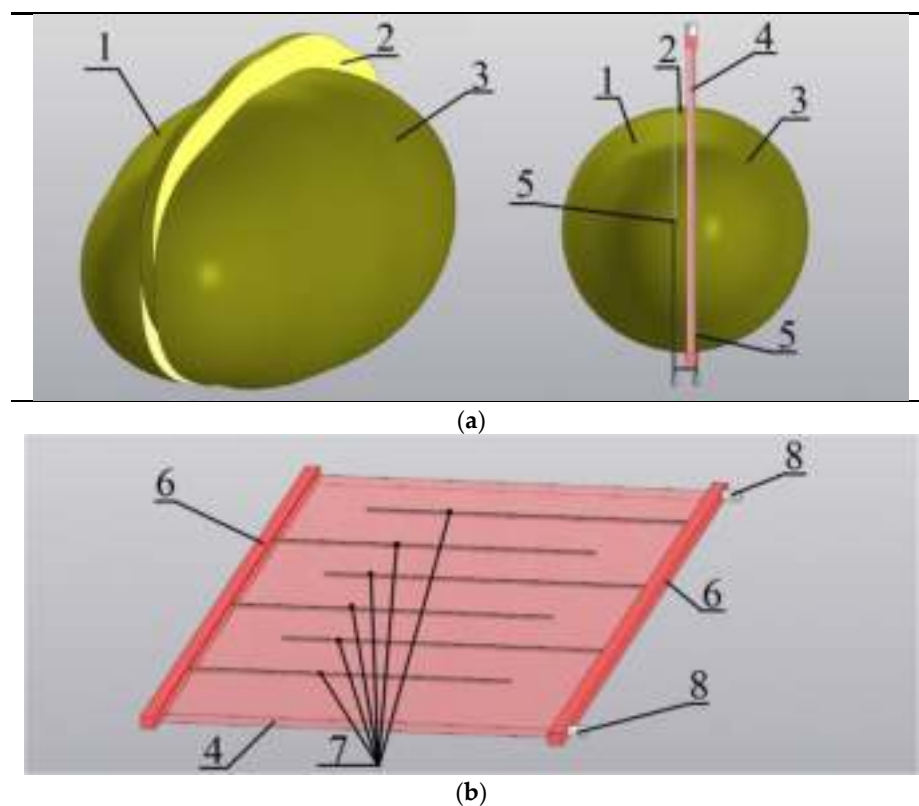


Figure 2. A physical model of the measuring device (a) and electric heater (b).

When an electric voltage pulse is applied to the heater, the temperature field in sample 2 can be written in the form of a boundary value problem:

$$\frac{\partial T(x, \tau)}{\partial \tau} = a \frac{\partial^2 T(x, \tau)}{\partial x^2}, \tau > 0, 0 < x < \infty, \quad (6)$$

$$T(x, 0) = T_0 = 0, \quad (7)$$

$$-\lambda \frac{\partial T(0, \tau)}{\partial x} = q[h(\tau) - h(\tau - \tau_{imp})], \quad (8)$$

$$T(\infty, \tau) = T_0 = 0, \quad (9)$$

where $T(x, \tau)$ is temperature of the sample under study in the plane with coordinate x at time τ , counted from the moment the thermal pulse begins to be supplied; T_0 is the initial temperature; q is the heat flux supplied to the sample over a period of time $0 < \tau \leq \tau_{imp}$; $h(\tau)$, $h(\tau - \tau_{imp})$ are unit asymmetric step functions specified by the following relations:

$$h(\tau) = \begin{cases} 0, & \tau < 0, \\ 1, & \tau \geq 0; \end{cases} h(\tau - \tau_{imp}) = \begin{cases} 0 & \tau < \tau_{imp}, \\ 1 & \tau \geq \tau_{imp}; \end{cases} \quad (10)$$

τ_{imp} is the duration of the thermal pulse when determining TPCs. We neglected the dependence of the desired thermophysical properties on temperature, since the temperature of plant tissues in the area of thermal exposure during the experiment changed by no more than 5 °C.

The solution to the boundary value problem (6)–(9) has the form $T(x, \tau) - T_0 = \frac{2q}{\lambda} \cdot f(x, \tau, \tau_{imp}, a)$,
where

$$f(x, \tau, \tau_{imp}, a) = \begin{cases} \sqrt{a\tau} \cdot ierfc\left(\frac{x}{2\sqrt{a\tau}}\right) & \text{if } 0 < \tau \leq \tau_{imp}, \\ \sqrt{a\tau} \cdot ierfc\left(\frac{x}{2\sqrt{a\tau}}\right) - \sqrt{a(\tau - \tau_{imp})} \cdot ierfc\left(\frac{x}{2\sqrt{a(\tau - \tau_{imp})}}\right) & \text{if } \tau > \tau_{imp}, \end{cases}$$

$ierfc(u) = \int_u^\infty erfc(W) dW = \frac{1}{\sqrt{\pi}} e^{-u^2} - uerfc(u)$ is a special function that is an integral of the function $erfc(W) = 1 - erf(W)$;
 $erf(W) = \frac{2}{\sqrt{\pi}} \int_0^W e^{-W^2} dW$ is Gaussian error function.

For time $\tau \geq \tau_{imp}$, the above solution will take the form

$$T(x, \tau) - T_0 = \frac{qx}{\lambda} \left[\frac{ierfc[U(x, \tau)]}{U(x, \tau)} - \frac{ierfc[U(x, \tau - \tau_{imp})]}{U(x, \tau - \tau_{imp})} \right], \quad (11)$$

where $U(x, \tau) = \frac{x}{2\sqrt{a\tau}}$, $U(x, \tau - \tau_{imp}) = \frac{x}{2\sqrt{a(\tau - \tau_{imp})}}$ are dimensionless functions depending on x, τ, τ_{imp}, a and $U(x, \tau - \tau_{imp}) = \frac{x}{2\sqrt{a(\tau - \tau_{imp})}} = \frac{x}{2\sqrt{a\tau} \sqrt{\frac{\tau - \tau_{imp}}{\tau}}} = U(x, \tau) \sqrt{\frac{\tau - \tau_{imp}}{\tau}}$.

From the analysis of Expression (11), it is clear that, at some point in time, $\tau = \tau_{max}$ the change in the temperature difference $[T(x, \tau) - T_0]$ reached its maximum value $[T_{max} - T_0] = [T(x, \tau_{max}) - T_0]$. And, at the specified point in time, $\tau = \tau_{max}$ corresponds to a certain value of the dimensionless function $U^m = U(x, \tau_{max}) = \frac{x}{2\sqrt{a\tau_{max}}}$.

3. Results

The algorithm used to conduct the experiment and determine the desired TPCs ensures that, during time τ_{imp} , a constant power P is supplied to a flat electric heater of area S ($P = 2qS$), and the change in time corresponding to temperature differences $T(l, \tau) - T_0$ is recorded using a signal from a differential thermocouple. Based on the experimental data obtained, the maximum value of the temperature difference $T_{max} - T_0$ is found and

the experiment is completed when the condition $(T(l, \tau) - T_0)/(T_{max} - T_0) > 0.8 \dots 0.9$ is met. The experimental data obtained are used to determine thermal conductivity and thermal diffusivity. To achieve this, for a set of values λ_x, a_x , selected from the ranges $a_{min} < a_x < a_{max}, \lambda_{min} < \lambda_x < \lambda_{max}$ in which the required TPCs may lie, the values of the functional $F(a_x, \lambda_x) = \sum_{i=1}^N [T_t(l, \tau_i) - T(l, \tau_i)]^2$ are determined, where $T_t(l, \tau_i)$ is the theoretical temperature calculated at time τ_i using model (11) for a given TPC; N is the number of temperature values recorded during the experimental measurement of thermophysical properties. For the pair $a_x, \lambda_x, F \rightarrow min$ is taken as the desired TPC.

Using the described algorithm, the TPCs of plant tissues of fruits and vegetables were obtained. As an example, Table 1 shows the mathematical expectation M and standard deviation D of the results of five measurements of the TPCs of healthy and disease-affected potato plant tissues. Using the obtained TPC values and the mathematical ones described above, the modeling of temperature fields in spherical objects with defects of various shapes was performed.

Table 1. TPCs of plant tissues of fruits and vegetables.

Test Sample	λ , W/(m·K)	a , ($\times 10^7$) m ² /s	$c\rho$, ($\times 10^3$) kJ/(m ³ ·K)
M (healthy tissue)	0.507	1.417	3.56
D (healthy tissue)	0.010	0.022	0.011
M (late blight tissue)	0.550	1.420	3.800
D (late blight tissue)	0.009	0.046	0.017
M (dry rot tissue)	0.384	1.272	3.015
D (dry rot tissue)	0.007	0.053	0.178

As an example, we consider the procedure for validating a mathematical model of heat transfer in an ellipsoid-shaped object (Figure 3a) with the major and minor semi-axes having dimensions of 64–70 mm and 23–65 mm, respectively. The indicated size ranges correspond to the range of average sizes of potato tubers used as seeds for planting. The defect has a spherical shape and is located at a depth h of up to 3 mm below the surface. Defects located at a depth of more than three millimeters cannot be detected by the proposed thermal method. The ratio of the volume of defect V_3 to the total volume V_0 of the simulated object was in the range of 1–3%, which approximately corresponds to the range of radii of spherical defects of 2.3–10.5 mm in the smallest and largest tubers. Calculations of the temperature field on the surface of the model were carried out at the object's initial temperature of $T_0 = 20^\circ\text{C}$. On the external surface of the test object, a heat flux q with a specific power of 50–3000 W/m² was set, acting for a time τ_m in the range 1–5 s. When choosing the specified ranges of heat flux density and time of thermal exposure for the test object, we were guided by the following considerations. The longer we act on a potato tuber and the greater the heat flux density, the more the tuber heats up. This leads to the thermal destruction of its tissues and reduces its shelf life. The lower the heat flux density on the surface of the tuber and the shorter the period for which heat is applied to the tuber, the less likely it is that a defect will be detected. Taking this into account, it was experimentally found that the choice of the indicated ranges will allow, on the one hand, for a defect to be detected; on the other hand, it will not lead to the potato tissue heating above 5°C relative to the initial temperature of 20°C . For given values of τ_m and q , the temperature field on the surface of the test object was calculated (Figure 3b), and then the absolute value of the maximum temperature difference ΔT_{max} in the defect region and in the region of undamaged tissue was calculated. An experiment was conducted in which a real test object, a potato tuber (Figure 3c), was exposed to a heat flux q for time τ_m and the temperature field was measured using a Fluke Ti32 thermal imager (Figure 3d). Next, ΔT_{max} was calculated.

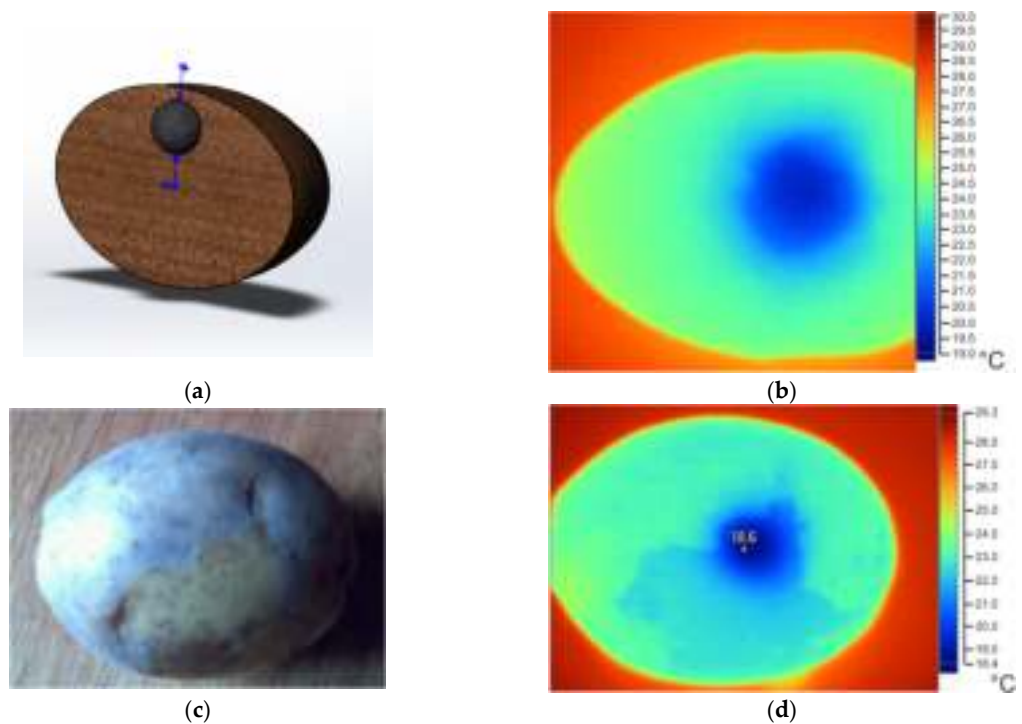


Figure 3. A 3D model of spherical object with a subsurface spherical defect (a), the temperature field on its surface in the defect area (top view) (b), a real potato tuber with an internal defect (c) and the temperature field on its surface (d).

Figure 4 shows examples of a comparison of theoretical (solid line) and experimental (dotted lines) values of ΔT_{max} , calculated for test objects with spherical defects such as dry rot (Figure 4a) and late blight type (Figure 4b), obtained for heat flux q in the range of 50–1500 W/m² at time τ_m of the thermal exposure of the object in the range of 1–5 s.

The obtained calculation data make it possible, depending on the sensitivity of the thermal imaging equipment used for thermal inspection, to select the minimum power of thermal exposure and its duration, ensuring defect recognition. For example, as is shown in Figure 4, it is clear that when a defect occurs at a depth of up to 3 mm at $q = 300$ W/m² and $\tau_1 = 1$ s, a temperature difference ΔT_{max} of about 0.5 °C will be observed in the region where the defect occurs, which is sufficient for detection by a thermal imager with a sensitivity of 0.05 °C.

The developed recommendations made it possible to select the optimal parameters of thermal impact on the test object. As an example, Figure 5 shows photographs of defect potatoes obtained in the visible and infrared regions of the radiation spectrum when objects were exposed to a heat flux with a density of 300 W/m² for a time of 2.4 s.

Figure 5 shows that the temperature in the defect area and in the defect-free area of a potato tuber tissue after thermal exposure differs by 1.4 °C. The temperature field of defect-free potato tissue is uniform. However, a temperature gradient is observed in the defect area. This allows for the use of a thermal imager to detect defects such as dry rot (Figure 5a) or late blight (Figure 5b). This can be achieved as follows. When a potato tuber is heated briefly, for up to 5 s, by an external heat source with a power of up to 1500 W/m², a temperature contrast in the defect area will be observed on the thermal imaging image, which allows it to be detected.

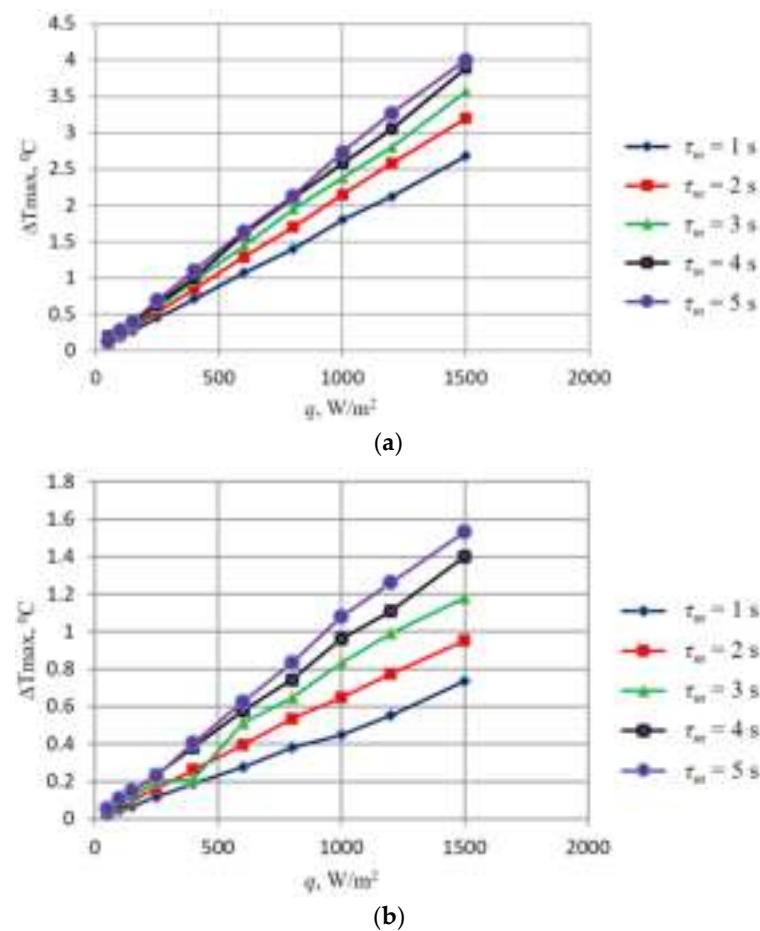


Figure 4. Dependency graphs $|\Delta T_{max}| = f(q)$ for spherical subsurface: (a) dry rot; (b) late blight defect.

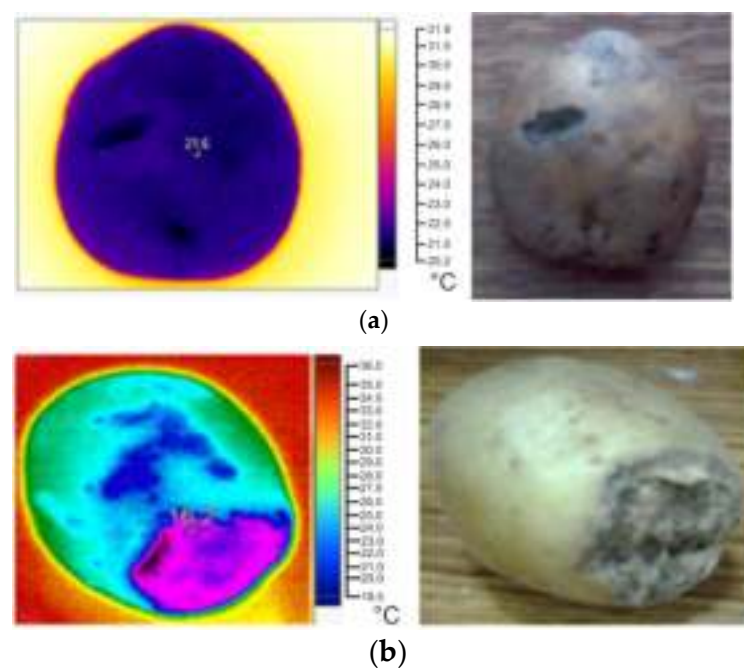


Figure 5. Photographs of defect-free and defect tissues of potato in the infrared and visible regions of the spectrum: (a) dry rot; (b) late blight.

4. Conclusions

When thermally inspecting defects in objects of plant origin, there are risks of thermal damage to plant tissue due to prolonged or high-intensity thermal exposure. On the other hand, insufficient power of the thermal influence may lead to some defects, especially those located under the surface of test objects, not being detected. Therefore, it is necessary to determine the optimal parameters of thermal exposure, particularly the exposure time and power of radiation sources. To solve the problem, based on well-known mathematical models for objects of predominantly flat, cylindrical and spherical shapes containing flat, spherical and cylindrical areas of defects, numerical solutions of three-dimensional non-stationary temperature fields were obtained, which made it possible to determine, using the finite element method, the power and time of the thermal impact on the surface of the sample due to radiation from infrared lamps. This made it possible to ensure reliable detection by modern thermal imaging cameras of a temperature contrast of up to 4 °C between the defect and defect-free zones of the test object. In this case, subsurface defects can lie at a depth of up to 3 mm from the surface.

To determine the parameters of mathematical models of temperature fields, such as the thermal conductivity and the coefficient of thermal diffusivity of plant tissues, a new method of a pulsed source of heat flow from a flat heater is proposed, which differs in the method of processing experimental data and makes it possible to determine the required characteristics with high accuracy during the active stage of the experiment, in a period not exceeding 1–3 min.

The experiments conducted to test the thermal testing method for detecting damaged plant tissues of a potato tuber showed that when exposed to heat from infrared lamps with a power of up to 1500 W for 3–5 s, surface and subsurface defects are quite distinguishable in infrared images obtained using a Fluke Ti32 thermal imager technical vision systems, which allows for the use of thermal testing to sort potatoes before storing them or before sending them to the consumer.

Author Contributions: P.B., A.D., A.E., S.P. and H.A.Z.: conceptualization, software, data curation, writing—original draft preparation, writing—review and editing, supervision and funding acquisition; P.B., V.Y. and S.B.: validation, formal analysis, investigation, visualization and writing—original draft preparation. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by financial support from the Ministry of Science and Higher Education of the Russian Federation within the framework of the project “Development of a robotic complex of ground and air unmanned platforms for use in agricultural technologies” (124062100023-3).

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Samanian, N.; Mohebbi, M. Thermography a New Approach in Food Science Studies: A Review. *MOJ Food Process. Technol.* **2016**, *2*, 110–119. [\[CrossRef\]](#)
2. Ballester, C.; Jiménez-Bello, M.A.; Castel, J.R.; Intrigliolo, D.S. Usefulness of thermography for plant water stress detection in citrus and persimmon trees. *Agric. For. Meteorol.* **2013**, *168*, 120–129. [\[CrossRef\]](#)
3. Zhao, L.; Gray, D.M.; Male, D.H. Numerical analysis of simultaneous heat and mass transfer during infiltration into frozen ground. *J. Hydrol.* **1997**, *200*, 345–363. [\[CrossRef\]](#)
4. Lykov, A.V. *Theory of Thermal Conductivity*; Vysshaya Shkola: Moscow, Russia, 1967; p. 600. (In Russian)
5. Zarubin, V.S.; Rodikov, A.V. Mathematical simulation of the temperature state of an inhomogeneous body. *High Temp.* **2007**, *45*, 243–254. [\[CrossRef\]](#)
6. Kurmachev, Y.U.F. *Mathematical Modeling of the Thermal Field of a Multilayer Inhomogeneous Ball in an Unsteady Mode*; Science & Education Scientific Edition of Bauman; MSTU: Murfreesboro, TN, USA, 2007. Available online: <https://cyberleninka.ru/article/n/matematicheskoe-modelirovanie-teplovogo-polya-mnogosloynnogo-neodnorodnogo-shara-pri-nestatsionarnom-rezhime/viewer> (accessed on 5 July 2024).

7. Sinitsyn, N.N.; Shestakov, I.V. Mathematical modeling of the temperature field of mixer armor taking into account a defect on the inner surface. *Cherepovets State Univ. Bull.* **2008**, *4*, 118–120.
8. Sinitsyn, N.N.; Makonkov, A.V.; Belozor, M.Y. Modeling of a non-stationary temperature field of asphalt pavement with rubber filler. In *Modern Materials, Equipment and Technology, Proceedings of the 4th International Scientific and Practical Conference, Kursk, Russia, 25–26 December 2014*; University Book Closed Joint Stock Company: Kursk, Russia, 2014; pp. 400–403.
9. Stankevich, I.V. Mathematical modeling of the temperature state of spatial rod structures made of inhomogeneous materials. *Symb. Sci.* **2016**, *13*, 53–57.
10. Li, J.; Yang, Y.; Liu, M. Research progress in the prediction of heat transfer properties of fabrics based on structural characteristics. *Adv. Text. Technol.* **2022**, *30*, 18–25. [[CrossRef](#)]
11. Kotel'nikov, V.V. Mathematical modeling of the process of formation of a temperature field on a defect in the form of a crack in the region of a stress raiser. *Occup. Saf. Ind.* **2008**, *5*, 51–58. (In Russian)
12. Rittel, D. Thermomechanical aspects of dynamic crack initiation. *Int. J. Fract.* **1999**, *99*, 201–212. [[CrossRef](#)]
13. Zvyagin, A.V.; Udalov, A.S.; Shamina, A.A. Numerical modeling of heat conduction in bodies with cracks. *Acta Astronaut.* **2024**, *214*, 196–201. [[CrossRef](#)]
14. Esfarjani, A.A.; Nazari, M.B.; Bayat, S.H. Dynamic crack propagation in variable stiffness composite laminates under thermal shock. *Theor. Appl. Fract. Mech.* **2024**, *130*, 104325. [[CrossRef](#)]
15. Salman, M.; Schmauder, S. Multiscale modeling of shape memory polymers foams nanocomposites. *Comput. Mater. Sci.* **2024**, *232*, 112658. [[CrossRef](#)]
16. Jin, L.; Zhang, R.; Du, X. Characterisation of temperature-dependent heat conduction in heterogeneous concrete. *Mag. Concr. Res.* **2018**, *70*, 325–339. [[CrossRef](#)]
17. She, Z.; Wang, K.; Li, P. Hybrid Trefftz polygonal elements for heat conduction problems with inclusions/voids. *Comput. Math. Appl.* **2019**, *78*, 1978–1992. [[CrossRef](#)]
18. Marin, M.; Hobiny, A.; Abbas, I. The effects of fractional time derivatives in porothermoelastic materials using finite element method. *Mathematics* **2021**, *9*, 1606. [[CrossRef](#)]
19. Chen, Y.; Sandhofer, S.N.; Wong, B.M. SHORYUKEN: An open-source software package for calculating nonlocal exchange interactions in nanowires. *Comput. Phys. Commun.* **2024**, *300*, 109197. [[CrossRef](#)]
20. Zhou, L.; Zhu, Z.; Que, X. Simulation of non-fourier heat conduction in discontinuous heterogeneous materials based on the peridynamic method. *Therm. Sci.* **2023**, *27*, 917–931. [[CrossRef](#)]
21. Wei, S.; Wang, Z.; Wang, F.; Xie, W.; Chen, P.; Yang, D. Simulation and experimental studies of heat and mass transfer in corn kernel during hot air drying. *Food Bioprod. Process.* **2019**, *117*, 170–182. [[CrossRef](#)]
22. Hii, C.L.; Law, C.L.; Law, M.C. Simulation of heat and mass transfer of cocoa beans under stepwise drying conditions in a heat pump dryer. *Appl. Therm. Eng.* **2013**, *54*, 264–271. [[CrossRef](#)]
23. Heshmati, M.K.; Khiavi, H.D.; Dehghannya, J.; Baghban, H. 3D simulation of momentum, heat and mass transfer in potato cubes during intermittent microwave-convective hot air drying. *Heat Mass Transf.* **2023**, *59*, 239–254. [[CrossRef](#)]
24. Divin, A.G.; Karpov, S.V. Using Laser Point Scanning Thermography for Quality Monitoring of Products Made of Composite Materials. *Eng. Technol. Syst.* **2024**, *34*, 145–163. [[CrossRef](#)]
25. Divin, A.G.; Ponomarev, S.V. Application of Laser Scanning Thermography and Regression Analysis to Determine Characteristics of Defects in Polymer Composite Materials. *Russ. J. Nondestruct. Test.* **2024**, *60*, 40–48. [[CrossRef](#)]
26. Nikolaev, A.A. Mathematical Modeling in Problems of Thermal Diagnostics and Forecasting the Durability of Composite Structural Elements with Defects. Tver: Ph.D. Diss. in Phys. and Math. Sci.: 05.13.18 Mathematical Modeling, Numerical Methods and Software Packages. 2010, p. 148. Available online: <https://www.dissercat.com/content/matematicheskoe-modelirovanie-v-zadachakh-teplovoy-diagnostiki-i-prognozirovaniya-dolgovechn> (accessed on 5 July 2024). (In Russian)
27. Savatorova, V.L.; Talonov, A.V.; Vlasov, A.N. Homogenization of thermoelasticity processes in composite materials with periodic structure of heterogeneities. *ZAMM Zeitschrift Angew. Math. Mech.* **2013**, *93*, 575–596. [[CrossRef](#)]
28. Vlasov, A.N.; Savatorova, V.L.; Talonov, A.V. Asymptotic averaging technique for heat conduction problems with phase transitions in layered media. *J. Appl. Mech. Tech. Phys.* **1995**, *36*, 773–780. [[CrossRef](#)]
29. Savatorova, V.L.; Talonov, A.V. Heat transfer in a composite material with structural hierarchy. *Compos. Mech. Comput. Appl.* **2021**, *12*, 45–59. [[CrossRef](#)]
30. Donsì, G.; Ferrari, G.; Nigro, R. Experimental determination of thermal conductivity of apple and potato at different moisture contents. *J. Food Eng.* **1996**, *30*, 263–268. [[CrossRef](#)]
31. Muniandy, A.; Benyathiar, P.; Mishra, D.K.; Ozadali, F. Dynamic thermal properties estimation using sensitivity coefficients for rapid heating process. *Foods* **2021**, *10*, 1954. [[CrossRef](#)] [[PubMed](#)]
32. Sablani, S.S.; Rahman, M.S. Using neural networks to predict thermal conductivity of food as a function of moisture content, temperature and apparent porosity. *Food Res. Int.* **2003**, *36*, 617–623. [[CrossRef](#)]
33. Carslaw, H.S.; Jaeger, J.C. *Conduction of Heat in Solids*; Clarendon Press: Oxford, UK, 1984; p. 532.

34. Polyanin, A.D. *Handbook of Linear Partial Differential Equations for Engineers and Scientists*; Chapman & Hall/CRC: Boca Raton, FL, USA, 2002; p. 735. [[CrossRef](#)]
35. Ponomarev, S.V.; Divin, A.G.; Ponomareva, E.S. Mathematical Models of Temperature Fields of Potato Tubers with Surface and Internal Defects. *Adv. Mater. Technol.* **2017**, *4*, 65–75. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Multilevel Quasi-Interpolation on Chebyshev Sparse Grids

Faisal Alsharif 

Department of Mathematics, College of Science, Taibah University,
Al-Madinah Al-Munawarah 30002, Saudi Arabia; fsharif@taibahu.edu.sa

Abstract: This paper investigates the potential of utilising multilevel quasi-interpolation techniques on Chebyshev sparse grids for complex numerical computations. The paper starts by laying down the motivations for choosing Chebyshev sparse grids and quasi-interpolation methods with Gaussian kernels. It delves into the practical aspects of implementing these techniques. Various numerical experiments are performed to evaluate the efficiency and limitations of the multilevel quasi-sparse interpolation methods with dimensions two dimension and three dimension. The work ultimately aims to provide a comprehensive understanding of the computational efficiency and accuracy achievable through this approach, comparing its performance with traditional methods.

Keywords: multilevel; quasi-interpolation; sparse; Chebyshev; Gaussian kernel; meshless grid

1. Introduction

Traditional interpolation methods, such as grid-based or radial basis function interpolation, encounter computational difficulties when dealing with multidimensional datasets. The computational complexity of these methods grows exponentially with the number of dimensions, rendering them similarly unfeasible on reasonable timeframes. Consequently, novel approaches are necessary to address the challenge of interpolation in multidimensional spaces.

The need for innovative solutions has driven the development of novel numerical methods that are capable of efficiently handling 2D, 3D, or high-dimensional data using higher-performance computers [1,2]. As a result, researchers have been able to make significant advancements in the various scientific domains mentioned, enabling the development of innovative solutions to complex problems [3].

2. Chebyshev Sparse Grids Technique

In the literature, a sparse grid technique has been proposed to overcome the curse of dimensionality with equally spaced grids for higher dimensional domains [4]. Georgoulis et al. [4] proposed an algorithm that reduces the requirement of a large number of data points on an equally spaced grid but keeps the simplicity of a grid structure in evaluating the interpolant. The construction of a sparse grid is defined as follows.

Define multi-index $\ell = (l_1, \dots, l_d) \in \mathbb{N}^d$ and $(i_1, \dots, i_d) \in \mathbb{N}^d$ as spatial positions. Define the point for that position as follows:

$$x_{\ell,i} = (x_{l_1,i_1}, \dots, x_{l_d,i_d}) \quad \text{with} \quad x_{l_j,i_j} = 2(i_j 2^{-l_j} - 0.5) \quad \text{for} \quad j = 1, \dots, d.$$

We can now define the family of uniform grids $\mathbb{G}_\ell$ as the following set of points:

$$\mathbb{G}_\ell = \{x_{\ell,i} : i = (i_1, \dots, i_d), i_j = 0, 1, \dots, 2^{l_j}, j = 1, \dots, d\}.$$

The number of points N_ℓ in $\mathbb{G}_\ell$ is given as follows:

$$N_\ell = \prod_{j=1}^d (2^{l_j} + 1).$$



Citation: Alsharif, F. Multilevel Quasi-Interpolation on Chebyshev Sparse Grids. *Computation* **2024**, *12*, 149. <https://doi.org/10.3390/computation12070149>

Received: 24 May 2024

Revised: 25 June 2024

Accepted: 3 July 2024

Published: 18 July 2024



Copyright: © 2024 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

If we have $\ell = (n, \dots, n)$, $\mathbb{G}_\ell$ is the uniform full grid of equally spaced points of level n with distances between the points given by $h = 2^{-n}$. The size of the grid $\mathbb{G}_\ell$ is $N = (2^n + 1)^d$. We denote this full grid as $\mathbb{G}_{n,d}$. Now, we consider a sparse subset of uniform full grids, $\mathbb{G}_{n,d}$, as follows:

$$\tilde{\mathbb{G}}_{n,d} = \bigcup_{|\ell|=n+(d-1)} \mathbb{G}_\ell, \quad (1)$$

with $|\ell| = l_1 + l_2 + \dots + l_d$. We refer to $\tilde{\mathbb{G}}_{n,d}$ as the sparse grid of level in dimension d . Figure 1 shows a visual representation of the sparse grid for $n = 4$ and $d = 2$. Notice that some of the grid points are included in more than one sub-grid; hence, there are some redundancies.



Figure 1. Sparse grid $\tilde{\mathbb{G}}_{4,2}$.

In an alternative formulation of the sparse grid, there are arguments that instead of using an equally spaced grid, one can use Chebyshev points to create grids on $[-1, 1]$. For the multi-index $\ell = (l_1, \dots, l_d) \in \mathbb{N}^d$ and $(i_1, \dots, i_d) \in \mathbb{N}^d$, we can define

$$x_{l_j, i_j}^c = -\cos\left(\frac{2i_j - 1}{2 \times 2^{l_j}}\right) \quad \text{for } j = 1, \dots, d$$

to have the point $x_{\ell, i}^c = (x_{l_1, i_1}^c, \dots, x_{l_d, i_d}^c)$. Similarly, we can define the family of Chebyshev grids as follows:

$$\mathbb{G}_\ell^c = \{x_{\ell, i}^c : i = (i_1, \dots, i_d), i_j = 0, 1, \dots, 2^{l_j}, j = 1, \dots, d\},$$

and the sparse subset of Chebyshev grids as follows:

$$\tilde{\mathbb{G}}_{n,d}^c = \bigcup_{|\ell|=n+(d-1)} \mathbb{G}_\ell^c, \quad (2)$$

where $\tilde{\mathbb{G}}_{n,d}^c$ is referred to as the sparse Chebyshev grid of level n in dimension d . Figure 2 shows a visual representation of the sparse Chebyshev grid for $n = 4$ and $d = 2$. We observe that the sparse grid points are not equally spaced.



Figure 2. Sparse Chebyshev grid $\tilde{\mathbb{G}}_{4,2}^c$.

3. Quasi-Sparse Interpolation with Gaussian Kernels

In the previous formulation of the interpolation using RBF, we observed that it is required to solve a system of N linear equations, where N is the total number of points. Using a sparse grid of points as defined in the previous sub-section will reduce the total number of points and, subsequently, the total number of equations for the higher dimension compared to a full grid of points. However, that does not fully eliminate the ill-conditioning of the problems. Another method proposed in the literature to avoid solving the system of linear equations is to use quasi-interpolation. There is extensive discussion on error estimates in quasi-interpolation in the literature [5–7].

We use the idea of quasi-interpolation on the sparse grid with anisotropic Gaussian basis functions. If $c_{i,j} = |x_{i+1,j} - x_{i,j}|$, the distance between points x_i and x_{i+1} in dimension j for $j = 1, \dots, d$. Then, the diagonal matrix can be defined as follows:

$$A_i = \text{diag}(c_{i_1,1}, \dots, c_{i_d,d}),$$

where $i = (i_1, \dots, i_d)$. For each multi-index ℓ , we can now define the quasi-interpolant anisotropically:

$$Q_\ell f(x) = \sum_{x_{\ell,i} \in \mathbb{G}_\ell} f(x_{\ell,i}) \phi_c(A_i^{-1}(x - x_{\ell,i})), \quad x \in \Omega. \quad (3)$$

Here, we consider the Gaussian kernel

$$\phi_c(x) = \frac{1}{\sqrt{2\pi}c} \exp\left(-\frac{x^2}{2c^2}\right),$$

with c determining the degree of smoothness of the interpolant.

To construct the quasi-sparse kernel-based interpolant on the sparse grid $\tilde{\mathbb{G}}_{n,d}$, the anisotropic interpolants on the sub-grids $\mathbb{G}_\ell$ must be linearly combined using the following formula:

$$Q_{n,d}f(x) = \sum_{j=0}^{d-1} (-1)^j \binom{d-1}{j} \sum_{|\ell|=n+(d-1)-j} Q_\ell f(x). \quad (4)$$

For example, the quasi-sparse kernel-based interpolation for $d = 2$ would be

$$Q_{n,2}f(x) = \sum_{|\ell|=n+1} Q_\ell f(x) - \sum_{|\ell|=n} Q_\ell f(x),$$

and for $d = 3$, it would be

$$Q_{n,3}f(x) = \sum_{|\ell|=n+2} Q_\ell f(x) - 2 \sum_{|\ell|=n+1} Q_\ell f(x) + \sum_{|\ell|=n} Q_\ell f(x).$$

In a similar fashion, we can construct the quasi-sparse kernel-based interpolation formula for the sparse grids based on Chebyshev points by replacing grids $\mathbb{G}_\ell$ with $\mathbb{G}_\ell^c$.

4. Multilevel Quasi-Sparse Interpolation

From the construction of sparse grids, it is obvious that sparse grids from the lower level to the upper level are nested; that is,

$$\tilde{\mathbb{G}}_{n,d} \subset \tilde{\mathbb{G}}_{n+1,d}$$

for all $n \geq 1$. We have illustrated this nested nature in Figure 3 for equally spaced grid points. For Chebyshev points, we have also shown the nested nature of the grids, which is presented in Figure 4. Moreover, while combining kernel-based anisotropic interpolants for each subgrid, they were appropriately scaled, with the scaling being proportional to the density of the corresponding constituent subgrid. One can utilize these nested properties of the subgrids to propose a multilevel method of interpolation with them without adding any further complexity to the quasi-sparse kernel-based interpolation introduced earlier. In the earlier literature [8–10], multilevel methods for radial basis functions were used to combine stationary and non-stationary interpolants to accelerate convergence and to improve the numerical stability of ill-conditioned problems.

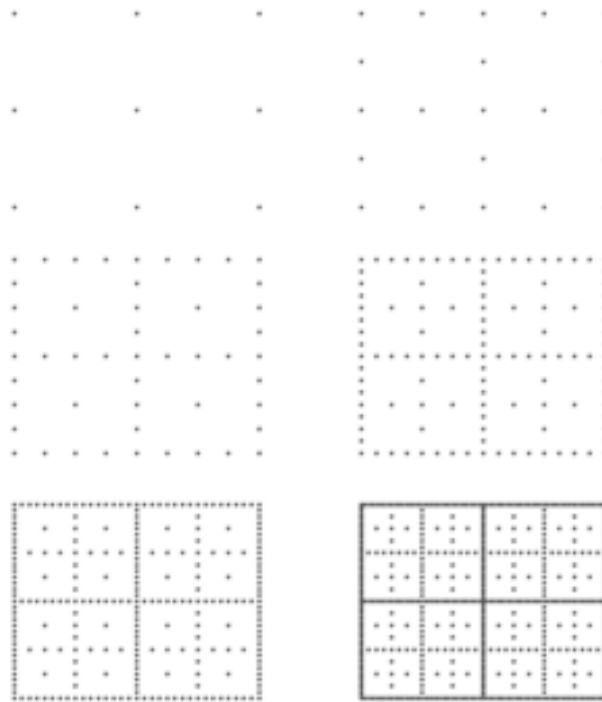


Figure 3. Nested nature of sparse grids $\tilde{\mathbb{G}}_{n,2}$ with equally spaced points for $n = 1, 2, \dots, 6$.

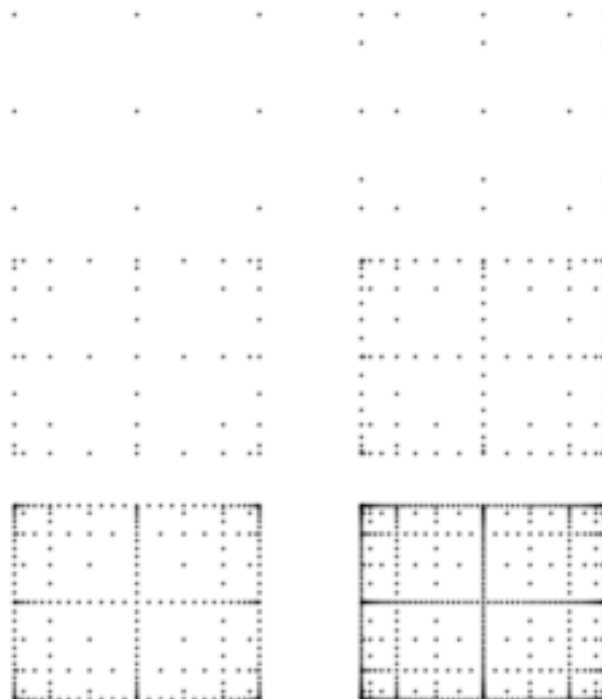


Figure 4. Nested nature of sparse grids $\tilde{\mathbb{G}}_{n,2}^c$ with Chebyshev points for $n = 1, 2, \dots, 6$.

The main idea of multilevel sparse grid-based quasi-interpolation is to interpolate the function at the lowest level, that is, for $n = 1$, and then update at each level of the sub-grids by computing the residuals at every level, interpolating them at the higher level. More formally, we obtain the quasi-sparse kernel-based interpolation $Q_{n_0,d}f(x)$ at the coarsest level 1; that is, the coarsest level approximation is $S_{0,d}f(x) = Q_{n_0,d}u(x)$. Then, we define

the residual function at level j as $r_j(x) = f(x) - S_{j-1,d}f(x)$ and interpolate the residual function as

$$\Delta_j(x) = Q_{n_0+j,d}r_j(x)$$

on the level $n_0 + j$ sparse grid $\tilde{\mathbb{G}}_{n_0+j,d}$ and then update the multilevel interpolant as follows:

$$S_{j,d}f(x) = S_{j-1,d}f(x) + \Delta_j(x), \quad \text{for } 1 \leq j \leq n.$$

We refer to this algorithm as the multilevel quasi-sparse kernel-based interpolation. As the quasi-sparse kernel-based interpolation is being used at every level, the time-complexity of the multilevel method is linear with these levels. However, it takes more time than the usual quasi-sparse interpolation for a fixed level as it evaluates the residuals at every grid point. Again, the algorithm is easily amenable to parallel computing.

5. Implementation

In this article, all algorithms and numerical experiments are implemented in MATLAB software 23.2. This is necessary to be done, otherwise it will affect publication. environment. The implementation is broken down into three main parts:

- Creating sparse grids using either equally spaced grids or Chebyshev points for each level of the multilevel grid;
- Anisotropic quasi-interpolation $Q_\ell f(x)$ using a Gaussian kernel;
- The development of an algorithm for quasi-sparse interpolation, which combines anisotropic quasi-interpolants, as described earlier.

For multilevel quasi-interpolation, we implement another function to apply the algorithm.

Multilevel Quasi-Sparse Interpolation

The outline for the algorithm of multilevel quasi-sparse interpolation is given in Algorithm 1. In the implementation of the multilevel quasi-sparse interpolation, after initial checks and creating a sparse grid of points, if necessary, we compute the quasi-interpolation approximation using the function QSIK for level $n = 1$. For all other levels $i > 1$, we evaluate the quasi-sparse interpolation approximation at all sparse grid points and then subtract that from the original function value to compute the residual function and then again compute the quasi-sparse approximation for the residual function at level i and add that estimate to the function approximation, where we already have up to level i . We continue computing the residual functions and updating our approximation at each level until the maximum level of evaluation n is reached. This multilevel approximation returns the function approximations at all levels $1 \leq i \leq n$. Though it requires a lot of quasi-sparse interpolation approximations at all levels, we can reduce the computation time by computing the function approximations at all levels together rather than calling the function separately for each level.

Algorithm 1: Multilevel Quasi-Sparse Interpolation Algorithm

Data: Sparse grid data decomposition

Result: The multilevel quasi-sparse interpolation $S_{n,d}f(x)$.

1: Initialize the first interpolation value at zero, that is $S_{0,d}f(x) = 0$;

2. Construct the nested sparse grids as $\tilde{\mathbb{G}}_{j,d} \subset \tilde{\mathbb{G}}_{j+1,d}$;

3. **for** $j = 1, \dots, n$ **do**

 Compute $r_j(x) = f(x) - S_{j-1,d}f(x)$ for all $x \in \tilde{\mathbb{G}}_{j,d}$;

 Compute $\Delta_j(x) = Q_{j,d}r_j(x)$ using Algorithm described in Section 4;

 Update $S_{j,d}f(x) = S_{j-1,d}f(x) + \Delta_j(x)$.

end

6. Numerical Experiments for Quasi-Sparse Interpolation with $d = 2$

We now illustrate the performance of the proposed quasi-sparse kernel-based interpolation for the following test functions:

- $F_1^{2d}(x, y)$ (Franke's function)

$$F_1^{2d}(x, y) = 0.75e^{-((9x-2)^2+(9y-2)^2)/4} + 0.75e^{-(9x+1)^2/49-(9y+1)^2/10} + 0.5e^{-((9x-7)^2+(9y-3)^2)/4} - 0.2e^{-((9x-4)^2+(9y-7)^2)}.$$

- $F_2^{2d}(x, y) = (1 - x^2)(1 - y^2).$
- $F_3^{2d}(x, y) = \sqrt{\frac{18}{\pi}}e^{-(x^2+81y^2)}.$
- $F_4^{2d}(x, y) = \frac{1.25+\cos(5.4y)}{6+6(3x-1)^2}.$

Franke's function is very commonly used as a test function in the interpolation literature with RBF kernels [11]. The others can be found in [12–14], respectively. These test functions are illustrated in Figure 5.

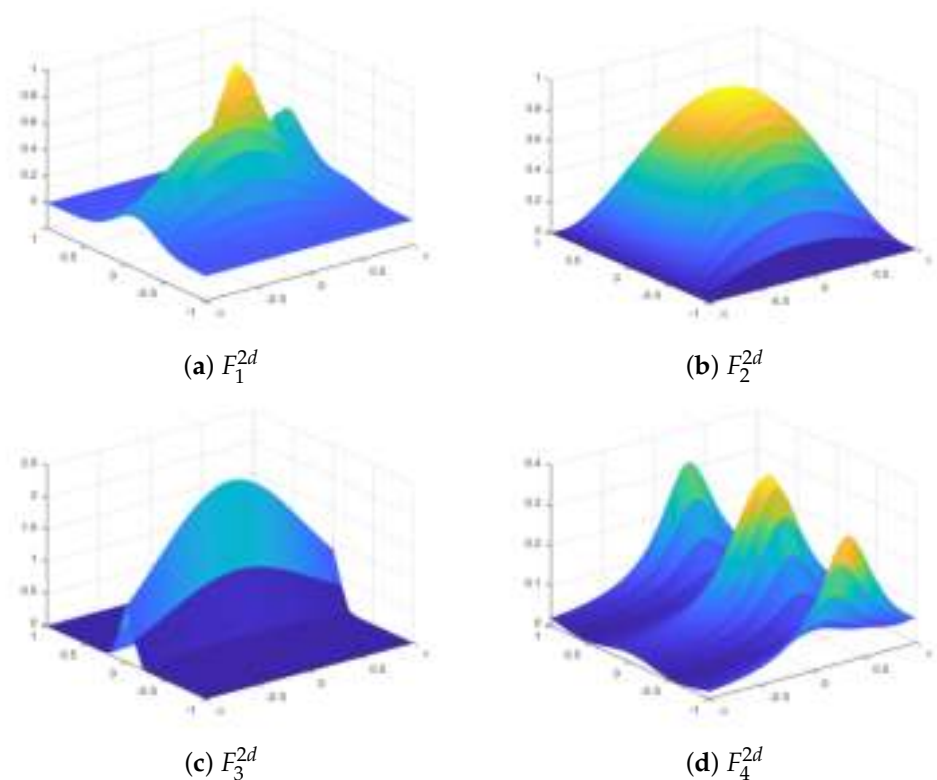


Figure 5. Test functions in two dimensions.

In all of these numerical experiments, we have used a 50×50 uniform grid of points in the rectangle $[-1, 1] \times [-1, 1]$. The error in approximation is reported as the maximum modulus error; that is, $\|f - Q_{n,d}f\|$, and root mean square error is defined as follows:

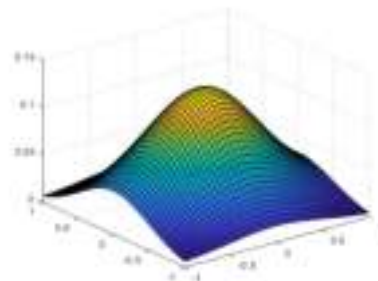
$$\sqrt{\frac{1}{N_e} \sum_{i=1}^n (f(x_i) - Q_{n,d}f(x_i))^2},$$

where $x_i, i = 1, \dots, N_e$ are the evaluation points, and N_e is the total number of evaluation points. We also report the number of points in the sparse grid $\tilde{G}_{n,d}$ (or $\tilde{G}_{n,d}^c$) as $SGnode$ and the number of points actually used in the computation as some of the points are revisited in the sparse grid computations as DoF .

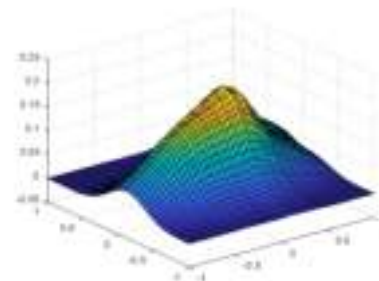
Figure 6 shows the approximation using our proposed quasi-sparse kernel-based interpolation with an equally spaced grid $\tilde{G}_{n,2}$ for different values of n . Table 1 also shows the error estimates. We observe that the maximum error or the RMS error does not decrease substantially as we increase the sparse grid level n , but the plots show that approximations are very close to the true function F_1^{2d} for $n > 6$. With Chebyshev points, we observe much better approximations with both the maximum error and the RMS error decreasing with the increase in level n . The plots of the interpolating function evaluated at the 50×50 grid points for the algorithm with Chebyshev points are shown in Figure 7. Again, we observe that the approximations are quite good for $n > 6$.

Table 1. Quasi-sparse interpolation with a Gaussian kernel using equally spaced grids and Chebyshev point grid results for the test function F_1^{2d} .

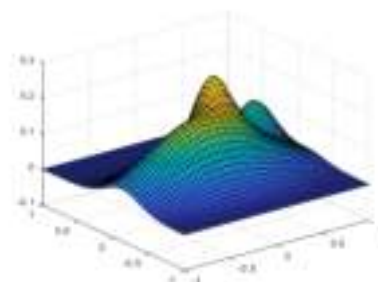
SGnodes	DoFs	Equal			Chebyshev		
		Max Error	RMS Error	Time	Max Error	RMS Error	Time
9	9	0.92616	0.24884	0.018	0.89474	0.231770	0.0179
21	39	0.92911	0.24940	0.025	0.91436	0.233400	0.024
49	109	0.88190	0.23617	0.047	0.76886	0.180060	0.048
113	271	0.86792	0.23205	0.079	0.64397	0.132990	0.076
257	641	0.82000	0.22755	0.149	0.55427	0.103100	0.145
577	1475	0.80022	0.22624	0.313	0.43634	0.070535	0.292
1281	3333	0.78678	0.22575	0.664	0.30363	0.043283	0.682
2817	7431	0.78586	0.22560	1.401	0.18151	0.024121	1.305
6145	16,393	0.77768	0.22554	2.822	0.09293	0.012131	2.678
13,313	35,851	0.78174	0.22553	5.536	0.04147	0.005629	5.116



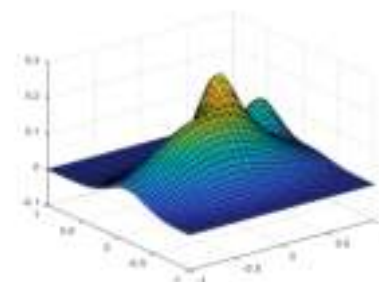
(a) $n = 1$



(b) $n = 4$



(c) $n = 7$



(d) $n = 10$

Figure 6. Approximations using an equally spaced sparse grid with level n and a Gaussian kernel for the test function F_1^{2d} .

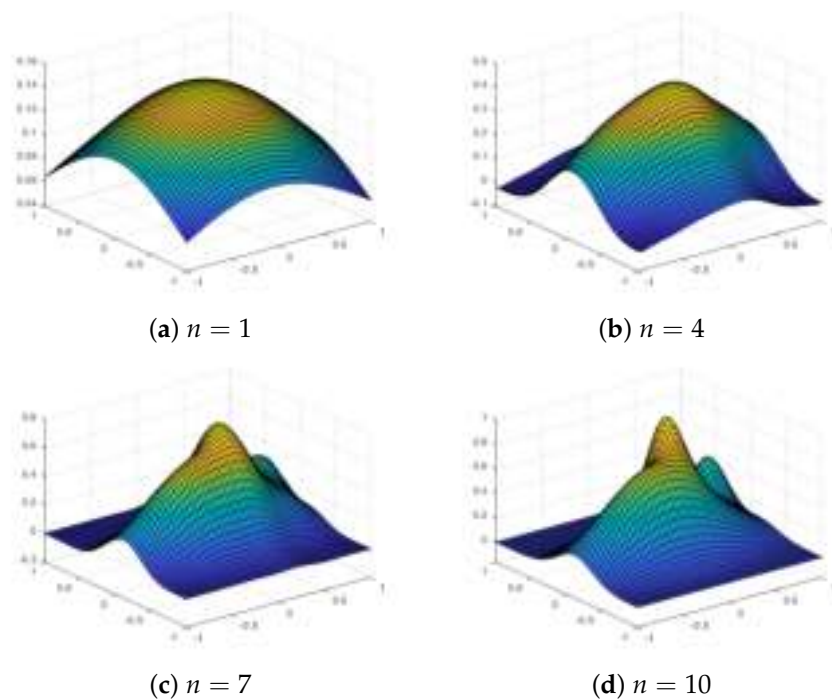


Figure 7. Approximations using Chebyshev sparse grids with level n and a Gaussian kernel for the test function F_1^{2d} .

Figure 8 shows the approximations using our proposed quasi-sparse kernel-based interpolation with equally spaced grids $\tilde{G}_{n,2}$ for different values of n for the test function $F_2^{2d}(x,y)$. Table 2 also shows the error estimates. We observe that the maximum error or the RMS error does not decrease substantially as we increase the grid level n , but the plots show that approximations are very close to the true function F_2^{2d} for $n > 2$ when we use equally spaced grids. With Chebyshev points, we observe much better approximations, with both the maximum error and the RMS error decreasing with the level n . The plots of the interpolating function evaluated at 50×50 grid points for the algorithm with Chebyshev points are shown in Figure 9. For $n = 1$, the approximations are poor, but it approximates nicely as n increases. On the other hand, when there are equally spaced grids, the interpolant starts to overfit the surface and lacks smoothness for larger values of n . Thus, by increasing the number of grid levels, the approximation actually deteriorates for equally spaced grids.

Table 2. Quasi-sparse interpolation with a Gaussian kernel using equally spaced grids and Chebyshev point grid results for the test function F_2^{2d} .

SGnodes	DoFs	Equal			Chebyshev		
		Max Error	RM SError	Time	Max Error	RMS Error	Time
9	9	0.84805	0.45491	0.018	0.8409	0.41803	0.017
21	39	0.77623	0.41629	0.028	0.8236	0.38014	0.023
49	109	0.75404	0.40043	0.072	0.5678	0.25995	0.050
113	271	0.75256	0.39481	0.085	0.4225	0.17896	0.079
257	641	0.75263	0.39297	0.186	0.2591	0.10923	0.141
577	1475	0.74977	0.39240	0.375	0.1347	0.05796	0.298
1281	3333	0.75337	0.39223	0.683	0.0593	0.02757	0.648
2817	7431	0.74458	0.39218	1.294	0.0229	0.01259	1.321
6145	16,393	0.74666	0.39217	2.749	0.0096	0.00589	2.667
13,313	35,851	0.75301	0.39216	5.027	0.0045	0.00290	5.115

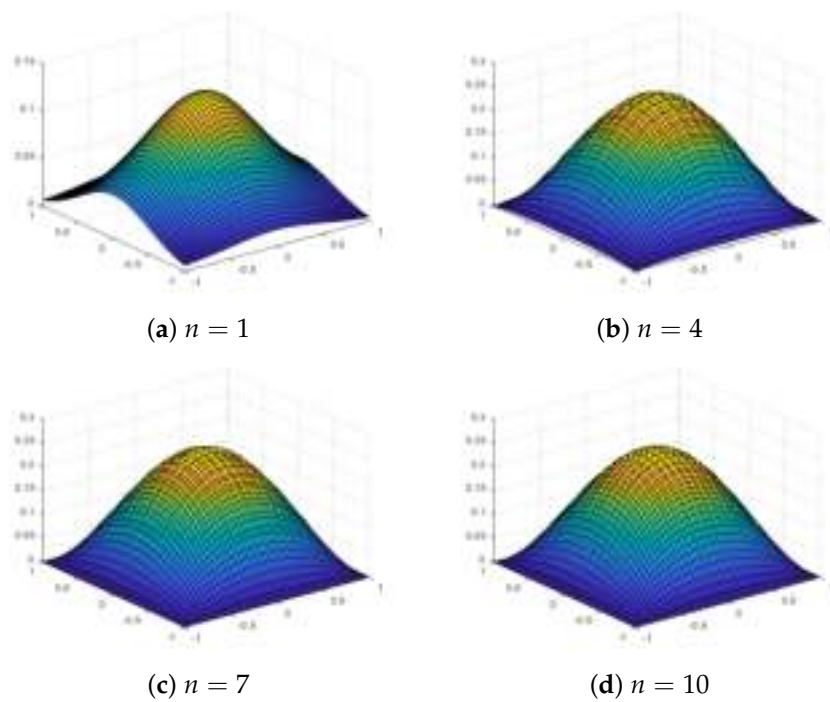


Figure 8. Approximations using an equally spaced sparse grid with level n and a Gaussian kernel for the test function F_2^{2d} .

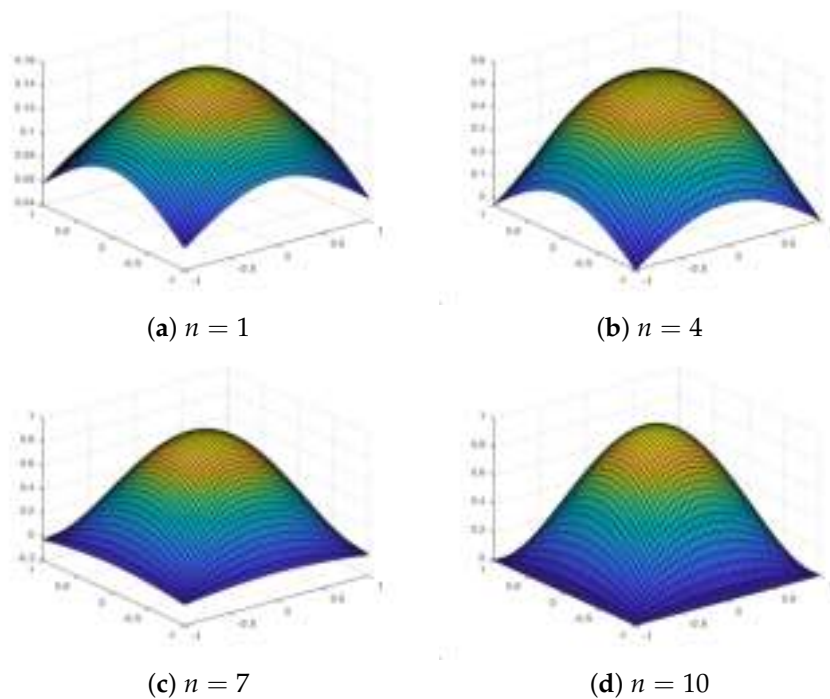


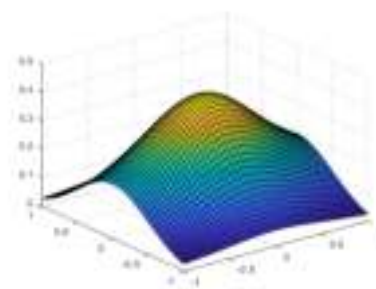
Figure 9. Approximations using a Chebyshev sparse grid with level n and a Gaussian kernel for the test function F_2^{2d} .

Figure 10 shows the approximations using our proposed quasi-sparse kernel-based interpolation with equally spaced grids $\tilde{\mathbb{G}}_{n,2}$ for different values of n for the test function $F_3^{2d}(x,y)$. Table 3 shows the error estimates. We observe that the maximum error or the RMS error does not decrease substantially as we increase the grid level n , but the plots show that approximations are very close to the true function F_3^{2d} for $n > 4$ when we use equally spaced grids. With Chebyshev points, we observe much better approximations, with

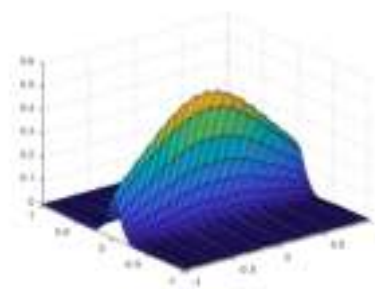
both the maximum error and the RMS error decreasing with the level n . The plots of the interpolating function evaluated at 50×50 grid points for the algorithm with Chebyshev points are shown in Figure 11. For $n = 1$, the approximations are poor, and the overall shape is not similar to the shape of the function, but it approximates nicely as n increases. On the other hand, when there are equally spaced grids, the interpolant starts to overfit the surface and lacks smoothness for larger values of n . Thus, by increasing the number of grid levels, the approximation actually deteriorates for equally spaced grids.

Table 3. Quasi-sparse interpolation with a Gaussian kernel using equally spaced grids and Chebyshev point grid results for the test function F_3^{2d} .

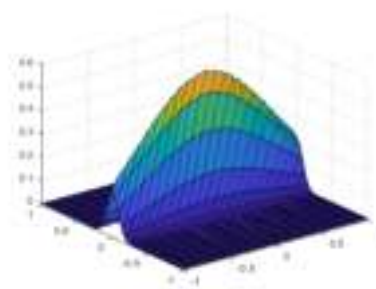
SGnodes	DoFs	Equal			Chebyshev		
		Max Error	RMS Error	Time	Max Error	RMS Error	Time
9	9	1.97480	0.40313	0.016	1.8427	0.48854	0.017
21	39	1.93230	0.37292	0.024	1.9868	0.45415	0.022
49	109	1.91530	0.36748	0.048	1.6822	0.39947	0.051
113	271	1.88590	0.37309	0.073	1.4987	0.34353	0.078
257	641	1.83230	0.36540	0.141	1.1621	0.25966	0.142
577	1475	1.80050	0.36168	0.308	0.8929	0.16772	0.300
1281	3333	1.79530	0.36006	0.626	0.6563	0.10807	0.682
2817	7431	1.78100	0.35929	1.319	0.3949	0.06122	1.325
6145	16,393	1.78460	0.35987	2.616	0.1928	0.03013	2.674
13,313	35,851	1.79160	0.36007	5.084	0.0797	0.01421	5.113



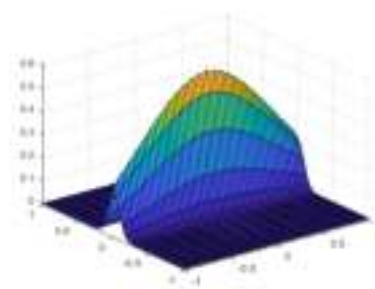
(a) $n = 1$



(b) $n = 4$



(c) $n = 7$



(d) $n = 10$

Figure 10. Approximations using an equally spaced sparse grid with level n and a Gaussian kernel for the test function F_3^{2d} .

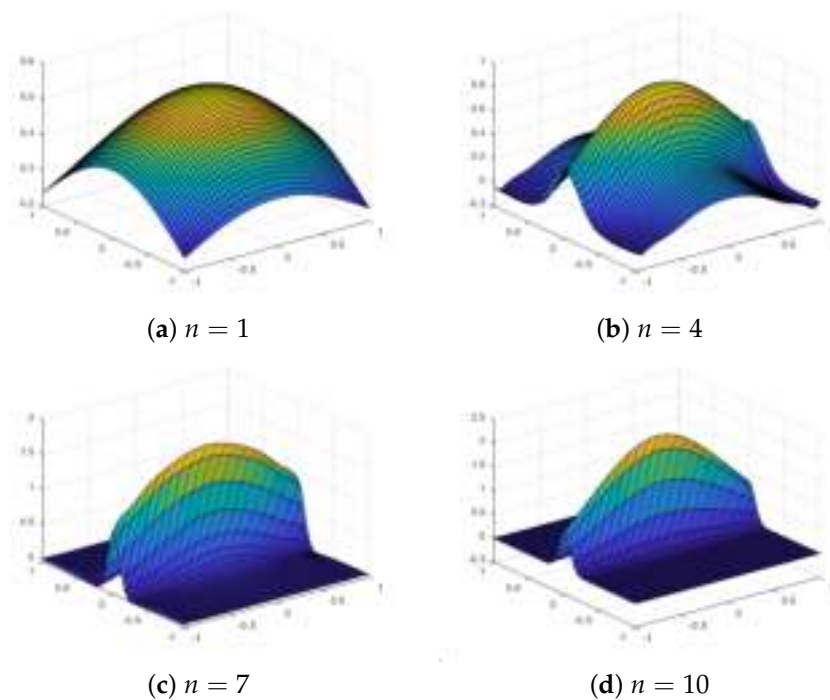


Figure 11. Approximations using a Chebyshev sparse grid with level n and a Gaussian kernel for the test function F_3^{2d} .

Figure 12 shows the approximations using our proposed quasi-sparse kernel-based interpolation with equally spaced grids $\tilde{G}_{n,2}$ for different values of n for the test function $F_4^{2d}(x, y)$. Table 4 shows the error estimates. We observe that the maximum error or the RMS error does not decrease substantially as we increase the grid level n , but the plots show that approximations are very close to the true function F_4^{2d} for $n > 4$ when we use equally spaced grids. With Chebyshev points, we observe much better approximations, with both the maximum error and the RMS error decreasing with the level n . The plots of the interpolating function evaluated at 50×50 grid points for the algorithm with Chebyshev points are shown in Figure 13. For $n = 2$, the approximations are poor, and the overall shape is not similar to the shape of the function, but it approximates nicely as n increases.

Table 4. Quasi-sparse interpolation with a Gaussian kernel using equally spaced grids and Chebyshev point grid results for the test function F_4^{2d} .

SGnodes	DoFs	Equal			Chebyshev		
		Max Error	RMS Error	Time	Max Error	RMS Error	Time
9	9	0.33863	0.08989	0.019	0.2952	0.07502	0.019
21	39	0.31804	0.08456	0.023	0.3350	0.08348	0.021
49	109	0.29950	0.08309	0.059	0.2294	0.05323	0.044
113	271	0.29077	0.08263	0.091	0.1873	0.04437	0.073
257	641	0.28290	0.08220	0.146	0.1491	0.03358	0.146
577	1475	0.28154	0.08200	0.299	0.1084	0.02364	0.299
1281	3333	0.28141	0.08192	0.615	0.0671	0.01448	0.748
2817	7431	0.27816	0.08187	1.372	0.0396	0.00817	1.339
6145	16,393	0.27869	0.08187	2.632	0.0305	0.00467	2.664
13,313	35,851	0.27887	0.08187	5.023	0.0265	0.00315	5.177

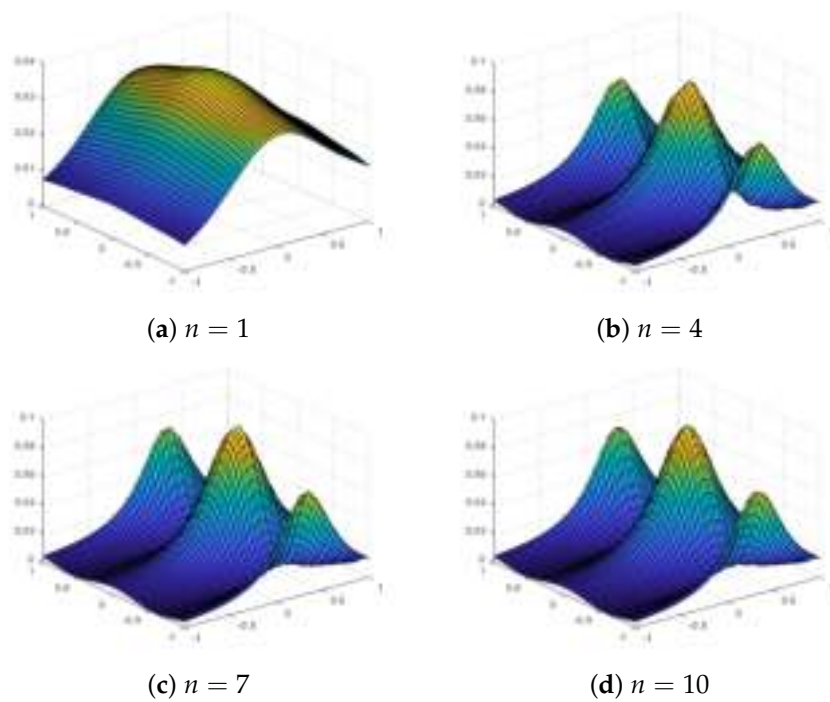


Figure 12. Approximations using an equally spaced sparse grid with level n and a Gaussian kernel for the test function F_4^{2d} .

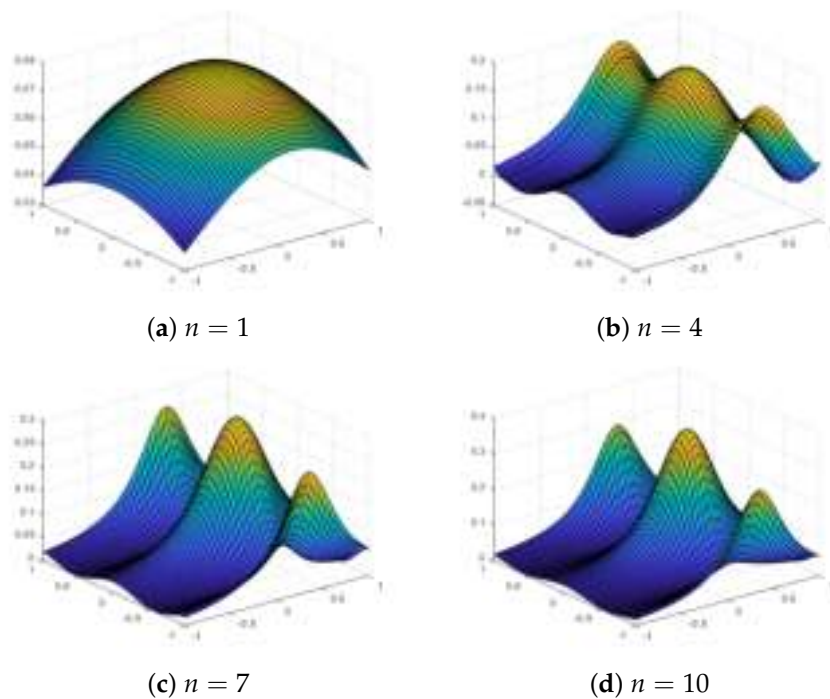


Figure 13. Approximations using a Chebyshev sparse grid with level n and a Gaussian kernel for the test function F_4^{2d} .

In these results, quasi-sparse interpolation with equally space grids reaches a stable maximum absolute error and root mean square errors at a lower level with a small number of sparse grid points. However, the quasi-sparse interpolation with Chebyshev points attains much smaller values for maximum errors and RMS errors for all test functions and stabilizes for larger values of the level n . Additionally, since the errors remain stable when the number of grid points increases, the quasi-sparse interpolation goes well with

the multilevel quasi-sparse algorithm, which will be examined in a subsequent sub-section. The main observation in these results is that quasi-sparse interpolation does not converge for equally spaced grids, but it does converge for Chebyshev grid points. The rate of convergence depends on the test function.

7. Numerical Experiments for Quasi-Sparse Interpolations with $d = 3$

In this section, we present the performance analysis for some test functions with domain $\Omega = [-1, 1]^3$ in dimension $d = 3$. The following functions are used as test functions:

- $F_1^{3d}(x, y)$ (Franke's function)
$$F_1^{3d}(x, y) = 0.75e^{-((9x-2)^2+(9y-2)^2+(9z-2)^2)/4} + 0.75e^{-(9x+1)^2/49-(9y+1)^2/10-(9z+1)^2/10} + 0.5e^{-((9x-7)^2+(9y-3)^2+(9z-5)^2)/4} - 0.2e^{-((9x-4)^2+(9y-7)^2+(9z-5)^2)}.$$
- $F_2^{3d}(x, y) = (1 - x^2)(1 - y^2)(1 - z^2).$
- $F_3^{3d}(x, y) = \sqrt{\frac{18}{\pi}}e^{-(x^2+81y^2+z^2)}.$
- $F_4^{3d}(x, y) = \cos(6z)\frac{1.25+\cos(5.4y)}{6+6(3x-1)^2}.$

The performance of our quasi-sparse kernel-based interpolant is evaluated in a grid of $20 \times 20 \times 20$ equally spaced points in the cube $[-1, 1]^3$. The results are summarised in Tables 5–8. Due to the long computing time and the unavailability of a parallel computing environment, we evaluate only $n = 7$ levels for the multilevel algorithm. The results are similar to our observation for the numerical experiments in $d = 2$ as we do not observe any convergence of the quasi-sparse interpolation with equally spaced grids, but the approximations converge for Chebyshev points in all test functions.

Table 5. Quasi-sparse interpolation with a Gaussian kernel using equally spaced grids and Chebyshev point grid results for the test function F_1^{3d} .

SGnodes	DoFs	Equal			Chebyshev		
		Max Error	RMS Error	Time	Max Error	RMS Error	Time
27	27	0.82746	0.11080	0.051	0.8143	0.10688	0.444
81	162	0.83700	0.11181	0.168	0.8221	0.10694	0.168
225	630	0.83172	0.10873	0.552	0.7755	0.09306	0.553
593	1997	0.82756	0.10771	1.647	0.7349	0.08037	1.608
1505	5687	0.80391	0.10663	4.165	0.7023	0.06938	4.117
3713	15,188	0.79746	0.10614	10.798	0.6484	0.05622	10.655
8961	38,868	0.77528	0.10574	26.174	0.5836	0.04364	25.607

Table 6. Quasi-sparse interpolation with a Gaussian kernel using equally spaced grids and Chebyshev point grid results for the test function F_2^{3d} .

SGnodes	DoFs	Equal			Chebyshev		
		Max Error	RMS Error	Time	Max Error	RMS Error	Time
27	27	0.9365	0.34503	0.055	0.9365	0.32901	0.045
81	162	0.8978	0.33059	0.217	0.9262	0.31016	0.171
225	630	0.8802	0.32260	0.557	0.7728	0.25254	0.560
593	1997	0.8733	0.31900	1.630	0.6776	0.20495	1.616
1505	5687	0.8708	0.31756	4.172	0.5138	0.15122	4.154
3713	15,188	0.8699	0.31703	10.468	0.3677	0.10426	10.389
8961	38,868	0.8697	0.31684	25.953	0.2402	0.06723	25.705

Table 7. Quasi-sparse interpolation with a Gaussian kernel using equally spaced grids and Chebyshev point grid results for the test function F_3^{3d} .

SGnodes	DoFs	Equal			Chebyshev		
		Max Error	RMS Error	Time	Max Error	RMS Error	Time
27	27	2.2099	0.32267	0.046	2.0758	0.33152	0.045
81	162	2.1726	0.31437	0.169	2.2421	0.32432	0.183
225	630	2.1558	0.31386	0.559	1.8688	0.29868	0.555
593	1997	2.1364	0.31547	1.625	1.7660	0.27446	1.614
1505	5687	2.1096	0.31310	4.150	1.4888	0.22923	4.176
3713	15,188	2.0936	0.31163	10.760	1.2753	0.17821	10.662
8961	38,868	2.0862	0.31130	25.750	1.0220	0.13504	25.441

Table 8. Quasi-sparse interpolation with a Gaussian kernel using equally spaced grids and Chebyshev point grid results for the test function F_4^{3d} .

SGnodes	DoFs	Equal			Chebyshev		
		Max Error	RMS Error	Time	Max Error	RMS Error	Time
27	27	0.3849	0.07847	0.045	0.4322	0.09322	0.042
81	162	0.3662	0.07535	0.202	0.3769	0.07620	0.166
225	630	0.3578	0.07125	0.583	0.4280	0.07657	0.543
593	1997	0.3394	0.06952	1.647	0.3923	0.06564	1.616
1505	5687	0.3339	0.06877	4.163	0.3448	0.05411	4.116
3713	15,188	0.3282	0.06837	10.471	0.3018	0.04450	10.294
8961	38,868	0.3249	0.06827	25.972	0.2531	0.03517	25.899

8. Numerical Experiments for Multilevel Quasi-Sparse Interpolation

In this section, we illustrate the performance of our proposed multilevel quasi-sparse kernel b-sed interpolation method for the same test functions, F_1^{2d} , F_2^{2d} , F_3^{2d} , and F_4^{2d} , as introduced earlier. We evaluate the interpolations at a grid of 50×50 points in the rectangle $[-1, 1] \times [-1, 1]$. The results are shown in Tables 9–12 for the test functions in $d = 2$. Some of the approximate functions are also plotted in Figures 14–21. For both equally spaced sparse grids and Chebyshev point-based sparse grids, we observe that the maximum error and the root mean square error decrease with the increase in level n . The multilevel quasi-sparse interpolation results are even better with Chebyshev points instead of equally spaced grids.

Table 9. Multilevel quasi-sparse interpolation with a Gaussian kernel using equally spaced grids and Chebyshev grid results for the test function F_1^{2d} .

SGnodes	DoFs	Equal			Chebyshev		
		Max Error	RMS Error	Time	Max Error	RMS Error	Time
9	9	0.92616	0.24884	0.009	0.89474	0.23177	0.011
21	39	0.83928	0.21052	0.042	0.79237	0.19255	0.043
49	109	0.73335	0.16992	0.078	0.66953	0.15399	0.075
113	271	0.64322	0.13556	0.158	0.52317	0.10671	0.161
257	641	0.5213	0.10474	0.342	0.43208	0.07463	0.350
577	1475	0.4117	0.07972	0.865	0.33932	0.04972	0.789
1281	3333	0.3143	0.06017	2.569	0.22871	0.03020	2.321
2817	7431	0.2393	0.04526	7.121	0.11902	0.01538	7.031
6145	16,393	0.1799	0.03400	25.948	0.04341	0.00589	26.286
13,313	35,851	0.1361	0.02553	103.060	0.01552	0.00155	104.900

Table 10. Multilevel quasi-sparse interpolation with a Gaussian kernel using equally spaced grids and Chebyshev grid results for the test function F_2^{2d} .

SGnodes	DoFs	Equal			Chebyshev		
		Max Error	RMS Error	Time	Max Error	RMS Error	Time
9	9	0.8409	0.45491	0.009	0.8481	0.41803	0.009
21	39	0.6510	0.36356	0.043	0.6838	0.31352	0.039
49	109	0.4929	0.27946	0.080	0.4557	0.19141	0.079
113	271	0.3731	0.21150	0.172	0.2341	0.09706	0.158
257	641	0.2813	0.15917	0.341	0.0856	0.03728	0.349
577	1475	0.2130	0.11955	0.862	0.0278	0.01046	0.812
1281	3333	0.1596	0.08974	2.274	0.0087	0.00236	2.229
2817	7431	0.1196	0.06735	7.312	0.0038	0.00075	7.041
6145	16,393	0.0904	0.05054	26.093	0.0028	0.00034	26.036
13,313	35,851	0.0675	0.03793	104.410	0.0012	0.00015	106.930

Table 11. Multilevel quasi-sparse interpolation with a Gaussian kernel using equally spaced grids and Chebyshev grid results for the test function F_3^{2d} .

SGnodes	DoFs	Equal			Chebyshev		
		Max Error	RMS Error	Time	Max Error	RMS Error	Time
9	9	1.97480	0.40313	0.009	1.8427	0.48854	0.009
21	39	1.60630	0.35940	0.044	1.4970	0.56101	0.041
49	109	1.31510	0.30034	0.080	1.3375	0.50115	0.074
113	271	1.06830	0.24349	0.159	1.1287	0.38064	0.195
257	641	0.83075	0.18909	0.375	0.8010	0.26008	0.332
577	1475	0.63185	0.14348	0.866	0.5141	0.12958	0.805
1281	3333	0.47352	0.10812	2.259	0.2914	0.05399	2.226
2817	7431	0.35422	0.08115	7.053	0.1212	0.01992	6.980
6145	16,393	0.26586	0.06086	26.270	0.0339	0.00579	31.276
13,313	35,851	0.19815	0.04569	104.750	0.0065	0.00116	104.840

Table 12. Multilevel quasi-sparse interpolation with a Gaussian kernel using equally spaced grids and Chebyshev grid results for the test function F_4^{2d} .

SGnodes	DoFs	Equal			Chebyshev		
		Max Error	RMS Error	Time	Max Error	RMS Error	Time
9	9	0.3386	0.08989	0.011	0.2952	0.07502	0.025
21	39	0.2906	0.07151	0.043	0.2692	0.06758	0.038
49	109	0.2348	0.05599	0.101	0.1914	0.05441	0.109
113	271	0.1847	0.04338	0.162	0.1257	0.04157	0.227
257	641	0.1407	0.03316	0.345	0.1080	0.02857	0.332
577	1475	0.1068	0.02514	0.880	0.0690	0.01629	0.886
1281	3333	0.0811	0.01900	2.230	0.0317	0.00701	2.396
2817	7431	0.0615	0.01434	7.056	0.0104	0.00223	7.005
6145	16,393	0.0477	0.01082	25.989	0.0024	0.00029	25.746
13,313	35,851	0.0370	0.00817	107.280	0.0008	0.00009	104.010

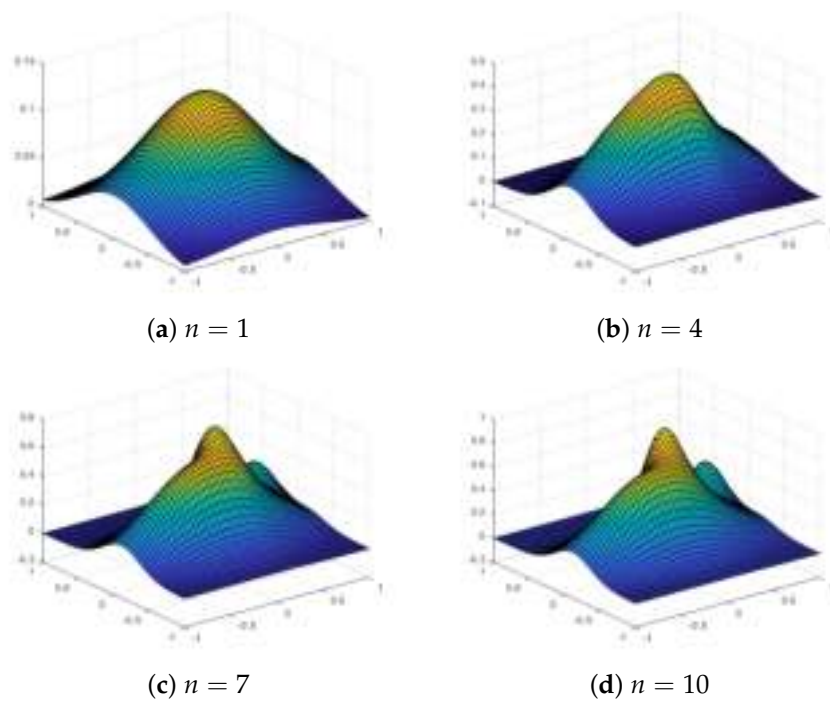


Figure 14. Multilevel quasi-sparse approximations using an equally spaced grid with level n and a Gaussian kernel for the test function F_1^{2d} .

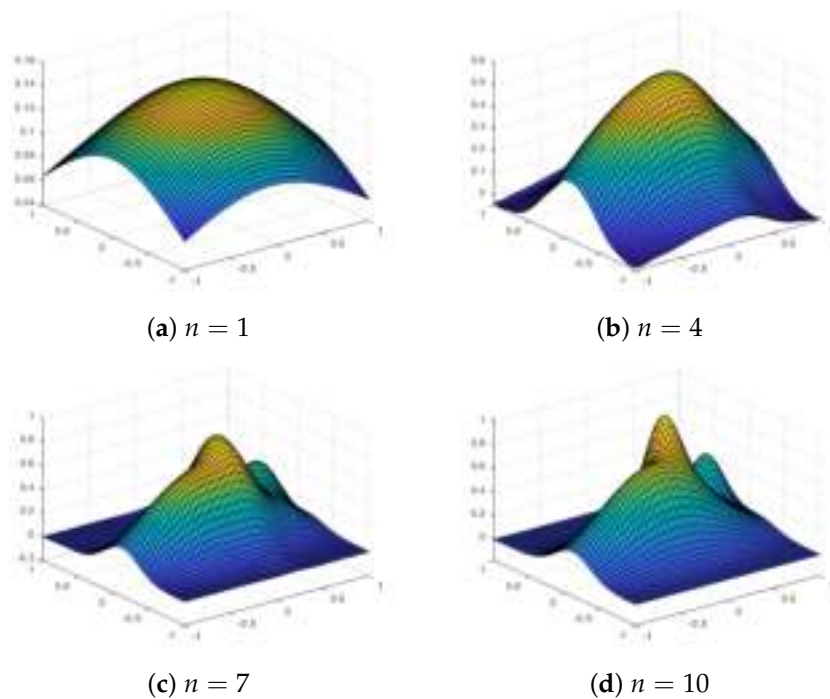


Figure 15. Multilevel quasi-sparse approximations using a Chebyshev grid with level n and a Gaussian kernel for the test function F_1^{2d} .

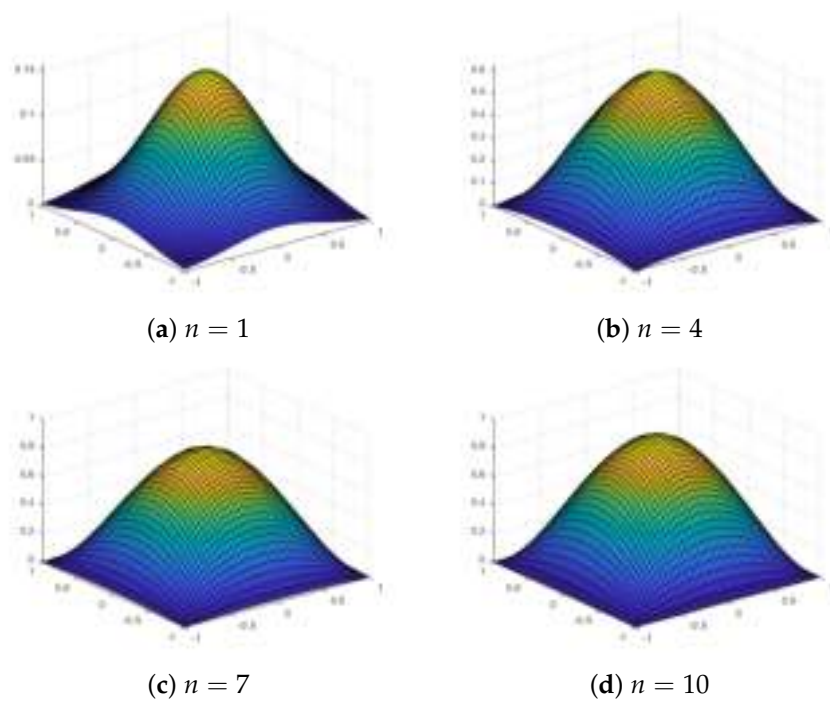


Figure 16. Multilevel quasi-sparse approximations using an equally spaced grid with level n and a Gaussian kernel for the test function F_2^{2d} .

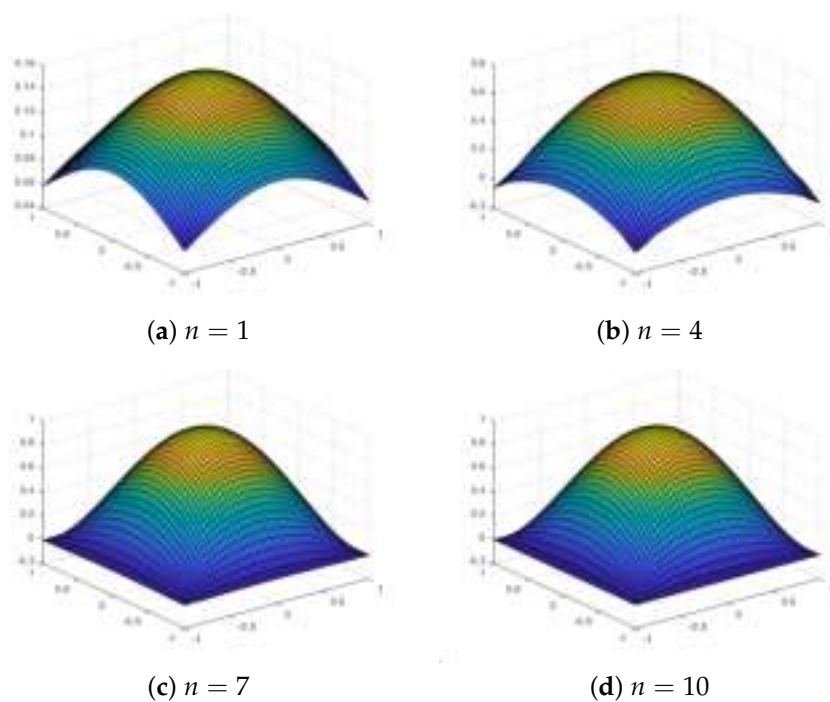


Figure 17. Multilevel quasi-sparse approximations using a Chebyshev grid with level n and a Gaussian kernel for the test function F_2^{2d} .

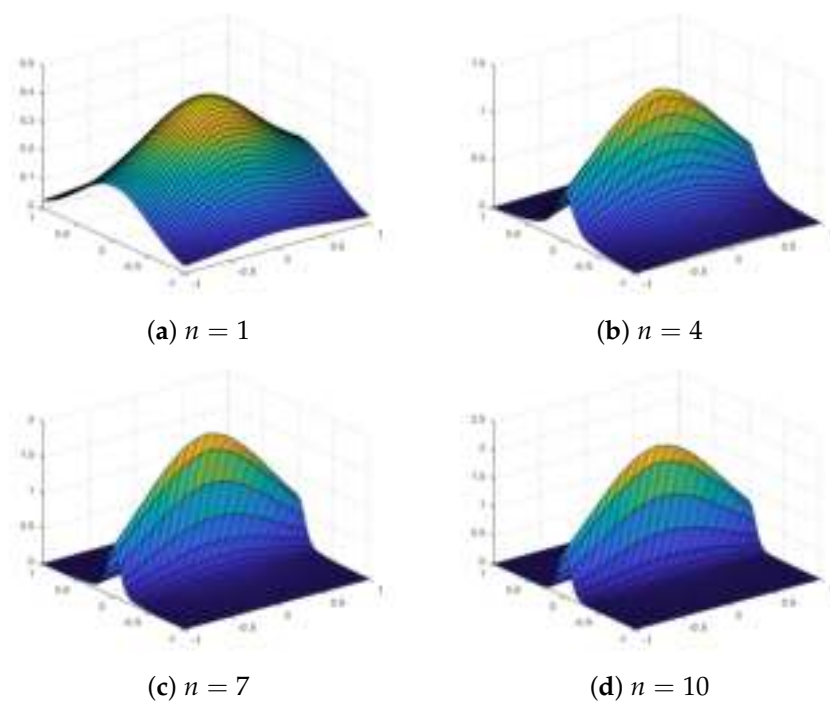


Figure 18. Multilevel quasi-sparse approximations using an equally spaced grid with level n and a Gaussian kernel for the test function F_3^{2d} .

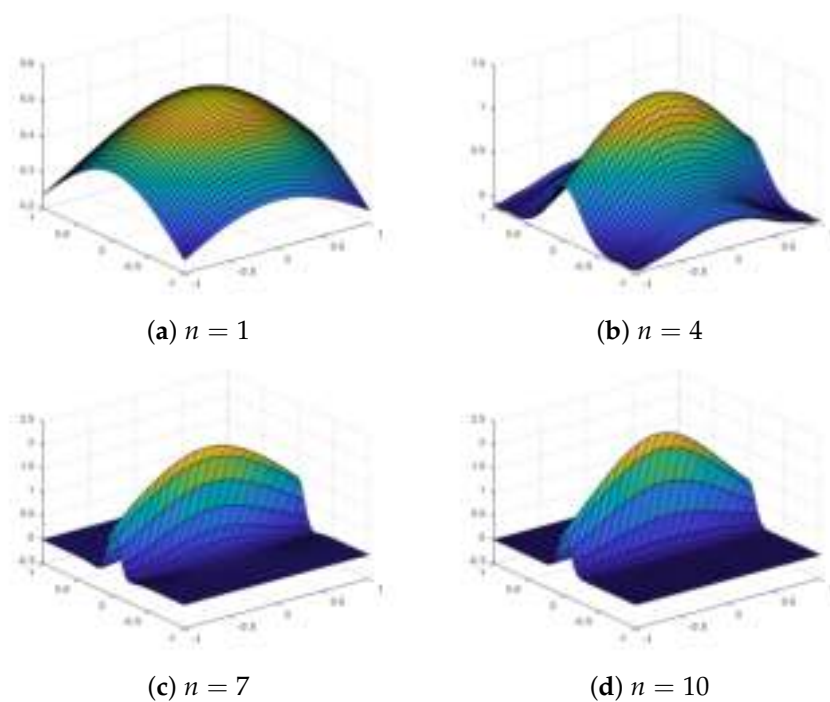


Figure 19. Multilevel quasi-sparse approximations using a Chebyshev grid with level n and a Gaussian kernel for the test function F_3^{2d} .

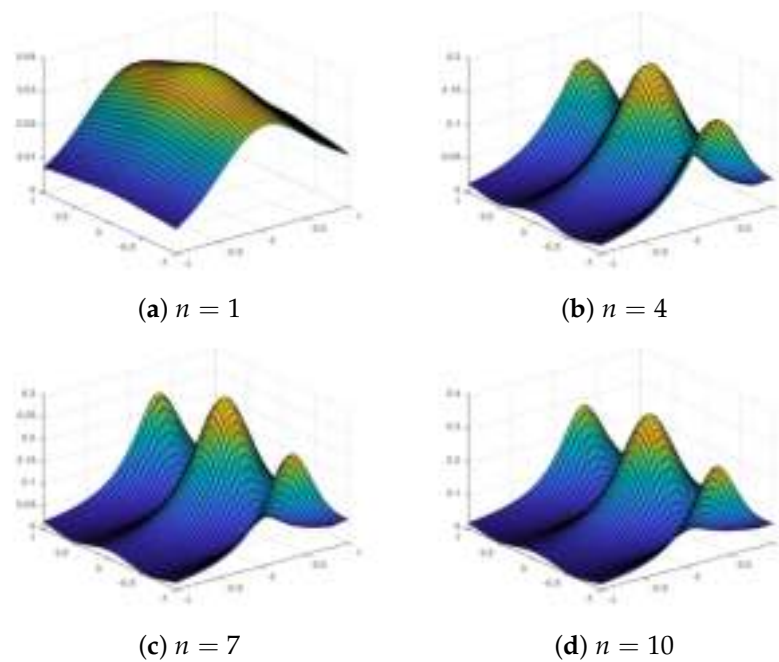


Figure 20. Multilevel quasi-sparse approximations using an equally spaced grid with level n and a Gaussian kernel for the test function F_4^{2d} .

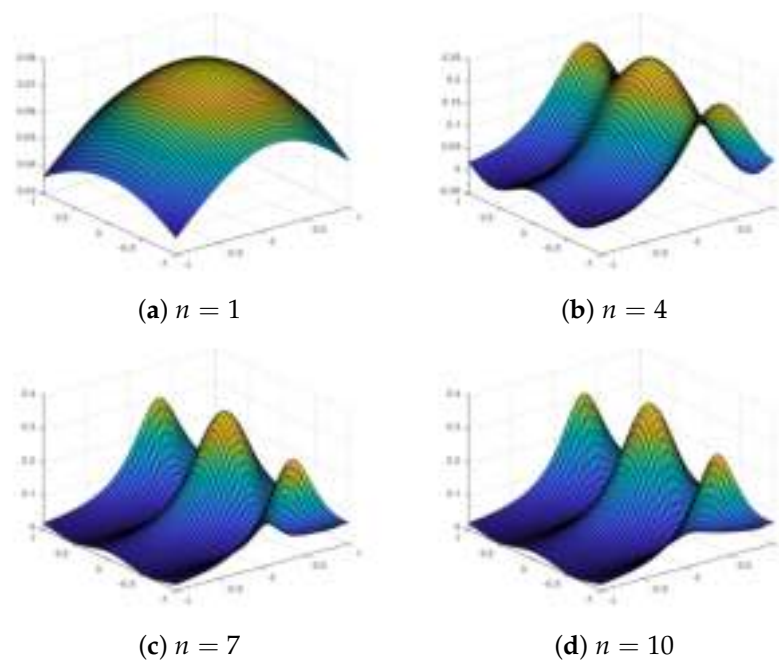


Figure 21. Multilevel quasi-sparse approximations using a Chebyshev grid with level n and a Gaussian kernel for the test function F_4^{2d} .

Next, we illustrate the application of our multilevel quasi-sparse interpolation with a Gaussian kernel for the test functions F_1^{3d} , F_2^{3d} , F_3^{3d} , and F_4^{3d} in dimension $d = 3$, as introduced earlier. For $d = 3$, we evaluate the multilevel interpolation using a grid of $20 \times 20 \times 20$ points in the cube $[-1, 1]^3$. Tables 13–16 show the error estimates with maximum absolute error and root mean square errors for the proposed interpolant for levels up to $n = 7$. Again, we observe that the errors decrease with the level n for the multilevel procedure for both equally spaced grids and Chebyshev points. However, errors are much smaller for Chebyshev points at higher levels.

Table 13. Multilevel quasi-sparse interpolation with a Gaussian kernel using equally spaced grids and Chebyshev grid results for the test function F_1^{3d} .

SGnodes	DoFs	Equal			Chebyshev		
		Max Error	RMS Error	Time	Max Error	RMS Error	Time
27	27	0.82746	0.11080	0.040	0.8143	0.10688	0.037
81	162	0.80642	0.10328	0.236	0.7753	0.09751	0.241
225	630	0.78322	0.09384	0.801	0.7306	0.08611	0.821
593	1997	0.75882	0.08465	2.480	0.6650	0.07091	2.469
1505	5687	0.71458	0.07554	8.196	0.6213	0.05736	8.085
3713	15,188	0.66788	0.06704	30.099	0.5785	0.04466	29.483
8961	38,868	0.60584	0.05910	128.210	0.5211	0.03287	127.010

Table 14. Multilevel quasi-sparse interpolation with a Gaussian kernel using equally spaced grids and Chebyshev grids results for the test function F_2^{3d} .

SGnodes	DoFs	Equal			Chebyshev		
		Max Error	RMS Error	Time	Max Error	RMS Error	Time
27	27	0.9365	0.34503	0.041	0.9365	0.32901	0.049
81	162	0.8405	0.31540	0.202	0.8602	0.28440	0.201
225	630	0.7393	0.28145	0.782	0.7125	0.21744	0.783
593	1997	0.6451	0.24828	2.514	0.5097	0.14621	2.475
1505	5687	0.5636	0.21794	8.222	0.3128	0.08456	8.113
3713	15,188	0.4938	0.19096	29.847	0.1556	0.04162	29.519
8961	38,868	0.4335	0.16720	128.580	0.0600	0.01689	126.620

Table 15. Multilevel quasi-sparse interpolation with a Gaussian kernel using equally spaced grids and Chebyshev grid results for the test function F_3^{3d} .

SGnodes	DoFs	Equal			Chebyshev		
		Max Error	RMS Error	Time	Max Error	RMS Error	Time
27	27	2.2099	0.32267	0.039	2.0758	0.33152	0.037
81	162	2.0083	0.29220	0.237	1.9006	0.35681	0.204
225	630	1.8151	0.26318	0.833	1.6722	0.36067	1.108
593	1997	1.6279	0.23656	2.493	1.3596	0.32133	2.630
1505	5687	1.4391	0.20954	8.325	1.0261	0.25248	8.110
3713	15,188	1.2623	0.18412	29.474	0.7688	0.16061	29.341
8961	38,868	1.1086	0.16129	130.150	0.5329	0.08978	127.010

Table 16. Multilevel quasi-sparse interpolation with a Gaussian kernel using equally spaced grids and Chebyshev grid results for the test function F_4^{3d} .

SGnodes	DoFs	Equal			Chebyshev		
		Max Error	RMS Error	Time	Max Error	RMS Error	Time
27	27	0.3849	0.07847	0.064	0.4322	0.09322	0.064
81	162	0.3815	0.07424	0.252	0.4136	0.08504	0.207
225	630	0.3697	0.06821	0.805	0.4424	0.08104	0.777
593	1997	0.3409	0.06137	2.488	0.4245	0.06963	2.491
1505	5687	0.3097	0.05468	8.152	0.3540	0.05359	8.124
3713	15,188	0.2746	0.04841	29.676	0.2721	0.04173	29.474
8961	38,868	0.2425	0.04275	128.450	0.1914	0.03210	126.960

9. Convergence with Computation Time

In Figure 22, the root mean square (RMS) error is plotted against the computation time for the same test functions in $d = 2$, and similar plots of RMS error against time for the test functions in $d = 3$ are plotted in Figure 23. These figures show that the quasi-sparse interpolation with equally spaced grids does not converge in error as computational time increases. On the other hand, the errors decrease quite fast and converge for quasi-sparse interpolation with Chebyshev points. Multilevel quasi-sparse interpolation with an equally spaced grid improves errors in approximation and shows some convergence. Still, Chebyshev points present much better convergence results than equally spaced points. It is also interesting to note that, at a similar computational time (in seconds), quasi-sparse interpolation with Chebyshev points yields better results than the multilevel method, though multilevel methods execute only at a smaller number of levels at a comparable time.

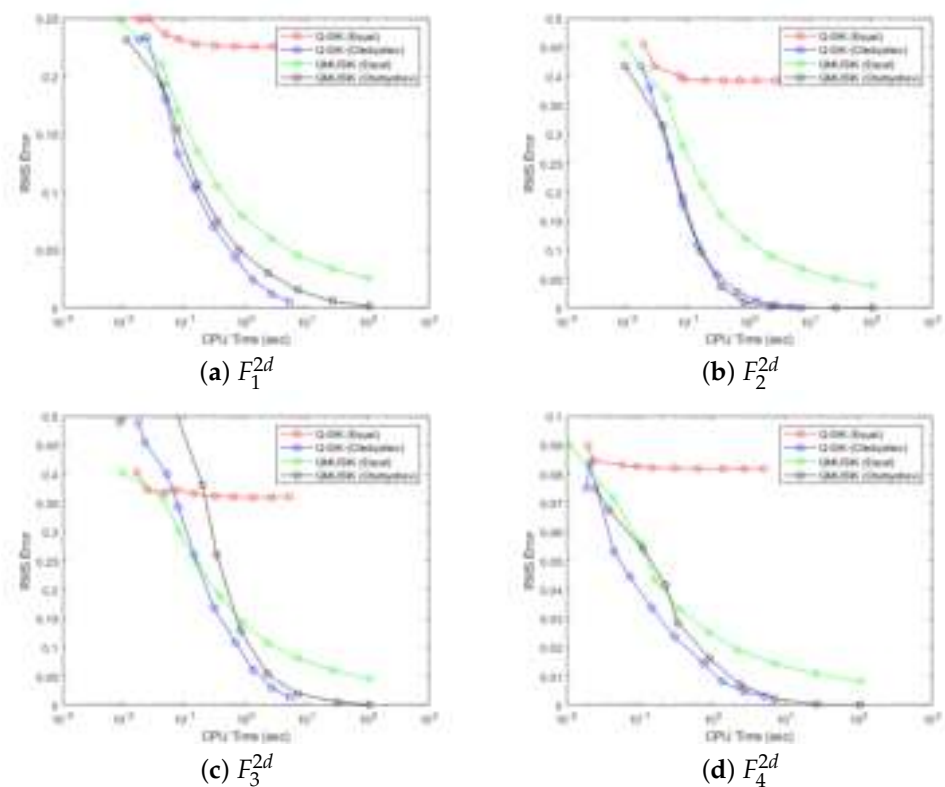


Figure 22. RMS error against the computation time for the test functions in $d = 2$.

Figure 23 shows the RMS error in approximation against the computational time (in seconds) for the test functions in dimension $d = 3$. Here, we do not observe the convergence in any of the algorithms as all except equally spaced points with quasi-sparse interpolation are show a decreasing trend in errors with computational time. Here, better performance is also obtained for Chebyshev points and Chebyshev points with only quasi-sparse interpolations, which are competitive with multilevel algorithms for a similar computational time and better than multilevel algorithms with equally spaced points.

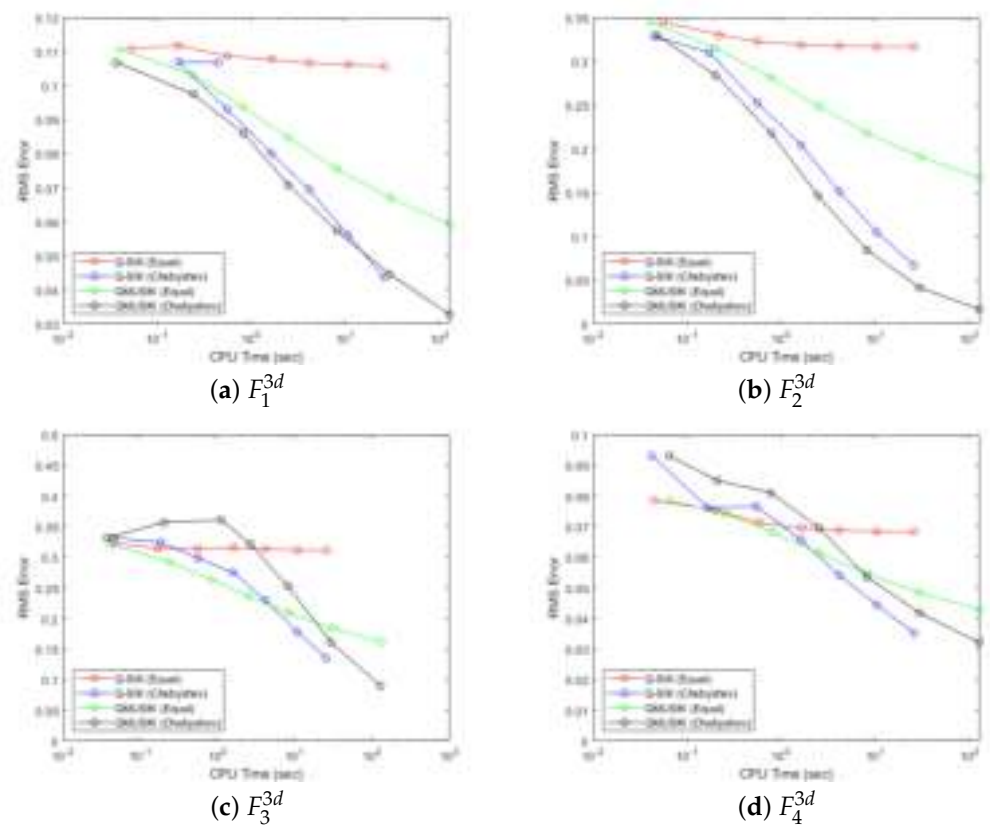


Figure 23. RMS error against the computation time for the test functions in $d = 3$.

10. Discussion and Findings

In this article, we explored an algorithm for quasi-interpolation with a Gaussian kernel based on sparse grids. We observed that quasi-sparse kernel-based interpolation approximates two- and three-dimensional functions quite reasonably. However, quasi-interpolation is not guaranteed to converge. We observed that even for large values of the level n of the grids, the root mean square error of the approximation does not converge to 0 if we use equally spaced sparse grids. However, if we use Chebyshev points, the root mean square error of the approximation steadily decreases with level n . We obtain a much better approximation with Chebyshev points than an equally spaced grid for larger values of n . However, Chebyshev points are more dense along the boundary of the domain and more sparse in the centre. For this reason, Chebyshev points yield a very poor approximation in the centre for smaller values of n , which is evident from the plots with Chebyshev approximations with $n = 1$ or $n = 2$.

The way grids $\mathbb{G}_\ell$ are constructed, it is possible to implement the algorithm in parallel to make it faster. Otherwise, the repeat of evaluation of the anisotropic interpolant at many grid points for different values of n slows down the execution. We can see that for $d = 2$, the level $n = 10$ has 13313 distinct points only, but these points are revisited many times, yielding a total number of 35851 points used for computation. This increases the function evaluation by nearly three times. For $d = 3$ and the level $n = 7$, the number of unique grid points is only 8961, but the points used for computation reach 38868. Due to this, the time-complexity of the algorithm increases by many times; one should either implement the algorithm for parallel computing only or improve the algorithm to avoid revisiting grid points for different multi-index ℓ .

In this article, we also proposed multilevel quasi-sparse kernel-based interpolation based on equally spaced grids as well as sparse grids based on Chebyshev points. We illustrated our proposed methods using some numerical test functions. We observed that the proposed multilevel quasi-procedure performs better than the proposed quasi-

sparse interpolation method. However, the computational time for the multilevel method is considerably higher than for simple quasi-sparse interpolation. This is because the multilevel method requires residual functions to be evaluated at every sparse grid point, and as the level n increases, the number of sparse grid points also increases exponentially, which leads to a huge increase in computational time. On a positive note, we would like to mention that due to the nature of sparse grid construction, it is possible to implement the algorithm for parallel computing, which will reduce the computational time considerably.

Furthermore, the multilevel quasi-sparse algorithm requires substantially less computational time compared to other multilevel methods proposed in the literature as it does not require solving a system of linear equations at every level of approximations with the residual functions. Due to the implementation of the quasi-sparse algorithm at every level, the proposed multilevel quasi-sparse method will have linear time complexity with the level n .

11. Conclusions

The paper has successfully demonstrated the viability and advantages of using multilevel quasi-interpolation techniques on Chebyshev sparse grids. In comparison to conventional methods, the multilevel quasi-sparse interpolation with Gaussian kernels showed promising performance metrics. This paper also elaborated on the implementation complexities and challenges, providing an efficient way to construct sparse grids and compute anisotropic quasi-interplants. While the methodology was effective in lower dimensions, i.e., two dimensions and three dimensions, limitations began to appear as the computational time and complexity increased, indicating areas for future research and optimisation.

The numerical experiments yielded insightful data on the performance of this technique, showcasing its capacity for faster convergence with computation time in specific settings. The study acts as a foundational resource for future research aiming to employ multilevel quasi-interpolation techniques for various applications, especially in high-dimensional computational problems.

Funding: This research received no external funding.

Data Availability Statement: Data are contained within the article.

Acknowledgments: The author is thankful to the University of Leicester for providing the academic environment and resources. Additionally, sincere appreciation goes to his supervisors, Ruslan Davidchack and Jeremy Levesley, for their invaluable guidance and mentorship. Special thanks are also extended to Taibah University for their generous financial support throughout his doctoral studies.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Alsharif, F. Multilevel Quasi-Interpolation With a Gaussian Kernel Using Chebyshev Points—Numerical and Theoretical Studies Using the Radial Basis Function Approach. Ph.D. Thesis, University of Leicester, Leicester, UK, 2023. [\[CrossRef\]](#)
2. Alsharif, F. Quasi-Interpolation on Chebyshev Grids with Boundary Corrections. *Computation* **2024**, *12*, 100. [\[CrossRef\]](#)
3. Franz, T.; Wendland, H. Multilevel quasi-interpolation. *IMA J. Numer. Anal.* **2022**, *43*, 2934–2964. [\[CrossRef\]](#)
4. Georgoulis, E.; Levesley, J.; Subhan, F. Multilevel sparse kernel based interpolation. *SIAM J. Sci. Comput.* **2013**, *35*, 815–832. [\[CrossRef\]](#)
5. Beatson, R.; Powell, M. Univariate multiquadric approximation: quasi-interpolation to scattered data. *Constr. Approx* **1992**, *8*, 275–288. [\[CrossRef\]](#)
6. Wu, Z.; Robert, S. Shape preserving properties and convergence of univariate multiquadric quasi-interpolation. *Acta Math. Appl. Sin.* **1994**, *10*, 441–446. [\[CrossRef\]](#)
7. Wu, Z.; Liu, J. Generalized strang-fix condition for scattered data quasi-interpolation. *Adv. Comput. Math.* **2005**, *23*, 201–214. [\[CrossRef\]](#)
8. Floater, M.S.; Iske, A. Multistep scattered data interpolation using compactly supported radial basis functions. *J. Comput. Appl. Math* **1996**, *73*, 65–78. [\[CrossRef\]](#)
9. Hales, S.J.; Levesley, J. Error estimates for multilevel approximation using polyharmonic splines. *Numer. Algorithms* **2002**, *30*, 1–10. [\[CrossRef\]](#)
10. Hubbert, S.; Levesley, J. Convergence of Multilevel Stationary Gaussian Quasi-Interpolation. *arXiv* **2017**, arXiv:1609.02457.
11. Franke, R. Scattered data interpolation: tests of some methods. *Math. Comput.* **1982**, *38*, 181–200.

12. Fasshauer, G.E. *Meshfree Approximation Methods with MATLAB of Interdisciplinary Mathematical Sciences*; World Scientific Publishing Co. Pte. Ltd.: Hackensack, NJ, USA, 2007; Volume 6.
13. Beatson, R.; Davydov, D.; Levesley, J. Error bounds for anisotropic rbf interpolation. *J. Approx. Theory* **2010**, *162*, 512–527. [[CrossRef](#)]
14. Casciola, G.; Montefusco, L.B.; Morigi, S. The regularizing properties of anisotropic radial basis functions. *Appl. Math. Comput.* **2007**, *190*, 1050–1062. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Using Machine Learning Algorithms to Develop a Predictive Model for Computing the Maximum Deflection of Horizontally Curved Steel I-Beams

Elvis Ababu ^{1,*}, George Markou ^{1,2,*}  and Sarah Skorpen ¹¹ Civil Engineering Department Hatfield Campus, University of Pretoria, Pretoria 0028, South Africa; sarah.skorpen@up.ac.za² Civil Engineering Department, Neapolis University Pafos, Danais Avenue, Paphos 8042, Cyprus

* Correspondence: u17061785@tuks.co.za (E.A.); george.markou@up.ac.za (G.M.)

Abstract: Horizontally curved steel I-beams exhibit a complicated mechanical response as they experience a combination of bending, shear, and torsion, which varies based on the geometry of the beam at hand. The behaviour of these beams is therefore quite difficult to predict, as they can fail due to either flexure, shear, torsion, lateral torsional buckling, or a combination of these types of failure. This therefore necessitates the usage of complicated nonlinear analyses in order to accurately model their behaviour. Currently, little guidance is provided by international design standards in consideration of the serviceability limit states of horizontally curved steel I-beams. In this research, an experimentally validated dataset was created and was used to train numerous machine learning (ML) algorithms for predicting the midspan deflection at failure as well as the failure load of numerous horizontally curved steel I-beams. According to the experimental and numerical investigation, the deep artificial neural network model was found to be the most accurate when used to predict the validation dataset, where a mean absolute error of 6.4 mm (16.20%) was observed. This accuracy far surpassed that of Castigliano's second theorem, where the mean absolute error was found to be equal to 49.84 mm (126%). The deep artificial neural network was also capable of estimating the failure load with a mean absolute error of 30.43 kN (22.42%). This predictive model, which is the first of its kind in the international literature, can be used by professional engineers for the design of curved steel I-beams since it is currently the most accurate model ever developed.

Keywords: structural engineering; structural steel; curved beams; machine learning; finite element modelling



Citation: Ababu, E.; Markou, G.; Skorpen, S. Using Machine Learning Algorithms to Develop a Predictive Model for Computing the Maximum Deflection of Horizontally Curved Steel I-Beams. *Computation* **2024**, *12*, 151. <https://doi.org/10.3390/computation12080151>

Academic Editors: Manolis Georgioudakis, Vagelis Plevris and Mahdi Kioumars

Received: 6 March 2024

Revised: 26 May 2024

Accepted: 27 May 2024

Published: 24 July 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Although curved beams have been in existence for centuries, minimal guidance has been available on the design of curved steel elements. The fabrication of curved steel beams began in the 19th century when steel members were cast in a curved profile or built up from components into a curved profile. Fabrication advanced, utilising three-roll pressing or induction bending in situations where residual stresses were particularly significant or smaller radii were required [1]. However, the majority of steel used for construction is formed through the use of roller bending, which is a cold process. When curving open sections using cold processes, the flanges exert a significant force on the web, which could lead to local buckling. Therefore, additional rolls are provided inside the tension flange for sections that are susceptible to local buckling (Figure 1).

Horizontally curved beams experience significant residual stresses due to the manner in which they are formed. A steel member rolled at room temperature experiences mechanical residual stresses (compressive and tensile) that reduce the ultimate strength of the member by increasing its flexibility, which subsequently causes a decrease in buckling strength [2]. Tensile residual stresses are particularly detrimental as they are often the cause

of fatigue failure and stress corrosion cracking, whereas compressive residual stresses can be somewhat beneficial as they can mitigate the origination and propagation of cracks [3].

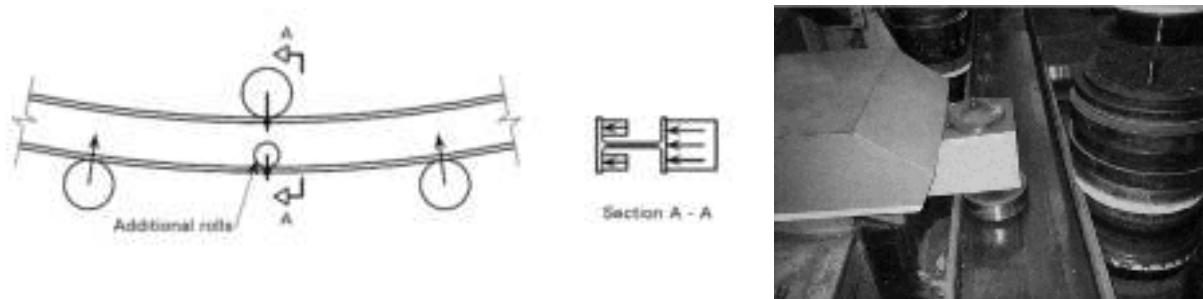


Figure 1. Additional rolls to prevent web buckling on an I-section [4].

The main challenge encountered with horizontally curved beams is that due to the geometry of the beam when normal loads are applied, the beam experiences a complex combination of bending, shear, and torsion simultaneously. Torsional moments can be visualised from the deflected shape as the compression flange tends to deflect laterally away from the centre of curvature [5]. When bending and torsional moments are applied simultaneously to a beam, the coupling of these forces tends to reduce the carrying capacity of the member. Research has found that the behaviour of horizontally curved beams is dependent on the R/L ratio, which can also be represented as the span angle. It has been noted that when the span angle is less than 1° , the beam responds similarly to a straight beam and is dominated by flexure (bending). When the span angle is larger than 20° , the beam behaviour of the beam is primarily dominated by torsion. When the span angle is between 1° and 20° , both bending and torsion significantly impact the behaviour of the beam [6].

In straight I-beams, lateral torsional buckling is easily observable through the lateral displacement and rotation of the member. In curved I-beams, however, this behaviour is always present due to the torsional moments experienced. This behaviour was termed lateral–torsional–vertical behaviour by Lee, et al. [7]. Therefore, due to initial curvature, similar to initial out-of-straightness in columns, bifurcation-type lateral torsional buckling may not be observed in curved steel I-beams. However, it has been noted that, as seen with straight beams, the flexural strength of curved beams decreases for members that are susceptible to lateral torsional buckling [6]. The effect of lateral torsional buckling is negligible when the angle between torsional restraints is less than 22.5° . Should this angle be larger than 22.5° , the American Institute of Steel Construction (AISC) recommends that the beam be handled as a straight beam with an adjusted lateral–torsional buckling modification factor that accounts for the curvature.

Over the years, numerous researchers have attempted to provide analytical formulae for horizontally curved members; however, very few have focused on equations regarding deflection and rotation. Wong [8] modified Castigliano’s second theorem, which was originally developed for straight beams, and made it applicable to horizontally curved beams. Castigliano’s second theorem is a strain energy method that is used to calculate deflection. The equation generated can be seen in Equation (1). C_d can be calculated from Figure 2. It is important to note here that Equation (1) is limited to beams that are fixed on both sides and a point load is applied at the centre point. Dahlberg [9] created various equations that can be used for various boundary and load conditions. However, these equations are quite difficult to implement and are therefore rarely used in practice. Furthermore, Dahlberg [9] did not compare the formula with experimental or finite element results.

$$\Delta_{0.5\theta} = \frac{P_r^3}{EI} C_d \quad (1)$$

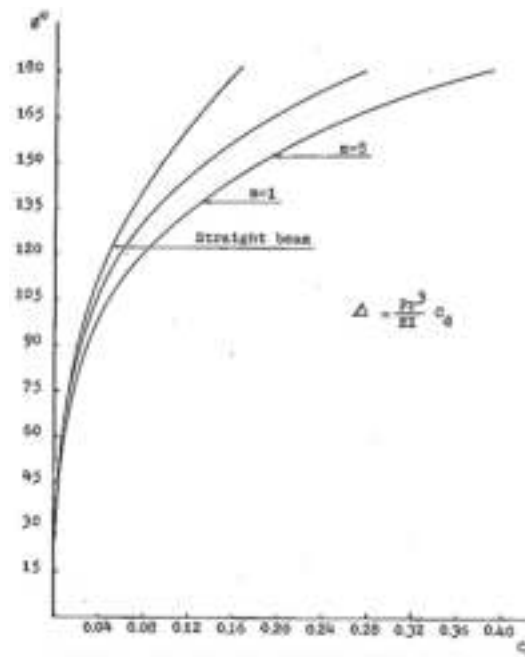


Figure 2. Variation of deflection coefficients with span angle for curved beams loaded with a concentrated load [8].

Castigliano's second theorem can also be applied to determine the rotation of curved beams; however, this is often even more tedious than the equations necessary for deflection. However, a simplified equation is provided by AISC (Equation (2)). The symbol h_0 represents the height of the member and $\Delta_{\max}$ represents the maximum displacement of the member [6].

This equation is derived using the M/R method, which is used extensively in the design of horizontally curved beams, where the curved beam is modelled as a straight beam with a length equal to the curved member length. The torsional moment is then accounted for by a separate equation:

$$\theta_{\max} = \tan^{-1} \left(\frac{2\Delta_{\max}}{h_0} \right) \quad (2)$$

Various assumptions have been during the derivation of this formula, including that the thickness of each plate element is small relative to the width, which is, in turn, small relative to the span. The stresses due to warping are assumed to be negligible.

Modelling curved beams in FEM causes further complexities. The majority of commercially available software applications model curved beams as a series of straight beams, which provides sufficient accuracy for most design purposes. If the model experiences significant nonlinear behaviour, a convergence study is required in order to optimise the number of elements required, while the development of torsion further complicates the analysis process. Conventional beam finite elements (FEs) cannot be used to model horizontally curved steel I-beams that experience torsion, due to the existence of warping torsion. Conventional beam FEs account only for St Venant stiffness, which causes numerical error regarding computing the torsional deformation. Current beam FEs that incorporate warping stiffness are available; however, these are rarely incorporated in commercially available software and also tend to be less accurate. Three-dimensional FEs such as shell or solid elements are commonly used by researchers attempting to model the behaviour of curved beams. These elements accurately represent torsional behaviour without the need for numerous torsional constants. Previous researchers found that using solid FEs paired with nonlinear analyses led to accurate estimates of experimental results for both ultimate load and midspan deflection ([10–13]).

The combination of bending and shear as well as displacements and twisting rotations causes second-order bending actions in the plane of a curved I-beam. With an increase in load, these interactions grow rapidly and may ultimately cause early nonlinear behaviour and yielding, leading to reductions in the ultimate load-carrying capacity of the beam. This therefore necessitates the usage of nonlinear analyses when analysing horizontally curved beams, as opposed to general static (linear) analysis. Given the complexity of the problem, the use of FEs that are integrated with 3D numerical material models is also recommended, as in this research work.

2. Machine Learning Algorithms

This section briefly presents the various ways machine learning (ML) has been used to develop formulae that allow the design or mechanical investigation of civil engineering structures. ML and artificial intelligence (AI) have been implemented over recent decades in different fields as efficient tools used to predict analysis outputs for engineering problems that are deemed computationally demanding and highly complicated to solve analytically. The use of ML has largely eliminated the need for large numerical analysis, as it contains the ability to provide adequate estimates of the desired outputs. The main issue at hand, however, is that to train an ML algorithm, a large enough dataset is required. Therefore, the primary task of the majority of research work at this stage is the generation of datasets for various reinforced concrete and steel-related problems using either physical or validated numerical experiments. A meta-analysis of various case studies was presented by Markou, et al. [14]; however, the current paper focused on a single mechanical response problem in an attempt to provide a solution to a problem that has never been solved in the past. It is important to note that all ML algorithms that were used to perform the training, testing, and validation for the needs of this research work can be found on GitHub (<https://github.com/nbakas/nbml/>, accessed on 21 June 2023), through downloading nbml freeware. Furthermore, the datasets that were developed for the needs of this research work can be found through the following link (<https://github.com/nbakas/nbml/tree/a0d27c94dd590688815180ebf6428963a24ca245/datasets>, accessed on 1 July 2024), whereas the proposed models can be developed directly by the reader.

Various ML algorithms were used in this research, namely linear regression (LR), polynomial regression with hyperparameter tuning (POLYREG-HYT), hyperparameter tuning of extreme gradient boosting (XGBoost-HYT-CV), and parallel deep learning artificial neural networks with hyperparameter tuning (DANN-MPIH-HYT). The LR method was used as a point of comparison for the other ML algorithms (Markou, Bakas, Papadrakakis, and Chatzichristofis [14]).

POLYREG-HYT is useful for generating relatively accurate closed-form formulae and is applied to develop predictive models in higher-order classes. This relatively simplistic method provides a formula based on the nonlinear combination of all independent variables [14]. This model is based on the creation of nonlinear terms that are based on the independent variables up to the third degree. The algorithm then selects the nonlinear features that correspond to the minimum error. Originally, this methodology was utilised by Gravett, et al. [15], who made use of a simplistic approach in determining the number of features to use during training. Thereafter, that algorithm was improved and the use of hyperparameter tuning was introduced, as shown in Markou, Bakas, Papadrakakis, and Chatzichristofis [14], showcasing the proposed ML algorithm that has been used for the needs of this research work. The improved algorithm is outlined in Algorithm 1.

XGBoost-HYT-CV is a modification to the currently open-source extreme gradient boosting algorithm. The original XGBoost is a gradient-boosting library for ML problems such as classifying and regressing. The algorithm implements the gradient boosting framework, which is designed to be fast and scalable, making it suitable for large datasets [14]. For improving the current XGBoost algorithm, hyperparameter tuning was used. This tuning was found to exhibit accurate results compared to those of deep learning and require

less computing demand [14]. One benefit of this algorithm is its ability to locate and replace missing values in the training and testing datasets.

Algorithm 1: Feature selection algorithm for polynomial regression [14]

Data: $\mathbf{X}, \mathbf{y}, m_f$ (maximum number of features)
Result: Initialize $[o] = 1$ with the constant term $\in [p]$
 Solve Linear System $\mathbf{X}' \times \mathbf{a} = \mathbf{y}$, where $\mathbf{X}' \subset \mathbf{X}$, with $[o]$ columns.
 Compute regression errors e_1 .
 Set as optimal error $\hat{e} \leftarrow e_1$.
 Set as optimal indices $[\hat{o}] \leftarrow [o]$.
for $i \in [1, 2, \dots, l]$ **do**
 repeat
 Select an index $d \in [p]$ randomly.
 if $d \in [o]$ **then**
 $r \leftarrow \mathcal{U}(0, 1)$
 if $r < \frac{1}{2}$ **then**
 Select randomly $o_d \in [p] : o_d \notin [o]$
 $[o] \leftarrow ([o] \setminus d) \cup o_d$;
 else
 $[o] \leftarrow [o] \setminus d$;
 end
 else
 if $o < m_f$ **then**
 $[o] \leftarrow [o] \cup d$;
 else
 Select randomly $o_d \in [o]$
 $[o] \leftarrow ([o] \setminus o_d) \cup d$;
 end
 end
 until $\text{rank}(\mathbf{X}') \equiv o$;
 Solve Linear System $\mathbf{X}' \times \mathbf{a} = \mathbf{y}$.
 Compute regression error e_i .
 if $e_i < \hat{e}$ **then**
 $\hat{e} \leftarrow e_i$
 $[\hat{o}] \leftarrow [o]$
 else
 $[o] \leftarrow [\hat{o}]$
 end
end

The final algorithm to be engaged was DANN-MPIH-HYT, which has the ability to train on both small and large datasets through the use of parallel processing. The algorithm was programmed to train on both computer processing units (CPUs) and graphics processing units (GPUs) and the use of distributed computations was also provided. This led to faster training and testing times when using the algorithm, as deep learning is known to be particularly slow when dealing with training processes, especially in cases where large datasets are used [14]. A combination of Horovod with MPI was used to optimise the numerical procedure. The Horovod library was implemented for multi-GPU training. Using Horovod, one is able to take a single GPU training script and run it across numerous GPUs in parallel. Through the use of message-passing interface (MPI) commands, each process is initialised and assigned its MPI rank in a straightforward manner, which is achieved in fewer code changes compared with other approaches. Various experimental algorithms were tested by Markou, Bakas, Papadrakakis and Chatzichristofis [14] on the Cyclone Supercomputer, utilising PyTorch for computer vision as well as regression tasks, highlighting the efficiency of data parallelism.

ML combined with FEM modelling has been used to create accurate formulae for various engineering applications in recent years. Markou and Bakas [16] created formulae to determine the shear capacity of concrete slender beams without stirrups. In their research work, a total of 35,849 beams were created and four ML algorithms were used, namely linear regression, polynomial regression, XGBoost, and deep learning neural networks. In that study, it was noted that the XGBoost algorithm was the most accurate, boasting an error of 5.82%. A similar methodology was followed by [17–20], where design formulae and predictive models were developed in relation to RC and steel structural problems. In this context, the current research work validates numerical models through the use of results from experiments conducted on curved steel I-beams at the laboratories of the University of Pretoria, and thereafter, the development of a relatively large dataset through nonlinear analyses is described. The next step will involve the development of the proposed predictive models that are validated and presented in the present paper.

3. Experimental Investigation

Experimental studies were performed in order to validate the finite element models that were used to develop the datasets for the needs of this research work. Experiments were also used to determine the appropriate finite element type that would yield the most accurate results regarding the deflection and rotation of the I-beam. The experiments involved the investigation of a horizontally curved steel I-beam (IPE 100) where the two ends were fully fixed and a vertically upward load was applied at the midspan. The results measured were vertical deflection at the midspan, rotation at the midspan and quarter points, and strains at the midspan and the support.

The beam that was analysed was a 3.5 m long IPE100 beam. This beam was selected to minimise the applied forces in order to maintain a safe working environment. A simplified schematic outlining how the beam was supported and loaded can be seen in Figure 3.

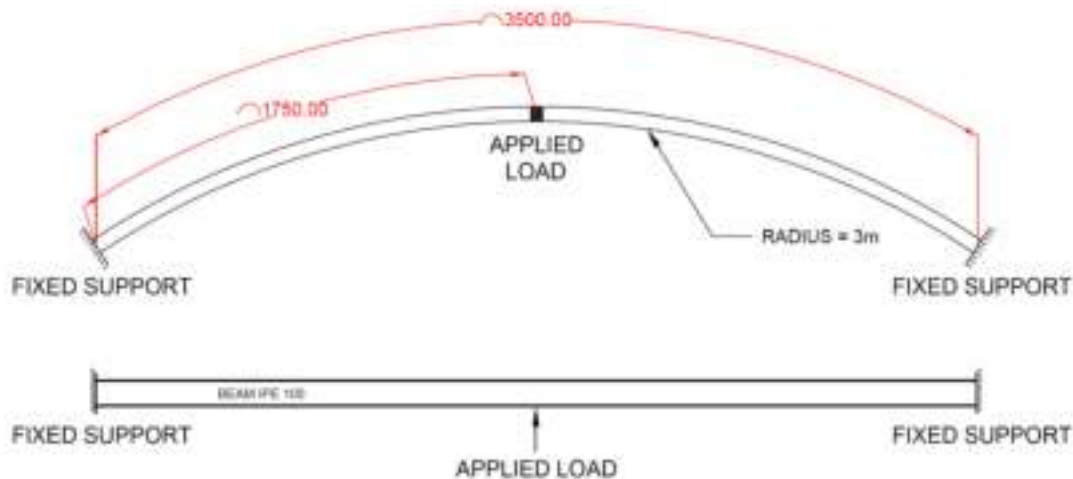


Figure 3. Basic schematic of the curved beam experimental setup (all dimensions are in millimetres).

To create a fully fixed support, the cross-section of the IPE 100 beam was welded onto two 12 mm thick plates. These 12 mm thick plates were then welded onto a stiff beam, which was then bolted onto the test floor using M24 Gr8.8 bolts. The entire system was assumed rigid and minimal deflections and rotations were expected at this point. Additional linear variable differential transformers (LVDTs) were placed at the supports. A schematic of the support conditions as well as a photograph can be seen in Figures 4 and 5, respectively.

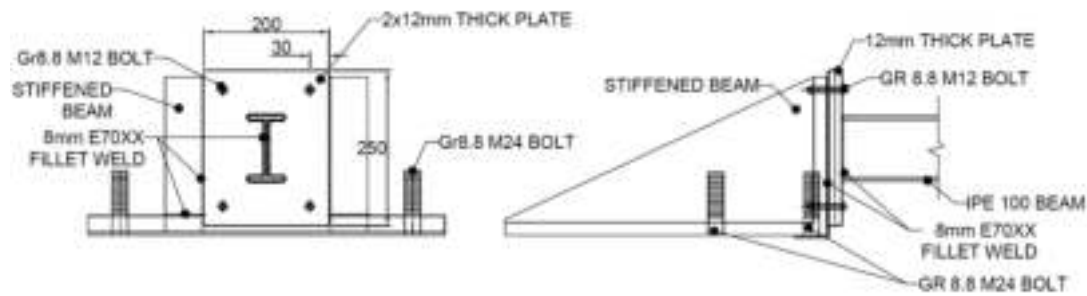


Figure 4. Schematic of the experiment with fixed support (all dimensions are in millimetres).



Figure 5. Experiment with fixed support.

A simple load application mechanism was used, which allowed a quicker and easier test set-up. The beam was loaded with a 40 mm steel bar that had a flat edge. This flat edge was loaded onto the beam using an overhead crane. The load cell then calculated the load experienced by the beam every 100 milliseconds, which was seen to provide sufficient data for analysis purposes. The loading equipment had a capacity of 50 tonnes, which was sufficient for the expected failure load of the beam. The steel bar also allowed the beam to rotate freely at the point of load application, which was required given the large expected rotations of this horizontally curved beam.

LVDTs were used to measure beam deflection only in the vertical direction. Two different LVDT types were used, one with a 250 mm range, which was placed at the supports and the other with a 1000 mm range, which was placed at the midspan. The LVDT at the midspan was placed 160 mm away from the point of load application so measurements could be taken at the same position as the strain gauges. The strain gauges were offset in order to not be influenced by the applied point load. LVDTs were used due to their high accuracy and good long-term stability as opposed to potentiometers, which have lower accuracy and precision but are typically more versatile. A schematic of the placement of LVDTs can be seen in Figure 6.

Inclinometers were used to measure the rotation along the beam as well as across the beam. Two dual-axis inclinometers were used. The inclinometers used made use of an electrolytic level which was capable of measuring inclination along two axes (pitch and roll). One inclinometer was placed 100 mm from the point of load application but on the opposite half of the beam from where the LVDT and strain gauges were placed, and the other was placed at the quarter point of the same half. A schematic outlining the position of the inclinometers can be seen in Figure 7.

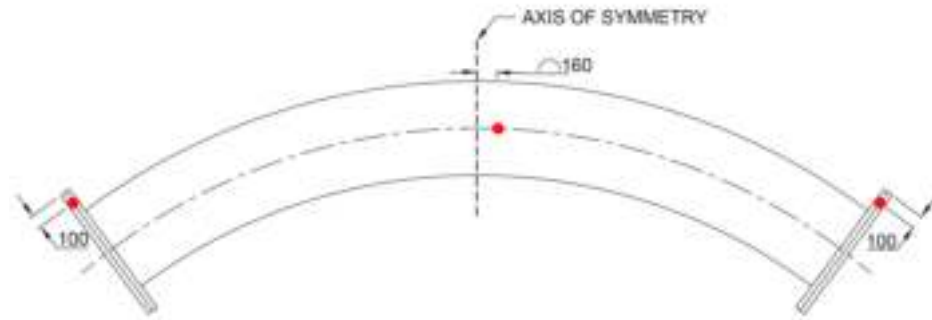


Figure 6. Schematic indicating points of displacement measurements (all dimensions are in millimetres).

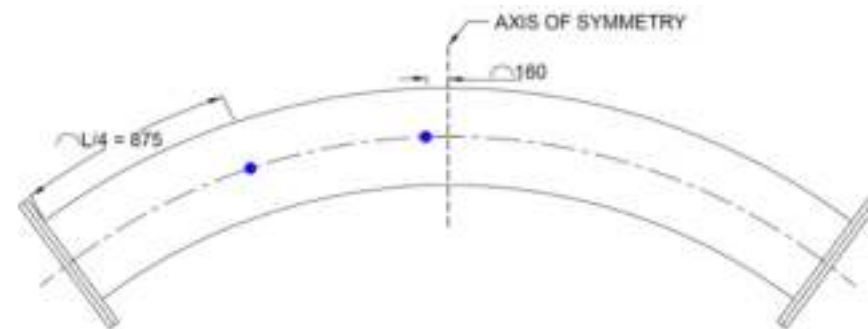


Figure 7. Schematic indicating points of rotation measurements (all dimensions are in millimetres).

The majority of the setup of the instrumentation involved accurately placing the strain gauges. Strain gauges are very sensitive. Therefore, the surface where a strain gauge is applied is required to be extremely smooth and free of all impurities prior to the placement of the strain gauge. The placement of these strain gauges was critical so that comparison between the strain gauge and the FE analysis could be as accurate as possible. This was particularly true for the strain gauges placed at 45° , which were required to accurately measure torsional strain within the beam. A total of 16 strain gauges were used to measure the strain at various points on one half of the beam. At the midspan, only longitudinal strain gauges were provided at the top and the bottom flange. These longitudinal flanges were placed at the far edge of the beam to read the maximum longitudinal strain at the midspan. The same setup was followed at the support, which led to another four longitudinally placed strain gauges being placed at this point. Two strain gauges were also placed transversely at the top and bottom flanges to investigate the shear flow in the beam. A further two vertical strain gauges were placed on either side of the web to measure the shear flow of the beam across the web. Four diagonal strain gauges were placed to measure the torsional stresses experienced at these points. A schematic outlining the position of the strain gauges can be seen in Figure 8.

All instrumentation was connected to an HBX Quantum logger. This logger allowed conversion from voltage to the appropriate units for the various instrumentation after calibration (e.g., millimetres for the LVDT, kilonewtons for the load cell, and micrometres/metres for the strain gauges). During the experiment, the load was applied to the beam in increments up to the final failure load past the point of yielding. During loading, a video was taken to record the experiment, which was used for verification when analysing the results. Photos were then also taken after the final load was applied and yielding had been experienced for further analysis (Figure 9).

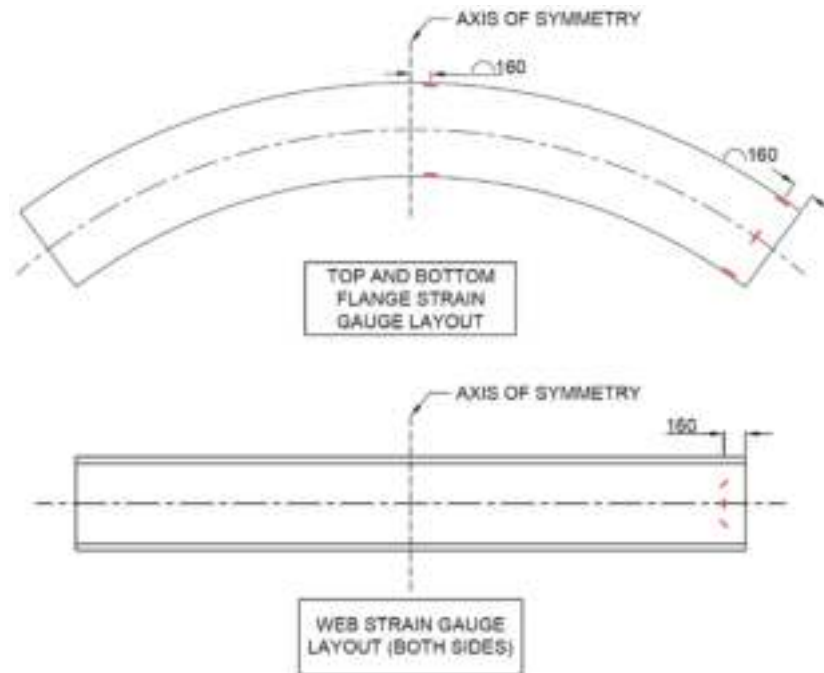


Figure 8. Schematic indicating points of strain measurements (all dimensions are in millimetres).



Figure 9. Deflected shape of the beam during the loading phase.

4. Analysis of Experimental Results

During testing, a total of five variables were measured. These variables were the load applied at the midspan, the displacement, the tilt/rotation, and the strain. The initial readings prior to the load application were taken as a reference point for all future readings. As previously mentioned, the load was applied incrementally in steps of 3 kN. This was carried out so it could be accurately determined at what load yielding occurred (through measuring permanent deformation). It was noted that after the 12 kN load was removed, permanent deformation occurred at the midpoint of the beam. The deflection for this load was approximately 55 mm.

The deflection measured at the supports was negligible, reaching up to 0.07 mm (upwards). This was seen as negligible and therefore was not factored into the calculations of the midspan deflection. Figure 10 shows the load–deflection curve of the beam. Results showed that the beam deflected linearly up to 55.05 mm at a load of 12.42 kN. The experiment was stopped right at the point of yielding; therefore, a plateau is not indicated in this graph, which would be a further indication that yielding had occurred.

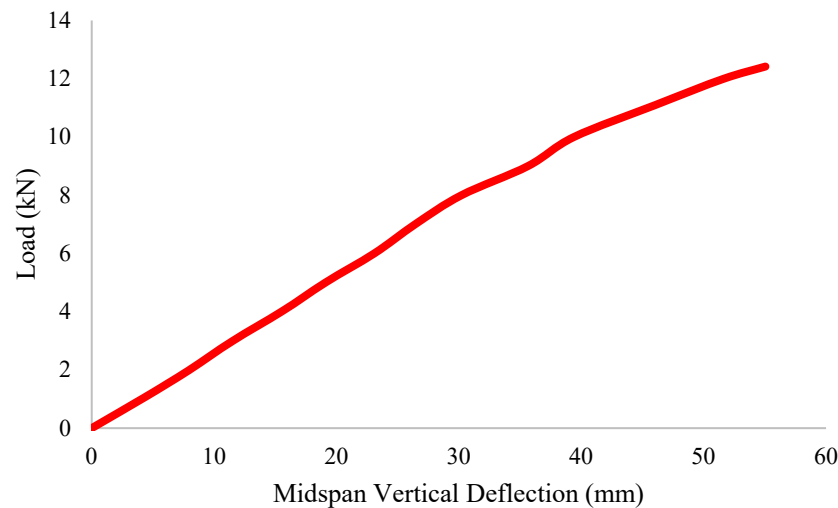


Figure 10. Experimental beam load–deflection curve (midspan).

When investigating the tilt results, errors were noted. The results showed that tilt did not consistently increase as load increased. At some points, a decrease in tilt was measured with an increase in load. This can be seen in Figure 11, where it can be seen that as the load increased from 2 kN to 4 kN, the tilt was measured to have decreased slightly. Practically, this is not possible. It is believed that this error was due to incorrect use of the measuring apparatus. The apparatus used to measure tilt made use of an electrolytic level. Electrolytic levels are typically poor when it comes to dynamic/cyclical loading, due to the conductive fluid found within the sealed glass. These results can therefore be used to obtain a rough estimate of how tilt increases as the load increases; however, these results cannot be used to determine the accuracy of FE models or analytical formulae. A line of best fit has been included in Figure 11 for further interpretation, which shows that tilt increased as load increased.

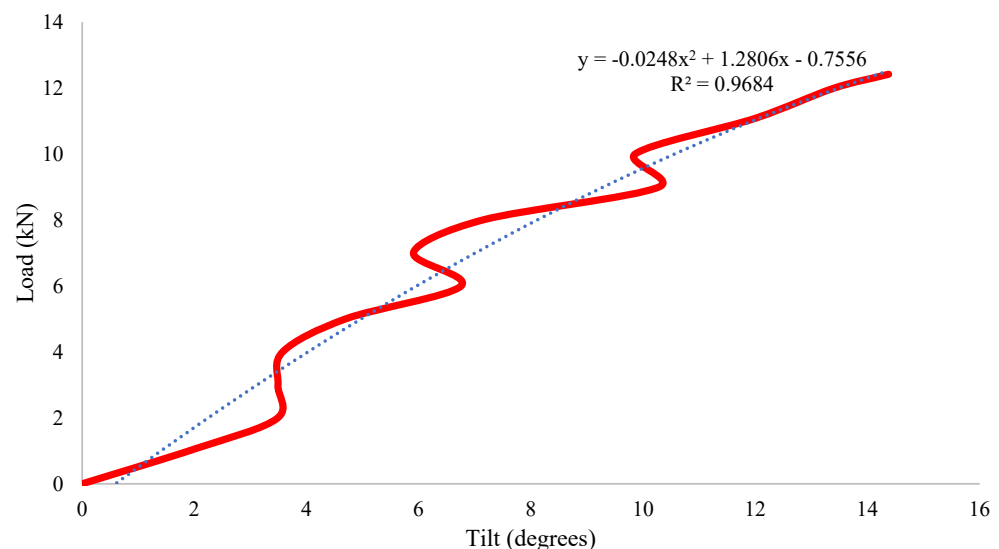


Figure 11. Experimental beam load–tilt diagram.

The strain results were quite interesting and showed the implication of the support conditions utilised. The support was intended to be fully fixed; however, practically, the support behaved in a way between that of a fully fixed connection and a torsionally pinned connection, as was seen with the strain gauge results. The strain gauge pattern was symmetrical with the cross section (both vertically and horizontally); however, the results were not seen to be perfectly symmetrical. The major reason for the discrepancies was most likely to have been due to inaccuracies during the setting out of the strain gauges. Material and geometric imperfections also played a role. The strain gauge results can be seen from Figures 12–14.

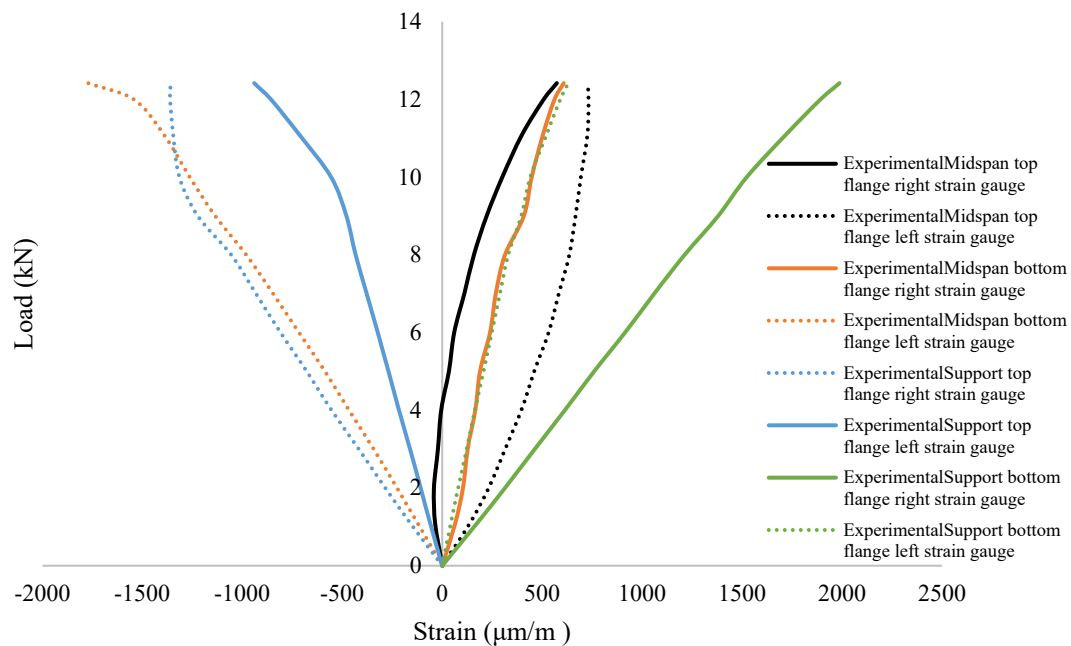


Figure 12. Graphical summary of experimental beam flange longitudinal strain gauge results.

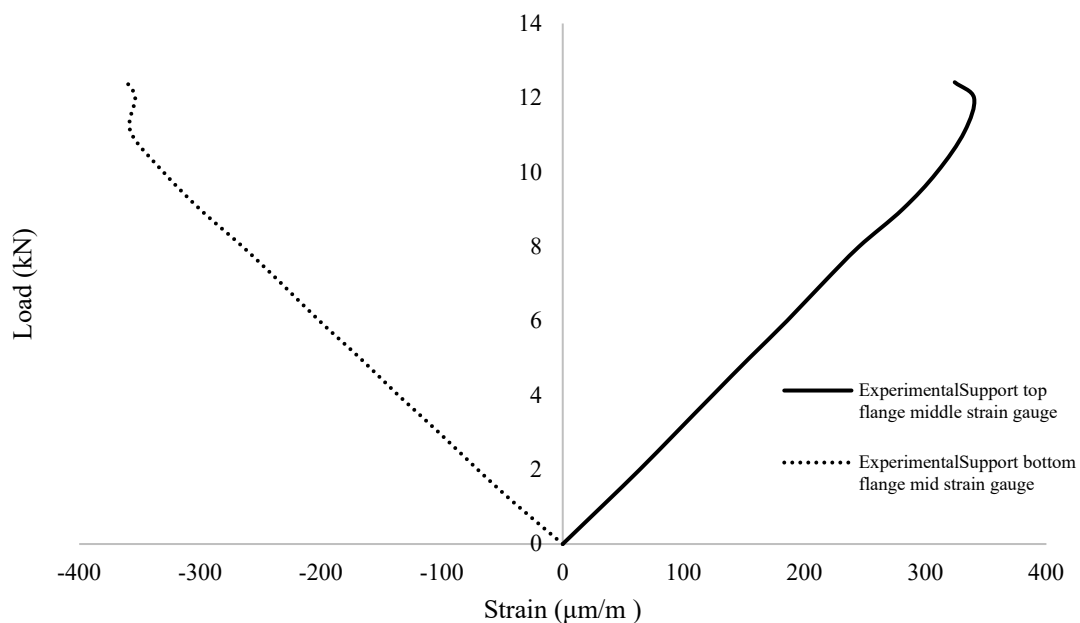


Figure 13. Graphical summary of experimental beam flange transverse strain gauge results.

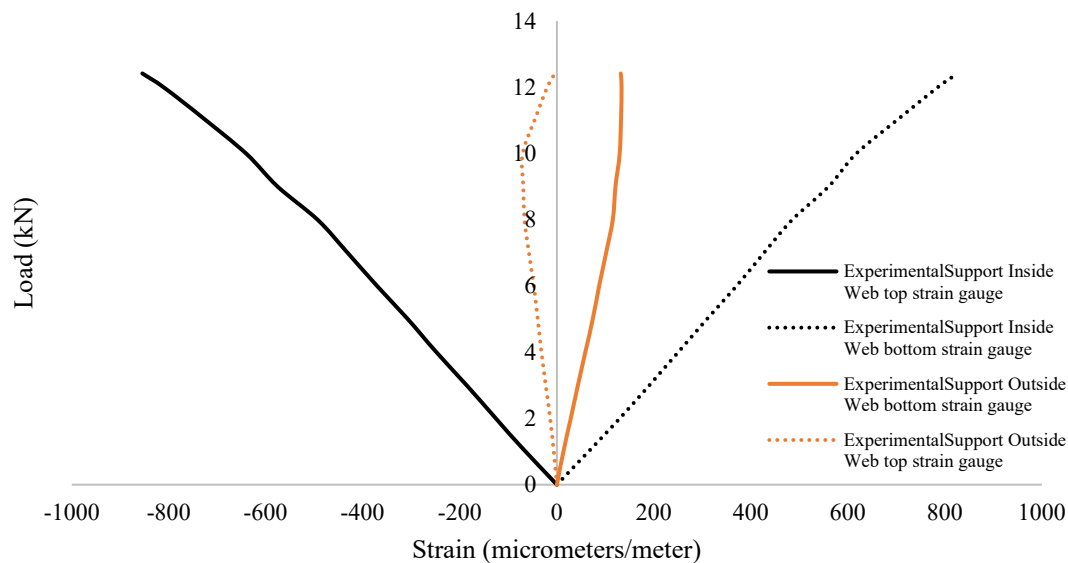


Figure 14. Graphical summary of experimental beam 45° web strain gauge results.

5. Comparison of Experimental Data with Finite Element Analysis Results

Various FE models were used and compared to determine which model most accurately represented the behaviour of the experimental beam. The focus was to determine the optimum FE model to be used for the development of the datasets. The datasets were used to train various ML algorithms for the development of the proposed predictive models related to the calculation of the ultimate strength of curved steel I-beams that are fixed at their ends and the respective deflection.

CivilFEM 2021 was used to develop the numerous FE meshes that were used to construct the dataset. This software was selected due to its interface, which allowed the creation of curved steel I-beam meshes in an automated manner. Solid and shell FEs were considered in this research work, where the mesh size was determined based on a previous research study where a mesh sensitivity analysis was conducted by the authors [21]. Linear and quadratic elements were also considered for each of the FEs used to reproduce the experimental results. It is important to note at this stage that [21] used experimental results found in the literature to calibrate their model that involved steel I-beams with a different support system compared with the fixed-end beams that are investigated herein. That pilot project showed that hexahedral isoparametric FEs were able to accurately reproduce experimental results [21].

Therefore, a total of four models were created for the needs of this research work, namely, a linear solid element, a quadratic solid element, a linear shell element, and a quadratic shell element. All material-related parameters were kept constant between the four models to maintain consistency. A typical nonlinear material model with anticipated bilinear behaviour was assumed. The hardening type selected was isotropic, indicating that yielding occurred when the effective or equivalent stress was equal to the specified yield stress. This selected yielding criterion was set to be that of von Mises. Furthermore, the material properties were assumed to be that of an S355 steel material (355 MPa yield stress, 210 GPa elastic modulus, a Poisson ratio of 0.3, and a density of 7850 kg/m³). The experimental deflection and rotation at the midspan were then compared with the midspan deflection and rotation calculated using the various FE models. Analytical formulae were also included for comparative purposes.

Figure 15 shows that the linear solid FE derived the worst prediction when estimating midspan deflection. The analytical method, Castigliano's second theorem, was also unable to accurately represent the behaviour of the beam and performed quite poorly compared with the more advanced FE models. The quadratic solid model most accurately calculated

the midspan deflection, with an average error of only 1.83 mm. A summary of the computed mean absolute error (MAE) can be seen in Table 1.

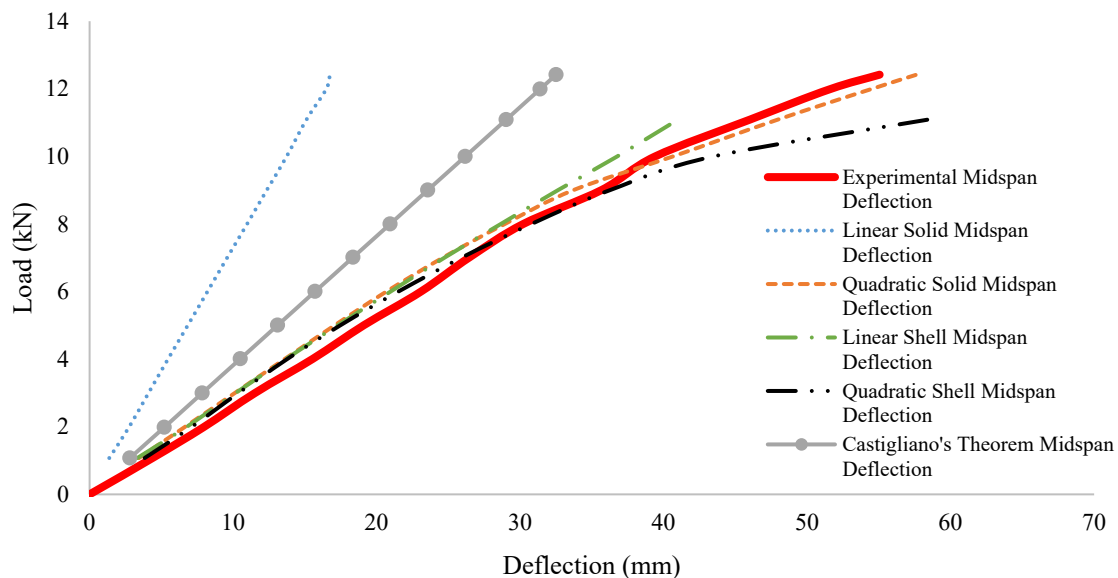


Figure 15. Load–deflection graph displaying experimental deflection results, FE results, and results obtained with the analytical formula.

Table 1. Mean absolute error of the FEM deflection estimate and Castigliano’s theorem compared with the experimental data.

Deflection Estimation Method	MAE
Linear solid finite element model	18.64 mm
Quadratic solid finite element model	1.83 mm
Linear shell finite element model	2.12 mm
Quadratic shell finite element model	2.33 mm
Castigliano’s theorem	9.91 mm

Furthermore, the resulting section rotations can be seen in Figure 16. Similar to the deflection results, experimental data and the results derived from FE analyses were compared with those of a commonly used analytical method, as discussed previously (Equation (2)). The experimental deflection was used as an input to achieve the highest accuracy possible when implementing the analytical method. Due to the fact that solid FEs do not contain rotational degrees of freedom, the rotation was calculated from first principles through making use of the vertical deflection at opposite ends of the bottom flange and using trigonometry to calculate the rotation. This method only holds given that the cross-section does not warp significantly. From the graphic, it was noted that the linear solid FEs provided the least accuracy, underestimating the midspan tilt quite significantly.

The M/R performed slightly better compared with the linear solid Fes; however, these results were still exceedingly conservative compared with the more advanced FE models. The more advanced FE models correlated well with the line of best fit of the experimental data. However, it seems the quadratic shell model was more accurate up until approximately 10 kN. From 10 kN to failure, the linear shell and the quadratic solid models both displayed lower errors compared with the experimental data. A summary of error metrics (MAE) can be seen in Table 2.

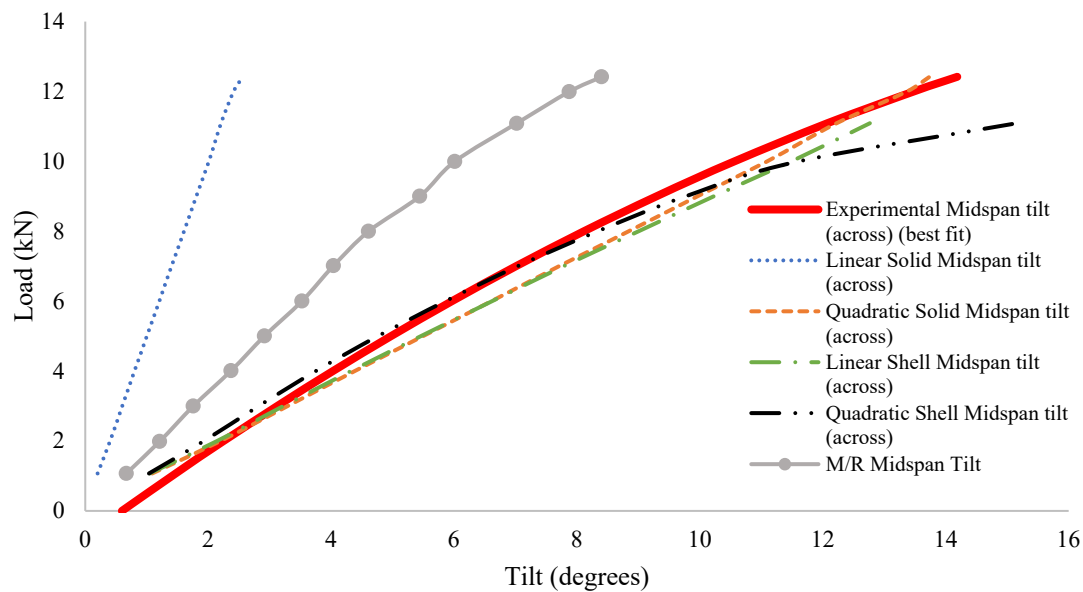


Figure 16. Load–rotation graph displaying experimental, FE, and the M/R analytical formula results.

Table 2. Mean absolute error of FEM rotation estimate and the analytical formula for experimental data.

Rotation Estimation Method	MAE
Linear solid finite element model	4.98°
Quadratic solid finite element model	1.31°
Linear shell finite element model	1.52°
Quadratic shell finite element model	1.33°
M/R method	2.31°

It is therefore concluded that FE models with the numerical modelling technique as outlined in this section are capable of accurately representing the behaviour of horizontally curved steel I-beams. These models have also been shown to outperform quite significantly the existing analytical formulae for estimating both deflection and rotation. It was noted that of all the FE models, the quadratic solid FE performed best in both estimating deflection and rotation. Therefore, the numerical modelling technique outlined in this section was utilised in developing a large dataset as described in the subsequent sections.

6. Numerical Campaign

A relatively large dataset consisting of 864 models was created through the use of CivilFEM. A total of six sections were considered, namely, IPE100, IPE200, IPE300, IPE400, IPE500, and IPE600. This was done to encompass the entire IPE section list according to European standards. The variables considered were the moment of inertia around the x axis (I_{xx}), the moment of inertia around the y axis (I_{yy}), polar moment of inertia (J), section height (h), section width (b), section area (A), beam curved length (L), beam radius of curvature (R), material yield strength (f_y), material Young's modulus (E), material Poisson ratio (ν), and material shear modulus (G). For each section considered, three yield strengths were assumed (235 MPa, 275 MPa, and 355 MPa), to account for the various structural steel strengths commercially available in South Africa. Three Young's moduli were considered, namely, 190 GPa, 200 GPa, and 210 GPa, to account for variances in material stiffness. The Poisson ratio was fixed at 0.28 and the equation of G was calculated based on E and ν . Four lengths and four radii were considered per section in order to cover the broad range of geometries that the proposed predictive models will be applicable for. Given that R/L ratios

have been found to control the behaviour of curved beams, four specific R/L ratios were considered on all beams, namely, $R/L = 1, 2, 4$, and 8 . All the beams were fully fixed on both ends and a load was applied normally downwards at the midspan.

To maintain consistency, the models were automatically generated using a Python script within CivilFEM. This ensured all material and geometrical properties were maintained and only a single variable was modified, allowing automatic mesh generation. A Python script was developed and optimised to ensure consistency and minimise running time through decreasing the number of variables and the overall performance of the code. The models all made use of quadratic solid elements with a mesh size of up to 50 mm. A typical nonlinear material model with predicted bilinear behaviour was assumed. The hardening type selected was isotropic, which indicates that yielding occurred when the effective or equivalent stress was equal to the specified yield stress. This selected yielding criterion was set to be that of von Mises.

Table 3 indicates different statistical parameters, including skewness. A value of 0 indicates that the dataset contained no skewness and had a perfectly normal distribution. According to Aminu and Shariff [22], a range from -3 to 3 can be considered as a cutoff, and based on this, the results show that the data were not significantly skewed. Kurtosis is another parameter that assists further in determining whether a dataset is normally distributed or not. This is performed by determining whether a dataset is “heavy-tailed” or “light-tailed”. “Heavy-tailed” implies that a dataset contains numerous data points in outlier positions, whereas a “light-tailed” dataset contains minimal to no outliers. Once again, there is no accepted convention on what is deemed “heavy-tailed”; however, Aminu and Shariff [22] state that a range from $+10$ to -10 is deemed light-tailed and contains minimal outliers. In this context, it was observed that the dataset created for the needs of this research work contained minimal outliers with a good distribution.

Table 3. Horizontally curved beam dataset descriptive statistics.

	Mean	Median	STD	Min	Max	Skewness	Kurtosis
I_{xx} (mm ⁴)	291,198,612	231,300,000	326,514,119	1,710,000	920,800,000	0.977	−0.455
I_{yy} (mm ⁴)	12,720,819	13,180,000	1,1960,328	159,000	33,870,000	0.625	−0.960
J (mm ⁴)	692,957	841,000	553,277	12,100	1,540,000	0.142	−1.343
h (mm)	349.80	400	172.05	100	600	−0.011	−1.284
b (mm)	150.59	180	58.02	55	220	−0.472	−1.181
A (mm ²)	7490	8450	5048	1030	15,600	0.285	−1.223
L (m)	7.48	4	6.44	1	24	1.346	0.953
R (m)	27.99	16	35.26	1	192	2.653	8.242
F_y (MPa)	289.37	275	50.23	235	355	0.343	−1.543
E (GPa)	199.93	200	8.14	190	210	0.012	−1.491

The correlation matrix for the displacement dataset can be found in Figure 17a and for the failure load dataset in Figure 17b. It was clear from the correlation matrix results that all cross-sectional properties had a strong positive correlation with the midspan deflection and the failure load. It was also clear that, of all the properties, the curved length had the largest correlation with the midspan vertical deflection. This implies the potential to translate to positive results in the sensitivity analysis; this is discussed in Section 7.

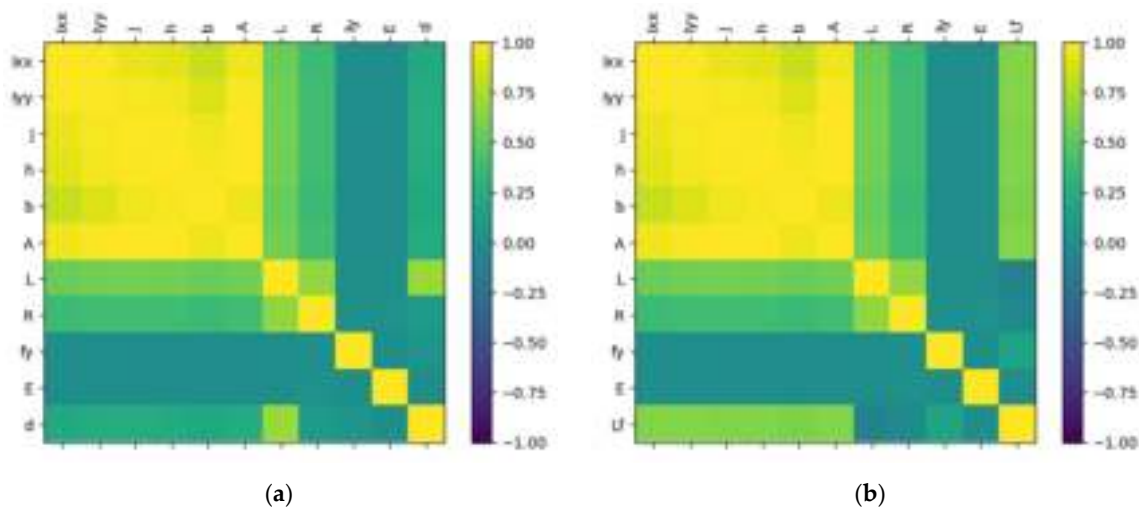


Figure 17. Correlation matrix of (a) midspan deflection dataset and (b) failure load dataset.

7. Machine Learning Training and Testing

This section outlines how numerous ML algorithms were applied in order to create formulae that can outperform current analytical methods used to estimate the deflection of horizontally curved steel I-beams. The ML algorithms considered were linear regression (LR), polynomial regression with hyperparameter tuning (POLYREG-HYT), deep artificial neural networks with MPI, Horovod, and hyperparameter tuning (DANN-MPIH-HYT), and extreme gradient boosting with hyperparameter tuning and cross-validation (XGBoost-HYT-CV) [14]. It must be noted here that for the ML analyses performed, 85% of the dataset was used to train the ML algorithm and 15% was used for testing. The performance of ML algorithms varies; therefore, to quantify accuracy, numerous error metrics were considered. The error metrics considered in this research work were the mean absolute percentage error (MAPE), the mean absolute mean percentage error (MAMPE), the mean absolute error (MAE), and the root-mean-square error (RMSE). The Pearson correlation coefficient was also considered (R) for determining the similarity between the PV (predicted value) and the DV (dependent variable). An in-depth discussion of the error metrics can be found in [23].

7.1. Proposed Predictive Models for the Case of Deflection

This section outlines the midspan deflection predictions provided by the various ML algorithms. The numerically obtained error metrics were analysed to determine the performance of the various proposed predictive models. The correlation for the train and test datasets can be seen in Figures 18–21.

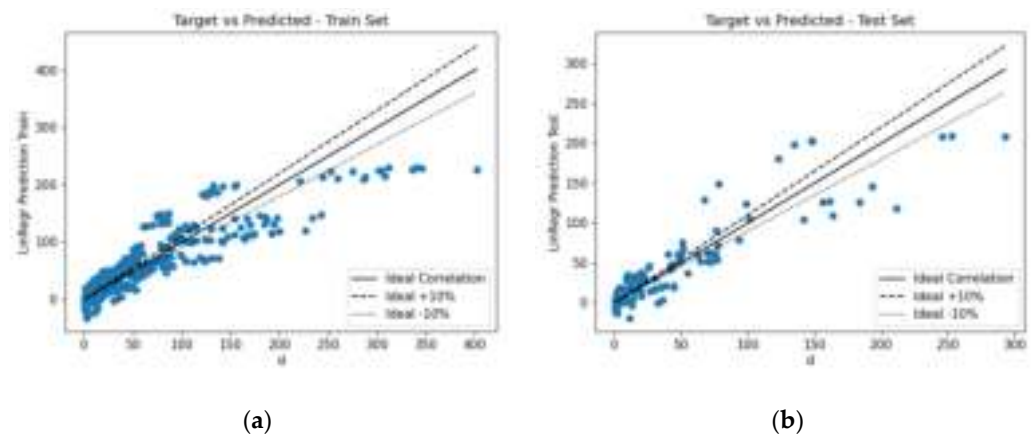


Figure 18. Correlation between the deflection determined using FEM and the deflection estimated using LR (a) on the training dataset and (b) on the testing dataset.

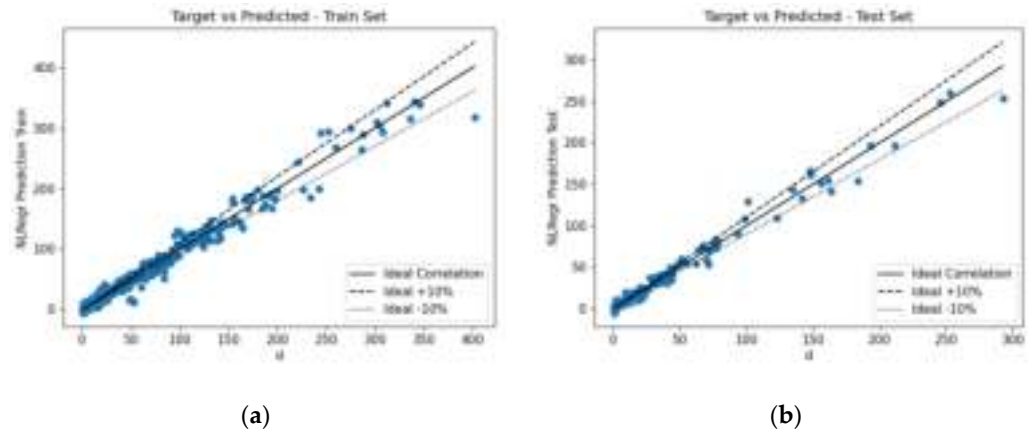


Figure 19. Correlation between the deflection determined using FEM and the deflection estimated using POLYREG-HYT (a) on the training dataset and (b) on the testing dataset.

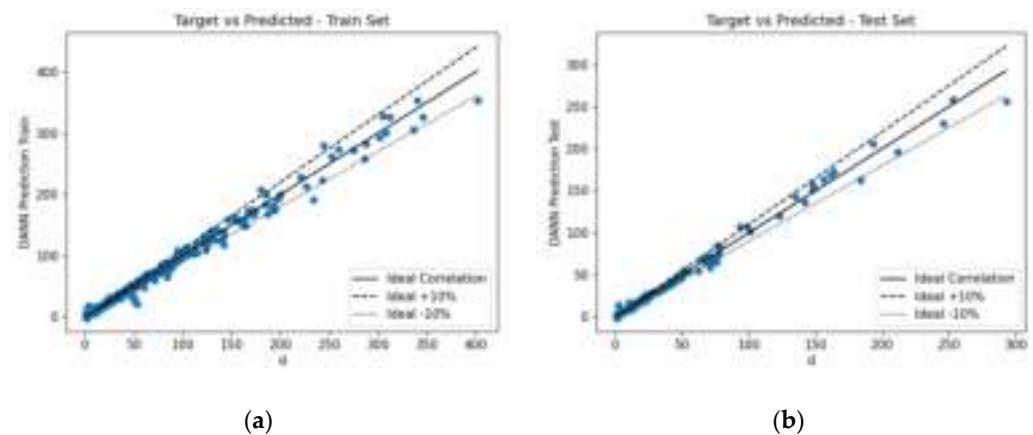


Figure 20. Correlation between the deflection determined using FEM and the deflection estimated using the DANN-MPIH-HYT algorithm (a) on the training dataset (b) on the testing.

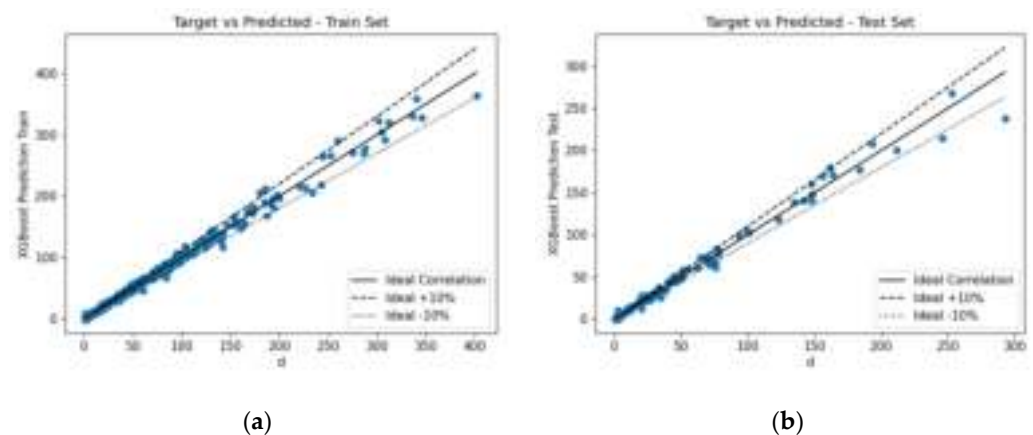


Figure 21. Correlation between the deflection determined using FEM and the deflection estimated using the XGBoost-HYT-CV algorithm (a) on the training dataset (b) on the testing dataset.

LR and POLYREG-HYT are the only ML algorithms considered in this study that are capable of providing closed-form formulae. Equation (3) shows the formula generated with the LR algorithm, which can be used to estimate the vertical midspan deflection, whereas Equation (4) provides the proposed predictive model derived using POLYREG-HYT. When implementing the formulae, all cross-sectional properties are in $\text{mm}/\text{mm}^2/\text{mm}^3/\text{mm}^4$, and the span is given in m. The yield strength is provided in MPa and the Young's modulus in GPa.

$$d = -2.9147 \times 10^{-8}E - 08 \times I_{xx} + 6.085 \times 10^{-6} \times I_{yy} + 7.153 \times J + 8.6563 \times 10^{-1} \times h - 6.7409 \times 10^{-1} \times b - 4.496 \times 10^{-2} \times A + 1.1562 \times 10^1 \times L - 1.1261 \times R + 5.1585 \times 10^{-2} \times f_y - 6.87948 \times 10^{-2}E \quad (3)$$

$$\begin{aligned} d = & -1.10850 \times 10^{-2} \times h \times L + 5.54753 \times L^2 - 2.59341 \times 10^{-2} \times b \times L^2 + \\ & 8.71342 \times 10^{-1} \times L + 4.36940 \times 10^{-1} \times R - 3.72902 \times 10^{-4} \times b^2 + 7.64167 \times \\ & 10^{-5} \times h \times R^2 + 4.14113 \times 10^{-7} \times h \times A \times L + 1.20492 \times 10^{-1} \times R^2 - \\ & 2.10080 \times 10^{-9} \times J \times L \times f_y + 2.36217 \times 10^{-8} \times A^2 \times R + 8.06401 \times 10^{-3} \times b \times L \times \\ & R - 3.01987 \times 10^{-5} \times A \times L \times R - 1.61891 \times L \times R - 2.78160 \times 10^{-2} \times L^3 - \\ & 3.98564 \times 10^{-4} \times R^3 - 4.35591 \times 10^{-9} \times b \times A^2 - 6.98688 \times 10^{-4} \times b \times R^2 + \\ & 4.76077 \times 10^{-3} \times L \times R^2 + 9.28207 \times 10^{-12} \times I_{yy} \times b \times f_y - 2.05690 \times 10^{-7} \times A \times \\ & R \times f_y + 2.68695 \times 10^{-7} \times h^3 + 1.26135 \times 10^{-10} \times I_{yy} \times R^2 + 2.50786 \times 10^{-5} \times R^2 \times \\ & f_y + 8.68248 \times 10^{-4} \times L^2 \times f_y + 6.78638 \times 10^{-11} \times I_y \times R \times f_y - 1.03722 \times 10^{-6} \times \\ & h \times b \times f_y + 1.76647 \times 10^{-5} \times h \times b \times R - 3.55305 \times 10^{-7} \times A \times E^2 - \\ & 9.13425 \times 10^{-7} \times h \times f_y^2 + 1.87704 \times 10^{-7} \times h \times A \times E - 8.00661 \times 10^{-7} \times h \times A \times \\ & R + 4.02981 \times 10^{-6} \times h \times f_y \times E - 2.62557 \times 10^{-4} \times L \times R \times f_y - 1.01188 \times 10^{-2} \times \\ & L^2 + 1.18944 \times 10^{-6} \times J \times L^2 - 7.59868 \times 10^{-10} \times I_{yy} \times L \times E \end{aligned} \quad (4)$$

Visually, it was difficult to discern which ML algorithm performed the best. Therefore, as previously mentioned, numerous error metrics were used to measure the performance of the various algorithms. Table 4 summarises the results. As can be seen, XGBoost-HYT-CV was the most accurate proposed model when looking at the training dataset, whereas when comparing the error metrics obtained from the testing dataset, that was not the case. The DANN-MPIH-HYT proposed predictive model was found to be slightly more accurate. It should also be noted that, even though DANN-MPIH-HYT was the most accurate in the testing phase, this algorithm did require the longest computation time (27 times slower).

Table 4. Comparison of the ML algorithms trained on the deflection dataset.

Model	Dataset	R	MAPE	MAMPE	MAE	RMSE	Computation Time (s)
LR	Train	0.8934	225.93%	45.08%	17.6614	25.6709	0.0013
POLYREG-HYT	Train	0.9874	69.91%	13.90%	5.4447	9.0447	10.91
DANN-MPIH-HYT	Train	0.9950	39.64%	7.59%	2.9731	5.7272	2019.76
XGBoost-HYT-CV	Train	0.9968	22.97%	5.92%	2.3199	4.5566	75.27
LR	Test	0.9064	168.22%	39.74%	17.1810	24.4716	0.00
POLYREG-HYT	Test	0.9905	70.38%	12.49%	5.3997	7.9865	0.00
DANN-MPIH-HYT	Test	0.9950	35.82%	8.04%	3.4776	5.8759	0.0004
XGBoost-HYT-CV	Test	0.9922	23.89%	8.44%	3.6504	7.2498	0.0033

Ultimately, it was apparent that accurate formulae can be created using ML algorithms. A later section of this paper compares these results with out-of-sample data, further validating the proposed predictive models' ability to capture unknown results.

Sensitivity analyses were conducted on each of the independent variables to determine which variable affected the dependent variable the most. This will assist future studies in determining which variables to exclude in order to increase efficiency in the proposed predictive models. This analysis also provides engineers with a deeper understanding of

the behaviour of horizontally curved steel I-beams. Figure 22 summarises the findings of the sensitivity analysis. As can be seen, the three most influential variables on midspan vertical deflection were the curved length (L), the moment of inertia around the minor axis (I_{yy}), and the radius of curvature (R). L was by far the most influential variable, according to the obtained results. The cross-section area, cross-section width, and cross-section height had practically no impact on the midspan deflection, according to the sensitivity analysis performed for the needs of this research work related to the beam deflection.

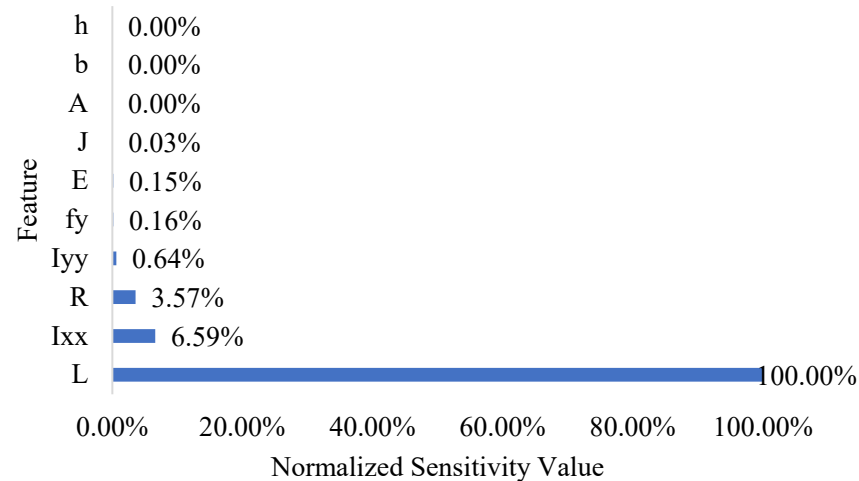


Figure 22. Summary of sensitivity analysis findings (XGBoost-HYT-CV).

It is important to show here that the predictive models did not overfit and that the solution obtained through the ML analysis led to an objective predictive model that did not overfit. In addition to the validation presented in Section 8, Figure 23 shows the tuned cross-validation histories for the DANN-MPIH-HYT and the XGBoost-HYT-CV ML algorithms. It was evident that the algorithms were able to derive models that were optimised through the training and testing procedure.

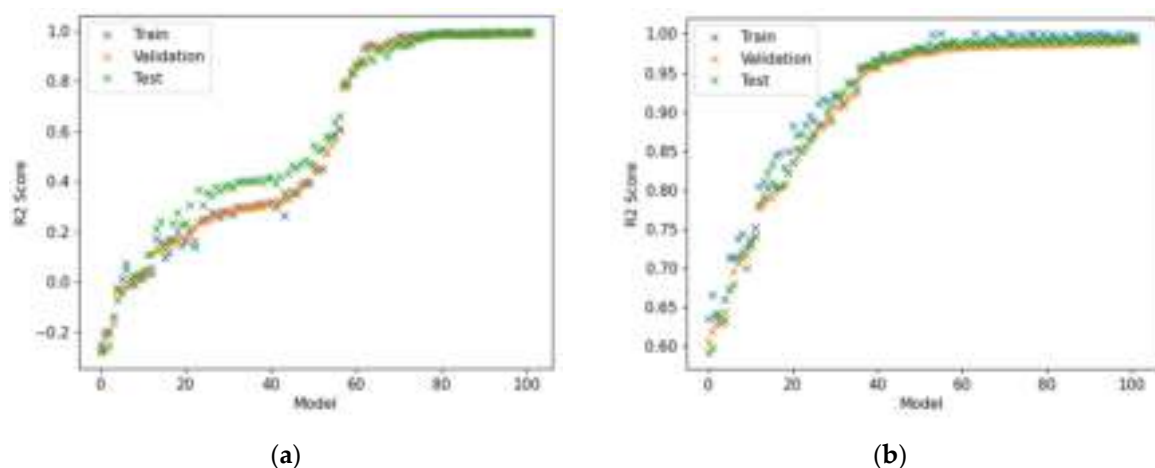


Figure 23. Case of deflection. Tuned cross-validation history for the case of (a) DANN-MPIH-HYT and (b) XGBoost-HYT-CV.

7.2. Proposed Predictive Models for the Case of Failure Load

This section outlines the failure load predictions provided by the various ML algorithms. An analysis was conducted to determine the performance of the various proposed predictive models. The correlation for the train and test datasets can be seen in Figures 24–27.

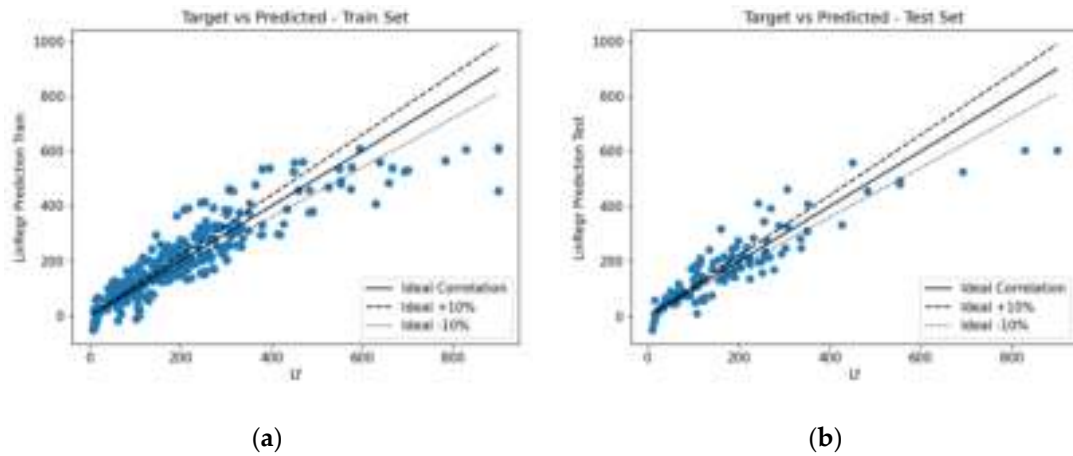


Figure 24. Correlation between the failure load determined using FEM and the failure load estimated using LR (a) on the training dataset (b) on the testing dataset.

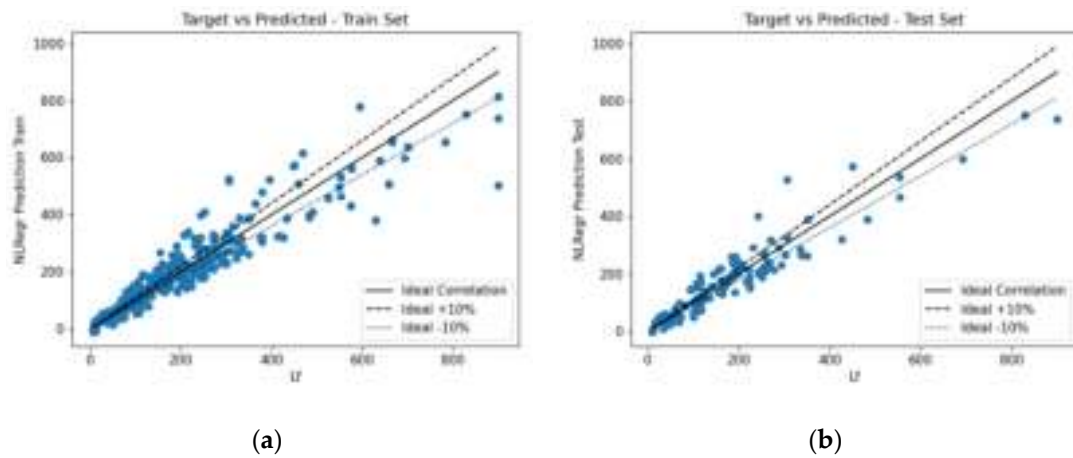


Figure 25. Correlation between the failure load determined using FEM and the failure load estimated using POLYREG-HYT (a) on the training dataset (b) on the testing dataset.

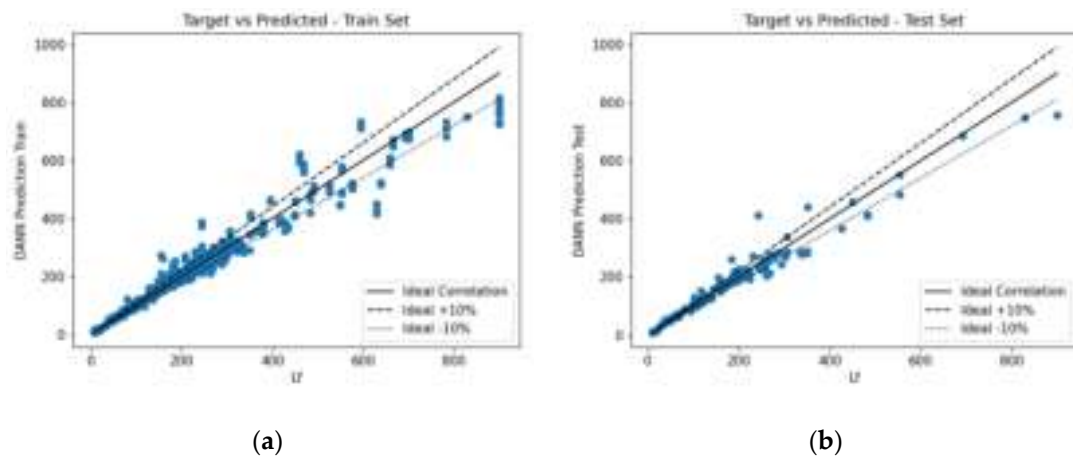


Figure 26. Correlation between the failure load determined using FEM and the failure load estimated using the DANN-MPIH-HYT algorithm (a) on the training dataset (b) on the testing dataset.

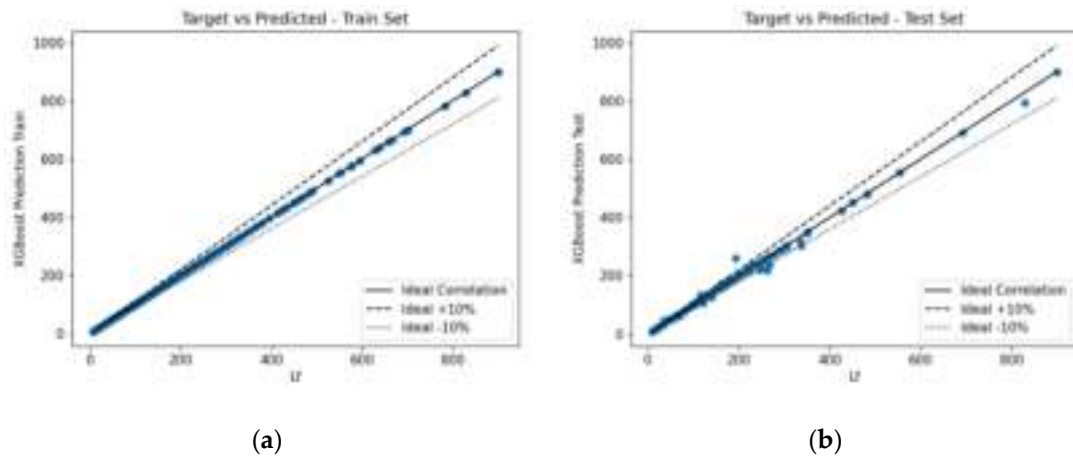


Figure 27. Correlation between the failure load determined using FEM and the failure load estimated using the XGBoost-HYT-CV algorithm (a) on the training dataset (b) on the testing dataset.

As stated previously, it was necessary to use error metrics to determine which ML algorithm performed the best. Therefore, a summary of error metrics can be seen in Table 5. As can be seen, the proposed XGBoost-HYT-CV model was the most accurate when looking at the training and testing datasets. This indicates the fact that no ML algorithm was capable of providing accurate results for all datasets, and different algorithms were used in allocating the best fit to a specific dataset.

Table 5. Comparison of the ML algorithms trained on the failure load dataset.

Model	Dataset	R	MAPE	MAMPE	MAE	RMSE	Computation Time (s)
LR	Train	0.8939	53.64%	27.45%	54.597	76.468	0.001
POLYREG-HYT	Train	0.9369	28.10%	19.26%	38.308	59.633	11.09
DANN-MPIH-HYT	Train	0.9811	10.29%	8.90%	17.693	34.055	2014.47
XGBoost-HYT-CV	Train	1.0000	0.00%	0.00%	0.0025	0.0041	75.07
LR	Test	0.8909	46.06%	27.56%	49.658	67.978	0.00
POLYREG-HYT	Test	0.9412	29.15%	20.86%	37.593	50.776	0.00
DANN-MPIH-HYT	Test	0.9791	10.59%	10.12%	18.241	32.1723	0.0008
XGBoost-HYT-CV	Test	0.9978	3.79%	2.61%	4.694	10.225	0.003

For the cases of the closed-form solutions, Equation (5) shows the formula generated with LR, and Equation (6) provides the proposed predictive formula generated through the POLYREG-HYT ML algorithm:

$$L_f = 2.0577 \times 10^{-6} \times I_{xx} - 1.5497 \times 10^{-4} \times I_{yy} + 7.7002 \times 10^{-4} \times J + 1.9789 \times h - 8.2286 \times b + 2.1078 \times 10^{-1} \times A - 2.0335 \times 10^1 \times L + 2.5421 \times 10^{-1} \times R + 5.8542 \times 10^{-1} \times f_y - 9.3538 \times 10^{-2} \times E \quad (5)$$

$$\begin{aligned}
L_f = & 1.2526 \times 10^{-3} \times L \times R^2 - 2.153 \times 10^{-4} \times I_{yy} - 1.455 \times 10^{-3} \times b \times A \\
& + 1.8898 \times 10^{-5} \times h \times L \times E - 1.2804 \times 10^{-9} \times I_{xx} \times h \times L \\
& - 7.8792 \times 10^{-5} \times h \times R^2 + 1.1314 \times 10^{-11} \times I_{xx} \times h \times R \\
& + 6.9077 \times 10^{-3} \times L \times f_y - 1.9573 \times 10^{-3} \times R \times f_y \\
& + 1.6174 \times 10^{-8} \times I_{yy} \times L^2 + 2.9069 \times 10^{-11} \times I_{yy} \times A \times f_y \\
& + 2.2861 \times 10^{-9} \times I_{yy} \times A \times L - 2.2406 \times 10^{-9} \times I_{xx} \times b \times L \\
& + 3.1524 \times 10^{-8} \times J \times h \times f_y + 3.8102 \times 10^{-4} \times h \times b^2 \\
& + 8.3911 \times 10^{-5} \times b \times R \times E - 2.3274 \times 10^{-6} \times J \times b \\
& - 1.1303 \times 10^{-2} \times A \times L - 6.1153 \times 10^{-10} \times I_{yy} \times R \times E \\
& - 4.5723 \times 10^{-4} \times b \times L \times f_y - 1.7705 \times 10^{-11} \times J^2 \times f_y \\
& + 2.7631 \times 10^{-3} \times f_y^2 - 6.0399 \times 10^{-6} \times f_y^3 + 7.0583 \times 10^{-5} \times J \times L \\
& + 1.36958 \times 10^{-10} \times I_{xx} \times A - 2.8373 \times 10^{-9} \times J \times R \times E
\end{aligned} \tag{6}$$

Sensitivity analyses were also conducted on the failure load dataset. A graphical summary of the findings can be seen in Figure 28. As can be seen, the three most influential variables on the midspan failure load were the moment of inertia around the major axis (I_{xx}), the curved length (L), and the radius of curvature (R). I_{xx} was by far the most influential variable, due to the fact that as the stiffness of the beam in bending increased, a larger load was required to reach failure. The cross-sectional area (A), flange width (b), cross-sectional height (h), polar moment of inertia (J), and moment of inertia around the minor axis (I_{yy}) had practically no impact on the failure load according to the XGBoost-HYT-CV ML algorithm. These findings can be used to create an improved, simpler dataset in the future.

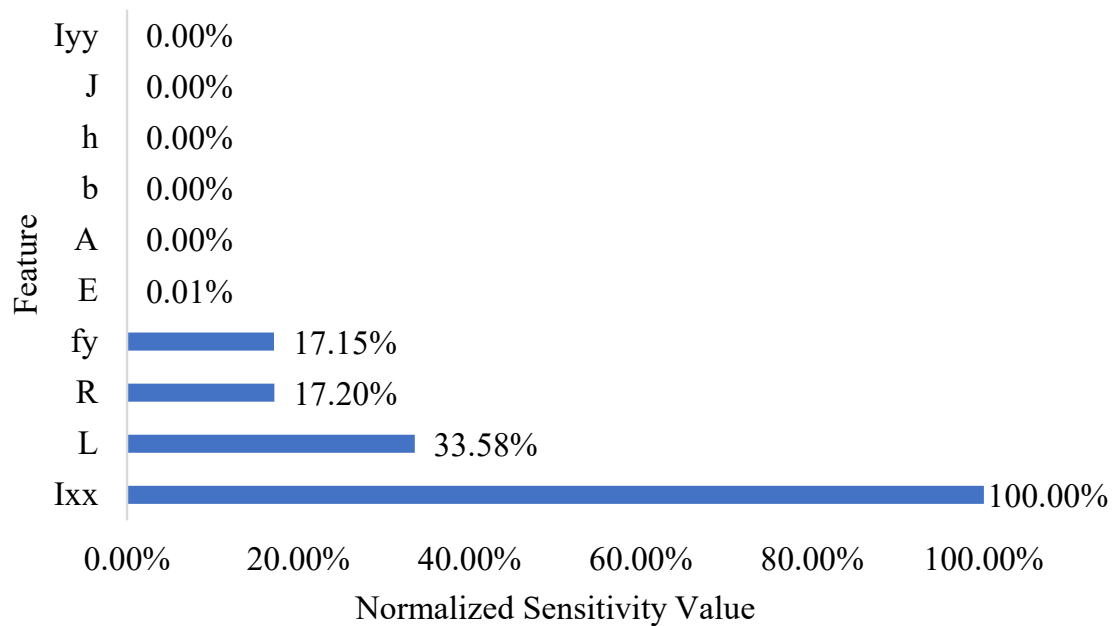


Figure 28. Results of sensitivity analysis (XGBoost-HYT-CV).

Before moving to the validation section presented next, Figure 29 shows the tuned cross-validation history for the case of the DANN-MPIH-HYT and XGBoost-HYT-CV algorithms resulting from the analysis. Once more, it is easy to observe the numerical response of the two histories that converge to $R^2 = 1.0$, which represents maximum data correlation.

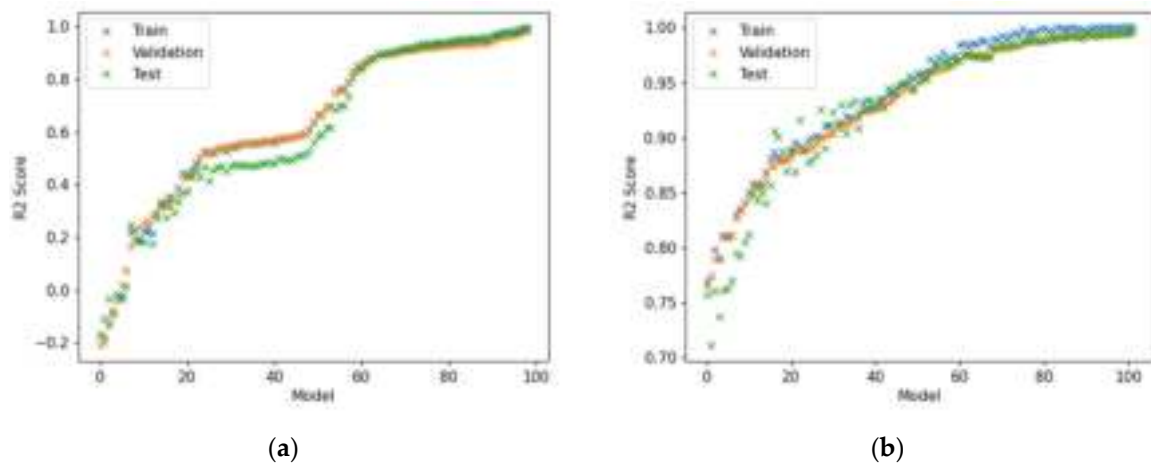


Figure 29. Case of ultimate load. Tuned cross-validation history for the case of (a) DANN-MPIH-HYT and (b) XGBoost-HYT-CV.

8. Validation

This section outlines the process of validation that was performed through the use of out-of-sample FE models. The validation aimed to evaluate the proposed predictive models' abilities to capture data that were not used during training or testing. It must be noted here that the ML algorithms were trained on IPE sections (IPE 100, IPE 200, IPE 300, IPE 400, IPE 500, and IPE 600) of various geometries. For validation purposes, it was decided to use sections as per the *South African Steel Construction Handbook* (SASCH). The sections considered were $203 \times 133 \times 25$, $305 \times 165 \times 40$, and $457 \times 191 \times 67$. Therefore, the proposed predictive models had never been exposed to this type of sectional geometry. Additionally, out-of-sample R/L ratios of 2.5 and 5 were considered and the overall span of the section varied depending on the section depth. Out-of-sample yield strengths of 285 MPa and 325 MPa were considered and the Young's modulus values considered were 195 GPa and 205 GPa. This led to a total of 48 out-of-sample beams being created for validation purposes. The descriptive statistics relating to the new validation dataset can be seen in Table 6.

Table 6. Validation models descriptive statistics.

	Mean	Median	Std	Min	Max	Skewness	Kurtosis
$I_{xx} \text{ (mm}^4\text{)}$	134,333,333	85,500,000	116,928,298	23,500,000	294,000,000	0.58	−1.53
$I_{yy} \text{ (mm}^4\text{)}$	8,416,667	7,660,000	4,738,358	3,090,000	14,500,000	0.25	−1.53
$J \text{ (mm}^4\text{)}$	194,667	149,000	134,794	59,000	376,000	0.49	−1.53
$h \text{ (mm)}$	320.20	303.8	103.97	203.2	453.60	0.24	−1.53
$b \text{ (mm)}$	162.8	165.1	23.37	133.4	189.90	−0.15	−1.53
$A \text{ (mm}^2\text{)}$	5643.33	516	2225.95	3220	8550	0.33	−1.53
$L \text{ (m)}$	9.17	8.5	5	3	18	0.56	−0.75
$R \text{ (m)}$	34.38	27.5	22.9	7.5	90	1.11	0.69
$f_y \text{ (MPa)}$	305	305	20.21	285	325	0.00	−2.09
$E \text{ (GPa)}$	200	200	5.05	195	205	0.00	−2.09

8.1. Validation of Deflection ML Results

The proposed predictive models that were developed to estimate the deflection of the beams were used to predict the deflections of the out-of-sample data. This section focuses only on the two algorithms, namely, DANN-MPIH-HYT and XGBoost-HYT-CV, that were

found to outperform the rest of the ML-generated predictive models. Correlation plots of the two ML models can be seen in Figures 30 and 31. These results were compared with the current analytical method, Castigliano's second theorem. The graph showcasing the correlation of Castigliano's second theorem can be seen in Figure 32.

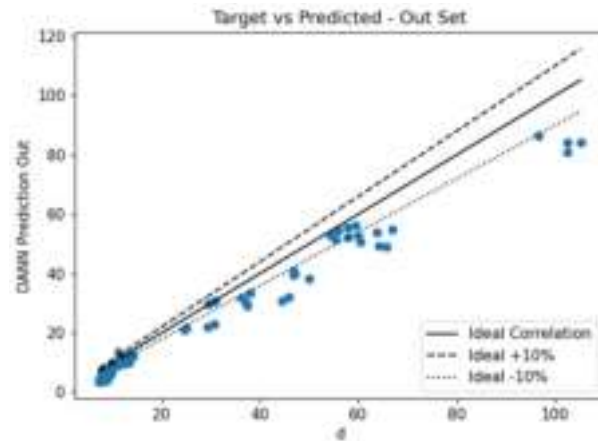


Figure 30. Correlation between the deflection determined using FEM and the deflection estimated using DANN-MPIH-HYT on the validation dataset.

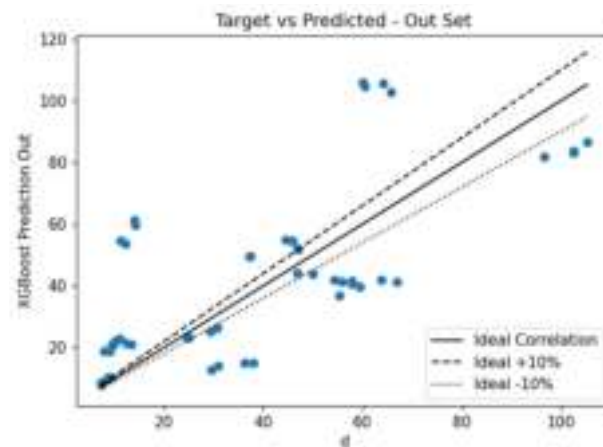


Figure 31. Correlation between the deflection determined using FEM and the deflection estimated using XGBoost-HYT-CV on the validation dataset.

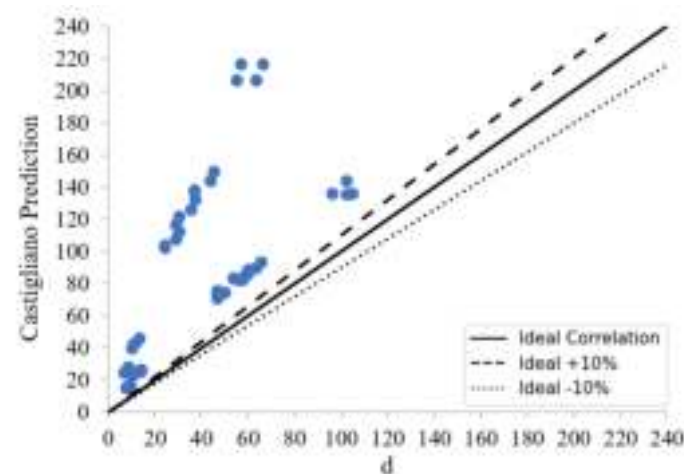


Figure 32. Correlation between the deflection determined using FEM and the deflection estimated using Castigliano's theorem on the validation dataset.

A summary of the error metrics of all ML algorithms can be seen in Table 7. As can be seen, the DANN-MPIH-HYT algorithm outperformed all ML algorithms and the currently used analytical formula. The POLYREG-HYT formula performed the worst. This poor performance was attributed to over-fitting, as the LR formula was quite accurate. It should be noted that even though these results were verified using FEM models, the FE modelling technique used was experimentally verified using an experimental beam where the error experienced was negligible.

Table 7. Summary of the error metrics on the validation dataset (case of deflection).

Algorithm	R	MAPE	MAMPE	MAE	RMSE
LR	0.9267	82.37%	51.84%	20.471	24.426
POLYREG-HYT	0.5174	122.73%	131.81%	52.047	72.419
DANN-MPIH-HYT	0.988	19.08%	16.19%	6.393	8.389
XGBoost-HYT-CV	0.7007	68.37%	42.40%	16.744	21.509
Castigliano	0.6781	155.42%	126.23%	49.842	65.574

8.2. Validation of Failure Load ML Results

The proposed predictive models previously discussed were used to predict the failure load of the out-of-sample data. This section outlines the findings. No analytical formulae are discussed in this section, as the focus of this article is more on deflection estimation and not failure mode investigation. Due to the complexities associated with the failure modes of horizontally curved beams, numerous equations were consulted, and their inclusion would overwhelm the article. The focus herein is on DANN-MPIH-HYT and XGBoost-HYT-CV since they were found to outperform the other predictive models. Correlation plots can be seen in Figures 33 and 34.

The results paint a different picture from what was seen during training and testing. The XGBoost-HYT-CV did not seem to achieve better results than those generated with the DANN-MPIH-HYT algorithm; however, a closer look at the error metrics was required. A summary of the error metrics can be seen in Table 8.

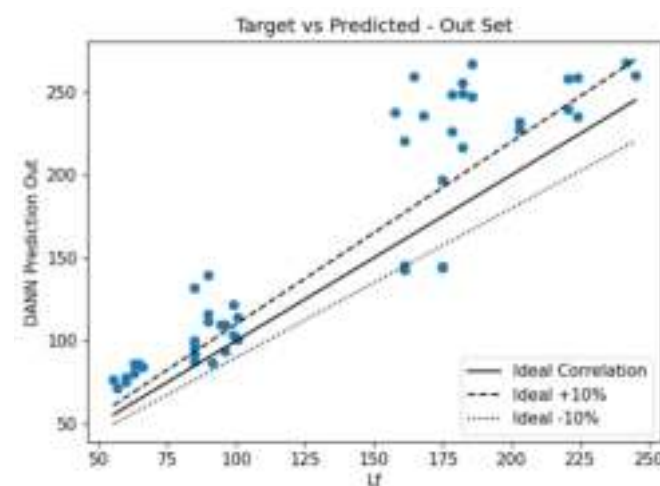


Figure 33. Correlation between the failure load determined using FEM and the failure load estimated using DANN-MPIH-HYT on the validation dataset.

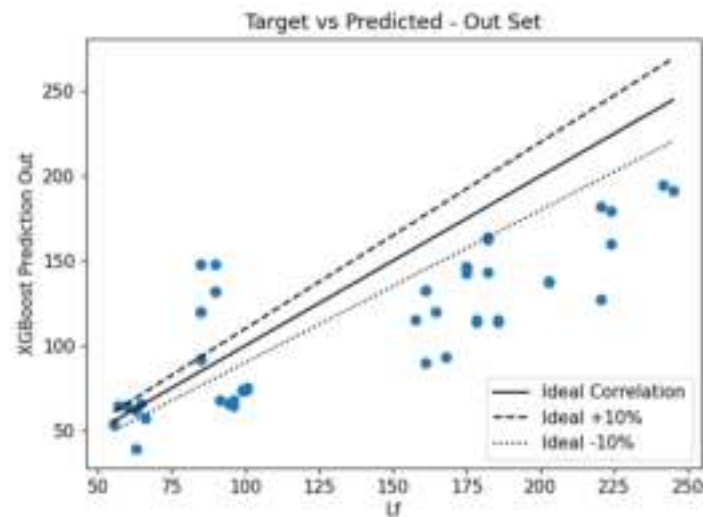


Figure 34. Correlation between the failure load determined using FEM and the failure load estimated using XGBoost-HYT-CV on the validation dataset.

Table 8. Summary of error metrics on the validation dataset (case of failure load).

Algorithm	R	MAPE	MAMPE	MAE	RMSE
LR	0.1440	478.86%	403.82%	548.0527	560.4963
POLYREG-HYT	0.2843	539.71%	500.12%	678.7503	885.5885
DANN-MPIH-HYT	0.9235	23.22%	22.42%	30.4302	38.4831
XGBoost-HYT-CV	0.7914	26.42%	26.87%	36.4645	42.9955

As can be seen, the correlation was significantly worse and all the error metrics showed that the predictions provided through the XGBoost-HYT-CV were less accurate than the predictions provided through the DANN-MPIH-HYT algorithm. This is the same phenomenon as was experienced in the deflection dataset where the DANN-MPIH-HYT algorithm performed better with the out-of-sample data compared with all the other ML algorithms. Therefore, the extended duration required to train the DANN-MPIH-HYT algorithm can be said to translate to improved accuracy during validation.

9. Conclusions and Recommendations

This study showed that FE modelling was able to replicate the experimental results acquired with horizontally curved steel I-beams. For the needs of this research work, an experiment was performed that involved the loading of an IPE100 curved steel I-beam under a vertical load. The experimental data obtained were then used to validate numerical models through the use of CivilFEM software. The usage of quadratic solid FEs with a hexahedral mesh size of 50 mm as well as appropriate material constitutive models paired with nonlinear analyses led to the most accurate results when compared with the experimental data.

A parametric investigation was conducted in which it was noted that the FE modelling technique outlined in this research work was capable of accurately estimating the midspan deflection and rotation of horizontally curved steel I-beams. The MAE was noted to be 3.32% when estimating deflection and 9.17% when estimating rotation, which far outperformed analytical methods, which had an error of 19.80% when estimating deflection and 16.07% when estimating rotation.

The experimentally validated model was then used to develop numerous FE meshes that were analysed under ultimate limit state loading conditions. The numerically obtained results were used to develop a large dataset. A total of 864 models were created that

encompassed the entire IPE cross-section list (IPE100 to IPE600). A total of 10 independent variables were considered in this study.

Numerous ML algorithms were used, namely, LR, POLYREG-HYT, DANN-MPIH-HYT, and XGBoost-HYT-CV, for developing the proposed predictive models. During the validation phase, experimentally validated FE models that were outside the training dataset were created to validate the various proposed predictive models proposed in this research work. To further evaluate the accuracy of the available analytical methods for computing the deflection and ultimate load of curved steel I-beams, the validation data were used to assess Castigliano's analytical formulae as well. According to the numerical findings, the DANN-MPIH-HYT algorithm was the most accurate in estimating both deflection and failure load. When estimating deflection, the DANN-MPIH-HYT proposed predictive model had a MAMPE of 16.19%. This was found to be a significant improvement compared with the analytical method, Castigliano's second theorem, which derived an extremely large MAMPE of 126.23%, highlighting the need for more accurate and objective predictive models. It is also safe to conclude that based on the findings of this dissertation, the proposed ML-generated predictive models are far more accurate than the current analytical methods for estimating the deflection of curved steel I-beams.

In addition to the above, when estimating the failure load of curved steel I-beams, the DANN-MPIH-HYT generated model was found to be significantly more accurate than the other proposed models. This is most likely to be due to the complexities associated with horizontally curved steel I-beams and the nature of the datasets. Horizontally curved steel I-beams have various failure modes, given that they can fail due to flexure, torsion, shear, or a combination of these. The beams may also fail due to lateral torsional buckling. Therefore, a larger dataset is required for the ML algorithms to derive the patterns that are connected to the failure modes of the beams. At this stage, the international literature includes numerous equations for each of the failure modes. Therefore, there is no individual formula to compare the estimates with. It is believed that the DANN-MPIH-HYT algorithm, which resulted in a MAMPE of 22.42%, can be further improved in the future by either increasing the dataset or providing different formulae for the different failure modes, as is currently the case in all design codes.

Finally, this research study is able to propose, for the first time in the international literature, accurate and objective predictive models that outperform any known formula used to compute the deflection of curved steel I-beams and their ultimate capacity. The largest datasets currently available in the international literature were also developed for the training and testing of the proposed predictive models. This pilot project paves the way for the development of future design formulae that can be more accurately applied to a larger range of beams for different boundary conditions.

This research was limited to evaluating the feasibility of using FE models to estimate the midspan deflection and failure load of horizontally curved steel I-beams with a single point load at midspan and fully fixed at both ends. The development of the datasets foresaw the use of minimum and maximum geometrical features. Therefore, the proposed predictive models should not be used for beams with dimensions that are larger or smaller than the respective maximum and minimum geometrical values of the beams found in the datasets developed herein. Furthermore, recommendations for future work should include the following:

- Perform more experiments on curved steel I-beams to further validate the proposed predictive models;
- Consider various boundary conditions, such as torsionally pinned beams, and determine the influence this has on deflection, failure load, rotation, and stress distribution. In theory, this should not significantly impact the failure load. However, this is expected to have a drastic impact on the stress distribution and rotation of the section;
- Consider various loading conditions, such as uniformly distributed loads;
- Consider the impact residual stresses have on stress distribution and investigate whether this has an impact on deflection and failure loads. Residual stresses are far

greater in horizontally curved steel I-beams compared with straight beams, due to the initial cold-forming process. Therefore, residual stresses vary depending on the R/L ratio at hand. Currently, there is no formula available in the international literature that allows accounting for how the residual stresses vary throughout the length of the beam and within the cross-section. Therefore, a detailed investigation is required.

- Develop an accurate formula to determine the midspan rotation of horizontally curved steel I-beams using a relevant dataset.

Author Contributions: Conceptualization, G.M. and E.A.; Methodology, G.M., E.A. and S.S.; Software, E.A. and G.M.; Validation, E.A. and G.M.; Formal Analysis, E.A.; Investigation, E.A.; Resources, G.M. and E.A.; Data curation, E.A.; Writing—original draft preparation, E.A.; Writing—Reviewing and editing, G.M. and S.S.; Visualization, E.A. and G.M.; Supervision, G.M. and S.S.; Project Administration, G.M.; Funding Acquisition, E.A. and G.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Research Fund (NRF South Africa [MND21062361 5086]) and the APC was funded by the journal.

Data Availability Statement: The datasets that were developed for the needs of this research work can be found through the following link (<https://github.com/nbakas/nbml/tree/a0d27c94dd590688815180ebf6428963a24ca245/datasets>, accessed on 1 July 2024).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Liew, J.R.; Thevendran, V.; Shanmugam, N.; Tan, L. Behaviour and design of horizontally curved steel beams. *J. Constr. Steel Res.* **1995**, *32*, 37–67. [\[CrossRef\]](#)
2. Sonck, D.; Van Impe, R. Study of residual stresses in I-section members and cellular members. In Proceedings of the Annual Stability Conference Structural Stability Research, St. Louis, MO, USA, 16–20 April 2013.
3. Rossini, N.; Dassisti, M.; Benyounis, K.; Olabi, A.-G. Methods of measuring residual stresses in components. *Mater. Des.* **2012**, *35*, 572–588. [\[CrossRef\]](#)
4. King, C.; Brown, D. *Design of Curved Steel*; Steel Construction Institute Ascot: Ascot, UK, 2001.
5. Rettie, N.J. Lateral Torsional Buckling Resistance of Horizontally Curved Steel I-Girders. MSc Thesis, University of Alberta, Edmonton, AB, Canada, 2015.
6. Dowsell, B. *Design Guide 33: Curved Member Design*; American Institute of Steel Construction: Chicago, IL, USA, 2018.
7. Lee, K.; Davidson, J.S.; Choi, J.; Kang, Y. Ultimate strength of horizontally curved steel I-girders with equal end moments. *Eng. Struct.* **2017**, *153*, 17–31. [\[CrossRef\]](#)
8. Wong, Y.-C. Horizontally Curved Beam Analysis and Design. MSc Thesis, Oregon State University, Corvallis, OR, USA, 1969.
9. Dahlberg, T. Procedure to calculate deflections of curved beams. *Int. J. Eng. Educ.* **2004**, *20*, 503–513.
10. Shanmugam, N.; Thevendran, V.; Liew, J.R.; Tan, L. Experimental study on steel beams curved in plan. *J. Struct. Eng.* **1995**, *121*, 249–259. [\[CrossRef\]](#)
11. Shanmugam, N.; Mahendrakumar, M.; Thevendran, V. Ultimate load behaviour of horizontally curved plate girders. *J. Constr. Steel Res.* **2003**, *59*, 509–529. [\[CrossRef\]](#)
12. Dabaon, M.; El-Naggar, M.; Yossef, N. Experimental and theoretical study of curved rolled and castellated composite beams. *Alex. Eng. J.* **2003**, *42*, 219–230.
13. Lim, J.; Kang, Y.-J.; Lee, J.; Kim, S.; Lee, K. Experimental and Analytical Study of Horizontally Curved I-Girders Subjected to Equal End Moments. *Metals* **2021**, *11*, 1132. [\[CrossRef\]](#)
14. Markou, G.; Bakas, N.; Papadrakakis, M.; Chatzichristofis, S. A General Framework of High-Performance Machine Learning Algorithms: Application in Structural Mechanics. *Comput. Mech.* **2024**, *73*, 705–729. [\[CrossRef\]](#)
15. Gravett, D.Z.; Moulas, C.; Taljaard, V.-L.; Bakas, N.; Markou, G.; Papadrakakis, M. New fundamental period formulae for soil-reinforced concrete structures interaction using machine learning algorithms and ANNs. *Soil Dyn. Earthq. Eng.* **2021**, *144*, 106656. [\[CrossRef\]](#)
16. Markou, G.; Bakas, N.P. Prediction of the shear capacity of reinforced concrete slender beams without stirrups by applying artificial intelligence algorithms in a big database of beams generated by 3D nonlinear finite element analysis. *Comput. Concr.* **2021**, *28*, 433–447.
17. AlHamaydeh, M.; Markou, G.; Bakas, N.; Papadrakakis, M. AI-based shear capacity of FRP-reinforced concrete deep beams without stirrups. *Eng. Struct.* **2022**, *264*, 114441. [\[CrossRef\]](#)
18. Calitz, D.; Markou, G.; Bakas, N.; Papadrakakis, M. Developing Fundamental Period Formulae for Steel Framed Structures through Machine Learning and Automated Algorithms. In Proceedings of the Compdyn 2023, Athens, Greece, 12–14 June 2023.

19. Chibaya, V.T.; Markou, G.; Bakas, N. Development of Formulae for the Section Rotations Due to Bending of Curved Steel I-Beams through Ai and MI Algorithms. In Proceedings of the Compdyn 2023, Athens, Greece, 12–14 June 2023.
20. van der Westhuizen, A.M.; Bakas, N.; Markou, G. Developing an Artificial Neural Network Model That Predicts the Fundamental Period of Steel Structures Using a Large Dataset. In Proceedings of the Compdyn 2023, Athens, Greece, 12–14 June 2023.
21. Ababu, E.; Markou, G.; Bakas, N. Using Machine Learning and Finite Element Modelling to Develop a Formula to Determine the Deflection of Horizontally Curved Steel I-Beams. In Proceedings of the 14th International Conference on Agents and Artificial Intelligence, Virtual, 3–5 February 2022; pp. 958–963. [[CrossRef](#)]
22. Aminu, I.M.; Shariff, M.N.M. Strategic orientation, access to finance, business environment and SMEs performance in Nigeria: Data screening and preliminary analysis. *Eur. J. Bus. Manag.* **2014**, *6*, 124–132.
23. Dimopoulos, T.; Tyrallis, H.; Bakas, N.P.; Hadjimitsis, D. Accuracy measurement of Random Forests and Linear Regression for mass appraisal models that estimate the prices of residential apartments in Nicosia, Cyprus. *Adv. Geosci.* **2018**, *45*, 377–382. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Modelling the Impact of Cloud Storage Heterogeneity on HPC Application Performance

Jack Marquez  and Oscar H. Mondragon * 

Faculty of Engineering, Universidad Autonoma de Occidente, Cali 760030, Colombia; jdmarquez@uao.edu.co

* Correspondence: ohmondragon@uao.edu.co

Abstract: Moving high-performance computing (HPC) applications from HPC clusters to cloud computing clusters, also known as the HPC cloud, has recently been proposed by the HPC research community. Migrating these applications from the former environment to the latter can have an important impact on their performance, due to the different technologies used and the suboptimal use and configuration of cloud resources such as heterogeneous storage. Probabilistic models can be applied to predict the performance of these applications and to optimise them for the new system. Modelling the performance in the HPC cloud of applications that use heterogeneous storage is a difficult task, due to the variations in performance. This paper presents a novel model based on Extreme Value Theory (EVT) for the analysis, characterisation and prediction of the performance of HPC applications that use heterogeneous storage technologies in the cloud and high-performance distributed parallel file systems. Unlike standard approaches, our model focuses on extreme values, capturing the true variability and potential bottlenecks in storage performance. Our model is validated using return level analysis to study the performance of representative scientific benchmarks running on heterogeneous cloud storage at a large scale and gives prediction errors of less than 7%.

Keywords: HPC cloud; heterogeneous storage; performance modelling; extreme value theory



Citation: Marquez, J.; Mondragon, O.H. Modelling the Impact of Cloud Storage Heterogeneity on HPC Application Performance. *Computation* **2024**, *12*, 150. <https://doi.org/10.3390/computation12070150>

Academic Editor: Demos T. Tsahalidis

Received: 15 May 2024

Revised: 22 June 2024

Accepted: 11 July 2024

Published: 19 July 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

High-performance computing (HPC) can be defined as the use of supercomputers to efficiently solve complex computational problems [1], while cloud computing is defined as ubiquitous and on-demand access to configurable computing resources [2]. Although these two technologies were initially designed for different purposes, recent attempts to integrate them have given rise to the new concept of the HPC cloud. Netto et al. [3] defined the HPC cloud as “the use of cloud resources to run HPC applications”. Both new challenges and new opportunities have emerged from this work.

From around 2009 onwards, some HPC users started to consider the cloud as a cost-effective alternative to high-cost HPC clusters. The idea of a pay-on-demand business model that was flexible and provided more customer control over resources was tempting [4,5]. Cloud computing allows for the flexible provision of computing resources such as CPUs, memory, storage, networks and graphics processing units. These resources are elastic, meaning that they can be scaled according to the requirements of each application. Instantaneous access to computational resources can help users to deploy or test their application when they need it [6].

The deployment of HPC applications over cloud computing clusters presents several challenges that have yet to be resolved. One potential problem concerns storage systems, as cloud clusters do not use the same types of storage systems as HPC clusters. Storage systems can also be considered an issue in the adoption of cloud systems for HPC applications [3,7,8]. Anticipating the performance of large-scale applications in heterogeneous storage systems is challenging. Technical documentation on aspects such as the throughput

and latency of storage technologies is not sufficient to predict the performance of HPC applications, given the impacts of multiple sources of interference.

In this paper, we describe a statistical model based on Extreme Value Theory (EVT) that can be used to characterise and predict the performance of HPC applications that rely on heterogeneous storage in cloud systems. Our paper makes the following contributions:

- We present an EVT-based model for characterising the performance of HPC applications that make use of heterogeneous storage technologies in cloud computing systems.
- We develop a method for predicting the performance of HPC applications based on return level analysis with different numbers of storage nodes, which can inform storage algorithms and hence improve the performance of applications.
- We evaluate the proposed model by using it to predict the performance of HPC applications running on heterogeneous cloud storage at large scales.

This paper is organised as follows: Section 2 describes the main concepts associated with heterogeneous storage in cloud computing infrastructures, parallel file systems for supporting HPC applications in the cloud, and Extreme Value Theory. In Section 3, we describe our approach to modelling the impact of using heterogeneous cloud storage on the performance of HPC applications, and we validate this model in Section 4 using return level analysis. Section 5 describes some of the main related works. Finally, we present conclusions and suggestions for future work in Section 6.

2. Background

In this section, we describe the storage systems currently used in general-purpose cloud computing environments, and show how these can be leveraged efficiently in HPC. We then focus on BeeGFS as a flexible and scalable alternative for the deployment of a parallel file system that is optimised for HPC on cloud infrastructure. Finally, we describe key concepts related to the modelling of application performance using EVT.

2.1. Heterogeneous Storage Systems in Cloud Computing

In cloud computing clusters, the performance of data-intensive applications is limited by disk data transfer rates, among other factors. To mitigate the impact on performance, cloud systems that offer hierarchical and heterogeneous storage architectures are becoming commonplace. The integration of different storage alternatives such as solid-state drives (SSDs), hard disk drives (HDDs), or even RAMDISK (a block of RAM used as volatile storage) may improve the performance of applications by taking advantage of the characteristics of each type of storage [9].

Table 1 presents the storage technologies and services offered by the top three cloud providers. It can be seen that cloud providers offer a variety of heterogeneous storage devices, each of which has certain particularities. For instance, Azure provides disks with high throughput, high IOPS and low latency (Ultra disk), but also provides a low-cost disk with standard throughput, IOPS and latency (Standard HDD). In addition to a general-purpose SSD, AWS provides a type of SSD with configurable IOPS (io2-io1), and another optimised for speed. Each of these disks is designed for special cases, and offers different capabilities such as volume size, maximum IOPS and maximum throughput. GCP also provides different types of disks and offers zonal and regional replication, as well as high-performance disks (Extreme Persistent Disk). The three main types of storage service are object, block and file storage, which make use of the available disk types.

Table 1. Storage services offered by the top three cloud providers.

Cloud Platform	Disk Types	Storage Services/File Systems
Azure	Premium SSD	Blob Storage
	Standard SSD	Azure Files
	Standard HDD	Queues
	Ultra Disk	Tables
Amazon Web Services		Disks
	SSD-io2-Provisioned IOPS	S3
	SSD-io2 Block Express	Elastic Block Store
	SSD-io1	Elastic File System
	SSD-gp3-General Purpose	Glacier
	SSD-gp2	Storage Gateway
	HDD-st1-Optimized Speed	FSx Windows
Google Cloud Platform	HDD-sc1-Cold	FSx Lustre
	SSD-NVMe	Block Storage-Persistent disk
	SSD	Cloud Storage
	HDD	Standard storage
	Zonal Persistent Disk	Nearline storage
	Regional Persistent Disk	Coldline storage
	Extreme Persistent Disk	Archive
		Filestore

2.2. Leveraging Heterogeneous Cloud Storage for HPC

Modern HPC systems are designed to provide high performance for scientific applications by offering low-latency networks such as InfiniBand and optimised distributed parallel file systems. LustreFS [10] and BeeGFS [11] are two of the most commonly used parallel file systems in the top500 HPC clusters. Both of these parallel systems are being increasingly deployed on cloud platforms [11–13], with the aim of achieving fast access to large amounts of data in cloud clusters.

The use of Lustre and BeeGFS helps to maximise the writing and reading process of large amounts of data in cloud clusters. In this work, we focus on the BeeGFS parallel file system, as it has certain features that make it particularly suitable for cloud environments, such as its simple installation. It is also available for different Linux distributions, is hardware agnostic, and supports high concurrency. An additional benefit is that it is unnecessary to host the management service on a dedicated machine.

BeeGFS Parallel File System

BeeGFS is a client-server model file system that was developed with a focus on performance and scalability. Figure 1 shows the architecture of BeeGFS, which is composed of three types of nodes: a management server, a metadata server, and an object storage server. The management server is the node in charge of the configuration of the BeeGFS file system and its other components; there is usually only one management server in a BeeGFS configuration. The metadata server contains the metadata target, a storage device that carries the structure of the file system and the file names. This server also manages indexes and namespaces. The object storage server is responsible for receiving the data sent from the client and storing it in the object storage target (OST). A given BeeGFS configuration may involve a high number of OSTs, and BeeGFS will try to store data efficiently by splitting the workload among them.

BeeGFS was developed with easy-to-use feature installation and management. Two of the most important aspects of this simple installation are that BeeGFS does not require a kernel patch, and that it comes with graphical Grafana [14] dashboards. BeeGFS also has a striping feature, and the chunk size, the numbers of storage nodes to use (targets) and the heterogeneous storage support can be specified.

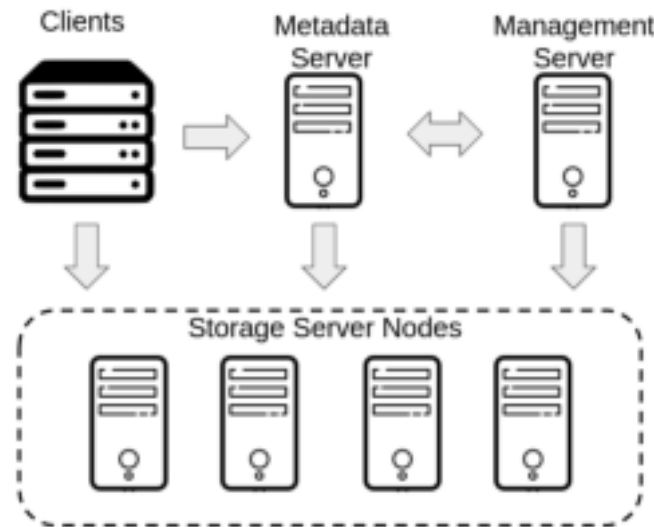


Figure 1. BeeGFS architecture.

2.3. Modelling Extreme Values

EVT is a field of statistics that relates to the behaviour of exceptional or extreme values of a set of random variables, i.e., those which deviate from the median of the probability distribution [15]. EVT has been applied in areas such as statistical quality control [16], finance [17–19], transportation [20], study of calcium content [21] and in hydrology to calculate the probability of floods in a certain period [22–24]. In computer science, EVT has been used to predict HPC applications' performance [25] and the impact of interference sources on HPC applications' performance [26,27] on HPC clusters.

Our work focuses on the impact of storage actions on HPC applications provisioned in cloud environments. Standard statistical analyses focus on average performance and may overlook significant outliers that impact HPC applications in cloud environments. Extreme Value Theory (EVT) is necessary to model these outliers and provide accurate predictions of worst-case performance scenarios. Measuring extreme execution times is crucial because it allows EVT to capture the system's true variability. Unlike traditional methods focused on averages, EVT prioritises these outliers, which can be hidden bottlenecks or signs of future issues. By understanding these extremes, we can proactively allocate resources for peak loads and prevent outages. EVT even allows us to quantify the likelihood of such events, enabling systems design with sufficient capacity. This focus on extremes, rather than just averages, ensures more reliable and efficient resource provisioning.

EVT was developed based on the results of work by Fisher and Tippett [28,29]. These authors showed that the distributions of a set of random and independent, identically distributed (i.i.d.) variables with an unknown underlying distribution and outliers or extremes could converge to one of three possible asymptotic distributions: Frechet, Gumbel or Weibull. Jenkinson [30] introduced a generalised extreme value (GEV) distribution that combined these three distributions. Equation (1) shows the cumulative distribution (cdf) for the GEV distribution.

$$F(x|\xi, \mu, \sigma) = e^{-[1+\xi(\frac{x-\mu}{\sigma})]^{-\frac{1}{\xi}}} \quad (1)$$

Equation (1) has three parameters: scale σ , location μ and shape ξ , where, $-\infty < \xi < \infty$, $-\infty < \mu < \infty$ and $\sigma > 0$. In this equation, the scale parameter defines the dispersion of the distribution or the variability, whereas the location parameter defines where the distribution is centred on the real axis, and the shape parameter, also known as the extreme value index, defines whether the distribution is a Gumbel ($\xi = 0$), Fréchet ($\xi > 0$) or Weibull distribution ($\xi < 0$) [31] (see Figure 2).

To fit the GEV distribution, samples of extreme values from the measured values must be obtained. There are two common methods for carrying out this process: peaks over threshold (POT) and the block maxima method (BMM). In the POT approach, a threshold is established and samples are selected from above that threshold. BMM consists of dividing the measured values into blocks of size n and selecting the maximum value from each block. BMM has been widely used in hydrology, where the block size is set to obtain the GEV parameters for seasonal variation over several years [22].

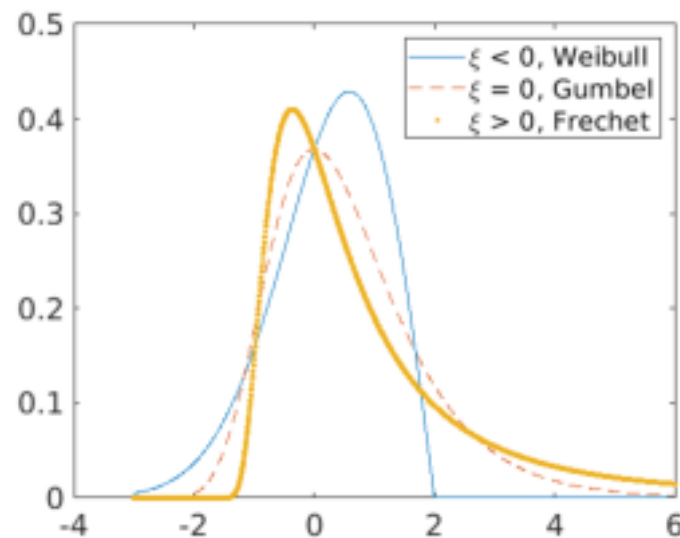


Figure 2. Probability density function.

Three-parameter estimation has been widely studied in relation to GEV [32]. The most well-known methods are maximum likelihood estimation (MLE) [33] and the method of L moments (LMOM) [34]. MLE is the most frequently used method of estimating GEV parameters in conjunction with BMM. Its main advantage is based on the asymptotic normality and parameter estimation when there is a known prior distribution. According to Smith [35], to establish the asymptotic property of GEV and obtain suitable estimators, ξ must be a value in the range $(-0.5, \infty)$. If ξ is in the range $(-1, -0.5)$, estimators can be obtained but the asymptotic property is not established. Finally, if ξ is in the range $(-\infty, -1)$, estimators can not be obtained.

Once the GEV model has been determined and the parameters have been calculated, it is possible to calculate the return level values of the extremes. A return level is defined as the value that will be exceeded on average only once for every N samples, or for every N blocks of the distribution when using the block maxima method [36]. We can calculate this return value as:

$$RL = F^{-1}(P) \quad (2)$$

In Equation (2), F represents the GEV distribution, and P is given by $P = 1 - 1/i$, where i , is the return period, defined as the average length of time between events with the same or a more significant value. Return level analysis has historically been applied more often in domains such as finance [37] and hydrology [38] than in computer science [26].

3. Modelling Heterogeneous Cloud Storage Impact on HPC Application Performance

This section describes our proposed stochastic model for the analysis of heterogeneous storage performance and the way in which we estimate the parameters for this model, which allows us to explore the impact of heterogeneous storage on HPC cloud systems. Unlike previous approaches, our model leverages Extreme Value Theory (EVT) to address the variability and extremities in storage performance without assuming any a priori distribution of the storage times. This novel application of EVT provides a more flexible

and accurate representation of storage performance in parallel distributed systems such as BeeGFS, thereby offering new insights into the performance dynamics of HPC applications.

3.1. Modelling Approach

The storage time in distributed parallel file systems such as BeeGFS is dominated by the time taken by the slowest storage node in the cluster. We use EVT to model the data storage performance in this type of system. To do this, we assume that the time that a distributed parallel file system takes to store each chunk of a file is i.i.d.

We use BMM to model the extremes. The block and sample sizes must be large enough to obtain a good model fitting. Several methods have been proposed for testing the goodness of fit of a model [39,40]; these studies have presented approaches for the selection of the size of the blocks, using BMM and the sample, and have also suggested techniques for testing the asymptotic property using graphs or by analysing the values resulting from the estimation of parameters. We describe a rule of thumb for testing the asymptotic property in Section 2.3.

3.2. Estimating the Model Parameters

In our block maxima formulation, the block size is the number of storage nodes configured in BeeGFS, and we sample the value of the execution time from the slowest node. After selecting a set of block maxima, we need to estimate the parameters of the GEV distribution. Henwood et al. [41], suggested a minimum block number of 60; we use a block number of 100 and follow a method based on the normality test suggested in [40] to determine whether the values are normally distributed with a 95% confidence interval and to avoid unbiased estimations. We then apply MLE to estimate the GEV parameters.

3.3. Predicting Performance on Heterogeneous Storage

Predicting the performance of applications in cloud environments is essential to achieve efficient resource provisioning. We use return level analysis and our model to predict the performance of scientific applications using cloud infrastructure at scale. After obtaining estimations for the three GEV parameters, we can use return level analysis based on Equation (2), as described in Section 2.3. In our model, the return period in Equation (2) is the block number used in the block maxima method.

4. Validation of the Model

In this section, we present a validation of our proposed model using representative scientific benchmarks. We study the performance of these benchmarks running on heterogeneous cloud storage, and use our model to predict performance at larger scales. The results are validated against simulations and observed values. To collect the initial data to fit our GEV model, we configured a BeeGFS cluster consisting of eight nodes. We used a block size equal to the number of nodes to apply BMM and MLE for GEV parameter estimation. Subsequently, we expanded the cluster up to 16, 32, 64 and 128 nodes to collect performance data at those scales and compare against the results of the return level analysis.

4.1. Experimental Setup

For our experiments, we used nodes from CloudLab [42], a project of the University of Utah (Salt Lake City, UT, USA), Clemson University (Clemson, SC, USA), the University of Wisconsin Madison (Madison, WI, USA), the University of Texas at Austin (Austin, TX, USA), the University of Massachusetts Amherst (Amherst, MA, USA), and US Ignite (Washington, DC, USA). Each of the configured nodes had three available storage targets (RAMDISK, SSD, HDD) (see Figure 3). For these experiments, we used c220g2 nodes. Each of these nodes had two Intel E5-2660 v3 10-core CPUs at 2.60 GHz (Haswell EP), 160 GB ECC Memory (10 × 16 GB DDR4 2133 MHz dual-rank RDIMMs), one Intel DC S3500 480 GB 6 G SATA SSDs, two 1.2 TB 10K RPM 6G SAS SFF HDDs, and a dual-port Intel X520

10 Gb NIC (PCIe v3.0, eight lanes), and frequency scaling was disabled. The operating system was CentOS 7 with kernel version 3.10.0-1127.19.1.el7.x86_64.

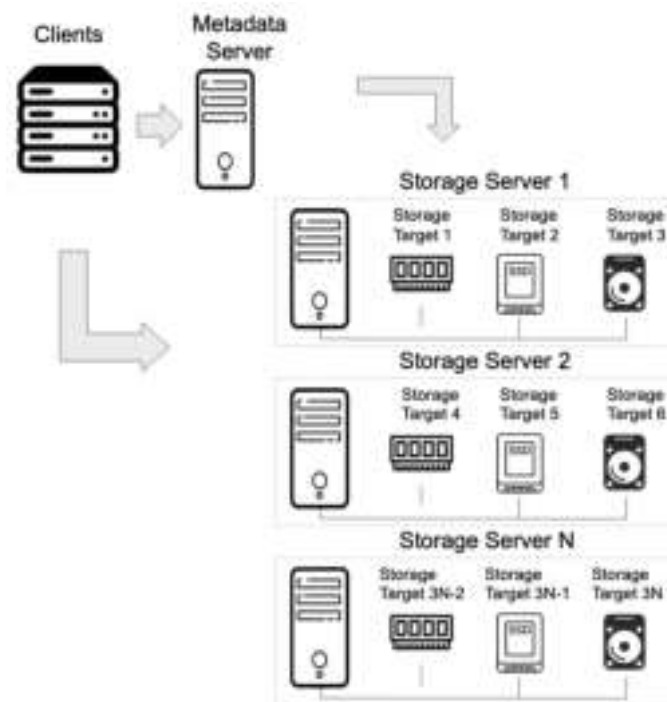


Figure 3. The architecture of BeeGFS.

4.1.1. Benchmarks

We selected several representative benchmarks (see Table 2) to validate our model. These benchmarks copy fundamental parts of applications and simulate their behaviour:

- Iozone is a tool used for the analysis of file systems. It includes different operations such as write, read, re-write, and re-read, and allows for latency and bandwidth analysis. We use the write operation with a single stream measurement in order to allow BeeGFS to receive all the files as a single unit, and then split this over all the storage nodes.
- Fallocate is a Linux function that is widely used as a benchmark for file systems, as it can preallocate new spaces for files of a specific size. This tool first analyses whether the required space is sufficient for the file system, and then reserves that space for the file if required.
- BT was developed by NAS Parallel Benchmark (NPB) and is extensively used for testing HPC clusters. It solves a highly configurable block-tridiagonal (BT) problem using MPI.
- PIOS is a test tool created to work as an I/O simulator on file systems. This tool simulates a load from many clients that generate I/O in file systems. Due to its parallel nature, PIOS can write to the same or different files at the same time, but in this study it is used to write only to a single file.

Table 2. Selected benchmarks for performance analysis.

Benchmark	Ref.
Iozone	[43,44]
Fallocate	[45]
BT-NPB	[46,47]
PIOS	[48]

4.1.2. Data Collection

For data collection, we followed an experimental methodology according to the recommendations in [49], to avoid wrong measures. Some of the recommendations are about record keeping, in which they recommend adding labels and metadata to the data collected to be able to quickly retrieve and check them later. We used four benchmarks (Fallocate, Iozone, BTC, and PIOS) and three types of storage (RAMDISK, SSD, and HDD), and aimed to ensure that each run of each benchmark had the same testing conditions in the cluster. We therefore developed a script to sample the execution times of the slowest node for each benchmark by repeating every run 100 times and collecting all the possible benchmark-to-storage times at each iteration. The process of benchmark-to-storage is the execution time per node, in this case the slowest node in our configuration that stores a piece of data.

4.2. Estimation of GEV Parameters

In order to estimate the GEV parameters for each benchmark, we used BMM and MLE with 100 blocks and a block size of eight. Table 3 shows the parameters obtained with this approach.

The results in Table 3 allow us to determine which of the three GEV types of distributions gives the best fit. The results indicate that all the distributions are the Weibull type ($\xi < 0$). The shape parameters in all cases are greater than -0.5 , thus meeting the asymptotic condition for the GEV distribution. The runtime distributions of the selected benchmarks using BeeGFS on heterogeneous storage devices show similar tailedness, and therefore fit a Weibull distribution. Also, by checking Figure 4, we can see that although all are upper bound, BT_C on RAMDISK, BT_C on HDD, and PIOS on SSD are close to light tail (Gumbel).

Table 3. Estimated GEV parameters for block maxima sample.

	Shape (ξ)	Scale (σ) (s)	Location (μ) (s)
Fallocate RAMDISK	−0.16	154.47	566.95
Fallocate SSD	−0.17	177.52	711.15
Fallocate HDD	−0.17	192.99	785.41
Iozone RAMDISK	−0.22	70.03	265.10
Iozone SSD	−0.38	75.98	339.61
Iozone HDD	−0.39	103.81	467.63
BT_C RAMDISK	−0.05	116.63	1006.72
BT_C SSD	−0.11	89.46	1268.03
BT_C HDD	−0.07	498.15	2026.47
PIOS RAMDISK	−0.17	21.99	57.42
PIOS SSD	−0.08	25.67	69.20
PIOS HDD	−0.46	36.71	126.54

The skewness values show that all benchmarks except PIOS have the same location of the tail for the SSD and HDD experiments. PIOS has a left-tailed skewness for HDD, and a right-tailed skewness for SSD. This is an interesting outcome, as PIOS on SSD also has the highest coefficient of variation, indicating that this experiment has the most variation in its data; this generates more extreme values, making the prediction process more challenging.

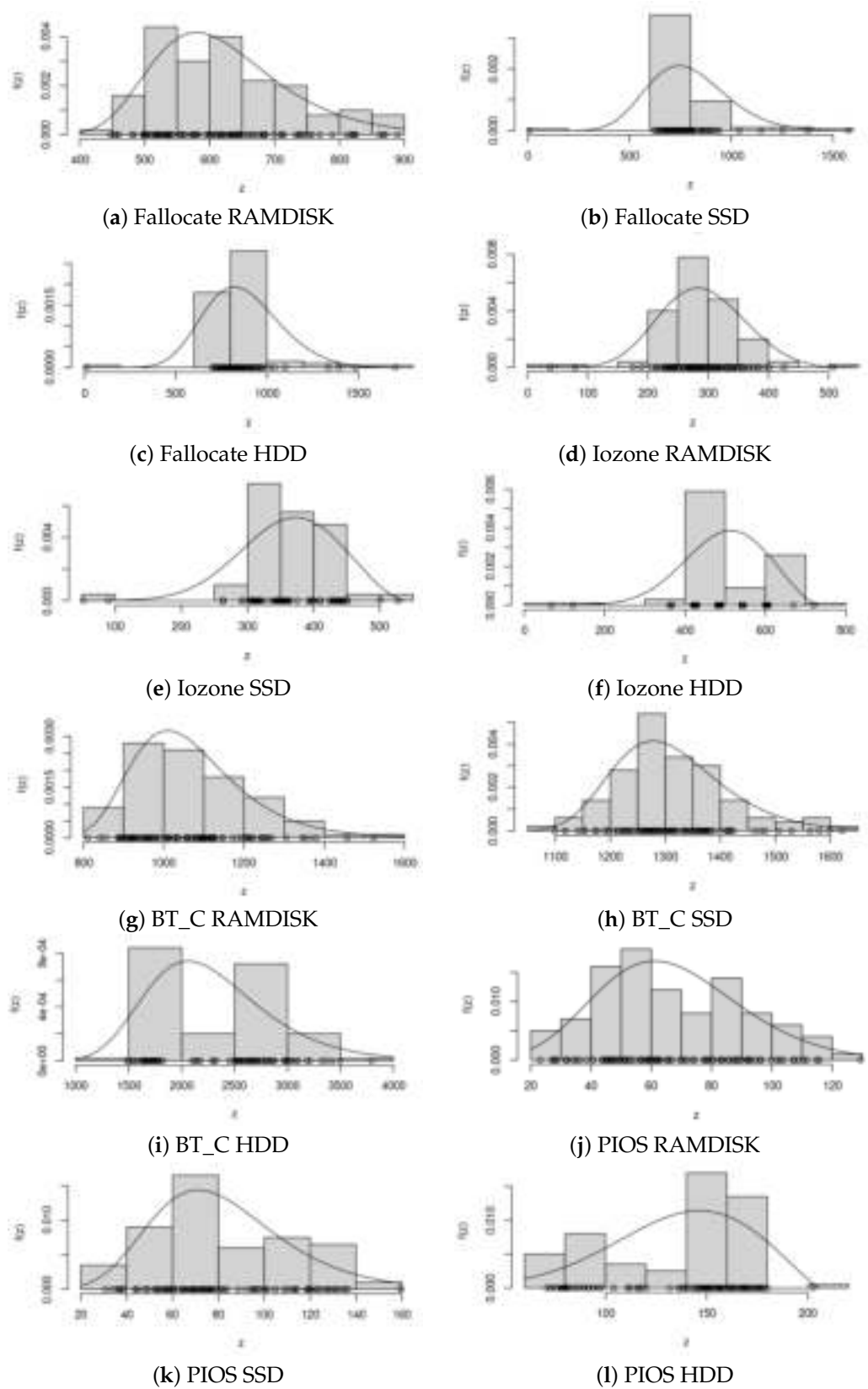


Figure 4. Density Plots (in seconds) for Block Maxima Sample.

4.3. Return Level Analysis

In this section, we use return level analysis to predict the performance of HPC applications on BeeGFS. Figures 5–8 show the return value, computed using Equation (2), for 16,

32, 64 and 128 BeeGFS object storage servers, respectively. These figures present the results of the return level analysis compared to the observed values. This allows us to predict the performance of the application at scale and to compare this to the observed performance. The largest prediction error was 6.64% and the smallest was 0.46%, for PIOS on 64 nodes with RAMDISK and BT_C on 32 nodes with SSD, respectively.

The variations in the return levels for 16 and 128 nodes were larger for BT_C on HDD, PIOS on RAMDISK and PIOS on SSD (See Figures 7c and 8a,b), which were the experiments with shape parameters closer to zero, which often indicates a higher variability in performance outcomes. For instance, the higher variability observed for BT_C on HDD could be due to the inherent slower performance and higher latency of HDDs compared to SSDs or RAMDISK. Similarly, the performance variability for PIOS on RAMDISK might be influenced by the memory management and data handling differences inherent to RAMDISK storage.

The return level values for Iozone did not produce differences as large as those for BT_C on HDD, PIOS on RAMDISK, and PIOS on SSD (see Figure 6). The return level behaviour of Iozone exhibits less variability than other tools because of its single-stream measurement approach and the efficient file distribution handled by BeeGFS. This leads to more consistent performance metrics, as the tool focuses on straightforward write operations without the added complexity of parallel I/O patterns as in PIOS or the computation and communication overheads introduced by BT. The inherent design of Iozone and its specific use case in this study (i.e., writing a single stream) contribute to its stable and predictable performance, reflected in the less variable return levels observed in the analysis.

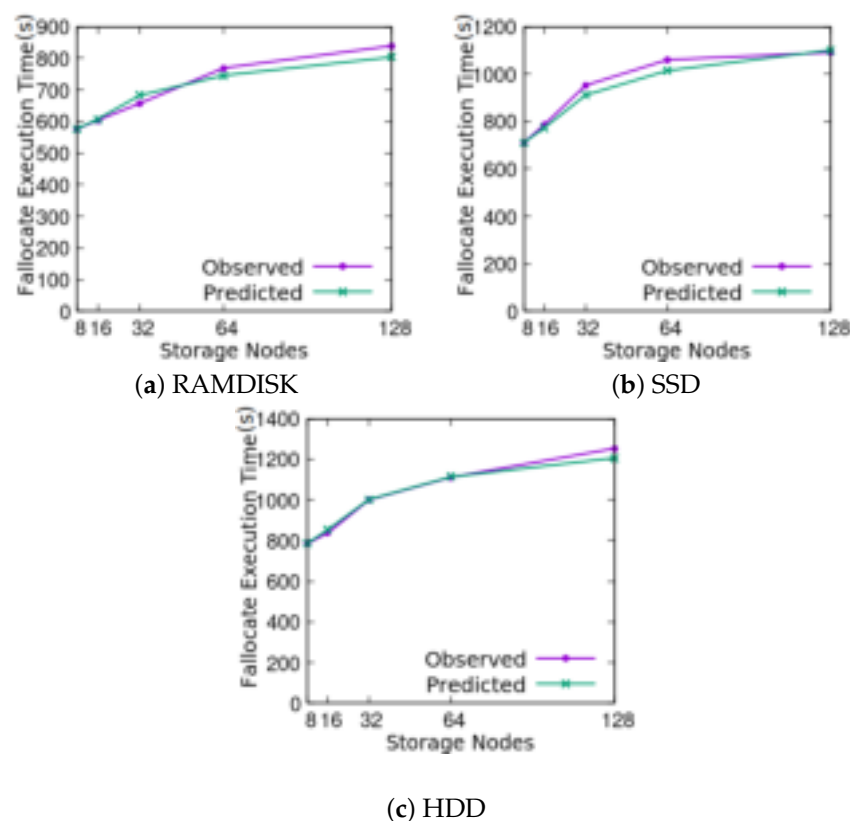


Figure 5. Performance prediction for Falloccate on BeeGFS using heterogeneous storage.

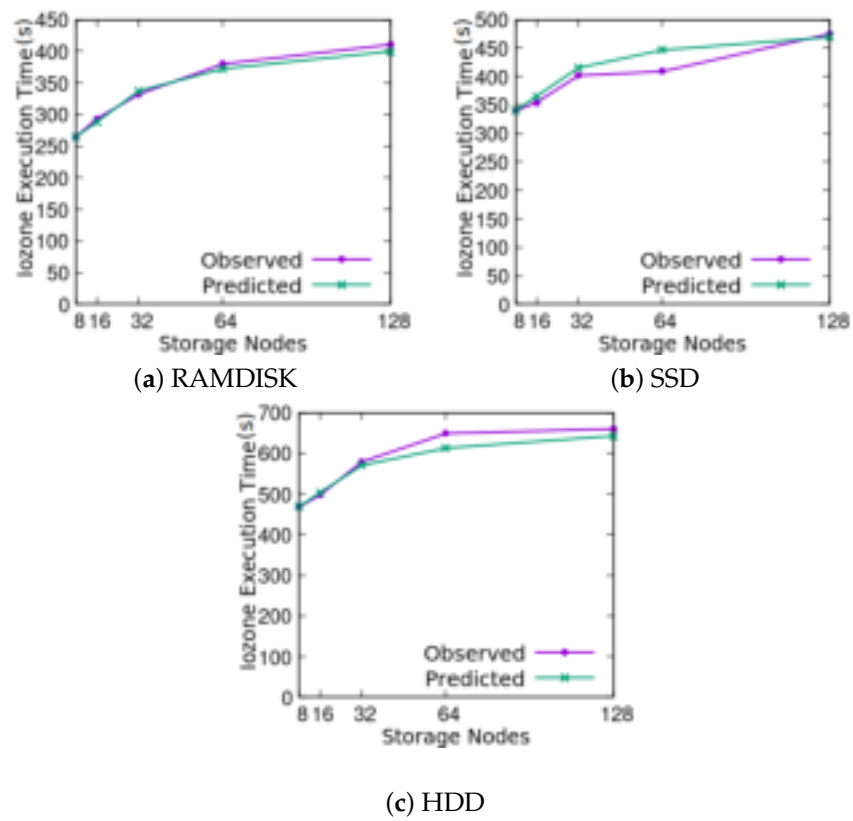


Figure 6. Performance prediction for Iozone on BeeGFS using heterogeneous storage.

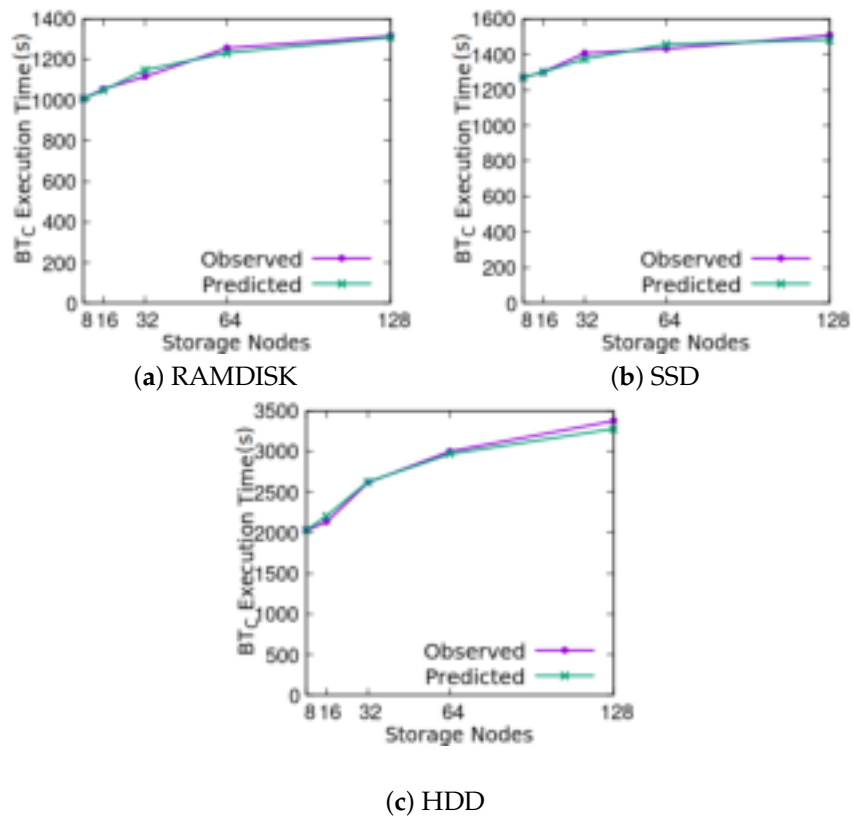


Figure 7. Performance prediction for BT_C on BeeGFS using heterogeneous storage.

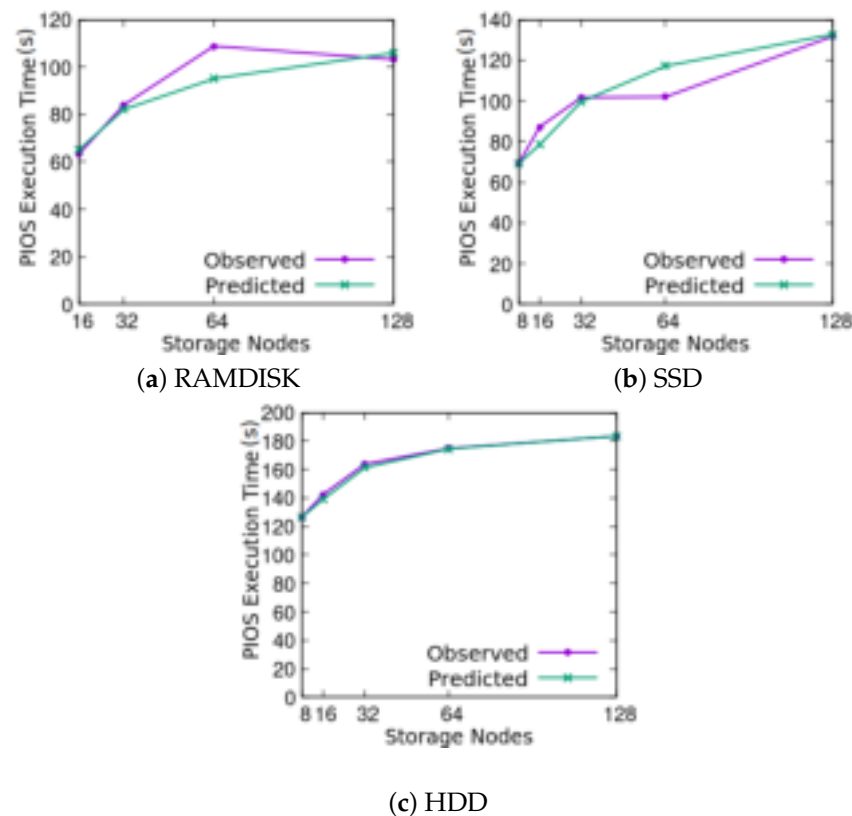


Figure 8. Performance prediction for PIOS on BeeGFS using heterogeneous storage.

5. Related Work

Cloud platforms are increasingly incorporating SSDs into their storage systems. Huang et al. [50] proposed a black-box model to predict the performance of SSDs in terms of the latency, bandwidth and throughput, by applying statistical machine learning algorithms. They evaluated their model using micro-benchmarks and real-world traces from online transaction processing (OLTP) applications, and recorded errors of 9% in the latency prediction and 1% for the bandwidth and throughput. Unlike our work, in which we consider different storage types in our model, this work focused on predicting performance in SSDs but not HDDs.

Mondragon et al. [26] presented a model for analysing the performance of bulk synchronous HPC applications based on the use of EVT. They used their model to characterise the impact of next-generation interference sources on applications and predicted the performance of applications at large scales. Their model obtained a prediction error of less than 7.4% for HPC applications running on HPC clusters. We also apply EVT to characterise and predict the performance of HPC applications using heterogeneous storage in cloud systems.

Dominguez-Trujillo et al. [51] presented an approach for modelling variations in the performance of large-scale HPC systems using EVT through a study of the maximum length of the distributed workload time interval for bulk synchronous HPC applications, using parametric and non-parametric ping. This work focused on the variability generated by the hardware and software used in HPC clusters. Our work instead focuses on analysing and predicting performance variations for heterogeneous storage systems in cloud systems.

Another approach that used EVT was an analysis of the worst-case execution time (WCET) [52], where the authors studied the accuracy of EVT in detecting the WCET for several processes. This approach has also been used for CUDA kernel tasks [53] and the analysis of automotive applications in embedded safety-critical systems [54], but not for the performance of HPC applications using heterogeneous storage systems in the cloud.

6. Conclusions

An increasing number of researchers are considering the use of the cloud to run their HPC applications. This work contributes to the adoption of the cloud as a feasible environment for HPC applications by providing a model for HPC applications in cloud systems and the integration of a high-performance storage file system such as BeeGFS into the cloud. In this paper, we have presented an EVT-based model for analysing, characterising and predicting the performance of HPC applications that use heterogeneous storage technologies in the cloud.

Modelling and understanding the performance of HPC applications that use heterogeneous storage can benefit both cloud providers and cloud users. An accurate prediction of the performance of HPC applications that takes into consideration the storage performance can help cloud providers to offer this feature in their platforms, and can help cloud users to identify the resources they will need more specifically.

In addition to predicting storage performance, our extreme value model can be used as guidance to design intelligent data placement algorithms that consider heterogeneous storage infrastructure, increasingly incorporated by cloud providers, seeking to take advantage of the characteristics of each storage type as well as data locality and data access patterns.

Although our model obtained accurate results with low prediction errors, modelling applications that exhibit heavy-tail storage time distributions such as PIOS is challenging. In those cases, the model might benefit from the use of techniques such as statistical bootstrapping or smoothing sample extremes, in order to more precisely fit collected data to distributions using a small number of samples.

One future direction for this work would be to include in the model other storage devices such as the new NVRAM. The development of a model of the performance of HPC applications based on the use of heterogeneous storage in cloud systems represents only one piece of a larger challenge, and additional cloud resources or even multi-cloud resources could be integrated into the model in future work.

Author Contributions: Conceptualization, J.M. and O.H.M.; Investigation, J.M. and O.H.M.; Methodology, J.M. and O.H.M.; Software, J.M.; Supervision, O.H.M.; Validation, J.M. and O.H.M.; Visualization, J.M.; writing—original draft preparation, J.M. and O.H.M.; writing—review and editing, J.M. and O.H.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: All data were presented in main text.

Acknowledgments: Results presented in this paper were obtained using CloudLab testbed supported by the U.S. National Science Foundation.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Neuwirth, S.; Paul, A.K. Parallel i/o evaluation techniques and emerging hpc workloads: A perspective. In Proceedings of the 2021 IEEE International Conference on Cluster Computing (CLUSTER), Portland, OR, USA, 7–10 September 2021; pp. 671–679.
2. Mell, P.; Grance, T. *The NIST Definition of Cloud Computing*; NIST: Boulder, CO, USA, 2011.
3. Netto, M.A.S.; Calheiros, R.N.; Rodrigues, E.R.; Cunha, R.L.F.; Buyya, R. HPC Cloud for Scientific and Business Applications: Taxonomy, Vision, and Research Challenges. *ACM Comput. Surv.* **2018**, *51*, 1–29. [[CrossRef](#)]
4. Borin, E.; Drummond, L.M.A.; Gaudiot, J.L.; Melo, A.; Alves, M.M.; Navaux, P.O.A. *High Performance Computing in Clouds: Moving HPC Applications to a Scalable and Cost-Effective Environment*; Springer Nature: Berlin/Heidelberg, Germany, 2023.
5. Dancheva, T.; Alonso, U.; Barton, M. Cloud benchmarking and performance analysis of an HPC application in Amazon EC2. *Clust. Comput.* **2024**, *27*, 2273–2290. [[CrossRef](#)]
6. Aithal, P. Information communication & computation technology (ICCT) as a strategic tool for industry sectors. *Int. J. Appl. Eng. Manag. Lett. (IJAEML)* **2019**, *3*, 65–80.
7. dos Santos, M.A.; Cavalheiro, G.G.H. Cloud infrastructure for HPC investment analysis. *Rev. Informática Teórica E Apl.* **2020**, *27*, 45–62. [[CrossRef](#)]
8. Cheriére, N.; Dorier, M.; Antoniu, G. How fast can one resize a distributed file system? *J. Parallel Distrib. Comput.* **2020**, *140*, 80–98. [[CrossRef](#)]

9. Subramanyam, R. HDFS Heterogeneous Storage Resource Management Based on Data Temperature. In Proceedings of the 2015 International Conference on Cloud and Autonomic Computing, Boston, MA, USA, 21–25 September 2015; pp. 232–235. [CrossRef]
10. Braam, P. The Lustre storage architecture. *arXiv* **2019**, arXiv:1903.01955.
11. Heichler, J. An introduction to BeeGFS. 2014. Available online: http://www.beegfs.de/docs/whitepapers/Introduction_to_BeeGFS_by_ThinkParQ.pdf (accessed on 1 April 2024).
12. Souza Filho, P.; Felipe, L.; Aragão, P.; Bejarano, L.; de Paula, D.T.; Sardinha, A.; Azambuja, A.; Sierra, F. Large Scale Seismic Processing in Public Cloud. In Proceedings of the 82nd EAGE Annual Conference & Exhibition, Amsterdam, The Netherlands, 8–11 June 2020; Volume 2020, pp. 1–5.
13. Rao, M.V. Data duplication using Amazon Web Services cloud storage. In *Data Deduplication Approaches: Concepts, Strategies, and Challenges*; Academic Press: Cambridge, MA, USA, 2020; p. 319.
14. Chakraborty, M.; Kundan, A.P. Grafana. In *Monitoring Cloud-Native Applications: Lead Agile Operations Confidently Using Open Source Software*; Springer: Berlin/Heidelberg, Germany, 2021.
15. Haan, L.; Ferreira, A. *Extreme Value Theory: An Introduction*; Springer: Berlin/Heidelberg, Germany, 2006; Volume 3.
16. Reghenzani, F.; Massari, G.; Fornaciari, W. Probabilistic-WCET reliability: Statistical testing of EVT hypotheses. *Microprocess. Microsyst.* **2020**, *77*, 103135. [CrossRef]
17. Omar, C.; Mundia, S.; Ngina, I. Forecasting value-at-risk of financial markets under the global pandemic of COVID-19 using conditional extreme value theory. *J. Math. Financ.* **2020**, *10*, 569–597. [CrossRef]
18. Embrechts, P.; Klüppelberg, C.; Mikosch, T. In *Modelling Extremal Events: For Insurance and Finance*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2013; Volume 33.
19. Coles, S.; Bawa, J.; Trenner, L.; Dorazio, P. In *An Introduction to Statistical Modeling of Extreme Values*; Springer: Berlin/Heidelberg, Germany, 2001; Volume 208.
20. Wang, C.; Xu, C.; Xia, J.; Qian, Z.; Lu, L. A combined use of microscopic traffic simulation and extreme value methods for traffic safety evaluation. *Transp. Res. Part C Emerg. Technol.* **2018**, *90*, 281–291. [CrossRef]
21. Beirlant, J.; Goegebeur, Y.; Segers, J.; Teugels, J.L. In *Statistics of Extremes: Theory and Applications*; John Wiley & Sons: Hoboken, NJ, USA, 2006.
22. Ouellette, P.; El-Jabi, N.; Rousselle, J. Application of extreme value theory to flood damage. *J. Water Resour. Plan. Manag.* **1985**, *111*, 467–477. [CrossRef]
23. Merz, B.; Basso, S.; Fischer, S.; Lun, D.; Blöschl, G.; Merz, R.; Guse, B.; Viglione, A.; Vorogushyn, S.; Macdonald, E.; et al. Understanding heavy tails of flood peak distributions. *Water Resour. Res.* **2022**, *58*, e2021WR030506. [CrossRef]
24. Tabari, H. Extreme value analysis dilemma for climate change impact assessment on global flood and extreme precipitation. *J. Hydrol.* **2021**, *593*, 125932. [CrossRef]
25. Haskins, K.; Wofford, Q.; Bridges, P.G. Workflows for performance predictable and reproducible hpc applications. In Proceedings of the 2019 IEEE International Conference on Cluster Computing (CLUSTER), Albuquerque, NM, USA, 23–26 September 2019; pp. 1–2.
26. Mondragon, O.H.; Bridges, P.G.; Levy, S.; Ferreira, K.B.; Widener, P. Understanding performance interference in next-generation HPC systems. In Proceedings of the SC'16: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, Salt Lake City, UT, USA, 13–18 November 2016; pp. 384–395.
27. Seelam, S.; Fong, L.; Tantawi, A.; Lewars, J.; Divirgilio, J.; Gildea, K. Extreme scale computing: Modeling the impact of system noise in multicore clustered systems. In Proceedings of the 2010 IEEE International Symposium on Parallel & Distributed Processing (IPDPS), Atlanta, GA, USA, 19–23 April 2010; pp. 1–12.
28. Fisher, R.A.; Tippet, L.H.C. Limiting forms of the frequency distribution of the largest or smallest member of a sample. In *Proceedings of the Mathematical Proceedings of the Cambridge Philosophical Society*; Cambridge University Press: Cambridge, UK, 1928; Volume 24, pp. 180–190.
29. Gnedenko, B. Sur La Distribution Limite Du Terme Maximum D'Une Série Aléatoire. *Ann. Math.* **1943**, *44*, 423–453. [CrossRef]
30. Jenkinson, A.F. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Q. J. R. Meteorol. Soc.* **1955**, *81*, 158–171. [CrossRef]
31. Markose, S.; Alentorn, A. The generalized extreme value distribution, implied tail index, and option pricing. *J. Deriv.* **2011**, *18*, 35–60. [CrossRef]
32. Lu, L.H.; Stedinger, J.R. Variance of two-and three-parameter GEV/PWM quantile estimators: Formulae, confidence intervals, and a comparison. *J. Hydrol.* **1992**, *138*, 247–267. [CrossRef]
33. Hirose, H. Maximum likelihood estimation in the 3-parameter Weibull distribution. A look through the generalized extreme-value distribution. *IEEE Trans. Dielectr. Electr. Insul.* **1996**, *3*, 43–55. [CrossRef]
34. Hosking, J.R. L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *J. R. Stat. Soc. Ser. B (Methodol.)* **1990**, *52*, 105–124. [CrossRef]
35. Smith, R.L. Extreme value theory based on the r largest annual events. *J. Hydrol.* **1986**, *86*, 27–43. [CrossRef]
36. McNeil, A.J. *Calculating Quantile Risk Measures for Financial Return Series Using Extreme Value Theory*; Technical Report; ETH Zurich: Zurich, Switzerland, 1998.

37. Mehta, N.J.; Yang, F. Portfolio optimization for extreme risks with maximum diversification: An empirical analysis. *Risks* **2022**, *10*, 101. [\[CrossRef\]](#)
38. Gu, X.; Ye, L.; Xin, Q.; Zhang, C.; Zeng, F.; Nerantzaki, S.D.; Papalexiou, S.M. Extreme precipitation in China: A review on statistical methods and applications. *Adv. Water Resour.* **2022**, *163*, 104144. [\[CrossRef\]](#)
39. Beretta, S. More than 25 years of extreme value statistics for defects: Fundamentals, historical developments, recent applications. *Int. J. Fatigue* **2021**, *151*, 106407. [\[CrossRef\]](#)
40. Cai, Y.; Hames, D. Minimum sample size determination for generalized extreme value distribution. *Commun. Stat. Comput.* **2010**, *40*, 87–98. [\[CrossRef\]](#)
41. Henwood, R.; Watkins, N.W.; Chapman, S.C.; McLay, R. A parallel workload has extreme variability in a production environment. *arXiv* **2018**, arXiv:1801.03898.
42. Duplyakin, D.; Ricci, R.; Maricq, A.; Wong, G.; Duerig, J.; Eide, E.; Stoller, L.; Hibler, M.; Johnson, D.; Webb, K.; et al. The Design and Operation of CloudLab. In Proceedings of the 2019 USENIX Annual Technical Conference (ATC 2019), Renton, WA, USA, 10–12 July 2019; pp. 1–14.
43. Fragalla, J. Configure, Tune, and Benchmark a Lustre Filesystem. In *2014 Oil & Gas HPC Workshop*; 2014. Available online: http://rice2014oghpc.blogs.rice.edu/files/2014/03/Fragalla-Xyratex_Lustre_PerformanceTuning_Fragalla_0314.pdf (accessed on 1 April 2024).
44. NORCOTT. Iozone Filesystem Benchmark. 2003. Available online: <http://www.iozone.org/> (accessed on 1 April 2024).
45. Conway, A.; Bakshi, A.; Jiao, Y.; Jannen, W.; Zhan, Y.; Yuan, J.; Bender, M.A.; Johnson, R.; Kuszmaul, B.C.; Porter, D.E.; et al. File systems fated for senescence? nonsense, says science! In Proceedings of the 15th USENIX Conference on File and Storage Technologies (FAST 17), Santa Clara, CA, USA, 27 February–2 March 2017; pp. 45–58.
46. Yu, W.; Vetter, J.; Canon, R.S.; Jiang, S. Exploiting lustre file joining for effective collective io. In Proceedings of the Seventh IEEE International Symposium on Cluster Computing and the Grid (CCGrid'07), Rio de Janeiro, Brazil, 14–17 May 2007; pp. 267–274.
47. Wong, P.; Der Wijngaart, R. *NAS Parallel Benchmarks I/O*, Version 2.4; Tech. Rep. NAS-03-002; NASA Ames Research Center: Moffet Field, CA, USA, 2003.
48. Oracle. Lustre 1.6 Operations Manual. 2010. Available online: <https://docs.oracle.com/cd/E19091-01/lustre.fs16/820-3681-11/820-3681-11.pdf> (accessed on 6 April 2024).
49. Amaral, J.N. About Computing Science Research Methodology. 2011. Available online: <https://webdocs.cs.ualberta.ca/~amaral/courses/MetodosDePesquisa/papers/Amaral-research-methods.pdf> (accessed on 2 April 2024).
50. Huang, H.H.; Li, S.; Szalay, A.; Terzis, A. Performance modeling and analysis of flash-based storage devices. In Proceedings of the 2011 IEEE 27th Symposium on Mass Storage Systems and Technologies (MSST), Denver, CO, USA, 23–27 May 2011; pp. 1–11.
51. Dominguez-Trujillo, J.; Haskins, K.; Khouzani, S.J.; Leap, C.; Tashakkori, S.; Wofford, Q.; Estrada, T.; Bridges, P.G.; Widener, P.M. Lightweight Measurement and Analysis of HPC Performance Variability. In *Proceedings of the 2020 IEEE/ACM Performance Modeling, Benchmarking and Simulation of High Performance Computer Systems (PMBS)*; IEEE: New York, NY, USA, 2020; pp. 50–60.
52. Lima, G.; Dias, D.; Barros, E. Extreme value theory for estimating task execution time bounds: A careful look. In Proceedings of the 2016 28th Euromicro Conference on Real-Time Systems (ECRTS), Toulouse, France, 5–8 July 2016; pp. 200–211.
53. Berezovskyi, K.; Santinelli, L.; Bletsas, K.; Tovar, E. WCET measurement-based and extreme value theory characterisation of CUDA kernels. In Proceedings of the Proceedings of the 22nd International Conference on Real-Time Networks and Systems, Versailles, France, 8–10 October 2014; pp. 279–288.
54. Castillo, J.D.; Padilla, M.; Abella, J.; Cazorla, F.J. Execution time distributions in embedded safety-critical systems using extreme value theory. *Int. J. Data Anal. Tech. Strateg.* **2017**, *9*, 348–361. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Ultrashort Echo Time and Fast Field Echo Imaging for Spine Bone Imaging with Application in Spondylolysis Evaluation

Diana Vucevic ¹ , Vadim Malis ¹, Yuichi Yamashita ², Anya Mesa ¹, Tomosuke Yamaguchi ¹, Suraj Achar ³, Mitsue Miyazaki ¹  and Won C. Bae ^{1,4,*} 

¹ Department of Radiology, University of California-San Diego, San Diego, CA 92103, USA

² Canon Medical Systems Corporation, Otawara 324-0036, Tochigi, Japan

³ Department of Family Medicine, University of California-San Diego, San Diego, CA 92103, USA

⁴ Department of Radiology, VA San Diego Healthcare System, San Diego, CA 92161, USA

* Correspondence: wbae@health.ucsd.edu

Abstract: Isthmic spondylolysis is characterized by a stress injury to the pars interarticularis bones of the lumbar spines and is often missed by conventional magnetic resonance imaging (MRI), necessitating a computed tomography (CT) for accurate diagnosis. We compare MRI techniques suitable for producing CT-like images. Lumbar spines of asymptomatic and low back pain (LBP) subjects were imaged at 3-Tesla with multi-echo ultrashort echo time (UTE) and field echo (FE) sequences followed by simple post-processing of averaging and inverting to depict spinal bones with a CT-like appearance. The contrast-to-noise ratio (CNR) for bone was determined to compare UTE vs. FE and single-echo vs. multi-echo data. Visually, both sequences depicted cortical bone with good contrast; UTE-processed sequences provided a flatter contrast for soft tissues that made them easy to distinguish from bone, while FE-processed images had better resolution and bone–muscle contrast, which are important for fracture detection. Additionally, multi-echo images provided significantly ($p = 0.03$) greater CNR compared with single-echo images. Using these techniques, progressive spondylolysis was detected in an LBP subject. This study demonstrates the feasibility of using spine bone MRI to yield CT-like contrast. Through the employment of multi-echo UTE and FE sequences combined with simple processing, we observe sufficient enhancements in image quality and contrast to detect pars fractures.

Keywords: MRI; low back pain; bone fracture; pars interarticularis



Citation: Vucevic, D.; Malis, V.; Yamashita, Y.; Mesa, A.; Yamaguchi, T.; Achar, S.; Miyazaki, M.; Bae, W.C. Ultrashort Echo Time and Fast Field Echo Imaging for Spine Bone Imaging with Application in Spondylolysis Evaluation. *Computation* **2024**, *12*, 152. <https://doi.org/10.3390/computation12080152>

Academic Editor: Anando Sen

Received: 8 June 2024

Revised: 15 July 2024

Accepted: 22 July 2024

Published: 24 July 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Low back pain (LBP) affects a significant proportion of young athletes, with up to 36% reporting it annually [1]. Isthmic spondylolysis [2] is characterized by a stress injury to the pars interarticularis of the vertebral arch of the lumbar spine [3,4]. It is a prevalent condition among young athletes engaged in activities that place repetitive stress on the lumbar spine. While the exact incidence of isthmic spondylolysis is not well documented, studies have reported it as a frequent cause of low back pain in young athletes, with prevalence rates ranging from 20% to 50% [5]. These findings highlight the need for accurate, radiation-free diagnostic tools for young individuals.

Isthmic spondylolysis presents with varying severity, necessitating different imaging techniques for its accurate assessment. In the initial stage of stress injuries to the pars interarticularis characterized by bone marrow edema without fracture, magnetic resonance imaging (MRI) is the most sensitive tool for detecting bone stress reactions [6,7]. As the condition progresses, both bone stress reactions and fractures may be present, while the terminal stage involves only fractures. Computed tomography (CT) is superior for visualizing bone fractures. This variability in presentation makes it challenging for a single imaging modality to fully characterize the stage of isthmic spondylolysis. Both MRI and

CT play crucial roles in the evaluation, and frequently, both modalities are necessary for a definitive diagnosis.

Due to concerns about ionizing radiation exposure in young patients, particularly those of reproductive age, MRI is the preferred initial imaging modality for detecting isthmic spondylolysis. However, due to the limited ability to visualize bone detail, conventional MRI frequently fails to identify a significant number of bone fractures that are later detected by CT scans [6–8]. A follow-up CT scan, providing excellent contrast and visualization for fine bone details [9], is then needed to ascertain the diagnosis for bone fracture at the pars. While multiple CT scans are detrimental for the long-term health of adolescent subjects, increasing the risk of cancer [10–12], it can be unavoidable in those who sustain recurrent low back injury, known to occur in as many as 33% of sport back injury patients within a year [13]. Reduction or avoidance of radiation would be highly beneficial for the imaging of young athletes suffering from low back pain.

Newer MRI techniques are gradually being adapted for bone imaging, with the development of sequences like ultrashort echo time (UTE) [14] and zero echo time (ZTE) [15]. These techniques have shown potential in visualizing the bone structure in several joints, including the knee [15,16], shoulder [17], and lumbar spine [18]. Nonetheless, they are hampered by inherent limitations, such as a low signal-to-noise ratio (SNR) and lower resolution, compared with CT. Thus, the continued study and development of bone-oriented MR techniques is warranted.

In this study, our objective was to introduce and compare two MRI techniques for imaging bone—multi-echo UTE and multi-echo field echo (FE) acquisitions—combined with simple post-processing to enhance appearance, SNR, and image contrast, with the aim of providing CT-like image characteristics. The use of multi-echo images to improve SNR has been proposed previously [19,20], and it will likely benefit UTE imaging for depicting cortical bone. FE sequences have also been suggested for bone imaging [21], with notable strengths including the wide availability of the sequence from major MRI vendors, high spatial resolution, and high contrast between muscle and bone, which are useful for bone visualization. In this study, we combined UTE and FE acquisitions with simple post-processing techniques to provide a CT-like appearance and to improve image quality.

2. Materials and Methods

2.1. Subjects

Lumbar spines of four subjects (1 male, 3 females; 19.5 ± 5.5 years old) were imaged at 3-T (Canon Galan) with a posterior spine coil. Two health volunteers were recruited with the inclusion criteria of no low back pain or history of low back surgery. Two of the subjects were adolescent (16 yo female, 14 yo male) athletes with non-specific and non-radiating low back pain for the past 4–8 weeks. The symptoms did not improve with conservative treatment, and the patients' primary care physician referred them to be evaluated for the presence of spondylolysis.

2.2. MRI

The subjects were imaged with a 3-Tesla MRI scanner (Galan, Canon Medical Systems Corp., Otawara, Japan) fitted with a posterior receive-only spine coil. One anatomical sequence and two bone-oriented sequences were used to image each subject. The anatomical sequence was a sagittal T2-weighted fast spin echo (FSE T2) with fat suppression with the following scanning parameters: repetition time (TR) = 5600 milliseconds (ms), echo time (TE) = 80 ms, field of view (FOV) = 220 millimeters (mm), image matrix = 320×416 , slice thickness = 3 mm, and echo train length (ETL) = 19. The bone-oriented sequences were (1) an axial 3D UTE multi-echo sequence with scan parameters of TR = 16.7 ms; TE = 0.1, 2.7, 5.3 ms; number of projections = 19,968; FOV = 300 mm; matrix = 320×320 ; and slice = 1 mm, reformatted to the sagittal plane and (2) a sagittal 3D field echo (FE) multi-echo sequence with scan parameters of TR = 21.8 ms; TE = 4, 8.6, 13.2 ms; FOV = 170 mm; matrix = 320×240 ; and slice = 1 mm. The source images are shown in Figure 1.

While no formal optimization was performed, we initially tried different numbers of radial projections for the 3D UTE, and different TE values for the 3D FE.

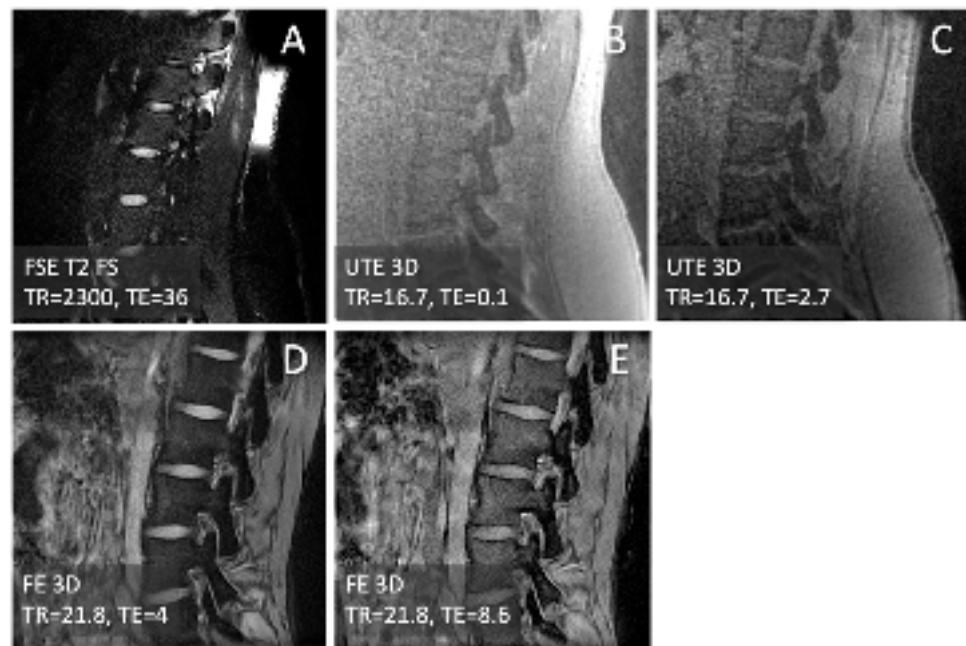


Figure 1. Raw MRI images of a lumbar spine acquired by (A) T2-weighted fat-suppressed fast spin echo (FSE T2 FS), (B,C) 3D ultrashort echo time (UTE) at varying echo times (TEs), and (D,E) 3D field echo (FE) at varying TEs.

2.3. CT-like Image Processing

To create a CT-like appearance from the bone-oriented MRIs, and to demonstrate the advantage of using multiple images from multi-echo acquisition, the images were processed directly on the scanner console in the following 2 ways. First, a set of CT-like images were created from the first-echo images from the UTE (TE = 0.1 ms, Figure 1B) and FE (TE = 4 ms; Figure 1D) were simply inverted (Equation (1)) to make the bone tissues bright and the surrounding tissues dark (Figure 2A,C).

$$\text{First Echo CT-like Image} = \frac{1}{\text{Source Image}} \quad (1)$$

A second set of CT-like images were created from all images (with varying TEs) in the multi-echo series. For each sequence of the UTE and the FE, all of the source images from different TEs were averaged and then inverted to yield the final CT-like images (Figure 2B,D) following Equation (2), where i = the source images corresponding to each TE and n = the number of TEs.

$$\text{Multi-Echo CT-like Image} = \frac{n}{\sum_{i=1}^n \text{Source Image}_i} \quad (2)$$

The intention was that the averaging would reduce the noise, and the inversion would provide high signal intensity to the bone. Representative processed images from both approaches using the UTE and the FE are shown in Figure 2.

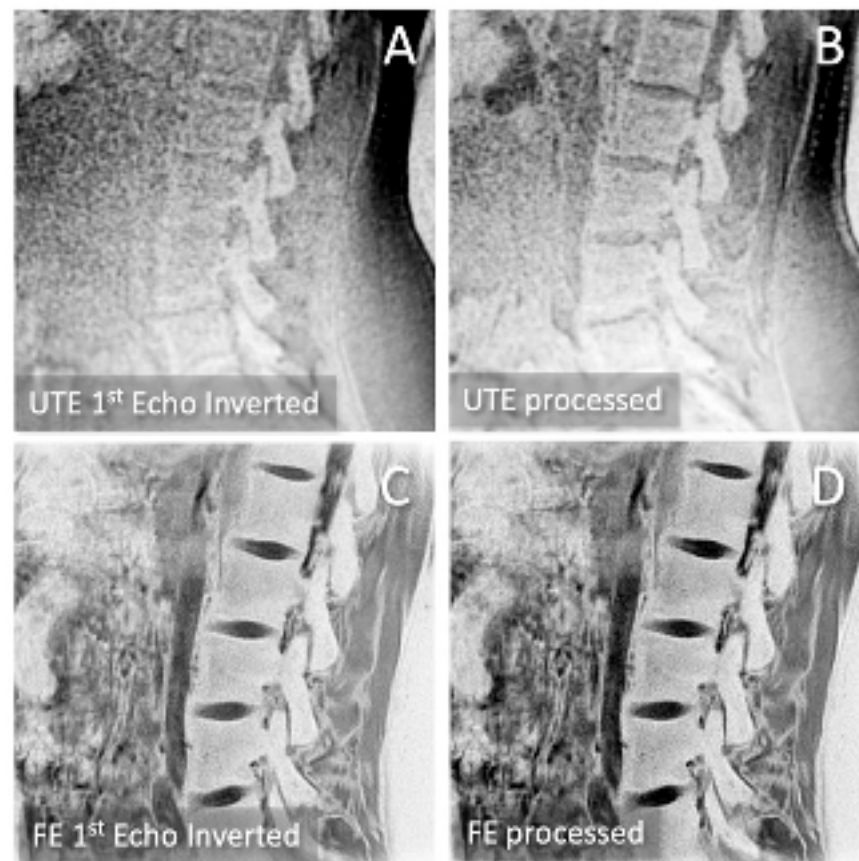


Figure 2. Processed images to create CT-like contrast from (A) the UTE 1st-echo image, (B) UTE multi-echo images, (C) the FE 1st-echo image, and (D) FE multi-echo images. Differences between sequences (UTE vs. FE) and improvements in contrast and image quality from multi-echo processing are apparent.

2.4. Signal-to-Noise Ratio (SNR) and Contrast-to-Noise Ratio (CNR)

Regions of interest (ROI) were drawn on a sagittal slice of the processed CT-like images that depicted cortical bone of the pars interarticularis and the surrounding paraspinal muscles (Figure 3). The ROIs were placed manually by a scientist with over 10 years of experience in imaging research. ROIs for the bone were placed along the pars interarticularis, and ROIs for the muscle were placed in the adjacent paraspinal muscle posterior to the bone ROI. Within each ROI, we determined the mean signal intensity (SI_{mean}). An ROI was also placed in the background (air) to determine the standard deviation (SI_{SD}) of the noise signal intensity. We evaluated slice-to-slice variability by the coefficient of variation in the mean signal intensity of bone ROIs in 5 consecutive slices of two datasets and found less than 2% variability, regardless of the sequence used.

The SNR was determined from the mean signal intensity divided by the standard deviation of the noise in the air [22] (Equation (3)).

$$SNR^1 = \frac{SI_{mean}^1}{SI_{SD}^{noise}} \quad (3)$$

The CNR was determined from the difference in mean signal intensity between selected ROIs divided by the standard deviation of the noise [23] (Equation (4)).

$$CNR^{1-2} = \frac{SI_{mean}^1 - SI_{mean}^2}{SI_{SD}^{noise}} \quad (4)$$

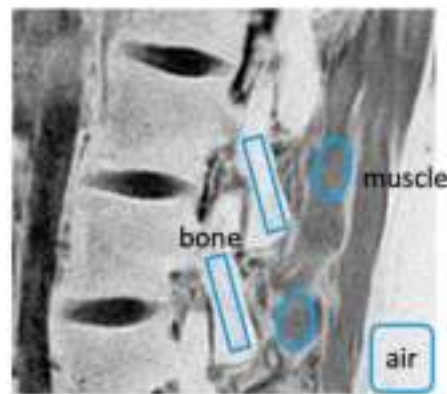


Figure 3. Regions of interest (blue-enclosed areas), including bone of the pars interarticularis, paraspinal muscles, and air, which were analyzed to determine the SNR and CNR.

2.5. Statistics

SNR and CNR values between the four types of CT-like images (processed from UTE 1st-echo, UTE multi-echo, FE 1st-echo, and FE multi-echo images) were compared using fully-factorial two-way ANOVA [24,25] to determine the effects of the sequence (UTE vs. FE) and the use of 1st-echo image vs. all three images from the multi-echo processing. The Systat statistical analysis software (v12, Grafiti LLC, Palo Alto, CA, USA) was used. The significance level was set at 5%.

2.6. Spondylolysis Evaluation

For the two subjects with low back pain, we compared the conventional SE T2, multi-echo-processed UTE, and multi-echo-processed FE images to evaluate the presence of spondylolysis (Figure 4).

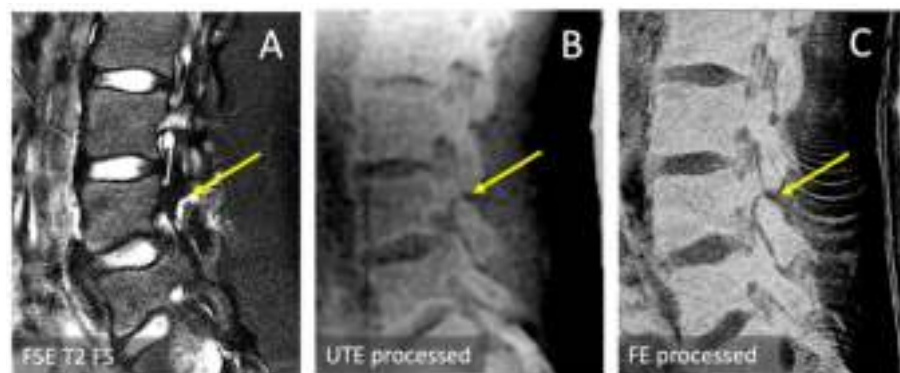


Figure 4. Detection of a moderately sized pars defect (arrows) in an adolescent athlete with persistent low back pain using FSE T2 FS (A), UTE multi-echo processing (B), and FE multi-echo processing (C). The processed images may enable an easier and more confident diagnosis of isthmic spondylolysis.

3. Results

3.1. Observations

Compared with FSE T2 (Figure 1A), CT-like processed images from first-echo (Figure 2A,C) and multi-echo (Figure 2B,D) UTE and FE images all depicted spinal bone distinctly with high signal intensity, similar to CT. UTE (Figure 2A,B) and FE (Figure 2B,C) images, while both providing CT-like contrast, had notable differences. UTE images were softer, while providing a flatter contrast for the soft tissue, making it easier to distinguish the bone if the reader is unfamiliar with the anatomy. In contrast, FE images were markedly sharper and depicted bone with a higher contrast. But FE images depicted many non-bone tissues (e.g., fascia of muscle) with a similarly high signal intensity as bone. This makes FE imaging less desirable for tasks such as the automatic segmentation and visualization

of bone but more desirable for evaluating small features such as hairline fractures. When comparing CT-like images from first-echo (Figure 2A,C) vs. multi-echo (Figure 2B,D), the multi-echo images showed an improvement, with less noise and better contrast for bone overall.

3.2. SNR and CNR

Table 1 summarizes the SNR and CNR values for the CT-like processed images. CT-like images from the UTE data had mean SNR values for the bone ranging from 90 to 100, which were significantly higher ($p = 0.0002$) than the values from the FE data (mean SNR range of 38 to 53). However, the CNR (bone–muscle) values were slightly higher ($p = 0.09$) for FE (mean CNR ranging from 15 to 22) compared with UTE (CNR of 13 to 16). The use of multi-echo images, for both UTE and FE data improved the CNR values significantly ($p = 0.03$): for UTE, the CNR on average improved by 2.7, while for FE, the CNR improved by 6.9. This corroborates the visual improvements in bone contrast observed in the image comparison in Figure 2.

Table 1. (Top) Mean and standard deviation of SNR and CNR values from various regions of interest. (Bottom) p -values from two-way ANOVA to assess the effects of sequence (UTE vs. FE) and processing (first-echo vs. multi-echo) on the SNR and CNR. **Bold faced** p -value indicates statistical significance ($p < 0.05$).

Measurement	Mean (+/− Std. Dev.) Values for Each Sequence			
	UTE 1st Echo	UTE Multi	FE 1st Echo	FE Multi
Bone SNR	107 (66)	89.8 (29.5)	38.1 (10.3)	53.3 (16.3)
Muscle SNR	94.4 (65.4)	74.0 (32.8)	23.4 (7.0)	31.9 (14.2)
Bone–Muscle CNR	13.0 (5.2)	15.7 (5.0)	14.6 (4.2)	21.5 (8.5)
Measurement	Two-Way ANOVA: Effect of			
	UTE vs. FE	1 st - vs. Multi-echo	Interaction	
Bone SNR	0.0004	0.9289	0.2235	
Muscle SNR	0.0002	0.6556	0.2864	
Bone–Muscle CNR	0.0935	0.0311	0.3327	

3.3. Spondylolysis Depiction

In one of the low back pain subjects, we detected progressive spondylolysis with a moderate-sized bone defect seen as a gap with high signal intensity in FSE T2 (Figure 4A, arrow) and a clear non-union in both the multi-echo-processed UTE (Figure 4B, arrow) and multi-echo-processed FE (Figure 4C, arrow) images. We also measured the signal profile across the bone defect in UTE-processed and FE-processed images. The defect widths, measured as full width at half maximum, were 2.75 mm and 2.39 mm for UTE and FE, respectively.

4. Discussion

This study has demonstrated the feasibility of spine bone MRI to yield CT-like contrast. Through the employment of multi-echo UTE and FE sequences combined with straightforward image processing techniques, we have observed discernible enhancements in image quality and bone contrast compared with using single-echo approaches. Our findings are encouraging, suggesting that these approaches can be used to obtain CT-like images useful for evaluating spondylolysis, albeit with limitations.

Our study compared multi-echo UTE and multi-echo 3D FE sequences and demonstrated that, while both sequences can depict cortical bone reasonably well, UTE imaging provided a flatter contrast for soft tissues that made distinguishing bone (appearing dark) from surrounding tissues (appearing bright) easier, while FE images had greater spatial resolution and bone–muscle contrast than UTE, which will be invaluable for detecting finer

fractures. We also showed that processing all images from multi-echo acquisition provided significantly a greater CNR for bone for both UTE and FE acquisitions.

Our study builds upon past studies with a similar purpose of imaging bone using MRI. Many utilized a similar approach of acquiring images that depict bone with a signal void and surrounding tissue with higher signal intensity and inverting the image to provide a CT-like appearance [21,26,27]. These studies have demonstrated that these conventional sequences can indeed provide a good depiction of cortical bone structure, and in one study [27], high accuracy (>90%) in detecting bone fractures was demonstrated. Our result generally corroborated these past results and suggested the benefit of using multi-echo images for improving image quality. While we did not formally investigate the effect of different processing techniques, our simple averaging and inverting approach can be performed on any MRI console using built-in image processing tools. Various other processing techniques such as rescaled echo subtraction [28], simple inversion of a single echo [27], or summation of all but the last echo followed by the last echo [21] have been attempted with varying degrees of success. There are even more advanced approaches available, such as using deep learning to create CT-like (pseudo-CT) images from MRI [18,29,30], but these currently require extensive off-line processing that may delay clinical evaluation.

This is an early study with many limitations. While the feasibility has been shown, a comparative study using CT reference images is needed. Since it is difficult to acquire both CT and MRI images without a clinical need, an initial study could utilize cadaveric specimens, as performed previously [18]. Also, the number of subjects used was very small, but this is typical in early technical development. Nonetheless, a larger study will be needed in the future to confirm the present findings and to make the results generalizable for the intended application. We found that the differences in the SNR and the CNR between MR sequences were substantial, and the trend will likely be similar with a greater number of subjects, since the signal intensity of cortical bone and muscle may not vary markedly between subjects without a particular pathology. While it would have been ideal to involve radiologists to compare the image quality and the diagnostic performance for detecting pars fractures, that would have required quite a large number of subjects, as well as reference CT images. This is a likely future extension of the current work. As an early study, we were able to enroll two young athletes with persistent low back pain and positively showed spondylolysis in one of them in Figure 4, demonstrating the clinical feasibility.

Despite advances in pulse sequences and post-processing techniques, there is still no MRI-based method that can produce images equivalent to CT scans. In terms of techniques appropriate for bone fracture evaluation at the pars, both techniques discussed in this paper, along with past similar approaches based on conventional image processing [19,21,27], might provide value. The use of artificial intelligence to synthesize pseudo-CT images from MRI images [18,31,32], while yielding the best bone contrast (reaching CNRs of 100 to 150 [18], compared with about 20 in this study) and CT-like appearance, are not yet practical. Future work will explore the possibility of improving bone imaging by using more advanced image processing and/or deep learning techniques.

5. Conclusions

In conclusion, this study has demonstrated the feasibility of two MRI techniques to provide a radiation-free alternative for the evaluation of spondylolysis, with their clinical application shown on a young subject with a pars defect. Further advancements in MRI technology with multi-echo UTE and FE have the potential to transform the standard of care for adolescent athletes and others with isthmic spondylolysis, offering a safer, equally effective, and radiation-free imaging alternative.

Author Contributions: Conceptualization: W.C.B., S.A., V.M. and M.M.; methodology, W.C.B., V.M., Y.Y. and M.M.; software, W.C.B., V.M. and Y.Y.; validation, W.C.B. and S.A.; formal analysis, D.V. and W.C.B.; investigation, W.C.B., M.M., S.A., Y.Y., T.Y., A.M. and D.V.; resources, A.M., T.Y. and D.V.; data

curation, W.C.B.; writing—original draft preparation, W.C.B., S.A. and M.M.; writing—review and editing, V.M., A.M., T.Y., S.A., D.V. and W.C.B.; visualization, W.C.B.; supervision, M.M. and W.C.B.; project administration, A.M. and W.C.B.; funding acquisition, M.M. and W.C.B. All authors have read and agreed to the published version of the manuscript.

Funding: Won C. Bae received research funding from Canon Medical Systems, USA. Mitsue Miyazaki received research funding from Canon Medical Systems Corporation, Japan (35938). The funding agencies had no role in the study design, data collection, interpretation, or the decision to submit this work for publication.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of University of California, San Diego (protocol #200335, approval date 17 October 2023).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data that support the findings of this study are not publicly available due to reasons of sensitivity. Anonymized data may be available from the corresponding author upon a review of the request. Data are located in controlled-access data storage at the corresponding author's institution.

Conflicts of Interest: Yuichi Yamashita is an employee of Canon Medical Systems, Japan. The remaining authors declare no conflicts of interest.

References

- Olsen, T.L.; Anderson, R.L.; Dearwater, S.R.; Kriska, A.M.; Cauley, J.A.; Aaron, D.J.; LaPorte, R.E. The epidemiology of low back pain in an adolescent population. *Am. J. Public Health* **1992**, *82*, 606–608. [[CrossRef](#)] [[PubMed](#)]
- Wiltse, L.L.; Newman, P.H.; Macnab, I. Classification of spondylolysis and spondylolisthesis. *Clin. Orthop. Relat. Res.* **1976**, *117*, 23–29. [[CrossRef](#)]
- Fredrickson, B.E.; Baker, D.; McHolick, W.J.; Yuan, H.A.; Lubicky, J.P. The natural history of spondylolysis and spondylolisthesis. *J. Bone Jt. Surg. Am.* **1984**, *66*, 699–707. [[CrossRef](#)]
- Bechtel, W.; Griffiths, H.; Eisenstadt, R. The Pathogenesis of Spondylolysis. *Investig. Radiol.* **1982**, *17*, S29. [[CrossRef](#)]
- Li, J.; Liang, J.; Xu, Y.; Du, D.; Feng, F.; Shen, J.; Cui, Y. Incidence of lumbar spondylolysis in athletes with low back pain: A systematic evaluation and single-arm meta-analysis. *Medicine* **2023**, *102*, e34857. [[CrossRef](#)]
- West, A.M.; d'Hemecourt, P.A.; Bono, O.J.; Micheli, L.J.; Sugimoto, D. Diagnostic Accuracy of Magnetic Resonance Imaging and Computed Tomography Scan in Young Athletes With Spondylolysis. *Clin. Pediatr.* **2019**, *58*, 671–676. [[CrossRef](#)] [[PubMed](#)]
- Yamane, T.; Yoshida, T.; Mimatsu, K. Early diagnosis of lumbar spondylolysis by MRI. *J. Bone Jt. Surg. Br.* **1993**, *75*, 764–768. [[CrossRef](#)] [[PubMed](#)]
- Ulmer, J.L.; Mathews, V.P.; Elster, A.D.; Mark, L.P.; Daniels, D.L.; Mueller, W. MR imaging of lumbar spondylolysis: The importance of ancillary observations. *AJR Am. J. Roentgenol.* **1997**, *169*, 233–239. [[CrossRef](#)]
- Leone, A.; Cianfoni, A.; Cerase, A.; Magarelli, N.; Bonomo, L. Lumbar spondylolysis: A review. *Skelet. Radiol.* **2011**, *40*, 683–700. [[CrossRef](#)]
- Journy, N.; Roue, T.; Cardis, E.; Le Pointe, H.D.; Brisse, H.; Chateil, J.F.; Laurier, D.; Bernier, M.O. Childhood CT scans and cancer risk: Impact of predisposing factors for cancer on the risk estimates. *J. Radiol. Prot.* **2016**, *36*, N1–N7. [[CrossRef](#)]
- Berrington de Gonzalez, A.; Pasqual, E.; Veiga, L. Epidemiological studies of CT scans and cancer risk: The state of the science. *Br. J. Radiol.* **2021**, *94*, 20210471. [[CrossRef](#)] [[PubMed](#)]
- Wang, W.H.; Sung, C.Y.; Wang, S.C.; Shao, Y.J. Risks of leukemia, intracranial tumours and lymphomas in childhood and early adulthood after pediatric radiation exposure from computed tomography. *CMAJ* **2023**, *195*, E575–E583. [[CrossRef](#)]
- Carey, T.S.; Garrett, J.M.; Jackman, A.; Hadler, N. Recurrence and care seeking after acute back pain: Results of a long-term follow-up study. North Carolina Back Pain Project. *Med. Care* **1999**, *37*, 157–164. [[CrossRef](#)] [[PubMed](#)]
- Bae, W.C.; Biswas, R.; Chen, K.; Chang, E.Y.; Chung, C.B. UTE MRI of the Osteochondral Junction. *Curr. Radiol. Rep.* **2014**, *2*, 35. [[CrossRef](#)] [[PubMed](#)]
- Bharadwaj, U.U.; Coy, A.; Motamedi, D.; Sun, D.; Joseph, G.B.; Krug, R.; Link, T.M. CT-like MRI: A qualitative assessment of ZTE sequences for knee osseous abnormalities. *Skelet. Radiol.* **2022**, *51*, 1585–1594. [[CrossRef](#)] [[PubMed](#)]
- Cheng, K.Y.; Moazamian, D.; Ma, Y.; Jang, H.; Jerban, S.; Du, J.; Chung, C.B. Clinical application of ultrashort echo time (UTE) and zero echo time (ZTE) magnetic resonance (MR) imaging in the evaluation of osteoarthritis. *Skelet. Radiol.* **2023**, *52*, 2149–2157. [[CrossRef](#)] [[PubMed](#)]
- Breighner, R.E.; Endo, Y.; Konin, G.P.; Gulotta, L.V.; Koff, M.F.; Potter, H.G. Technical Developments: Zero Echo Time Imaging of the Shoulder: Enhanced Osseous Detail by Using MR Imaging. *Radiology* **2018**, *286*, 960–966. [[CrossRef](#)] [[PubMed](#)]
- Achar, S.; Hwang, D.; Finkenstaedt, T.; Malis, V.; Bae, W.C. Deep-Learning-Aided Evaluation of Spondylolysis Imaged with Ultrashort Echo Time Magnetic Resonance Imaging. *Sensors* **2023**, *23*, 8001. [[CrossRef](#)]

19. Jang, U.; Hwang, D. High-quality multiple T_2^* contrast MR images from low-quality multi-echo images using temporal-domain denoising methods. *Med. Phys.* **2012**, *39*, 468–474. [[CrossRef](#)]
20. Jang, U.; Nam, Y.; Kim, D.H.; Hwang, D. Improvement of the SNR and resolution of susceptibility-weighted venography by model-based multi-echo denoising. *Neuroimage* **2013**, *70*, 308–316. [[CrossRef](#)]
21. Johnson, B.; Alizai, H.; Dempsey, M. Fast field echo resembling a CT using restricted echo-spacing (FRACTURE): A novel MRI technique with superior bone contrast. *Skelet. Radiol.* **2021**, *50*, 1705–1713. [[CrossRef](#)] [[PubMed](#)]
22. Goerner, F.L.; Clarke, G.D. Measuring signal-to-noise ratio in partially parallel imaging MRI. *Med. Phys.* **2011**, *38*, 5049–5057. [[CrossRef](#)]
23. Edelstein, W.; Bottomley, P.; Hart, H.; Leue, W.; Schenck, J.; Redington, R. NMR imaging at 5.1 MHz: Work in progress. In *International Symposium on NMR Imaging*; Witcofski, R., Karstaedt, N., Partain, C., Eds.; Bowman Gray School of Medicine: Winston-Salem, NC, USA, 1982; pp. 139–145.
24. Mishra, P.; Singh, U.; Pandey, C.M.; Mishra, P.; Pandey, G. Application of student's t-test, analysis of variance, and covariance. *Ann. Card. Anaesth.* **2019**, *22*, 407–411. [[CrossRef](#)] [[PubMed](#)]
25. McHugh, M.L. Multiple comparison analysis testing in ANOVA. *Biochem. Med.* **2011**, *21*, 203–209. [[CrossRef](#)] [[PubMed](#)]
26. Tsuchiya, K.; Gomyo, M.; Katase, S.; Hiraoka, S.; Tateishi, H. Magnetic resonance bone imaging: Applications to vertebral lesions. *Jpn. J. Radiol.* **2023**, *41*, 1173–1185. [[CrossRef](#)] [[PubMed](#)]
27. Schwaiger, B.J.; Schneider, C.; Kronthaler, S.; Gassert, F.T.; Bohm, C.; Pfeiffer, D.; Baum, T.; Kirschke, J.S.; Karampinos, D.C.; Makowski, M.R.; et al. CT-like images based on T1 spoiled gradient-echo and ultra-short echo time MRI sequences for the assessment of vertebral fractures and degenerative bone changes of the spine. *Eur. Radiol.* **2021**, *31*, 4680–4689. [[CrossRef](#)] [[PubMed](#)]
28. Du, J.; Bydder, M.; Takahashi, A.M.; Carl, M.; Chung, C.B.; Bydder, G.M. Short T2 contrast with three-dimensional ultrashort echo time imaging. *Magn. Reson. Imaging* **2011**, *29*, 470–482. [[CrossRef](#)]
29. Largent, A.; Barateau, A.; Nunes, J.C.; Mylona, E.; Castelli, J.; Lafond, C.; Greer, P.B.; Dowling, J.A.; Baxter, J.; Saint-Jalmes, H.; et al. Comparison of Deep Learning-Based and Patch-Based Methods for Pseudo-CT Generation in MRI-Based Prostate Dose Planning. *Int. J. Radiat. Oncol. Biol. Phys.* **2019**, *105*, 1137–1150. [[CrossRef](#)] [[PubMed](#)]
30. Bourbonne, V.; Jaouen, V.; Hognon, C.; Boussion, N.; Lucia, F.; Pradier, O.; Bert, J.; Visvikis, D.; Schick, U. Dosimetric Validation of a GAN-Based Pseudo-CT Generation for MRI-Only Stereotactic Brain Radiotherapy. *Cancers* **2021**, *13*, 1082. [[CrossRef](#)]
31. Florkow, M.C.; Willemsen, K.; Mascarenhas, V.V.; Oei, E.H.G.; van Stralen, M.; Seevinck, P.R. Magnetic Resonance Imaging Versus Computed Tomography for Three-Dimensional Bone Imaging of Musculoskeletal Pathologies: A Review. *J. Magn. Reson. Imaging* **2022**, *56*, 11–34. [[CrossRef](#)]
32. Morbee, L.; Vereecke, E.; Laloo, F.; Chen, M.; Herregods, N.; Jans, L. MR Imaging of the Pelvic Bones: The Current and Cutting-Edge Techniques. *J. Belg. Soc. Radiol.* **2022**, *106*, 123. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Identification and Dynamics Understanding of Novel Inhibitors of Peptidase Domain of Collagenase G from *Clostridium histolyticum*

Farah Anjum ^{1,*} , Ali Hazazi ^{2,3}, Fouzeyyah Ali Alsaeedi ¹, Maha Bakhuraysah ¹, Alaa Shafie ¹, Norah Ali Alshehri ^{4,5}, Nahed Hawsawi ¹ , Amal Adnan Ashour ⁶ , Hamsa Jameel Banjar ¹ , Afaf Alharthi ¹  and Maryam Ishrat Niaz ⁷

- ¹ Department of Clinical Laboratory Sciences, College of Applied Medical Sciences, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia; fouzeyyah@tu.edu.sa (F.A.A.); maha1@tu.edu.sa (M.B.); a.shafie@tu.edu.sa (A.S.); hawsawi.nm@gmail.com (N.H.); h.banjer@tu.edu.sa (H.J.B.); a.awwadh@tu.edu.sa (A.A.)
 - ² Department of Pathology and Laboratory Medicine, Security Forces Hospital Program, P.O. Box 3643, Riyadh 11481, Saudi Arabia; ahazazi@sfh.med.sa
 - ³ College of Medicine, Alfaisal University, P.O. Box 50927, Riyadh 11533, Saudi Arabia
 - ⁴ Department of Family and Community Medicine, College of Medicine, King Saud University (KSU), Riyadh 11481, Saudi Arabia; drnora@ksu.edu.sa
 - ⁵ University Family Medicine Center, King Saud University Medical City, King Saud University (KSU), Riyadh 11481, Saudi Arabia
 - ⁶ Department of Oral and Maxillofacial Surgery and Diagnostic Sciences, Faculty of Dentistry, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia; a.a.ashour@tu.edu.sa
 - ⁷ Department of Basic Sciences for Nursing, Nursing College, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia; mimohammad@tu.edu.sa
- * Correspondence: farahanjum@tu.edu.sa



Citation: Anjum, F.; Hazazi, A.; Alsaeedi, F.A.; Bakhuraysah, M.; Shafie, A.; Alshehri, N.A.; Hawsawi, N.; Ashour, A.A.; Banjar, H.J.; Alharthi, A.; et al. Identification and Dynamics Understanding of Novel Inhibitors of Peptidase Domain of Collagenase G from *Clostridium histolyticum*. *Computation* **2024**, *12*, 153. <https://doi.org/10.3390/computation12080153>

Academic Editor: Claudio Amovilli

Received: 27 June 2024

Revised: 18 July 2024

Accepted: 21 July 2024

Published: 25 July 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: *Clostridium histolyticum* is a Gram-positive anaerobic bacterium belonging to the *Clostridium* genus. It produces collagenase, an enzyme involved in breaking down collagen which is a key component of connective tissues. However, antimicrobial resistance (AMR) poses a great challenge in combating infections caused by this bacteria. The lengthy nature of traditional drug development techniques has resulted in a shift to computer-aided drug design and other modern drug discovery approaches. The above method offers a cost-effective means for gathering comprehensive information about how ligands interact with their target proteins. The objective of this study is to create novel, explicit drugs that specifically inhibit the *C. histolyticum* collagenase enzyme. Through structure-based virtual screening, a library containing 1830 compounds was screened to identify potential drug candidates against collagenase enzymes. Following that, molecular dynamic (MD) simulation was performed in an aqueous solution to evaluate the behavior of protein and ligand in a dynamic environment while density functional theory (DFT) analysis was executed to predict the molecular properties and structure of lead compounds, and the WaterSwap technique was utilized to obtain insights into the drug–protein interaction with water molecules. Furthermore, principal component analysis (PCA) was performed to reveal conformational changes, salt bridges to express electrostatic interaction and protein stability, and absorption, distribution, metabolism, excretion, and toxicity (ADMET) to assess the pharmacokinetics profile of top compounds and control molecules. Three potent drug candidates were identified MSID000001, MSID000002, MSID000003, and the control with a binding score of -10.7 kcal/mol, -9.8 kcal/mol, -9.5 kcal/mol, and -8 kcal/mol, respectively. Furthermore, Molecular Mechanics Poisson–Boltzmann Surface Area (MMPBSA) analysis of the simulation trajectories revealed energy scores of -79.54 kcal/mol, -73.99 kcal/mol, -62.26 kcal/mol, and -70.66 kcal/mol, correspondingly. The pharmacokinetics properties exhibited were under the acceptable range. The compounds hold the potential to be novel drugs; therefore, further investigation needs to be conducted to find out their anti-collagenase action against *C. histolyticum* infections and antibiotic resistance.

Keywords: antimicrobial resistance; molecular dynamic simulation; principal component analysis; salt bridges; WaterSwap; structure-based drug design

1. Introduction

The upsurge in drug-resistant bacteria is a severe threat to public health and the worldwide economy, evidenced by the recent spike in bacterial fatalities [1]. Simultaneously, the slow pace by which new antibiotic discoveries occur worsens the problem and could ultimately render current minor disease fatalities [2]. To prevent this progression, prompt action must be taken, requiring the investigation of non-antibiotic alternative therapy approaches. Antivirulence drugs, commonly referred to as “pathoblockers”, stand out among them as a viable option. These drugs may effectively avoid or delay diseases by targeting and preventing the pathogenicity factors of bacteria [3].

The Gram-positive anaerobic bacterium *Clostridium histolyticum* belongs to the Clostridium genus, which comprises a broad category of bacteria characterized by having the ability to make an extensive number of toxins as well as enzymes [4]. *C. histolyticum* is a member of the class Clostridia and the phylum Firmicutes [5]. Mainly, *C. histolyticum* has drawn much focus in biomedical research since it has robust enzymatic activity along with associations to various kinds of medical conditions and diseases in humans. It may survive in a broad range of environments, including water, soil, and the guts of both humans and animals, because of its rod-shaped structure along with its ability to produce endospores amid challenging environmental conditions [6]. Peyronie’s disease and Dupuytren’s contracture are two types of diseases in which *C. histolyticum* collagenases have produced irregular collagen metabolism [7]. During these conditions, a high concentration of collagen deposits leads to tissue contractures and fibrosis, which limit function and cause pain [7].

Collagenases, hyaluronidases, proteases, and hemolysins represent a few of the virulence factors that *C. histolyticum* produces. These enzymes serve as vital in tissue invasion, host cell lysis, and pathogenesis [8]. *C. histolyticum* produces five distinct types of toxins. Of them, β -toxins are the primary cause of its pathogenesis whilst collagenases serve as crucial for clostridial infection by the virus [9]. Collagenases stand out among these virulence factors due to their unique ability to break down collagen, the key structural protein found in connective tissues [1]. Proteolytic enzymes called clostridial collagenases attack collagen molecules specifically. They induce the breakage of collagen fibers and the disruption of connective tissue by cutting peptide bonds inside the triple helical structure of collagen fibrils [10]. These enzymes exhibit a high degree of substrate specificity for numerous forms of protein collagen, which includes types I, II, III, and IV [11]. There are two types of collagenase enzymes, ColG and ColH [1]. Collagenase H (ColH) and Collagenase G (ColG) belong to the metalloprotease M9 family. Collagenases are considered to be key targets for the development of other antibiotic-resistance treatment approaches as they adversely affect the development of bacterial infections caused by *C. histolyticum* [9].

A study carried out by [12] utilized synthetic peptides and non-proteogenic amino acids aimed to find inhibitors against *Clostridium histolyticum* collagenases. The study aimed to find novel inhibitors against trypsin and collagenases. Through molecular docking, the interaction between non-proteinogenic amino acids and peptides with *C. histolyticum* collagenase and trypsin has been studied. Based on molecular docking analysis, N-tert-butoxycarbonylglycyl-(S)-2-amino-3-(4-((4-fluorophenyl)-ethynyl)-phenyl)- α -alanine exhibited promising interaction with collagenases. The result shows that dipeptide is a good inhibitor of collagenases.

Researchers have been able to develop novel drugs and vaccines that could decrease both the time and cost related to developing new drugs through advancements in bioinformatics and computational approaches [13]. Developing a completely novel drug via experimental approaches is a difficult task. The repurposing of drugs is an emerging practice that involves repurposing commonly used drugs that have previously been shown

to be safe in human trials for treating different diseases. However considering the high cost as well as the lengthy time frame, the approach is not sustainable [14]. The rapid rise of bacterial resistance compels researchers to emphasize on developing novel, potent drugs [15]. Considering the many downsides associated with traditional drug discovery, computer-aided drug design might be a more useful alternative for research due to its lower price and consumption of few resources [16]. A popular computational technique uses a structural drug design tool to virtually screen multiple drug libraries against a specified target. This could speed the drug design process by saving time in developing new leads to target the specific enzyme or protein [17]. Given this, multiple cheminformatics and biophysics techniques were employed during this study to discover potential lead drugs against *C. histolyticum* collagenases.

2. Methodology

Figure 1 depicts a schematic overview of the complete research work.

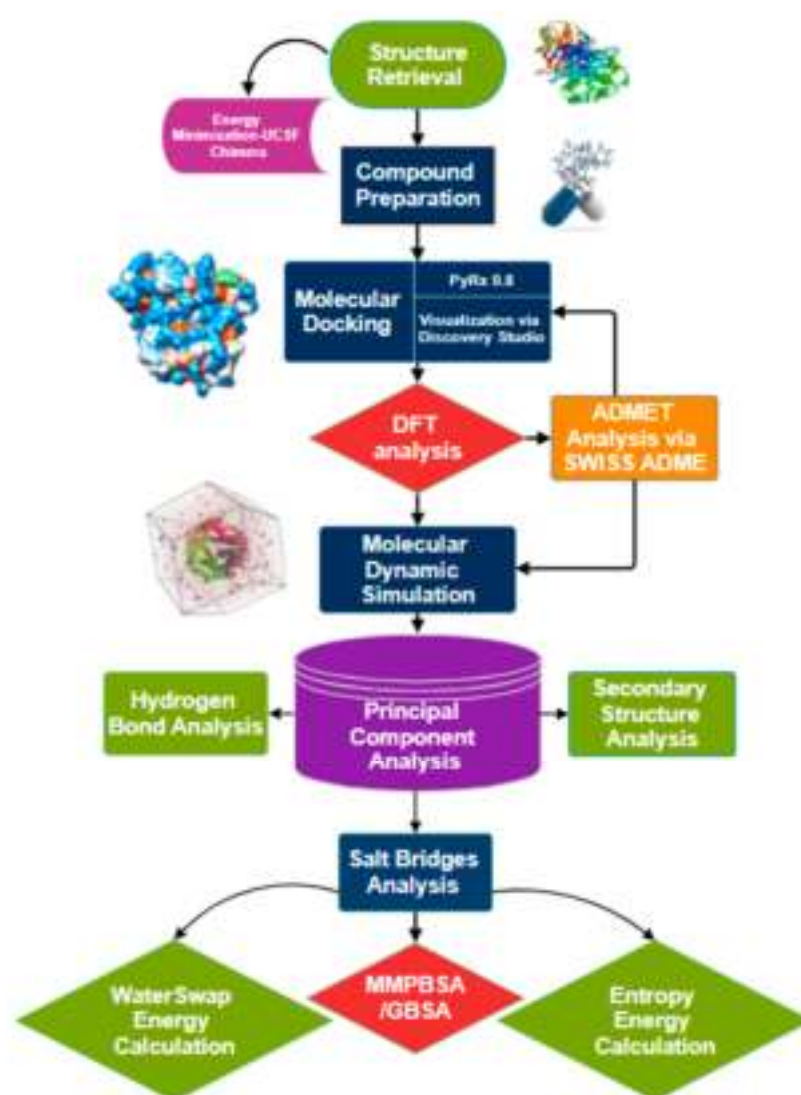


Figure 1. The integrated computational workflow begins with (i) structure retrieval and preparation followed by (ii) compound preparation and (iii) molecular docking. Next, DFT analysis to estimate electronic properties and structure (iv), ADMET to assess pharmacokinetics profile (v), and MD simulation, (vi) PCA analysis as well as (vii) secondary structure analysis, (viii) hydrogen bond analysis. (ix) Salt bridges, (x) MMPBSA, (xi) Entropy energy calculation, and finally (xii) water swap energy estimation.

2.1. Collagenase Enzyme Crystal Structure Retrieval and Preparation

The crystal structure of *C. histolyticum* collagenase was retrieved from the Protein Data Bank (PDB) https://www.rcsb.org/?ref=nav_home with PDB ID; 7ZBV. Subsequently, The structure was then loaded into UCSF Chimera v1.17 for early visualization, missing residue detection, and structural error detection [18]. To complete the structure, hydrogen atoms and charges were added via Chimaera v1.17 by using the “Add charge” and “AddH” functions [19]. Following that, two algorithms were applied to minimize energy: the conjugate gradient algorithm, run for 5000 steps, and the steepest descent algorithm, which ran for 5000 cycles with a step size of 0.02 Å [20]. These cyclic procedures aim to minimize steric issues in the structure of an enzyme. After minimizing energy, the enzyme’s improved structure was saved in a pdb format.

2.2. Ligands Library Selection and Preparation

To identify novel and potent inhibitors against the collagenase enzyme, which contains 1830 compounds, using several natural sources. The Medicinal Fungi Secondary metabolites And Therapeutics MeFSAT library was downloaded in SDF format [21]. The Library was then imported to PyRx 0.8 software for the energy minimization process [22]. The MM2 force field was applied to initially minimize the compound’s energy [23]. The compounds were converted to pdbqt format for structure-based virtual screening [24].

2.3. Molecular Docking Studies

Structure-based virtual screening is an effective technique for screening drug libraries against any target biomolecule [25]. Using Auto Dock in PyRx 0.8, a virtual screening of compounds targeting a specific protein was carried out [26]. After loading the target protein and chemical library into PyRx for further energy minimization, all ligands were converted to Auto Dock format. Furthermore, Auto Dock made it considerably quicker for the top three ligands to dock. Based on the binding energies score, the molecules were arranged. Those with the lowest energy score were chosen for stable binding conformation and following that, three-dimensional (3D) structural alignment, analysis, and visualization were carried out via Discovery Studio 2.4 [27].

2.4. Density Functional Theory (DFT)

All the studied structures were optimized using the B3LYP functional with GD3 correction and a 6–311+G (d, p) basis set using the DFT calculations method [28]. Geometry optimization was further used to check the stable conformation of studied structures along with chemical descriptors and molecular reactivity parameters, such as energy gap, chemical and ionization potential, frontier molecular orbital (FMO), electron affinity, electronegativity chemical hardness, softness, and electrophilic index was also estimated to find the nucleophilic and electrophilic nature, chemical reactivity, and stability [29]. Gaussian16 suite (<https://gaussian.com/gaussian16/>) was used for all the calculations. The GaussView6 Molecular Visualization Tool was employed to visualize and generate contour maps and molecular electrostatic potential (MEP) maps [30]. These maps could aid in the identification of possible areas that are engaged in chemical reactions or are crucial for drug–receptor interactions [31].

2.5. ADME and Pharmacokinetics Profile

ADME and the pharmacological activity of selected leads play an important role in the drug development toolbox [32]. The adsorption, distribution, metabolism, and excretion analysis profile was determined by using SwissADME <http://www.swissadme.ch/>. For toxicity prediction analysis of the compounds, a database called pkCSM <https://biosig.lab.uq.edu.au/> was utilized [33].

2.6. Molecular Dynamic Simulation

Since docking analysis can be inaccurate, it is necessary to perform additional validation, particularly through the MD simulation technique [34]. The initial stage in preparing protein and ligand complexes for MD simulation was docking analysis [35]. The binding site of a ligand into the protein's active pocket under static conditions was determined through molecular docking studies [36]. It was enabled to predict the ligand binding position during tough physiological conditions by using MD simulations [37]. The complexes were initially processed using AMBER22's Antechamber program. The receptor protein was created by the FF14Sb force field, whereas the compounds were made using the GAFF force field 5 [38]. The MD simulation was performed in three phases, which included the generation of the prmtop file, the preprocessing stage, and the production step [39]. After adding counter ions, the specifically chosen docked complexes were initially immersed in the TIP3P water box. Following that, each complex went through a 200 ns production run upon being gradually heated to 310 K and allowed for stabilization [40]. To study the structure-based screening of complexes 57, the CPPTRAJ was used. The simulation graph 60 has been created using the XMGRACE v5.1 [41].

2.7. Hydrogen Bond Analysis

It is important to count the number of H-bonds formed between the compounds and the residues in the active site of the enzyme after the simulation since this indicates how firmly the compounds bind [42]. In this research, we studied the formation of hydrogen bonds between the compounds and the active site of the enzyme residues through molecular dynamics simulations. The AMBER22 program was used for system setup, equilibration, and trajectory development. Subsequently AMBER's cpptraj module was applied to study H-bonds. In addition, an angle cutoff of 120° and a distance cutoff of 3.5 Å between hydrogen donor and acceptor atoms were employed to determine H-bonds [43]. Visualization of H-bonds was performed by using molecular visualization software as visual molecular dynamics (VMDs)v1.9.3 [44].

2.8. Calculating Binding Affinities

Getting insights about the binding free energy is vital for understanding how ligands and proteins interact [45]. The AMBER22 MM/PBSA techniques were used to calculate the binding free energies (G binding) of the complexes. Using the MMPBSA, the binding free energy between the compounds and the main collagenase enzyme was estimated [46]. The binding energies were calculated by using the following equation:

$$\begin{aligned} \text{EMM} &= \Delta E_{\text{int}} + \Delta E_{\text{ele}} + \Delta E_{\text{vdw}} \\ \Delta G_{\text{sol}} &= \Delta G_{\text{p}} + \Delta G_{\text{np}} \\ \Delta G_{\text{total}} &= \Delta \text{EMM} + \Delta G_{\text{sol}} \\ \Delta G_{\text{bind}} &= \Delta \text{EMM} + \Delta G_{\text{sol}} - T \end{aligned} \quad (1)$$

where ΔEMM stands for a total change of molecular mechanics energy, ΔE_{int} for a total change of internal interactions energy, ΔE_{ele} for a total change of electrostatic energy, ΔE_{vdw} for a total change of van der Waals energy, ΔG_{sol} for a total change of solvation energy, ΔG_{p} shows the total change of polar solvation energy, and ΔG_{np} for a total change of non-polar solvation energy.

2.9. Entropy Energy Calculation

A normal mode estimation method in the AMBER software v22 was used to estimate the entropy energy of each complex [47]. To ensure representative conformational sampling, ten frames were chosen from the MD simulation trajectories. These frames then underwent normal mode analysis to independently calculate the entropy energy for each complex.

2.10. WaterSwap Absolute Energy Estimation

Using an advanced WaterSwap technique, the absolute binding free energy of complexes was calculated to validate the formation of strong intermolecular complexes [48]. This technique is more precise and effective than MMPB/GBSA as it shows the vital function that water molecules play in promoting the interaction between ligands and the active site residues in a specific enzyme [49].

2.11. Secondary Structure Analysis

Secondary structure analysis was performed for the targeted proteins to identify the differences in secondary structure patterns [50]. The VMD module of AMBER22 was used for secondary structure analysis of the targeted enzyme.

2.12. Principal Component Analysis (PCA)

PCA, often referred to as Essential Dynamics (EDs), is a technique that simplifies the data and explains the protein's observable motional changes during the simulation period [51]. In the present study, PCA was attained by diagonalizing the covariance matrix constructed using the atoms C α of enzyme recoded in the SMT using the following equation:

$$C = \langle (q_i - \langle q_i \rangle)(q_j - \langle q_j \rangle) \rangle T \quad (2)$$

whereas $\langle q_i \rangle$ and $\langle q_j \rangle$ represent their mean positions over conformational groups derived from MIMD simulations, and q_i and q_j are the Cartesian coordinates of the i th and j th C α atoms in targeted enzyme, respectively. Collectively, the eigenvalue and eigenvector resulting from diagonalization indicate the coordinated motion of the structural domain and their fluctuating amplitude on an eigenvector, respectively. The CPPTRAJ program in Amber was used to carry out the PCA [52].

2.13. Salt Bridges

The AMBER22 software program was used to perform a computational evaluation of salt bridges [53]. Performing salt bridge analysis is vital for various reasons in computational drug design. Protein structures were stabilized and molecular interactions were influenced by salt bridges, which form between oppositely charged amino acid residues in a protein [45]. By computational analysis of salt bridges, researchers may learn more about the dynamics and stability of protein–ligand complexes, which aids in the identification of important interactions needed for drug binding [54].

3. Results

3.1. Structure Retrieval and Initial Preparation

The recent study was conducted to identify potent inhibitors against *Clostridium histolyticum* with the main collagenase enzyme. The structure was retrieved from PDB by using the PDB ID; 7ZBV. Collagenases have global Symmetry: Asymmetric—C1 and Global Stoichiometry: Monomer—A1. The X-ray diffraction method was used to determine the enzyme structure which is accessible at a resolution of 1.95 Å. The crystal structure was visualized via UCSF Chimera v.1.17. Figure 2 shows the overall three-dimensional structure (3D) collagenase structure depicting the alpha helix, catalytic residues, and beta strands. Additionally, Figure 2 also depicts the close-up view of the active sites involved in ligand binding.

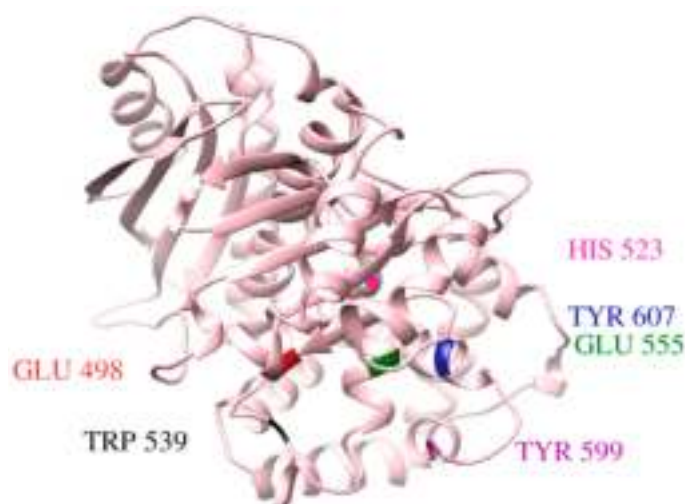


Figure 2. Highlights the labeled active site residues GLU498, TRP539, HIS523, TYR607, GLU555, and TYR599 in the three-dimensional (3D) structure of the collagenase enzyme.

3.2. Molecular Docking and Binding Interaction/Poses Analysis

Molecular docking provides a better understanding of protein and ligand interaction in the drug development process [32]. In this study, a library was retrieved in structure data file (SDF) format. The Medicinal Fungi Secondary Metabolites and Therapeutics MeFSAT library having 1830 compounds was chosen to identify the potential inhibitors against the *C. histolyticum* collagenase enzyme. Structure-based virtual screening was performed for 1830 compounds against the targeted enzyme by using the PyRx 0.8 AutoDock Vina tool. Based on the binding scores, 10 compounds that displayed strong interaction with the collagenase enzyme were selected as hits. Their structure and chemical names along with their binding scores are given in Table 1. To validate the docking results and evaluate their potential as inhibitors of collagenase enzyme, 3/10 compounds were selected for further computational analysis, i.e. molecular dynamic simulation. The docking analysis revealed that the top three compounds (MSID1) 2-(3-hydroxy-4,4,10,13,14-pentamethyl-2,3,4,5,6,7,10,11,12,13,14,15,16,17-tetradecahydro-1H-cyclopenta[a]phenanthren-17-yl)-6-methylhept-6-enoic acid, (MSID2) 4,4,10,13,14-pentamethyl-17-(6-methyl-5-methyleneheptan-2-yl)-2,3,4,5,6,7,10,11,12,13,14,15,16,17-tetradecahydro-1H-cyclopenta[a]phenanthren-3-ol, and (MSID3) (E)-17-(5,6-dimethylhept-3-en-2-yl)-10,13-dimethyl-2,3,4,5,6,9,10,11,12,13,14,15,16,17-tetradecahydro-1H-cyclopenta[a]phenanthrene-3,5,6-triol exhibited promising binding scores of -10.7 kcal/mol, -9.8 kcal/mol, and -9.5 kcal/mol, respectively, while the control showed a binding score of -8 with the target collagenase enzyme. The interaction and binding poses of the hits compounds are shown in Figure 3. Upon analysis of the binding poses and interaction, it was revealed that MSID1, with the highest binding score of -10.7 , formed conventional hydrogen bonds with Tyr496, and made van der Waals interaction with Glu498, Tyr599, Trp604, Ile497, Glu524, Leu495, Glu555, Gly493, and Asn492. Carbon–hydrogen bond interaction was shown by HIS527 while Alkyl and Pi-Alkyl interaction was seen with Trp39 and Ala531, respectively (Figure 4). MSID2 with the second-highest binding score of -9.8 revealed van der Waals interaction with Ile576, Glu559, and Arg573. His523, Glu524, Asn 492, Glu555, Trp539, Tyr599, Trp604, Glu519, and Gly 493 while Tyr607 pair Alky and PHE515 Pi-sigma interaction respectively (Figure 4). MSID3, having the third highest binding score -9.5 formed van der Waals interaction with Leu520, Tyr607, Glu524, Tyr599, His527, Ile497, Ala531, Glu498, Tyr496, Leu495, Asn492, Gly494, Phe515, and Trp539 pair Pi-Sigma interaction with the receptor protein (Figure 4). However, the control molecule was also seen to form conventional hydrogen bonds with His527, Glu498, and Tyr496, while Pro499, Ala531, Tyr528, Glu555, and Ile497 showed van der Waals interaction. Furthermore, Leu495 and Trp539 pair Pi-Alkyl and Pi-Pi stacked

interaction (Figure 4). The collagenase enzyme residues that were seen in interaction with the top 3 compounds are given in Table 2.

Table 1. Top 10 compounds with the highest binding affinities, chemical names coupled with their chemical structures.

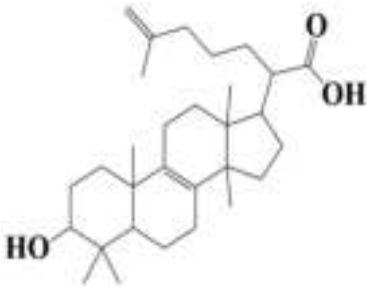
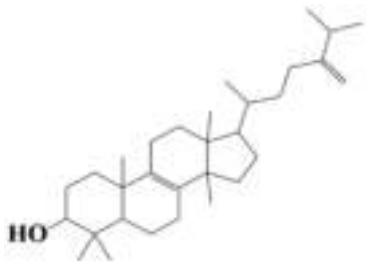
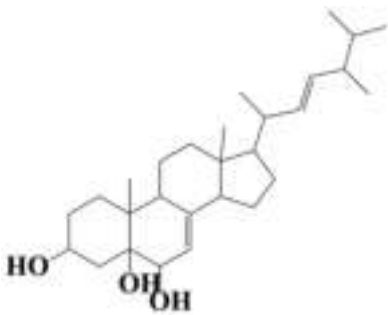
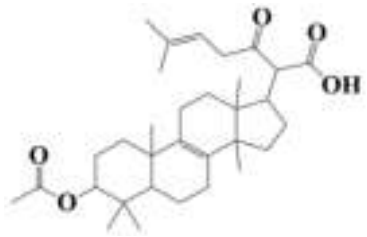
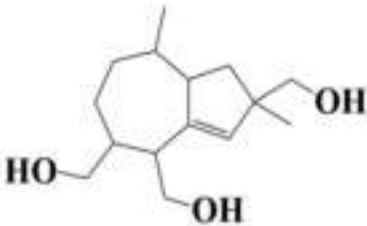
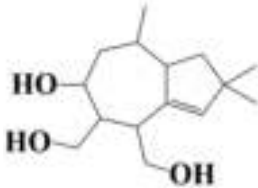
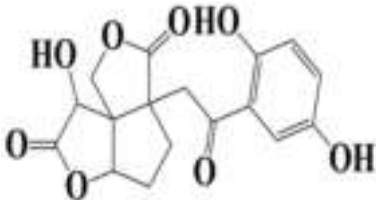
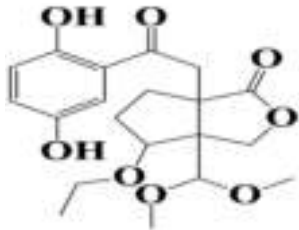
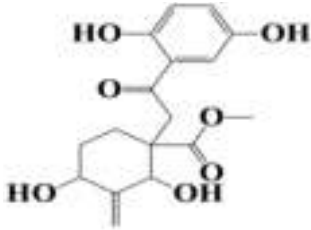
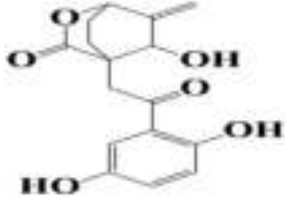
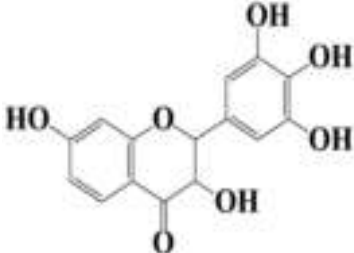
S.No	Compounds	Structure	Binding Affinity
1	MSID000001 2-(3-hydroxy-4,4,10,13,14-pentamethyl-2,3,4,5,6,7,10,11,12,13,14,15,16,17-tetradecahydro-1H-cyclopenta[a]phenanthren-17-yl)-6-methylhept-6-enoic acid		−10.7 kcal/mol
2	MSID000002 4,4,10,13,14-pentamethyl-17-(6-methyl-5-methyleneheptan-2-yl)-2,3,4,5,6,7,10,11,12,13,14,15,16,17-tetradecahydro-1H-cyclopenta[a]phenanthren-3-ol		−9.8 kcal/mol
3	MSID000003 (E)-17-(5,6-dimethylhept-3-en-2-yl)-10,13-dimethyl-2,3,4,5,6,9,10,11,12,13,14,15,16,17-tetradecahydro-1H-cyclopenta[a]phenanthrene-3,5,6-triol		−9.5 kcal/mol
4	MSID000004 2-(3-acetoxy-4,4,10,13,14-pentamethyl-2,3,4,5,6,7,10,11,12,13,14,15,16,17-tetradecahydro-1H-cyclopenta[a]phenanthren-17-yl)-6-methyl-3-oxohept-5-enoic acid		−9.3 kcal/mol
5	MSID000006 2,8-dimethyl-1,2,4,5,6,7,8,8a-octahydroazulene-2,4,5-triyl)trimethanol		−9.2 kcal/mol

Table 1. Cont.

S.No	Compounds	Structure	Binding Affinity
6	MSID000009 (6-hydroxy-2,2,8-trimethyl-1,2,4,5,6,7,8,8a-octahydroazulene-4,5-diyl)dimethanol		−9 kcal/mol
7	MSID000010 3a-(2-(2,5-dihydroxyphenyl)-2-oxoethyl)-8-hydroxytetrahydrocyclopenta[1,2-b:2,3-c']difuran-3,7(1H,8H)-dione		−8.9 kcal/mol
8	MSID000016 6a-(2-(2,5-dihydroxyphenyl)-2-oxoethyl)-3a-(dimethoxymethyl)-4-ethoxyhexahydro-1H-cyclopenta[c]furan-1-one		−8.6 kcal/mol
9	MSID000020 methyl 1-(2-(2,5-dihydroxyphenyl)-2-oxoethyl)-2,4-dihydroxy-3-methylenecyclohexanecarboxylate		−8.5 kcal/mol
10	MSID000022 4-(2-(2,5-dihydroxyphenyl)-2-oxoethyl)-5-hydroxy-6-methylene-2-oxabicyclo[2.2.2]octan-3-one		−8.4 kcal/mol
11	Control 3,7-dihydroxy-2-(3,4,5-trihydroxyphenyl)chroman-4-one		−8 kcal/mol

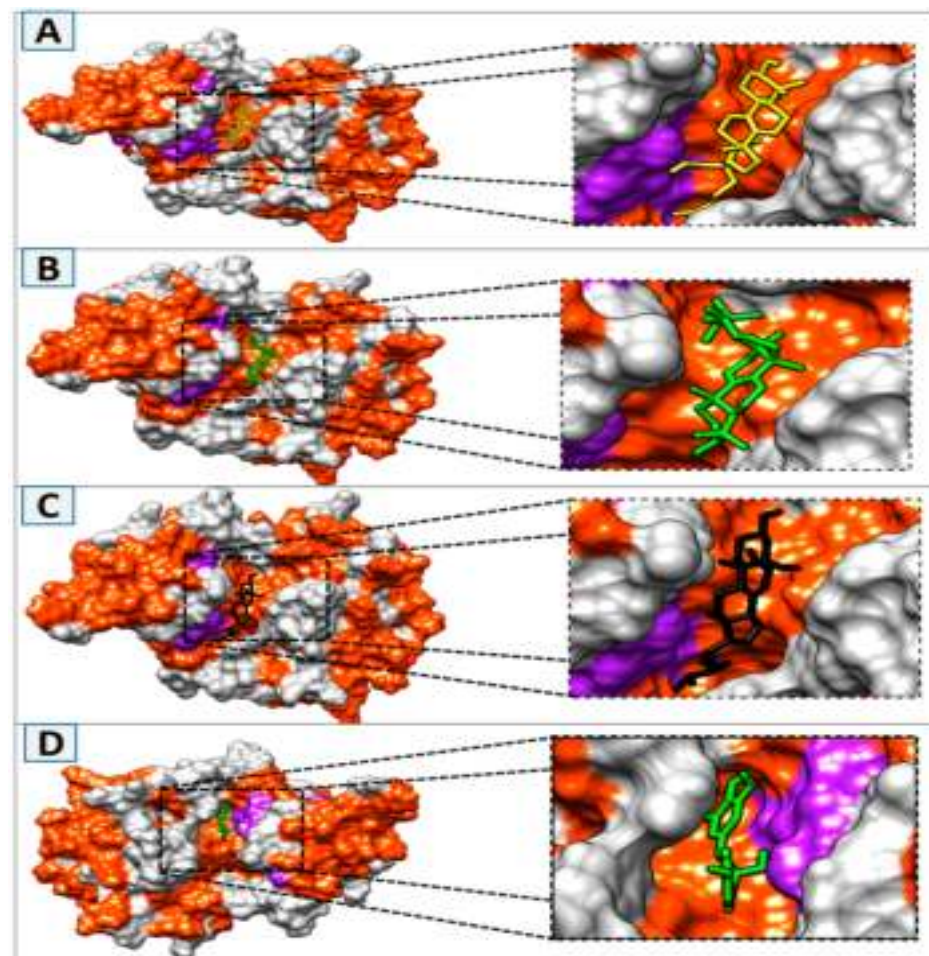


Figure 3. The three-dimensional (3D) docking interaction and binding poses of target enzyme with ligands as MSID000001 (A), MSID000002 (B), MSID000003 (C), and Control (D).

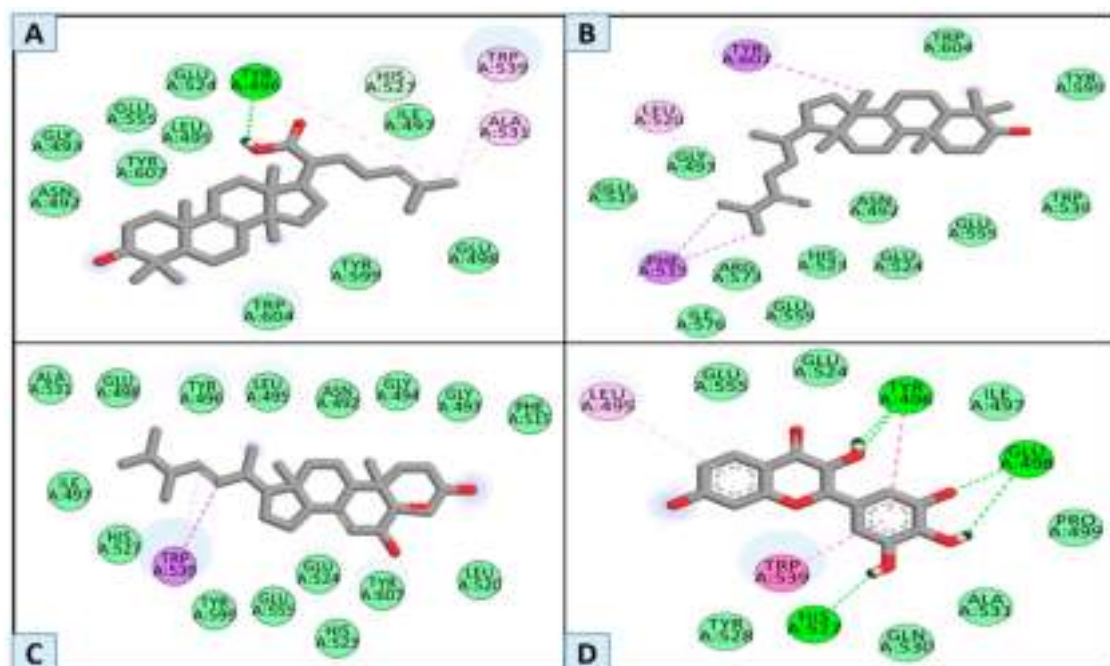


Figure 4. The two-dimensional (2D) interaction between the target enzyme collagenase and ligands MSID000001 (A), MSID000002 (B), MSID000003 (C), and Control (D).

Table 2. The table displays the amino acid interaction profile of hit compounds and control molecules with the target enzyme.

Compounds	H-Bond	Van der Waals	Pi-Alkyl	Alkyl	Carbon-Hydrogen Bond	Pi-Sigma	Pi-Pi Stacked and Pi-Pi T-Shaped
MSID000001	Tyr496	Glu498, Tyr599, Trp604, Ile497, Glu524, Leu495, Glu555, Gly493, Asn492	Trp530	Ala531	His527	-	-
MSID000002	-	Ile576, Glu559, Arg573, His523, Glu524, Asn 492, Glu555, Trp539, Tyr599, Trp604, Glu519, Gly 493	-	Leu520	-	Phe515, Tyr607	-
MSID000003	-	Leu520, Tyr607, Glu524, Tyr599, His527, Ile497, Ala531, Glu498, Tyr496, Leu495, Asn492, Gly494, Phe515	-	-	-	Trp539	
Control	His527, Glu498, Tyr496	Pro499, Ala531, Gln530, Tyr528, Glu555, Ile497, Glu524	Leu495	-	-	-	Trp539, Leu495

3.3. Density Functional Theory (DFT)

Table 3 provides a comprehensive analysis that includes various features such as optimization energies measured in atomic units, dipole moments measured in Debye, polarizability (α) measured in atomic units, and HOMO/LUMO energies and their relative energy gap, measured in electron volts (eV). These features are presented for the compounds under investigation. The range of Hartree optimization energies presented, ranging from a.u. −1105.73 (control) to −a.u.1398.26 (MSID000001), indicates the varying levels of stability among the structures being investigated. The compound with the most stable conformation is MSID000001, which has a high negative energy value of a.u. −1398.26. This is followed by MSID000003 with a value of a.u. −1320.80, then MSID000002 with a value of a.u. −1288.26. The control has the least stable conformation, with a value of a.u. −1105.73.

Table 3. Optimization energies, HOMO and LUMO energies, and their gap calculated in the gas phase at B3LYP/6–311+G(d, p) level of DFT calculations.

Ligand Code	Optimization Energy (a.u.)	Dipole Moment (debye)	Polarizability (α) (a.u.)	E _H (eV)	E _L (eV)	E _g (eV)
Control	−1105.73	6.10	207.35	−6.45	−1.86	4.59
MSID000001	−1398.26	3.08	347.85	−6.13	−0.43	5.71
MSID000002	−1288.26	1.80	352.91	−6.01	−0.21	5.81
MSID000003	−1320.80	2.30	338.32	−6.53	−0.51	6.02

Figure 5 depicts the optimal structures of the compounds under investigation, providing a visual representation of their chemical topologies and spatial arrangements. All

the compounds that were investigated exhibited stable conformation as real local minima, as evidenced by the measured positive imaginary frequencies. The MSID000001 exhibits a significant dipole moment, suggesting a strong electrostatic potential that may impact its binding affinity with the chosen target protein. The measured polarizability values ranged from 207.35 (control) to 347.85 (MSID000001), to 352.91 (MSID000002), to 338.32 (MSID000003) atomic units (a.u.). This suggests a spectrum of molecular reactivity among the chemicals tested. The nucleophilic and electrophilic properties are correlated with the energy values of the HOMO and the LUMO. The molecule MSID000002 has a significant electron-donating property due to its high HOMO energy value of e.V. -6.01 . The LUMO density cloud is a vacant cloud that exhibits an electron-accepting property. The MSID000001 chemical exhibits a low energy gap of e.V. 5.71, which is higher than the control value of e.V. 4.59. This indicates that MSID000001 has a greater reactivity compared to the other compounds investigated.

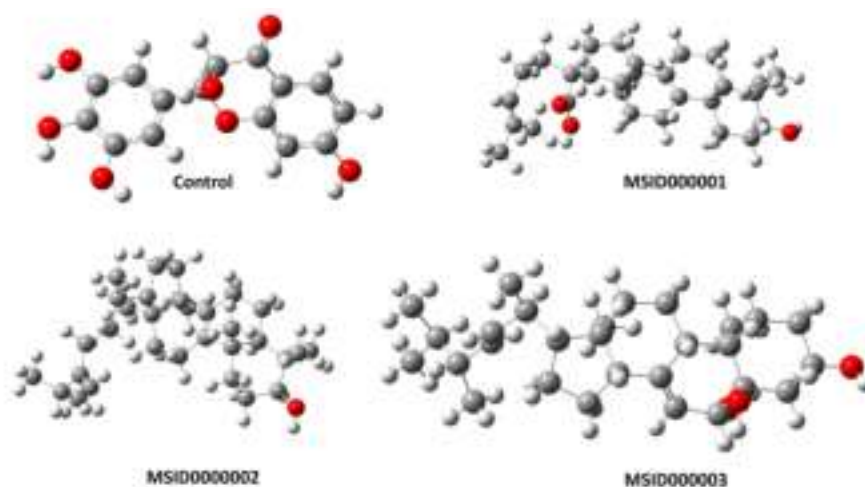


Figure 5. Optimized structures of the studied compounds at the B3LYP/6-311+G(d, p) level of DFT analysis in the gas phase.

Figure 6 displays the contour maps of HOMO-LUMO orbitals along with their corresponding energy gap values. The HOMO/LUMO energy values were utilized to create the global reactivity descriptors, which functioned as important instruments for comprehending medication properties and their potential interactions with biological targets. The compound MSID000003 exhibited a high ionization potential ($I = 6.53$), surpassing the control (6.45), MSID000001 (6.13), and MSID000002 (6.01) accordingly. This indicates that MSID000003 possesses a notable capacity to donate electrons in comparison to the other compounds under investigation.

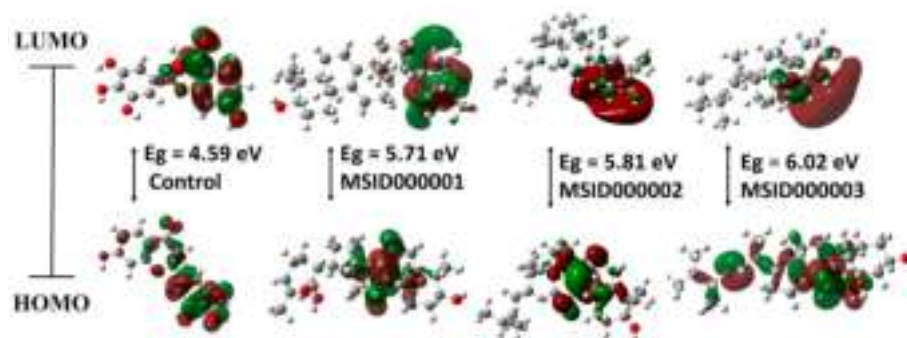


Figure 6. The contour plots of HOMOs and LUMOs of studied compounds.

Conversely, compound MSID000002 had a low electron affinity value ($A = 0.21$), indicating that it has the lowest capacity among the compounds to accept electrons. The

hardness and softness values range from 1.37 to 2.80 and 0.68 to 1.40, respectively, and were calculated for the compounds under investigation. The molecule MSID000002 exhibits a high level of reactivity, as seen by its elevated softness value ($S = 1.40$). The electrophilicity value (ω) ranged from 11.76 (control) to 17.08 (MSID000003), indicating their ability to interact with nucleophilic scaffolds of biological targets. The reactivity descriptors of the substances under investigation are provided in Table 4.

Table 4. Global reactivity descriptors of the targeted compounds.

Ligand Code	Chemical Potential μ (eV)	Electronegativity χ (eV)	Hardness η (eV)	Softness S (eV)	Electrophilicity ω (eV)	Ionization Potential (I)	Electron Affinity (A)
Control	4.15	−4.15	1.37	0.68	11.76	6.45	1.86
MSID000001	3.28	−3.28	2.64	1.32	14.20	6.13	0.43
MSID000002	3.11	−3.11	2.80	1.40	13.53	6.01	0.21
MSID000003	3.52	−3.52	2.75	1.38	17.08	6.53	0.51

Molecular electrostatic potential maps are essential tools for visualizing the charge distribution pattern of the investigated molecules. In the context of MEP, as depicted in Figure 7, the red areas indicate locations that have a higher affinity for electron-rich centers. These regions are considered susceptible to electrophilic attacks. Conversely, the blue spots indicate regions where molecules have a higher affinity for electron-deficient centers, which facilitates nucleophilic assaults. The electrophilic and nucleophilic regions of a molecule play a crucial role in forming polar contacts with specific amino acid residues. The yellowish-green parts exhibited the neutral aspect of the examined chemicals, indicating their potential van der Waals interactions with the target protein.

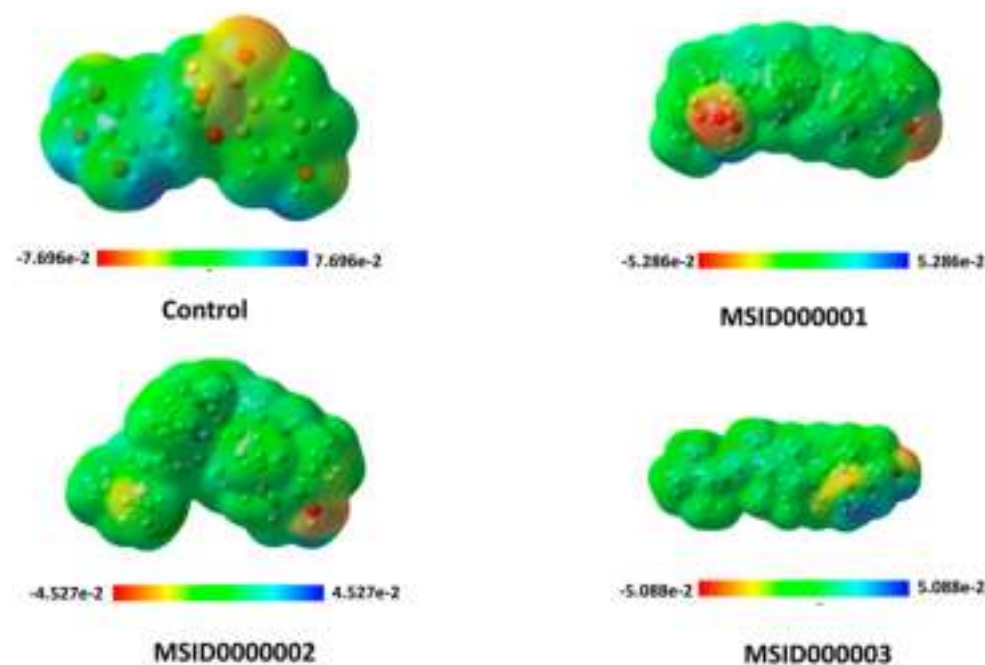


Figure 7. Molecular Electrostatic Potential (MEP) maps for the studied compounds.

Figure 7 illustrates the molecular electrostatic potential maps of the substances under investigation. The presence of negative charge potential, indicated by red patches on oxygen atoms in the analyzed structures, suggests that these locations are susceptible to electrophilic attacks. These locations can potentially operate as sites for polar interactions with important amino acid residues, as they are capable of accepting hydrogen bonds.

3.4. Absorption, Distribution, Metabolism, Excretion, and Toxicity (ADMET) Profiling

The selected compounds' physiochemical characteristics were also evaluated. The compound's chemical formula, molecular weight, number of aromatic heavy atoms, proportion Csp³, the physiochemical properties of refractivity, TPSA, and number of H-bond donors and acceptors were analyzed. Furthermore, pharmacokinetic, water solubility studies, and lipophilicity of the compounds were evaluated. The GI absorption, P-gp substrate, BBB permeant, medicinal chemistry features, and drug-likeness of the identified compounds were next studied in the pharmacokinetic profile analysis. Tables S1–S4 include all of this information. The compounds were identified as drug-like molecules since they have suitable physiochemical characteristics. It was also noted that the compounds had an acceptable range of pharmacokinetics, GI absorption oral bioavailability. The compounds have a high synthetic accessibility score, making them simple to synthesize. Similarly, the compounds had no PAINS assay and were therefore specific in their action.

3.5. Molecular Dynamic Simulation

With the “AMBER20 software”, the top-docked complexes and control molecule was subjected to a 500 ns MD simulation. The simulation trajectories generated data for (i) root mean square deviation (RMSD), (ii) root mean square fluctuation (RMSF), (iii) radius of gyration (RoG), (iv) beta factor analysis, and (v) solvent accessible surface area (SASA).

The protein–ligand complex's stability during the simulation period can be observed by the RMSD figure. If the RMSD increases steadily, it indicates structural deviation; whereas if it plateaus, it represents equilibrium [55]. The three lead complexes MSID000001, MSID000002, and MSID000003 along with the control molecule showed stability throughout the simulation period with no major deviation and changes observed in the ligand binding poses with average RMSD of 1.96 Å, 2.01 Å, 1.59 Å, and 1.60 Å, whereas maximum values noted were 2.76 Å, 3.07 Å, 2.31 Å, and 2.27 Å, respectively (Figure 8A). The flexibility of the residues of amino acid in the complex of protein–ligand can be seen by RMSF analysis (Figure 8B). In the RMSF graphic, troughs denote rigid regions while peaks indicate regions with greater flexibility [56]. RMSF analysis showed notable fluctuation at the start of the simulation period; however, stabilization can occur after 200 ns. The RMSF minimum values showed were 0.95 Å, 0.88 Å, 0.87 Å, and 0.82 Å with maximum values being 3.28 Å, 4.41 Å, 5.44 Å, and 3.74 Å respectively. The lowest values for the top three complexes and control were 0.44 Å, 0.40 Å, 0.44 Å, and 0.38 Å sequentially for MSID000001, MSID000002, MSID000003, and Control. Over the simulation, the rigidity of the protein–ligand combination is evaluated through the radius of gyration analysis (RoG). Whenever RoG increases, it indicates unwinding or expansion; when it declines, it indicates compactness and rigidity [38]. In accordance with the formation of a stable protein–ligand complex, our simulation results indicate that each complex remained rigid and compact over the simulation periods, suggesting the complexes stabilization and compaction (Figure 8C). Their average values observed were 21.88 Å, 21.77 Å, 21.65 Å, and 21.68 Å MSID000001, MSID000002, MSID000003, while maximum values of 22.19 Å, 22.16 Å, 22.04 Å, and 22.03 Å and lowest of 21.45 Å, 21.43 Å, 21.31 Å, and 21.35 Å were observed. Low RoG values suggest a compact structure while an increase shows expansion and folding.

β-factors analysis obtains insights into the flexibility and mobility of individual atoms inside the protein–ligand complex [57]. Beta factor results indicate mean values for the top three complexes and control are 29.55 Å, 27.96 Å, 26.06 Å, and 22.17 Å, respectively. Their maximum values obtained were 283.7 Å, 513.97 Å, 780.8 Å, and 370.05 Å in a similar manner. Moreover, their lowest values noted were 5.12 Å, 4.39 Å, 5.28 Å, and 3.92 Å for MSID000001, MSID000002, and MSID000003, correspondingly. Close the ligand binding position with values ranging from 22 to 29, reflecting better flexibility in residues involved in ligand binding and interaction (Figure 8D).

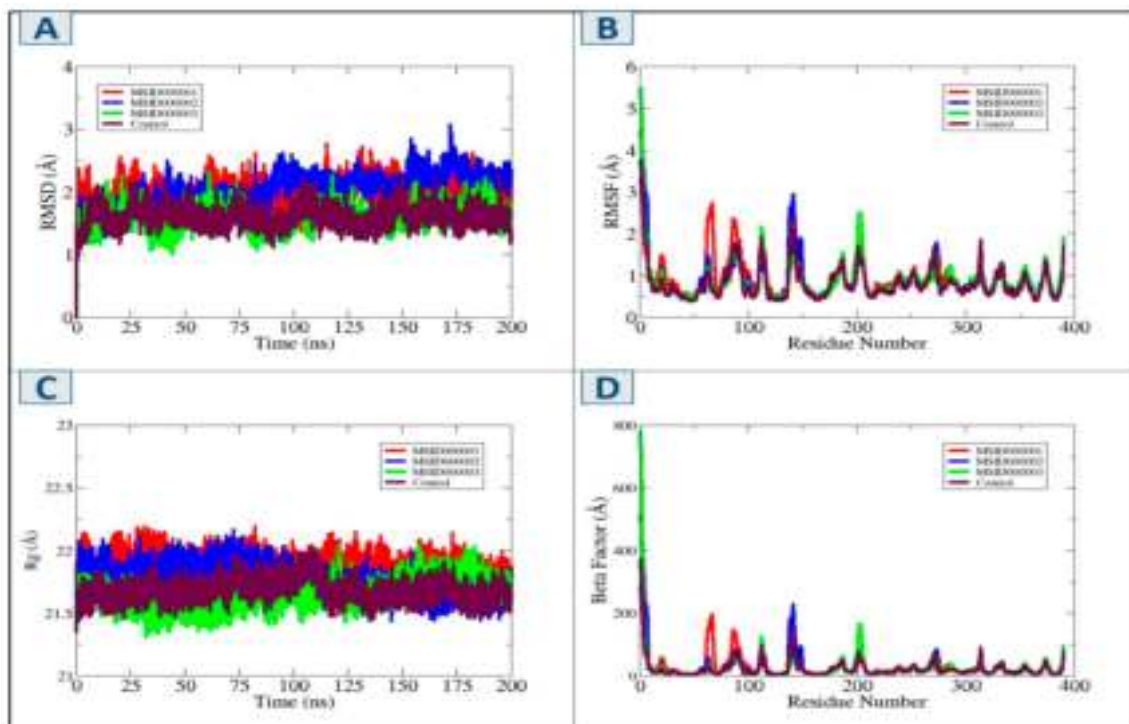


Figure 8. The MSID000001, MSID000002, and MSID000003 complexes and control molecule flexibility, compactness, and stability through (A) RMSD, (B) RMSF, (C) RoG, and (D) β -factors.

3.6. Solvent Accessible Surface Area (SASA)

An analysis of solvent-accessible surface area (SASA) was conducted on the top complexes to find out the interaction between the solvent molecules on the surface of the target protein. The average values for MSID000001, (1722.1 nm²), MSID000002 (16,649.3 nm²), MSID000003 (16,502 nm²), and control (16,625 nm²) were obtained, while maximum values were revealed as (18,620.3 nm²), (17,906.5 nm²), (18,158.8 nm²), and the control was (18,013.1 nm²), respectively. Conversely, the lowest values observed were MSID000001 (14,723.8 nm²), MSID000002 (14,772.3 nm²), MSID000003 (14,46.2 nm²), and the control (14,633.3 nm²). Plots illustrate that following ligand binding, no major deviations were noticed as shown in the following Figure 9.

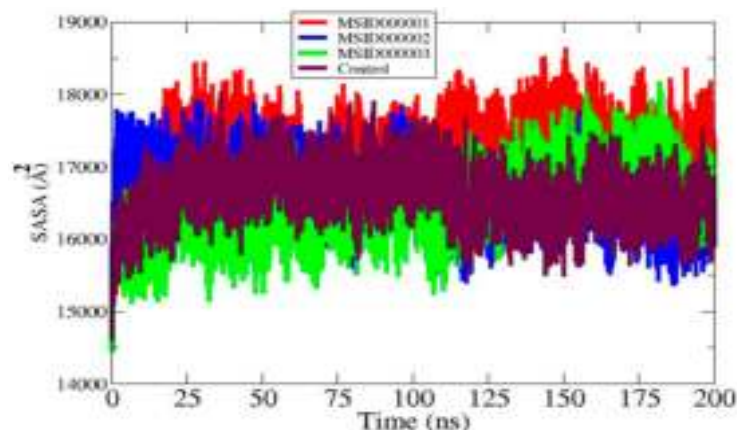


Figure 9. Insights into the top three complexes and control with target enzyme collagenase.

3.7. H-Bonding Analysis

It was vital to emphasize the important role of hydrogen bonds in maintaining complexes between ligands and docked receptors. Identifying which of the enzyme residues

formed hydrogen bonds with the compounds was particularly necessary. As demonstrated in Table 5, the hydrogen atom that forms the bond is referred to as the “Donor”, whereas the negatively charged atom that takes the bond is labeled as the “Acceptor”. The percentage of the simulation time where the hydrogen bond exists can be seen in the “Occupancy” column. The hydrogen bonds made by the receptor’s enzyme with the top three ligand molecules are given in the table.

Table 5. The hydrogen bonds formed between ligands and the active site of *Clostridium histolyticum* collagenase.

MSID000001		
Donor	Acceptor	Occupancy
HIE128-Side	LIG391-Main	0.30%
HIE128-Side	LIG391-Main	0.10%
LIG391-Side	GLU156-Side	0.10%
MSID000002		
LIG391-Main	ASP184-Side	29.80%
MSID000003		
LIG391-Main	TYR208-Side	1–10%
LIG391-Side	GLU156-Side	5.40%
LIG391-Main	TYR200-Side	0.30%
LIG391-Main	GLU156-Side	0.10%
LIG391-Side	TYR208-Side	1.50%
TYR208-Side	LIG391-Side	0.10%
LIG391-Side	TYR200-Side	0.10%
Control		
LIG391-Side	GLU125-Side	70.00%
LIG391-Side	GLU99-Side	2.90%

3.8. Principal Component Analysis (PCA)

Multivariate statistical methods such as PCA are employed to better understand the vast dataset’s complexity. It transforms the measured variables into principal components (PCs), that represent unrelated variables [58]. Each PC is complementary, indicating they are orthogonal to each other. PC1 represents the dataset’s highest variability, subsequently followed by PCs (PC2, PC3, etc.) [59]. PCA was conducted and visualized to identify the most notable structural changes within the proposed protein–ligand complexes exhibited during binding to each other system [60]. Principal component analysis (PCA) was employed to obtain data on the conformation phases and layout of the target enzyme with the ligands and control molecule by using the 100 ns molecular dynamics (MD) simulation trajectories.

According to the study’s findings, the majority of the compounds under this research are physiologically active molecules that interact with nuclear receptor ligands and other enzymes to induce physiological activities by inhibiting collagenase enzyme. After analyzing these graphs, it is apparent that the eigenvalue distributions of the complexes shown in Figure 10 differ from one another. Specifically, the eigenvalues of the control (D) complex show a much wider dispersion followed by MSID000003 than those of MSID000001 (A) and the MSID000002 (B) complex. This suggests that there is more variability in the conformational states of the interactions involving the *Clostridium histolyticum* collagenase enzyme in Control (D). This variability implies that the structural dynamics and poten-

tially the functional implications of the enzyme's interactions vary significantly among these complexes.

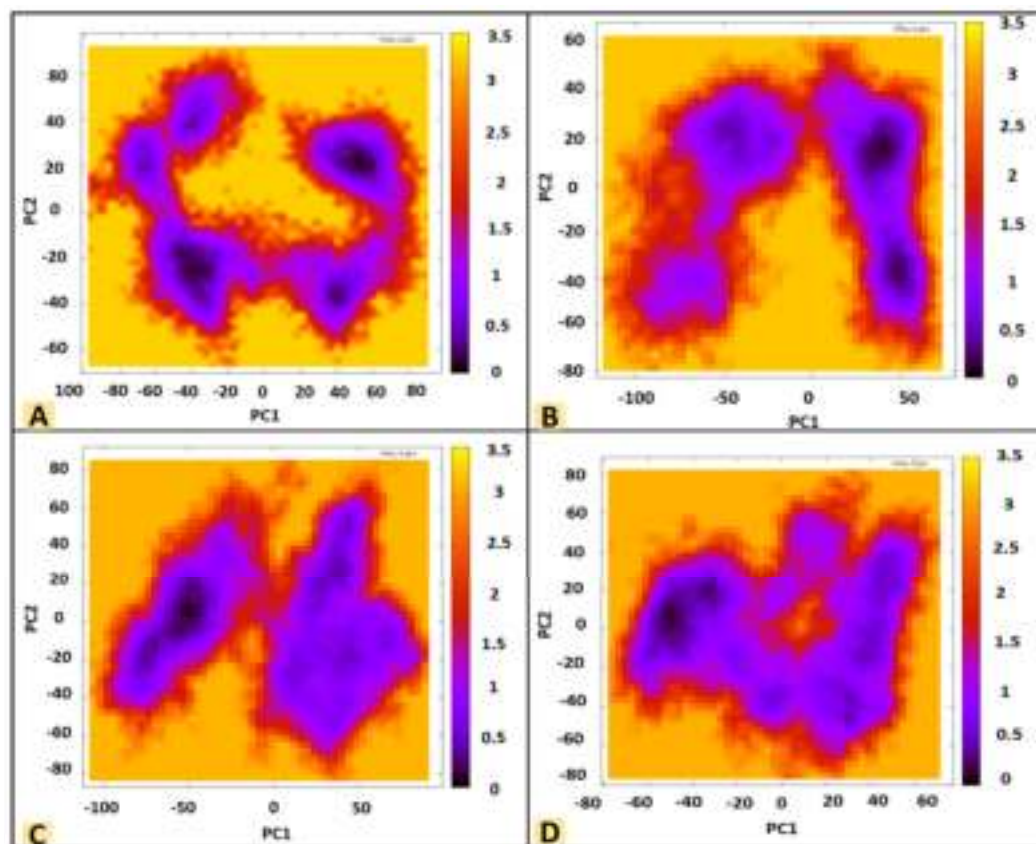


Figure 10. The principal component analysis of MSID000001 (A), MSID000002 (B), MSID000003 (C) and Control (D), respectively.

3.9. Secondary Structure Analysis

The overall structure of the protein–enzyme complex during analysis is made up of secondary structural elements including alpha helices and beta sheets, which provide complex stability and contain the active sites required for enzymatic activity [61]. Secondary structure analysis was performed to understand and investigate the structural basis of *C. histolyticum* collagenase function. Tight turns and loops can be seen within the enzyme component, which aids in the binding of substrates and catalytic turnover. The number of β sheets and alpha helix formed over the simulation period was calculated and depicted in Figure S1 given. Upon analyzing, different secondary structure elements were shown such as T stands for turn and loops in the structure, E represents β -strands, B for isolated H-bonding β bridge between β strands, H shows α helix in SS motif where protein backbone facilitates a helical structure, G represents a type of helical Sec-Structure, Pi–Pi indicated a rare form helical structure, and C shows coils in the protein structure. As shown in Figure S1, where the secondary structure of MSID000001, MSID000002, MSID000003, and Control are depicted.

3.10. MMPBSA/GSA Calculations

Using AMBER22's MM-GBSA and MM-PBSA modules, the binding free energies for the MSID000001, MSID000002, MSID000003, and Control docked complexes were estimated (Table 6). For the MSID000001 docked complex, the MM-GBSA calculated a total energy of -78.58 kcal/mol, for the MSID000002 complex, -70.44 kcal/mol, and for the MSID000003 docked complex, -59.32 kcal/mol. As shown in Table 3, the calculated net free energy for complexes in the MM-PBSA results was -79.54 kcal/mol, -73.99 kcal/mol,

and -62.22 kcal/mol, correspondingly. It is anticipated that the complexes develop strong and stable intermolecular interactions based the low net binding energy scores.

Table 6. MMGBSA and MMPBSA analysis of top three complexes along with control molecule.

Energy Parameter	MSID000001	MSID000002	MSID000003	Control
MMGBSA				
van der Waals energy	-65.14	-61.23	-55.91	-60.99
Energy electrostatic	-24.01	-20.81	-15.49	-17.67
Total gas phase energy	-89.15	-82.04	-71.4	-78.66
Total solvation energy	10.57	11.60	12.08	10.46
Net energy	-78.58	-70.44	-59.32	-68.2
MMPBSA				
Energy van der Waals	-65.14	-61.23	-55.91	-60.99
Energy electrostatics	-24.01	-20.81	-15.49	-17.67
Total gas phase energy	-89.15	-82.04	-71.4	-78.66
Total energy solvation	9.61	8.05	9.14	8.00
Net energy	-79.54	-73.99	-62.26	-70.66

3.11. WaterSwap Energy Estimation

WaterSwap energy estimation in computer-aided drug design is carried out to obtain insights into the interaction of water molecules with target proteins and drugs [49]. By studying these interactions, we can make better drugs that bind more tightly and specifically to their targets, improving their effectiveness [62]. To revalidate the GBSA/MMPB estimations, the binding free energy for each docked complex has been calculated with the WaterSwap technique. Figure 11 illustrates the predicted WaterSwap binding energy from multiple steps. Three different approaches were used, particularly Bennetts, thermodynamic integration (TI), and free energy perturbation (FEP). The lead compounds MSID000001, MSID000002, and MSID000003, and the control are stable in binding energies and appear to be promising drug candidates. The Bennetts, TI, and FEP-binding free energy values for each complex are outlined in Figure 11 below.

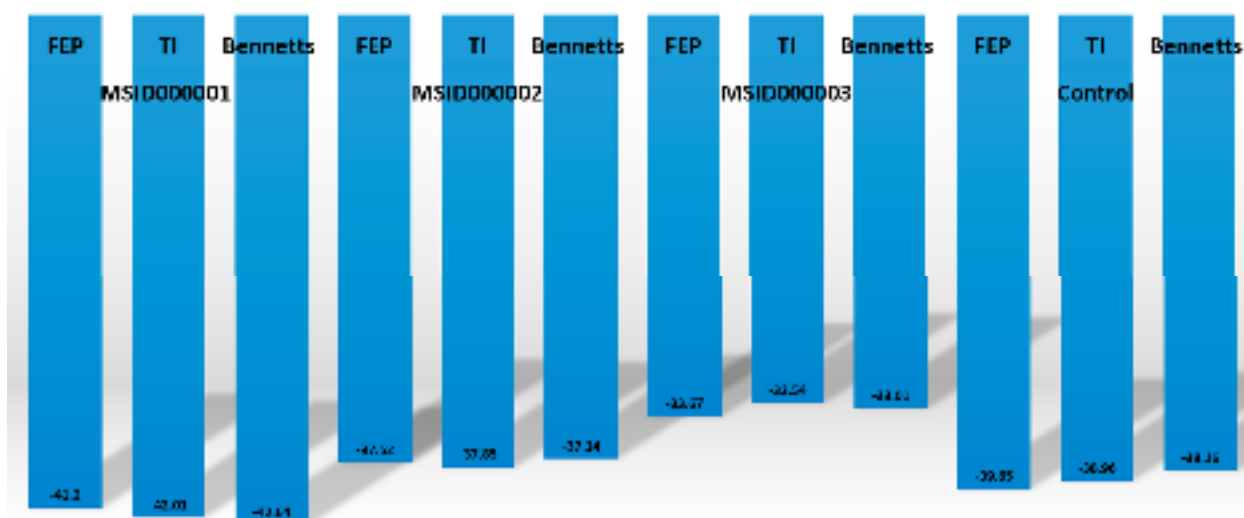


Figure 11. Predicted WaterSwap binding energy for MSID000001, MSID000002, and MSID000003, and Control.

3.12. Entropy Energy Estimation

Energy calculations are performed to assess ligand–target complexes' stability and binding affinity in the development of drugs, offering details about the complex's structural features, stability, and binding affinity [63]. To obtain insights into the net binding energy, entropy energy contribution was calculated using AMBER22. The entropy energy contribution of the top three complexes and control is outlined in the table (Table 7), showing values of -5.96 kcal/mol, -2.85 kcal/mol, -1.36 kcal/mol, and -2.60 kcal/mol for MSID000001, MSID000002, and MSID000003, and the control, respectively.

Table 7. The entropy energy calculations of the top three complexes as well as the control molecule.

Complex	Translational	Vibrational	Rotational	ΔS Total
MSID000001	5.01	7.86	1147.09	-5.96
MSID000002	10.85	12.66	1269.48	-2.85
MSID000003	15.96	13.05	1566.12	-1.36
Control	10.53	12.04	1428.64	-2.60

3.13. Salt Bridges Studies

The strongest non-covalent interaction found in nature is known as salt bridges, which are involved in molecular recognition, protein–protein interactions, and protein folding [64]. Two types of amino acid side chains are involved in salt bridges interaction, Glu or Glu if the ligand is positively charged, whereas negatively charged amino acid chains are Arg or Lys [65] (Table 8). In this technique, the electrostatic interaction between the oppositely charged ions was observed in the docking region of MSID000001, MSID000002, MSID000003, and Control. (See Figure S2).

Table 8. Salt bridge interaction between MSID000001, MSID000002, MSID000003, and Control with the target enzyme can be seen in the following table.

Complexes	Salt Bridges Interaction
MSID000001	Glu34-Lys37, Glu147-Lys148, Glu220-Lys221, Asp1-Arg101, Asp279-Arg38, Asp222-Lys254, Glu120-Arg174, Glu61-Arg133, Asp338-Lys335, Asp345-Arg52, Asp67-Arg133, Asp187-Lys185, Glu8-Arg26, Glu45-Arg123, Asp6-Lys2, Glu33-Lys14, Glu292-Lys289, Asp345-Arg52, Asp204-Lys194, Asp280-Lys360, Glu339-Lys335, Asp6-Lys9, Glu33-Lys37, Glu78-Lys81, Asp19-Arg44, Glu264-Lys268, Asp92-Lys81, Glu339-Lys331, Glu-Arg44
MSID000002	Glu349-Arg167, Asp6-Lys9, Glu328-Lys331, Asp380-Arg370, Asp222-Lys254, Glu34-Arg38, Asp187-Lys185, Glu108-Lys35, Glu339-Lys342, Asp5-Lys72, Glu8-Lys72, Glu8-Lys72, Glu34-Arg38, Glu45-Arg123, Asp66-Arg133, Glu220-Lys268, Glu160-Arg174, Glu33-Lys14, Glu292-Lys289, Glu160-Arg174, Glu119-Arg123, Glu264-Lys221, Asp338-Lys342, Asp5-Arg101, Asp184-Lys183, Asp19-Arg44, Asp57-Arg322, Asp242-Arg150, Asp184-Lys183, Asp280-Lys283, Glu108-Lys35, Glu339-Lys331, Asp220-Lys254
MSID000003	Glu34-Lys37, Asp6-Lys9, Asp57-Lys58, Glu8-Arg26, Asp237-Lys239, Asp380-Arg370, Asp29-Lys30, Asp374-Lys376, Glu292-Lys299, Glu339-Lys342, Glu45-Arg123, Asp66-Arg133, Glu220-Lys268, Asp5-Lys9, Asp267-Lys268, Asp345-Arg52, Glu339-Lys343, Asp250-Lys264, Asp5-Lys9, Asp338-Lys342, Asp57-Arg322, Asp242-Arg150, Glu34-Lys37, Asp280-Lys283, Asp242-Arg150, Glu257-Lys254, Asp29-Lys30, Glu349-Arg167, Asp242-Arg150, Asp5-Lys2, Asp222-Lys254
Control	Asp6-Lys2, Glu264-Lys268, Glu34-Arg38, Glu328-Lys331, Glu243-Lys264, Asp279-Arg38, Glu333-Lys231, Asp187-Lys185, Glu33-Lys14, Glu311-Arg44, Glu45-Arg123, Asp242-Lys148, Asp338-Lys335, Asp279-Arg38, Asp66-Arg133, Glu34-Arg38, Glu333-Lys228, Glu243-Lys239, Glu292-Lys289, Asp250-Lys193, Glu119-Arg44, Asp338-Lys342, Asp250-Lys264, Asp242-Arg150, Asp57-Arg322, Asp57-Arg322, Glu339-Lys331, Glu257-Lys254, Asp336-Lys331, Glu108-Lys35, Glu243-Lys246, Glu311-Arg44

There were significant differences in the distribution of charged residues involved upon comparing the salt bridges of the hit compounds compared to the control. When it came to creating salt bridges, the control displayed an increased number of positively

charged residues, such as arginine (Arg) and lysine (Lys). On the contrary, aspartate (Asp) and glutamate (Glu), are examples of negatively charged residues that are primarily involved in salt bridge interactions in the hit compounds. The variation suggests that whereas the control mostly comprises positively charged residues, the hit compounds largely form salt bridges with negatively charged residues. The differences in residue involvement increase the possibility that the hit compounds' binding affinities and specifics differ from those of the control, which might enhance the stability and total binding efficiency to the target protein.

4. Discussion

Experimental studies that seek to comprehend a disease's molecular basis at the structural level usually demand a long time and a lot of effort. Since the experimental techniques have many downsides and limitations, there exists a market for *in silico* approaches, which may identify potential inhibitors with greater efficiency and precision [38]. Multiple sequence-based and structure-based prediction approaches, which comprise several algorithms, combine to develop an effective tool that delivers precise and accurate predictions [66].

The main collagenase enzyme produced by *Clostridium histolyticum* draws a lot of attention as a promising target for drug interventions due to its key role in the breakdown of tissues and the healing process of wounds [67]. The protein that is most common in the extracellular matrix, or outer layer, of connective tissues, collagen, is broken down by the main collagenase enzyme produced by *Clostridium histolyticum* [10]. The underlying causes of particular disorders like Dupuytren's contracture, Peyronie's disease, as well as wound healing and tissue remodeling, all are significantly affected by the enzyme's activity [68]. Targeting the main collagenase enzymes via specific inhibitors has considerable potential in regulating tissue breakdown, changing wound healing mechanisms, and possibly lowering symptoms related to conditions affecting collagen [69]. So, developing drugs that efficiently target the main collagenase enzyme of *Clostridium histolyticum* does not come with any obstacles. Selection against host enzymes must be maintained, and adverse reactions need to be kept to a minimum [70].

Using computational techniques, a comprehensive *in silico* study was conducted to identify potential inhibitors against *Clostridium histolyticum* with the main collagenase enzyme. By employing molecular docking modeling, we performed the screening of Medicinal Fungi Secondary Metabolites And Therapeutics MeFSAT library having 1800 compounds and identified three lead compounds. MSID000001, MSID000002, MSID000003 exhibited promising binding scores of -10.7 kcal/mol, -9.8 kcal/mol, and -9.5 kcal/mol, respectively, while the control showed a binding score of -8 kcal/mol to the enzyme's active site. The binding mode and interactions were not changed in any apparent way during the simulated time. These compounds met the criteria needed to be considered drug-like and also exhibited good pharmacokinetic properties. Further, MD simulations were employed to show the dynamics and stability of the protein–ligand complexes during the simulation period. Trajectory data analysis was utilized to identify good binding conformations and structural modifications within the collagenase active site on ligand binding. Further, the dynamics of the enzyme–ligand systems were investigated by principal component analysis (PCA), which showed insight into the conformational changes caused by inhibitor binding. Moreover, the research highlighted the effect of inhibitor interaction on the collagenases' secondary structural elements as well as the role of salt bridges in maintaining the stability of enzyme–ligand complexes. Understanding the mechanism of action and optimizing lead compound development for greater specificity and efficacy requires knowledge of the mentioned data. Considering their potential as novel drugs, these compounds offer further experimental research to determine how effectively they act against *Clostridium histolyticum*. Taking everything considered, these findings indicate that the protein remained structurally unaltered whereas the selected ligands are found, despite no obvious conformational shifts occurred.

Our findings align with previous studies investigating inhibitors against collagenases from *Clostridium histolyticum*. A study was reported by [71] to develop a virtual screening technique that integrated ligand-based and structure-based discovery of drug methods to identify potential inhibitors of bacterial collagenase A. Three FDA-approved drugs, benzthiazide, entacapone, and lodoxamide, were identified to be strong candidates for inhibiting bacterial collagenase. These drugs were selected based on their good simulated interactions with bacterial collagenase and their assessed potential for blocking the enzyme. The study emphasized the need for further investigations to verify the predicted biological function of these specific compounds as well as assess their efficacy in managing infections caused by bacterial collagenase.

Identifying possible inhibitors of *Clostridium histolyticum* collagenases could potentially be helpful to address diseases like fibrosis and chronic wounds which are connected to the degradation of collagen. The enzyme activity of collagenases is the target of these inhibitors, which could reduce tissue damage and promote tissue regeneration and wound healing. The lead compounds found by in silico analysis, according to our findings, should be subjected to extensive validation experiments to assess the effectiveness and safety features. Whereas in silico analysis has tremendous promise, this study includes certain limitations. Simulation variables and the quality of the protein's structure are key factors in computational accuracy. Protein flexibility and effects of solvent might be missed by docking analysis, raising the likelihood of false positive results. The demand for multidisciplinary cooperation and validation to confirm inhibitory potency is highlighted by the limited experimental data.

The optimization of inhibitor features and experimental validation should be the primary fields of future study. Integrating experimental and computational data may advance the development of new drugs for conditions linked to collagen. In a nutshell, our findings facilitate the discovery of drugs by improving treatment options for disorders linked to collagen degradation, although further validation needs to be conducted to combat antimicrobial resistance.

5. Conclusions

This research aimed to identify *C. histolyticum* collagenase G inhibitors. By using computer modeling and an in silico approach, we were able to identify the top three compounds as MSID000001, MSID000002, and MSID000003 that showed potential in efficiently inhibiting the activity of the enzyme. These findings offered a step towards the development and characterization of novel inhibitors targeting the *C. histolyticum* Peptidase G domain. Further investigation needs to be conducted to efficiently inhibit the said enzyme. Moving forward, experimental investigation needs to be performed to validate the computational findings of this research.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/computation12080153/s1>, Figure S1: The secondary structure of the top three complexes MSID000001 (A), MSID000002 (B), MSID000003 (C), and Control (D), respectively. Figure S2: The salt bridges interaction of top three complexes MSID000001, MSID000002, MSID000003 and Control, respectively are depicted in the figure given above. Table S1: Pharmacokinetics profile of the MSID000001, MSID000002, MSID000003 and Control. Table S2: Lipophilicity properties of the investigated compounds. Table S3: Physiochemical properties of the selected compounds. Table S4: Drug-likeness of the studied compounds.

Author Contributions: Conceptualization, F.A.; methodology, F.A., A.H., N.A.A. and F.A.A.; software, N.H.; validation, A.A.A.; formal analysis, A.A.A.; investigation, M.B.; resources, A.S.; data curation, N.A.A. and A.A.A.; writing—original draft preparation, H.J.B.; writing—review and editing, A.A.; visualization, M.I.N.; supervision, F.A.; project administration, F.A.; funding acquisition, F.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Taif University, Saudi Arabia, project number (TU-DSPP-2024-140).

Data Availability Statement: All the data generated in the current research is presented in the manuscript.

Acknowledgments: The authors extend their appreciation to Taif University, Saudi Arabia, for supporting this work through project number (TU-DSPP-2024-140).

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Alhayek, A.; Abdelsamie, A.S.; Schönauer, E.; Camberlein, V.; Hutterer, E.; Posselt, G.; Serwanja, J.; Blöchl, C.; Huber, C.G.; Haupenthal, J.; et al. Discovery and Characterization of Synthesized and FDA-Approved Inhibitors of Clostridial and Bacillary Collagenases. *J. Med. Chem.* **2022**, *65*, 12933–12955. [\[CrossRef\]](#) [\[PubMed\]](#)
- Muteeb, G.; Rehman, M.T.; Shahwan, M.; Aatif, M. Origin of Antibiotics and Antibiotic Resistance, and Their Impacts on Drug Development: A Narrative Review. *Pharmaceuticals* **2023**, *16*, 1615. [\[CrossRef\]](#) [\[PubMed\]](#)
- Calvert, M.B.; Jumde, V.R.; Titz, A. Pathoblockers or Antivirulence Drugs as a New Option for the Treatment of Bacterial Infections. *Beilstein J. Org. Chem.* **2018**, *14*, 2607–2617. [\[CrossRef\]](#) [\[PubMed\]](#)
- Popoff, M.R.; Bouvet, P. Genetic Characteristics of Toxigenic Clostridia and Toxin Gene Evolution. *Toxicon* **2013**, *75*, 63–89. [\[CrossRef\]](#) [\[PubMed\]](#)
- Brown, R.L.; Larkinson, M.L.Y.; Clarke, T.B. Immunological Design of Commensal Communities to Treat Intestinal Infection and Inflammation. *PLoS Pathog.* **2021**, *17*, e1009191. [\[CrossRef\]](#) [\[PubMed\]](#)
- Palmer, J.S.; Hough, R.L.; West, H.M.; Avery, L.M. A Review of the Abundance, Behaviour and Detection of Clostridial Pathogens in Agricultural Soils. *Eur. J. Soil Sci.* **2019**, *70*, 911–929. [\[CrossRef\]](#)
- Gabrielson, A.T.; Spitz, J.T.; Hellstrom, W.J.G. Collagenase Clostridium Histolyticum in the Treatment of Urologic Disease: Current and Future Impact. *Sex. Med. Rev.* **2018**, *6*, 143–156. [\[CrossRef\]](#) [\[PubMed\]](#)
- Popoff, M.R. Overview of Bacterial Protein Toxins from Pathogenic Bacteria: Mode of Action and Insights into Evolution. *Toxins* **2024**, *16*, 182. [\[CrossRef\]](#) [\[PubMed\]](#)
- Kaya, C.; Hirsch, A.K.H. Targeting Extracellular Bacterial Proteases for the Development of Novel Antivirulence Agents. *Chimia* **2022**, *76*, 402–408. [\[CrossRef\]](#) [\[PubMed\]](#)
- Eckhard, U.; Huesgen, P.F.; Brandstetter, H.; Overall, C.M. Proteomic Protease Specificity Profiling of Clostridial Collagenases Reveals Their Intrinsic Nature as Dedicated Degradors of Collagen. *J. Proteom.* **2014**, *100*, 102–114. [\[CrossRef\]](#) [\[PubMed\]](#)
- Shekhter, A.B.; Balakireva, A.V.; Kuznetsova, N.V.; Vukolova, M.N.; Litvitsky, P.F.; Zamyatnin Jr, A.A. Collagenolytic Enzymes and Their Applications in Biomedicine. *Curr. Med. Chem.* **2019**, *26*, 487–505. [\[CrossRef\]](#) [\[PubMed\]](#)
- Sargsyan, A.; Hakobyan, H.; Mardiyani, Z.; Jamharyan, S.; Dadayan, A.; Sargsyan, T.; Hovhannisyanyan, N. Modeling, Synthesis and in Vitro Screening of Unusual Amino Acids and Peptides As Protease Inhibitors. *J. Chem. Technol. Metall.* **2023**, *58*, 615–620. [\[CrossRef\]](#)
- María, R.R.; Arturo, C.J.; Alicia, J.; Paulina, M.G.; Gerardo, A. The Impact of Bioinformatics on Vaccine Design and Development. *Vaccines* **2017**, *2*, 3–6.
- dos Santos Nascimento, I.J.; da Silva Rodrigues, É.E.; da Silva, M.F.; de Araújo-Júnior, J.X.; de Moura, R.O. Advances in Computational Methods to Discover New NS2B-NS3 Inhibitors Useful against Dengue and Zika Viruses. *Curr. Top. Med. Chem.* **2022**, *22*, 2435–2462. [\[CrossRef\]](#) [\[PubMed\]](#)
- Saha, M.; Sarkar, A. Review on Multiple Facets of Drug Resistance: A Rising Challenge in the 21st Century. *J. Xenobiotics* **2021**, *11*, 197–214. [\[CrossRef\]](#) [\[PubMed\]](#)
- Kore, P.P.; Mutha, M.M.; Antre, R.V.; Oswal, R.J.; Kshirsagar, S.S. Computer-Aided Drug Design: An Innovative Tool for Modeling. *Open J. Med. Chem.* **2012**, *2*, 26238. [\[CrossRef\]](#)
- Sarkar, S.; Gulati, K.; Kairamkonda, M.; Mishra, A.; Poluri, K.M. Elucidating Protein-Protein Interactions through Computational Approaches and Designing Small Molecule Inhibitors against Them for Various Diseases. *Curr. Top. Med. Chem.* **2018**, *18*, 1719–1736. [\[CrossRef\]](#)
- Palchevskyi, S.; Czarnocki-Cieciura, M.; Vistoli, G.; Gervasoni, S.; Nowak, E.; Beccari, A.R.; Nowotny, M.; Talarico, C. Structure of Human TRPM8 Channel. *Commun. Biol.* **2023**, *6*, 1065. [\[CrossRef\]](#)
- Kaliappan, S.; Bombay, I.I.T. UCSF Chimera—Overview. 2018. Available online: <https://www.cgl.ucsf.edu/chimera/> (accessed on 20 July 2024).
- Pencheva, T.; Lagorce, D.; Pajeva, I.; Villoutreix, B.O.; Miteva, M.A. AMMOS: Automated Molecular Mechanics Optimization Tool for in Silico Screening. *BMC Bioinform.* **2008**, *9*, 438. [\[CrossRef\]](#) [\[PubMed\]](#)
- Bhowmick, S.; Saha, A.; Osman, S.M.; Alasmary, F.A.; Almutairi, T.M.; Islam, M.A. Structure-Based Identification of SARS-CoV-2 Main Protease Inhibitors from Anti-Viral Specific Chemical Libraries: An Exhaustive Computational Screening Approach. *Mol. Divers.* **2021**, *25*, 1979–1997. [\[CrossRef\]](#) [\[PubMed\]](#)
- Zothantluanga, J.; Aswin, S.K.; Rudrapal, M.; Chetia, D. Antimalarial Flavonoid-Glycoside from Acacia Pennata with Inhibitory Potential against PfDHFR-TS: An In-Silico Study. *Biointerface Res. Appl. Chem* **2021**, *12*, 4871–4887.
- Halgren, T.A. MMFF VII. Characterization of MMFF94, MMFF94s, and Other Widely Available Force Fields for Conformational Energies and for Intermolecular-interaction Energies and Geometries. *J. Comput. Chem.* **1999**, *20*, 730–748. [\[CrossRef\]](#)
- Olawale, F.; Iwaloye, O.; Elekofehinti, O.O. Virtual Screening of Natural Compounds as Selective Inhibitors of Polo-like Kinase-1 at C-Terminal Polo Box and N-Terminal Catalytic Domain. *J. Biomol. Struct. Dyn.* **2022**, *40*, 13606–13624. [\[CrossRef\]](#) [\[PubMed\]](#)

25. Lavecchia, A.; Di Giovanni, C. Virtual Screening Strategies in Drug Discovery: A Critical Review. *Curr. Med. Chem.* **2013**, *20*, 2839–2860. [[CrossRef](#)] [[PubMed](#)]
26. Dallakyan, S.; Olson, A.J. Small-Molecule Library Screening by Docking with PyRx. In *Chemical Biology. Methods in Molecular Biology*; Humana Press: New York, NY, USA, 2015; pp. 243–250.
27. Kumar, S.P.; Girija, A.S.S.; Priyadharsini, J.V. Targeting NM23-H1-Mediated Inhibition of Tumour Metastasis in Viral Hepatitis with Bioactive Compounds from *Ganoderma Lucidum*: A Computational Study. *Indian J. Pharm. Sci.* **2020**, *82*, 300. [[CrossRef](#)]
28. Bilal, M.S.; Ejaz, S.A.; Zargar, S.; Akhtar, N.; Wani, T.A.; Riaz, N.; Aborode, A.T.; Siddique, F.; Altwaijry, N.; Alkahtani, H.M.; et al. Computational Investigation of 1, 3, 4 Oxadiazole Derivatives as Lead Inhibitors of VEGFR 2 in Comparison with EGFR: Density Functional Theory, Molecular Docking and Molecular Dynamics Simulation Studies. *Biomolecules* **2022**, *12*, 1612. [[CrossRef](#)] [[PubMed](#)]
29. Masnabadi, N.; Thalji, M.R.; Alhasan, H.S.; Mahmoodi, Z.; Soldatov, A.V.; Ali, G.A.M. Structural, Electronic, Reactivity, and Conformational Features of 2, 5, 5-Trimethyl-1, 3, 2-Diheterophosphinane-2-Sulfide, and Its Derivatives: DFT, MEP, and NBO Calculations. *Molecules* **2022**, *27*, 4011. [[CrossRef](#)] [[PubMed](#)]
30. Hashem, H.E.; Ahmad, S.; Kumer, A.; Bakri, Y. El In Silico and in Vitro Prediction of New Synthesized N-Heterocyclic Compounds as Anti-SARS-CoV-2. *Sci. Rep.* **2024**, *14*, 1152. [[CrossRef](#)]
31. Fatima, A.; Pooja, K.; Savita, S.; Singh, M.; Verma, I.; Siddiqui, N.; Javed, S. Quantum Chemical, Experimental Spectroscopic, Hirshfeld Surface and Molecular Docking Studies of the Anti-Microbial Drug Sulfathiazole. *J. Mol. Struct.* **2021**, *1245*, 131118. [[CrossRef](#)]
32. Bruno, A.; Costantino, G.; Sartori, L.; Radi, M. The in Silico Drug Discovery Toolbox: Applications in Lead Discovery and Optimization. *Curr. Med. Chem.* **2019**, *26*, 3838–3873. [[CrossRef](#)] [[PubMed](#)]
33. Lohohola, P.O.; Mbala, B.M.; Bambi, S.-M.N.; Mawete, D.T.; Matondo, A.; Mvondo, J.G.M. In Silico ADME/T Properties of Quinine Derivatives Using SwissADME and PKCSM Webservers. *Int. J. Trop. Dis. Health* **2021**, *42*, 1–12.
34. Makeneni, S.; Thieker, D.F.; Woods, R.J. Applying Pose Clustering and MD Simulations to Eliminate False Positives in Molecular Docking. *J. Chem. Inf. Model.* **2018**, *58*, 605–614. [[CrossRef](#)] [[PubMed](#)]
35. Shukla, R.; Tripathi, T. Molecular Dynamics Simulation of Protein and Protein–Ligand Complexes. In *Computer-Aided Drug Design*; Springer: Singapore, 2020; pp. 133–161.
36. Pujadas, G.; Vague, M.; Ardevol, A.; Blade, C.; Salvado, M.J.; Blay, M.; Fernandez-Larrea, J.; Arola, L. Protein-Ligand Docking: A Review of Recent Advances and Future Perspectives. *Curr. Pharm. Anal.* **2008**, *4*, 1–19. [[CrossRef](#)]
37. Alamri, M.A.; Mirza, M.U.; Adeel, M.M.; Ashfaq, U.A.; Tahir Ul Qamar, M.; Shahid, F.; Ahmad, S.; Alatawi, E.A.; Albalawi, G.M.; Allemailem, K.S.; et al. Structural Elucidation of Rift Valley Fever Virus L Protein towards the Discovery of Its Potential Inhibitors. *Pharmaceuticals* **2022**, *15*, 659. [[CrossRef](#)] [[PubMed](#)]
38. Miandad, K.; Ullah, A.; Bashir, K.; Khan, S.; Abideen, S.A.; Shaker, B.; Alharbi, M.; Alshammari, A.; Ali, M.; Haleem, A.; et al. Virtual Screening of *Artemisia Annua* Phytochemicals as Potential Inhibitors of SARS-CoV-2 Main Protease Enzyme. *Molecules* **2022**, *27*, 8103. [[CrossRef](#)] [[PubMed](#)]
39. Rieder, S.R.; Ries, B.; Schaller, K.; Champion, C.; Barros, E.P.; Hünenberger, P.H.; Riniker, S. Replica-Exchange Enveloping Distribution Sampling Using Generalized AMBER Force-Field Topologies: Application to Relative Hydration Free-Energy Calculations for Large Sets of Molecules. *J. Chem. Inf. Model.* **2022**, *62*, 3043–3056. [[CrossRef](#)] [[PubMed](#)]
40. Fratev, F.; Sirimulla, S. An Improved Free Energy Perturbation FEP+ Sampling Protocol for Flexible Ligand-Binding Domains. *Sci. Rep.* **2019**, *9*, 16829. [[CrossRef](#)] [[PubMed](#)]
41. Tiwari, P.; Sharma, P.; Kumar, M.; Kapil, A.; Abdul Samath, E.; Kaur, P. Identification of Novel Natural MurD Ligase Inhibitors as Potential Antimicrobial Agents Targeting *Acinetobacter Baumannii*: In Silico Screening and Biological Evaluation. *J. Biomol. Struct. Dyn.* **2022**, *40*, 14051–14066. [[CrossRef](#)] [[PubMed](#)]
42. Honarparvar, B.; Govender, T.; Maguire, G.E.M.; Soliman, M.E.S.; Kruger, H.G. Integrated Approach to Structure-Based Enzymatic Drug Design: Molecular Modeling, Spectroscopy, and Experimental Bioactivity. *Chem. Rev.* **2014**, *114*, 493–537. [[CrossRef](#)] [[PubMed](#)]
43. Zschau, R.L.; Zacharias, M. Mechanism of B-hairpin Formation in AzoChignolin and Chignolin. *J. Comput. Chem.* **2023**, *44*, 988–1001. [[CrossRef](#)] [[PubMed](#)]
44. Justino, G.C.; Nascimento, C.P.; Justino, M.C. Molecular Dynamics Simulations and Analysis for Bioinformatics Undergraduate Students. *Biochem. Mol. Biol. Educ.* **2021**, *49*, 570–582. [[CrossRef](#)] [[PubMed](#)]
45. Brandsdal, B.O.; Österberg, F.; Almlöf, M.; Feierberg, I.; Luzhkov, V.B.; Åqvist, J. Free Energy Calculations and Ligand Binding. *Adv. Protein Chem.* **2003**, *66*, 123–158. [[PubMed](#)]
46. Gul, S.; Ozcan, O.; Asar, S.; Okyar, A.; Baris, I.; Kavakli, I.H. In Silico Identification of Widely Used and Well-Tolerated Drugs as Potential SARS-CoV-2 3C-like Protease and Viral RNA-Dependent RNA Polymerase Inhibitors for Direct Use in Clinical Trials. *J. Biomol. Struct. Dyn.* **2021**, *39*, 6772–6791. [[CrossRef](#)] [[PubMed](#)]
47. Miller III, B.R.; McGee Jr, T.D.; Swails, J.M.; Homeyer, N.; Gohlke, H.; Roitberg, A.E. MMPBSA. Py: An Efficient Program for End-State Free Energy Calculations. *J. Chem. Theory Comput.* **2012**, *8*, 3314–3321. [[CrossRef](#)] [[PubMed](#)]
48. Woods, C.J.; Malaisree, M.; Michel, J.; Long, B.; McIntosh-Smith, S.; Mulholland, A.J. Rapid Decomposition and Visualisation of Protein–Ligand Binding Free Energies by Residue and by Water. *Faraday Discuss.* **2014**, *169*, 477–499. [[CrossRef](#)] [[PubMed](#)]

49. Bergström, C.A.S.; Larsson, P. Computational Prediction of Drug Solubility in Water-Based Systems: Qualitative and Quantitative Approaches Used in the Current Drug Discovery and Development Setting. *Int. J. Pharm.* **2018**, *540*, 185–193. [\[CrossRef\]](#) [\[PubMed\]](#)
50. Montgomerie, S.; Sundararaj, S.; Gallin, W.J.; Wishart, D.S. Improving the Accuracy of Protein Secondary Structure Prediction Using Structural Alignment. *BMC Bioinform.* **2006**, *7*, 301. [\[CrossRef\]](#) [\[PubMed\]](#)
51. Tanwar, H.; Kumar, D.T.; Doss, C.G.P.; Zayed, H. Bioinformatics Classification of Mutations in Patients with Mucopolysaccharidosis IIIA. *Metab. Brain Dis.* **2019**, *34*, 1577–1594. [\[CrossRef\]](#) [\[PubMed\]](#)
52. Wang, J.; Yang, W.; Zhao, L.; Wei, B.; Chen, J. Binding Mechanism of Inhibitors to BRD4 and BRD9 Decoded by Multiple Independent Molecular Dynamics Simulations and Deep Learning. *Molecules* **2024**, *29*, 1857. [\[CrossRef\]](#) [\[PubMed\]](#)
53. Henderson, R.; Anasti, K.; Manne, K.; Stalls, V.; Saunders, C.; Bililign, Y.; Williams, A.; Bubphamala, P.; Montani, M.; Kachhap, S. Engineering Immunogens That Select for Specific Mutations in HIV Broadly Neutralizing Antibodies. *bioRxiv* **2023**. [\[CrossRef\]](#)
54. Mortier, J.; Rakers, C.; Bermudez, M.; Murgueitio, M.S.; Riniker, S.; Wolber, G. The Impact of Molecular Dynamics on Drug Design: Applications for the Characterization of Ligand–Macromolecule Complexes. *Drug Discov. Today* **2015**, *20*, 686–702. [\[CrossRef\]](#) [\[PubMed\]](#)
55. Ahmed, S.A.; Ali, N.; Qureshi, U.; Khalil, R.; Qasmi, Z.-U.H. Molecular Dynamics Simulation of Human Pancreatic Lipase and Lipase-Colipase Complex: Insight into the Structural Fluctuations and Conformational Changes. *Int. J. Comput. Theor. Chem.* **2020**, *8*, 19. [\[CrossRef\]](#)
56. Elmaaty, A.A.; Darwish, K.M.; Khatat, M.; Elhady, S.S.; Salah, M.; Hamed, M.I.A.; Al-Karmalawy, A.A.; Saleh, M.M. In a Search for Potential Drug Candidates for Combating COVID-19: Computational Study Revealed Salvianolic Acid B as a Potential Therapeutic Targeting 3CLpro and Spike Proteins. *J. Biomol. Struct. Dyn.* **2022**, *40*, 8866–8893. [\[CrossRef\]](#) [\[PubMed\]](#)
57. Sun, Z.; Liu, Q.; Qu, G.; Feng, Y.; Reetz, M.T. Utility of B-Factors in Protein Science: Interpreting Rigidity, Flexibility, and Internal Motion and Engineering Thermostability. *Chem. Rev.* **2019**, *119*, 1626–1665. [\[CrossRef\]](#) [\[PubMed\]](#)
58. Gewers, F.L.; Ferreira, G.R.; De Arruda, H.F.; Silva, F.N.; Comin, C.H.; Amancio, D.R.; Costa, L.d.F. Principal Component Analysis: A Natural Approach to Data Exploration. *ACM Comput. Surv.* **2021**, *54*, 1–34. [\[CrossRef\]](#)
59. Yoo, C.; Shahlai, M. The Applications of PCA in QSAR Studies: A Case Study on CCR5 Antagonists. *Chem. Biol. Drug Des.* **2018**, *91*, 137–152. [\[CrossRef\]](#)
60. Karnati, K.R.; Wang, Y. Structural and Binding Insights into HIV-1 Protease and P2-Ligand Interactions through Molecular Dynamics Simulations, Binding Free Energy and Principal Component Analysis. *J. Mol. Graph. Model.* **2019**, *92*, 112–122. [\[CrossRef\]](#) [\[PubMed\]](#)
61. Chaudhary, K.; Kumar, K.; Venkatesu, P.; Masram, D.T. Protein Immobilization on Graphene Oxide or Reduced Graphene Oxide Surface and Their Applications: Influence over Activity, Structural and Thermal Stability of Protein. *Adv. Colloid Interface Sci.* **2021**, *289*, 102367. [\[CrossRef\]](#) [\[PubMed\]](#)
62. Jamal, S.B.; Ismail, S.; Yousaf, R.; Qazi, A.S.; Iftkhar, S.; Abbasi, S.W. Exploring Novel 1-Hydroxynaphthalene-2-Carboxanilides Based Inhibitors Against C-Jun N-Terminal Kinases Through Molecular Dynamic Simulation and WaterSwap Analysis. *Appl. Biochem. Biotechnol.* **2023**, *196*, 1803–1819. [\[CrossRef\]](#) [\[PubMed\]](#)
63. Panday, S.K.; Ghosh, I. In Silico Structure-Based Prediction of Receptor–Ligand Binding Affinity: Current Progress and Challenges. In *Structural Bioinformatics: Applications in Preclinical Drug Discovery Process. Challenges and Advances in Computational Chemistry and Physics*; Springer: Cham, Switzerland, 2019; pp. 109–175.
64. Kirby, J.P. *The Amidinium-Carboxylate Salt Bridge and Electron Transfer Reactions*; Michigan State University: East Lansing, MI, USA, 1997; ISBN 0591744015.
65. Gupta, M.N.; Uversky, V.N. Biological Importance of Arginine: A Comprehensive Review of the Roles in Structure, Disorder, and Functionality of Peptides and Proteins. *Int. J. Biol. Macromol.* **2023**, *257*, 128646. [\[CrossRef\]](#) [\[PubMed\]](#)
66. Pakhrin, S.C.; Shrestha, B.; Adhikari, B.; Kc, D.B. Deep Learning-Based Advances in Protein Structure Prediction. *Int. J. Mol. Sci.* **2021**, *22*, 5553. [\[CrossRef\]](#) [\[PubMed\]](#)
67. Gomes, D.; Silvestre, S.; Duarte, A.P.; Venuti, A.; Soares, C.P.; Passarinha, L.; Sousa, Â. In Silico Approaches: A Way to Unveil Novel Therapeutic Drugs for Cervical Cancer Management. *Pharmaceuticals* **2021**, *14*, 741. [\[CrossRef\]](#) [\[PubMed\]](#)
68. Al-Thakafi, S.; Al-Hathal, N. Peyronie’s Disease: A Literature Review on Epidemiology, Genetics, Pathophysiology, Diagnosis and Work-Up. *Transl. Androl. Urol.* **2016**, *5*, 280. [\[CrossRef\]](#) [\[PubMed\]](#)
69. Alipour, H.; Raz, A.; Zakeri, S.; Djadid, N.D. Therapeutic Applications of Collagenase (Metalloproteases): A Review. *Asian Pac. J. Trop. Biomed.* **2016**, *6*, 975–981. [\[CrossRef\]](#)
70. Supuran, C.T.; Scozzafava, A.; Mastrolorenzo, A. Bacterial Proteases: Current Therapeutic Use and Future Prospects for the Development of New Antibiotics. *Expert Opin. Ther. Pat.* **2001**, *11*, 221–259. [\[CrossRef\]](#)
71. Nitulescu, G.; Nitulescu, G.M.; Zangfirescu, A.; Mihai, D.P.; Gradinaru, D. Candidates for Repurposing as Anti-Virulence Agents Based on the Structural Profile Analysis of Microbial Collagenase Inhibitors. *Pharmaceutics* **2021**, *14*, 62. [\[CrossRef\]](#) [\[PubMed\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Bayesian Approach to Stochastic Estimation of Population Survival Curves in Chile Using ABC Techniques and Its Impact over Social Structures

Rolando Rubilar-Torrealba ¹ , Karime Chahuán-Jiménez ^{2,*} , Hanns de la Fuente-Mella ³ 
and Claudio Elórtogui-Gómez ⁴ 

¹ Departamento de Industrias, Universidad Técnica Federico Santa María, Valparaíso 2090123, Chile; rolando.rubilar@usm.cl

² Escuela de Auditoría, Centro de Investigación en Negocios y Gestión Empresarial, Universidad de Valparaíso, Valparaíso 2361891, Chile

³ Instituto de Estadística, Facultad de Ciencias, Pontificia Universidad Católica de Valparaíso, Valparaíso 2340031, Chile; hanns.delafuente@pucv.cl

⁴ Escuela de Periodismo, Facultad de Ciencias Económicas y Administrativas, Pontificia Universidad Católica de Valparaíso, Valparaíso 2373223, Chile; claudio.elortegui@pucv.cl

* Correspondence: karime.chahuan@uv.cl

Abstract: In Chile and worldwide, life expectancy has consistently increased over the past six decades. Thus, the purpose of this study was to identify, measure, and estimate the population mortality ratios in Chile, mortality estimates are used to calculate life expectancy when constructing life tables. The Bayesian approach, specifically through Approximate Bayesian Computation (ABC) is employed to optimize parameter selection for these calculations. ABC corresponds to a class of computational methods rooted in Bayesian statistics that could be used to estimate the posterior distributions of the model parameters. For this research, ABC was applied to estimate the mortality ratios in Chile, using information available from 2004 to 2021. The results showed heterogeneity in the results when selecting the best model. Additionally, it was possible to generate projections for the next 10 years for the series analysed in the research. Finally, the main contribution of this research is that we measured and estimated the population mortality rates in Chile, defining the optimal selection of parameters, in order to contribute to creating a link between social and technical sciences for the advancement and implementation of current knowledge in the field of social structures.

Keywords: mortality; older population; life expectancy; stochastic processes; Bayesian analysis



Citation: Rubilar-Torrealba, R.; Chahuán-Jiménez, K.; de la Fuente-Mella, H.; Elórtogui-Gómez, C. Bayesian Approach to Stochastic Estimation of Population Survival Curves in Chile Using ABC Techniques and Its Impact over Social Structures. *Computation* **2024**, *12*, 154. <https://doi.org/10.3390/computation12080154>

Academic Editors: Minzhang Zheng and Pedro D. Manrique

Received: 2 July 2024

Revised: 19 July 2024

Accepted: 25 July 2024

Published: 29 July 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Health is increasingly seen as a critical aspect of economic growth, and tracking healthy life expectancy is essential for assessing the financial sustainability of health and social care systems [1]. The authors of [2] found that life expectancy was more frequently incorporated into screening and treatment guidelines as evidence to facilitate clinical decision making related to the patient, and that life expectancy incorporating health status differed substantially from the life tables and health status for standard United States cases.

According to [3], the increase in life expectancy in high-income countries in Asia-Pacific is largely due to a shift in mortality from noncommunicable diseases to old age, which has led to rapid population aging and calls for policies that focus on healthy years in old age. On the other hand, projections of life expectancy in Western countries indicate that life expectancy will indeed continue to increase in most Western countries, but an important open question is whether longer life expectancy will be associated with more or fewer years of ill health. It is, therefore, useful to supplement projections of life expectancy with projections of health expectancy. The authors of [4] point out that when choosing a policy, analytical and applied factors should be considered, as well as their methods

of calculation for life expectancy and mortality tables, focusing on needs, audience and communication strategy; community values and norms to explicitly address inequalities; and analytical capacity.

Demographic change, with the aging of societies, and epidemiological change, with an increase in chronic diseases, is unquestionable globally. Among the consequences of these changes is the increase in multimorbidity, fragility, and, on the other hand, the increase in life expectancy—which means an increase in demand for health and social services. This scenario involves complexity in the management and organization of care, which require a structural and functional adaptation of the available health resources if we want to conserve them and make them sustainable. However, the effects of decision making in the medium- and long-term in health systems pose a great challenge without the help of technology, that is, without the use of computational tools and mathematical models. This is not a new challenge—scientists from different areas of knowledge have questioned the feasibility and sustainability of health systems for several decades. In this context, it is important to highlight the link between social and technical sciences for the advancement and implementation of current knowledge in this field.

Monitoring life expectancy in small areas is important for health policy, but poses difficulties for conventional life table analysis [5]. Moreover, the implicit model underlying conventional life table analysis involves a parameterised fixed-effects approach. According to the social class ranking, there is a clear relationship between educational attainment and life expectancy. The social class rank is even higher for health expectancy than for life expectancy.

To point out that there is a gap in the demographic literature associated with the interpretation of periodic trends in life expectancy, the tempo [6] effects are presented; this is the case when a large number of deaths are suddenly postponed. Under the above conditions, the life table increases the longevity of the population by weighting deferred deaths with the remaining life expectancy. The study by [7] shows that life expectancy at age 50 differs by 8.6 years for men and 5.5 years for women between the highest and lowest income quintiles. The difference in disability-free life expectancy is 12.8 and 11.0 years for men and women, respectively. The mortality effect contributes equally to the difference in disability-free life expectancy for men and slightly more for women.

The research by [8] shows that women live longer, but spend more years in poor health than men. The same authors emphasize the need to consider gender differences in the demand for health care. It is also important to consider the need for policies that increase the number of years that older people can live in good health. The results of [9] expose the existence of ethno-racial inequalities in life expectancy in Chile, and show that Mapuche indigenous peoples are more disadvantaged in terms of survival than other indigenous and non-indigenous groups, with implications for decision making on ethnic policies.

In the case of life expectancy, current health policies need to take into account the growing number of beneficiaries with multiple chronic diseases when determining population projections and funding solvency [10]. According to [11], life expectancy is a measure used to compare differences in health between populations. The use of combined flexible parametric methods to estimate life expectancy in small samples has yielded promising results, allowing for the modelling of life expectancy by exact age, with a higher statistical precision and lower bias, and a prediction of different patterns of covariates without stratification.

In terms of life risk measurement, [12] indicates that each individual has one life and that this life has an age-specific mortality risk. If one adjusts this mortality risk of an individual with their age-specific probability of living, the life expectancy would change.

Research by [13] shows that the life expectancy curves of homogeneous and inhomogeneous populations intersect at the average age at which the frailest would have died and at the maximum age lived in the population (when the last members of the cohort would have died).

An analysis of the life tables suggests that some immigrant groups have different life expectancies because the standardized mortality ratio is sensitive to demographic differences between the groups being compared [11]. For [14], the methods associated

with life tables do not usually involve an explicit statistical model, but the underlying statistical densities are applied when estimates of standard errors are required. The authors propose a Bayesian random-effects method [15–18] that aggregates the force into areas, ages, and groups using random-effects methods that do the following: (a) identify correlations between neighboring areas and ages and (b) detect similarities in spatial and age effects between stratified demographic groups (e.g., gender).

One of the first attempts to model human mortality was that of the English actuary Benjamin Gompertz, who argued that, above a certain age, the logarithm of the mortality force is a linear function of age. This hypothesis, later called Gompertz's law, has been widely used in demographic and actuarial projections of mortality over the last two centuries [19]. The authors of [20] applied the Gompertz–Makeham mortality model framework relationship commonly used in human and wildlife studies, assuming that the Gompertz rate parameter is affected by individual heterogeneity, and show that in the results of a simulation study, the model adequately recovers the parameters used for the simulation.

According to [21], mortality models can be divided into static (functions of age only) versus dynamic (functions of age and current year), and deterministic versus stochastic models. The authors considered the Gompertz and Makeham models, which are deterministic and static, and the Lee–Carter method, which predicts mortality stochastically as a random walk with drift. In addition, the authors pointed out that models based on Gompertz and Makeham are widely used for educational, forecasting and risk assessment purposes, and the most significant recent advances in mortality modeling and forecasting: the Lee–Carter method.

Mortality estimates are used to construct life expectancies in the construction of life tables. When comparing different estimates of life expectancy, the results of [22] show that the Lee–Carter model without the long-memory component can provide underestimates of life expectancy, while extensions to the structure of the long-memory model reduce this effect. However, the empirical results show that the estimation using the Gompertz–Makeham methodology as opposed to the Lee–Carter methodology is similar [21]; for this reason, the widely known Gompertz–Makeham methodology is used in this research.

According to [23], mortality improvement is a challenge for public pension planning and for the private annuity business. Anticipating future mortality is important for public policy as well as for the management of financial institutions. As population mortality declines, national social security systems and insurance companies in most developed countries are reassessing their mortality tables to reflect longevity risk [24].

Taking into account life expectancy and its impact on government policies, as well as the existing life table, this research aims to estimate the mortality curves of the male and female population, based on projections of the parameters that determine the mortality of each cohort, using the ABC (Approximate Bayesian Computation) method. The ABC approach provides us with a flexible methodology that can be adapted to highly complex estimation problems [6,25–29], such as estimating the mortality law of a population. In this regard, we can test different models of population mortality behavior without significantly altering the structure of the methodology that will be proposed in this research work.

This manuscript is divided into four main sections, which include the Materials and Methods, the Results, the Discussion and, finally, the Conclusions.

2. Materials and Methods

In this research, we used official data from the Chilean government on vital statistics obtained from <https://www.ine.gob.cl/estadisticas/sociales/demografia-y-vitales/nacimientos-matrimonios-y-defunciones> (accessed on 29 April 2024). Specifically, we utilized statistics on the number of deaths for each cohort, with information available from 2004 to 2021, which considered the consolidated official data for the estimates and projections to be made.

Figure 1 shows the evolution of life expectancy over time for the Chilean population. We can observe that life expectancy has consistently increased over the past six decades. This growth in life expectancy is not exclusive to Chile, it is a trend observed worldwide. The reasons for the improvement in life expectancy are centered around sustained improvements in quality of life and increased spending on healthcare, among other social variables that have been extensively studied in the literature [30,31].

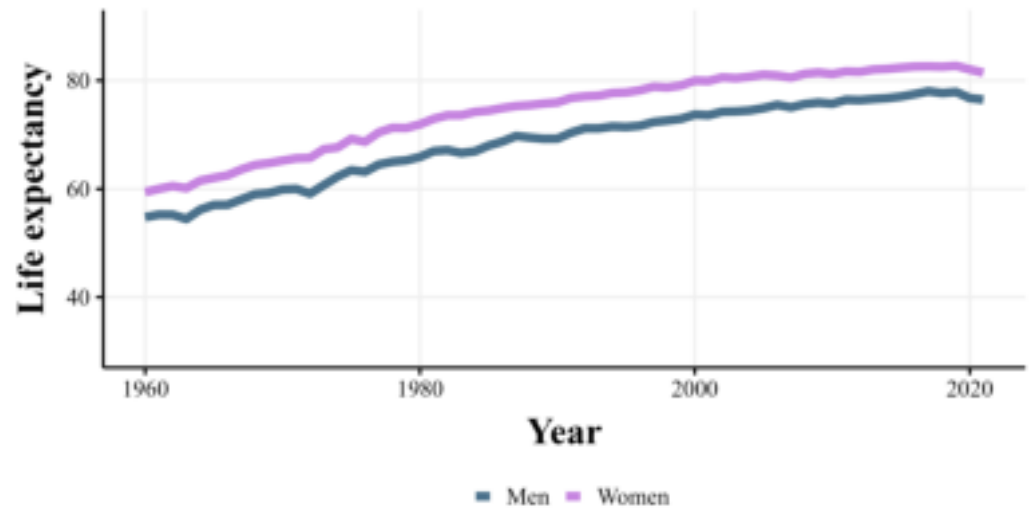


Figure 1. Evolution of life expectancy in Chile.

Focusing on life expectancy raises the need to model survival for each specific cohort and thus generate various public policies oriented towards demographic aspects. However, if we extend the static analysis to a temporal analysis, this prompts questions about the stability of the parameters that define population mortality and creates an opportunity to develop methodological improvements for projecting population life expectancy and its evolution over time.

2.1. Gompertz–Makeham Law

One of the most well-known and widely used models for estimating population mortality rates corresponds to those originally developed by Gompertz [32] and Makeham [33], and is constructed by using the instantaneous force of mortality $\lambda(i)$, where i corresponds to the age of the specific cohort. Estimating the Gompertz law requires the estimation of two parameters that can be estimated using standard statistical techniques related to the evolution of the age of the population. However, Makeham’s development adds a third parameter in the estimation process, which requires numerical approximation techniques but yields a better fit to population data by considering an extrinsic mortality effect [34,35]. In this research, we use a version of the model of Gompertz–Makeham, which we describe as

$$\lambda(i) = C + \frac{e^{(i-m)/b}}{b}, \quad (1)$$

where parameter C corresponds to the extrinsic effect of mortality, b corresponds to an exponential adjustment constant, and m corresponds to a factor that we can interpret as a population decay law over time and that is related to the life expectancy of the population. For further references on modeling the Gompertz–Makeham law, see [36,37].

For the calculation of the probability of death in the following period $q_i^t = 1 - p_i^t$, where $p_i^t = \Pr(T_i \geq t)$, we use the following definition of the survival function

$$p_i^t = \exp\left(-\int_i^{i+t} \lambda(s)ds\right). \quad (2)$$

Proposition 1. Under Gompertz–Makeham’s mortality law (Equation (1)), with $C = 0$, the survival probability to period t is

$$p_i^t = \exp\left\{e^{(i-m)/b}\left(1 - e^{t/b}\right)\right\}. \quad (3)$$

Proof. From Equations (1) and (2), we have the following definition

$$p_i^t = \exp\left\{-\int_i^{i+t}\left(C + \frac{e^{(s-m)/b}}{b}\right)ds\right\}.$$

Solving the definite integral for the above equation, we have

$$p_i^t = \exp\left\{-Ct + \left(e^{(i-m)/b}\right)\left(1 - e^{t/b}\right)\right\}.$$

Now, assuming $C = 0$, the survival probability is

$$p_i^t = \exp\left\{e^{(i-m)/b}\left(1 - e^{t/b}\right)\right\}.$$

□

Figure 2 shows the evolution of the population decay under different parameterizations, starting with an initial age of 50 years. This figure illustrates how the population declines as age increases and it enables the projection of the population’s age composition, facilitating the establishment of specific demographic policies tailored to the social reality.

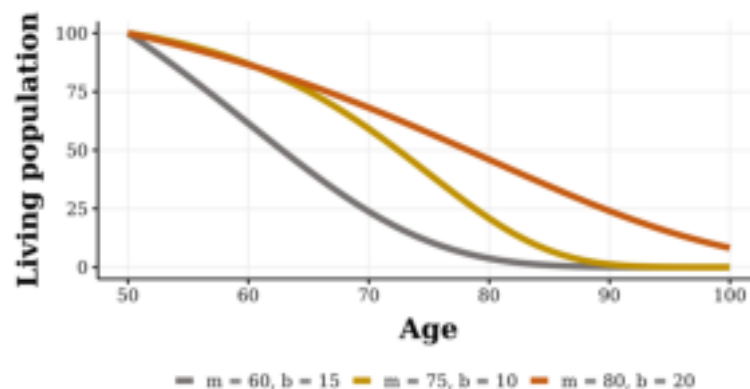


Figure 2. Different survival curves with synthetic data.

Within the essential elements for estimating the parameters that define the Gompertz–Makeham law, it is necessary to understand the mortality dynamics of different population cohorts. With mortality data from various cohorts, we generate adjustments and estimates of the parameter values for the Equation (3).

2.2. Mortality Death Rates

The mortality death rate tables are used in various applications such as the estimation of life expectancy and calculation of life pension amounts, among many other applications. The mortality death rate tables correspond to a matrix of values belonging to a specific cohort i at a specific time t , which allows us to understand the temporal evolution of death rates. We define $X_{i,t} = (x_{1,i,t}, x_{2,i,t}, \dots, x_{N_i,i,t})$ as a vector that represents the proportion of survivors in the following year for the number of people in the specific cohort i at a moment in time t . The probability density of each element of vector $X_{i,t}$ is defined as follows:

$$x_{n,i,t} = F(\Theta_{i,t}), \quad (4)$$

where $\Theta_{i,t} = (\theta_{1,i,t}, \theta_{2,i,t}, \dots, \theta_{j,i,t})$ corresponds to the set of parameters of cohort i in period t that define the stochastic behavior of the elements of vector $X_{i,t}$. We describe the evolution of the parameter values that define the density function as

$$\theta_{j,i,t} = H_j(i, t) + v_{i,t}, \quad (5)$$

where $H_j(\cdot)$ corresponds to a specific function of parameter j that depends on cohort i in period t ; $v_t \sim N(0, \sigma_v^2)$ corresponds to an error term with a normal distribution of variance σ_v^2 . This description allows us to understand the evolution of the parameter values that define the survival behavior of members of a specific cohort as a random phenomenon. As examples of functions that model the stochastic temporal evolution of parameters, we can mention [38,39], among other authors.

For this research, we assume that the function describing the survival of an individual from cohort i for the following year corresponds to a random variable that is Bernoulli distributed with parameter $g_{i,t}$, where parameter $g_{i,t}$ can be interpreted as the approximate proportion of individuals in cohort i who will die in the next period, which we describe as

$$x_{j,i,t} \sim \text{Bernoulli}(g_{i,t}). \quad (6)$$

On the other hand, we propose that parameter $g_{i,t}$ evolves over time according to an autoregressive integrated moving average model (ARIMA) with respect to the same cohort. In this sense, the autoregressive factor of the same cohort corresponds to the natural survival factor at a certain age and the moving average is related to the behaviour of the error term. We describe the ARIMA(S, 0, Z) model and the ARIMA(S, 1, Z) model in the following equations

$$g_{i,t} = \mu_i + \sum_{s=1}^S \phi_{i,s} g_{i,t-s} + \sum_{z=1}^Z \kappa_{i,z} v_{i,t-z} + v_{i,t}, \quad (7)$$

$$\Delta g_{i,t} = \sum_{s=1}^S \phi_{i,s} \Delta g_{i,t-s} + \sum_{z=1}^Z \kappa_{i,z} v_{i,t-z} + v_{i,t}. \quad (8)$$

This formulation allows us to understand the evolution of parameters over time and helps us comprehend their stability as a random phenomenon. Another important aspect is that it provides us with an interesting tool for projecting individual survival and, consequently, the age composition of society. One element to consider corresponds to the limitations of time series techniques in forecasting, especially when extending over a very large time span.

2.3. Estimation of Parameters Using ABC Techniques

In this section, we define the parameter estimation procedure for the Gompertz law using techniques based on Bayesian inference. In Bayesian analysis, the unknown parameter that defines the mortality behavior of the population is represented as a random variable η with a probability distribution $\pi(\eta)$, known as the prior distribution. The prior distribution encapsulates the initial beliefs about the parameter values held by the researcher, which must be updated based on the evidence revealed by the observed data.

Given a set of parameters η , the observed data $X = x_1, x_2, \dots, x_n$ follows a density function $f(X|\eta)$ that defines a parametric model with the parameters value η , which determines the behavior of the random variable. With this in mind, we can define the joint distribution as

$$f(X, \eta) = f(X|\eta)\pi(\eta), \quad (9)$$

and the marginal density of X is defined as follows:

$$f(X) = \int f(X, \eta) d\eta = \int f(X|\eta) f(\eta) d\eta. \quad (10)$$

The conditional density of η given the observed values of X corresponds to $f(\eta|X)$ and is known as the posterior distribution. The posterior distribution is defined as

$$\text{Posterior density} \propto \text{Likelihood} \times \text{Prior density}. \quad (11)$$

Within modern techniques for estimating the posterior distributions of parameters that define random phenomena, we can name methods like Approximate Bayesian Computation (ABC). In the literature, notable mentions include [6,25–29]. ABC techniques comprise a series of acceptance–rejection algorithms that utilize summary statistics from a random sample clustered around a target value for the estimation of the posterior distribution of parameters. These techniques have demonstrated significant flexibility due to the ease of generating independent random samples and allow for approximations of the true parameter distributions when analytical estimation is complex or not feasible.

Algorithm 1 shows how to apply ABC techniques for estimating population decay, using Equation (3) as a basis. The algorithm takes the following requirements to start: Nm which corresponds to the number of samples needed to calculate a probability density of the parameters; ψ_m , which corresponds to the a priori density of parameter m ; ψ_b , which corresponds to the a priori density of parameter b ; *initial*, which corresponds to the initial age at which the population decline is to be estimated; *terminal*, which corresponds to the terminal age for the assessment of population decline; *Mortality*, which corresponds to the annual mortality on which the estimate will be made, which can be obtained from the observed mortality data of a population or its projection; and ϵ , which corresponds to the error value that we are willing to accept for the generation of the probability density of the parameters.

Algorithm 1 Modelling of population decay.

Require: $N, \psi_m, \psi_b, \text{initial}, \text{terminal}, \text{Mortality}, \epsilon$.

$j \leftarrow 0$

while $j < N$ **do**

$m \leftarrow \text{Uniform}(0, \psi_m)$

$b \leftarrow \text{Uniform}(0, \psi_b)$

$\text{population} \leftarrow 1$

$\text{popsim} \leftarrow 1$

for $i = \text{initial}$ **to** $(\text{terminal} - 1)$ **do**

$\text{population}[i + 1] \leftarrow \text{population}[i] \times (1 - \text{Mortality}[i])$

$\text{popsim}[i + 1] \leftarrow \text{popsim}[i] \times (1 - \exp\{e^{(i-m)/b} (1 - e^{1/b})\})$

end for

if $\sum_{i=\text{initial}}^{\text{terminal}} |\text{population}[i] - \text{popsim}[i]| < \epsilon$ **then**

$M[j] \leftarrow m$

$B[j] \leftarrow b$

$j \leftarrow j + 1$

end if

end while

One of the most important steps in Bayesian methodologies is the selection of the prior m and b parameters. An appropriate selection of the prior improves the estimation results of the posterior distribution [40] and accelerates computational calculations, making the use of these methodologies feasible in many cases. However, often, there is no information regarding the best prior to be used, and a non-informative distribution such as a uniform distribution must be used, assuming the possibility of biases in the results that must be

corrected for by proper selection of the prior distribution. In our procedure, we suggest using a uniform distribution for the characterization of the prior, as we do not have information on the behavior of all mortality parameters as a whole.

In the following step, it is necessary to randomly select values for m and b , using the priors determined for the parameters. Subsequently, the decay of real mortality and that estimated based on the selected parameters is calculated. Once the decays, both real and based on the selected parameters, are obtained, the difference is measured. If the difference is below the ϵ value, it is accepted and incorporated into the M and B vectors. Finally, we can calculate the posterior density using the M and B vectors.

Figure 3 shows a summary of the methodology used for estimating the survival curves. First, it is necessary to obtain historical mortality data of the population. This allows us to estimate the value of the mortality parameters for each cohort for each year, thus obtaining the time series of the estimated parameter values. Second, we proceed to project the parameter values for a certain number of years, based on the methodology described in Equations (7) and (8). Third, we use the projections of each cohort as inputs for Algorithm 1, thereby obtaining the parameter densities used to calibrate the mortality curve described in Equation (3). Finally, we proceed to obtain random values for the parameters in Equation (3) using the calculated densities, in order to obtain the survival projection of the population. This provides a survival density for each of the projected ages, which can be used for some social application in the context of population projection.

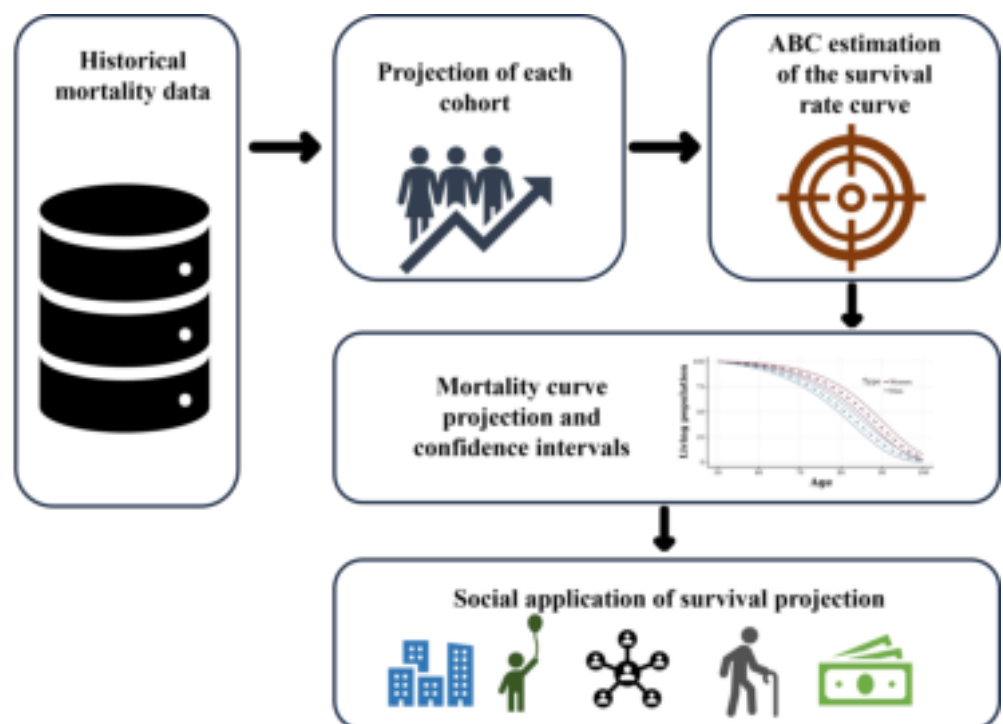


Figure 3. Summary of the methodology used in the research.

3. Results

In this section, we proceed to present the results of the methodology proposed in the previous sections. Firstly, we display the results in Table 1 for some of the regressions performed when estimating the projections of the parameter evolution over time, based on the general expression of Equations (7) and (8) for 2030. In this case, we selected the ARIMA model with the lowest AIC criteria, which allows for combining a model with low variance and a limited number of parameters to be estimated, a fundamental element due to the limited number of observations for each cohort. Columns AR, I, and MA correspond to the number of autoregressive, integrated, and moving average parameters in the model. We also incorporate the information of the AIC and BIC values for the selected models.

Table 1. Parameter estimation for a selected sample of cohorts.

		AR	I	MA	AIC	BIC
Female	Age 50	0	1	0	−223.06	−222.29
	Age 55	0	0	0	−241.18	−239.51
	Age 60	0	1	0	−193.29	−192.81
	Age 65	0	1	0	−195.04	−194.27
	Age 70	0	1	0	−179.70	−178.92
	Age 75	0	1	0	−158.23	−157.46
	Age 80	0	1	0	−138.95	−138.17
	Age 85	0	1	0	−121.86	−121.09
	Age 90	0	1	0	−96.52	−95.75
Male	Age 50	0	0	1	−218.45	−215.95
	Age 55	0	0	0	−203.69	−202.03
	Age 60	0	1	0	−176.22	−175.45
	Age 65	0	1	0	−172.06	−171.28
	Age 70	0	1	0	−155.15	−154.38
	Age 75	0	1	0	−135.20	−134.42
	Age 80	0	1	0	−129.22	−128.45
	Age 85	0	1	0	−102.55	−101.78
	Age 90	1	1	0	−89.32	−87.77

The results show heterogeneity of the results for the selection of the best model. In some cases, eliminating the trend using differencing is sufficient to obtain a stationary process, while in others it is necessary to consider the autoregressive and moving average components to obtain a stationary process.

Based on the selected models, we then generate projections for the next 10 years, which gives us the possible results for the parameter that defines the mortality of each cohort value with 95% confidence. This projection is shown in Figure 4, which considers projections for men and women aged 50, 70, and 90. In general, we can observe that for all projections, the mortality rate for men is higher than for women, indicating a significant difference in the age composition of the elderly population. With these projections, we can generate the survival curve described in Equation (3), using the methodology of parameter estimation using the ABC techniques described in Algorithm 1, which is a necessary element for the generation of social applications requiring the use of population projection.

Table 2 shows the estimation of the parameters for Equation (2) describing the behavior of population decay for men and women, obtained as the main results of Algorithm 1. The estimation uses the expected value of the projection of the parameters that determine the mortality rate for each cohort; the 5th percentile case corresponds to the best case scenario for each of the cohorts as it marks the lowest mortality scenario to which the population could be subjected and the 95th percentile case corresponds to the worst case mortality rate for each cohort. The columns identify the average value and standard deviation for parameters m and b for each of the cases. We can observe that the scenarios for parameter m for men range between 79.34 and 90.14, while for parameter b it ranges between 11.12 and 11.79; for women, parameter m ranges between 85.48 and 95.30 and parameter b between 9.57 and 11.30.

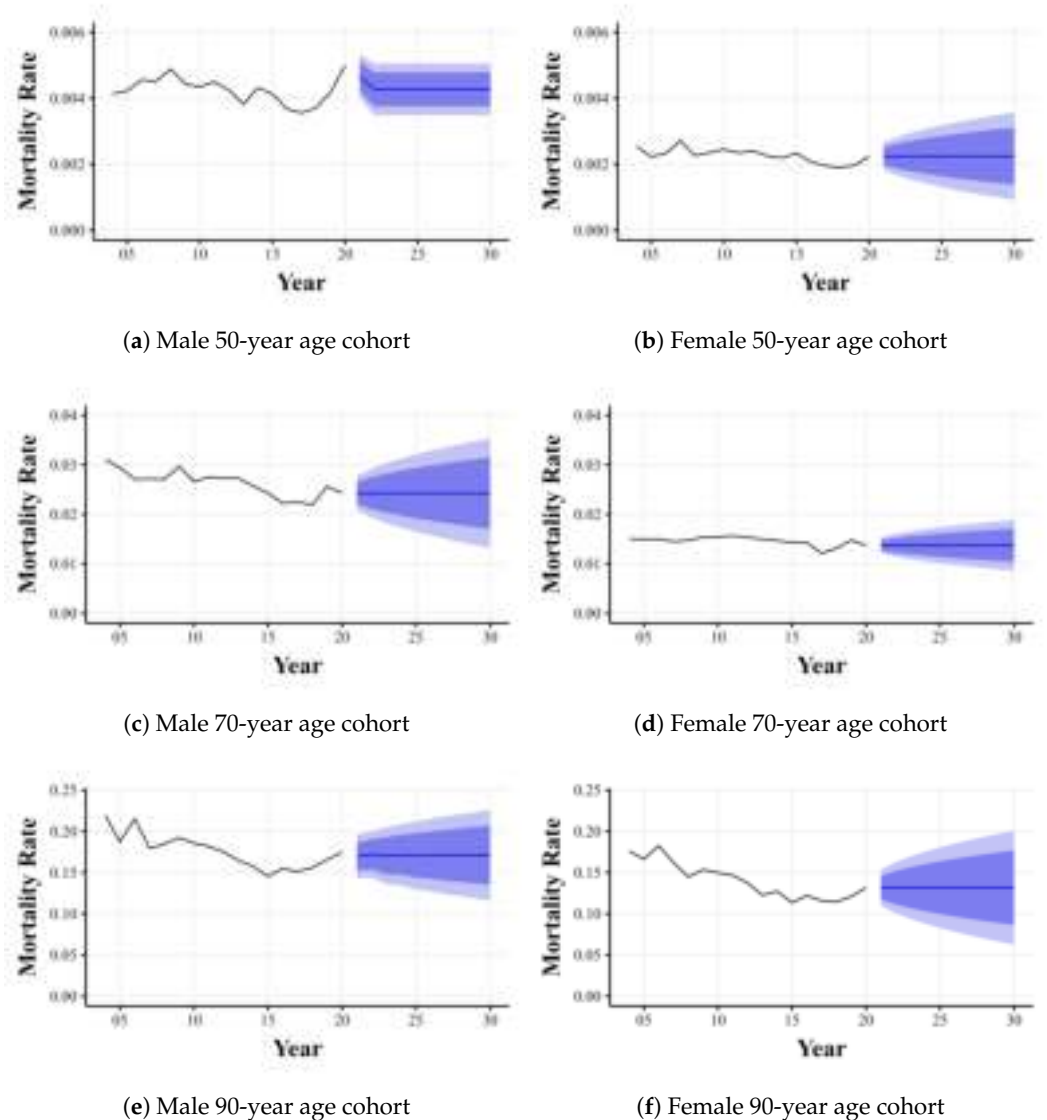


Figure 4. Projection of the parameter $\hat{p}$ for different ages and different sexes.

Table 2. Estimation of population decay parameters for men and women.

	m		b	
	Mean	Sd	Mean	Sd
Male mean	83.39	0.60	11.12	0.96
Male low	79.34	0.56	11.21	1.04
Male high	90.14	0.56	11.79	0.86
Female mean	88.92	0.64	9.95	0.79
Female low	85.48	0.62	9.57	0.77
Female high	95.30	0.96	11.30	1.08

Figure 5 shows the displacement of the survival curve for males and females, using the mean value of the parameters shown in Table 2 and compared with the estimated mortality curve for 2000 as a baseline. It is observed that for both cases, there is an improvement in life expectancy, which impacts the age composition of the elderly population and will have an important impact on other related elements, such as the estimation of pensions and health services, to name a few.

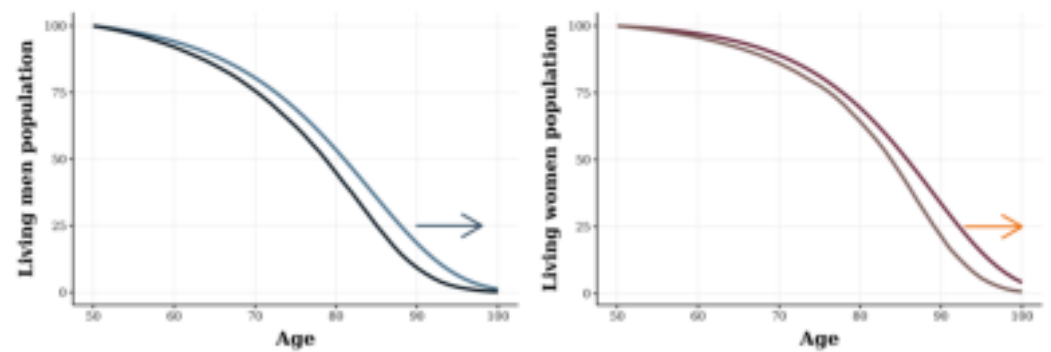


Figure 5. Evolution of the projection for the expected value of population decay. The shift to the right implies improvements in the life expectancy of the population.

Finally, we generate confidence bands for the population decay estimate using the values associated with the 5th percentile and 95th percentile of the projections of the parameters defined above, and we use Algorithm 1 for the estimation of their parameters. Figure 6 shows the population decay projection for males and females, with the respective confidence bands. It is worth mentioning that this procedure generates confidence bands, considering the “best” and “worst” case in terms of the estimation of the mortality rate parameters, which makes these confidence bands conservative scenarios in terms of the estimation of the population decay projection. For the generation of confidence bands considering specific percentiles, independent simulations must be performed for each of the cohorts. Then, the parameter estimation procedure defined in Algorithm 1 must be performed on a sufficiently large number of each of the cohorts to obtain the density of the decay projection.

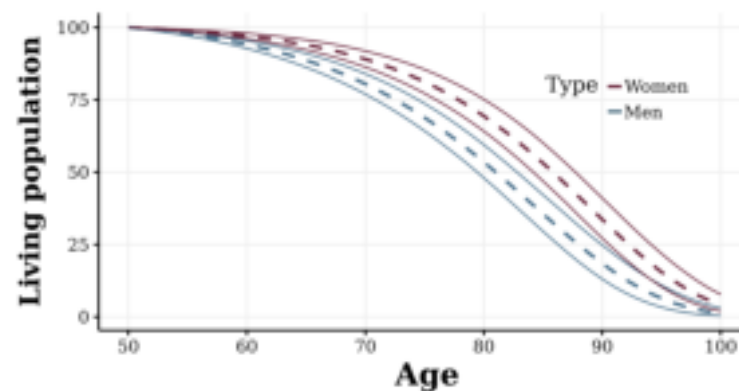


Figure 6. Estimation of population projections until 2030. The dotted lines correspond to the mean value of the projections.

Determination of Life Annuity

To understand how the proposed methodology significantly impacts various applications, we proceed with a simple example of determining a citizen’s life annuity. In this context, we consider a man who retires at the age of 65 with an accumulated amount of USD 100,000. Additionally, we use the 2010 mortality data as the baseline and contrast it with the 2030 population density projections generated and shown in the results. The life annuity of the citizen s is calculated as

$$life\_annuity_s = \frac{W_s / (1 + \gamma)}{a_i}, \quad (12)$$

where W_s corresponds to the level of wealth that individual s has at the time of retirement; γ corresponds to the commission of the insurance companies when offering the annuity,

which for this example we will take as zero; a_i corresponds the actuarial annuity factor (see [41]) at age i , which can be calculated as

$$a_i = \sum_{t=0}^{\infty} \frac{p_i^t}{(1+r)^t} \quad (13)$$

where p_i^t corresponds to the function defined in Equation (3) that defines the probability of survival in retirement age i at t years in the future, and r represents the effective annual valuation rate used to discount the cash flows, which for this example we consider a value of 3%.

Table 3 shows the estimated values for the calculation of the life annuity for a man retiring at age 65. The first and fifth columns correspond to the age of the individual; the second and sixth correspond to the discount factor used (3% per year); the third and seventh columns correspond to the estimated survival of the individual to survive to the age of the corresponding row, considering the mortalities of 2010; and, finally, the fourth and eighth columns are equivalent to the third and seventh, but for the projection of 2030.

Table 3. Values for life annuity estimation.

Age	$1/(1+i)^t$	$p_{i,2010}^t$	$p_{i,2030}^t$	Age	$1/(1+i)^t$	$p_{i,2020}^t$	$p_{i,2030}^t$
65	0.971	0.982	0.984	83	0.570	0.404	0.454
66	0.943	0.963	0.966	84	0.554	0.364	0.415
67	0.915	0.942	0.948	85	0.538	0.321	0.376
68	0.888	0.920	0.928	86	0.522	0.279	0.337
69	0.863	0.896	0.906	87	0.507	0.237	0.299
70	0.837	0.869	0.883	88	0.492	0.198	0.262
71	0.813	0.843	0.859	89	0.478	0.159	0.228
72	0.789	0.816	0.833	90	0.464	0.129	0.195
73	0.766	0.786	0.805	91	0.450	0.102	0.164
74	0.744	0.755	0.776	92	0.437	0.078	0.137
75	0.722	0.722	0.746	93	0.424	0.058	0.112
76	0.701	0.687	0.714	94	0.412	0.040	0.090
77	0.681	0.649	0.680	95	0.400	0.028	0.071
78	0.661	0.611	0.645	96	0.388	0.018	0.055
79	0.642	0.569	0.609	97	0.377	0.012	0.041
80	0.623	0.527	0.571	98	0.366	0.007	0.031
81	0.605	0.487	0.533	99	0.355	0.004	0.022
82	0.587	0.444	0.494	100	0.345	0.004	0.016

Finally, Table 4 shows the results of the calculation of the life annuity for the 65-year-old pensioner. The first column corresponds to the estimation of the actuarial annuity factor, considering the 2010 mortalities; the second column corresponds to the estimation of the actuarial annuity factor, but now considering the 2030 projection; the third and fourth columns correspond to the calculated annuity; and the fifth column shows the percentage variation between annuities. It is observed that as life expectancy improves, based on the proposed methodology, annuities decrease as people live longer. The amount must be distributed over a longer expected life span.

Table 4. Results of the impact of the evolution of mortality.

$a_{i,2010}$	$a_{i,2030}$	$life\_annuity_{2010}$	$life\_annuity_{2030}$	Variation %
11.887	12.571	8412	7955	−5.43%

4. Discussion

The debate on health and its various implications for the population has been an aspect of recurring and significant importance in citizen, political, and economic emergencies in different countries around the world.

The literature shows a growing need for the adequate estimation of life expectancy to try to project population dynamics and make decisions on public policy [42,43]. However, the implementation of new estimation methodologies provides a new framework on which to make these projections.

In this sense, the three stages of the methodology proposed in this research, (1) estimation of the population survival curve, (2) projection of the parameters that define the mortality of each cohort, and (3) estimation of the parameters using a Bayesian methodology, allow us to understand the projections of the survival curves as a random phenomenon, unlike most of the approximations that are made in the literature [36,37,44].

The appropriate estimation of survival curves presents an intrinsic problem due to the heterogeneity of the population. High-income countries have a longer life expectancy due to the treatment of diseases [3] and the case of rural populations versus urban populations [45], among other sources of variation, which makes it complex to have a single function that describes the evolution of the population as it becomes older.

The proposed methodology allows us to generate a variety of possible results for the population projection by means of the Bayesian methodology presented in this research and, therefore, absorbs technological changes that can cause drastic improvements in life expectancy, as well as catastrophic events that can induce a rise in mortality.

The results shown in the research show a difference in the estimation of the parameters that define the survival curve of men and women, showing a differentiated behavior that implies a new projection for the population pyramid, which are essential elements for the generation of gender public policies focused on the elderly. These general results imply a significant proportion of elderly women who will require specialized social services in the near future.

On the other hand, an intermediate result that we can observe is the estimation of models that define the time evolution of the mortality parameters of each cohort of the population. In general, we can observe that in a significant proportion of cohorts, the best model defining the time series process of the parameters is based on a simple differencing; however, other cohorts establish more complex models for the dynamics.

In general, we can mention that the methodology allows for flexibility within different populations and other applications that require establishing the survival dynamics of a population.

Our approach differs from other significant methodologies such as the estimation of mortality curves [19,46–48]. In our case, we consider the projection of the parameters that define mortality as the fundamental input for estimating mortality curves. Furthermore, the proposed methodology, conceived as modular components—with the use of Gompertz–Makeham’s law being one of those modules—allows us the possibility to employ other estimation approaches for mortality curves that can serve as model contrasts.

An important element to consider is the selection of the prior distribution implemented in the algorithm. In our case, a non-informative distribution (uniform distribution) is used for the methodology, assuming computational costs and potential biases due to the inadequate selection of the distribution [40]. This point should be considered in future research with a larger amount of data available for estimations because small fluctuations in the estimates can have a significant impact on the projections.

5. Conclusions

In the present research we used official data from the Government of Chile on vital mortality statistics. Specifically, statistics were used on the number of deaths for each cohort, with information available from 2004 to 2021, considering the consolidated official data for the estimates and projections to be made.

It can be observed that life expectancy has increased steadily over the last six decades. This growth in life expectancy is not unique to Chile; it is a global phenomenon. The reasons for the improvement in life expectancy are centred on the sustained increase in the quality of life and the increase in spending on medical services, among other social variables, which have been widely studied in the literature.

We have developed a methodology for estimating the curves that determine life expectancy, based on the temporal evolution of the parameters that determine the mortality of the population, using an ARIMA model. This methodology is based on the projection of the parameters to subsequently estimate the Gompertz–Makeham curve, using a Bayesian approach known as Approximate Bayesian Computation (ABC).

In fact, research such as that presented here can help to define priorities for work, monitoring, and the implementation of measures to improve the weaknesses observed in public systems that use the life expectancy of the population to define their services, which have a significant impact on the economy and that, today, represent a high level of public attention and scrutiny for the politicians and employers responsible for these services in relation to pensions and health systems.

The methodology presented makes it possible to establish the mean value of the shift of the survival curve for the population over the age of 50 and the projections of the best and worst scenarios, which can be interpreted as confidence bands in the prognosis of the survival of the population.

According to the literature, the relevance of the health area for economic growth could be a path of work and planning for the parties involved, with the purpose of focusing on one of their priority problems, such as economic reactivation, with the aim of achieving growth rates, being a reason it is an area of interest in the international arena, drastically reducing poverty and raising per capita income.

Regarding limitations and future directions, we plan to address extreme conditions such as financial crises or pandemic conditions, utilising alternative entropy estimation methods, such as the histogram-based method, to evaluate the predictability of the survival curves in Chile through the maximum entropy method, incorporating extreme volatility data influenced by social contexts. Additionally, as a comparative projection, this study could analyze the strategic weight that a series of variables that have been in the focus of public and private discussions, such as education, sustainability, environmental balance, mental health, gender equality, the modernization of the State, and the democratization of digital technologies.

Author Contributions: data curation, H.d.l.F.-M., K.C.-J., C.E.-G. and R.R.-T.; formal analysis H.d.l.F.-M., K.C.-J. and R.R.-T.; investigation, H.d.l.F.-M., K.C.-J., C.E.-G. and R.R.-T.; methodology, H.d.l.F.-M., K.C.-J., C.E.-G. and R.R.-T.; writing—original draft, H.d.l.F.-M., K.C.-J., C.E.-G. and R.R.-T.; writing—review and editing, H.d.l.F.-M., K.C.-J., C.E.-G. and R.R.-T. All authors have read and agreed to the published version of the manuscript.

Funding: The research work of H. de la Fuente-Mella was partially supported by Proyecto FONDECYT Regular. Código del Proyecto: 1230881. Agencia Nacional de Investigación y Desarrollo de Chile (ANID).

Data Availability Statement: The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Alaimo, L.S.; Levantesi, S.; Nigri, A. Fuzzy clustering of the healthy life expectancy decomposition: A multi-population analysis. *Socio-Econ. Plan. Sci.* **2024**, *92*, 101805. [[CrossRef](#)]
2. Cho, H.; Wang, Z.; Yabroff, K.R.; Liu, B.; McNeel, T.; Feuer, E.J.; Mariotto, A.B. Estimating life expectancy adjusted by self-rated health status in the United States: National health interview survey linked to the mortality. *BMC Public Health* **2022**, *22*, 141. [[CrossRef](#)] [[PubMed](#)]

3. Jung, M.; Ko, W.; Muhwava, W.; Choi, Y.; Kim, H.; Park, Y.S.; Jambere, G.B.; Cho, Y. Mind the gaps: Age and cause specific mortality and life expectancy in the older population of South Korea and Japan. *BMC Public Health* **2020**, *20*, 819. [[CrossRef](#)] [[PubMed](#)]
4. Roubal, A.M.; Pollock, E.A.; Gennuso, K.P.; Blomme, C.K.; Givens, M.L. Comparative methodologic and practical considerations for life expectancy as a public health mortality measure. *Public Health Rep.* **2022**, *137*, 255–262. [[CrossRef](#)] [[PubMed](#)]
5. Congdon, P. Life expectancies for small areas: A Bayesian random effects methodology. *Int. Stat. Rev.* **2009**, *77*, 222–240. [[CrossRef](#)]
6. Peters, G.W.; Fan, Y.; Sisson, S.A. On sequential Monte Carlo, partial rejection control and approximate Bayesian computation. *Stat. Comput.* **2012**, *22*, 1209–1222. [[CrossRef](#)]
7. Brønnum-Hansen, H.; Foverskov, E.; Andersen, I. Income inequality in life expectancy and disability-free life expectancy in Denmark. *J. Epidemiol. Community Health* **2021**, *75*, 145–150. [[CrossRef](#)]
8. Camargos, M.C.S.; do Nascimento Rodrigues, R.; Machado, C.J. Healthy life expectancy to Brazilian elders, 2003. *Ciência Saúde Coletiva* **2009**, *14*, 1903. [[CrossRef](#)]
9. Sandoval, M.H.; Alvear Portaccio, M.E.; Albala, C. Life expectancy by ethnic origin in Chile. *Front. Public Health* **2023**, *11*, 1147542. [[CrossRef](#)]
10. DuGoff, E.H.; Canudas-Romo, V.; Buttorff, C.; Leff, B.; Anderson, G.F. Multiple chronic conditions and life expectancy: A life table analysis. *Med. Care* **2014**, *52*, 688–694. [[CrossRef](#)]
11. Tyrer, F.; Chudasama, Y.V.; Lambert, P.C.; Rutherford, M.J. Flexible parametric methods for calculating life expectancy in small populations. *Popul. Health Metrics* **2023**, *21*, 13. [[CrossRef](#)] [[PubMed](#)]
12. Kameron, D.B. Mortality risk and life expectancy. *J. Oper. Res. Soc.* **1989**, *40*, 199–200. [[CrossRef](#)]
13. Wajiga, G.; Adekola, O.A. Life-expectancy in a nonhomogeneous population. *J. Oper. Res. Soc.* **1998**, *49*, 1011–1012. [[CrossRef](#)]
14. Ho Dang, P.; Nguyen, T.N. New methods of life expectancy estimation. *Environ. Ecol. Stat.* **2022**, *29*, 587–606. [[CrossRef](#)]
15. Cairns, A.J.; Blake, D.; Dowd, K.; Coughlan, G.D.; Khalaf-Allah, M. Bayesian stochastic mortality modelling for two populations. *ASTIN Bull. J. IAA* **2011**, *41*, 29–59.
16. Pham, H.; Pham, H.T. A Bayesian approach for multi-stage models with linear time-dependent hazard rate. *Monte Carlo Methods Appl.* **2019**, *25*, 307–316. [[CrossRef](#)]
17. Kim, Y.; Kang, S.B.; Berliner, L.M. Bayesian diffusion process models with time-varying parameters. *J. Korean Stat. Soc.* **2012**, *41*, 137–144. [[CrossRef](#)]
18. Poon, A.; Zhu, D. A new Bayesian model for contagion and interdependence. *Econom. Rev.* **2022**, *41*, 806–826. [[CrossRef](#)]
19. Li, H.; Tan, K.S.; Tuljapurkar, S.; Zhu, W. Gompertz law revisited: Forecasting mortality with a multi-factor exponential model. *Insur. Math. Econ.* **2021**, *99*, 268–281. [[CrossRef](#)]
20. Colchero, F.; Kiyakoglu, B.Y. Beyond the proportional frailty model: Bayesian estimation of individual heterogeneity on mortality parameters. *Biom. J.* **2020**, *62*, 124–135. [[CrossRef](#)]
21. Melnikov, A.; Romaniuk, Y. Evaluating the performance of Gompertz, Makeham and Lee–Carter mortality models for risk management with unit-linked contracts. *Insur. Math. Econ.* **2006**, *39*, 310–329. [[CrossRef](#)]
22. Yan, H.; Peters, G.W.; Chan, J. Mortality models incorporating long memory for life table estimation: A comprehensive analysis. *Ann. Actuar. Sci.* **2021**, *15*, 567–604. [[CrossRef](#)]
23. Delwarde, A.; Denuit, M.; Eilers, P. Smoothing the Lee–Carter and Poisson log-bilinear models for mortality forecasting: A penalized log-likelihood approach. *Stat. Model.* **2007**, *7*, 29–48. [[CrossRef](#)]
24. Giacometti, R.; Bertocchi, M.; Rachev, S.T.; Fabozzi, F.J. A comparison of the Lee–Carter model and AR–ARCH model for forecasting mortality rates. *Insur. Math. Econ.* **2012**, *50*, 85–93. [[CrossRef](#)]
25. Beaumont, M.A.; Cornuet, J.M.; Marin, J.M.; Robert, C.P. Adaptive approximate Bayesian computation. *Biometrika* **2009**, *96*, 983–990. [[CrossRef](#)]
26. Toni, T.; Welch, D.; Strelkowa, N.; Ipsen, A.; Stumpf, M.P. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *J. R. Soc. Interface* **2008**, *6*, 187–202. [[CrossRef](#)] [[PubMed](#)]
27. Sisson, S.A.; Fan, Y.; Tanaka, M.M. Sequential monte carlo without likelihoods. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 1760–1765. [[CrossRef](#)] [[PubMed](#)]
28. Marjoram, P.; Molitor, J.; Plagnol, V.; Tavaré, S. Markov chain Monte Carlo without likelihoods. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 15324–15328. [[CrossRef](#)]
29. Rubilar-Torrealba, R.; Chahuán-Jiménez, K.; de la Fuente-Mella, H. A Stochastic Analysis of the Effect of Trading Parameters on the Stability of the Financial Markets Using a Bayesian Approach. *Mathematics* **2023**, *11*, 2527. [[CrossRef](#)]
30. Roffia, P.; Buccioli, A.; Hashlamoun, S. Determinants of life expectancy at birth: A longitudinal study on OECD countries. *Int. J. Health Econ. Manag.* **2023**, *23*, 189–212. [[CrossRef](#)]
31. Kabir, M. Determinants of life expectancy in developing countries. *J. Dev. Areas* **2008**, *41*, 185–204. [[CrossRef](#)]
32. Gompertz, B. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. In a letter to Francis Baily, Esq. FRS &c. By Benjamin Gompertz, Esq. FR S. *Philos. Trans. R. Soc. Lond.* **1825**, *115*, 252–253.
33. Makeham, W.M. On the law of mortality and the construction of annuity tables. *J. Inst. Actuar.* **1860**, *8*, 301–310. [[CrossRef](#)]
34. Hallén, A. Makeham’s addition to the Gompertz law re-evaluated. *Biogerontology* **2009**, *10*, 517–522. [[CrossRef](#)] [[PubMed](#)]

35. Golubev, A. How could the Gompertz–Makeham law evolve. *J. Theor. Biol.* **2009**, *258*, 1–17. [[CrossRef](#)]
36. Missov, T.I. Gamma-Gompertz life expectancy at birth. *Demogr. Res.* **2013**, *28*, 259–270. [[CrossRef](#)]
37. Missov, T.I.; Lenart, A. Gompertz–Makeham life expectancies: Expressions and applications. *Theor. Popul. Biol.* **2013**, *90*, 29–35. [[CrossRef](#)] [[PubMed](#)]
38. Stock, J.H.; Watson, M.W. Median unbiased estimation of coefficient variance in a time-varying parameter model. *J. Am. Stat. Assoc.* **1998**, *93*, 349–358. [[CrossRef](#)]
39. Zhou, X.; Sun, L. Additive hazards regression with missing censoring information. *Stat. Sin.* **2003**, 1237–1257.
40. van de Schoot, R.; Depaoli, S.; King, R.; Kramer, B.; Märtens, K.; Tadesse, M.G.; Vannucci, M.; Gelman, A.; Veen, D.; Willemsen, J.; et al. Bayesian statistics and modelling. *Nat. Rev. Methods Prim.* **2021**, *1*, 1. [[CrossRef](#)]
41. Winklevoss, H.E. *Pension Mathematics with Numerical Illustrations*; University of Pennsylvania Press: Philadelphia, PA, USA, 1993.
42. Rostan, P.; Belhachemi, R.; Rostan, A. Appraising the financial sustainability of a pension system with signal processing. *Stud. Appl. Econ.* **2015**, *33*, 801–816. [[CrossRef](#)]
43. Queiroz, B.L.; Ferreira, M.L.A. The evolution of labor force participation and the expected length of retirement in Brazil. *J. Econ. Ageing* **2021**, *18*, 100304. [[CrossRef](#)]
44. Salinari, G.; De Santis, G. One or more rates of ageing? The extended gamma-Gompertz model (EGG). *Stat. Methods Appl.* **2020**, *29*, 211–236. [[CrossRef](#)]
45. Baeten, S.; Van Ourti, T.; Van Doorslaer, E. Rising inequalities in income and health in China: Who is left behind? *J. Health Econ.* **2013**, *32*, 1214–1229. [[CrossRef](#)]
46. Carter, L.R.; Lee, R.D. Modeling and forecasting US sex differentials in mortality. *Int. J. Forecast.* **1992**, *8*, 393–411. [[CrossRef](#)]
47. Conde-Gutiérrez, R.; Colorado, D.; Hernández-Bautista, S. Comparison of an artificial neural network and Gompertz model for predicting the dynamics of deaths from COVID-19 in México. *Nonlinear Dyn.* **2021**, *104*, 4655–4669. [[CrossRef](#)]
48. Mohamed, H.S.; Ali, M.M.; Yousof, H.M. The Lindley Gompertz Model for Estimating the Survival Rates: Properties and Applications in Insurance. *Ann. Data Sci.* **2023**, *10*, 1199–1216. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

A Dynamic Analysis of a Poly-Articulated Robot

Sorin Dumitru, Cristian Copilusi , Nicolae Dumitru *  and Ionut Geonea 

Faculty of Mechanics, University of Craiova, 200512 Craiova, Romania; sorin.dumitru@edu.ucv.ro (S.D.); cristian.copilusi@edu.ucv.ro (C.C.); ionut.geonea@edu.ucv.ro (I.G.)

* Correspondence: nicolae.dumitru@edu.ucv.ro

Abstract: This paper studies the kinematics and dynamics of a poly-articulated robot. The robot can be used in hardly accessible places and special environments. The poly-articulated robot includes two main parts: a flexible unit and an actuation unit. The flexible unit consists of three modules specially designed for serving in a complex 3D workspace. Each module has flexible vertebrae and rigid disks. The poly-articulated robot simulation is achieved with the MSC Adams 2012 and ANSYS R14.5 software. Thus, we aim to determine whether the variation laws depend on the time of the kinematic parameters for each part in a specific motion, considering each part has to act as a rigid body or a deformable body. Using the finite element method, the stress and deformations for normal and critical positions are calculated for the poly-articulated robot. To validate the simulation models designed in this research, an experimental analysis of the proposed poly-articulated robot is developed. The command and control unit was equipped with motion sensors that allow to identify the position of each flexible unit module.

Keywords: dynamics; simulation; poly-articulated robot



Citation: Dumitru, S.; Copilusi, C.; Dumitru, N.; Geonea, I. A Dynamic Analysis of a Poly-Articulated Robot. *Computation* **2024**, *12*, 156. <https://doi.org/10.3390/computation12080156>

Academic Editor: Demos T. Tsahalidis

Received: 12 June 2024

Revised: 18 July 2024

Accepted: 26 July 2024

Published: 2 August 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Nowadays, poly-articulated robots are widely used in domains like industrial, medical, military and aerospace, due to their specific design for accomplishing tasks in a complex workspace.

Their conceptual design is inspired from elephant trunks, snakes and even octopus tentacles like the ones from [1–6]. These are characterized by a large number of DOFs due to their highly articulated structure. Thus, by considering the previously mentioned fundamental domains, these robots are well suited for exploration operations in highly constrained environments.

The large DOF is determined by the use of a number of serially connected links and actuated joints, like [7,8], when designing such robots.

An important design task, specific to the ones used in the medical domain, is to have a miniaturized structure size for use in endoscopy [9] or cholecystectomy [10]. Thus, this design task will trigger new problems to be solved, like stiffness, increased dexterity and precise control guidance [11].

Despite the imposed tasks, these robots need to go through several design stages that involve complex kinematic and dynamic analyses of the poly-articulated structure. On the other hand, the command and control unit should have a specific design architecture in order to work perfectly with a poly-articulated structure like the ones from [12,13].

By considering design concepts similar to the ones from [14–17], a major and common design objective imposed for the poly-articulated robot structure can be remarked. This design objective is to obtain an S-shaped spatial trajectory developed by the poly-articulated structure in a Cartesian coordinate system.

Another argument of this design task is to have a parametrized S-shaped spatial trajectory. This will involve complex dynamic analyses, like the ones performed in [18–20].

Also, this can be accomplished by designing a continuum robot concept characterized by combining rigid bodies with flexible ones like [21–23].

By having in sight the ones presented in [20–22], our research aim is to design a robot with a poly-articulated structure that can develop a parametrized S-shaped spatial trajectory by using a combination of rigid bodies with flexible ones.

Thus, our contribution to this research field is to develop, design and test a prototype able to accomplish the desired aims. Another research peculiarity is to achieve the proposed aims by using modern design techniques and software, such as MSC Adams and ANSYS. Thus, the research core is based on the Craig–Bampton method, presented in [23], for elaborating the mathematical models used in kinematic and dynamic analyses. Other original contribution is the architecture, though similar poly-articulated robots can be found, like the ones from [24–26].

In order to achieve the mentioned objectives, this research is organized as follows. Starting from our own original contribution, we drew up a robot design principle characterized by nine DOFs, which combines rigid bodies with flexible ones. Based on this principle, two different flexible units—a cylindrical and a tronconic one—will be designed and experimentally tested. This is presented in the second part of the research, along with the working principle.

The novelty brought about through this research emphasizes an elaborated prototype characterized by constructive simplicity of the flexible unit, the design of a flexible unit in a modular form that can be combined anytime (by lowering or increasing the DoF), and a new shape of the flexible unit, namely a tronconic shape of this, which will contribute to the improvement and control of a variable stiffness. By having in sight the processed virtual models, there will be mentioned several parameters for virtual simulations setup, and these are valuable for other future research.

As previously mentioned, the research core focuses on a dynamic analysis of the proposed poly-articulated unit, and the approach consists of elaborating a mathematical model based on the Craig–Bampton method from [23], presented in the third part of this research. This mathematical model represents the background for virtual simulations of the poly-articulated robot and numerical processing in dynamic conditions with the aid of the MSC Adams and ANSYS software.

To validate the elaborated poly-articulated concept, several virtual simulations in different environments have been performed. One will be accomplished in MSC Adams, where the poly-articulated unit was imported and the combination between flexible and rigid bodies was properly defined. This is presented in the fourth part of the research, where significant results were obtained by considering imposed constraints similar to reality.

Another interface was created in the ANSYS environment, and several virtual simulations in dynamic conditions will be performed to study the contact between the unit components of the poly-articulated robot. This case will be presented in the fifth part of the research, where more attention will be given to the contact definition during virtual simulations in order to achieve a good behavior of the entire robotic system and to validate the proposed concept.

Based on the significant results that will have been obtained, the last part of the research will present the prototype, and a range of experimental tests will prove the efficacy of the dynamic analyses performed with the aid of the Craig–Bampton method, combined with modern techniques.

2. Virtual Prototyping of the Poly-Articulated Robot

A poly-articulated robotic system was designed that has two major units: a flexible unit and an actuation unit.

The flexible unit structure has three modules, and each module contains three flexible vertebrae and rigid disks. These modules are connected through a flexible central element. Thus, the entire flexible unit has nine flexible vertebrae that are actuated through nine

electric motors. These electric motors represent the actuation unit, and the translational motion is assured using screw and nut joint transmissions.

Two types of flexible units—a cylindrical and a tronconic one—were designed. The latter, i.e., the tronconic form, has a major advantage, which consists of an easy and accurate control of the stiffness of the flexible unit across all three modules. The tronconic concept of the flexible unit assures a variable stiffness across its own length, and this will ensure enhanced flexibility onto the distal module, according to Figure 1.

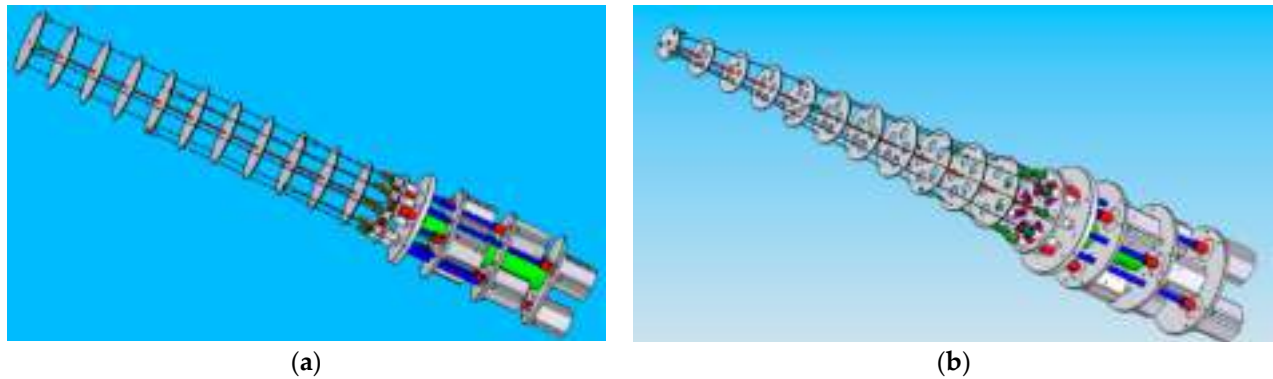


Figure 1. Flexible unit types proposed through this research: (a) flexible unit with a cylindrical architecture, and (b) flexible unit with a tronconic form.

The cylindrical form of the flexible unit implies technological facilities, so that the manufacturing processes are much easier than in the case of the tronconic form, and it was also seen that this can be done with low costs.

The robot workspace was defined by the mobility limits of the flexible unit, and this was designed in a modular form in order to have a simplified control for the robot configuration and its evaluated workspace. The deformed shape of the flexible unit seems to be a snake-like form, and thus, the term “snake-like robot” can be assigned to the entire designed system.

The robot’s mechanical structure is designed so that the axial displacements can be adjusted, and this ensures the robot’s kinematic precision. In particular, the flexible unit was designed for two types: one with a cylindrical structure and the other with a conic structure. The robot’s mechanical parts were built at the laboratories of the Faculty of Mechanics, and the command and control unit was designed at the Faculty of Automation, Computers and Electronics of the University of Craiova [21].

Thus, the poly-articulated robot was designed according to the 3D model presented in Figure 2. Figure 3 shows an exploded view with a proper identification of the components of the poly-articulated robotic system 3D model.

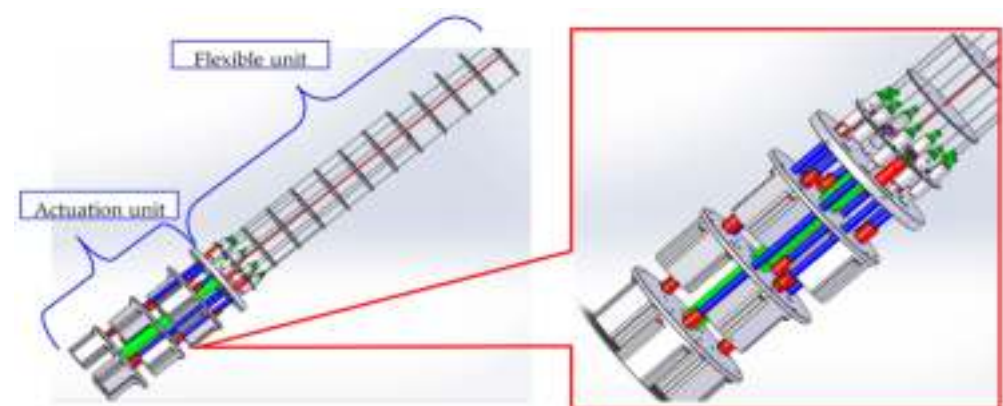


Figure 2. Virtual model of the poly-articulated robot in a parametrized form.

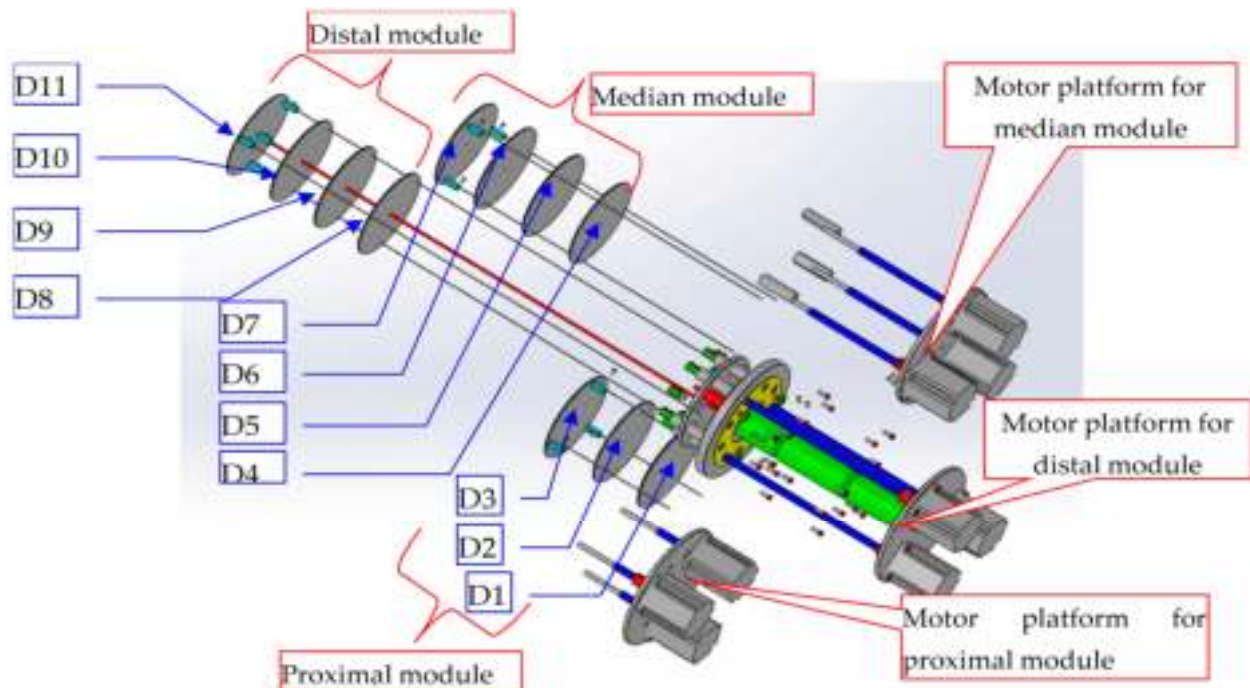


Figure 3. A virtual model of the poly-articulated robot—expanded view.

The working principle of the robot depends on the motor actuation control, which, in this case, is represented by the nine stepper motors. The robot architecture consists of three modules (proximal, median and distal), and each module is actuated by three stepper motors in order to develop combined motions in a 3D working space. The stepper motors are directly connected to power screw transmissions, which will pull or release the actuation cables. Thus, the power screw transmissions will transform the rotational motion of the stepper motor shaft in a translational motion imposed on the nut component connected to an actuation cable. Thus, the motion is based on screwing and unscrewing the power transmissions, which will determine the shortening or prolonging of the continuum robot motion. In this way, the tentacle curvature will be changed with the variation of the actuation cables.

Figure 3 shows the backbone of the tentacle—the central component marked in red, which represents an elastic rod manufactured from steel, which sustains the whole structure and permits the bending sequence. Thus, the working principle depends on which cable is shortened or extended, and the tentacle bends in different planes. Each cable will contribute to making a motion with different angles, depending on the initial coordinate frame attached to the manipulator segment, for example, allowing a 3D motion. Thus, Figure 1 shows both flexible unit models, and it can be seen that they use the same number of components, the only difference being given by the shape of the disks. This was designed in a parametrized form, considering the definition of the precise contour conditions. The obtained model respects the kinematic conditions and was also designed in two variants, compatible with MSC Adams 2012 and ANSYS R 14.5 software. These are shown in Figures 4 and 5.

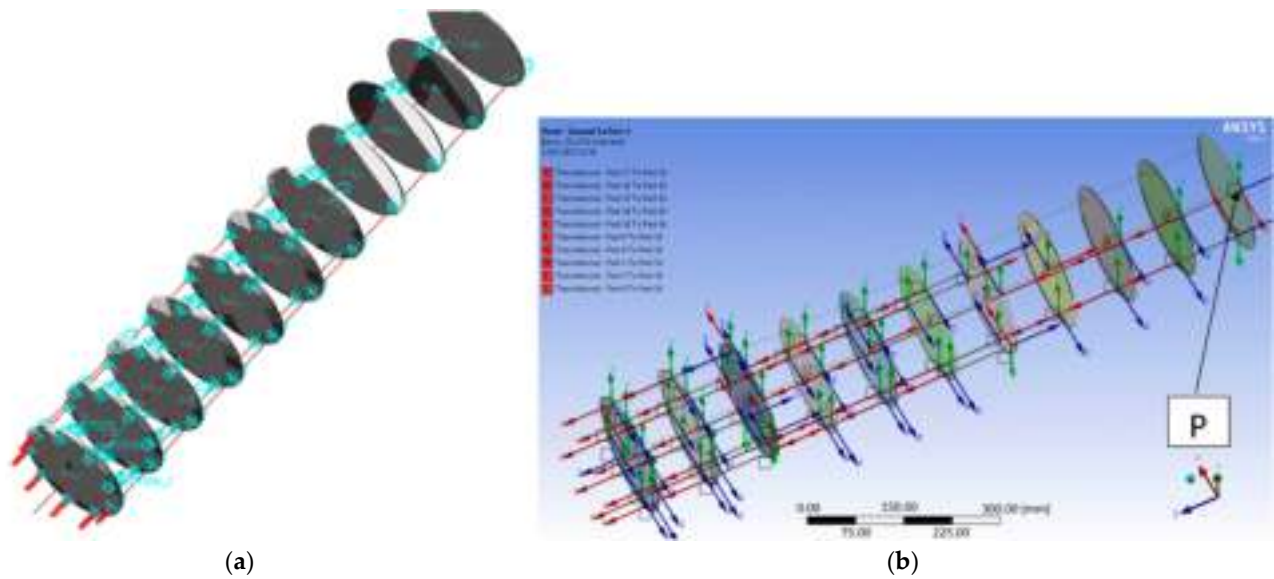


Figure 4. Poly-articulated cylindrical robot kinematic models using (a) MSC Adams and (b) ANSYS software.

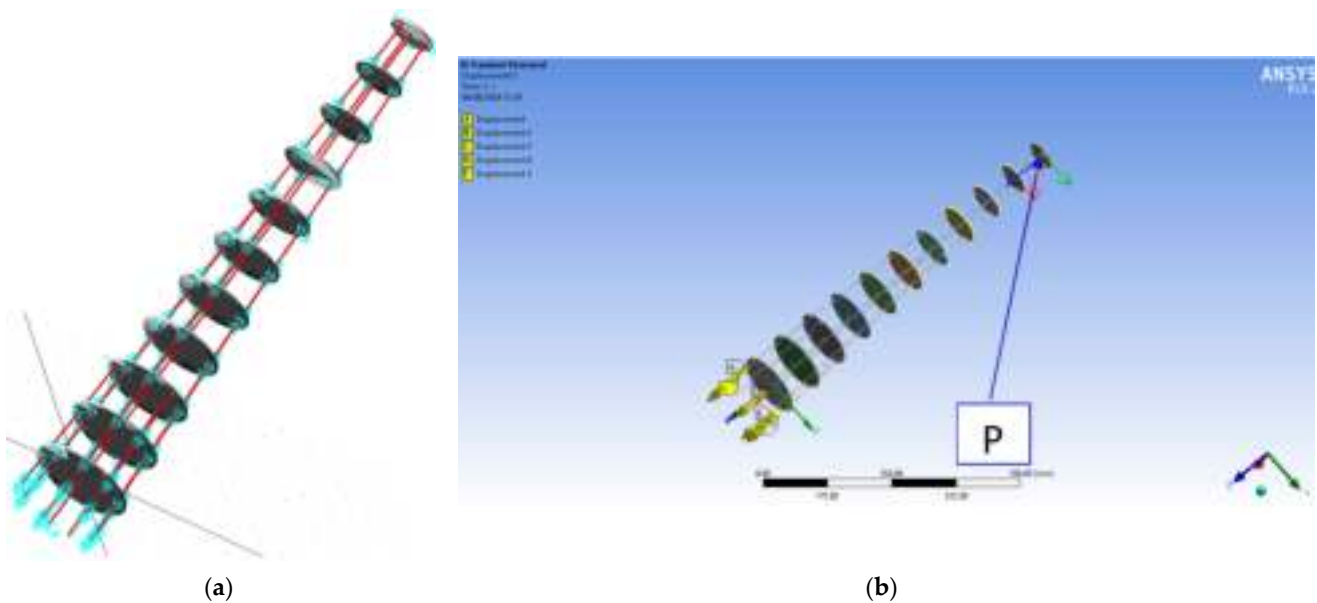


Figure 5. Poly-articulated tronconic robot kinematic models using (a) MSC Adams and (b) ANSYS software.

3. Mathematical Models

In the case of the MSC Adams software robot analysis, the mathematical model is based on the Craig–Bampton method. The principle of the method is presented in [23] and assumes the following theoretical considerations: by transforming a rigid body into a flexible one, a procedure is characterized by taking into account only the small linear strain of the proposed body, depending on a local reference coordinate system; the proposed rigid body will be meshed into finite elements. In this way, a finite number of DOFs will be identified.

Nodal displacements can be expressed like a linear combination between a small number of proper vectors, named modal forms and assigned to a $\emptyset$ parameter. For a deformable body “ i ”, these are represented by

$$u = \sum_{i=1}^M \varnothing_j \cdot q_j \quad (1)$$

where u —nodal displacements; M —number of mode shapes; q_j —modal coordinates.

Equation (1) also represents the connection between a wide set of physical coordinates and a small set of modal coordinates, namely q_j .

The Craig–Bampton method allows optimizing the obtained modal base by considering two DOF sets, i.e., contour degrees of freedom u_B , which are used for load applications or for connecting a flexible body to a rigid one; internal degrees of freedom u_I , i.e., the internal degrees of freedom of the flexible element, which can be condensed using super-element techniques.

For the kinematic analysis of a flexible body (element), two reference coordinate systems will be considered as follows: $T'_i (x_i, y_i, z_i)$ —the reference coordinate system attached to a flexible body; $T_0 (X, Y, Z)$ —the global reference coordinate system.

Some interest nodes can be identified by meshing up the flexible body into finite elements. Markers with proper local reference coordinate systems can be placed on the identified interest nodes.

Thus, a “ P_i ” marker can be assigned on a flexible body, and its position can be defined in a global coordinate reference system as follows:

$$\vec{r}_{pi} = \vec{r}_i + \vec{S}_{pi} + \vec{U}_{pi} \quad (2)$$

where $\vec{r}_i$ —the position vector of the T'_i local coordinate system origin, depending on the T_0 global reference system; $\vec{S}_{pi}$ —position vector of the P_i —marker (undeformed position), depending on the T'_i local coordinate system; $\vec{U}_{pi}$ —the position vector of the P'_i —marker, depending on the T_{Pi} —reference system, which is centered in the P_i —marker.

A scheme of the coordinate system position placed on a simplified model is shown in Figure 6. The coordinate transformation matrices will be identified by taking into account Equation (2), and the following results will be obtained:

$$r_{Pi} = r_i + A_{Oi} \cdot S'_{Pi} + A_{Oi} \cdot U'_{Pi} \quad (3)$$

$$r_{Pi} = r_i + A_{Oi} \cdot (S'_{Pi} + U'_{Pi}) \quad (4)$$

where $\vec{r}_i = [x_i, y_i, z_i]^T$ represents the origin position of the T'_i local reference system, depending on the T_0 global reference system.

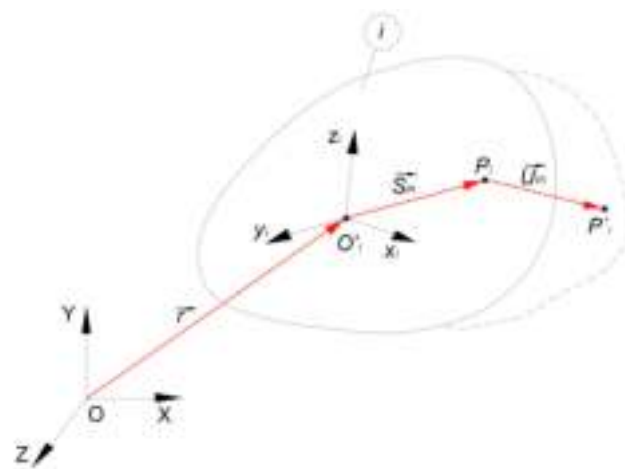


Figure 6. Coordinate systems location on a flexible body for the Craig–Bampton method.

Thus, the generalized coordinate system vector of the flexible body in an undeformed situation is

$$q_{ri} = [x_i, y_i, z_i, \psi_i, \theta_i, \varphi_i]^T \quad (5)$$

where x_i, y_i, z_i —the coordinates of the original mobile reference system, namely $T'_i(x_i, y_i, z_i)$ —fixed, with a deformable body— i ; ψ_i, θ_i and φ_i —Euler angles describing the space orientation of the i —body; A_{0i} —matrix of the coordinate transformation created from the director cosines of the T'_i , local reference system, depending on the T_0 global reference system; U'_{Pi} —deformation vector expressed in the T'_i local reference system.

The deformation displacement vector can be expressed as a modal superposition in a matrix form as

$$U_{Pi} = \phi_{Pi} \cdot q_i \quad (6)$$

where ϕ_{Pi} —is a part of the modal matrix that corresponds to P_i —marker translational DOF. This matrix has a $3 \times M$ dimension, where M represents the modal forms number; q_i —modal generalized coordinates vector.

Using the previous equations, the generalized coordinate vector can be defined for a flexible body as follows:

$$\xi = [x_i, y_i, z_i, \psi_i, \theta_i, \varphi_i, q_i]^T; i = 1 \dots M \quad (7)$$

Equation (7) shows the q_i modal coordinates ($i = 1 \dots M$), which characterize the generalized coordinates of the flexible element.

The angular deformation of a flexible element can be obtained by using a modal superposition:

$$\theta_{Pi} = \Phi_{Pi}^* \cdot q_i \quad (8)$$

where Φ_{Pi}^* is a part of the modal matrix that corresponds to rotational degrees of freedom for the P_i —node. It has a $3 \times M$ dimension, where M is the mode number.

The solid-deformable motion is analyzed for a K_i marker that performs rotations with small angles, depending on its local reference system.

Each marker is defined through a proper coordinate system. In the case of the K_i —marker orientation, depending on the T_0 global reference system, a coordinate transformation matrix was given, namely A_{0Ki} , where

$$A_{0Ki} = A_{0i} \cdot A_{iPi} \cdot A_{PiKi} \quad (9)$$

where A_{0i} is the transformation matrix obtained by crossing from the local reference system to the i body, namely T'_i , to the global reference system T_0 ; A_{iPi} is the transformation matrix that results from changing the orientation due to P_i —node deformation; A_{PiKi} is the constant transformation matrix that was defined when the marker was placed on the flexible body. The presented elements represent input data for the used procedure for transforming a rigid body into a deformable one, with the aid of MSC Adams software.

4. Dynamic Modeling with MSC Adams

The flexible unit was analyzed with MSC Adams 2012 software. Thus, it was created as a flexible interface for importing the flexible unit CAD geometry in 2D and 3D in MSC Adams. There was defined the material properties for the proper components (steel for vertebrae and aluminum for the disks), and also, there will be allocated local reference systems for each component of the flexible unit. There were defined kinematic joints together with the attached reference systems. These consist of 21 kinematic elements, namely 10 flexible elements and 11 rigid elements represented by the disks, according to Figure 2. The links between these kinematic elements were provided by 75 kinematic joints, namely 21 fixed joints and 54 translational joints with a single DOF. This unit was actuated through 6 joint motors. Figure 7 shows the motor joint definition. The motion is represented by the vectors MOTION_1, MOTION_2 and MOTION_3 (Figure 7). The displacement U_y ,

in Figure 7 represent the motion along the central axis shown in Figure 3. The motion sequence is as follows: IF (time—1: 0, 0, $-2 \times \text{time}$)—motion_1—proximal module; IF (time—1:0, 0, $-1.5 \times \text{time}$)—motion_2—median module; IF (time—2: $-2 \times \text{time}$, 0, 0)—motion_3—distal module.

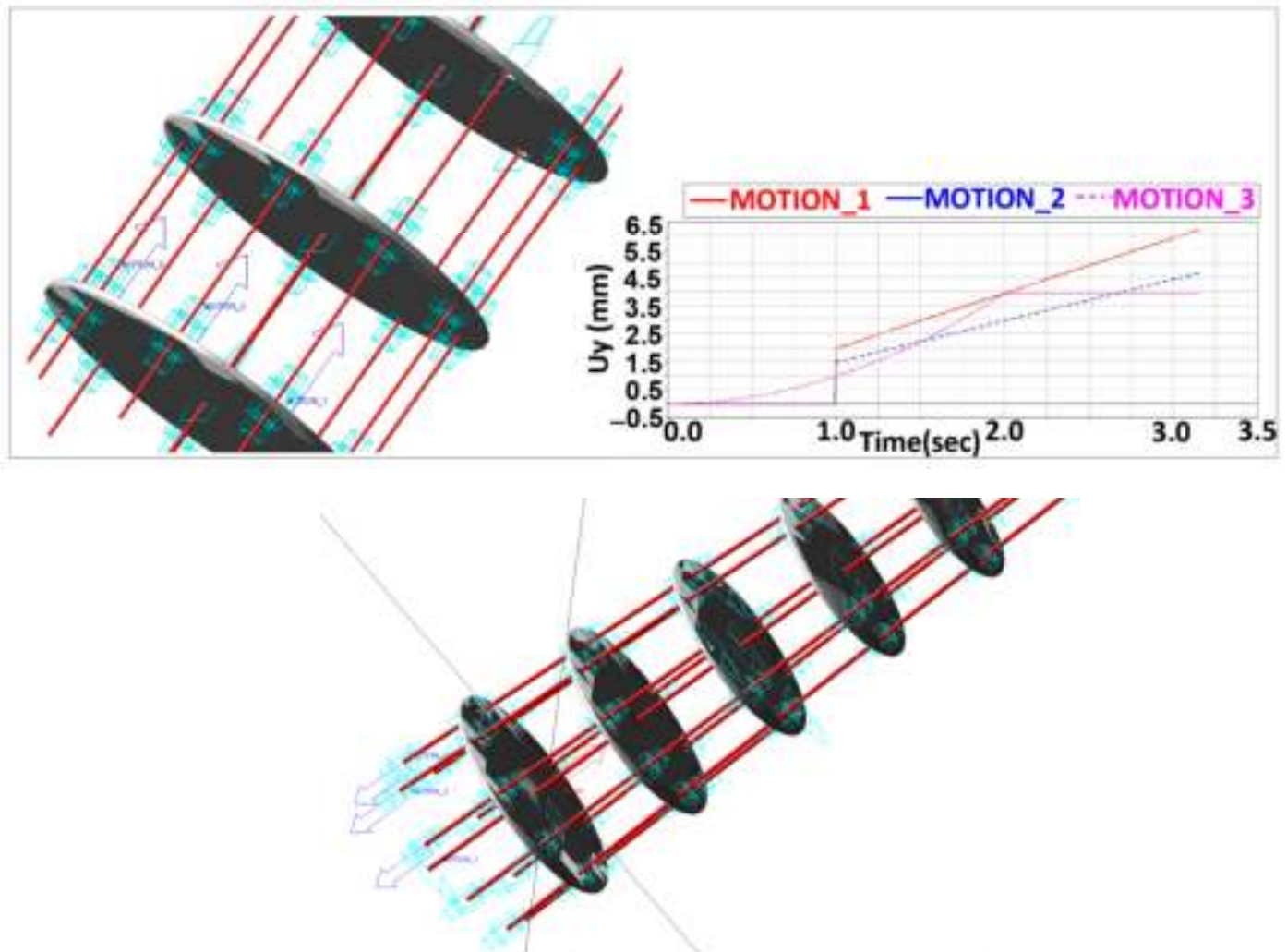


Figure 7. Motor joint displacement setup on a 3-s time interval for both analyzed flexible units.

For this, a robot motion analysis in a kinematic or dynamic mode will be possible by considering the flexible unit as a mechanism with deformable elements. Considering the elaborated Craig and Bampton mathematical models [9,20], ADAMS 2012 software will allow performing a modal dynamic analysis of any flexible mechanical system by superimposing the rigid-solid motion onto a deformable-solid one.

There were transformed nine actuation vertebrae from rigid bodies to flexible bodies as follows: each body was meshed in finite elements (Element type: solid, Element shape: tetrahedral, Element order: quadratic and Edge shape: mixed); there were defined the connection conditions with other elements through a procedure that consists of nodes identification master–slave type.

The flexible elements—generically referred to as elastic vertebrae—were meshed into tetrahedral finite elements. The modal dynamic analysis was performed for 77 mode shapes.

After this, there were determined the natural frequencies and vibration proper modes, and in this way, there were created the proper conditions for applying the modal superposition method (Craig–Bampton) to achieve the mechanism dynamic analysis.

For solving the differential equations system, there were used the WSTIFF integration module, with SI2 formulations and an integrating step set at 0.01.

The contact from the translational joints was a set of solid rigid-solid deformable type.

During the virtual simulations, kinematic parameter motion laws were identified in the case of the node placed on the extremity of the central vertebrae, depending on the reference coordinate system placed on the mass center of the first disk, which was considered fixed in space for the actuation unit.

Figure 8 shows a few flexible unit-deformed shapes in a 3D workspace, which resulted during motor joint actuation in different actuation modes.

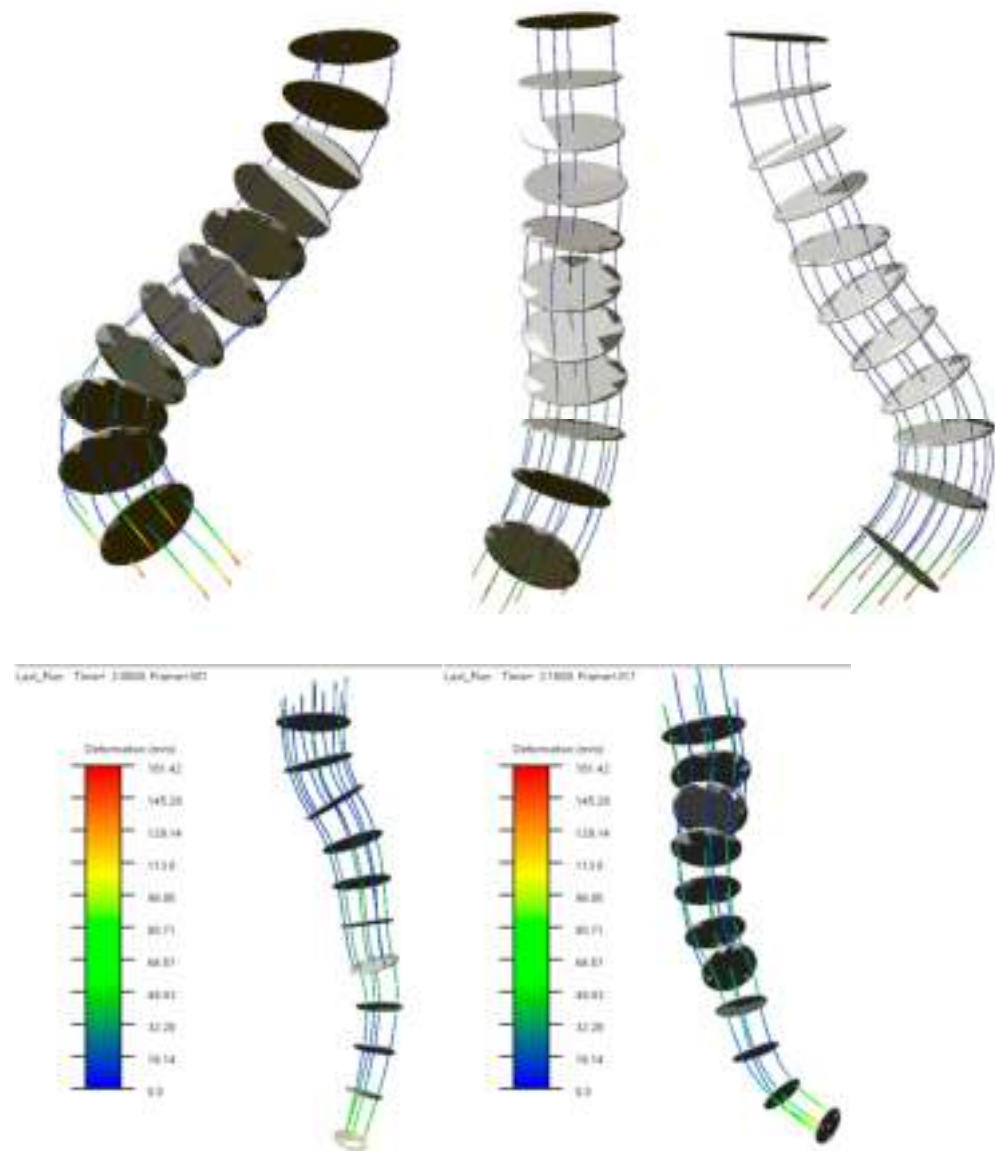


Figure 8. Deformed shapes of the flexible unit for both analyzed forms processed under the MSC Adams environment.

The obtained results are represented by displacement components of the end disk (D11) for distal module no.2 and the end disk (D7) for median module no.1, according to Figure 2. Some nodes located at the mass center of the proposed disks were considered for measurement during the deformable-solid virtual simulation.

Figure 9 represents the nodal translational deformation for node no. 14405 placed in the mass center of the final disk D11 across the distal module. It can be seen that the

maximum displacement values are recorded in a transversal plane defined by the x and y axes of the local reference system.

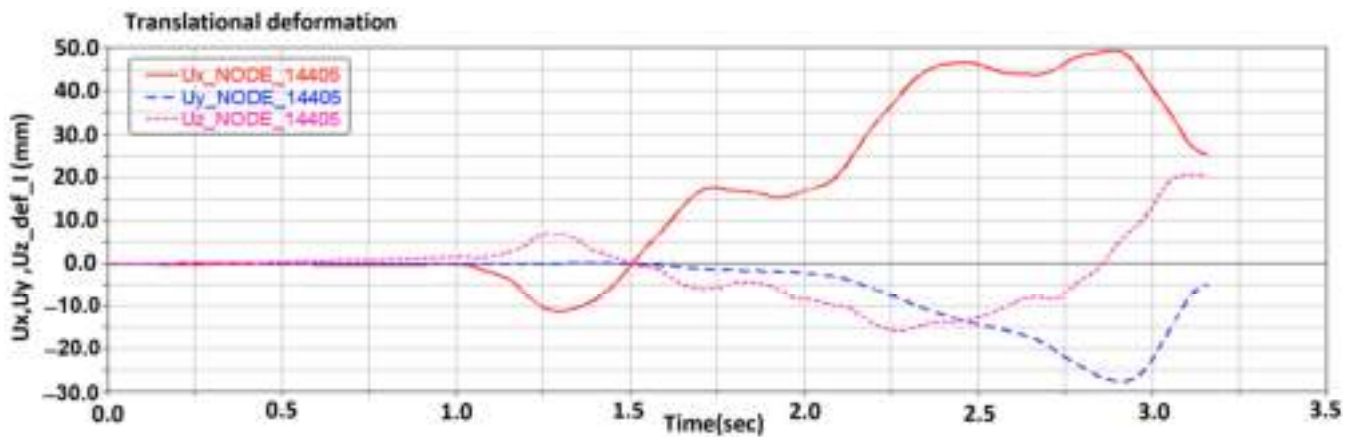


Figure 9. Translational deformation components for the distal module corresponding to the node from disk D11 vs. time.

Figure 10 represents the nodal translational deformation for node no. 14401 placed in the mass center of the final disk D7 across the median module. Also, in this case, it can be noticed that the maximum displacement values are recorded in a transversal plane of the flexible unit, but these values are smaller than the ones reported in Figure 7. This is caused by the median module stiffness, which is higher than the one for the distal module.

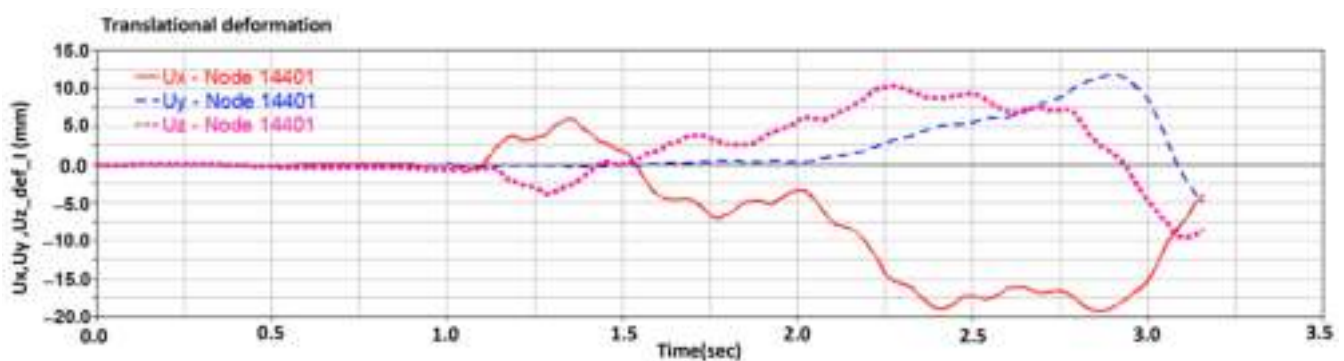


Figure 10. Translational deformation components for the distal module corresponding to the node from disk D7 vs. time.

5. Dynamic Modeling with ANSYS

In this case, the robotic system from Figure 2 was considered, and, for this, a dynamic analysis was performed with the ANSYS R14.5 package [8]. Firstly, the geometric model was imported, and the second step was to mesh this into finite elements, as it can be seen in Figure 9. The third step was to define the contact between disks and vertebrae, and the contour conditions were established at its end.

The kinematic links were considered to be flexible, and increased attention has been devoted to the network of finite elements in the contact areas where these were resized, according to Figure 11. Thus, this research frame presents, as a detailed example, the case of the flexible unit characterized by a cylindrical form. In the other case, i.e., the tronconic form of the flexible unit, a similar approach was implemented, with similar results.

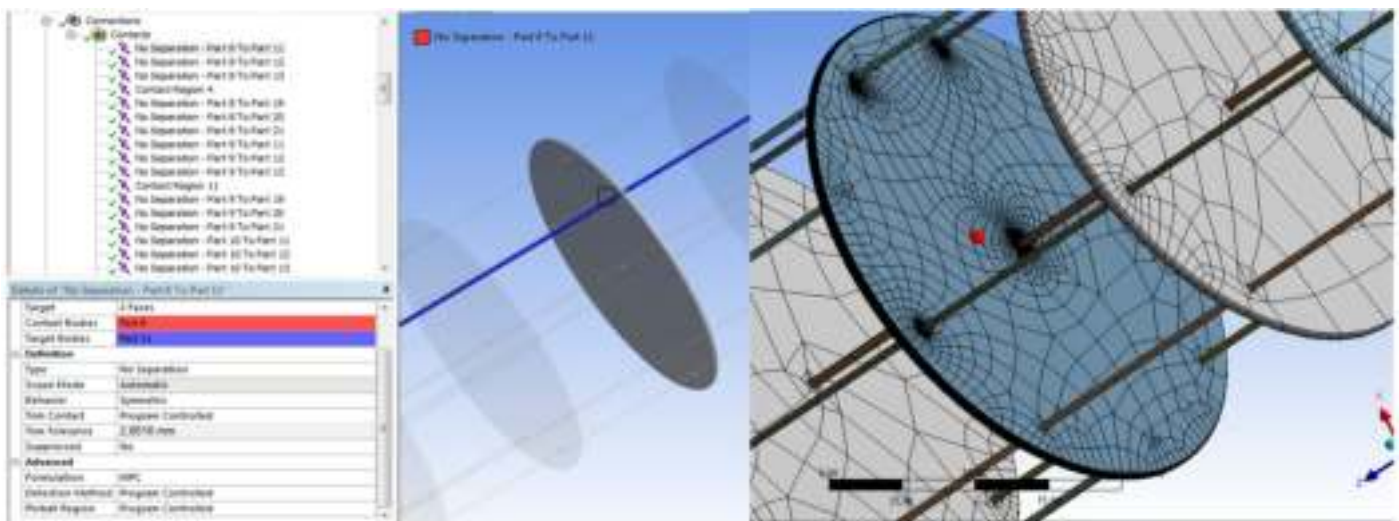


Figure 11. Contact definition between elastic vertebrae and flexible unit disks.

Regarding the initial setup of the analysis, there were defined the contact modeling parameters, the solver type and the time for sequence setup of the proper virtual processing. Thus, for the solver type and time setup, we define the following parameters: Number of steps: 3; Initial time step: 0.1 s; Minimum time step: 0.001; Maximum time step: 0.002; Solver type: Direct; Numerical damping: 0.1. In case of the contact modeling, there were defined contact parameters in case of the specific joints modeling, respectively, the nine actuating vertebrae and disks as follows: Type contact: No separation; Behavior: Symmetric; Formulation: MPC; Detection Method: Nodal-Normal to Target. The contact between the disks and the central flexible element was done as follows: Type: Bounded; Behavior: Symmetric; Formulation: Pure Penalty; Detection Method: Nodal-Normal to Target. According to Figure 2, the disk D1 was fixed in space.

In the fourth step, the motor parameters were defined, as it can be remarked in Figure 12.

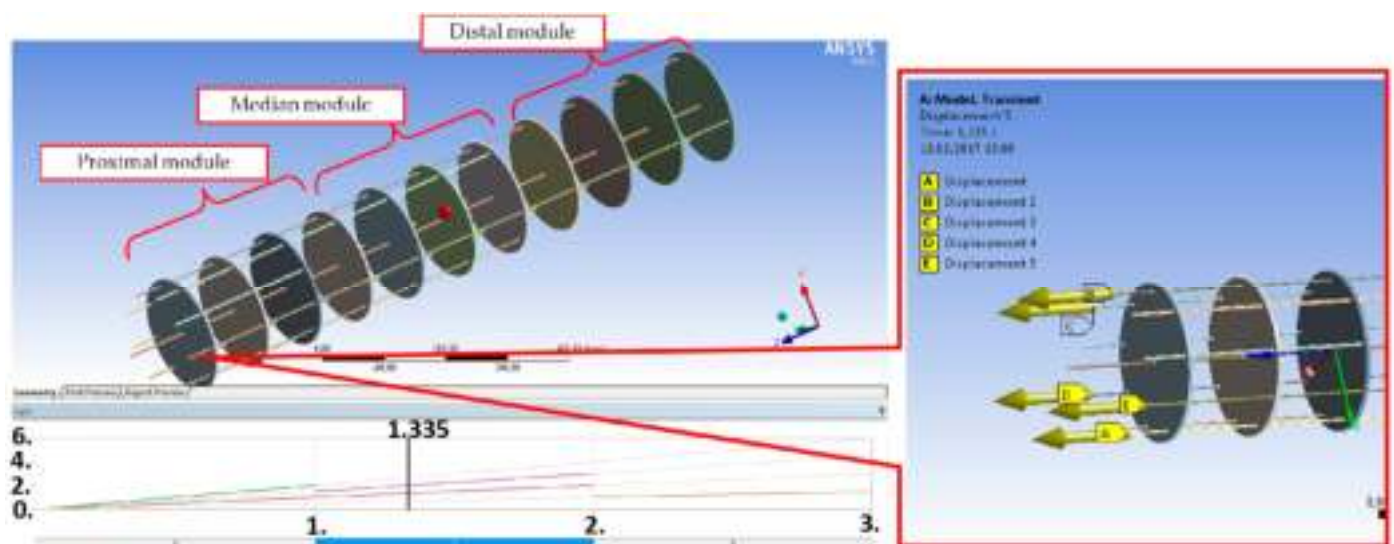


Figure 12. Definition of the flexible unit motor elements in case of the cylindrical form.

A simulation setup for the actuation mode during the time sequence was performed for all flexible unit modules according to Figure 12, with proper motions: median module, $uz1 = t$ ($t \leq 2$); distal (final) module $uz2 = 2t$ ($t \leq 1$); proximal (base) module $uz3 = 0.5t$

($1 \leq t \leq 2$); distal (final) module $uz4 = 1.5t$ ($1 \leq t \leq 2$); distal (final) module $uz5 = 0.5t$ ($2 \leq t \leq 3$).

After processing all the input data, increased attention was given to the motion of the components of the robotic system, according to two reference systems, i.e., a global coordinate system and a local one placed at the center of the base disk. Thus, the variation laws in the motion of kinematic parameters were identified, depending on the time for the chosen markers, namely the ones placed at the centers of reference disks for proximal, medial and distal modules. The motion developed by these markers was analyzed after three components of the proper coordinate system.

The reference disks are the ones actuated directly by the motor vertebrae (final disks from each module). Thus, several results were obtained, which are shown in the diagrams from Figures 13–15, and these represent the displacement components on the X and Z axes for the marker placed on the local coordinate system attached to the final disk of the proximal module.

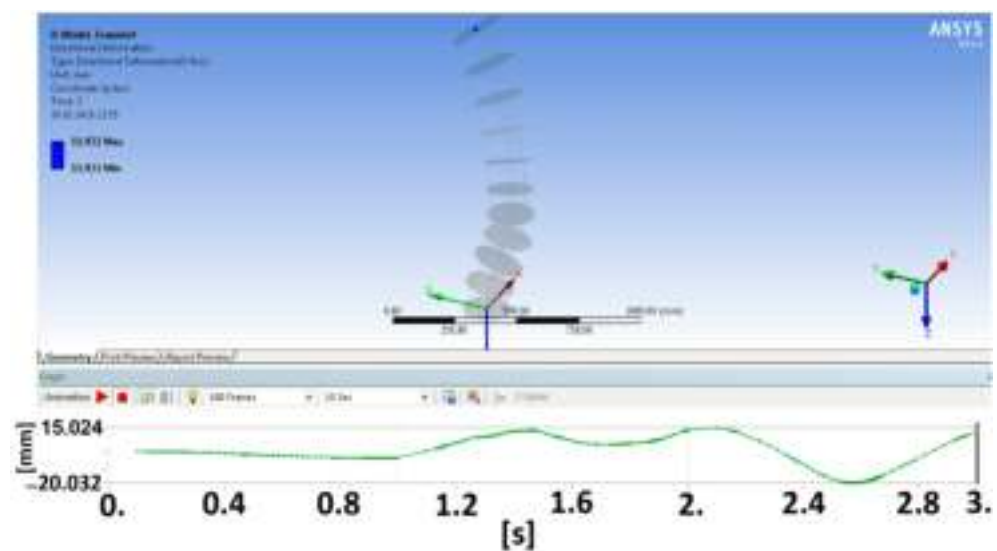


Figure 13. The displacement component variation on the X-axis for the final disk from the distal module in a local coordinate system vs. time.

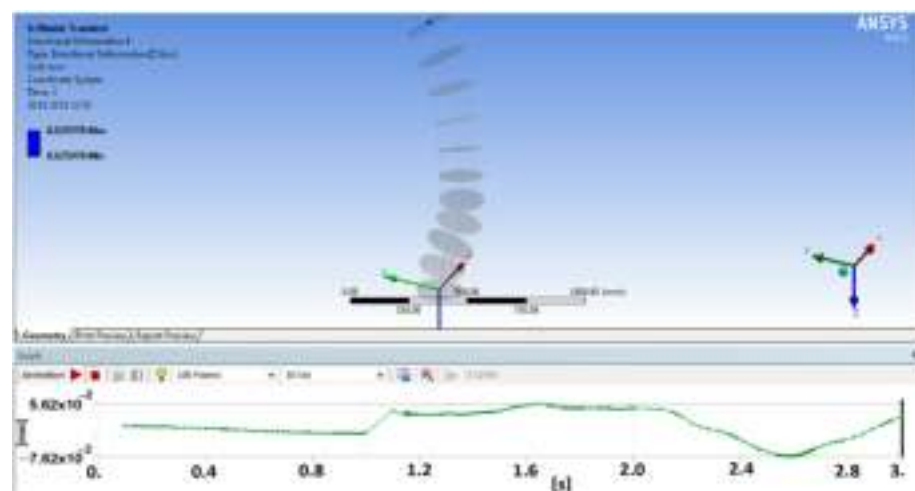


Figure 14. The displacement component variation on the Z-axis for the final disk from the distal module in a local coordinate system vs. time.

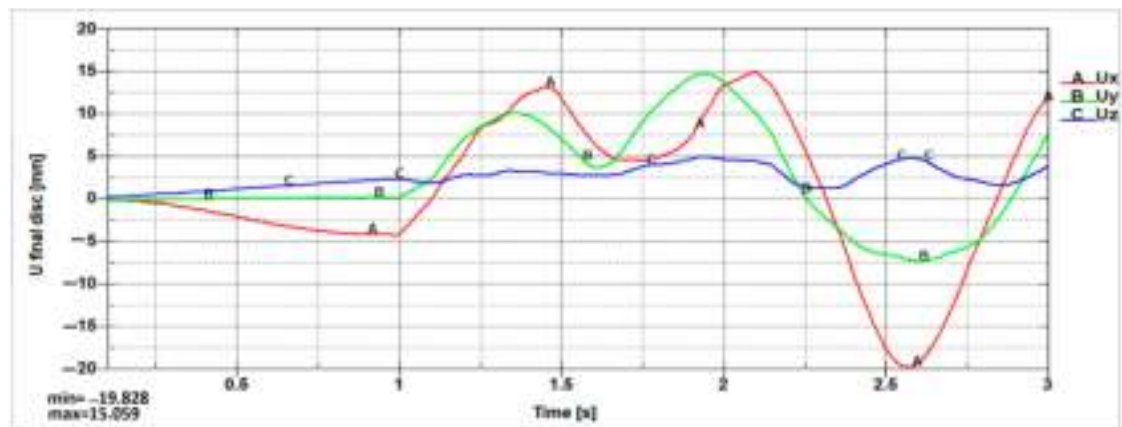


Figure 15. The elastic displacement variation of the final disk marker from the proximal module vs. time.

Each module is successively actuated or simultaneously with the other two, according to the desired configuration of the robotic system described in the proposed workspace.

The axes that define the analyzed flexible unit transversal plane are X and Y. Thus, Figure 13 shows a deformed shape of the analyzed flexible unit, depending on the time corresponding to the X-axis component for the final disk D11 mass center of the distal module. This variation was calculated according to a local reference system placed on the D7 base disk. It was recorded until a value of 15.059 mm was reached. Similarly, we calculated the displacement for the Z-axis variation, depending on the time for the mass center final disk D11 of the distal module. This is shown in Figure 14, and a maximum value of 0.056298 mm was obtained.

Figure 15 shows a total displacement ranging from a minimum value of -19.828 mm to a maximum of 15.059 mm. Thus, a total displacement amplitude of approximately 35 mm was obtained.

Figure 16 outlines the deformed shape of the flexible unit and the displacement variation on the Y-axis.

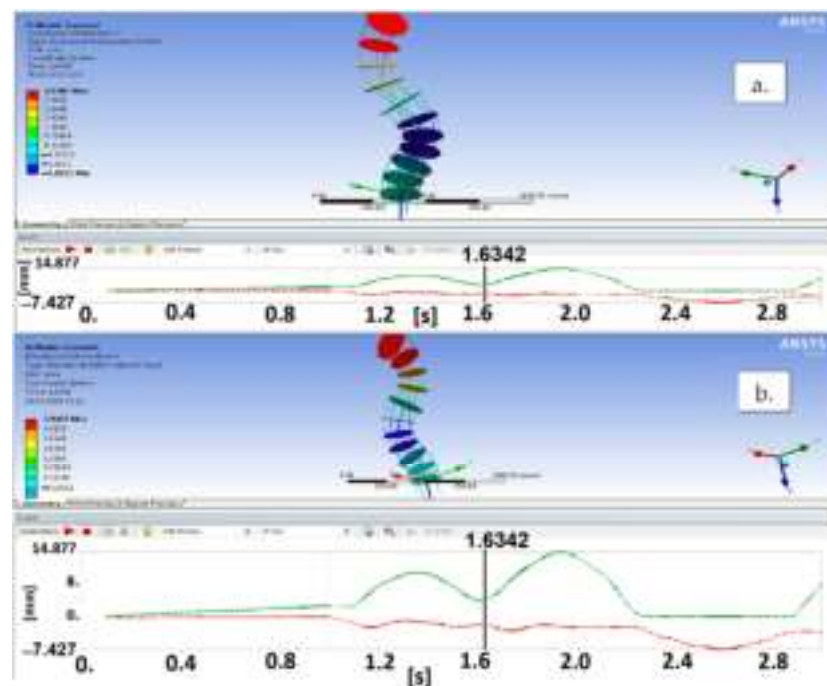


Figure 16. The variation of the deformed shape of the flexible unit across the Y-axis of the local coordinate system across time ((a)—isometric view; (b)—rotated view in the X–Y plane).

Based on this result, it can be seen that the transversal displacements are compatible with the deformed shape of the flexible unit. Also, the deformed shape can be evaluated by considering the elastic vertebrae from the flexible unit structure, namely Figure 17, where the resulting displacement variation is shown for two vertebrae, namely a motor one and a driven vertebra.

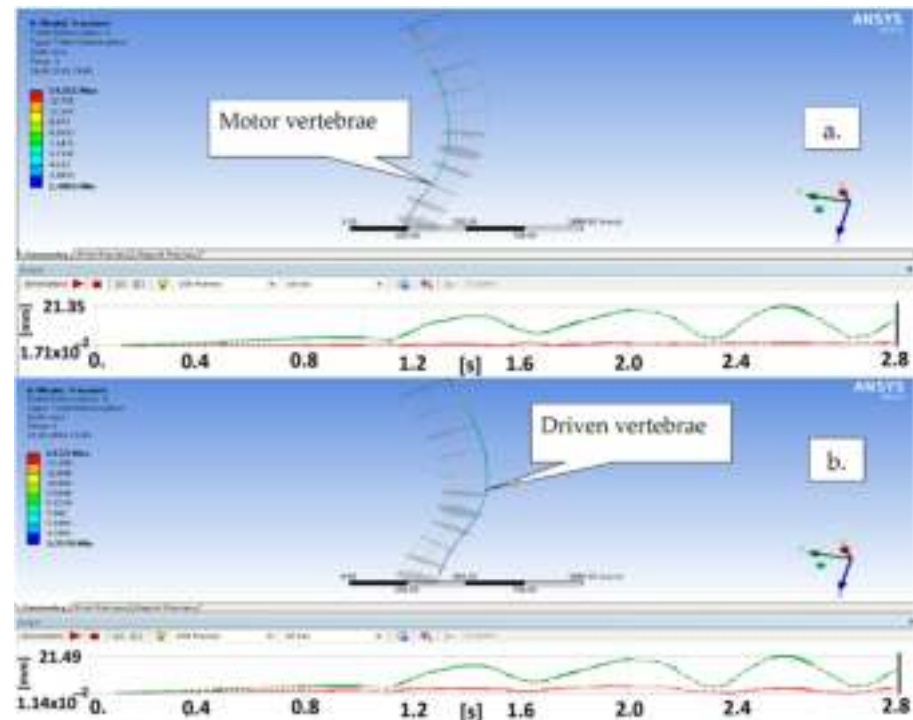


Figure 17. The resulting displacement variation for vertebrae in a global coordinate system vs. time ((a)—motor vertebrae; (b)—driven vertebrae).

Finally, in Figure 17 it was analyzed the displacement variations for distal module vertebrae.

Also, the deformations, displacements and von Mises stress were obtained in the case of each vertebra for two distinctive cases: when this was a motor element and when this was a driven element.

Thus, in Figure 18, we can see the resulting displacements for median module vertebrae over a time interval of 3 s. After numerical processing, values of 4.3 mm were obtained for the motor vertebra and 4.1 mm for the driven vertebra.

The oscillation amplitude is correlated with the stiffness of all the modules of the flexible unit and also with the setup of the actuation commands.

Based on Figure 19, the resulting displacements of vertebrae from the proximal module actuation are characterized by maximum displacement curves that are almost identical and reach values around 2.288 mm. Also, different displacement values were obtained for all the modules during numerical processing. This occurs due to the stiffness that will be increased by passing from the distal module to the median and then to the proximal module, which has a higher stiffness than the other two.

Other results, like the resulting elastic displacements for the entire flexible unit on different temporal time moments, are reported in Figure 20. Thus, in Figure 20, we can remark on the deformed shape that is compatible with the recorded variations in the case of the resulting elastic displacements. Also, the total elastic deformation has a maximum value of 21.83 mm, and its shape variation is compatible with the deformed flexible shape.

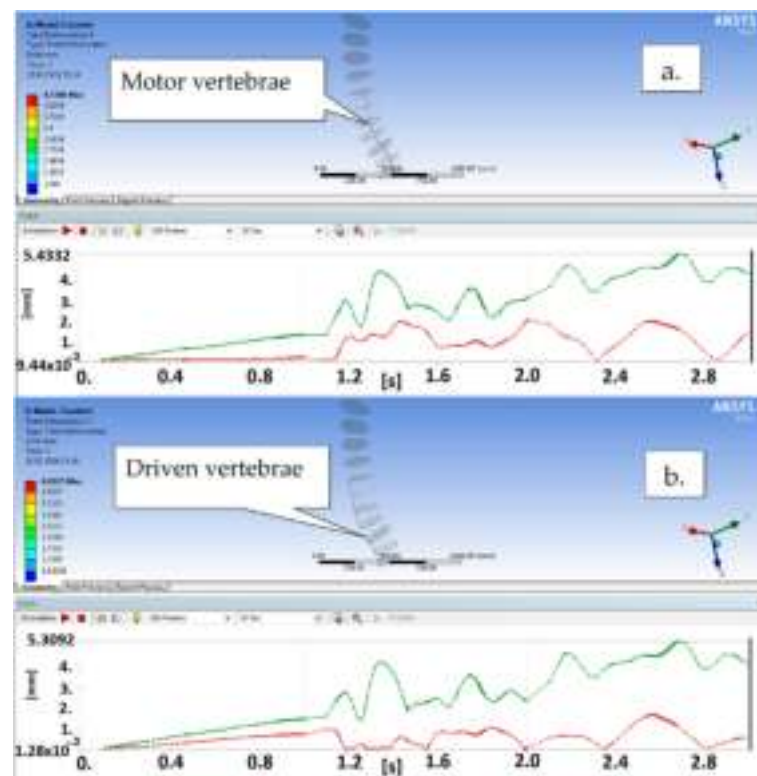


Figure 18. The resulting displacement variation for the vertebrae of the median module in the global coordinate system vs. time ((a)—motor vertebrae; (b)—driven vertebrae).

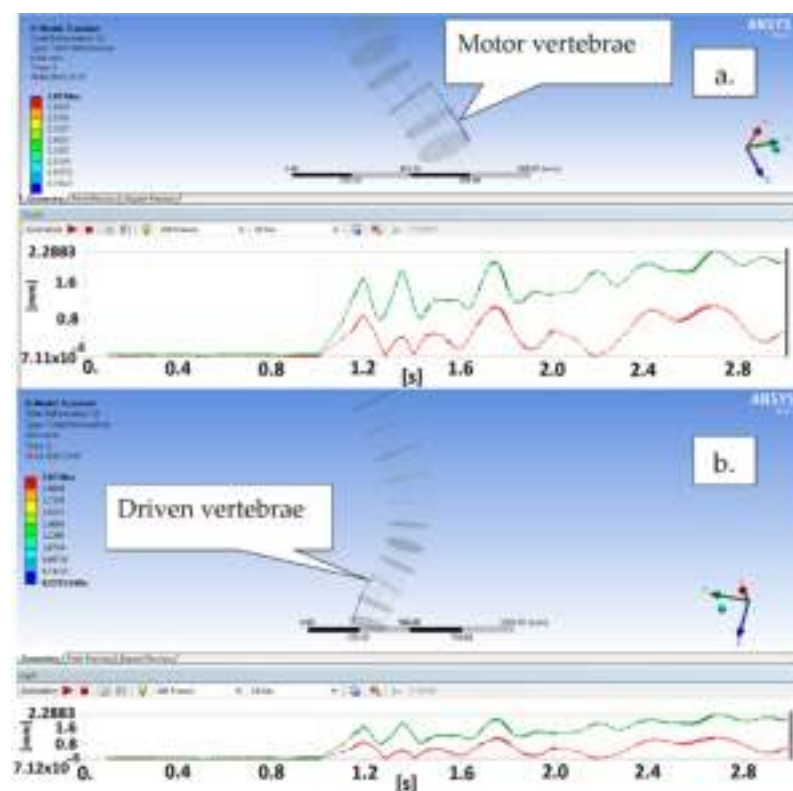


Figure 19. The resulting displacement variation for the vertebrae of the proximal module in the global coordinate system vs. time ((a)—motor vertebrae; (b)—driven vertebrae).

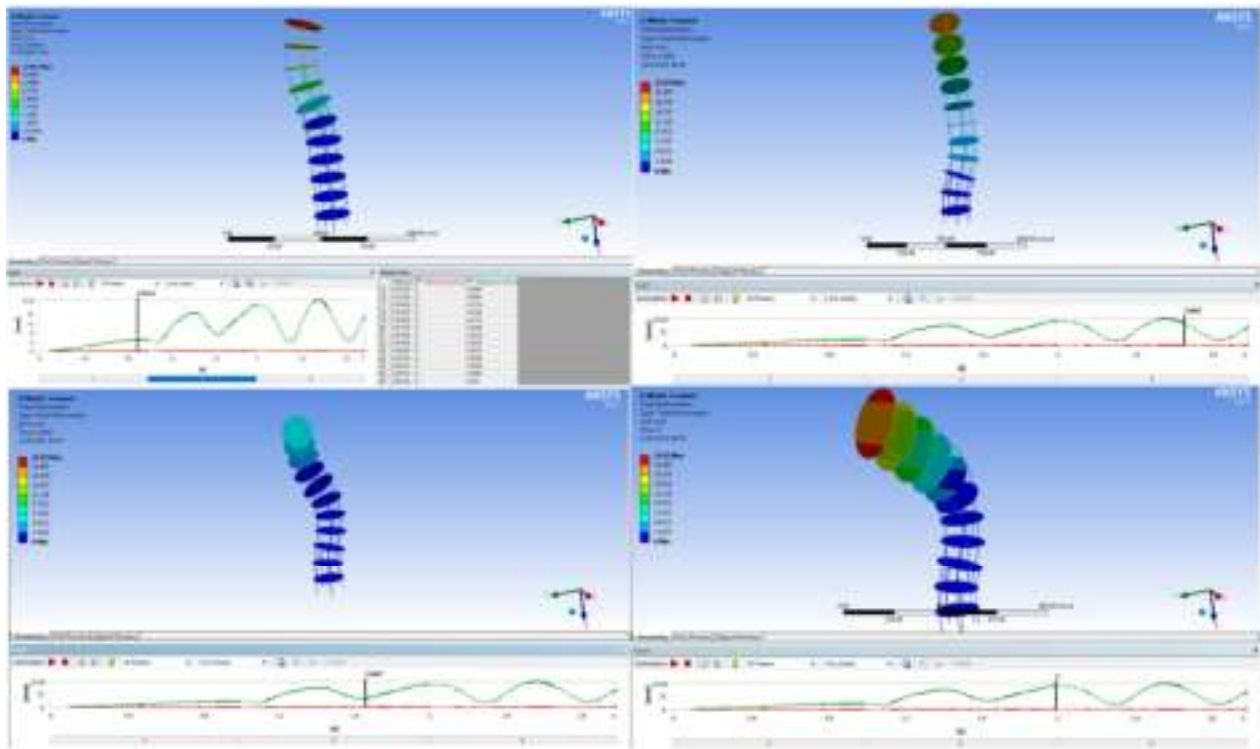


Figure 20. Snapshots of the acquired distribution of elastic displacement for various temporal moments vs. time.

The major results obtained through this analysis are represented by the von Mises stress and deformations of the flexible unit in the case of the interest markers allocated on the final disk of each module. These are reported in Figures 21–23.

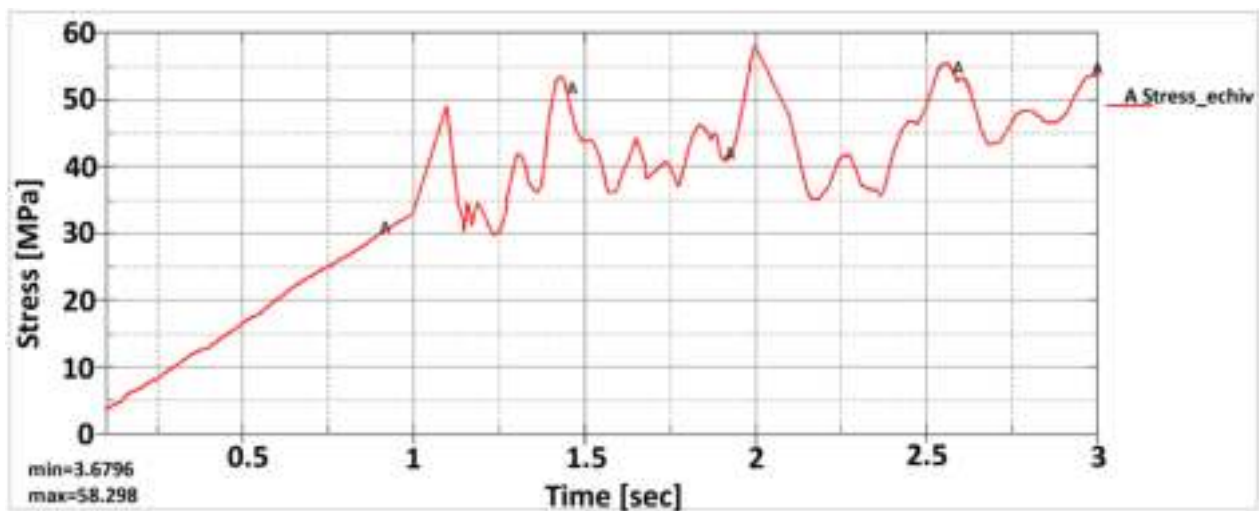


Figure 21. Equivalent von Mises stress distribution vs. time.

For the results of von Mises stress, the maximum value was 58.298 MPa, and the stress distribution was also influenced by the actuation setup and stiffness for all modules of the flexible unit structure. Thus, Figure 22 shows a smooth linear evolution. For the first step interval according to Figure 21, at a range of 32 MPa, the equivalent von Mises stress evolution will have an unbalanced behavior between 32 MPa and 58.298 MPa. Another particular aspect is that the equivalent von Mises stress starts from a minimum value of

3.67 MPa. This value results from an initial stiffness of the poly-articulated unit, even when it is not actuated.

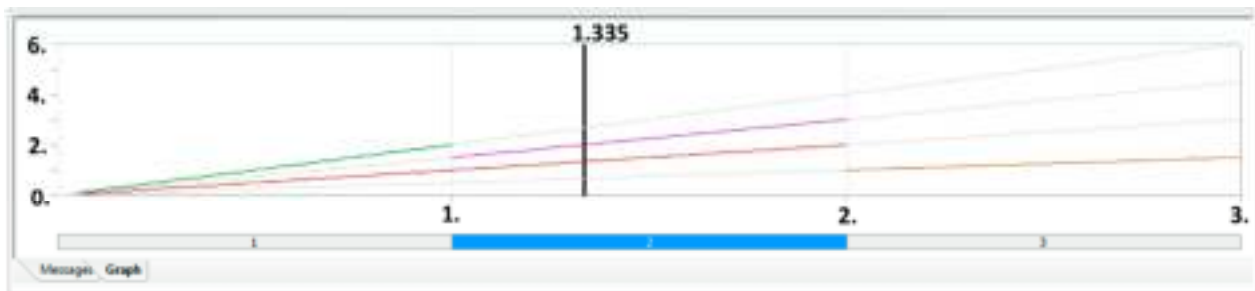


Figure 22. Actuation modes represented in steps vs. time for the whole flexible unit.

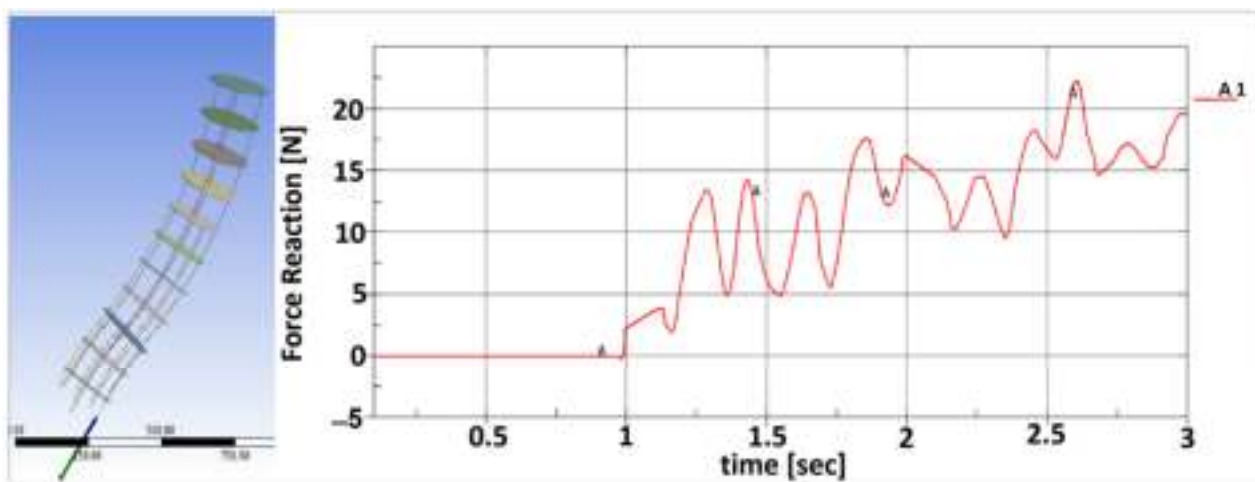


Figure 23. The reaction force distribution for the motor vertebra of the distal module depends on the time interval of 2 s (1 to 3 s).

Other results, like the reaction forces from several flexible unit components, are reported in diagrams from Figures 24–27. By having in sight these results, along the 1–3 s time interval, the force of the drive vertebra used for final disk actuation from the distal module, namely D11, reaches a value of 25 N, as shown in Figure 23.

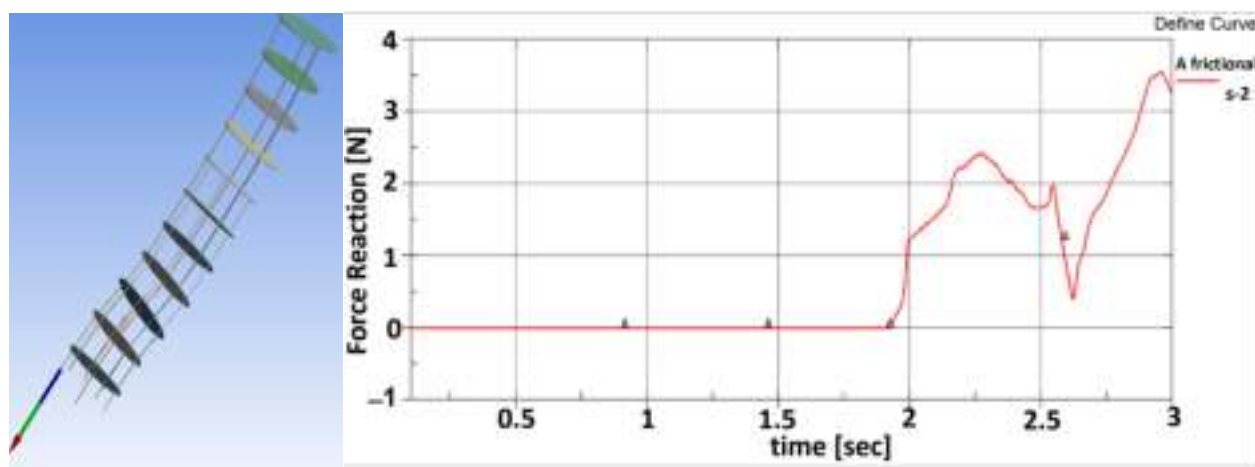


Figure 24. The reaction force distribution for a driven vertebra of the distal module depends on the time interval of 1 s.

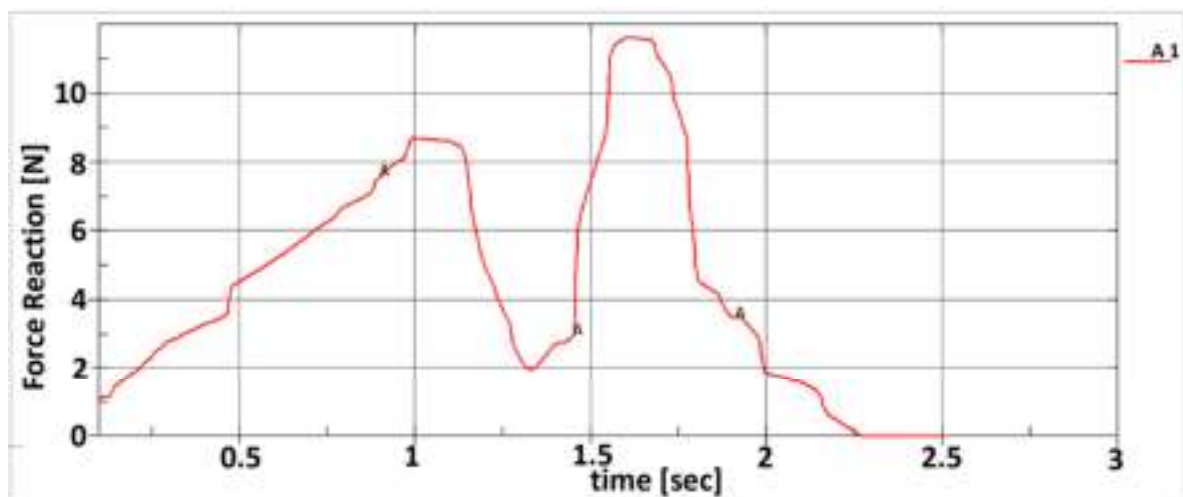


Figure 25. The reaction force distribution for the motor vertebra of the median module depends on the time interval of 2 s.

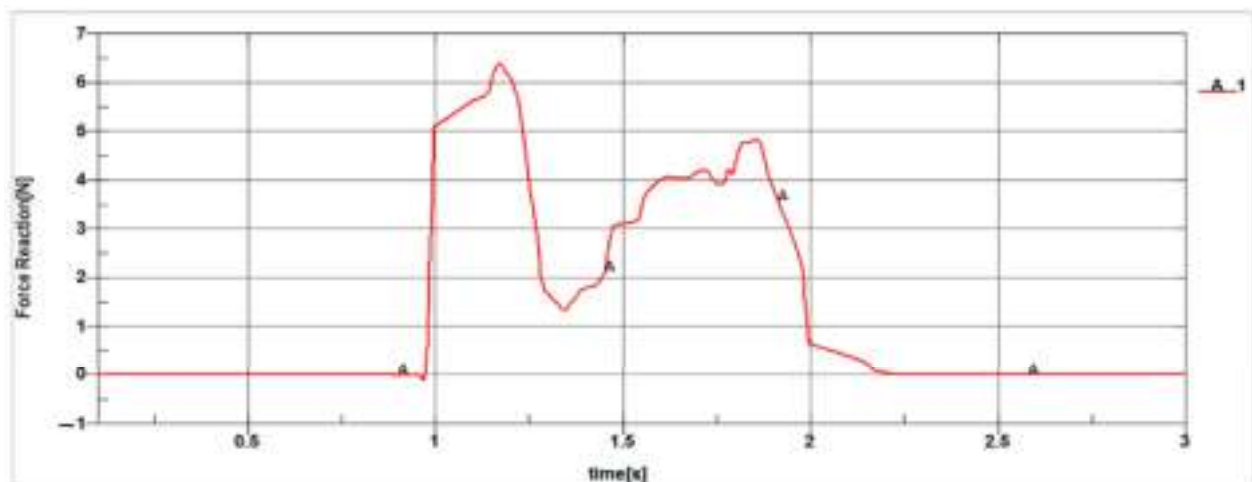


Figure 26. The reaction force distribution for a driven vertebra of the median module depends on the time interval of 1 s.

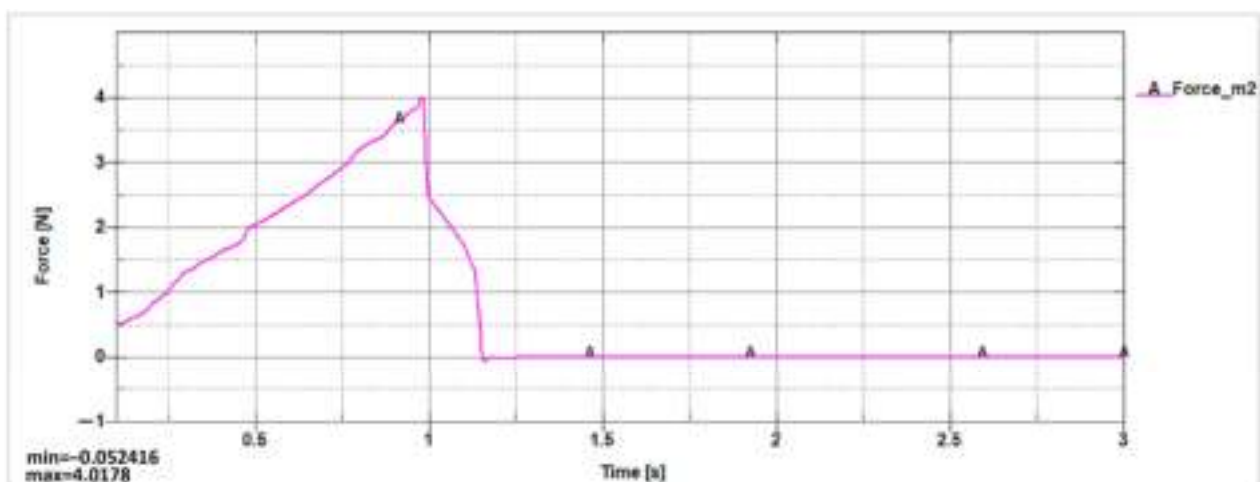


Figure 27. The reaction force distribution for a driven vertebra of the proximal module depends on a time interval of 1 s.

In Figure 24, the obtained reaction force for driven vertebra during the actuation of the distal module has a variation ranging from 0 to 3.5 N. This happens along a 2–3 s time interval.

For the actuation of the distal module, the drive vertebra actuation force records an almost linear time variation, as can be seen in Figure 25. This was produced in a period of 0 to 2 s and reached a maximum value of 11.8 N.

The reaction force from the median module-driven vertebra became active during the period of 1 to 2 s and reached values ranging from 0 to 6.5 N, according to Figure 26.

Figure 27 shows a reaction force on the first time interval ranging from 0 to 2 s. This reaches a maximum value of 4 N at the end of the actuation period. This also has a smooth linear evolution and starts from 0.5 N, when the entire poly-articulated unit will be actuated.

The distribution of the contact parameter results was analyzed for an intermediary disk and the central vertebrae. The obtained results were represented by frictional stress, penetration factor and contact pressure for a time interval of 3 s. These results are shown in Figures 28–30.

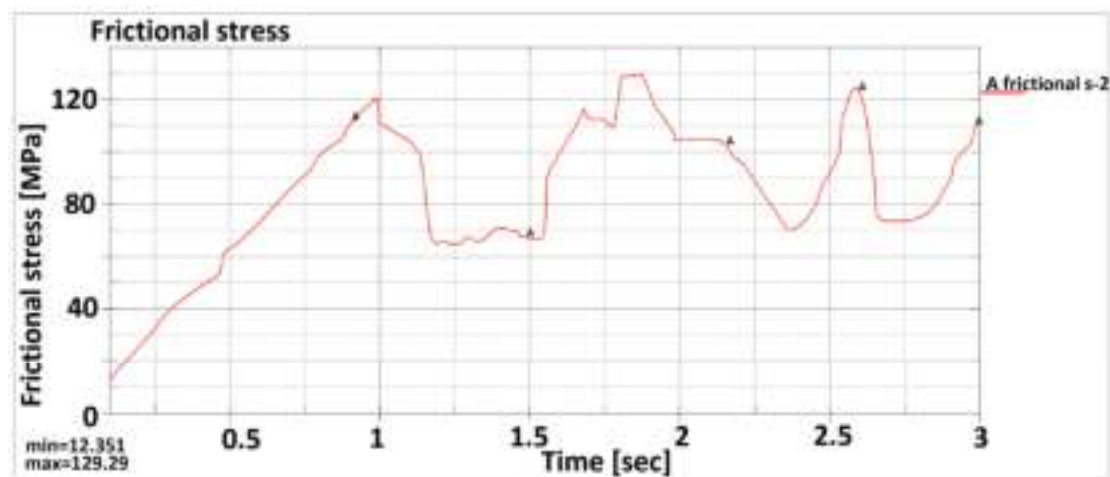


Figure 28. The frictional stress distribution depends on time.

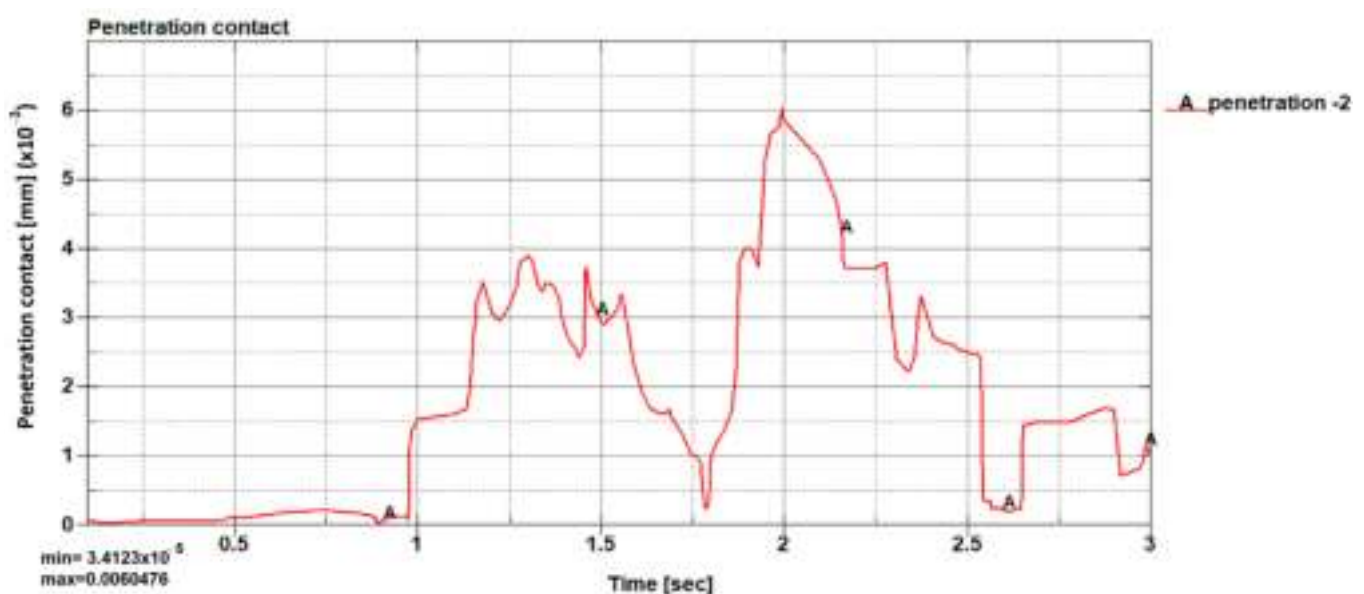


Figure 29. The penetration contact variation for the mechanical system contact definition vs. time.

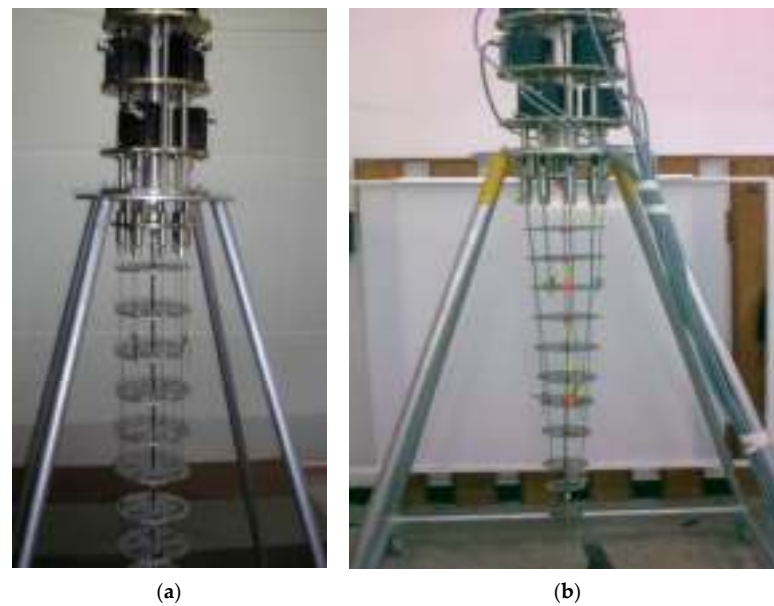


Figure 30. The poly-articulated robotic system prototype equipped with the two flexible unit forms: (a) cylindrical form and (b) tronconic form.

Figure 29 shows that the penetration factor between contact surfaces and the targeted surfaces is very small and can be neglected. According to Figure 29, this reaches a maximum value of 0.0060476 mm. During actuation, the frictional stress reaches a value of almost 130 MPa, and it has an unbalanced behavior due to the friction between elements and straight contact, as can be seen in Figure 28. During simulations, a symmetric pressure distribution can be noticed when a module is actuated, and another is released. The oscillations, which occur during virtual simulations that can be remarked on the variation curves from Figures 25–29, are caused mainly by contact definition parameters setup, contact type, integration step and method, correlated to the time interval for monitoring the system motion.

6. Prototype of the Poly-Articulated Robotic System

The experimental prototype was created according to the design parameters imposed by the 3D virtual model, but this was also validated through the virtual simulation outputs. This has two main units, namely one for the robotic system actuation and the other represented by the flexible unit. The translation motion was transformed and transmitted from the rotational motors through nine nut–screw mechanical transmissions to the considered module from the flexible unit. The actuation unit also has mechanical devices, which are used for axial displacement compensation. The designed prototypes for both flexible unit forms, i.e., the cylindrical and the tronconic one, are shown in Figure 30. The actuation unit, which is common to both solutions, is shown in Figures 31 and 32.



Figure 31. The poly-articulated robotic actuation unit.



Figure 32. Guiding systems for the nut–screw mechanical transmissions.

7. Experimental Motion Analysis

Both prototypes were experimentally evaluated through several tests in dynamic conditions. The purpose of this analysis is to validate the experimental prototype and also to analyze the command and control protocols. In order to achieve the proposed aims, CONTEMPLAS video motion equipment was used [27]. The principle of this analysis is shown in Figure 33. Thus, the experimental test results of the evaluated robotic system are presented for one solution, namely the one with a tronconic form.

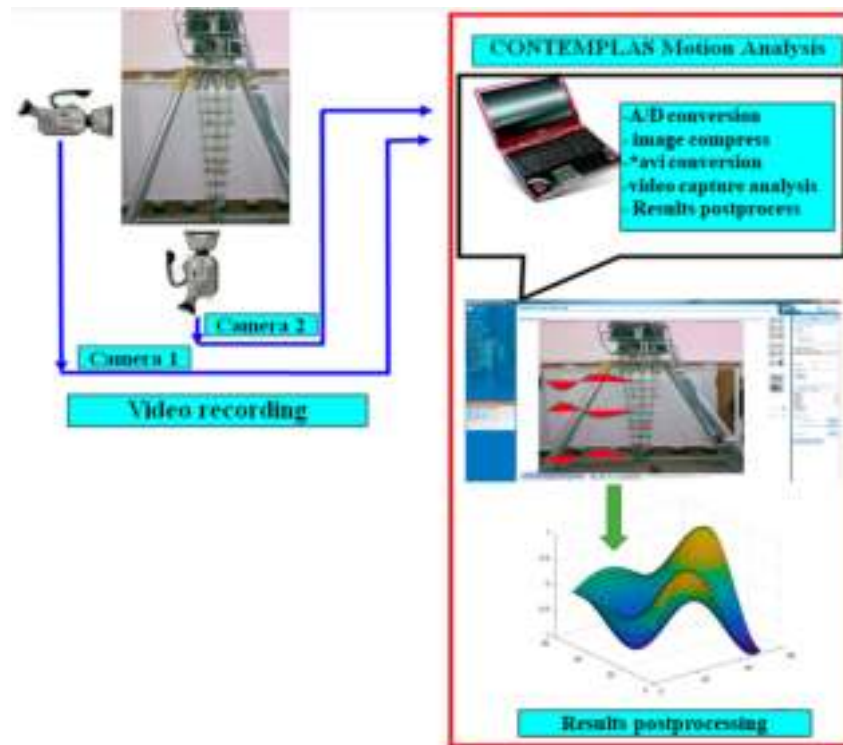


Figure 33. CONTEMPLAS motion analysis workflow.

This equipment has two cameras that can record video sequences up to 350 frames/second. The video recording was done by keeping the same conditions imposed through virtual simulations, namely period and working actuation modes for all modules. The first step of the proposed experimental analysis was to perform a setup for the equipment and also for the poly-articulated robotic system in order to evaluate its workspace. This is shown in Figure 34. The video cameras were placed at an angle of 90 degrees between them and a distance of 2.5 m from the analyzed prototype, as can be seen in Figure 34.

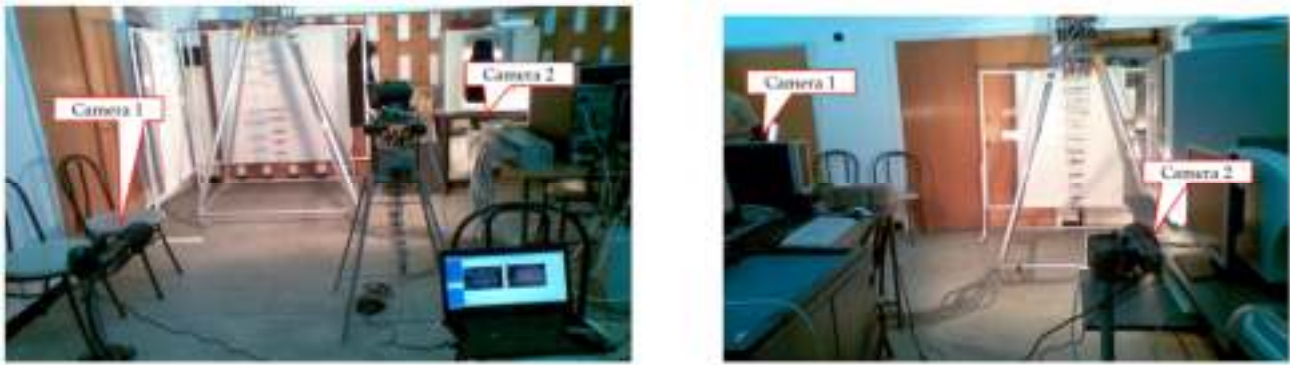


Figure 34. Experimental analysis setup for the analyzed robotic system.

Twelve reflexive markers were attached to the poly-articulated robotic system to evaluate the workspace, which was characterized by the developed experimental trajectories. The equipment hardware and Templo Plus V2010 software can identify the marker position in each sequence of the recorded video sequence. In this case, the task for the analyzed markers attached to the central vertebrae was trajectory variations, according to Figure 35. Some snapshots during trajectory evaluation can be observed in Figure 36.



Figure 35. Marker attachment on the poly-articulated robotic system.

By having in sight the obtained experimental analysis tests, a comparative study was done based on the obtained results through virtual simulations and the ones from experimental tests. This analysis was done with LS-Dyna R7.1 software in order to establish the accuracy of the performed motion. Thus, the most important results are represented through the total deformation of the entire poly-articulated robotic system. This comparative study will be focused on monitoring the central marker position situated on the end disk from the distal module. These results are shown in Figures 37–39, where the behavior of the deformation components represents the proof of validation of the poly-articulated robotic system prototype.



Figure 36. Some snapshots during the experimental tests for trajectory evaluation.

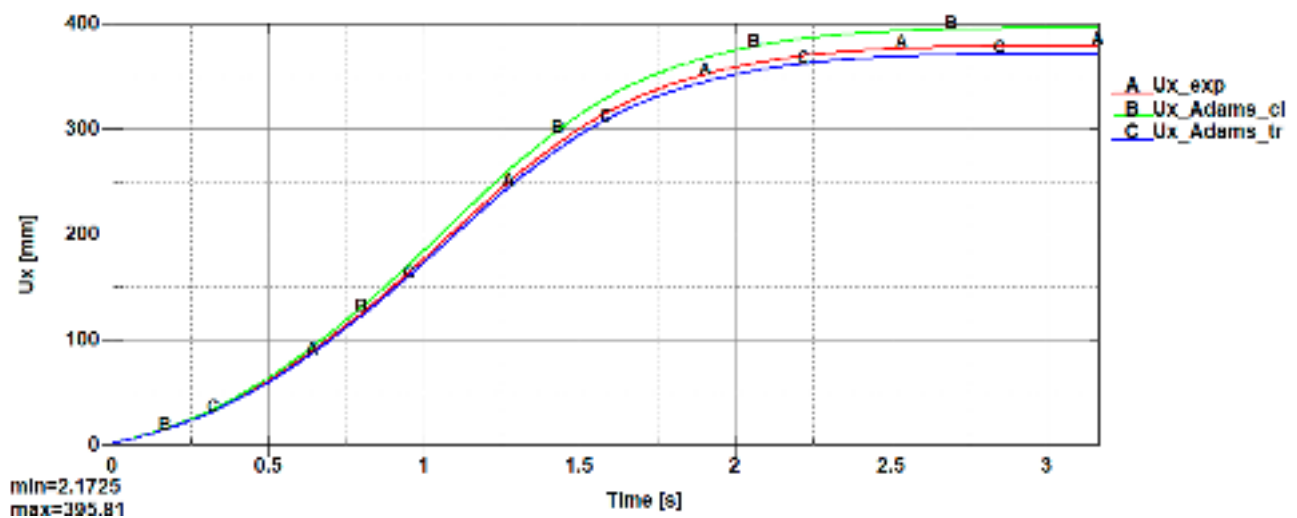


Figure 37. Total displacement on the X-axis of the poly-articulated robotic system vs time: A—Experimental tronconic model; B—Adams virtual model with a cylindrical flexible unit; C—Adams virtual model with a tronconic form of the flexible unit.

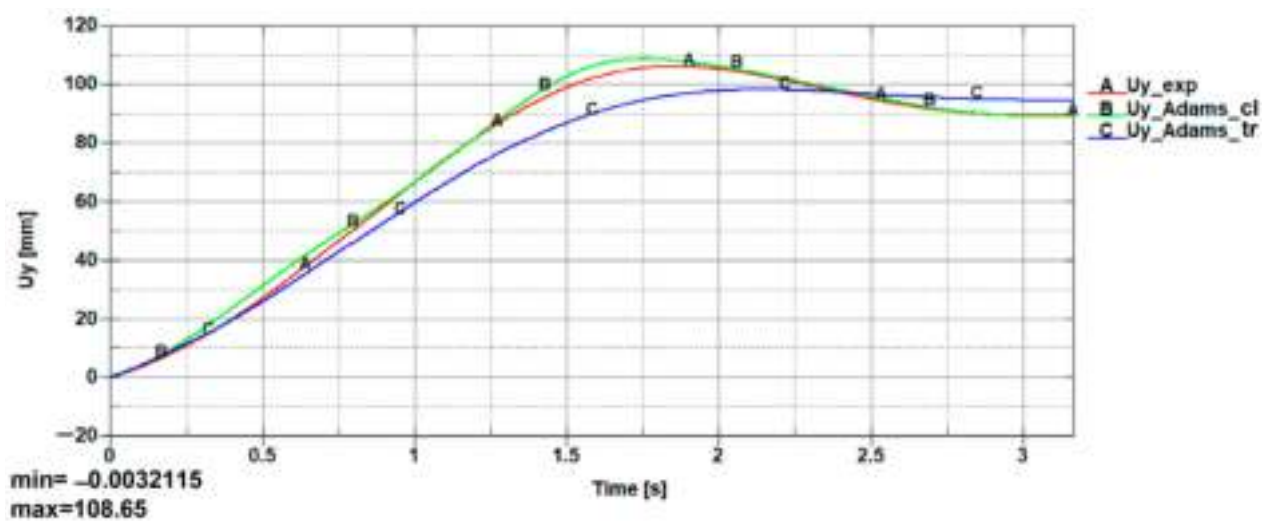


Figure 38. Total displacement on the Y-axis of the poly-articulated robotic system vs time: A—Experimental tronconic model; B—Adams virtual model with a cylindrical flexible unit; C—Adams virtual model with a tronconic form of the flexible unit.

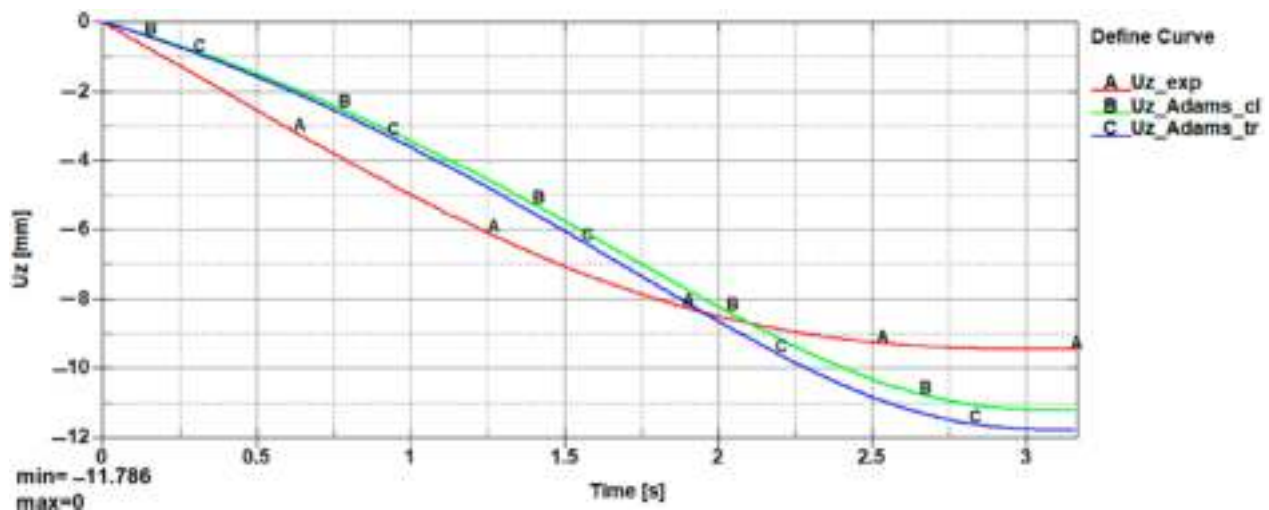


Figure 39. Total displacement on the Z-axis of the poly-articulated robotic system vs time: A—Experimental tronconic model; B—Adams virtual model with a cylindrical flexible unit; C—Adams virtual model with a tronconic form of the flexible unit.

Thus, Figure 37 shows a diagram of the X-direction total displacement for the obtained prototype, respectively, virtual simulations carried out with MSC Adams software. From this, it can be remarked that all models have the same trajectory, ranging from 0 to a maximum value of 398.22 mm obtained for the poly-articulated robot virtual model with a cylindrical form of the flexible unit. In the case of the experimental prototype, this reaches a value of 382.31 mm and is much closer to the third one, respectively, to the one characterized by a flexible unit with a tronconic form. The third one reaches a value of 380.77 mm. Thus, the accuracy is around 4.188%.

By having in sight the plot reported in Figure 38, it can be observed that all paths of the analyzed models have an appropriate trajectory in the case of the Y-axis total displacement, and a maximum value of 109.63 mm was recorded, which corresponds to the poly-articulated robot virtual model characterized by a flexible unit with a cylindrical form. In the case of the prototype analyzed during experimental tests, the maximum recorded value was 106.32 mm. In the third case, namely the virtual model characterized by the tronconic form, this reaches a maximum value of 99.64 mm. In this case, an accuracy of 6.064% was established.

A similar comparative analysis was done in the case of total displacement on the Z-axis, as it was reported in the diagram from Figure 39. In this case, all paths have appropriate trajectories, but different values were recorded for the analyzed models. By having in sight the poly-articulated virtual model simulated with MSC Adams, with a tronconic form of the flexible unit, a minimum value of −11.774 mm can be observed, but a minimum value of −9.678 mm was recorded for the experimental analysis of the poly-articulated robot prototype. In the latter case, namely the robotic system virtual model with a cylindrical form of the flexible unit, a minimum value of −11.33 millimeters was recorded, which results in an accuracy of 8.783%.

8. Conclusions

A prototype of a poly-articulated robotic system was validated in this research. This robot can be used in several fundamental domains like the automotive industry, medicine, military defense and even architecture. The virtual prototyping of the robotic system was developed with MSC Adams. An interface was defined for importing the proposed CAD model with all mechanism components considered as subassemblies and also considering the imposed contour conditions. An equivalent kinematic model of the proposed mechanical system was built by having in sight the kinematic joints and some local reference systems located in the mass centers of the flexible unit links, respectively, in the joint

centers. The motor joints were defined according to motion laws set for all actuators on a time interval of 3 s. A procedure was identified for transforming the solid bodies into deformable ones by meshing them in tetrahedral finite elements. This was done by taking into account the connection type between elements and the definition of the node setup in the master and slave mode.

An elastic dynamic analysis was performed, which implies crossing two stages: a modal analysis for evaluating the natural frequencies and the mode shapes for the flexible kinematic elements and a dynamic analysis of the flexible unit based on the Craig–Bampton method in order to establish the kinematic parameter variation, depending on time.

The prototype validation core focuses on the motion of central vertebrae as a rigid body but also as a deformable body. The motion analysis was performed on specific nodes/markers, depending on the global reference system and on other local reference systems.

The rigid body motion is defined through a generalized coordinate system set that represents the mobile reference system origin identification according to the global reference system and its orientation through Euler angles. The deformable body motion was defined through a modal coordinate set based on the Craig–Bampton method.

The elastic deformation variations were determined depending on the time for angular and linear displacements for the end node placed in the central vertebrae. It is important to remark that the diagram results represented through numerical values and curve trajectories from virtual simulations are similar to the ones obtained in the experimental analysis for the designed robotic system prototype.

The virtual prototyping created with Ansys has a complex nature, using the dynamic characterization of the contact models created between deformable elements and disks. The performed dynamic analysis has a nonlinear behavior.

For the contact definition, the following conditions were taken into account: the translational motion between actuation vertebrae and disks, the finite element network at the level of contact areas, penetration phenomena between the contact surfaces and targeted ones and differential equations between the contact type and integration method. The processed model evaluated with Ansys software is completely different from the one analyzed with MSC Adams, because all the elements are in contact, and these are considered deformable ones.

Variation laws of the kinematic parameters were identified through this analysis, depending on time and stress, deformations, strain and the resulting actuation forces from the drive elements. The flexible unit deformed shapes were obtained through numerical simulations. These shapes define the robotic system workspace and are influenced by the time setup for the command and control procedure. The deformation behavior of the central vertebrae was kept in sight through time variation diagrams.

The major contribution brought about by this research is represented by the contact model development with dumping control possibilities; the penetration depth between different surface types such as contact and target; the normal stiffness factor for tests and the choice of suitable ones from ANSYS-allowed formulations (Pure penalty, MPC, Augmented Lagrange and Normal Lagrange). Thus, the obtained flexible unit geometrical model was analyzed and parametrized with the aid of the DesignModeller module from ANSYS R14.5 software, with a view to envisaging future research like optimization processes of the components in a dynamic mode.

The virtual models of the flexible unit achieved in MSC Adams 2012 and ANSYS R14.5 software are parametrized, with optimization possibilities under kinematic- and dynamic-specific conditions.

The robotic system architecture, which includes the actuation system, i.e., the command and control unit, represents the fundamental base of a robotic system conceptual design for a future miniaturized robotic system especially designed for minimally invasive surgery procedures like [17].

Author Contributions: Conceptualization, N.D. and S.D.; methodology, S.D. and N.D.; software, I.G.; validation, C.C., S.D. and N.D.; formal analysis, I.G.; investigation, C.C.; resources, S.D.; data curation, S.D. and I.G.; writing—original draft preparation, C.C.; writing—review and editing, C.C. and N.D.; visualization, N.D.; supervision, C.C. and N.D.; project administration, N.D.; funding acquisition, N.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: All the data generated in the current research is pre-sented in the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Bajo, A.; Simaan, N. Hybrid Motion/Force Control of Multibackbone Robots. *Int. J. Robot. Res.* **2015**, *35*, 1–13. [\[CrossRef\]](#)
2. Hannan, M.; Walker, I. The ‘Elephant Trunk’ Manipulator, Design and Implementation. In Proceedings of the IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM), Como, Italy, 8–12 July 2001; pp. 14–19.
3. Hirose, S. *Biologically Inspired Robots: Snake-like Locomotors and Manipulators*; Applied Mechanics Reviews; Oxford University Press: Oxford, UK, 1995; Volume 48, p. 27.
4. Kim, Y.; Cheng, S.; Kim, S.; Iagnemma, K. Design of a Tubular Snake-Like Manipulator with Stiffening Capability by Layer Jamming. In Proceedings of the 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, Vilamoura, Portugal, 7–12 October 2012.
5. Roppenecker, D.B.; Schuster, L.; Coy, J.A.; Traeger, M.F.; Entsfellner, K.; Lueth, T.C. Modular Body of the Multi Arm Snake-Like Robot. In Proceedings of the International Conference on Robotics and Biomimetics, Bali, Indonesia, 5–10 December 2014.
6. Gu, X.; Ren, H. A Survey of Transoral Robotic Mechanisms: Distal Dexterity, Variable Stiffness, and Triangulation. *Cyborg Bionic Syst.* **2023**, *4*, 0007. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Shi, L.; Ma, Z.; Yan, S.; Zhou, Y. Cucker–Smale Flocking Behavior for Multiagent Networks With Coopetition Interactions and Communication Delays. *IEEE Trans. Syst. Man Cybern. Syst.* **2024**, 1–10. [\[CrossRef\]](#)
8. Shi, L.; Ma, Z.; Yan, S.; Ao, T. Flocking Dynamics for Cooperation–Antagonism Multi-Agent Networks Subject to Limited Communication Resources. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2024**, *71*, 1396–1405. [\[CrossRef\]](#)
9. Kim, S.-H.; Choi, H.-S.; Keum, B.; Chun, H.-J. Robotics in Gastrointestinal Endoscopy. *Appl. Sci.* **2021**, *11*, 11351. [\[CrossRef\]](#)
10. Clark, A.B.; Mathivannan, V.; Rojas, N. A Continuum Manipulator for Open-Source Surgical Robotics Research and Shared Development. *IEEE Trans. Med. Robot. Bionics* **2021**, *3*, 277–280. [\[CrossRef\]](#)
11. Ivanescu, M.; Cojocaru, D.; Bizdoaca, N.; Florescu, M.; Popescu, N.; Popescu, D.; Dumitru, S. Boundary Control by Boundary Observer for Hyper-redundant Robots. *Int. J. Comput. Commun. Control* **2010**, *5*, 755–767. [\[CrossRef\]](#)
12. Xu, S.; He, B.; Zhou, Y.; Wang, Z.; Zhang, C. A Hybrid Position/Force Control Method for a Continuum Robot With Robotic and Environmental Compliance. *IEEE Access* **2019**, *7*, 100467–100479. [\[CrossRef\]](#)
13. Wooten, M.B.; Walker, I.D. Environmental Interaction With Continuum Robots Exploiting Impact. *IEEE Robot. Autom. Lett.* **2022**, *7*, 10136–10143. [\[CrossRef\]](#)
14. Chen, Y.; Wang, L.; Galloway, K.; Godage, I.; Simaan, N.; Barth, E. Modal-Based Kinematics and Contact Detection of Soft Robots. *Soft Robot.* **2020**, *8*, 298–309. [\[CrossRef\]](#)
15. Liu, H.; Farvardin, A.; Pedram, S.A.; Iordachita, I.; Taylor, R.H.; Armand, M. Large Deflection Shape Sensing of a Continuum Manipulator for Minimally-Invasive Surgery. In Proceedings of the IEEE International Conference on Robotics and Automation ICRA, Seattle, WA, USA, 26–30 May 2015; pp. 201–206.
16. Desai, J.P.; Sheng, J.; Cheng, S.S.; Wang, X.; Deaton, N.J.; Rahman, N. Towards Patient-Specific 3D-Printed Robotic Systems for Surgical Interventions. *IEEE Trans. Med. Robot. Bionics* **2019**, *1*, 77–87. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Dumitru, N.; Ciurezu-Gherghe, L.; Copilusi, C.; Geonea, I.; Dumitru, S. Theoretical and experimental study methods for a robotic system with deformable elements used in minimally invasive surgery. *Mech. Mach. Theory* **2022**, *167*, 104459. [\[CrossRef\]](#)
18. Zhang, Z.; Dequidt, J.; Back, J.; Liu, H.; Duriez, C. Motion control of cable-driven continuum catheter robot through contacts. *IEEE Robot. Autom. Lett.* **2019**, *4*, 1852–1859. [\[CrossRef\]](#)
19. Oliver-Butler, K.; Till, J.; Rucker, D.C. Continuum robot stiffness under external loads and prescribed tendon displacements. *IEEE Trans. Robot.* **2019**, *35*, 403–419. [\[CrossRef\]](#)
20. Goldman, R.E.; Bajo, A.; Simaan, N. Compliant Motion Control for Multisegment Continuum Robots with Actuation Force Sensing. *IEEE Trans. Robot.* **2014**, *30*, 890–902. [\[CrossRef\]](#)
21. Dumitru, S.; Cojocaru, D.; Marghitu, D. Computer Aided Design of a Hyper-redundant Manipulator. *Adv. Electr. Comput. Eng.* **2013**, *13*, 51–56. [\[CrossRef\]](#)
22. Grazioso, S.; Di Gironimo, G.; Siciliano, B. A geometrically exact model for soft continuum robots: The finite element deformation space formulation. *Soft Robot.* **2019**, *6*, 790–811. [\[CrossRef\]](#)
23. Craig, R.R.; Bampton, C.C. Coupling of Substructures for Dynamics Analyses. *AIAA J.* **1998**, *6*, 1313–1319. [\[CrossRef\]](#)
24. Duan, J.; Zhang, K.; Qian, K.; Hao, J.; Zhang, Z.; Shi, C. An Operating Stiffness Controller for the Medical Continuum Robot Based on Impedance Control. *Cyborg Bionic Syst.* **2024**, *8*, 0110. [\[CrossRef\]](#) [\[PubMed\]](#) [\[PubMed Central\]](#)

25. Zhang, S.; Li, F.; Fu, R.; Li, H.; Zou, S.; Ma, N.; Qu, S.; Li, J. A Versatile Continuum Gripping Robot with a Concealable Gripper. *Cyborg Bionic Syst.* **2023**, *4*, 0003. [[CrossRef](#)]
26. Chirikjian, G.S.; Burdick, J.W. A Modal Approach to Hyper-Redundant Manipulator Kinematics. *IEEE Trans. Robot. Autom.* **1994**, *10*, 343–353. [[CrossRef](#)]
27. CONTEMPLAS User Manual. Available online: www.contemplas.com (accessed on 23 July 2018).

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

A Novel Mixed Finite/Infinite Dimensional Port–Hamiltonian Model of a Mechanical Ventilator

Milka C. I. Madahana ^{1,*}, John E. D. Ekoru ² and Otis T. C. Nyandoro ²¹ School of Mining Engineering, University of the Witwatersrand, 2nd Floor, ARM Building, 1 Jan Smuts Avenue, Braamfontein, Johannesburg 2017, South Africa² School of Electrical and Information Engineering, University of the Witwatersrand, 2nd Floor, ARM Building, 1 Jan Smuts Avenue, Braamfontein, Johannesburg 2017, South Africa; john.ekoru@wits.ac.za (J.E.D.E.); otis.nyandoro@wits.ac.za (O.T.C.N.)

* Correspondence: milka.madahana@wits.ac.za

Abstract: Mechanical ventilation is a life-saving treatment for critically ill patients who are struggling to breathe independently due to injury or disease. Globally, per year, there has always been a large number of individuals who have required mechanical ventilation. The COVID-19 pandemic brought to light the significance of mechanical ventilation, which played a significant role in sustaining COVID-19-infected critically ill patients who could not breathe on their own. The pandemic drew the attention of the world to the shortage of ventilators globally. Some of the challenges to providing an adequate number of ventilators include: increased demand for ventilators, supply chain disruptions, manufacturing constraints, distribution inequalities, financial constraints, maintenance and logistics difficulties, training and expertise shortages, and the lack of design and development of affordable mechanical ventilators that satisfy the stipulated requirements. This research work presents the formulation of a detailed Port–Hamiltonian model of a mechanical ventilator integrated with the human respiratory system. The interconnection and coupling conditions for the various subsystems within the mechanical ventilator and the coupling between the mechanical ventilator and the human respiratory system are also presented. Structure-preserving discretization is provided alongside numerical simulations and results. The obtained results are found to be comparable to results presented in the literature. Future work will include the design of suitable controllers for the system.



Citation: Madahana, M.C.I.; Ekoru, J.E.D.; Nyandoro, O.T.C. A Novel Mixed Finite/Infinite Dimensional Port–Hamiltonian Model of a Mechanical Ventilator. *Computation* **2024**, *12*, 155. <https://doi.org/10.3390/computation12080155>

Academic Editor: Simone Brogi

Received: 3 May 2024

Revised: 23 July 2024

Accepted: 25 July 2024

Published: 31 July 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: mechanical ventilator; Port–Hamiltonian; human respiratory system; Dirac structure

1. Introduction

The World Health Organization (WHO) has “access to oxygen” on its model list as one of the essentials required by an individual, especially when they are in a critical condition health-wise and are unable to breathe on their own. Oxygen is the only listed medicine that does not have a substitute. Access to oxygen even after the worldwide pandemic is still a major dilemma in middle- and low-income countries [1]. A mechanical ventilator is a medical device that is employed to assist patients in cases where their respiratory system is not functioning well and, thus, the patient has challenges with breathing or has shortness of breath. A mechanical ventilator can also be used in cases where a patient has been sedated and is undergoing surgery.

The basic operation of a mechanical ventilator is to control a high-pressure region during the inspiration stage. In the inspiration stage, the mechanical ventilator is paused. During expiration, air flows out due to the lung’s natural recoil, which creates higher pressure in the alveoli. Due to the high demand for mechanical ventilators both in the past and at present, it is imperative that efficient and affordable mechanical ventilators be researched, modelled, designed and implemented. To ensure that robust mechanical ventilators are designed, it is important to formulate new models that can be used in the research and testing stages of mechanical ventilators.

1.1. Mathematical Models of the Respiratory System

Sound mathematical formulations of the underlying physics of the respiratory system are very important for understanding the anatomy and physiology of the organs. Classical modelling methodologies that employ a physiological or clinical perspective have been previously used mainly for data-fitting for single-compartment model parameters: for example, [2]. Human lungs are usually mathematically viewed to be made up of intricate branching networks of compliant tubes spanning a broad range of scales and flow regimes. The flow may be a highly pulsatile flow or an almost laminar diffusion-dominated process. Flow enters a sponge-like microstructure of lung tissue from the terminal branches of the airway. The lung tissue material's behaviour is very nonlinear, and for efficient gaseous exchange, the topology is optimized with small embedded blood vessels. Current research developments in the mechanical and mathematical modelling of the lungs have mainly focused on studying isolated effects that are difficult to understand and assess *in vivo*. These studies seek to advance medical imaging and functional diagnostics and assist with patient-specific treatment planning and optimization. Bates [3] used a system identification approach to provide a mathematical description of lung mechanics, with evolving intricacies in the mathematical formulations. This ranged from single linear compartments to generalized multi-compartmental models. Maury [4] further developed the concept proposed by Bates [3] and described the lung as a resistive tree. Various levels of description have been proposed, including lumped models with a small number of parameters described by ordinary differential equations (ODEs) up to infinite-dimensional models based on partial differential equations (PDEs). Resistive networks and how they can be used to investigate the dependence of the lung's resistance on geometrical characteristics have also been presented. Analyses of the various models from a theoretical perspective have been provided. The approximate solution techniques that have been provided have shown good correlation with experimental measurements. Other researchers that have proposed similar works are Burrowes et al. [5], Tawhai et al. [6], and Berger et al. [7]. Research has also been conducted by Roth et al. [8] using Navier–Stokes equations for fluid flow from a very mathematical perspective. This research work also provides comprehensive and computationally fully resolved fluid representations of human respiratory models or parts thereof. Solutions of the three-dimensional Navier–Stokes equations on rigid geometries have also been formulated. Information from medical imaging alongside the fluid domain have been used to formulate real fluid–structure interaction effects [8]. As an approach for replicating real fluid–structure effects, techniques have been developed and published that demonstrate this concept in the fluid domain by the application of data from medical imaging information. To quantify particle deposition or drug delivery in the lower and upper airways and for investigation of patterns in the lungs such as jets or vortices, a fully resolved fluid model may be used. Such models are very complex, and in most cases, only parts of the human respiratory system are investigated, thus eliminating the important impact of the entire organ. A holistic human respiratory system model requires a consistent, physiologically realistic representation of the respiratory mechanics in addition to the purely mechanical or structural aspects because the basic physics of respiratory mechanics are governed by the interaction of the fluid with the anatomy. The normal functioning of the lung is possible due to the deformation of the associated muscles, airway tree and lung tissue. Indeed, tissue properties have an influence on airflow and the processes that occur during air transportation: for instance, during gaseous exchange or aerosol transportation.

1.2. Related Works

Tran et al. [9] conducted research work with the main goal of designing, controlling, modelling, and simulating a mechanical ventilator that is lightweight, portable, and suitable for use at home. Tran et al. [9] modelled a mechanical ventilator as a voltage source. El-Hadj et al. [10] applied the fluid–structure interaction (FSI) to couple the computational fluid dynamics used for fluid flow with finite element analysis for the solid domain. This facilitated the investigation of fluid behaviour, structural behaviour, and interactions. Two

designs were proposed for the mechanical ventilator. The first design was reported to be uncontrollable, and the second design considers computational fluid dynamics (CDF) with a moving boundary applied to a piston-cylinder-based ventilator. Tharton et al. [11] designed and developed a ventilator prototype to be used by professionals in medical emergencies and in any other context where the regular ventilator is not available. A mechanical ventilator was modelled using a crank rocker mechanism in order to meet specific requirements for mechanical ventilator design. Pivk et al. [12] developed an empirical model for a low-cost mechanical ventilator by observing the response of each of the ventilator subsystems. The current progress made globally in the development of affordable and effective designs is important. However, research in mechanical ventilator development lacks model development. The Port–Hamiltonian approach applied in this research work acts as a modelling template for future energy-based mechanical ventilator modelling and design. Lagrangian modelling and Kane’s approach do not emphasize energy routing in the system through a Dirac structure. Kane’s method is also limited with regard to modelling of the fluid interaction of the mechanical ventilator–human respiratory system. It is important to develop designs that rest on a solid understanding of a significant aspect of the design. Naggar [13] developed a piecewise model of a mechanical ventilator that describes the artificial behaviour of a mechanical ventilator. A pressure-controlled ventilator was created and simulated. The mechanical ventilator was modelled using a periodic function with inequalities to control the beginning of the inspiration and expiration periods. Shi et al. [14] developed a mathematical model of volume-controlled mechanical ventilation. The model is viewed as a pneumatic system wherein the ventilator is regarded as an air compressor. The exhalation valve is considered a throttle. The compressor and the container represent the ventilator. Naggar et al. [15] proposed an integrated mathematical model of the mechanical ventilator and the lung. Linear quadratic and exponential equations were used to model the system. The integrated model was used to simulate volume-controlled artificial ventilation. Giri et al. [16] proposed a simplified design of a mechanical ventilator to reduce cost and automate the mechanical ventilation process. The proposed design was simulated on the MATLAB platform. Hannon et al. [17] reviewed the progress in mathematical modelling and simulation of human anatomy, physiology, and pathophysiology using computers in the context of mechanical ventilation. This study emphasized the clinical applications of computer simulations in various disease states. The research work further discussed the present constraints and the possibility of *in silico* modelling. There are currently no existing Port–Hamiltonian models of mechanical ventilators.

1.3. Applications of Integrated Mechanical Ventilator–Human Lung System Models

Integrated mechanical ventilator–human lung system models can be used to simulate the behaviour of the respiratory system and to study the effects of mechanical ventilation on lung function. These models have several applications in research, education, and clinical practice. Overall, integrated mechanical ventilator–human lung system models provide a valuable tool for understanding the behaviour of the respiratory system under mechanical ventilation and for developing improved strategies for the diagnosis and treatment of respiratory disorders. The subsections to follow provide examples of applications of integrated models of the mechanical ventilator–respiratory system. In Mehedi et al. [18], the respiratory system was modelled as a combination of the blower-hose patient system with nonlinear lung compliance. There are currently very few studies that have conducted simulations during expiration since the availability of storage for the inhaled volume is required during expiration.

1.4. Motivation of the Port–Hamiltonian Approach

Both the mechanical ventilator and the human respiratory system are multi-physical in nature (they contain components from various domains in physics). As such, a more apt approach to understand the manner in which these systems work, the coupling between

them, etc., and, furthermore, to improve the efficiency of mechanical ventilation, is to make use of models that better represent these systems. Thus, a more representative modelling approach would be to account for the multi-physical nature of these systems [19,20]. The Port–Hamiltonian approach for modelling integrated systems has several benefits, including [19,20]:

1. The Port–Hamiltonian approach is modular and is based on the topology of the mechanical ventilator and the human respiratory system. Thus, various types of mechanical ventilators and human respiratory systems with different structure/topologies can be simulated/tested numerically under various operating conditions to analyse performance, etc. [19,20].
2. The modular nature of the Port–Hamiltonian method, together with its basis in graph theory, means that one can proceed to develop software to automate the modelling of various types of mechanical ventilator/human respiratory system topologies [19,20].
3. Properties such as the passivity of the integrated mechanical ventilator–human respiratory system can be shown and studied [19,20].

Currently, there are no energy-based models of MV–patient systems that can be used in systems that have dissipation of energy from the ports. Complex dynamical systems can be modelled and controlled in the Port–Hamiltonian framework because the framework provides a powerful tool for the analysis, modelling and control of such systems. The key attributes of the Port–Hamiltonian (pH) framework are: Firstly, it is established on the multi-domain bond-graph modelling method, which connects various physical elements via ports. The interconnection in the Port–Hamiltonian method is such that there is power conservation. Secondly, there is significance assigned to the energy variables as well as the total energy stored, which is also known as the Hamiltonian and is the foundation for multi-physics system modelling. The Port–Hamiltonian approach provides the basis for understanding energy transfer in the various physical domains that govern the respiratory system. Sampled data can be obtained when the Port–Hamiltonian structure is exploited and preserved in discretization. This research work applies the Port–Hamiltonian system approach to model the MV–human respiratory system. The resulting model explicitly shows the energy storage, energy dissipation and the interconnection underlying the system. The obtained Port–Hamiltonian model has numerous applications: for instance, it may be applied in the design of various medical devices and for respiratory disease diagnostics.

The main contribution in this research work is the development of an integrated Port–Hamiltonian model representation of a mechanical ventilator–human respiratory system. The model consists of electromechanical and electromagnetic components modelled in a finite-dimensional representation and interconnected with fluid components in an infinite-dimensional representation. As a result, this Port–Hamiltonian model is of a mixed finite/infinite dimensional nature. The remainder of this paper is structured as follows: The materials and methods are outlined in Section 2, Sections 3 and 4 present the detailed Port–Hamiltonian model of a mechanical ventilator and the solenoid valve subsystem. Sections 5 and 6 give the model network topology and model interconnection with coupling conditions. Sections 7 and 8 show the structure-preserving discretization, results and discussion. Finally, the conclusion and recommendations are presented in Section 9.

2. Materials and Method

This section presents the mathematical preliminaries of the Port–Hamiltonian approach that are followed in later sections. These definitions are standard definitions from the literature and are mostly derived from the following references: [21–23].

Definition 1. (Dirac structure):

A Dirac structure (DS) is a pair of elements, $f \in \mathbb{R}^n$ and $e \in \mathbb{R}^n$, that satisfies the set:

$$\mathcal{D} := \left\{ (f, e) \in \mathcal{D}, (\hat{f}, \hat{e}) \in \mathcal{D} \mid e^T \hat{f} = -\hat{e}^T f \right\}.$$

The DS is a subset $\mathcal{D} \subset \mathcal{F} \times \mathcal{E}$, where $\mathcal{F}$ and $\mathcal{E}$ represent the flows and efforts, respectively.

The DS in electric circuits can be represented as:

Proposition 1. The subspace

$$\mathcal{D} \subset \mathbb{R}^n \times \mathbb{R}^n,$$

is called a DS if and only if there exist $A, B \in \mathbb{R}^{n \times n}$ such that

$$AB^T + BA^T = 0 \text{ and } \text{rank}([A \ B]) = n,$$

satisfying the condition

$$\mathcal{D} = \{ (f, e) \in \mathbb{R}^n \times \mathbb{R}^n \mid Af = -Be \}.$$

Definition 2. (Resistive Relation).

Any relation $\mathcal{R} \subset \mathbb{R}^n \times \mathbb{R}^n$ is said to be resistive if $\forall (f_{\mathcal{R}}, e_{\mathcal{R}}) \in \mathcal{R}$,

$$e_{\mathcal{R}}^T f_{\mathcal{R}} \leq 0,$$

is satisfied.

Definition 3. (Port–Hamiltonian (pH) system):

The set $(\mathcal{D}, \mathcal{L}, \mathcal{R})$ defines a pH system where:

- $\mathcal{D} \subset (\mathcal{F}_{\mathcal{L}} \times \mathcal{F}_{\mathcal{R}} \times \mathcal{F}_{\mathcal{P}}) \times (\mathcal{E}_{\mathcal{L}} \times \mathcal{E}_{\mathcal{R}} \times \mathcal{E}_{\mathcal{P}})$ is a DS;
- $\mathcal{L} \subset \mathcal{F}_{\mathcal{L}} \times \mathcal{E}_{\mathcal{L}}$ is an LS;
- $\mathcal{R} \subset \mathcal{F}_{\mathcal{R}} \times \mathcal{E}_{\mathcal{R}}$ a resistive relation.

The elements of the sets:

- $\mathcal{F}_{\mathcal{L}} \in \mathbb{R}^{n_{\mathcal{L}}}$ and $\mathcal{E}_{\mathcal{L}} \in \mathbb{R}^{n_{\mathcal{L}}}$ are known as flows and efforts;
- $\mathcal{F}_{\mathcal{R}} \in \mathbb{R}^{n_{\mathcal{R}}}$ and $\mathcal{E}_{\mathcal{R}} \in \mathbb{R}^{n_{\mathcal{R}}}$ are known as resistive flows and efforts;
- $\mathcal{F}_{\mathcal{P}} \in \mathbb{R}^{n_{\mathcal{P}}}$ and $\mathcal{E}_{\mathcal{P}} \in \mathbb{R}^{n_{\mathcal{P}}}$ are known as external flows and efforts, respectively.

where $n_{\mathcal{L}}, n_{\mathcal{R}}$ and $n_{\mathcal{P}} \in \mathbb{N}_0$.

The dynamics of the pH system are given by the differential inclusion:

$$\left(-\frac{d}{dt}x(t), f_{\mathcal{R}}(t), f_{\mathcal{P}}(t), e_{\mathcal{L}}(t), e_{\mathcal{R}}(t), e_{\mathcal{P}}(t) \right) \in \mathcal{D},$$

where $(x(t), e_{\mathcal{L}}(t)) \in \mathcal{L}$, $(f_{\mathcal{R}}(t), e_{\mathcal{R}}(t)) \in \mathcal{R}$, and $(f_{\mathcal{P}}(t), e_{\mathcal{P}}(t)) \in \mathcal{P}$.

Definition 4. (Interconnection of n pH systems):

Let $(\mathcal{D}_i, \mathcal{L}_i, \mathcal{R}_i)$ denote the space of the i th pH system in a set of n pH systems that are to be interconnected. The space of flows is divided into an external part and a part to be linked and is given by

$$\mathcal{F}_i = \mathcal{F}_{\mathcal{L}_i} \times \mathcal{F}_{\mathcal{R}_i} \times \mathcal{F}_{\mathcal{P}_i} \times \mathcal{F}_{\mathcal{P}link}.$$

Similarly, the space of efforts is divided into an external part and a part to be linked and is expressed as:

$$\mathcal{E}_i = \mathcal{F}_{\mathcal{L}_i} \times \mathcal{E}_{\mathcal{R}_i} \times \mathcal{E}_{\mathcal{P}_i} \times \mathcal{E}_{\mathcal{P}link}.$$

The interconnection of two pH systems $(\mathcal{D}_1, \mathcal{L}_1, \mathcal{R}_1)$ and $(\mathcal{D}_2, \mathcal{L}_2, \mathcal{R}_2)$ with respect to the link $(F_{\mathcal{P}link}, E_{\mathcal{P}link})$ results in a new interconnected pH system: $(\mathcal{D}, \mathcal{L}, \mathcal{R})$. This interconnected system is given by the expression:

$$(\mathcal{D}_1, \mathcal{L}_1, \mathcal{R}_1) \circ (\mathcal{D}_2, \mathcal{L}_2, \mathcal{R}_2) := (\mathcal{D}, \mathcal{L}, \mathcal{R}).$$

Definition 5. (Directed graph):

A directed graph is a quadruple $\mathcal{G} := (V, E, l, r)$, where

1. V is a set of vertices;
2. E is a set of edges;
3. $l : E \rightarrow V$ maps each edge, e , to an initial vertex;
4. $r : E \rightarrow V$ maps each edge, e , to a terminal vertex.

Definition 6. (Loop-free directed graph):

If G is a directed graph, then G is said to be loop-free if for all $e \in E$,

$$l(e) \neq r(e).$$

Definition 7. (Subgraphs).

Given the graphs $\mathcal{G} := (V, E, l, r)$ and $\mathcal{G}' := (V', E', l, r)$, then $\mathcal{G}'$ is said to be a subgraph of $\mathcal{G}$ if $E' \subset E$ and $V' \subset V$. Furthermore,

1. A subgraph is said to be an induced subgraph on V' if $E' = E|_{V'}$;
2. A subgraph is said to be spanning if $V' = V$;
3. A subgraph is said to be a proper subgraph if $E' \neq E$;
4. If both V and E are finite, then $\mathcal{G}$ is said to be finite.

Definition 8. (Paths, connectivity and cycles):

Let $\mathcal{G} = (V, E, l, r)$ be a directed finite graph.

1. An n -tuple $e = (e_1, \dots, e_n) \in (E \cup -E)^n$ is called a path from v to ω if
 - (a) $l(e_1), \dots, l(e_n)$ are distinct;
 - (b) $r(e_i) = l(e_{i+1})$ for all $i \in \{1, \dots, n-1\}$;
 - (c) $l(e_1) = v \wedge r(e_n) = \omega$.
2. A path from v to v is called a cycle.
3. Two vertices, v and ω are said to be connected if there exists a path from v to ω .
4. The existence of paths from vertices gives an equivalence relation on the set of vertices.
5. A subgraph is a component of the graph.
6. A graph with only one component is said to be connected.

Definition 9. (Incidence matrix):

Let $\mathcal{G} = (V, E, l, r)$ be a directed graph that is finite and loop-free such that $E = e_1, \dots, e_m$ and $V = v_1, \dots, v_n$. Then the j th row and k th column of the incidence matrix $A_0 \in \mathbb{R}^{n \times m}$ of $\mathcal{G}$ is given by

$$a_{jk} = \begin{cases} 1 & l(e_k) = v_j, \\ -1 & r(e_k) = v_j, \\ 0 & \text{otherwise.} \end{cases}$$

If $\text{rank}(A_0) = n - k$, then graph $\mathcal{G}$ has $k \in \mathbb{N}$ components, such that k rows can be removed from A_0 , resulting in a matrix with the same rank.

Definition 10. (Kirchhoff–Dirac structure, Kirchhoff–Lagrange submanifold):
The Kirchhoff–Dirac structure of $\mathcal{G}$ can thus be defined by the set

$$\mathcal{D}_K^S(\mathcal{G}) := \left\{ (j, i, \phi, u) \in \mathbb{R}^{n-|S|} \times \mathbb{R}^m \times \mathbb{R}^{n-|S|} \times \mathbb{R}^m \left| \begin{bmatrix} I & A \\ 0 & 0 \end{bmatrix} \begin{pmatrix} j \\ i \end{pmatrix} + \begin{bmatrix} I & A \\ 0 & 0 \end{bmatrix} \begin{pmatrix} \phi \\ u \end{pmatrix} = 0 \right. \right\}, \quad (1)$$

where i and u are the currents and voltages, respectively, at the edges of the graph, whereas q and ϕ are the charges and potentials, respectively, at the vertices of the graph, respectively.

Assuming that $S = \{v_1, \dots, v_{|S|}\}$, the Kirchhoff–Lagrange submanifold of $\mathcal{G}$ with respect to S is defined as

$$\mathcal{L}_K^S(\mathcal{G}) := 0 \times \mathbb{R}^{n-|S|} \subset \mathbb{R}^{n-|S|} \times \mathbb{R}^{n-|S|}, \quad (2)$$

where $\mathcal{G} = (V, E, l, r)$ is a directed graph that

1. Is finite;
2. Is loop-free;
3. Has an incidence matrix $A_0 \in \mathbb{R}^{n \times m}$.

If $\mathcal{G}_1, \dots, \mathcal{G}_k$ are the components of $\mathcal{G}$ with corresponding vertices $V_1, \dots, V_k \subset V$ so that there exists a subset $S \subset V$ such that S contains at most one vertex from each component, i.e., $\forall s, s', i \leq k : v, v' \in V_i \Rightarrow v = v'$, then $A \in \mathbb{R}^{(n-k) \times m}$ can be constructed by deleting the rows corresponding to the vertices S from $A_0 \in \mathbb{R}^{n \times m}$.

Remark 1. According to Proposition 1, Equations (1) and (2) indicate that $\mathcal{D}_K^S(\mathcal{G})$ is a DS and $\mathcal{L}_K^S(\mathcal{G})$ is an LS in the space $\mathbb{R}^{n-|S|} \times \mathbb{R}^{n-|S|}$.

Definition 10 caters to an introduction of a pH system $(\mathcal{D}_K^S(\mathcal{G}), \mathcal{L}_K^S(\mathcal{G}), \{0\})$ with dynamics:

$$\left(-\frac{d}{dt} q(t), i(t), \phi(t), u(t) \right) \in \mathcal{D}_K^S(\mathcal{G}), \quad (3)$$

where $(q(t), \phi(t)) \in \mathcal{L}_K^S(\mathcal{G})$.

Using the equivalence of $(q(t), \phi(t)) \in \mathcal{L}_K^S(\mathcal{G})$ to $q(t) = 0$ and $\phi(t) \in \mathbb{R}^{n-|S|}$, we see that Equation (3) holds if and only if this condition holds:

$$q(t) = 0 \wedge A i(t) = 0 \wedge A^T \phi(t) = -u(t).$$

$$a_{jl} = \begin{cases} 1 & e_l \in C_j \text{ with orientations that coincide,} \\ -1 & e_l \in C_j \text{ with orientations that do not coincide,} \\ 0 & \text{otherwise.} \end{cases}$$

On this basis, one can define a DS

$$\mathcal{D}'_K(\mathcal{G}) := \left\{ (j, i, \phi, u) \in \mathbb{R}^{n-|S|} \times \mathbb{R}^m \times \mathbb{R}^{n-|S|} \times \mathbb{R}^m \left| \begin{bmatrix} I & A \\ 0 & 0 \end{bmatrix} \begin{pmatrix} j \\ i \end{pmatrix} + \begin{bmatrix} I & A \\ 0 & 0 \end{bmatrix} \begin{pmatrix} \phi \\ u \end{pmatrix} = 0 \right. \right\},$$

with dynamics

$$\left(-\frac{d}{dt} \psi(t), i(t), \iota(t), u(t) \right) \in \mathcal{D}'_K(\mathcal{G}),$$

and an LS

$$\mathcal{L}'_K(\mathcal{G}) := \{0\} \times \mathbb{R}^{n-m+k},$$

with dynamics

$$(\psi(t), \iota(t)) \in \mathcal{L}'_K(\mathcal{G}).$$

Together, they form a pH system defined by the set $(D'_K(\mathcal{G}), \mathcal{L}'_K(\mathcal{G}), \{0\})$.

Dirac Structure

A key feature of a Dirac structure is the fact that the standard composition of two Dirac structures is again a Dirac structure. The implication of this statement is that any power-conserving interconnection of a Port–Hamiltonian system is also a Port–Hamiltonian system itself. This constitutes the foundational feature of the Port–Hamiltonian approach to modelling, simulation and control of complex physical systems. The intricate Dirac structure is the guide to the algebraic constraints of the interconnected system as well as its Casimir functions [23]. The Casimir functions are significant for the set-point regulation of Port–Hamiltonian systems. The framework of the Port–Hamiltonian approach allows for port-based modelling. Port-based modelling means that we are interconnecting many different elements through ports. Dirac structures are the tools used to connect multiple elements. These various elements are energy-storing elements, energy-dissipating elements and external elements that can supply energy. A diagram to demonstrate the connection structure is given in Figure 1 [23].

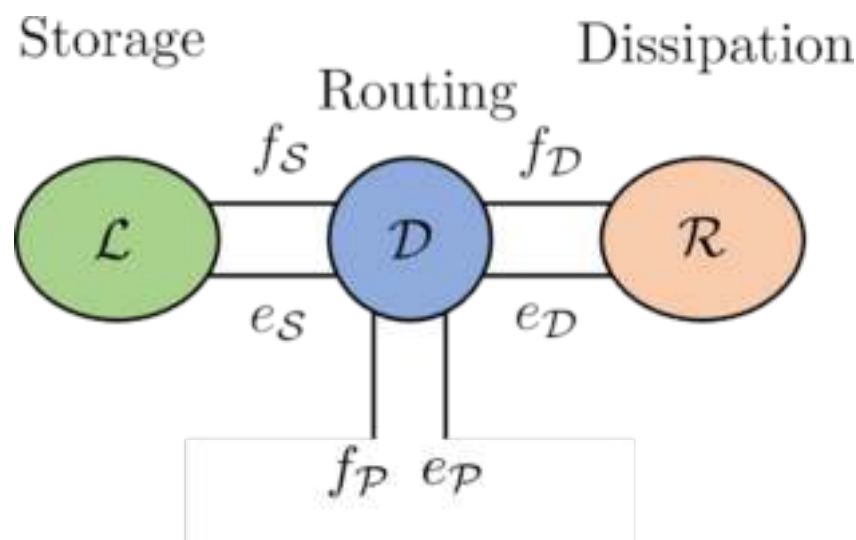


Figure 1. Energy storage, routing and dissipation.

3. Detailed Port–Hamiltonian Model of a Mechanical Ventilator

3.1. Description of the System

The overall mechanical ventilator system diagram is provided in Figure 2. The entire system consists of various subsystems. In this research work, the main subsystem that contributes to the flow of air and, consequently, the air pressure are discussed. Therefore, the following main subsystems will be discussed: the DC motor subsystem, the turbine pump/blower subsystem, the pump-shaft/impeller subsystem and the solenoid valve subsystem.

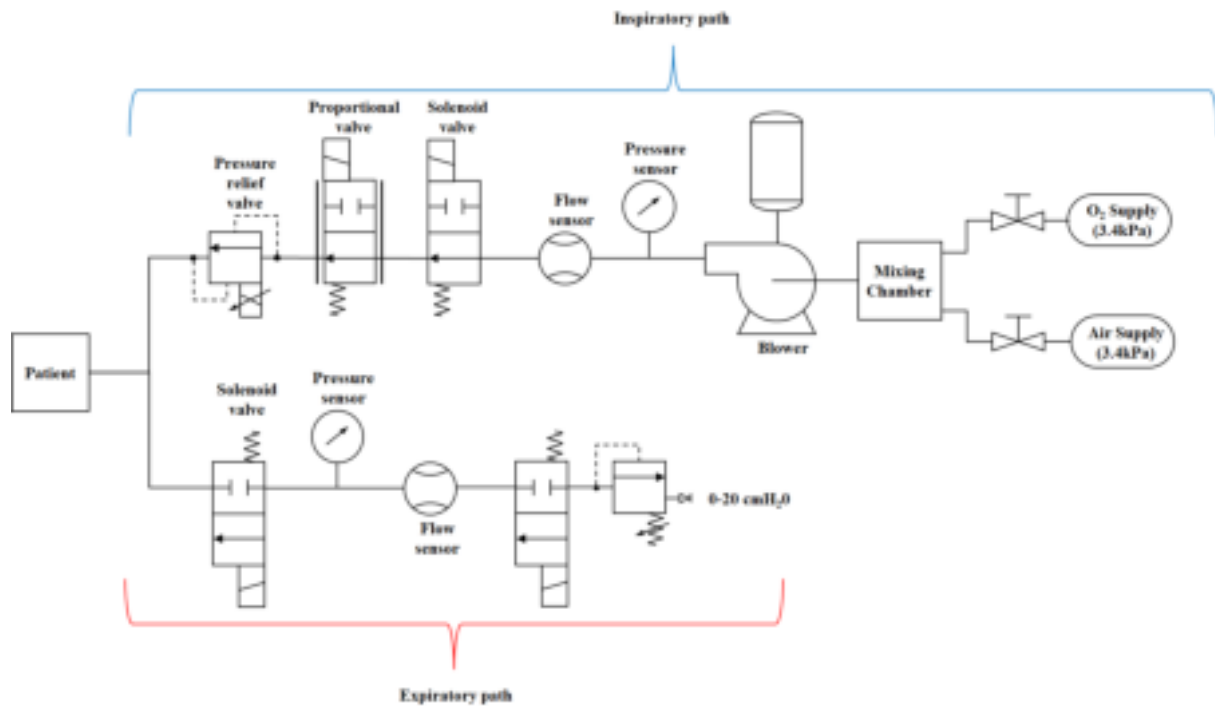


Figure 2. Schematic diagram of a mechanical ventilator

3.2. Blower Model

In this section, the dynamical system model equations for the blower model are presented in the Port–Hamiltonian framework. The blower model comprises three sub-systems: namely, a DC motor that drives the blower-shaft/impeller, a blower-shaft/impeller that couples the DC motor to the fluid by using the rotational motion of the motor to accelerate the fluid, and finally, the fluid being driven.

From the Port–Hamiltonian perspective, the state vector $\mathcal{X}_b \in \mathbb{R}^4$ is given by:

$$\mathcal{X}_b := [p_m, \phi_m, P_b, Q_b]^\top, \quad (4)$$

where $p_m \in \mathbb{R}^1$ and $\phi_m \in \mathbb{R}^1$ are the angular momentum and magnetic flux of the DC motor, respectively, while $P_b \in \mathbb{R}^1$ and $Q_b \in \mathbb{R}^1$ are the pressure and flow rate of the blower, respectively.

The total energy of the blower is given by the Hamiltonian:

$$\mathcal{H}_b[\mathcal{X}_b] = \frac{1}{2} \left(\frac{p_m^2}{I_m} + \frac{\phi_m^2}{L_m} + C_b P_b^2 + I_b Q_b^2 \right), \quad (5)$$

where $I_m \in \mathbb{R}_+^1$ is the inertia and $L_m \in \mathbb{R}_+^1$ is the inductance of the DC motor, while $C_b \in \mathbb{R}_+^1$ and $I_b \in \mathbb{R}_+^1$ are the hydraulic capacitance and inertance of the air in the blower, respectively.

Thus, the Port–Hamiltonian model of a blower is given by

$$\dot{\mathcal{X}}_b = (\mathcal{J}_b - \mathcal{R}_b) \nabla \mathcal{H}_b[\mathcal{X}_b] + \mathcal{G}_b u_b, \quad (6)$$

$$\mathcal{Y}_b = \mathcal{G}_b^* \nabla \mathcal{H}_b[\mathcal{X}_b], \quad (7)$$

where

$$\begin{aligned}\mathcal{J}_b &:= \begin{bmatrix} 0 & K_m & 0 & -(K_o p_m)/I_m \\ -K_m & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/(C_b I_b) \\ (K_o p_m)/I_m & 0 & -1/(C_b I_b) & 0 \end{bmatrix} \in \mathbb{R}^{4 \times 4}, \\ \mathcal{R}_b &:= \text{diag}([b_m \quad R_m \quad 0 \quad (R_b p_m)/(I_b I_m)]) \in \mathbb{R}^{4 \times 4}, \\ \nabla \mathcal{H}_b[\mathcal{X}_b] &:= [p_m/I_m \quad \phi_m/L_m \quad C_b P_b \quad I_b Q_b] \in \mathbb{R}^{4 \times 1}, \\ \mathcal{G}_b &:= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & -(1/C_b) & 0 \\ 0 & 0 & (1/I_b) \end{bmatrix} \in \mathbb{R}^{4 \times 3}, \\ u_b &:= [V_m \quad Q_{in} \quad P_{in}] \in \mathbb{R}^{3 \times 1},\end{aligned}$$

where $K_m \in \mathbb{R}_+^1$ is the motor torque constant, $b_m \in \mathbb{R}_+^1$ is the viscous damping, $R_m \in \mathbb{R}_+^1$ is the armature resistance, and $K_o \in \mathbb{R}_+^1$ is the motor's angular momentum/pressure coupling constant. The inputs to the system are the DC motor voltage, the input volumetric flow rate and the pressure, which are given by $V \in \mathbb{R}^1$, $Q_{in} \in \mathbb{R}^1$ and $P_{in} \in \mathbb{R}^1$, respectively. One can easily show that $\mathcal{J}_m = -\mathcal{J}_m^\top$ and $\mathcal{R}_m = \mathcal{R}_m^\top \geq 0$.

Taking the time derivative of the Hamiltonian gives

$$\dot{\mathcal{H}}_b[\mathcal{X}_b] = u_b \cdot \mathcal{Y}_b - \nabla \mathcal{H}_b[\mathcal{X}_b] \cdot (\mathcal{R}_b \nabla \mathcal{H}_b[\mathcal{X}_b]). \quad (8)$$

This system has power and resistive ports.

4. Solenoid Valve Subsystem

Figure 3 shows a solenoid valve in the open and closed positions. It is assumed that the air gaps are sufficiently small such that the effect of fringing of the magnetic flux is negligible. Consider a solenoid in which the permeability of the core and the length of the part of the magnetic circuit inside the core are denoted by $\mu_c \in \mathbb{R}_+^1$ and $l_c \in \mathbb{R}_+^1$, respectively. The equivalent length of the solenoid's magnetic circuit $l_{eq}[\cdot] : \mathbb{R} \rightarrow \mathbb{R}$ is dependent on the displacement of the spool $q_s \in \mathbb{R}$ and can be written as:

$$l_{eq}[q_s] = l_c + \frac{\mu_c}{\mu_0}(q_{stot} - q_s), \quad (9)$$

where $\mu_0 \in \mathbb{R}_+^1$ is the permeability of air, and $q_{stot} \in \mathbb{R}_+^1$ is the total air gap. The solenoid coordinate system is represented in Figure 4. Thus, the inductance of the solenoid varies with displacement of the spool and, hence, can be expressed by the function $L_s[\cdot] : \mathbb{R} \rightarrow \mathbb{R}$ given by:

$$L_s[q_s] = \frac{N^2 A_e \mu_c}{l_{eq}[q_s]}, \quad (10)$$

where $N \in \mathbb{R}_+^1$ is the number of turns in the coil of the solenoid, and $A_e \in \mathbb{R}_+^1$ is the effective cross-sectional area of the path of the magnetic flux. The magnetic and mechanical subsystems in the solenoid valve are therefore coupled magnetically due to the dependence of the inductance on the displacement of the spool.

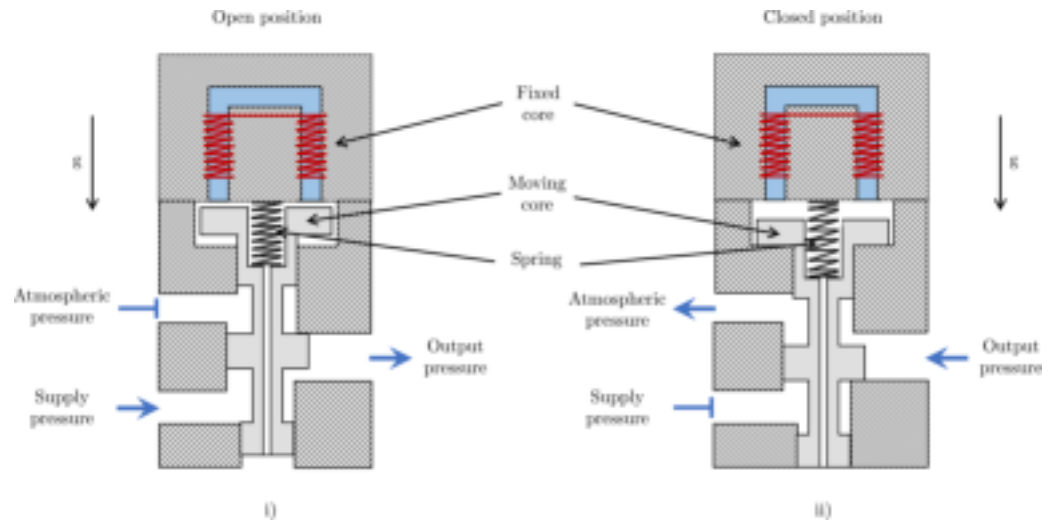


Figure 3. Diagram of a solenoid valve in (i) the open position, which allows fluid flow, and (ii) the closed position, which stops fluid flow.

The total energy of the solenoid is given by the Hamiltonian $\mathcal{H}_s[\cdot] : \mathbb{R}^3 \rightarrow \mathbb{R}$, which is a function of the state vector $\mathcal{X}_s = [\phi_s, p_s, q_s]^T \in \mathbb{R}^3$ and is expressed as the sum of the magnetic, kinetic and potential energies, which are denoted by $\mathcal{H}_{\text{magnetic}}[\phi_s, q_s] : \mathbb{R}^2 \rightarrow \mathbb{R}$, $\mathcal{H}_{\text{kinetic}}[p_s] : \mathbb{R} \rightarrow \mathbb{R}$ and $\mathcal{H}_{\text{potential}}[q_s] : \mathbb{R} \rightarrow \mathbb{R}$, respectively. Thus,

$$\mathcal{H}_s[\mathcal{X}_s] = \mathcal{H}_{\text{magnetic}}[\phi_s, q_s] + \mathcal{H}_{\text{kinetic}}[p_s] + \mathcal{H}_{\text{potential}}[q_s], \quad (11)$$

given a magnetic flux $\phi_s \in \mathbb{R}$.

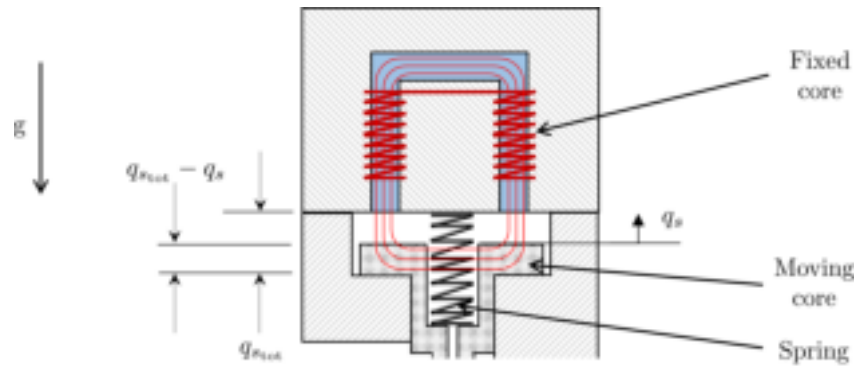


Figure 4. Solenoid coordinate system.

Assuming that the pretension of the spring is set to $q_0 \in \mathbb{R}$, the Hamiltonian is

$$\mathcal{H}_s[\mathcal{X}_s] = \frac{1}{2} \left(\frac{\phi_s^2}{L[q_s]} + \frac{p_s^2}{m_s} + k_s(q_s + q_{s0})^2 \right) + m_s q_s g, \quad (12)$$

where $p_s \in \mathbb{R}$ is the momentum of the spool, $m_s \in \mathbb{R}_+^1$ is the mass of the spool, $k_s \in \mathbb{R}_+^1$ is the spring stiffness, and $g \in \mathbb{R}_+^1$ is the acceleration due to gravity.

Hence, the solenoid's state and output dynamics are expressed in Port–Hamiltonian form in Equations (13) and (14), respectively:

$$\dot{\mathcal{X}}_s = (\mathcal{J}_s - \mathcal{R}_s) \nabla \mathcal{H}_s[\mathcal{X}_s] + \mathcal{G}_s u_s, \quad (13)$$

$$\mathcal{Y}_s = \mathcal{G}_s^* u_s, \quad (14)$$

where

$$\begin{aligned} \mathcal{J}_s &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}, \quad \mathcal{R}_s = \begin{bmatrix} R_s & 0 & 0 \\ 0 & b_s & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathcal{G}_s = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & (A_{s1} - A_{s2}) & (A_{s3} - A_{s4}) \end{bmatrix}, \\ \mathcal{G}_s^* &= \mathcal{G}_s^\top, \quad u_s = \begin{bmatrix} V_s \\ p_s \end{bmatrix} \quad \text{and} \quad \nabla \mathcal{H}_s[\mathcal{X}_s] = \begin{bmatrix} \frac{\phi_s}{A_e N^2 \mu_c} \left(l_c - \frac{\mu_c (q_s - q_{s\text{tot}})}{\mu_0} \right) \\ \frac{p_s}{m_s} \\ \frac{A_e N^2 \mu_c^2 \mu_0}{(\mu_0 l_c - \mu_c (q_s - q_{s\text{tot}}))^2} + k_s (q_s + q_{s0}) + m_s g \end{bmatrix}, \end{aligned}$$

where $R_s \in \mathbb{R}_+^1$ is the resistance of the coil, $b_s \in \mathbb{R}_+^1$ is the viscous damping acting on the spool, and ∇ is the gradient operator. $A_{s1}, A_{s2} \in \mathbb{R}_+^1$ are the various cross-sectional areas of the surfaces of the spool, and the input vector $u_s \in \mathbb{R}^2$ consists of the input voltage $V_s \in \mathbb{R}^1$ and the supply pressure $p_s \in \mathbb{R}^1$. It can be seen that $\mathcal{J}_s = -\mathcal{J}_s^\top \in \mathbb{R}^{3 \times 3}$ possesses skew symmetry, while $\mathcal{R}_s = \mathcal{R}_s^\top \in \mathbb{R}^{3 \times 3}$ is positive semi-definite.

Taking the time derivative of the Hamiltonian gives:

$$\dot{\mathcal{H}}_s[\mathcal{X}_s] = u_s \cdot \mathcal{Y}_s - \nabla \mathcal{H}_s[\mathcal{X}_s] \cdot (\mathcal{R}_s \nabla \mathcal{H}_s[\mathcal{X}_s]). \quad (15)$$

This system has power and resistive ports.

4.1. Pipe Model

In this section, a Port–Hamiltonian model of a single pipe segment is developed. The basis of these developments is the Navier–Stokes equations for one-dimensional non-stationary flow of gas in a pipe. The following assumptions are taken for the sake of model simplification [24,25]:

1. The pipe is taken as rigid (the cross-section does not expand as a result of fluid flow).
2. Frictional and gravitational effects are neglected (this will be relaxed in future works in this research area).
3. The model parameters of the gas remain constant along the pipe cross-section but vary in time along the pipe length. Thus, they can be averaged about the cross-section, and thus, the gas flow is one-dimensional.
4. The temperatures of the pipe walls are assumed to be constant and equal to the ambient room temperature. Hence, temperature effects are ignored.

Taking into account these assumptions, the coordinate system attached to a segment of pipe is illustrated in Figure 5. Thus, for a given time interval $t_s \leq t < t_f$ with start time t_s and finish time t_f , the length-normalized one-dimensional Euler equations for gas with a density $\rho(t, z) : (t_s, t_f] \times (0, \ell) \rightarrow \mathbb{R}^1$ flowing at a velocity $v(t, z) : (t_s, t_f] \times (0, \ell) \rightarrow \mathbb{R}^1$ through a pipe of length $\ell \in \mathbb{R}_+^1$ and cross-sectional area $A_p \in \mathbb{R}_+^1$ are given by:

$$\partial_t(\rho A_p) + \partial_z(\rho A_p v) = 0, \quad (16)$$

$$\partial_t(\rho A_p v) + \partial_z(\rho A_p v^2 + p A_p) = 0, \quad (17)$$

where $p(t, z) : (t_s, t_f] \times (0, \ell) \rightarrow \mathbb{R}^1$ is the pressure, and $\partial_i := \partial/\partial_i$ is the partial derivative with respect to the temporal and spatial variables given by the subscripts $i \in \{t, z\}$.

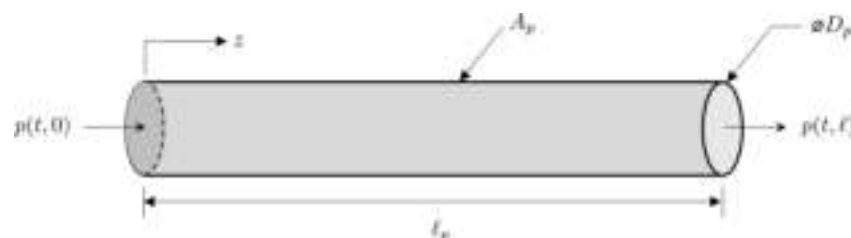


Figure 5. Pipe segment coordinate system.

Port–Hamiltonian Formulation of Pipe-Flow Model

The fluid dynamics can be written in terms of mass per unit length, i.e., $q := \rho A_p$, as well as the fluid momentum, $m := qv$. Thus, Equations (16) and (17) can be expressed as

$$\partial_t q = -\partial_z m, \quad (18)$$

$$\partial_t m = -\partial_z \left(\frac{m^2}{q} + p A_p \right), \quad (19)$$

Defining the state vector of the gas flow through a pipe segment as $\chi_p := \begin{bmatrix} q \\ m \end{bmatrix}$, the energy of the gas can be expressed in the form of a Hamiltonian $\mathcal{H}_p[\chi_p] : \mathbb{R}^2 \rightarrow \mathbb{R}^1$ given by

$$\mathcal{H}_p[\chi_p] = \int_0^\ell H_p[\chi_p] \, dx = \int_0^\ell \left(\frac{m^2}{2q} + qU[q/A_p] \right) dx, \quad (20)$$

where $H_p[\cdot] : \mathbb{R}^2 \rightarrow \mathbb{R}^1$ is the Hamiltonian density, and $U[\cdot] : \mathbb{R}^1 \rightarrow \mathbb{R}^1$ is the internal energy of the gas, which, in the case of an isentropic fluid, can be expressed as a function of density.

The Port–Hamiltonian dynamics take the following form:

$$\partial_t \chi_p = \mathcal{J}[\chi_p] \delta_{\chi_p} \mathcal{H}_p[\chi_p], \quad (21)$$

where $\mathcal{J}[\chi_p]$, which is the formally skew-symmetric operator, and $\delta_{\chi_p} \mathcal{H}[\chi_p]$, which is the variational derivative of the Hamiltonian density, are expressed as

$$\mathcal{J}[\chi_p] = -\begin{bmatrix} 0 & \partial_z \\ \partial_z & 0 \end{bmatrix} \quad \text{and} \quad \delta_{\chi_p} \mathcal{H}_p[\chi_p] = \begin{bmatrix} h - \frac{m^2}{2q^2} \\ \frac{m}{q} \end{bmatrix}, \quad (22)$$

where h is the enthalpy, and $\mathcal{J}[\chi_p]$ is a formally skew-symmetric operator.

The rate of change of the Hamiltonian can be found as

$$\dot{\mathcal{H}}_p = \mathbf{u}_p^\top \mathbf{y}_p, \quad (23)$$

where

$$\mathbf{u}_p = W_B R_{\text{ext}} \begin{bmatrix} \delta_{\chi_p} \mathcal{H}_p|_0 \\ \delta_{\chi_p} \mathcal{H}_p|_\ell \end{bmatrix} \quad \text{and} \quad \mathbf{y}_p = W_C R_{\text{ext}} \begin{bmatrix} \delta_{\chi_p} \mathcal{H}_p|_0 \\ \delta_{\chi_p} \mathcal{H}_p|_\ell \end{bmatrix}, \quad (24)$$

with components given by

$$W_B = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 \end{bmatrix}, \quad W_C = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix}, \quad R_{\text{ext}} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & -1 & 0 & 1 \\ -1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix},$$

and

$$\begin{bmatrix} \delta_{\chi_p} \mathcal{H}_p|_0 \\ \delta_{\chi_p} \mathcal{H}_p|_\ell \end{bmatrix} = \begin{bmatrix} \left(h - \frac{m^2}{2q^2} \right) \Big|_0 \\ \left(\frac{m}{q} \right) \Big|_0 \\ \left(h - \frac{m^2}{2q^2} \right) \Big|_\ell \\ \left(\frac{m}{q} \right) \Big|_\ell \end{bmatrix}.$$

Expanding $\mathbf{u}_p$ and $\mathbf{y}_p$ gives

$$\mathbf{u}_p = \begin{bmatrix} \left(\frac{m}{q}\right)\Big|_{\ell} \\ \left(h - \frac{m^2}{2q^2}\right)\Big|_0 \end{bmatrix} \quad \text{and} \quad \mathbf{y}_p = \begin{bmatrix} \left(h - \frac{m^2}{2q^2}\right)\Big|_{\ell} \\ -\left(\frac{m}{q}\right)\Big|_0 \end{bmatrix}. \quad (25)$$

Thus, the rate of change of the Hamiltonian is

$$\dot{\mathcal{H}}_p = \left(h - \frac{m^2}{2q^2}\right)\Big|_{\ell} \left(\frac{m}{q}\right)\Big|_{\ell} - \left(\frac{m}{q}\right)\Big|_0 \left(h - \frac{m^2}{2q^2}\right)\Big|_0. \quad (26)$$

4.2. Electric Circuit Model of the Lung

The Port–Hamiltonian formulation for nonlinear electric circuits is presented in this section. Since the main focus of this article is a mechanical ventilator model, the lung model is simplified by considering an electric circuit analogy. The model under consideration is that of a fully sedated patient who relies completely on the mechanical ventilator to breathe. The circuit model is shown in Figure 6. The circuit can be represented in the form of a network graph using graph theory. The graph of the circuit given in Figure 6 is given in Figure 7. The model has $n = |\mathcal{V}| = 7$ vertices.

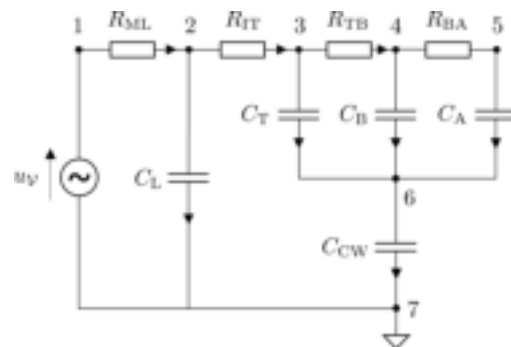


Figure 6. Circuit diagram of an electric model of a lung of a fully sedated patient [26].

The complete graph of the circuit, A_0 , is

$$\mathcal{A}_0 = \begin{bmatrix} R_{ML} & R_{IT} & R_{TB} & R_{BA} & C_T & C_B & C_A & C_L & C_{CW} & u_V \\ \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} & \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & -1 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} & \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} & \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \end{bmatrix}. \quad (27)$$

The ventilator input voltage is represented as a source $\mathcal{S} = (\mathcal{D}_S, \mathcal{L}_S, \mathcal{R}_S)$. The set of sources is $\mathcal{S} = \{v_7\}$, and the dimension of the source is $|\mathcal{S}| = 1$. The circuit also consists of $m_C = 5$ capacitors (storage elements) that are represented as $\mathcal{C}_i = (\mathcal{D}_{C_i}, \mathcal{L}_{C_i}, \mathcal{R}_{C_i})$, where the i th index used to distinguish between the components is $i = \{C_L, C_T, C_B, C_A, C_{CW}\}$. Furthermore there are $m_R = 4$ resistors (dissipative elements) that are represented as $\mathcal{R}_i = (\mathcal{D}_{R_i}, \mathcal{L}_{R_i}, \mathcal{R}_{R_i})$, where the i th index is given by $i = \{R_{ML}, R_{IT}, R_{TB}, R_{BA}\}$.

Selecting node 7 as the ground, the reduced incidence matrix, $\mathcal{A}$, is given by

$$A = [A_{\mathcal{R}} \quad A_C \quad A_S] = \left(\underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}}_{A_{\mathcal{R}}} \mid \underbrace{\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ -1 & -1 & -1 & 0 & 1 \end{pmatrix}}_{A_C} \mid \underbrace{\begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}}_{A_S} \right), \quad (28)$$

where the dimensions of the components of $\mathcal{A}$ are $\dim(A_{\mathcal{R}}) = (n - |S|) \times m_{\mathcal{R}} = 6 \times 4$, $\dim(A_C) = (n - |S|) \times m_C = 6 \times 5$ and $\dim(A_S) = (n - |S|) \times m_S = 6 \times 1$.

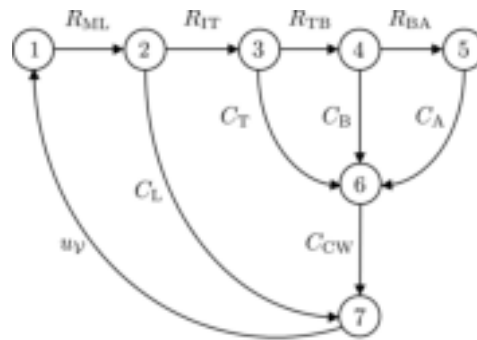


Figure 7. The graph associated with the circuit diagram given in Figure 6.

The total energy of the circuit is given by the Hamiltonian

$$\mathcal{H}_C = \sum_{\forall i} \mathcal{H}_{C_i}(q_{C_i}) = \frac{1}{2} \sum_{\forall i} \frac{q_{C_i}^2}{C_i}, \quad (29)$$

where $q_{C_i} \in \mathbb{R}^1$ is the charge of the i th capacitor for $i = \{C_L, C_T, C_B, C_A, C_{CW}\}$. Thus,

$$\begin{bmatrix} \mathbb{I} & A_C & A_{\mathcal{R}} & A_S \end{bmatrix} \begin{pmatrix} -\frac{d}{dt}q \\ -\frac{d}{dt}q_C \\ -i_{\mathcal{R}} \\ -i_S \end{pmatrix} + \begin{bmatrix} \mathbb{O} & \mathbb{O} & \mathbb{O} & \mathbb{O} \\ -A_{\mathcal{R}}^T & \mathbb{O} & \mathbb{I} & \mathbb{O} \\ -A_C^T & \mathbb{I} & \mathbb{O} & \mathbb{O} \\ -A_S^T & \mathbb{O} & \mathbb{O} & \mathbb{I} \end{bmatrix} \begin{pmatrix} \phi \\ u_C \\ u_{\mathcal{R}} \\ u_S \end{pmatrix} = \mathbb{O}, \quad (30)$$

where $\mathbb{I}$ and $\mathbb{O}$ are appropriately sized identity and zero matrices, respectively, $q = 0$, and ϕ is the node potential. Expanding Equation (30) gives

$$A_C \frac{d}{dt} q_C + A_{\mathcal{R}} (A_{\mathcal{R}}^T \phi) + A_S i_S = 0, \quad (31)$$

$$-A_C^T \phi + u_C = 0, \quad (32)$$

$$-A_S^T \phi + u_S = 0. \quad (33)$$

The Dirac structure is given by the set with the following attributes for $(\mathcal{D}, \mathcal{L}, \mathcal{R})$:

$$\left(-\frac{d}{dt}q, -\frac{d}{dt}q_C, -i_{\mathcal{R}}, -i_S, \phi, u_C, u_{\mathcal{R}}, u_S \right) \in \mathcal{D}, \quad (34)$$

$$(q, q_C, \phi, u_C) \in \mathcal{L}, \quad (35)$$

$$(-i_{\mathcal{R}}, e_{\mathcal{R}}) \in \mathcal{R}. \quad (36)$$

5. Model Network Topology

Definition 11. (Directed graph) [27]: A directed graph $\mathcal{G}$ denotes the pair $(\mathcal{V}, \mathcal{A})$, where $\mathcal{V}(\mathcal{D})$ and $\mathcal{A}(\mathcal{D})$ denote the set of vertices and arcs, respectively. An arc is a distinct ordered-pair of vertices.

The directed graph of a mechanical ventilator model is composed of the disjoint union of the vertices associated with the patient as well as the inspiratory and expiratory limbs of the ventilator, which are given by $\mathcal{V}_O, \mathcal{V}_I$ and $\mathcal{V}_E$, respectively. Similarly, the patient, inspiratory and expiratory arcs are given by $\mathcal{A}_O, \mathcal{A}_I$ and $\mathcal{A}_E$, respectively. Each arc and vertex can further be subdivided into those associated with the blowers, valves, pipes and electric circuits present in the mechanical ventilator. Thus, the totality of vertices and arcs is given by

$$\mathcal{V}(\mathcal{D}) = \{v_O\} \cup \mathcal{V}_I \cup \mathcal{V}_E = \{v_O\} \cup (\mathcal{V}_{I_b} \cup \mathcal{V}_{I_v} \cup \mathcal{V}_{I_p} \cup \mathcal{V}_{I_e}) \cup (\mathcal{V}_{E_b} \cup \mathcal{V}_{E_v} \cup \mathcal{V}_{E_p} \cup \mathcal{V}_{E_e})$$

and

$$\mathcal{A}(\mathcal{D}) = \{a_O\} \cup \mathcal{A}_I \cup \mathcal{A}_E = \{a_O\} \cup (\mathcal{A}_{I_b} \cup \mathcal{A}_{I_v} \cup \mathcal{A}_{I_p} \cup \mathcal{A}_{I_e}) \cup (\mathcal{A}_{E_b} \cup \mathcal{A}_{E_v} \cup \mathcal{A}_{E_p} \cup \mathcal{A}_{E_e}),$$

respectively, where the subscripts $i \in \{O, p, b, v, e\}$ are used to indicate the patient, blowers, valves, pipes and electric circuit components, respectively.

A simplified graph of the mechanical ventilator shown in Figure 2 is given in Figure 8. This simplified graph indicates the elements belonging to the patient, as well as the inspiratory arcs and vertices belonging to the mechanical ventilator. The arcs belonging to the patient are denoted by $\mathcal{A}_{O_{21}} \in \mathcal{A}_O$, while the vertices are denoted by $\{\mathcal{V}_{O_1}, \mathcal{V}_{O_2}\} \in \mathcal{V}_O$. The elements of the inspiratory and expiratory arcs and vertices, given by $\mathcal{V}_I, \mathcal{V}_E$ and $\mathcal{A}_{IO}, \mathcal{A}_{OE}$, respectively. A more detailed representation of the graph of the mechanical ventilator in Figure 8 is given in Figure 9.

Figure 9 shows a detailed graph of the mechanical ventilator given in Figure 2 that indicates the patient vertex $\mathcal{V}_O$ and the arcs and vertices along the inspiratory path, given by

$$\mathcal{A}_I := \{\mathcal{A}_{IV_{12}}, \mathcal{A}_{IV_{23}}, \mathcal{A}_{IV_{3B}}, \mathcal{A}_{IV_{23}}, \mathcal{A}_{IBV_4}, \mathcal{A}_{IV_{4S1}}, \mathcal{A}_{IBV_5}, \mathcal{A}_{IV_{5S2}}\}$$

and

$$\mathcal{V}_I := \{\mathcal{V}_{IV_1}, \mathcal{V}_{IV_2}, \mathcal{V}_{IV_3}, \mathcal{V}_{IB}, \mathcal{V}_{IV_4}, \mathcal{V}_{IS_1}, \mathcal{V}_{IV_5}, \mathcal{V}_{IS_2}\},$$

respectively, as well as the arcs and vertices along the expiratory path, given by

$$\mathcal{A}_E := \{\mathcal{A}_{OE_{V1}}, \mathcal{A}_{EV_{12}}, \mathcal{A}_{EV_{23}}\} \quad \text{and} \quad \mathcal{V}_E := \{\mathcal{V}_{EV_1}, \mathcal{V}_{EV_2}, \mathcal{V}_{EV_3}\},$$

respectively.

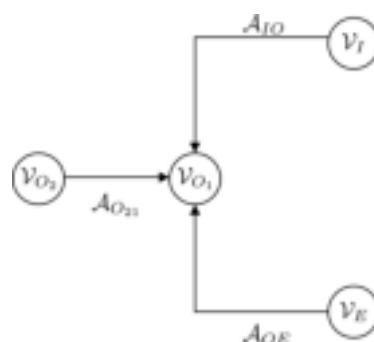
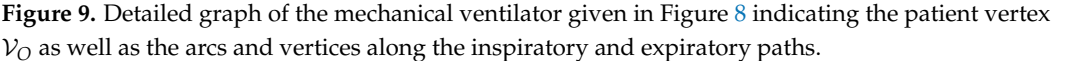


Figure 8. Simplified graph of the mechanical ventilator given in Figure 2 indicating the elements belonging to the patient, given by the arcs $\mathcal{A}_{O_{21}} \in \mathcal{A}_O$ and vertices $\{\mathcal{V}_{O_1}, \mathcal{V}_{O_2}\} \in \mathcal{V}_O$, as well as elements of the inspiratory and expiratory arcs and vertices, given by $\mathcal{V}_I, \mathcal{V}_E$ and $\mathcal{A}_{IO}, \mathcal{A}_{OE}$, respectively.



In this section, the coupling conditions of the Port-Hamiltonian network model of the mechanical ventilator are given. The types of interconnections occurring in this model are given below:

- $$P_b = p_0 \quad \text{and} \quad Q_b = m(0).$$

- $$P_v = p(0) \quad \text{and} \quad Q_v = m(0) \quad \text{at the inlet,}$$

$$P_v = p(\ell) \quad \text{and} \quad Q_v = m(\ell) \quad \text{at the outlet.}$$

- $$p(\ell) = v_{c_{\text{in}}} \quad \text{and} \quad m(\ell) = i_{c_{\text{in}}}.$$

On the other hand, the output voltage and current of a circuit can be interconnected with a fluid pipe at the inlet of the pipe. In this case, the output voltage and or current

of the circuit should be equal to the inlet pressure and inlet flow rate, respectively. This relation can be expressed mathematically as:

$$v_{c_{out}} = p(0) \quad \text{and} \quad i_{c_{out}} = m(0) \quad \text{at the inlet.}$$

The Hamiltonian of the complete system is given by the sum of the Hamiltonian's of the individual systems:

$$\mathcal{H} = \mathcal{H}_b + \mathcal{H}_s + \mathcal{H}_p + \mathcal{H}_c. \quad (37)$$

The rate of change of the energy of the complete system is

$$\begin{aligned} \dot{\mathcal{H}} = & \dot{P}_b(\partial_z p(0) + F_p) + \dot{Q}_b(\partial_z m(0) + M_p) \\ & + \dot{P}_{v_{in}}(\partial_z p(0) + F_{v_{in}}) + \dot{Q}_{v_{in}}(\partial_z m(0) + M_{v_{in}}) \\ & + \dot{P}_{v_{out}}(\partial_z p(\ell) + F_{v_{out}}) + \dot{Q}_{v_{out}}(\partial_z m(\ell) + M_{v_{out}}) \\ & + \dot{P}_{v_{out}}(\partial_z p(\ell) + F_{v_{out}}) + \dot{Q}_{v_{out}}(\partial_z m(\ell) + M_{v_{out}}) \\ & + \dot{\mathcal{H}}_b + \dot{\mathcal{H}}_s + \dot{\mathcal{H}}_c. \end{aligned} \quad (38)$$

The terms $F_p, M_p, F_{v_{in}}, M_{v_{in}}, F_{v_{out}}$ and $M_{v_{out}}$ are the external pressures and flow rates acting on the system. They should be equal to zero to complete the interconnection.

7. Structure-Preserving Discretization

The Port–Hamiltonian model of the pipe is a partial differential equation that is continuous in space. As such, it is difficult to simulate the dynamics of a pipe section. In order to do this, it is necessary to approximate the model with a discrete model: in this case, a finite difference model that is an approximation of the original system. Within the context of Port–Hamiltonian systems, an additional requirement is the need to ensure that the discrete approximation maintains the structural properties of the original system: e.g., skew symmetry, etc.

Each system state can be replaced with a discrete approximation consisting of a total on n elements, as can be seen in Figure 10. As such, $q \approx q_d = [q_1, q_2, \dots, q_n] \in \mathbb{R}^{n \times 1}$, and $m \approx m_d = [m_1, m_2, \dots, m_n] \in \mathbb{R}^{n \times 1}$. As such, state vector χ_p can be replaced by a discrete approximation $\chi_d = [q_1, \dots, q_n, m_1, \dots, m_n] \in \mathbb{R}^{2n \times 1}$. The i th element of $q_i, m_i \in \mathbb{R}^{1 \times 1}$ is located at $z = \{\Delta(i-1), \Delta(i-1/2)\}$, where Δ is the fixed discrete step size between points, and $i = 1, 2, \dots, n$. In addition, the efforts at the boundaries are given by $\delta_{\chi_0} \mathcal{H}_0$ and $\delta_{\chi_n} \mathcal{H}_n$. Thus, the Hamiltonian given in Equation (20) can be approximated by a discrete approximation such that $\mathcal{H}_p[\chi_p] \approx \Delta \mathcal{H}_d[\chi_d]$ so that the discrete system effort is now $\delta_{\chi_d}(\mathcal{H}_d)$. A finite difference approximation of the spatial derivatives at the i th point is

$$\left. \frac{\partial}{\partial z} q(t, z) \right|_i \approx \frac{1}{\Delta} (q(t, z_{i+0.5}) - q(t, z_{i-0.5})) \quad \text{and} \quad \left. \frac{\partial}{\partial z} m(t, z) \right|_i \approx \frac{1}{\Delta} (m(t, z_{i+1}) - m(t, z_i)). \quad (39)$$

The central difference approximation at the i th point is

$$\partial_t \begin{bmatrix} q_i \\ m_i \end{bmatrix} = \frac{1}{\Delta} \left(\begin{bmatrix} \delta_m \mathcal{H}[\chi_i] \\ \delta_q \mathcal{H}[\chi_{i+1}] \end{bmatrix} - \begin{bmatrix} \delta_m \mathcal{H}[\chi_{i-1}] \\ \delta_q \mathcal{H}[\chi_i] \end{bmatrix} \right). \quad (40)$$

In matrix form, this is

$$\partial_t q_d = \frac{1}{\Delta} \begin{bmatrix} -1 & & & \\ 1 & -1 & & \\ & \ddots & \ddots & \\ & & 1 & -1 \end{bmatrix} \delta_m \mathcal{H}[\chi_d] + \frac{1}{\Delta} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \delta_m \mathcal{H}[\chi_d] \quad (41)$$

and

$$\partial_t m_d = \frac{1}{\Delta} \begin{bmatrix} 1 & -1 & & \\ & \ddots & \ddots & \\ & & 1 & -1 \\ & & & 1 \end{bmatrix} \delta_\varrho \mathcal{H}[\chi_d] + \frac{1}{\Delta} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ -1 \end{bmatrix} \delta_\varrho \mathcal{H}[\chi_0], \quad (42)$$

which can be re-written as

$$\partial_t \varrho_d = D \delta_m \mathcal{H}[\chi_d] + \mathcal{G}_\varrho \delta_m \mathcal{H}[\chi_0] \quad \text{and} \quad \partial_t m_d = -D^\top \delta_\varrho \mathcal{H}[\chi_d] + \mathcal{G}_m \delta_\varrho \mathcal{H}[\chi_n], \quad (43)$$

where

$$D = \frac{1}{\Delta} \begin{bmatrix} -1 & & & \\ 1 & -1 & & \\ & \ddots & \ddots & \\ & & 1 & -1 \end{bmatrix}, \quad \mathcal{G}_\varrho = \frac{1}{\Delta} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{and} \quad \mathcal{G}_m = \frac{1}{\Delta} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ -1 \end{bmatrix}.$$

The skew-symmetric operator is clear from Equation (43).

$$\begin{bmatrix} \partial_t \varrho_d \\ \partial_t m_d \end{bmatrix} = \begin{bmatrix} D \\ -D^\top \end{bmatrix} \begin{bmatrix} \delta_\varrho \mathcal{H}[\chi_d] \\ \delta_m \mathcal{H}[\chi_d] \end{bmatrix} + \begin{bmatrix} \mathcal{G}_\varrho \\ \mathcal{G}_m \end{bmatrix} \begin{bmatrix} \delta_\varrho \mathcal{H}[\chi_0] \\ \delta_m \mathcal{H}[\chi_n] \end{bmatrix}. \quad (44)$$

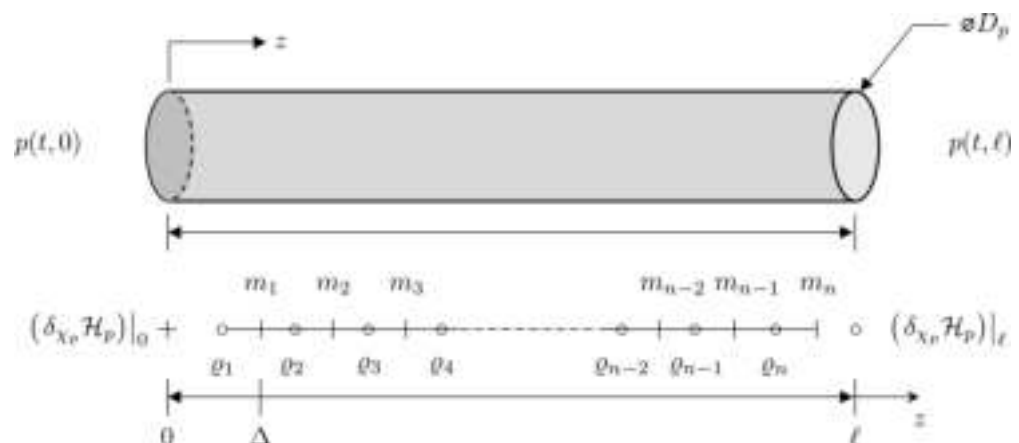


Figure 10. Staggered grid discretization of the one-dimensional Port-Hamiltonian pipe dynamic model.

8. Results and Discussion

The pipe and solenoid valve model parameters used in this work are given in Table 1 and Table 2, respectively. The Lung circuit model parameters are given in Table 3.

Table 1. Pipe model parameter values.

Parameter	Description	Value	Units
A_p	Pipe cross-sectional area	3.8013×10^{-4}	m^2
D_p	Pipe diameter	2.2×10^{-3}	m
ℓ	Pipe length	1.5×10^{-1}	m

Table 2. Solenoid model parameter values [28].

Parameter	Description	Value	Units
A_e	Effective cross-sectional area	8×10^{-5}	m^2
A_{s1}	Spool surface area	2.657×10^{-3}	m^2
A_{s2}	Spool surface area	3.525×10^{-3}	m^2
A_{s3}	Spool surface area	5.586×10^{-3}	m^2
A_{s4}	Spool surface area	5.586×10^{-3}	m^2
b_s	Viscous damping factor	2×10^{-1}	$\text{Ns}\cdot\text{m}^{-1}$
g	Acceleration due to gravity	9.81	$\text{m}\cdot\text{s}^{-2}$
k_s	Spring stiffness	1×10^{-4}	$\text{N}\cdot\text{m}^{-1}$
l_c	Length of the part of the magnetic circuit inside the core	1.15×10^{-2}	m
m_s		2.7×10^{-1}	kg
N	Number of turns in the coil	1250	turns
$q_{s\text{tot}}$	Total air gap	3.3×10^{-4}	m
q_{s0}	Pre-tension in the spring	1.1×10^{-3}	m
R_s	Resistance of the coil	13	Ω
μ_0	Permeability of air	$4\pi \times 10^{-7}$	$\text{N}\cdot\text{A}^{-2}$
μ_c	Permeability of the magnetic core	$4.8\pi \times 10^{-5}$	$\text{N}\cdot\text{A}^{-2}$

Table 3. Lung circuit model parameter values [26].

Parameter	Description	Value	Units
R_{ML}	Resistance from the mouth to larynx	1.021	$\text{cmH}_2\text{O}\cdot\text{s}\cdot\text{L}^{-1}$
R_{LT}	Resistance from the larynx to trachea	3.369×10^{-1}	$\text{cmH}_2\text{O}\cdot\text{s}\cdot\text{L}^{-1}$
R_{TB}	Resistance from the trachea to bronchi	3.063×10^{-1}	$\text{cmH}_2\text{O}\cdot\text{s}\cdot\text{L}^{-1}$
R_{BA}	Resistance from the bronchi to alveoli	8.17×10^{-2}	$\text{cmH}_2\text{O}\cdot\text{s}\cdot\text{L}^{-1}$
C_L	Compliance of the larynx	1.27×10^{-3}	$\text{L}/\text{cmH}_2\text{O}$
C_T	Compliance of the trachea	2.38×10^{-3}	$\text{L}/\text{cmH}_2\text{O}$
C_B	Compliance of the bronchi	1.31×10^{-2}	$\text{L}/\text{cmH}_2\text{O}$
C_A	Compliance of the alveoli	2×10^{-1}	$\text{L}/\text{cmH}_2\text{O}$
C_{CW}	Compliance of the chest wall	2.445×10^{-1}	$\text{L}/\text{cmH}_2\text{O}$

8.1. Model Validation

The model was validated using parameters obtained from the literature. The simulation results provided in the results and discussion section were found to be comparable to those in the existing literature.

8.2. Simulation Environment

Simulations were conducted using MATLAB Version: 9.13.0.2193358 (R2022b) Update 5 on an Intel Core i5 CPU with four cores and 8 GB of RAM. Figures 11 and 12 show ventilation curves for a lung under two conditions: a partially damaged lung with a compliance and resistance of 20 mL/cmH₂O and 20 cmH₂O-s/L, respectively, and a lung with complete damage, which is represented by a lung compliance and resistance of 10 mL/cmH₂O and 50 cmH₂O-s/L, respectively. The ventilation curves maintain the same sinusoidal behaviour. For the damaged lung case study, minimum ventilation values of flow over time and air volume are achieved. In the damaged lung, the compliance and resistance of the lung increase the pressure.

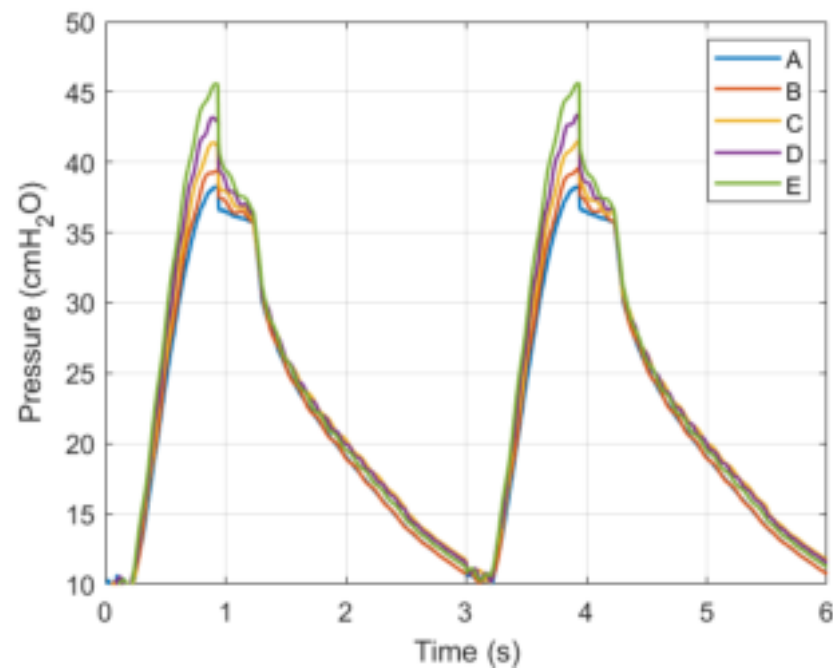


Figure 11. Graph of air pressure versus time.

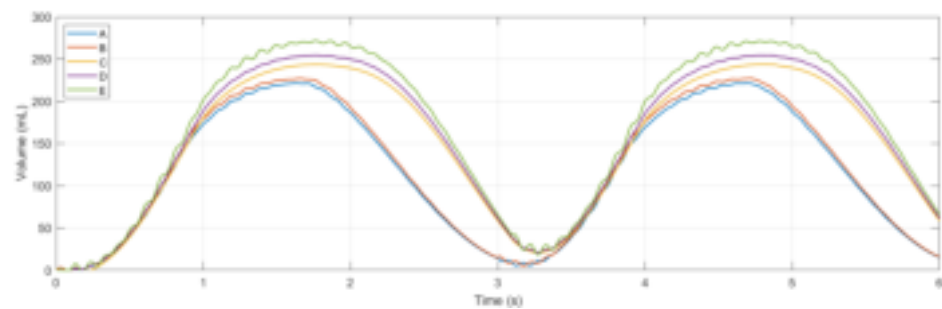


Figure 12. Graph of volume flow versus time.

9. Conclusions and Recommendations

In this research work, we formulate a detailed mechanical ventilator in the Port–Hamiltonian framework. This is followed by a Port–Hamiltonian model of the respiratory system, and thereafter, these two systems are integrated. The conducted work demonstrates that the Port–Hamiltonian approach is a valid method for the modelling of an integrated mechanical ventilator–human respiratory system. The Port–Hamiltonian framework is modular in nature. Furthermore, the complete model structure is based on the underlying graph of the mechanical ventilator and human respiratory system. Future research directions include:

- Developing a software framework for the automatic generation of the Port–Hamiltonian dynamics for various mechanical ventilator and human respiratory system configurations as well as their associated integrated models [19];
- Development of various reduced-order models;
- Development of an integrated framework for the modelling of mechanical ventilator/human respiratory systems to enable control and system optimization [19,20];
- Taking advantage of the modular and graph theoretic nature of the Port–Hamiltonian approach to incorporate machine learning and artificial intelligence [19];
- Investigation of the most appropriate structure-preserving discretizations and model-order reductions for various data-driven applications and control system designs.

Author Contributions: Conceptualization, M.C.I.M.; methodology, M.C.I.M.; software, M.C.I.M.; validation, M.C.I.M. and J.E.D.E.; formal analysis, M.C.I.M.; investigation, M.C.I.M.; resources, M.C.I.M.; data curation, M.C.I.M.; writing—original draft preparation, M.C.I.M.; writing—review and editing, M.C.I.M.; visualization, M.C.I.M.; supervision, O.T.C.N.; project administration, M.C.I.M.; funding acquisition, M.C.I.M. All authors have read and agreed to the published version of the manuscript.

Funding: I would like to thank the School of Mining, Faculty of Engineering, and the Built Environment, University of the Witwatersrand, for the financial support to publish this work.

Institutional Review Board Statement: Ethical review and approval were waived for this study because the study did not involve humans or animals.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Rubio, J.; Rojas, C.; Sanchez, M.; Gómez-Alzate, D.; Córdova, M.; Montoya, V.; Castaneda, B.; Chang, J.; Pérez-Buitrago, S. COVOX: Providing oxygen during the COVID-19 health emergency. *HardwareX* **2023**, *13*, e00383. [\[CrossRef\]](#)
2. Hickling, K.G. The Pressure–Volume Curve Is Greatly Modified by Recruitment. *Am. J. Respir. Crit. Care Med.* **1998**, *158*, 194–202. [\[CrossRef\]](#)
3. Bates, J.H.T. *Lung Mechanics: An Inverse Modeling Approach*; Cambridge University Press: Cambridge, UK, 2009.
4. Maury, B. *The Respiratory System in Equations*, 1st ed.; MS&A—Modeling, Simulation and Applications; Springer: Milan, Italy, 2013. [\[CrossRef\]](#)
5. Burrowes, K.S.; Hunter, P.J.; Tawhai, M.H. Anatomically based finite element models of the human pulmonary arterial and venous trees including supernumerary vessels. *J. Appl. Physiol.* **2005**, *99*, 731–738. [\[CrossRef\]](#)
6. Tawhai, M.H.; Bates, J.H.T. Multi-scale lung modeling. *J. Appl. Physiol.* **2011**, *110*, 1466–1472. [\[CrossRef\]](#)
7. Berger, L.; Bordas, R.; Burrowes, K.; Brightling, C.; Hartley, R.; Kay, D. Understanding the Interdependence Between Parenchymal Deformation and Ventilation In Obstructive Lung Disease. In *B30. Dynamics of Airway Narrowing in Asthma: Still Misunderstood?* American Thoracic Society: New York, NY, USA, 2014; Volume 189, p. A2677.
8. Roth, C.J.; Yoshihara, L.; Ismail, M.; Wall, W.A. Computational modelling of the respiratory system: Discussion of coupled modelling approaches and two recent extensions. *Comput. Methods Appl. Mech. Eng.* **2017**, *314*, 473–493. [\[CrossRef\]](#)
9. Tran, A.S.; Thinh Ngo, H.Q.; Dong, V.K.; Vo, A.H. Design, Control, Modeling, and Simulation of Mechanical Ventilator for Respiratory Support. *Math. Probl. Eng.* **2021**, *2021*, 2499804. [\[CrossRef\]](#)
10. El-Hadj, A.; Kezrane, M.; Ahmad, H.; Ameer, H.; Bin Abd Rahim, S.Z.; Younsi, A.; Abu-Zinadah, H. Design and simulation of mechanical ventilators. *Chaos Solitons Fractals* **2021**, *150*, 111169. [\[CrossRef\]](#)
11. Tharion, J.; Kapil, S.; Muthu, N.; Tharion, J.; Subramani, K. Rapid Manufacturable Ventilator for Respiratory Emergencies of COVID-19 Disease. *Trans. Indian Natl. Acad. Eng.* **2020**, *5*, 373–378. [\[CrossRef\]](#)
12. Pivik, W.J.; Clayton, G.M.; Jones, G.F.; Nataraj, C. Dynamic Modeling of a Low-cost Mechanical Ventilator. *IFAC-PapersOnLine* **2022**, *55*, 81–85. [\[CrossRef\]](#)
13. Al-Naggar, N. Modelling and Simulation of Pressure Controlled Mechanical Ventilation System. *J. Biomed. Sci. Eng.* **2015**, *8*, 707–716. [\[CrossRef\]](#)
14. Shi, Y.; Ren, S.; Cai, M.; Xu, W. Modelling and Simulation of Volume Controlled Mechanical Ventilation System. *Math. Probl. Eng.* **2014**, *2014*, 271053. [\[CrossRef\]](#)
15. Al-Naggar, N.Q.; Al-Hetari, H.Y.; Al-Akwaa, F.M. Simulation of Mathematical Model for Lung and Mechanical Ventilation. *J. Sci. Technol.* **2016**, *21*, 1–9. [\[CrossRef\]](#)
16. Giri, J.; Kshirsagar, N.; Wanjari, A. Design and simulation of AI-based low-cost mechanical ventilator: An approach. *Mater. Today Proc.* **2021**, *47*, 5886–5891. [\[CrossRef\]](#)
17. Hannon, D.M.; Mistry, S.; Das, A.; Saffaran, S.; Laffey, J.G.; Brook, B.S.; Hardman, J.G.; Bates, D.G. Modeling Mechanical Ventilation In Silico—Potential and Pitfalls. *Semin. Respir. Crit. Care Med.* **2022**, *43*, 335–345. [\[CrossRef\]](#)
18. Mehedi, I.M.; Shah, H.S.; Al-Saggaf, U.M.; Mansouri, R.; Bettayeb, M. Fuzzy PID control for respiratory systems. *J. Healthc. Eng.* **2021**, *2021*, 1926711. [\[CrossRef\]](#)
19. Mehrmann, V.; Unger, B. Control of port-Hamiltonian differential-algebraic systems and applications. *Acta Numer.* **2023**, *32*, 395–515. [\[CrossRef\]](#)
20. van der Schaft, A. Port-Hamiltonian Modeling for Control. *Annu. Rev. Control Robot. Auton. Syst.* **2020**, *3*, 393–416. [\[CrossRef\]](#)
21. Villegas, J.A. A Port-Hamiltonian Approach to Distributed Parameter Systems. Ph.D. Thesis, Department of Applied Mathematics, Faculty EWI, Universiteit Twente, Enschede, The Netherlands, 2007.
22. Le Gorrec, Y.; Zwart, H.; Maschke, B. Dirac structures and Boundary Control Systems associated with Skew-Symmetric Differential Operators. *SIAM J. Control Optim.* **2005**, *44*, 1864–1892. [\[CrossRef\]](#)

23. van der Schaft, A.J.; Maschke, B.M. Hamiltonian formulation of distributed-parameter systems with boundary energy flow. *J. Geom. Phys.* **2002**, *42*, 166–194. [[CrossRef](#)]
24. Anderson, J. *Computational Fluid Dynamics: The Basics with Applications*; McGraw-Hill International Editions: Mechanical Engineering; McGraw-Hill Inc.: New York, NY, USA, 1995.
25. Kamiński, Z. A simplified lumped parameter model for pneumatic tubes. *Math. Comput. Model. Dyn. Syst.* **2017**, *23*, 523–535. [[CrossRef](#)]
26. Albanese, A.; Cheng, L.; Ursino, M.; Chbat, N.W. An integrated mathematical model of the human cardiopulmonary system: Model development. *Am. J. Physiol.-Heart Circ. Physiol.* **2016**, *310*, H899–H921. [[CrossRef](#)] [[PubMed](#)]
27. Bondy, J.A.; Murty, U.S.R. *Graph Theory with Applications*; The Macmillan Press Ltd.: London, UK, 1976.
28. Taghizadeh, M.; Ghaffari, A.; Najafi, F. Modeling and identification of a solenoid valve for PWM control applications. *Comptes Rendus Mécanique* **2009**, *337*, 131–140. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

The Effect of Proportional, Proportional-Integral, and Proportional-Integral-Derivative Controllers on Improving the Performance of Torsional Vibrations on a Dynamical System

Khalid Alluhydan ¹, Ashraf Taha EL-Sayed ^{2,*} and Fatma Taha El-Bahrawy ²

¹ Department of Mechanical Engineering, College of Engineering, King Saud University, P.O Box 800, Riyadh 11421, Saudi Arabia; kalluhydan@ksu.edu.sa

² Department of Basic Science, Modern Academy for Engineering and Technology, Elmokattam 11439, Egypt; fatma.taha@eng.modern-academy.edu.eg

* Correspondence: atelsayed@eng.modern-academy.edu.eg

Abstract: The primary goal of this research is to lessen the high vibration that the model causes by using an appropriate vibration control. Thus, we begin by implementing various controller types to investigate their impact on the system's reaction and evaluate each control's outcomes. The controller types are presented as proportional (P), proportional-integral (PI), and proportional-integral-derivative (PID) controllers. We employed PID control to regulate the torsional vibration behavior on a dynamical system. The PID controller aims to increase system stability after seeing the impact of P and PI control. This kind of control ensures that there are no unstable components in the system. By using the multiple time scale perturbation (MTSP) technique, a first-order approximate solution has been obtained. Using the frequency response function approach, the stability and steady-state response of the system at the primary resonance scenario ($\Omega_1 \cong \omega_1, \Omega_2 \cong \omega_2$) are considered as the worst resonance and addressed. Additionally examined are the nonlinear dynamical system's chaotic response and the numerical solution for various parameter values. The MATLAB programs are utilized to attain simulation outcomes.

Keywords: MSPT; P; PI; PID controllers; nonlinear differential equations; active control; stability; torsional vibration



Citation: Alluhydan, K.; EL-Sayed, A.T.; El-Bahrawy, F.T. The Effect of Proportional, Proportional-Integral, and Proportional-Integral-Derivative Controllers on Improving the Performance of Torsional Vibrations on a Dynamical System. *Computation* **2024**, *12*, 157. <https://doi.org/10.3390/computation12080157>

Academic Editor:
Chathura Wanigasekara

Received: 15 July 2024
Revised: 24 July 2024
Accepted: 29 July 2024
Published: 3 August 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Torsional vibration occurs in some form in all rotating machinery. Even when the vibration is nearing destructive amplitude, there are situations in which it cannot be identified without specialized monitoring equipment. After the shaft bending stiffness and diametral moment of inertia have been replaced by the twisting stiffness and polar moment of inertia, respectively, many elements of torsional vibration are equivalent to shaft vibration. When something mediates the connection between the vibration and the ground, or when gear teeth or coupling jaws are empty, torsional vibration can be detected by the noise level and vibration (perceptible to touch). Torsional vibration can interact with the ground through gear sets used to change the speed of power transmission systems; in reciprocating machines, the path to the ground is provided by sliding crank mechanisms found in engines and compressors. Typically, torsional vibration manifests as a complicated vibration signal with numerous frequency components. Some systems experience brief torsional vibrations due to shock from sudden starts and the unloading of gear teeth; synchronous electric motor systems may also experience torsional resonance at startup.

Conventional PID controllers have been widely adopted in numerous industrial applications due to their simple design, affordable price, simplicity of maintenance, and there being ready-made modules [1–3]. Ref. [4] offers the implementation of nonlinear state-dependent (SDP-PID+) control employing the SDP transfer function model, a type

of nonlinear description of dynamical structures. Fine-tuning traditional PID control is problematic owing to the specific properties of numerous procedures, including binary sample time delays, longer durations, higher-order TF models, and nonlinearities [5–7]. Sayed et al. [8] developed a nonlinear resilient SDP PID control approach for a discrete transfer function with nonlinear properties. This research emphasized robust response processes, for which the robust PID and SDP-PID techniques were specifically designed. Dano and Julli'ere [9] studied how MFC actuators controlled oscillations in a composite structure. Kumar and Ray [10] examined vibrations in sandwich shells with between one and three piezoelectric composites and layering damping techniques. The PD controller is designed to decrease oscillations of a hung Jeffcott rotor through two pairs of poles [11]. A PD controller can effectively relax a beam system's steady-state amplitude vibrations [12]. Eissa et al. [13,14] suggested a PD controller and a time-delayed PD controller to reduce the vibrations of magnetic systems with cubic and quadratic nonlinear coefficients underneath parametric principal forces. Bauomy and El-Sayed employed a PD controller to control the behavior of the MFC laminated shell structure. The purpose of employing the PD controller is to improve system stability by increasing control, as it can forecast future errors in the framework response [15]. More than 90% of industries still utilize PID controllers due to their simplicity, functionality, and ease of usage [16]. PID controller gains are acquired by matching the frequency response of the closed-loop control system [17]. Recently, a systematic approach was used to select PID parameters for nonlinear uncertain structures [18]. PID controllers offer advantages over passive approaches for controlling semi-active suspension systems [19–21]. Ref. [22] investigates the control of a quarter-car semi-active suspension system utilizing a PID controller. The typical PID controller is constructed using the Ziegler–Nichols approach and is used to regulate the suspension system. The torsional vibrations of a one-degree-of-freedom nonlinear dynamical structure are regulated with active control [23]. Wenzhi and Zhiyong [24] suggested an active control to decrease torsional vibration in a big turbogenerator with a rotor shaft. Whole-state feedback control using a linear quadratic regulator (LQR) effectively reduces torsional vibration energy and response in the turbogenerator shaft system. El-Sayed and Bauomy [25] they succeeded in reducing the torsional vibration of a nonlinear dynamical system using passive and active control methods. The research study [26] proposes an adaptive PI event-triggered control approach for MIMO nonlinear systems with unpredictable input delay. An adaptive proportional/proportional-integral (P/PI) control strategy is proposed for a solar-driven volumetric methane/steam reforming reactor (SVMSR) with passive thermal management. The strategy aims to stabilize product components and reduce fluctuations in the hydrogen production rate under fluctuating radiation conditions [27]. Previous research, such as studies [28–30], has demonstrated the successful application of the multiple-scale perturbation technique to derive approximate solutions for various vibrating systems. These studies often utilize MATLAB programs to analyze and solve vibration problems.

A cantilever beam model was explored and derived to load an intermediate lumped mass under harmonic excitation. The vibration suppression can be succeeded by using IRC+NSC controller. To determine the unstable and stable zones for each frequency response curve, numerical stability research was carried out [31]. This research investigates the use of a nonlinear spring pendulum for the vibration control of ship roll motion. Three second-order nonlinear differential equations—one for the rotation angle, one for the relative elongation of the absorber spring, and one for the elongation of the absorber spring—make up the mathematical model that depicts the ship roll motion with the absorber. The authors use a series solution in which they account for terms up to the fourth order in trigonometric functions of the rotational angle. According to the authors, the nonlinear spring-pendulum system's two modes may be made to respond to multi-parametric excitation forces in a way that reduces their maximum values by 8.8% and 0.02% [32]. The PPF control of the nonlinear GMA framework has been described in [33]. The basic resonance and framework amplitude stability can be successfully constrained by tuning the PPF limitations. A few

experiments have been carried out to verify the accuracy of the findings. The investigation illustrated the growth of the framework and contrasted its usefulness to earlier research. The 3D plot improves and illustrates the work's correctness. Ref. [34] suggests using NIPPF controllers to manage the nonlinear vibration of a spinning shaft's primary resonance vibration. Among these is a comparison between the controllers for the FRCs under study, NIPPF, and ANIPPF. A two-degree-of-freedom system, counting quadratic and cubic nonlinearities among the parametric and external forces, demonstrate the calculated system. Multiple scales are joined in a connected manner to analyze the stability of the measured structure and obtain approximations of solutions. From the mathematical solution, every resonance is retrieved. The Runge–Kutta fourth-order process is used to gauge the system's performance. Within the numerical results, the examined structure's scheduled frequency response curves are examined for influences including significant coefficients [35]. A low-speed and high-torque permanent magnet synchronous motor powers the semi-direct drive-cutting gearbox system of a shearer, which uses nonlinear integral positive position feedback (NIPPF) and adaptive nonlinear integral positive position feedback (ANIPPF) controllers. Using the averaging technique to solve the nonlinear differential equations and modeling the system with controllers yields an analytic solution in the case of primary and 1:1 internal resonance. The MATLAB program was used to compare the numerical and analytical solutions for time history and FRCs in order to verify their comparability [36].

In this work, we utilized PID control to suppress the torsional vibration after studying the model on a website [37]. A two-degree-of-freedom system under multiple excitations results from this. To provide an estimated solution up to the first-order approximations, MSPT is employed throughout. Using frequency response functions, the stability of the system is examined in the vicinity of the principal resonance case. A few suggestions on the system's various parameters are given. Numerical examples are provided to show how active controllers affect the behavior of the system. A comparison is shown between PID control and additional controllers.

2. Mathematical Modeling

2.1. System Dynamics without Control

This section is presented to illustrate the investigation of the torsional vibration dynamical system. The dynamical model in this work consists of two coupled parts, as shown in Figure 1. θ_1 and θ_2 represent the angular positions (generalized coordinates) within the system. In this nonlinear dynamical system, the following hold:

- I_1 and I_2 represent the polar mass moments of inertia.
- k_1 is the linear spring stiffness of the first part.
- k_2 represents the spring stiffness of the second part, which comprises the following:
 - k_{21} : a linear component.
 - k_{22} : nonlinear quadratic and cubic components.
- $F_1^\bullet$ and $F_2^\bullet$ denote the generalized excitation forces.

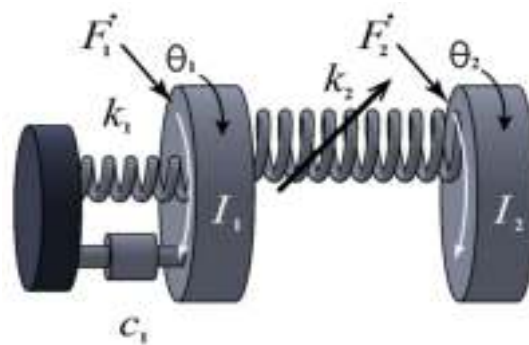


Figure 1. Model of schematic diagram of main system.

Therefore, the equations of the motions of the free body system found in Figure 1 can be constructed in the same manner as those in [25,37]:

$$I_1 \ddot{\theta}_1 + k_1 \theta_1 + c_1 \dot{\theta}_1 + k_{21}(\theta_1 - \theta_2) + k_{22}(\theta_1 - \theta_2)^2 + k_{23}(\theta_1 - \theta_2)^3 = F_1^\bullet, \quad (1)$$

$$I_2 \ddot{\theta}_2 + k_{21}(\theta_2 - \theta_1) - k_{22}(\theta_1 - \theta_2)^2 - k_{23}(\theta_1 - \theta_2)^3 = F_2^\bullet, \quad (2)$$

$$\ddot{\theta}_1 + \frac{k_1}{I_1} \theta_1 + \frac{c_1}{I_1} \dot{\theta}_1 + \frac{k_{21}}{I_1} (\theta_1 - \theta_2) + \frac{k_{22}}{I_1} (\theta_1 - \theta_2)^2 + \frac{k_{23}}{I_1} (\theta_1 - \theta_2)^3 = \frac{F_1^\bullet}{I_1}, \quad (3)$$

$$\ddot{\theta}_1 + \left(\frac{k_1 + k_{21}}{I_1} \right) \theta_1 + \frac{c_1}{I_1} \dot{\theta}_1 - \frac{k_{21}}{I_1} \theta_2 + \frac{k_{22}}{I_1} (\theta_1 - \theta_2)^2 + \frac{k_{23}}{I_1} (\theta_1 - \theta_2)^3 = \frac{F_1^\bullet}{I_1}, \quad (4)$$

$$\ddot{\theta}_2 + \frac{k_{21}}{I_2} (\theta_2 - \theta_1) - \frac{k_{22}}{I_2} (\theta_1 - \theta_2)^2 - \frac{k_{23}}{I_2} (\theta_1 - \theta_2)^3 = \frac{F_2^\bullet}{I_2}, \quad (5)$$

$$\frac{\ddot{\theta}_1}{\theta_0} + \left(\frac{k_1 + k_{21}}{I_1} \right) \frac{\theta_1}{\theta_0} + \frac{c_1}{I_1} \frac{\dot{\theta}_1}{\theta_0} - \frac{k_{21}}{I_1} \frac{\theta_2}{\theta_0} + \frac{k_{22} \theta_0}{I_1} \frac{(\theta_1 - \theta_2)^2}{\theta_0^2} + \frac{k_{23} \theta_0^2}{I_1} \frac{(\theta_1 - \theta_2)^3}{\theta_0^3} = \frac{F_1^\bullet}{I_1}, \quad (6)$$

$$\frac{\ddot{\theta}_2}{\theta_0} + \frac{k_{21}}{I_2} \frac{(\theta_2 - \theta_1)}{\theta_0} - \frac{k_{22} \theta_0}{I_2} \frac{(\theta_1 - \theta_2)^2}{\theta_0^2} - \frac{k_{23} \theta_0^2}{I_2} \frac{(\theta_1 - \theta_2)^3}{\theta_0^3} = \frac{F_2^\bullet}{I_2}. \quad (7)$$

We will now introduce the dimensionless forms of the parameters used in this analysis.

$$\varphi_j = \frac{\theta_j}{\theta_0} (j = 1, 2), \omega_1^2 = \frac{k_1 + k_{21}}{I_1}, \zeta = \frac{c_1}{I_1}, \beta = \frac{k_{21}}{I_1}, \alpha_1 = \frac{k_{22} \theta_0}{I_1}, \alpha_2 = \frac{k_{23} \theta_0^2}{I_1}, \omega_2^2 = \frac{k_{21}}{I_2}, \alpha_3 = \frac{k_{22} \theta_0}{I_2},$$

$$\alpha_4 = \frac{k_{23} \theta_0^2}{I_2}, f_j \sin(\Omega_j t) = \frac{F_j^\bullet}{I_j}.$$

Based on the above parameters, we can obtain the following equations of motion in their dimensionless forms:

$$\ddot{\varphi}_1 + \omega_1^2 \varphi_1 + \zeta \dot{\varphi}_1 - \beta \varphi_2 + \alpha_1 (\varphi_1 - \varphi_2)^2 + \alpha_2 (\varphi_1 - \varphi_2)^3 = f_1 \sin(\Omega_1 t), \quad (8)$$

$$\ddot{\varphi}_2 + \omega_2^2 (\varphi_2 - \varphi_1) - \alpha_3 (\varphi_1 - \varphi_2)^2 - \alpha_4 (\varphi_1 - \varphi_2)^3 = f_2 \sin(\Omega_2 t). \quad (9)$$

2.2. System Dynamics with PID Control

The goal of the present section is to propose PID controllers to reduce harmful torsional vibration on the dynamical system in this work at one of the worst resonance cases, as depicted in Figure 2. Furthermore, the control in our work is represented as two net control forces F_{C1} and F_{C2} that are generated to suppress the torsional oscillations in two directions.

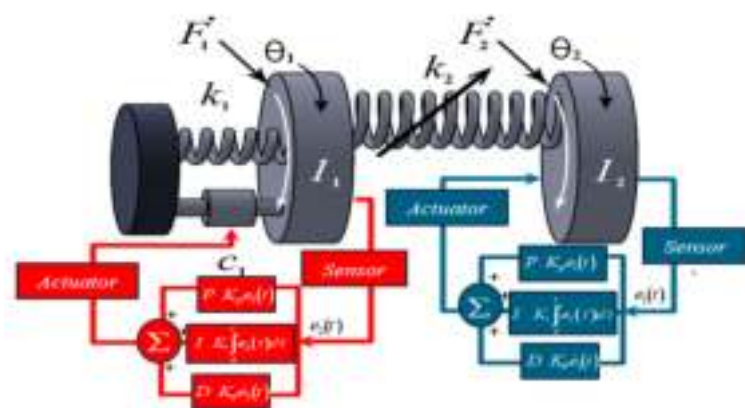


Figure 2. Model of schematic diagram of main system with PID controllers.

Here, K_p , K_i , and K_d denote the proportional feedback control gain, integral feedback control gain, and derivative feedback control gain, respectively. Also, $e_1(t)$ and $e_2(t)$ denote the error value with zero steady-state step error.

The equations of motion utilizing the PID controllers coupled to the nonlinear dynamical system as depicted in Figure 2 can be expressed as follows:

$$\ddot{\varphi}_1 + \omega_1^2 \varphi_1 + \varepsilon \zeta \dot{\varphi}_1 - \varepsilon \beta \varphi_2 + \varepsilon \alpha_1 (\varphi_1 - \varphi_2)^2 + \varepsilon \alpha_2 (\varphi_1 - \varphi_2)^3 = \varepsilon f_1 \sin(\Omega_1 t) + \varepsilon F_{C1}, \quad (10)$$

$$\ddot{\varphi}_2 + \omega_2^2 (\varphi_2 - \varphi_1) - \varepsilon \alpha_3 (\varphi_1 - \varphi_2)^2 - \varepsilon \alpha_4 (\varphi_1 - \varphi_2)^3 = \varepsilon f_2 \sin(\Omega_2 t) + \varepsilon F_{C2}, \quad (11)$$

$$F_{C1} = -K_p \varphi_1 - K_i \int_0^t \varphi_1(\tau) d\tau - K_d \dot{\varphi}_1, \quad (12a)$$

$$F_{C2} = -K_p \varphi_2 - K_i \int_0^t \varphi_2(\tau) d\tau - K_d \dot{\varphi}_2, \quad (12b)$$

$$\ddot{\varphi}_1 + \omega_1^2 \varphi_1 + \varepsilon \zeta \dot{\varphi}_1 - \varepsilon \beta \varphi_2 + \varepsilon \alpha_1 (\varphi_1 - \varphi_2)^2 + \varepsilon \alpha_2 (\varphi_1 - \varphi_2)^3 = \varepsilon f_1 \sin(\Omega_1 t) - \varepsilon K_p \varphi_1 - \varepsilon K_i \int_0^t \varphi_1(\tau) d\tau - \varepsilon K_d \dot{\varphi}_1, \quad (13)$$

$$\ddot{\varphi}_2 + \omega_2^2 (\varphi_2 - \varphi_1) - \varepsilon \alpha_3 (\varphi_1 - \varphi_2)^2 - \varepsilon \alpha_4 (\varphi_1 - \varphi_2)^3 = \varepsilon f_2 \sin(\Omega_2 t) - \varepsilon K_p \varphi_2 - \varepsilon K_i \int_0^t \varphi_2(\tau) d\tau - \varepsilon K_d \dot{\varphi}_2. \quad (14)$$

3. Analytical Investigations

3.1. Perturbation Analysis

The multiple-scales perturbation technique (MSPT) is applied within this section to obtain an approximation solution of the nonlinear dynamical system with the proposal control (i.e., PID control) given by Equations (13) and (14). Correspondingly, we were able to find a first-order approximate solution to Equations (13) and (14), proposed as follows [28,29]:

$$\varphi_n = \varphi_{n0} + \varepsilon \varphi_{n1} + O(\varepsilon^2), \quad (n = 1, 2) \quad (15)$$

where the minor perturbation parameter ε is located in the range of $0 < \varepsilon \ll 1$. Let us define two time scales, T_0 and T_1 , where $T_0 = t$ represents a fast scale while $T_1 = \varepsilon t$ is the slow one. The derivatives of time are converted into the following:

$$\frac{d}{dt} = \frac{dT_0}{dt} \frac{\partial}{\partial T_0} + \frac{dT_1}{dt} \frac{\partial}{\partial T_1} + \dots = D_0 + \varepsilon D_1 + \dots, \quad (16)$$

$$\frac{d^2}{dt^2} = \frac{d}{dt} \left(\frac{dT_0}{dt} \frac{\partial}{\partial T_0} + \frac{dT_1}{dt} \frac{\partial}{\partial T_1} + \dots \right) = D_0^2 + 2\varepsilon D_0 D_1 + \dots \quad (17)$$

where $D_j = \frac{\partial}{\partial T_j}$ ($j = 0, 1$).

The following set of ordinary differential equations was constructed by substituting Equations (15) and (16) into Equations (13) and (14) and equating the coefficients of the same power of ε in both sides:

$O(\varepsilon^0)$:

$$(D_0^2 + \omega_1^2) \varphi_{10} = 0, \quad (18a)$$

$$(D_0^2 + \omega_2^2) \varphi_{20} = \omega_2^2 \varphi_{10}. \quad (18b)$$

$O(\varepsilon^1)$:

$$(D_0^2 + \omega_1^2)\varphi_{11} = -2D_0D_1\varphi_{10} - \zeta D_0\varphi_{10} + \beta\varphi_{20} - \alpha_1(\varphi_{10} - \varphi_{20})^2 - \alpha_2(\varphi_{10} - \varphi_{20})^3 + f_1 \sin(\Omega_1 t) - K_p\varphi_{10} - K_i \int_0^t \varphi_{10}(\tau) d\tau - K_d D_0\varphi_{10}, \quad (19a)$$

$$(D_0^2 + \omega_2^2)\varphi_{21} = \omega_2^2\varphi_{11} - 2D_0D_1\varphi_{20} + \alpha_3(\varphi_{10} - \varphi_{20})^2 + \alpha_4(\varphi_{10} - \varphi_{20})^3 + f_2 \sin(\Omega_2 t) - K_p\varphi_{20} - K_i \int_0^t \varphi_{20}(\tau) d\tau - K_d D_0\varphi_{20}. \quad (19b)$$

The zeroth order approximation is represented by the general solution of Equations (18a) and (18b), which can be expressed as follows:

$$\varphi_{10} = A_1(T_1) \exp(i\omega_1 T_0) + cc., \quad (20a)$$

$$\varphi_{20} = A_2(T_1) \exp(i\omega_2 T_0) + \Gamma A_1(T_1) \exp(i\omega_1 T_0) + cc. \quad (20b)$$

where cc denotes the complex conjugate of the preceding components, A_m is the complex function in T_1 , and $\Gamma = \left(\frac{\omega_2^2}{\omega_2^2 - \omega_1^2}\right)$.

Substituting Equation (20) into Equation (19a), we have

$$\begin{aligned} (D_0^2 + \omega_1^2)\varphi_{11} = & \left[\begin{aligned} & -2i\omega_1 D_1 A_1 - i\omega_1 \zeta A_1 + \Gamma \beta A_1 - \alpha_2 \left(\frac{6(1-\Gamma)A_1 A_2 \bar{A}_2}{+3(1-\Gamma)^3 A_1^2 \bar{A}_1} \right) \\ & - \left(K_p - i\frac{K_i}{\omega_1} + K_d i\omega_1 \right) A_1 \end{aligned} \right] \exp(i\omega_1 T_0) \\ & + \left[-\alpha_1(1-\Gamma)^2 A_1^2 \right] \exp(2i\omega_1 T_0) + \left[-\alpha_2(1-\Gamma)^3 A_1^3 \right] \exp(3i\omega_1 T_0) + [\beta A_2 \\ & + \alpha_2(3A_2^2 \bar{A}_2 + 6(1-\Gamma)^2 A_1 \bar{A}_1 A_2)] \exp(i\omega_2 T_0) + [-\alpha_1 A_1^2] \exp(2i\omega_2 T_0) \\ & + [\alpha_2 A_2^3] \exp(3i\omega_2 T_0) + [2\alpha_1(1-\Gamma)A_1 A_2] \exp(i(\omega_1 + \omega_2)T_0) \\ & + [2\alpha_1(1-\Gamma)A_1 \bar{A}_2] \exp(i(\omega_1 - \omega_2)T_0) + [3\alpha_2(1-\Gamma)^2 A_1^2 A_2] \exp(i(2\omega_1 + \omega_2)T_0) \\ & + [3\alpha_2(1-\Gamma)^2 A_1^2 \bar{A}_2] \exp(i(2\omega_1 - \omega_2)T_0) + [-3\alpha_2(1-\Gamma)A_1 A_2^2] \exp(i(\omega_1 + 2\omega_2)T_0) \\ & + [-3\alpha_2(1-\Gamma)A_1 \bar{A}_2^2] \exp(i(\omega_1 - 2\omega_2)T_0) + \left[-\alpha_1 \left((1-\Gamma)^2 A_1 \bar{A}_1 + A_2 \bar{A}_2 \right) - iK_i \frac{A_1}{\omega_1} \right] \\ & + \left[\frac{-if_1}{2} \right] \exp(i\Omega_1 T_0) + cc., \end{aligned} \quad (21)$$

To obtain the solvability condition of Equation (21), the closeness term of the excitation frequencies (Ω_1) and natural frequency (ω_1) is described via employing the dimensionless detuning parameter σ_1 , as in the following:

$$\Omega_1 = \omega_1 + \varepsilon\sigma_1. \quad (22)$$

By substituting Equation (22) into Equation (21) and canceling the term that leads to secular ones, we obtain the particular solution of the first approximation (φ_{11}) as the following:

$$\begin{aligned} \varphi_{11} = & M_1 \exp(2i\omega_1 T_0) + M_2 \exp(3i\omega_1 T_0) + M_3 \exp(i\omega_2 T_0) + M_4 \exp(2i\omega_2 T_0) \\ & + M_5 \exp(3i\omega_2 T_0) + M_6 \exp(i(\omega_1 + \omega_2)T_0) + M_7 \exp(i(\omega_1 - \omega_2)T_0) \\ & + M_8 \exp(i(2\omega_1 + \omega_2)T_0) + M_9 \exp(i(2\omega_1 - \omega_2)T_0) + M_{10} \exp(i(\omega_1 + 2\omega_2)T_0) \\ & + M_{11} \exp(i(\omega_1 - 2\omega_2)T_0) + M_{12} + cc. \end{aligned} \quad (23)$$

Here, M_i ($i = 1, 2, \dots, 12$) is presented in Appendix A.

Substituting Equations (20) and (23) into Equation (19b), we have

$$\begin{aligned}
 (D_0^2 + \omega_2^2) \varphi_{21} = & \left[-2i\omega_1 D_1 \Gamma A_1 + \alpha_4 \left(3(1-\Gamma)^3 A_1^2 \bar{A}_1 + 6(1-\Gamma) A_1 A_2 \bar{A}_2 \right) \right. \\
 & - \left(K_p - i \frac{K_i}{\omega_1} + i\omega_1 K_d \right) \Gamma A_1 \left. \right] \exp(i\omega_1 T_0) + \left[\omega_2^2 M_1 + \alpha_3 (1-\Gamma)^2 A_1^2 \right] \exp(2i\omega_1 T_0) \\
 & + \left[\omega_2^2 M_2 + \alpha_4 (1-\Gamma)^3 A_1^3 \right] \exp(3i\omega_1 T_0) + \left[\omega_2^2 M_3 - 2i\omega_2 D_1 A_2 \right. \\
 & - \left(K_p - i \frac{K_i}{\omega_2} + i\omega_2 K_d \right) A_2 - \alpha_4 \left(3A_2^2 \bar{A}_2 + 6(1-\Gamma)^2 A_1 \bar{A}_1 A_2 \right) \left. \right] \exp(i\omega_2 T_0) \\
 & + \left[\omega_2^2 M_4 + \alpha_3 A_2^2 \right] \exp(2i\omega_2 T_0) + \left[\omega_2^2 M_5 - \alpha_4 A_2^3 \right] \exp(3i\omega_2 T_0) \\
 & + \left[\omega_2^2 M_6 - 2\alpha_3 (1-\Gamma) A_1 A_2 \right] \exp(i(\omega_1 + \omega_2) T_0) + \left[\omega_2^2 M_7 - 2\alpha_3 (1-\Gamma) A_1 \bar{A}_2 \right] \\
 & \times \exp(i(\omega_1 - \omega_2) T_0) + \left[\omega_2^2 M_8 - 3\alpha_4 (1-\Gamma)^2 A_1^2 A_2 \right] \exp(i(2\omega_1 + \omega_2) T_0) \\
 & + \left[\omega_2^2 M_9 - 3\alpha_4 (1-\Gamma)^2 A_1^2 \bar{A}_2 \right] \exp(i(2\omega_1 - \omega_2) T_0) + \left[\omega_2^2 M_{10} + 3\alpha_4 (1-\Gamma) A_1 A_2^2 \right] \\
 & \times \exp(i(\omega_1 + 2\omega_2) T_0) + \left[\omega_2^2 M_{11} + 3\alpha_4 (1-\Gamma) A_1 \bar{A}_2^2 \right] \exp(i(\omega_1 - 2\omega_2) T_0) \\
 & + \left[\omega_2^2 M_{12} + \alpha_3 \left((1-\Gamma)^2 A_1 \bar{A}_1 + A_2 \bar{A}_2 \right) - K_i \left(i \frac{A_2}{\omega_2} + i \frac{\Gamma A_1}{\omega_1} \right) \right] + \left[\frac{-if_2}{2} \right] \exp(i\Omega_2 T_0) + cc.
 \end{aligned} \tag{24}$$

To obtain the second solvability condition of Equation (24), the closeness term of excitation frequencies (Ω_2) and natural frequency (ω_2) is described via employing the dimensionless detuning parameter σ_2 as the following:

$$\Omega_2 = \omega_2 + \varepsilon \sigma_2. \tag{25}$$

By substituting Equation (25) into Equation (24) and canceling the term which leads to the secular ones, we obtain the particular solution of the first approximation (φ_{21}) as the following:

$$\begin{aligned}
 \varphi_{21} = & N_1 \exp(i\omega_1 T_0) + N_2 \exp(2i\omega_1 T_0) + N_3 \exp(3i\omega_1 T_0) + N_4 \exp(2i\omega_2 T_0) \\
 & + N_5 \exp(3i\omega_2 T_0) + N_6 \exp(i(\omega_1 + \omega_2) T_0) + N_7 \exp(i(\omega_1 - \omega_2) T_0) \\
 & + N_8 \exp(i(2\omega_1 + \omega_2) T_0) + N_9 \exp(i(2\omega_1 - \omega_2) T_0) + N_{10} \exp(i(\omega_1 + 2\omega_2) T_0) \\
 & + N_{11} \exp(i(\omega_1 - 2\omega_2) T_0) + N_{12} + cc.
 \end{aligned} \tag{26}$$

Here, N_i ($i = 1, 2, \dots, 12$) is presented in Appendix A.

After inserting Equations (22) and (25) into Equations (21) and (24) and deleting the secular term, the conditions of solvability can be gained as follows:

$$2i\omega_1 D_1 A_1 = \left[\begin{aligned} & \left(-i\omega_1 \zeta + \Gamma \beta - K_p + i \frac{K_i}{\omega_1} - K_d i\omega_1 \right) A_1 + (-6\alpha_2 (1-\Gamma)) A_1 A_2 \bar{A}_2 \\ & + (-3\alpha_2 (1-\Gamma)^3) A_1^2 \bar{A}_1 \end{aligned} \right] + \left[\frac{-if_1}{2} \right] \exp(i\sigma_1 T_1) \tag{27a}$$

$$2i\omega_2 D_1 A_2 = \left[\begin{aligned} & \left(-\Gamma \beta - K_p + i \frac{K_i}{\omega_2} - i\omega_2 K_d \right) A_2 + (-3\Gamma \alpha_2 - 3\alpha_4) A_2^2 \bar{A}_2 \\ & + (-6\Gamma \alpha_2 - 6\alpha_4) (1-\Gamma)^2 A_1 \bar{A}_1 A_2 \end{aligned} \right] + \left[\frac{-if_2}{2} \right] \exp(i\sigma_2 T_1) \tag{27b}$$

To obtain the amplitude-phase equations of the controlled system, we analyze the solution of Equation (27), exchanging $A_n(T_1)$ by the polar form as

$$A_n = \frac{1}{2} a_n(T_1) e^{i\gamma_n(T_1)}, \quad (n = 1, 2). \tag{28}$$

We obtain the governing equations of the amplitudes a_n and the phases γ_n by substituting from Equation (28) into Equation (27) and then separating the real and imaginary components.

$$\dot{a}_1 = \left(-\frac{\zeta}{2} + \frac{K_i}{2\omega_1^2} - \frac{K_d}{2} \right) a_1 + \left[\frac{-f_1}{2\omega_1} \right] \cos(\psi_1), \tag{29a}$$

$$a_1 \dot{\gamma}_1 = \left[\left(-\frac{\Gamma\beta}{2\omega_1} + \frac{K_p}{2\omega_1} \right) a_1 + \left(\frac{3\alpha_2(1-\Gamma)}{4\omega_1} \right) a_1 a_2^2 + \left(\frac{3\alpha_2(1-\Gamma)^3}{8\omega_1} \right) a_1^3 \right] + \left[\frac{-f_1}{2\omega_1} \right] \sin(\psi_1), \quad (29b)$$

$$\dot{a}_2 = \left(\frac{K_i}{2\omega_2^2} - \frac{K_d}{2} \right) a_2 + \left[\frac{-f_2}{2\omega_2} \right] \cos(\psi_2), \quad (30a)$$

$$a_2 \dot{\gamma}_2 = \left[\left(\frac{\Gamma\beta}{2\omega_2} + \frac{K_p}{2\omega_2} \right) a_2 + \left(\frac{3\Gamma\alpha_2+3\alpha_4}{8\omega_2} \right) a_2^3 + \left(\frac{3\Gamma\alpha_2+3\alpha_4}{4\omega_2} \right) (1-\Gamma)^2 a_1^2 a_2 \right] + \left[\frac{-f_2}{2\omega_2} \right] \sin(\psi_2) \quad (30b)$$

where

$$\psi_1 = \sigma_1 T_1 - \gamma_1, \psi_2 = \sigma_2 T_1 - \gamma_2. \quad (31)$$

The fixed points in Equations (29)–(30) correspond to the steady-state solutions of the system, which in turn correspond to $\dot{a}_n = 0$ and $\dot{\psi}_n = 0$.

From Equation (31), we indicate that $\dot{\gamma}_1 = \sigma_1$ and $\dot{\gamma}_2 = \sigma_2$.

As a result, $\dot{a}_n = 0$ and $\dot{\psi}_n = 0$; the practical case's frequency response equations (FRE) ($a_1 \neq 0, a_2 \neq 0$) are provided as

$$0 = \left(-\frac{\zeta}{2} + \frac{K_i}{2\omega_1^2} - \frac{K_d}{2} \right) a_1 + \left[\frac{-f_1}{2\omega_1} \right] \cos(\psi_1), \quad (32a)$$

$$a_1 \sigma_1 = \left[\left(-\frac{\Gamma\beta}{2\omega_1} + \frac{K_p}{2\omega_1} \right) a_1 + \left(\frac{3\alpha_2(1-\Gamma)}{4\omega_1} \right) a_1 a_2^2 + \left(\frac{3\alpha_2(1-\Gamma)^3}{8\omega_1} \right) a_1^3 \right] + \left[\frac{-f_1}{2\omega_1} \right] \sin(\psi_1), \quad (32b)$$

$$0 = \left(\frac{K_i}{2\omega_2^2} - \frac{K_d}{2} \right) a_2 + \left[\frac{-f_2}{2\omega_2} \right] \cos(\psi_2), \quad (33a)$$

$$a_2 \sigma_2 = \left[\left(\frac{\Gamma\beta}{2\omega_2} + \frac{K_p}{2\omega_2} \right) a_2 + \left(\frac{3\Gamma\alpha_2+3\alpha_4}{8\omega_2} \right) a_2^3 + \left(\frac{3\Gamma\alpha_2+3\alpha_4}{4\omega_2} \right) (1-\Gamma)^2 a_1^2 a_2 \right] + \left[\frac{-f_2}{2\omega_2} \right] \sin(\psi_2) \quad (33b)$$

3.2. Stability Analysis via Linearizing the above System

In order to examine the stability of the given fixed points nonlinear solution, let

$$\left. \begin{aligned} a_m &= a_{m0} + a_{m1}, \\ \psi_m &= \psi_{m0} + \psi_{m1} \end{aligned} \right\} \quad (34)$$

where a_{m1} and ψ_{m1} are perturbations that are thought to be tiny in comparison to a_{m0} and ψ_{m0} . Here, a_{m0} and ψ_{m0} are the solutions of Equations (29) and (30). Equation (34) is substituted into Equations (29) and (30) with Equation (31), retaining only the linear terms in a_{m1} and ψ_{m1} . Therefore, the linearized system of Equations (29) and (30) has the forms

$$\dot{a}_{11} = \left[-\frac{\zeta}{2} + \frac{K_i}{2\omega_1^2} - \frac{K_d}{2} \right] a_{11} + \left[\frac{f_1}{2\omega_1} \sin(\psi_{10}) \right] \psi_{11}, \quad (35a)$$

$$\dot{\psi}_{11} = \left[\left(\frac{\sigma_1}{a_{10}} + \frac{\Gamma\beta}{2a_{10}\omega_1} - \frac{K_p}{2a_{10}\omega_1} \right) + \left(-\frac{3\alpha_2(1-\Gamma)}{4a_{10}\omega_1} \right) a_{20}^2 + \left(-\frac{9\alpha_2(1-\Gamma)^3}{8\omega_1} \right) a_{10} \right] a_{11} + \left[\frac{f_1}{2a_{10}\omega_1} \cos(\psi_{10}) \right] \psi_{11} + \left[\left(-\frac{3\alpha_2(1-\Gamma)}{2\omega_1} \right) a_{20} \right] a_{21} \quad (35b)$$

$$\dot{a}_{21} = \left[\left(\frac{K_i}{2\omega_2^2} - \frac{K_d}{2} \right) \right] a_{21} + \left[\frac{f_2}{2\omega_2} \sin(\psi_{20}) \right] \psi_{21}, \quad (36a)$$

$$\dot{\psi}_{21} = \left[-\left(\frac{3\Gamma\alpha_2 + 3\alpha_4}{2\omega_2} \right) (1 - \Gamma)^2 a_{10} \right] a_{11} + \left[\left(\frac{\sigma_2}{a_{20}} - \frac{\Gamma\beta}{2a_{20}\omega_2} - \frac{K_p}{2a_{20}\omega_2} \right) - \left(\frac{9\Gamma\alpha_2 + 9\alpha_4}{8\omega_2} \right) a_{20} - \left(\frac{3\Gamma\alpha_2 + 3\alpha_4}{4a_{20}\omega_2} \right) (1 - \Gamma)^2 a_{10}^2 \right] a_{21} + \left[\frac{f_2}{2\omega_2 a_{20}} \cos(\psi_{20}) \right] \psi_{21} \quad (36b)$$

The above Equations (35) and (36) can be described as the following matrix:

$$\begin{bmatrix} \dot{a}_{11} & \dot{\psi}_{11} & \dot{a}_{21} & \dot{\psi}_{21} \end{bmatrix}^T = [J] \begin{bmatrix} a_{11} & \psi_{11} & a_{21} & \psi_{21} \end{bmatrix}^T \quad (37)$$

Here $[J]$ is represented by the appropriate portions of Equations (35) and (36).

The eigenvalues of $[J]$ ascertained using the subsequent equation are as follows:

$$\lambda^4 + \Gamma_1 \lambda^3 + \Gamma_2 \lambda^2 + \Gamma_3 \lambda + \Gamma_4 = 0. \quad (38)$$

The coefficients of Equation (38) are denoted as Γ_m ($m = 1, 2, \dots, 4$). The solutions of the system with PID control are stable if the roots' real parts of λ have negative values; if not, they are unstable. It is a necessary and sufficient requirement for a steady-state solution to use the Routh–Hurwitz criterion, which states that all of the roots of Equation (38) must have negative real parts if and only if all of the principal minors and the following determinant D are positive.

$$D = \begin{vmatrix} \Gamma_1 & 1 & 0 & 0 \\ \Gamma_3 & \Gamma_2 & \Gamma_1 & 1 \\ 0 & \Gamma_4 & \Gamma_3 & \Gamma_2 \\ 0 & 0 & 0 & \Gamma_4 \end{vmatrix}. \quad (39)$$

4. Results and Discussion

4.1. Time History Performance without Control

To analyze the behavior of the dynamical system the fourth-order Runge–Kutta algorithm (ode45 in MATLAB) [30] is applied to find the numerical solution of the given uncontrolled system of Equations (8) and (9).

The time history performance of the system without any controller at the primary resonance ($\Omega_1 \cong \omega_1, \Omega_2 \cong \omega_2$) is shown in Figure 3 at the chosen values ($\omega_1 = 3; \zeta = 0.2; \omega_2 = 2; \beta = 0.5\omega_2^2; \alpha_1 = 0.2; \alpha_2 = 0.3; f_1 = 10; \alpha_3 = 2\alpha_1; \alpha_4 = 2\alpha_2; f_2 = 5; \Omega_1 = \omega_1; \Omega_2 = \omega_2$).

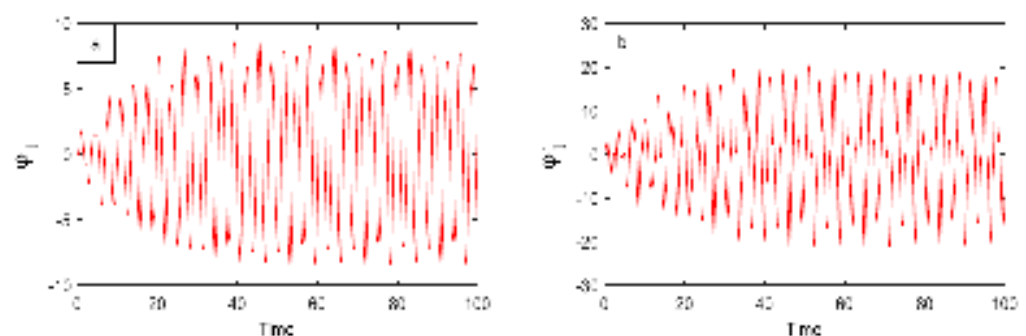


Figure 3. The time diagram without any controller (a) the first part of the system, ϕ_1 (b) the response of its velocity, $\dot{\phi}_1$.

Figure 3a,b show the steady-state time response against the amplitude ϕ_1 for the uncontrolled first system and the velocity $\dot{\phi}_1$, respectively. Similarly, Figure 4a,b show the uncontrolled second main system and its amplitude ϕ_2 and velocity $\dot{\phi}_2$ with the time response.

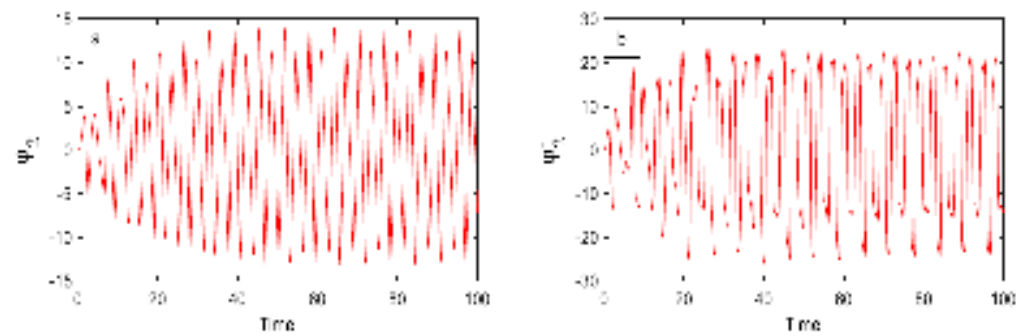


Figure 4. The time diagram without any controller (a) the second part of the system, φ_2 (b) response of its velocity, $\dot{\varphi}_2$.

4.2. Time History Performance with Different Control

The closed-loop performances of the P, PI, and PID controllers for the two systems φ_1 and φ_2 , along with their corresponding velocities $\dot{\varphi}_1$ and $\dot{\varphi}_2$, are displayed in Figure 5. The closed-loop responses of the traditional P and PI controllers are unstable and exhibit peak overshoot. Better closed-loop performance, such as in responses with less peak overshoot and settling time, was supplied by the redesigned PID controller structure, which conveys closed-loop responses with less peak and overshoot and a more stable curve than the others. The study was for the first main system and its velocity, as in Figure 5a,b, and the second main system shown in Figure 5c,d. Figure 5 shows evidence of both steady behavior and the absence of chaos in the produced wavelengths and amplitudes in every section. We utilized the gain values that are depicted in Table 1 to obtain these closed loops using the three control approaches.

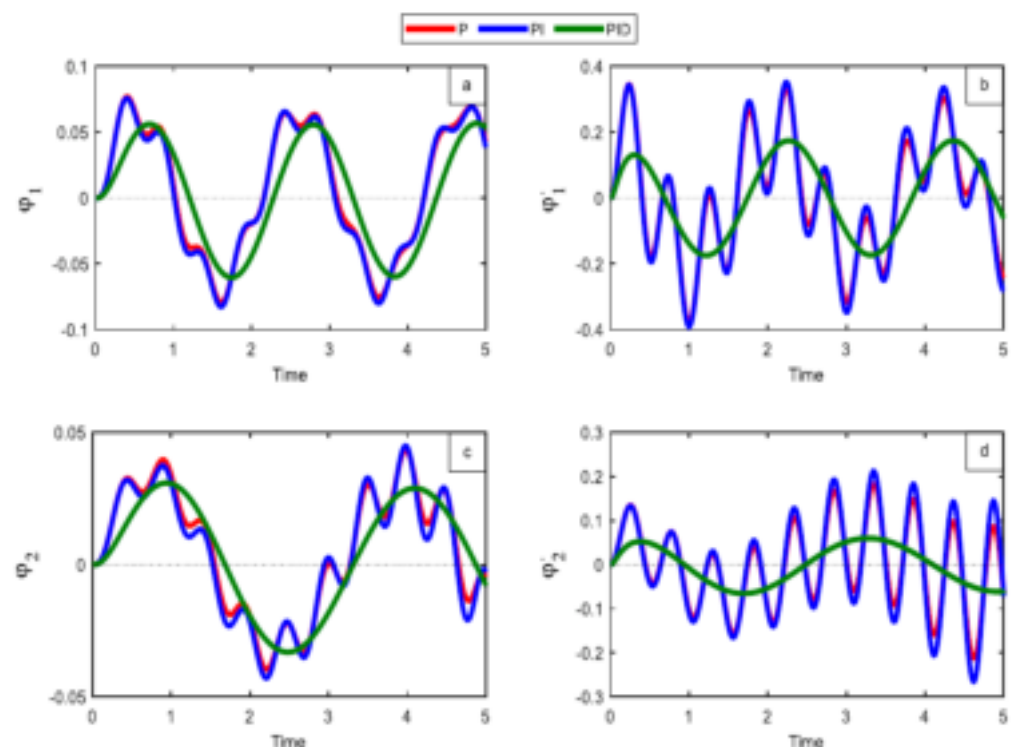


Figure 5. Responses in closed loop using P, PI, and PID controller. (a) the first part of the system φ_1 (b) the response of its velocity $\dot{\varphi}_1$ (c) the second part of the system φ_2 (d) the response of its velocity $\dot{\varphi}_2$.

Table 1. Parameters of a PID controller.

Controller		Gain		
		K_p	K_i	K_d
P	(Red)	150	-	-
PI	(Blue)	150	20	-
PID	(Green)	150	20	20

In Figure 6a,b, there is a depiction of the first main system's amplitude decreasing from 7.788 to 0.0569 after the utilization of a time PID controller. This indicates that the controller's effectiveness (E_a = amplitude without control/amplitude with) was equivalent to 136.87 for the first main system φ_1 , with a proportional reduction of 99.27%. Likewise, the amplitude φ_2 of the second system, which is equal to 13.59, as shown in Figure 7, decreased to 0.02886, where $E_a = 470.89$, reflecting a proportional reduction of 99.79%.

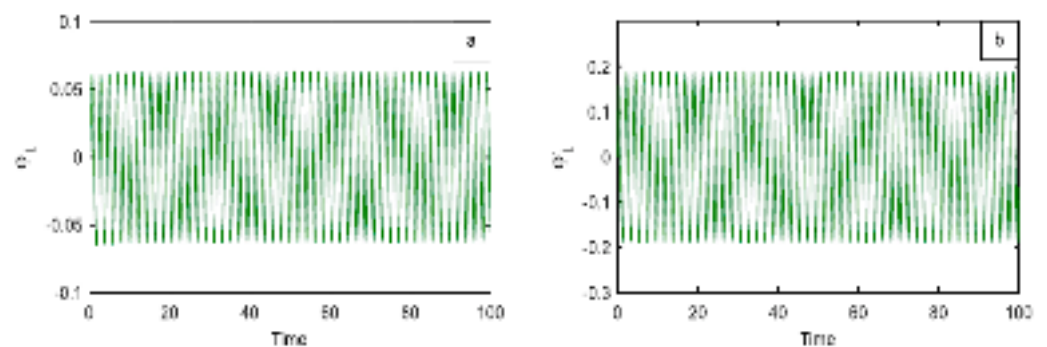
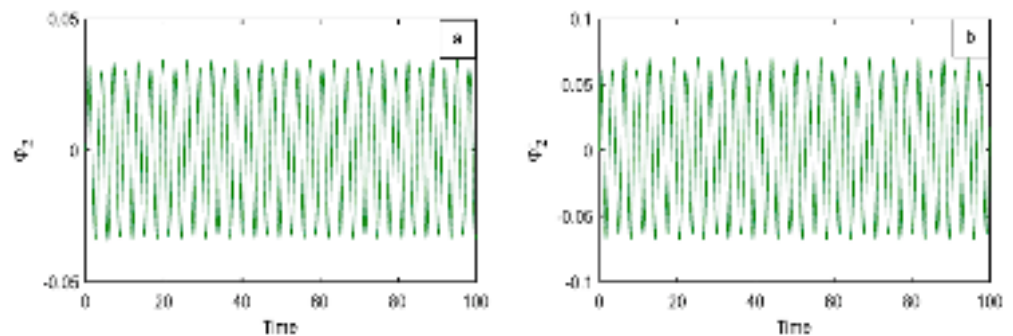
**Figure 6.** The time diagram in the primary resonance situation with the PID controller (a) the first part of the system φ_1 (b) the response of its velocity $\dot{\varphi}_1$.**Figure 7.** The time diagram in the primary resonance situation with the PID controller (a) the second part of the system φ_2 (b) the response of its velocity $\dot{\varphi}_2$.

Figure 8 illustrates the phase plane of both main systems before and after adding the PID controller. Figure 8a shows that the first main system φ_1 , with a multi-limit cycle in the case of not adding control, and as improved with a limited cycle after adding PID. With the second main system φ_2 , as shown in Figure 8b before adding the control, the phase plane is the multi-limit cycle, which decreased after adding PID.

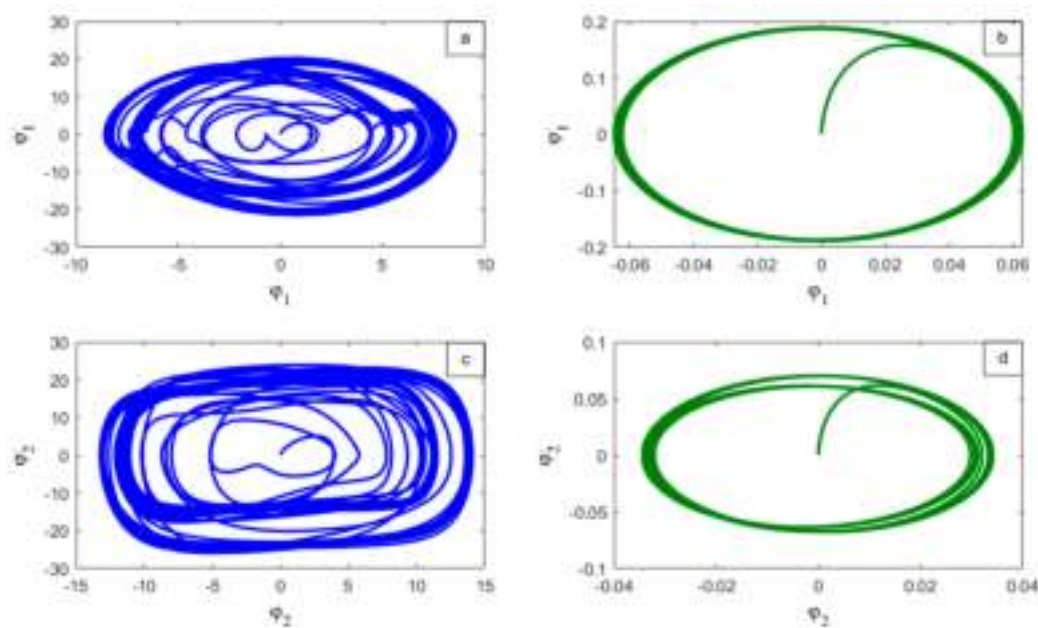


Figure 8. The phase plane of the system's (a) ϕ_1 with angular velocities $\dot{\phi}_1$ in the primary resonance situation without control (b) ϕ_1 with angular velocities $\dot{\phi}_1$ in the primary resonance situation with the PID controller (c) ϕ_2 with angular velocities $\dot{\phi}_2$ in the primary resonance situation without control (d) ϕ_2 with angular velocities $\dot{\phi}_2$ in the primary resonance situation with the PID controller.

4.3. Frequency Response Curves (FRC)

The frequency response curves of Equations (32) and (33) of the first main system a_1 , against the detuning parameter σ_1 , are illustrated in Figure 9a without the influence of any controller, where the solid black line represents the stable solution while the red line indicates the unstable solution of the same equation. The amplitude of the first main system is fully reduced after the P controller is added; however, as Figure 9b shows, this declining portion of the curve is unstable. The response curve in Figure 9c shows what happens when a PI controller is added. It is discovered that this fully reduces the amplitude of the primary system a_1 , yet the curve is absolutely unstable. After adding a PID controller, the first main system's behavior is shown in Figure 9d, where it is discovered that the curve has a monotonically decreasing amplitude and is completely stable. The response curves of the second main system a_2 , against the detuning parameter σ_2 , are illustrated in Figure 10a without the influence of any controller. The amplitude of the second main system is fully reduced after the P controller is added; however, as Figure 10b shows, this declining portion of the curve is unstable. The response curve in Figure 10c shows what happens when a PI controller is added. It is discovered that this controller reduces the amplitude of the second system a_2 , yet the curve behaves in an unstable way. After adding a PID controller, the second main system's behavior is shown in Figure 10d, where it is discovered that the curve has a monotonically decreasing amplitude and is completely stable. Figure 11 clearly illustrates the difference in the curves' responses before and after the PID controller was included for the two primary systems, a_1 and a_2 . Figure 11a,b demonstrated how the system is only stable and does not contain any unstable parts, as well as how the addition of a PID controller causes the amplitude to drop with a_1 and disappear with a_2 . The amplitude rose for increasing values of the external force f_1 for the first main system a_1 , as Figure 12a illustrates, with the first component of the system, a_1 . The amplitude of the first main system is shifted to the right and displays a monotonically declining curve as the values of natural frequency ω_1 increase, as seen in Figure 12b. The amplitude decreases monotonically when the damping coefficient ζ values increase, as shown in Figure 12c. As the values of the nonlinear parameter β increased, the amplitude of the first half of the system bent to the left, as illustrated in Figure 12d. Lastly, Figure 12e illustrates how the nonlinear parameter α_2

behaves for both positive and negative parameter values. For very tiny negative values of α_2 , the curve bends left, but for high positive values of α_2 , the curve bends right. The second main system's amplitude increased for large values of the external force f_2 , as shown in Figure 13a, with the second portion of the system, a_2 . As seen in Figure 13b, the second main system's amplitude is moved to the right and exhibits a monotonic declining curve for increased values of the natural frequency ω_2 . As seen in Figure 13c, the amplitude remains constant and shifts to the right for extraordinarily high positive values of the nonlinear parameter α_4 and to the left for extraordinarily low negative values of α_4 . The response of the gains with the amplitude is examined in Figure 14. Figure 14a establishes the values of K_i and K_d . Upon further examination of the influence of K_p , we discovered that when K_p values increase, the amplitude of both primary systems decreases while remaining steady. With different values of K_i , we observed that the amplitude of both main systems increased indistinguishably for the fixed values of K_p and K_d . However, when $K_i = 80$, both main systems behaved in an unstable manner, as shown in Figure 14b. Lastly, although the other two gains are fixed, Figure 14c shows the improvement of K_d . Painting made it evident that the amplitude is unstable at the initial K_d values and that, in both systems, the amplitude decreases as K_d increases.

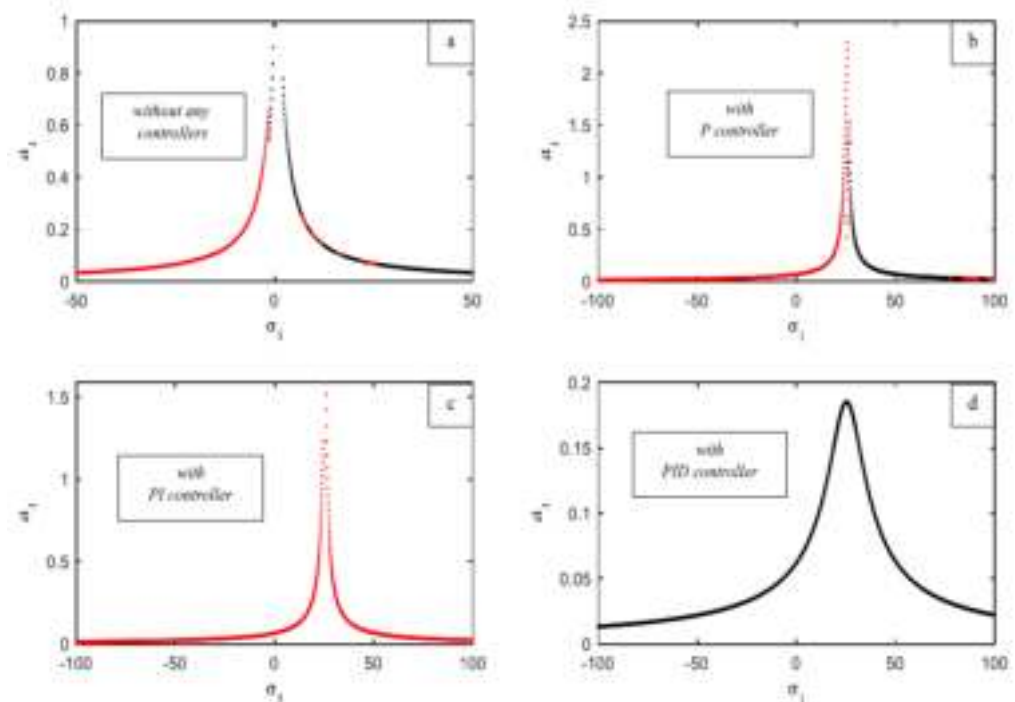


Figure 9. The frequency response curves on the plane σ_1 and the first part of the system (a_1) (a) without a controller, (b) with a P controller, (c) with a PI controller, and (d) with a PID controller. (red color) unstable region (black color) stable region

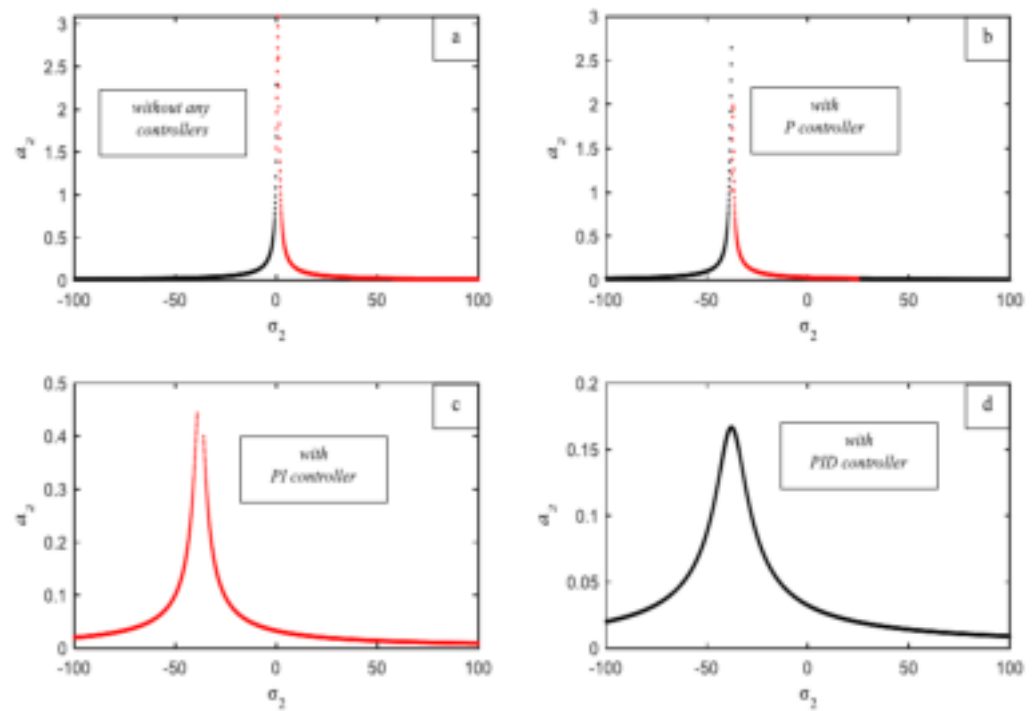


Figure 10. The frequency response curves on plane σ_2 and the first part of the system (a_2) (a) without a controller, (b) with a P controller, (c) with a PI controller, and (d) with PID controller. (red color) unstable region (black color) stable region.

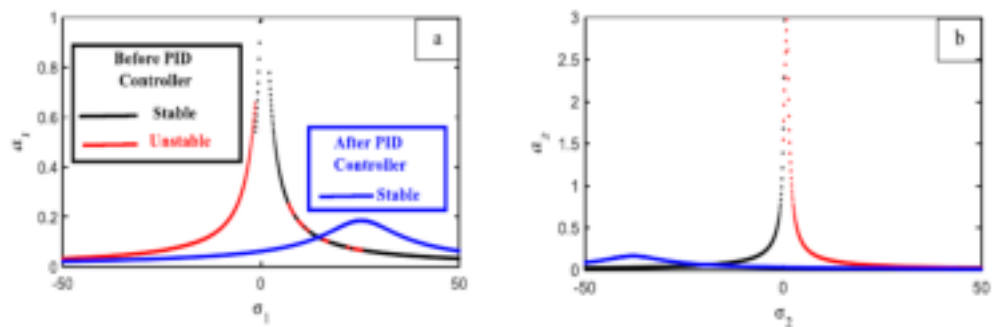


Figure 11. Frequency response before and after PID: (a) first main system a_1 and (b) first main system a_2 .

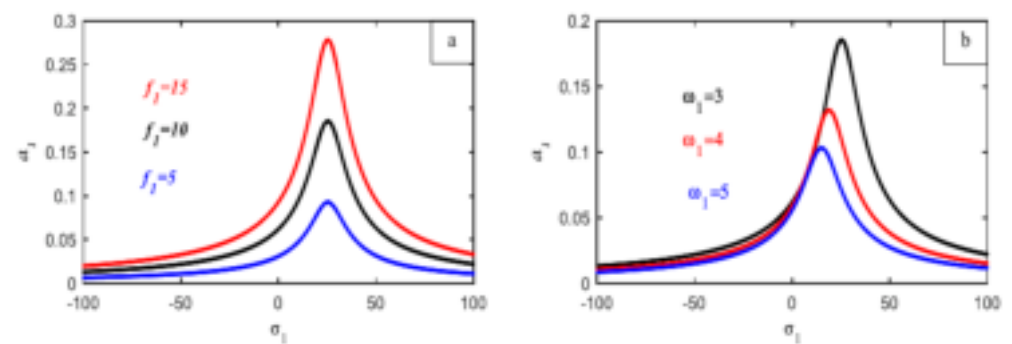


Figure 12. Cont.

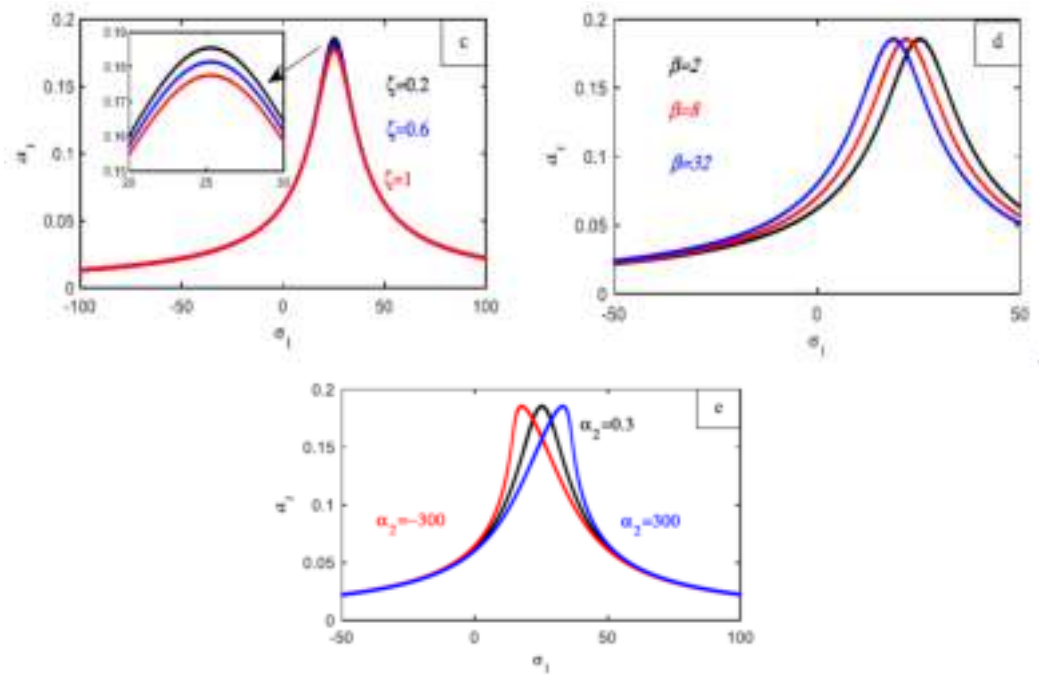


Figure 12. (a) The FRC of the external force action f_1 . (b) Natural frequency ω_1 . (c) Damping coefficient ζ . (d) Linear parameter β . (e) Nonlinear parameter α_2 . All of these regard the first part of the system a_1 .

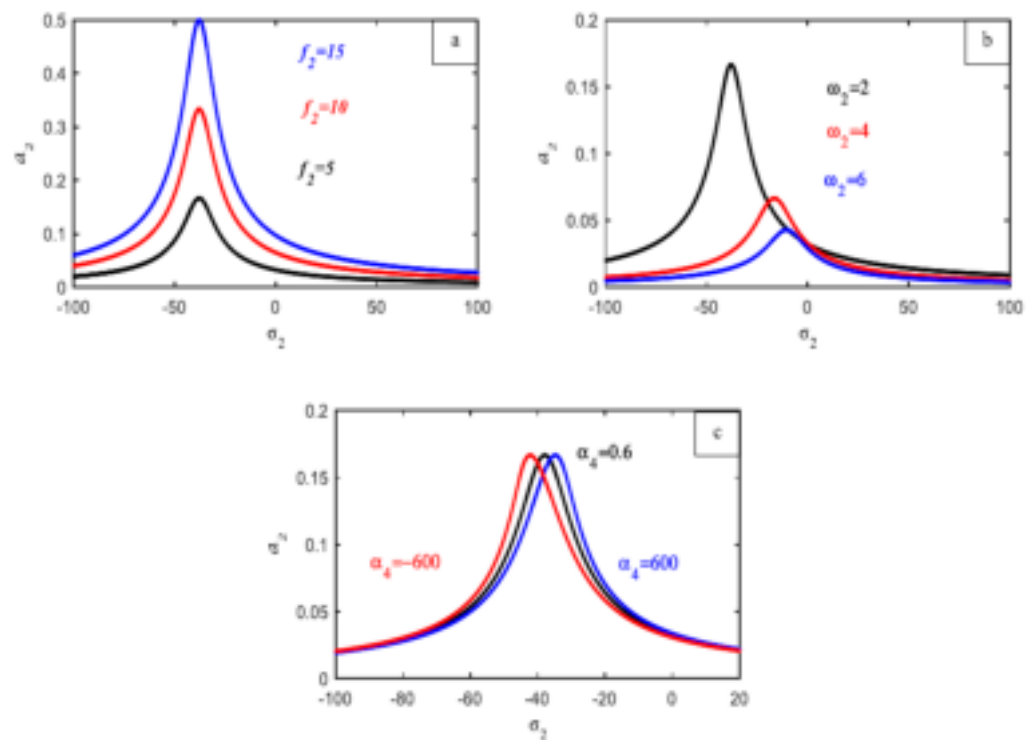


Figure 13. (a) The FRC of the external force action f_2 . (b) Natural frequency ω_2 . (c) Nonlinear parameter α_4 . All of these regard the second part of system a_2 .

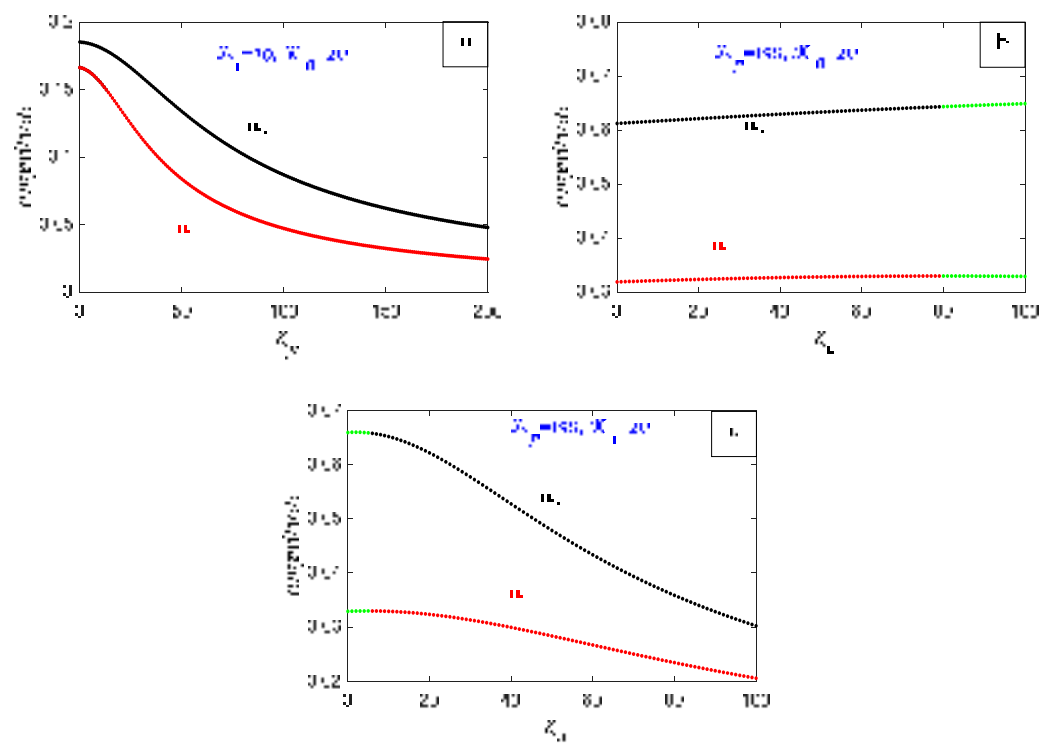


Figure 14. Response curve for the system amplitudes with (a) K_p (b) K_i (c) K_d .

5. Comparison

5.1. Comparison of the Perturbation's Temporal Response Solutions with Numerical Techniques

Analytical solutions to Equations (29) and (30) are graphically represented by (---) lines, which correspond with the numerical solutions of Equations (13) and (14), as displayed in Figure 15a for the first main system, and in Figure 15b for the second main system.

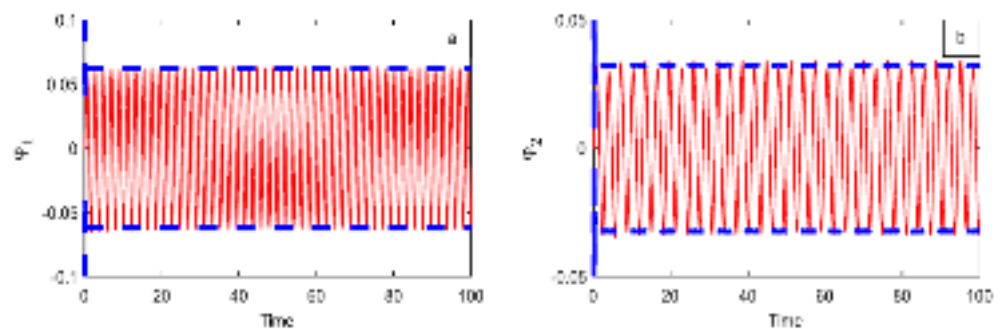


Figure 15. Comparison of numerical simulation and perturbation analysis for both framework modes in PID controllers. (a) the first part of the system φ_1 (b) the second part of the system φ_2

5.2. Comparison between RK-4 and FRC

Figure 16 shows an acceptable agreement with the numerical solutions of Equations (4) and (5), using (RK-4) highlighted by yellow circles and the frequency response curves (FRC) for the two main systems a_1 in Figure 16a and a_2 in Figure 16b, which are presented with the solid line.

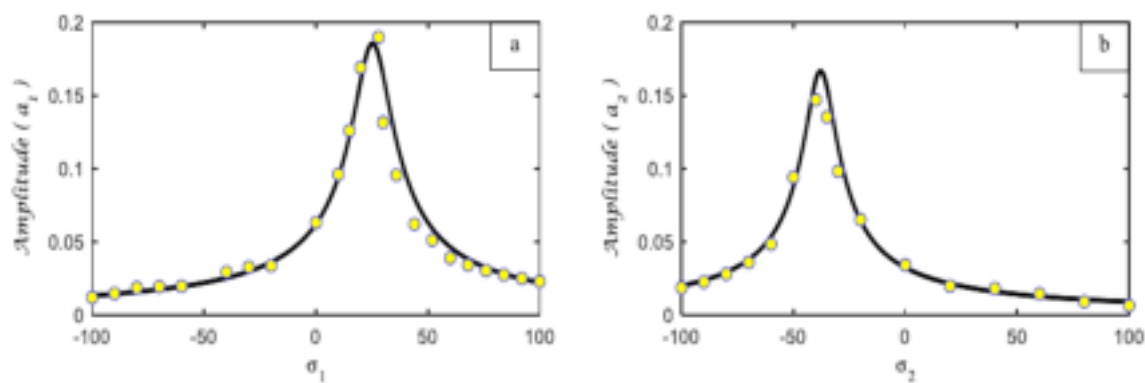


Figure 16. Contrast between the FRC solution and RK-4 solution for both systems (a) the first part of the system a_1 (b) the second part of the system a_2 .

5.3. Comparison with Published Work

For previous work, the following can be stated:

- (1) Torsional vibration was examined in a review of the literature [25] for active and passive control, which is a useful tool for managing the torsional vibration of a nonlinear dynamical system that is exposed to several parametric excitations.
- (2) They used the frequency response equations; the stability of the system is examined in the vicinity of the simultaneous sub-harmonic and internal resonances.
- (3) The behavior of the system with and without two controllers is numerically integrated and examined.

For this work, after adding control, we compared the time histories of our work and a prior study and discovered that the amplitudes and wavelengths showed that the solutions generated exhibited stable behavior and were devoid of chaos.

6. Conclusions

The torsional vibration control of a nonlinear dynamical system has been tackled within this article. Proportional-integral-derivative (PID) controllers have been proposed to control the harmful torsional vibration of the system, as shown in Figure 2 for one of the worst resonance cases. The multiple-scale perturbation approach is applied to obtain the approximation solution for the coupled controlled system. The time history is drawn to study the steady-state vibration amplitudes and the effectiveness of the applied control algorithms in suppressing vibration. In addition, the stability and effects of different system and control parameters are illustrated in frequency response curves according to the Routh–Hurwitz criterion. Based on the above discussion, the following conclusions can be drawn:

- (1) It is noted that the PID controller is operated to reduce the dangerous vibrations in a short time.
- (2) This work provided a comparison between P, PI, and PID controllers in the context of FRCs.
- (3) Numerous coefficients' effects are examined and shown numerically.
- (4) Despite the excitation frequency, the PID controller is the most effective control method for minimizing the vibrations in the framework.
- (5) The closed-loop response of relative displacement is obtained with the PID controller, which comprises the peak-overshoot.
- (6) The modified structure of the PID controller, such as with the P and PI controllers, is used for the control of the relative displacement of the suspension system. From the results, the PID controller provided better closed-loop performance in terms of peak overshoot and settling time, which are minimized.

Author Contributions: K.A.: resources; methodology; formal analysis; validation; visualization; writing—original draft; writing—review and editing (equal); funding acquisition. A.T.E.-S.: conceptualization; investigation; methodology; data curation; validation; writing—original draft; writing—review and editing. F.T.E.-B.: formal analysis; validation; investigation; methodology; data curation; conceptualization; writing—original draft; writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Researchers Supporting Project, grant number (RSPD2024R588), King Saud University, Riyadh, Saudi Arabia.

Data Availability Statement: All data generated or analyzed during this study are included in this published article.

Acknowledgments: The authors extend their appreciation to the Researchers Supporting Project, grant number (RSPD2024R588), King Saud University, Riyadh, Saudi Arabia.

Conflicts of Interest: The authors have not revealed any conflicting interests.

Nomenclature

I_1, I_2	Polar mass moment of the nonlinear dynamical system
k_1, k_2	Linear spring stiffness of the nonlinear dynamical system
k_{21}, k_{22}, k_{23}	Quadratic and cubic stiffness parts
F_1^*, F_2^*	Excitation forces
c_1	Linear damping coefficients of the nonlinear dynamical system
ω_1, ω_2	Natural frequencies
ζ	Damping coefficient
β	Linear parameter
$\alpha_i (i = 1, 2, 3, 4)$	Nonlinear parameter
F_{C1}, F_{C2}	Forced control
K_p	Proportional gain
K_d	Derivative gain
K_i	Integral gain
θ_0	Reference angle

Appendix A

$$M_1 = \left[\frac{\alpha_1(1-\Gamma)^2 A_1^2}{3\omega_1^2} \right], M_2 = \left[\frac{\alpha_2(1-\Gamma)^3 A_1^3}{8\omega_1^2} \right], M_3 = \left[\frac{\beta A_2 + \alpha_2(3A_2^2 \bar{A}_2 + 6(1-\Gamma)^2 A_1 \bar{A}_1 A_2)}{\omega_1^2 - \omega_2^2} \right]$$

$$M_4 = \left[\frac{-\alpha_1 A_1^2}{\omega_1^2 - 4\omega_2^2} \right], M_5 = \left[\frac{\alpha_2 A_2^3}{\omega_1^2 - 9\omega_2^2} \right], M_6 = \left[\frac{2\alpha_1(1-\Gamma) A_1 A_2}{\omega_1^2 - (\omega_1 + \omega_2)^2} \right], M_7 = \left[\frac{2\alpha_1(1-\Gamma) A_1 \bar{A}_2}{\omega_1^2 - (\omega_1 - \omega_2)^2} \right]$$

$$M_8 = \left[\frac{3\alpha_2(1-\Gamma)^2 A_1^2 A_2}{\omega_1^2 - (2\omega_1 + \omega_2)^2} \right], M_9 = \left[\frac{3\alpha_2(1-\Gamma)^2 A_1^2 \bar{A}_2}{\omega_1^2 - (2\omega_1 - \omega_2)^2} \right], M_{10} = \left[\frac{-3\alpha_2(1-\Gamma) A_1 A_2^2}{\omega_1^2 - (\omega_1 + 2\omega_2)^2} \right]$$

$$M_{11} = \left[\frac{-3\alpha_2(1-\Gamma) A_1 \bar{A}_2^2}{\omega_1^2 - (\omega_1 - 2\omega_2)^2} \right], M_{12} = \left[\frac{-\alpha_1 \left((1-\Gamma)^2 A_1 \bar{A}_1 + A_2 \bar{A}_2 \right) - iK_i \frac{A_1}{\omega_1}}{\omega_1^2} \right]$$

$$N_1 = \frac{1}{\omega_2^2 - \omega_1^2} \left[-2i\omega_1 D_1 \Gamma A_1 + \alpha_4 \left(3(1-\Gamma)^3 A_1^2 \bar{A}_1 + 6(1-\Gamma) A_1 A_2 \bar{A}_2 \right) - \left(K_p - \frac{K_i}{\omega_1} + \omega_1 K_d \right) \Gamma A_1 \right]$$

$$N_2 = \left[\frac{\omega_2^2 M_1 + \alpha_3(1-\Gamma)^2 A_1^2}{\omega_2^2 - 4\omega_1^2} \right], N_3 = \left[\frac{\omega_2^2 M_2 + \alpha_4(1-\Gamma)^3 A_1^3}{\omega_2^2 - 9\omega_1^2} \right]$$

$$\begin{aligned}
N_4 &= \left[\frac{-\omega_2^2 M_4 - \alpha_3 A_2^2}{3\omega_2^2} \right], N_5 = \left[\frac{-\omega_2^2 M_5 + \alpha_4 A_2^3}{8\omega_2^2} \right], N_6 = \left[\frac{\omega_2^2 M_6 - 2\alpha_3(1-\Gamma)A_1 A_2}{\omega_2^2 - (\omega_1 + \omega_2)^2} \right] \\
N_7 &= \left[\frac{\omega_2^2 M_7 - 2\alpha_3(1-\Gamma)A_1 \bar{A}_2}{\omega_2^2 - (\omega_1 - \omega_2)^2} \right], N_8 = \left[\frac{\omega_2^2 M_8 - 3\alpha_4(1-\Gamma)^2 A_1^2 A_2}{\omega_2^2 - (2\omega_1 + \omega_2)^2} \right] \\
N_9 &= \left[\frac{\omega_2^2 M_9 - 3\alpha_4(1-\Gamma)^2 A_1^2 \bar{A}_2}{\omega_2^2 - (2\omega_1 - \omega_2)^2} \right], N_{10} = \left[\frac{\omega_2^2 M_{10} + 3\alpha_4(1-\Gamma)A_1 A_2^2}{\omega_2^2 - (\omega_1 + 2\omega_2)^2} \right] \\
N_{11} &= \left[\frac{\omega_2^2 M_{11} + 3\alpha_4(1-\Gamma)A_1 \bar{A}_2^2}{\omega_2^2 - (\omega_1 - 2\omega_2)^2} \right], N_{12} = \left[\frac{\omega_2^2 M_{12} + \alpha_3 \left((1-\Gamma)^2 A_1 \bar{A}_1 + A_2 \bar{A}_2 \right) - K_i \left(\frac{A_2}{\omega_2} + \frac{\Gamma A_1}{\omega_1} \right)}{\omega_2^2} \right]
\end{aligned}$$

References

1. Sowmya, V.S.; Dharsini, S.P.; Dharshini, R.P.; Aravind, P. Application of various PID controller tuning techniques for a temperature system. *Int. J. Comput. Appl.* **2014**, *103*, 32–34.
2. Pradeepkannan, D.; Sathiyamoorthy, S. Implementation of Gain Scheduled PID Controller for a Nonlinear Coupled Spherical Tank Process. *Int. J. Mech. Mechatron. Eng.* **2014**, *14*, 93–98.
3. Sabri, L.A.; Hussein, A.A. Implementation of fuzzy and PID controller to water level system using LabView. *Int. J. Comput. Appl.* **2015**, *116*, 6–10.
4. Bashiri, A.H. Empirical study of robust/developed PID control for nonlinear time-delayed dynamical system in discrete time domain. *Heliyon* **2024**, *10*, e29749. [[CrossRef](#)] [[PubMed](#)]
5. Hamed, A.; Shaban, E.M.; Darwish, R.R.; Abdel Ghany, A.M. Design and implementation of discrete PID control applied to Bitumen tank based on new approach of pole placement technique. *Int. J. Dynam. Control* **2017**, *5*, 604–613. [[CrossRef](#)]
6. Shaban, E.M.; Hamed, A.; Darwish, R.R.; Abdel Ghany, A.M. New tuning approach of discrete PI/PID controller applied to bitumen system based on non-minimal state space formulation. In Proceedings of the 25th International Conference on Computer Theory and Applications (ICCTA), Alexandria, Egypt, 24–26 October 2015; pp. 52–57.
7. Hamed, A.R.; Shaban, E.M.; Abdelhaleem, A.M. Industrial implementation of state dependent parameter PID+ control for nonlinear time delayed bitumen tank system. *Iran J. Sci. Technol. Trans. Electr. Eng.* **2022**, *46*, 743–751. [[CrossRef](#)]
8. Sayed, H.; Shaban, E.M.; Abdelhamid, A. State-Dependent Parameter PID+ Control Applied to A Nonlinear Manipulator Arm. In *Recent Advances in Engineering Mathematics and Physics*; Farouk, M., Hassanein, M., Eds.; Springer: Cham, Switzerland, 2020.
9. Dano, M.L.; Julli'ere, B. Active control of thermally induced distortion in composite structures using macro fiber composite actuators. *Smart Mater. Struct.* **2007**, *16*, 2315–2322. [[CrossRef](#)]
10. Kumar, R.S.; Ray, M.C. Active control of geometrically nonlinear vibrations of doubly curved smart sandwich shells using 1–3 piezoelectric composites. *Compos. Struct.* **2013**, *105*, 173–187. [[CrossRef](#)]
11. Saeed, N.; Kamel, M. Nonlinear PD-controller to suppress the nonlinear oscillations of horizontally supported Jeffcott-rotor system. *Int. Non-Linear Mech.* **2016**, *87*, 109–124. [[CrossRef](#)]
12. Fey, R.H.B.; Wouters RM, T.; Nijmeijer, H. Proportional and derivative control for steady-state vibration mitigation in a piecewise linear beam system. *Nonlinear Dyn.* **2010**, *60*, 535–649. [[CrossRef](#)]
13. Eissa, M.; Kandil, A.; El-Ganaini, W.A.; Kamel, M. Analysis of a nonlinear magnetic levitation system vibrations controlled by a time-delayed proportional derivative controller. *Nonlinear Dyn.* **2015**, *79*, 1217–1233. [[CrossRef](#)]
14. Eissa, M.; Kandil, A.; Kamel, M.; El-Ganaini, W.A. On controlling the response of primary and parametric resonances of a nonlinear magnetic levitation system. *Meccanica* **2015**, *50*, 233–251. [[CrossRef](#)]
15. Bauomy, H.S.; EL-Sayed, A.T. Act of nonlinear proportional derivative controller for MFC laminated shell. *Phys. Scr.* **2020**, *95*, 095210. [[CrossRef](#)]
16. Hote, Y.V.; Jain, S. PID controller design for load frequency control: Past, Present and future challenges. *IFAC-Pap. Online* **2018**, *51*, 604–609. [[CrossRef](#)]
17. Darwish, N.M. PID controller design in the frequency domain for time-delay systems using direct method. *Trans. Inst. Meas. Control* **2018**, *40*, 940–950. [[CrossRef](#)]
18. Zhao, C.; Guo, L. PID controller design for second order nonlinear uncertain systems. *Sci. China Inf. Sci.* **2017**, *60*, 022201. [[CrossRef](#)]
19. Hanafi, D. PID controller design for semi-active car suspension based on model from intelligent system identification. In Proceedings of the IEEE Computer Society, Second International Conference on Computer Engineering and Applications, Bali, Indonesia, 19–21 March 2010; pp. 60–63.

20. Constantin, M.; Popescu, O.S.; Mastorakis, N.E. Testing and simulation of motor vehicle suspension. *Int. J. Syst. Appl. Eng. Dev.* **2009**, *3*, 74–83.
21. Kumar, M.S. Development of Active System for Automobiles using PID Controller. In Proceedings of the World Congress on Engineering, London, UK, 2–4 July 2008; Volume II.
22. Govinda, K.E.; Raghul, S.; Kishore, K.R.; Mathan, K.M.; Prasanth, C. Enhancing the Closed Loop Performance of Semi-active Suspension System with I-PD Controller. *Mater. Sci. Eng.* **2019**, *561*, 012085. [\[CrossRef\]](#)
23. Amer, Y.A.; EL-Sayed, A.T.; El-Bahrawy, F.T. Torsional vibration reduction for rolling mill's main drive system via negative velocity feedback under parametric excitation. *J. Mech. Sci. Technol.* **2015**, *29*, 1581–1589. [\[CrossRef\]](#)
24. Wenzhi, G.; Zhiyong, H. Active control and simulation test study on torsional vibration of large turbo-generator rotor shaft. *Mech. Mach. Theory* **2010**, *45*, 1326–1336. [\[CrossRef\]](#)
25. El-Sayed, A.T.; Bauomy, H.S. Passive and active controllers for suppressing the torsional vibration of multiple degree-of-freedom system. *J. Vib. Control* **2015**, *21*, 2616–2632. [\[CrossRef\]](#)
26. Jianhui, W.; Yushen, W.; Philip Chen, C.L.; Zhi, L.; Wenqiang, W. Adaptive PI event-triggered control for MIMO nonlinear systems with input delay. *Inf. Sci.* **2024**, *677*, 120817.
27. Yang, W.W.; Yang, Y.J.; Tang, X.Y.; Zhang, K.R.; Li, J.C.; Xu, C. An adaptive P/PI control strategy for a solar volumetric methane/steam reforming reactor with passive thermal management. *Chem. Eng. Sci.* **2023**, *281*, 119005. [\[CrossRef\]](#)
28. Kevorkian, J.; Cole, J.D. Multiple Scale and Singular Perturbation Methods. In *Applied Mathematical Sciences*; Springer: Berlin/Heidelberg, Germany, 1996.
29. Nayfeh, A. *Perturbation Methods*; Wiley: New York, NY, USA, 1973.
30. Dukkupati, R.V. *Solving Vibration Analysis Problems Using Matlab*; New Age International Pvt Ltd Publishers: New Delhi, India, 2007.
31. Bauomy, H.S.; El -Sayed, A.T. Mixed controller (IRC+ NSC) involved in the harmonic vibration response cantilever beam model. *Meas. Control* **2020**, *53*, 1954–1967. [\[CrossRef\]](#)
32. Kamel, M.; Eissa, M.; El-sayed, A.T. Vibration reduction of a nonlinear spring pendulum under multi-parametric excitation via a longitudinal absorber. *Phys. Scr.* **2009**, *80*, 025005. [\[CrossRef\]](#)
33. Bauomy, H.S.; El -Sayed, A.T.; Salem, A.M.; El-Bahrawy, F.T. The improved giant magnetostrictive actuator oscillations via positive position feedback damper. *AIMS Math.* **2023**, *8*, 16864–16886. [\[CrossRef\]](#)
34. Bauomy, H.S.; El -Sayed, A.T.; El-Bahrawy, F.T.; Salem, A.M. Safety of a continuous spinning Shaft's structure from nonlinear vibration with NIPPF. *Alex. Eng. J.* **2023**, *67*, 193–207. [\[CrossRef\]](#)
35. Bauomy, H.S.; El -Sayed, A.T. Vibration performance of a vertical conveyor system under two simultaneous resonances. *Arch. Appl. Mech.* **2018**, *88*, 1349–1368. [\[CrossRef\]](#)
36. El-Sayed, A.T.; Bauomy, H.S. NIPPF versus ANIPPF controller outcomes on semi- direct drive cutting transmission system in a shearer. *Chaos Solitons Fractals* **2022**, *156*, 111778. [\[CrossRef\]](#)
37. Forming Equations of Motion for Multiple Degree-of-Freedom Systems. Available online: www.efunda.com/formulae/vibrations/mdof_eom.cfm (accessed on 1 January 2013).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Review

Systematic Review of Forecasting Models Using Evolving Fuzzy Systems

Sebastian-Camilo Vanegas-Ayala ^{1,2,*} , Julio Barón-Velandia ^{1,†}  and Efren Romero-Riaño ^{3,†} 

- ¹ Doctorate in Engineering, Faculty of Engineering, Universidad Distrital Francisco José de Caldas, Bogotá 111611-111611537, Colombia; jbaron@udistrital.edu.co
- ² Systems Engineering Program, Faculty of Engineering and Basic Sciences, Fundación Universitaria Los Libertadores, Bogotá 111221-111221440, Colombia
- ³ Research Subdirector, Observatorio Colombiano de Ciencia y Tecnología (OCyT), Bogotá 111311-111311474, Colombia; eromero@ocyt.org.co
- * Correspondence: scvanegasa@libertadores.edu.co
- † These authors contributed equally to this work.

Abstract: Currently, the increase in devices capable of continuously collecting data on non-stationary and dynamic variables affects predictive models, particularly if they are not equipped with algorithms capable of adapting their parameters and structure, causing them to be unable to perceive certain time-varying properties or the presence of missing data in data streams. A constantly developing solution to such problems is evolving fuzzy inference systems. The aim of this work was to systematically review forecasting models implemented through evolving fuzzy inference systems, identifying the most common structures, implementation outcomes, and predicted variables to establish an overview of the current state of this technique and its possible applications in other unexplored fields. This research followed the PRISMA methodology of systematic reviews, including scientific articles and patents from three academic databases, one of which offers free access. This was achieved through an identification, selection, and inclusion workflow, obtaining 323 records on which analyses were carried out based on the proposed review questions. In total, 62 investigations were identified, proposing 115 different system structures, mainly focused on increasing precision, in addition to addressing eight main fields of application and some optimization techniques. It was observed that these systems have been successfully implemented in forecasting variables with dynamic behavior and handling missing values, continuous data flows, and non-stationary characteristics. Thus, their use can be extended to phenomena with these properties.

Keywords: evolving fuzzy system; forecasting; fuzzy; literature review; PRISMA; VOSviewer



Citation: Vanegas-Ayala, S.-C.; Barón-Velandia, J.; Romero-Riaño, E. Systematic Review of Forecasting Models Using Evolving Fuzzy Systems. *Computation* **2024**, *12*, 159. <https://doi.org/10.3390/computation12080159>

Academic Editor: Demos T. Tsahalidis

Received: 12 June 2024

Revised: 18 July 2024

Accepted: 27 July 2024

Published: 8 August 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

1.1. Forecasting Models

Forecasting models enable control and decision-making over the modeled environment. A model with a high degree of complexity can add significant computational difficulty to this task, while a model that is too simple may reduce the performance or incorrectly represent the phenomenon [1]. The dynamics of the modeled phenomena can vary over time, implying the need for models to adapt during their operation through data-based approaches and continuous learning [2]. This requires that models have evolutionary dynamic structures capable of capturing variations in the data flow [1,3,4].

Currently, most forecasting methods tend to use limited historical information to generate their initial configurations, and obtaining this information can be difficult or expensive. This approach can hinder models from responding to unexpected changes in the modeled phenomenon. Moreover, some machine learning approaches often rely on offline training, leading to long training cycles, increased computational complexity, and error propagation, which can impair the forecast accuracy [5,6].

Implementing an effective online prediction method becomes imperative to improve the model's response strategy, enabling the transformation of learned knowledge in a highly intuitive and interpretable way [3,6]. Models must respond quickly and learn from situations with varying system conditions, such as drifts, outliers, missing values, shifts, or non-stationary environments. They must also handle unlimited data streams, operating online and in real time [7].

1.2. Evolving Fuzzy Inference Systems

Evolving fuzzy inference systems (EFSs) have recently emerged as a modeling solution for highly complex, nonlinear, online, and non-stationary processes. They can adapt their parameters and alter their internal structures incrementally with new, incoming data samples—for example, by modifying the set of rules or the number of inputs or outputs of the system. This allows them to accommodate new states or changes in behavior over time, without necessitating time-intensive retraining phases or modeling interruptions [8–11].

EFSs provide a flexible structure with incremental and continuous learning, processing data flow samples only once, thus reducing the memory consumption [7]. Moreover, they are fuzzy-rule-based systems (FRSs) capable of self-updating along with their meta-parameters. This involves identifying and refining the antecedent and consequent parts of the rules, as well as dynamically adapting the parameters of the membership functions from the continuous data flow [3,9,12,13]. To date, various EFSs with different evolutionary structures and parameter learning schemes have been proposed to tackle real-world classification, identification, and regression problems, where model interpretability plays a key role [3,9,14].

EFSs are currently employed as practical tools in data streaming applications such as autonomous driving, weather forecasting, and high-frequency trading [15]. As a result, numerous EFSs have been introduced, with notable flexibility, linguistic interoperability, and dynamic autonomy. These systems enable the representation of information with efficient computational performance for continuous self-learning in non-stationary processes [10,11], in addition to offering fast training speeds and high precision compared to other nonlinear prediction strategies [6].

This work provides a systematic review of EFSs for forecasting problems, highlighting that their importance has been indicated by other review works that have covered other approaches, such as control theory [16], heuristic design [14], interpretability [17], their application in data flows [18], and the challenges at the beginning of their development [19]. Thus, a systematic review of forecasting models using EFSs is given to provide an overview of the current state of the art and explore potential applications in unexplored domains. To achieve this, the review aims to address the following questions:

1. Who are the authors with the greatest impact on the development of EFSs?
2. What are the trends in the topics, concepts, or keywords associated with forecasting models using EFSs?
3. How has the frequency of research on forecasting models using EFSs evolved?
4. Which variables or phenomena are predicted by EFSs?
5. What are the structures of EFSs applicable to forecasting?
6. Which optimization methods are implemented in EFSs to enhance their forecasting accuracy?
7. What performance results have been obtained in forecasting models using EFSs?

This manuscript presents an analysis of the state of the art of the technique of EFS for forecasting problems, describing the existing structures, the fields in which they have been applied, and their variables, along with the results of their implementation based on the answers to the questions considered in the review. This allows us to visualize new opportunities for research, innovation, and development and possible applications in problems that have the same characteristics.

This systematic review follows a clear and orderly structure, adapting the PRISMA methodology for systematic reviews, which is described in the following sections of the

document. The Materials and Methods section details the methodology used; in the Results and Discussion, the findings obtained are presented and analyzed; in the Conclusions, the main conclusions of the review are summarized; and, finally, the Future Directions section describes future fields of research and existing challenges.

2. Materials and Methods

The present review adheres to the PRISMA workflow guide for systematic reviews [20]. This process is organized into three phases: identification, selection, and inclusion (Figure 1).

1. The identification phase involves defining the eligibility criteria, selecting the information sources, and developing a search strategy.
2. The selection phase encompasses the selection and collection of data, including the description of the collected data items and the synthesis method.
3. The inclusion phase does not entail any specific development, as it focuses on conducting a thorough search to avoid overlooking relevant records throughout the workflow.

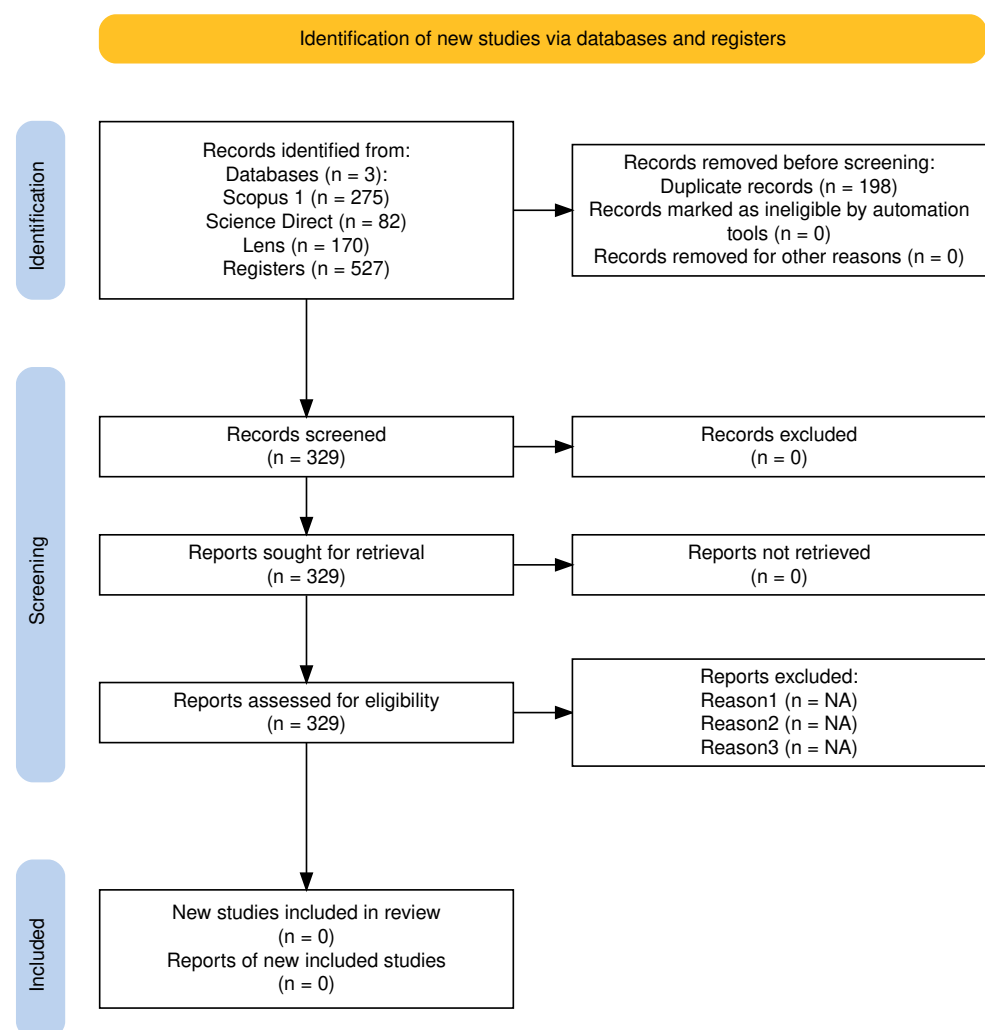


Figure 1. Description of the methodology following the systematic review flow chart according to PRISMA [20,21].

2.1. Eligibility Criteria, Information Sources, and Search Strategy

All studies and research conducted throughout the first two months of 2024 on forecasting models involving EFSs are considered under the admissibility criteria. The data sources used are the scientific databases Scopus, Science Direct, and Lens. In Table 1, the

corresponding search equation is provided for each source, along with the number of records found in each repository.

Table 1. Description of search strategy for the systematic review.

DB	Equation	Records
Scopus	TITLE-ABS-KEY ("Model*" OR "Prediction*" OR "Estimation*" OR "Mathematical Model*" OR "Computer Simulation*" OR "Forecasting*" OR "Prediction Model*" OR "Parameter Estimation*" OR "Parameters Estimation*") AND TITLE-ABS-KEY ("fuzzy" OR "ANFIS" OR "Membership Function*") AND TITLE-ABS-KEY ("Evolving fuzzy system*" OR "Evolving system*")	275
Science Direct	Find articles with these terms: ("Model" OR "Prediction" OR "Estimation" OR "Mathematical Model" OR "Computer Simulation" OR "Forecasting" OR "Prediction Model" OR "Parameter Estimation" OR "Parameters Estimation") Title, abstract or author-specified keywords: (("fuzzy" OR "ANFIS" OR "Membership Function") AND ("Evolving fuzzy system" OR "Evolving system"))	82
Lens	(Title: (("Model*" OR "Prediction*" OR "Estimation*" OR "Mathematical Model*" OR "Computer Simulation*" OR "Forecasting*" OR "Prediction Model*" OR "Parameter Estimation*" OR "Parameters Estimation*") AND ("fuzzy" OR "ANFIS" OR "Membership Function*") AND ("Evolving fuzzy system*" OR "Evolving system*"))) OR (abstract: (("Model*" OR "Prediction*" OR "Estimation*" OR "Mathematical Model*" OR "Computer Simulation*" OR "Forecasting*" OR "Prediction Model*" OR "Parameter Estimation*" OR "Parameters Estimation*") AND ("fuzzy" OR "ANFIS" OR "Membership Function*") AND ("Evolving fuzzy system*" OR "Evolving system*"))) OR (keyword: (("Model*" OR "Prediction*" OR "Estimation*" OR "Mathematical Model*" OR "Computer Simulation*" OR "Forecasting*" OR "Prediction Model*" OR "Parameter Estimation*" OR "Parameters Estimation*") AND ("fuzzy" OR "ANFIS" OR "Membership Function*") AND ("Evolving fuzzy system*" OR "Evolving system*")))	165
Lens Patent Search	(title: (("fuzzy" OR "ANFIS" OR "Membership Function*") AND ("Evolving fuzzy*"))) OR (abstract: (("fuzzy" OR "ANFIS" OR "Membership Function*") AND ("Evolving fuzzy*"))) OR (claim: (("fuzzy" OR "ANFIS" OR "Membership Function*") AND ("Evolving fuzzy*")))	5

After removing duplicates from the 527 initial records, 323 unique records remain for use in the review process.

2.2. Selection Process and Data Collection Process

Each document was exported from the database in comma-separated values and RIS format for bibliometric analysis. The method used to determine the inclusion criteria specifically for the investigative analysis is described as follows.

1. Only patents and articles are selected.
2. Articles must be published in scientific journals.
3. A time window indicating an increase or progress in the field of study is chosen based on bibliometric analysis.
4. Any records unrelated to the research topic—forecasting with EFSs—are excluded.

2.3. Data Items and Synthesis Methods

From the total number of records, the following data were extracted and grouped according to their relevance to the review.

1. Bibliometrics
 - Authors: Used to analyze collaborations and author relevance.
 - Keywords: Used for term co-occurrence analysis.
 - Year of publication: Used to review the temporal distribution of publications on the topic.
2. Investigative
 - Analyzed variable: Corresponds to the modeled variable or phenomenon.
 - Fuzzy inference system type: Indicates the type of fuzzy inference system used.

- Optimization: Indicates the method or type of optimization applied.
 - Model Description: Provides detailed information on the structure of the developed model.
 - Precision: Indicates the measurement and performance value of the obtained model.
3. Filtering fields for investigative analysis
- Title, Abstract: Used to determine whether the documents correspond to the topic.
 - Type of publication: Indicates the medium of publication, such as a journal, book, conference, or other.

2.4. Data Analysis Tools

The VOSviewer tool was utilized to conduct the bibliometric analysis. This tool enables the visualization of relationships that address the review questions. Specifically, it facilitates the analysis of co-authorship and keyword co-occurrence with a temporal focus. The investigative analysis also involves tabulating each article according to the investigative fields of the analyzed variables, the fuzzy inference system type, optimization, and the results.

3. Results

The results were categorized into two sections based on their objectives: one focusing on bibliometric data to address review questions 1 to 3 and the other requiring a more comprehensive investigative analysis to answer questions 4 to 7.

3.1. Bibliometric Analysis

Commencing with the co-authorship analysis depicted in Figure 2, all authors are listed, with three main clusters identified as the most prominent. These clusters represent the researchers Fernando Gomide (light blue), Edwin Lughofer (green), and Plamen Angelov (red).

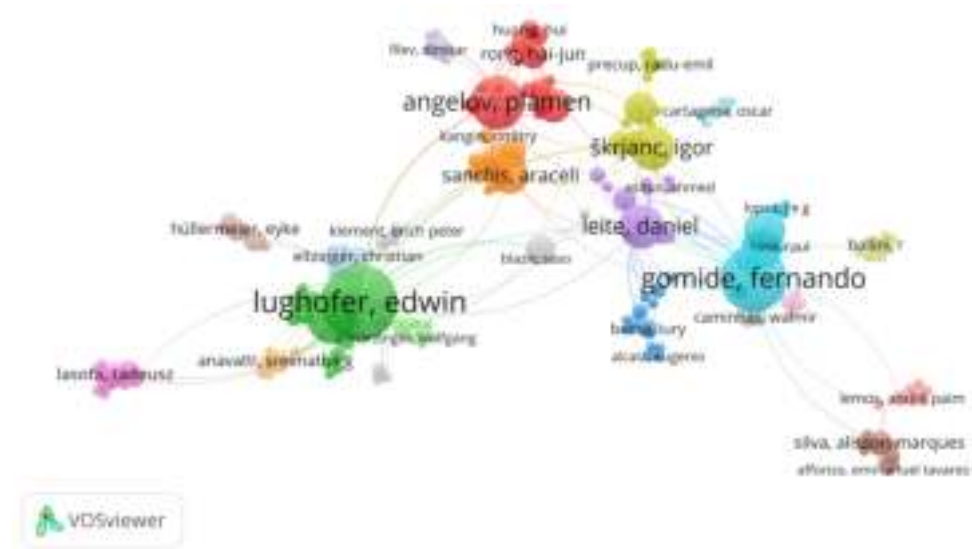


Figure 2. Co-authorship network by number of documents on EFSs.

In Figure 3, the co-occurrence of keywords is examined, with the minimum occurrence threshold set at 13 repetitions to identify the most relevant topics within the scope of the review. In addition to EFSs, it is observed that there are elements related to applications such as real-time systems, data streams, and communication systems (highlighted in red). Furthermore, implementations such as mathematical modeling, Takagi–Sugeno models,

and neural networks are noted (highlighted in blue), along with approaches such as prediction, classification, online learning, and artificial intelligence (highlighted in green).

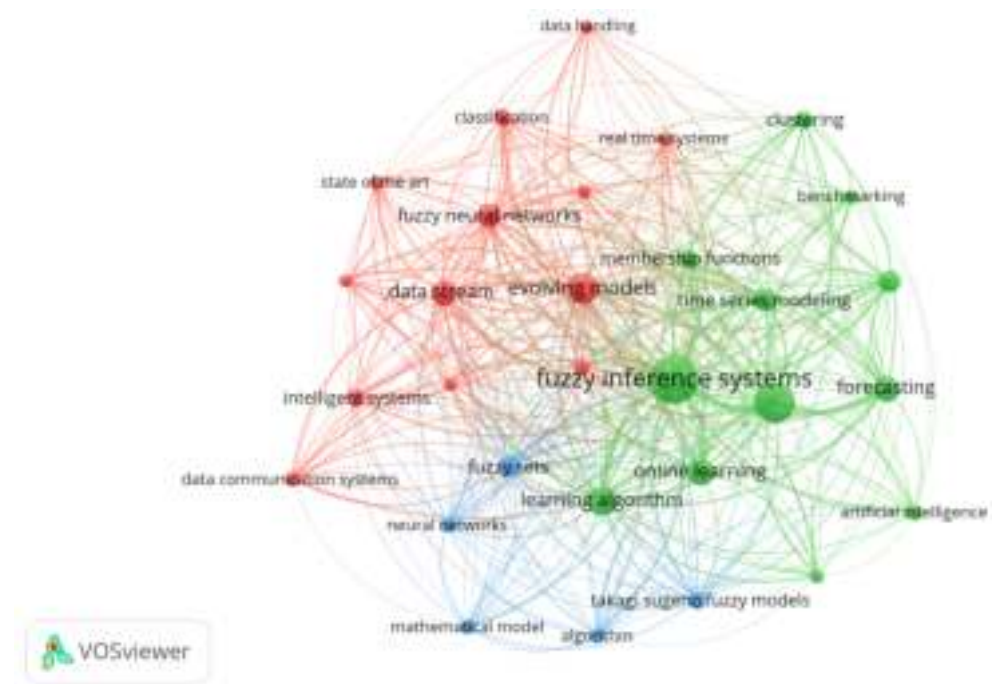


Figure 3. Keyword co-occurrence map of each document on EFSs.

In Figure 4, the co-occurrence of keywords over time is examined, with the minimum occurrence threshold set at 13 repetitions. Research topics with the most recent development include data streams, information classification, forecasting, and time series.

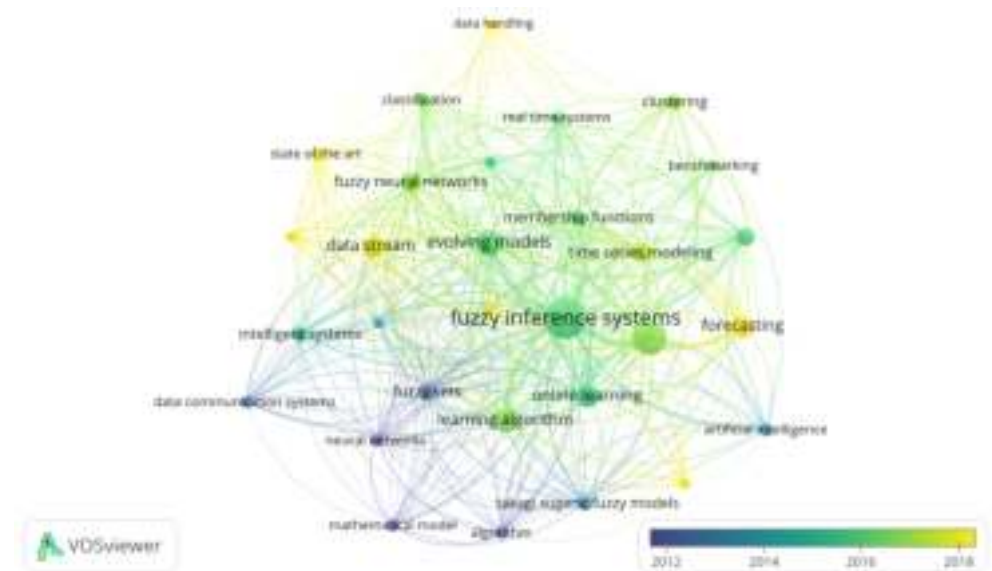


Figure 4. Keyword co-occurrence map of each document on EFSs over time.

Figure 5 illustrates the frequency of research on forecasting models using EFSs over time. A growth trend has been observed since 2010, with an approximate average of 20 studies per year. This growth is attributed to the maturation of the technology and supporting computer tools.

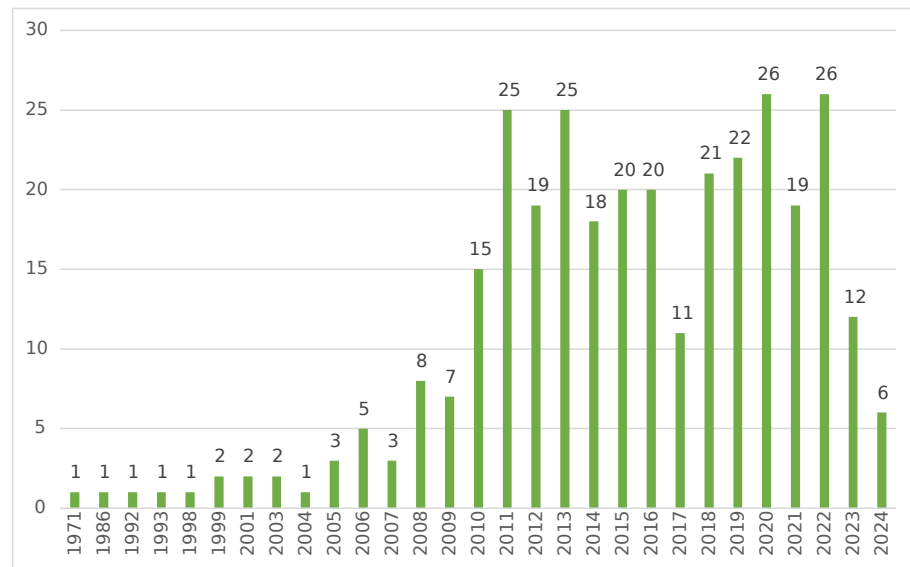


Figure 5. Histogram of the distribution of the amount of research on EFSs over time.

3.2. Investigative Analysis

Out of the total number of investigations, 82 studies have been selected through filtering since 2010, focusing on journal articles and patent registrations, as stipulated by the selection methodology. Among these, 20 studies were discarded due to not corresponding to forecasting models or not utilizing EFSs.

Table 2 analyzes the remaining 62 investigations and briefly describes the implemented models. Notably, the first five entries correspond to patents.

Table 2. Types of EFS implemented by each investigation.

Research	Type of EFS
[22]	Fuzzy logic DB
[23]	Evolving fuzzy neural network (EFuNN)
[24]	Evolving fuzzy neural network (EFuNN)
[25]	Evolving fuzzy neural network (EFuNN)
[26]	Fuzzy network
[27]	Self-organizing fuzzy inference system (SOFIS+)
[28]	Dynamic similarity weighted EFS (DSW-EFS)
[29]	Zero-order EFSs: Self-organizing fuzzy inference system (SOFIS+); Zero-order ALMMo (ALMMo0); Self-organizing fuzzy belief inference system (SOFBIS)
[30]	Fuzzy prediction interval (FPI)
[5]	Dual-model semi-supervised self-organizing fuzzy inference system (DMS3OF); self-organizing fuzzy inference system (SOFIS+); self-organizing fuzzy belief inference system (SOFBIS)
[31]	Evolving neuro-fuzzy (eNF)
[32]	Fuzzification DSP algorithm (FDSPA)
[33]	Adaptive evolving fuzzy (AEF)
[34]	Evolving ellipsoidal fuzzy information granules (EEFIGs); T-S fuzzy auto-regressive with exogenous inputs (TS fuzzy ARX)
[35]	Modified set membership eMG (MSM-eMG); variable step-size eMG (VS-eMG)
[36]	Evolving Neo-Fuzzy Neuron with Missing Data Procedure (eNFN-MDP)
[11]	Statistical evolving fuzzy inference system (SEFIS)
[37]	Online sequential ensembling of fuzzy systems (OS-FS)
[38]	Self-adaptive fuzzy learning (SAFL)
[39]	Weakly supervised scalable teacher forcing network (WeScatterNet)
[40]	Multi-layer ensemble evolving fuzzy inference system (MEEFIS); first-order evolving fuzzy inference systems (EFISs)
[41]	TSK streams
[42]	EFS with self-learning/adaptive thresholds (EFS-SLAT)
[43]	Evolving neuro-fuzzy Takagi–Sugeno network (Evolving NF-TS)

Table 2. Cont.

Research	Type of EFS
[44]	Enhanced set membership ePL-KRLS (ESM-ePL-KRLS); set membership ePL-KRLS (SM-ePL-KRLS)
[45]	Extended zero-order ALMMo (ALMMo-0*)
[46]	Incremental fuzzy C-regression (InFuR)
[47]	Generalized heterogeneous interval type-2 fuzzy classifier (GHIT2Class)
[48]	Evolving fuzzy granular predictor (eFGP)
[49]	TS evolving fuzzy Kalman filter
[50]	Adaptive fuzzy classifier based on gradient descent (AFCGD)
[51]	Evolving Takagi–Sugeno (eTS); extended Takagi–Sugeno (xTS); evolving fuzzy participatory learning (ePL)
[52]	Robust structure identification method (RSIM)
[53]	Parsimonious network based on fuzzy inference system (PANFIS)
[54]	Cloud-based identification procedure
[55]	Time series evolving fuzzy engine (TiSEFE)
[56]	Weighted neuro–neo–fuzzy–ANARX Model
[57]	GENERIC classifier (gClass)
[58]	Scaffolding Type-2 classifier (ST2Class)
[59]	Evolving heterogeneous fuzzy inference system (eHFIS)
[60]	Evolving fuzzy-GARCH
[61]	Evolving classifier (eClass)
[62]	Evolving output-context fuzzy system (EOCFS)
[63]	Evolving information granule (EIG)
[64]	Generalized smart EFS (Gen-Smart-EFS) (GS-EFS)
[65]	eTS anomaly detection (eTSAD)
[66]	Adaptive neuro–fuzzy inference system (ANFIS)
[67]	Evolving neuro–fuzzy recurrent network (ENFRN)
[68]	Evolving classifier (eClass)
[69]	Generic evolving neuro–fuzzy inference system (GENEFIS)
[70]	Evolving extended Takagi–Sugeno (exTS)
[71]	Evolving fuzzy participatory learning + (ePL+)
[72]	Gaussian participatory evolving clustering
[73]	Evolving classifier (eClass)
[74]	Evolving type-2 neural fuzzy inference system (eT2FIS)
[75]	eTS based on switching to neighboring (eTS-N)
[76]	Four-layer fuzzy multiagent system (FMAS)
[77]	Self-organizing FNN (SOFNN)
[4]	Evolving Takagi–Sugeno (eTS); flexible fuzzy inference systems (FLEXFIS)
[78]	Evolving Takagi–Sugeno (eTS)
[79]	Evolving classifier of agent behaviors (EvCAB)
[80]	eSensors

Table 3 displays the variables or phenomena predicted by EFSs, with each variable grouped according to its respective field of study. The fields with the most significant applications include economics or finance, health, manufacturing, and technology. Particularly in the technological field, a considerable portion is dedicated to applying synthetic variables to verify and validate new model developments.

Table 3. List of variables predicted using EFSs by number of investigations.

Field	Analyzed Variable	Count
Climatic—Meteorological	Heat exchanger system	1
Economic—Financial	Cryptocurrency	1
	Financial return volatility	1
	Housing prices	1
	Stock market volatility	2
	Yield curve	1

Table 3. Cont.

Field	Analyzed Variable	Count
Energy	Battery remaining useful life	1
	Hot-spot temperature	1
	Thermal modeling of power transformers	1
Health	Activity recognition	1
	Family trees	1
	Gene expression	3
	Heart sounds	2
Manufacturing	Fault detection and diagnosis	1
	Manufacturing process	1
	Real-time gear system fault	1
	Remaining useful life	1
	Robotic disassembly sequence planning (DSP)	1
	Water-chiller plant	1
Social	Real estate appraisal	1
	Web news	1
Technological	Example data	30
	Example image	1
	Fuzzy network structures	1
	Network traffic	1
	Non-Gaussian noise	1
	Sensor data	1
	UNIX commands	1
Transport and logistics	Helicopter rotor	1

Table 2 describes the 115 evolutionary fuzzy inference system structures applicable to forecasting. Additionally, Figure 6 illustrates the usage percentages of various EFSs, considering a minimum of two applications. eTS has the highest frequency of application, followed by eClass, eMG, and xTS, in predicting the variables indicated in Table 3.

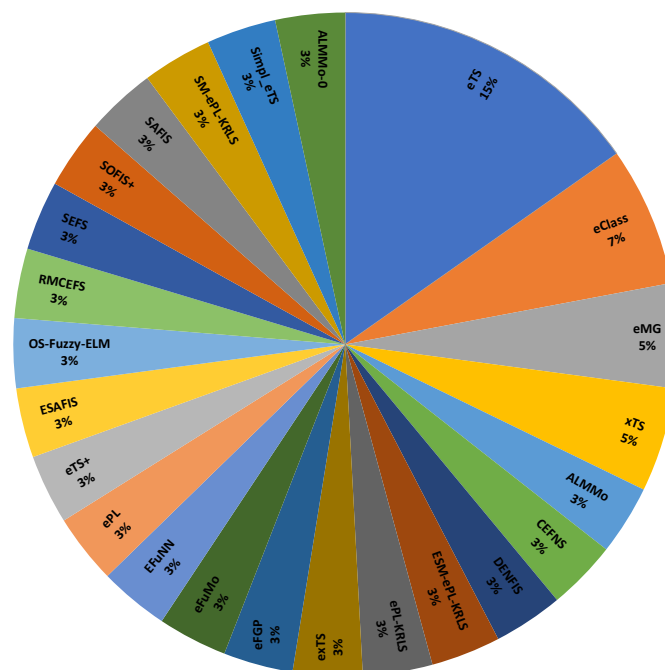


Figure 6. Pie chart of the distribution of EFS types applied to prediction problems.

Moreover, Figure 7 displays the different EFSs applied in forecasting the identified variables at least twice. Notably, example data, the thermal modeling of power transformers, and non-Gaussian noise are the variables predicted by the largest number of different EFSs.

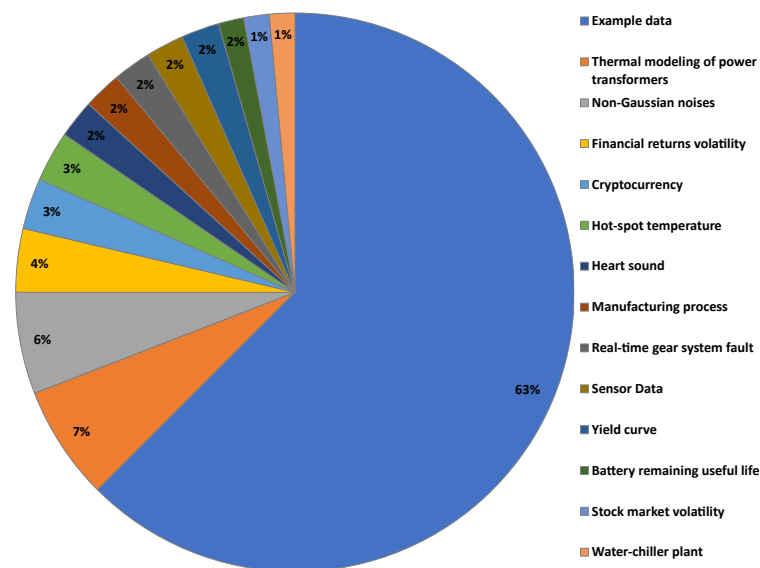


Figure 7. Pie chart of the distribution of the variable types in research that applies EFSs.

Regarding the optimization methods implemented in EFSs to enhance the forecasting precision, Table 4 shows that 12 investigations include optimization elements, specifying which methods have been implemented and their frequency.

Table 4. List of optimization methods by number of investigations that apply them.

Optimization	Count
Adadelata function (NaD)	1
Adaptive boosting (AdaBoost)	1
Adaptive particle filter (aPF)	1
Adaptive semi-supervised ensemble method (ASSEMBLE)	1
Adaptive wavelet	1
Boosting framework for semi-supervised learning (SemiBoost)	1
Decoupled extended Kalman filter (DEKF)	1
Drift detection and reaction	1
Extended binary search tree (E-BST)	1
Fuzzily weighted AdaBoost (FWAdaBoost)	1
Fuzzy weighted recursive least squares with variable-direction forgetting (FWRLS-VDF)	1
Genetic algorithm (GA)	1
Global least square	1
Gradient descent (GD)	2
Gradual forgetting	1
Hyperparameter optimization based on Hoeffding's inequality	1
MapReduce	1
Merging algorithm	1
ModifiedEvolving	1
Robust AdaBoost (RobAdaBoost)	1
Semi-supervised FWAdaBoost (SSFWAdaBoost)	1
Stagewise additive modeling using a multi-class exponential (SAMME)	1
Adadelata function (NaD)	1

Furthermore, forecasting models using EFSs employ various performance measures. Notably, the root mean squared error (RMSE) is mentioned in 23 investigations, while

the mean squared error (MSE) is considered in five. The obtained values are 0.0662 and 0.0561, respectively.

4. Discussion

In this section, we discuss relevant elements derived from the analyses carried out, focusing on the findings of the common structures, the results of the implementation of the models, and the variables used for forecasting with EFSs.

Within the types of EFS structures applicable to forecasting, the most used and referenced are eGNN, FLEXFIS, eTS, and eClass, proposed by Fernando Gomide, Edwin Lughofer, and Plamen Angelov, respectively, which provide a point of comparison for future research, in which they may be implemented, contrasted, and improved. Additionally, various authors have presented new proposals over the last decade that introduce novel approaches or enhancements to previous implementations, indicating a burgeoning and expanding field.

Even now, new proposals are being developed from the approaches analyzed—for example, seeking to increase their interpretability, such as X-Fuzz [81]; handling missing data [82]; using other approaches, such as the use of multiple layers in hierarchical evolving fuzzy systems [83]; or improving existing approaches, such as dynamic rule generation [84].

When analyzing the current trends, it is observed that the structures of the EFSs applicable to forecasting are predominantly based on Takagi–Sugeno-type fuzzy inference systems. These systems address approaches such as the creation or dynamic adjustment of rules [31,38,41,45] and the configuration or monitoring of input or output flows [59,65]. However, there are a few configurations focused on managing missing data [36,48] and likewise based on Mamdani-type fuzzy inference systems [62,63,74], which offer higher interpretability.

This trend shows that there are currently many proposals focused on precision but few on interpretability or the handling of missing data, providing a research opportunity. This is evidenced, for example, by the fact that of the 115 proposals analyzed, only 10 use Mamdani-type systems:

- Evolving output-context fuzzy system (EOCFS);
- Evolving construction scheme for fuzzy systems (ECSFS);
- Evolving information granule (EIG);
- Evolving type-2 neural fuzzy inference system (eT2FIS);
- Fuzzy adaptive learning control network applying adaptive resonance theory (FALCON-ART);
- Evolving fuzzy neural network (EFuNN);
- Evolving neural–fuzzy semantic memory (eFSM);
- Generic SOFNN (GenSoFNN);
- Rough-set-based pseudo outer product (RSPOP);
- Type-2 self-organizing neural fuzzy system (T2SONFS).

The precision values obtained in the various approaches addressed by EFSs demonstrate their robustness in implementation. Notably, these systems continually adapt their structures to accommodate new data flows or variations in the behavior of the analyzed variable.

It is observed that by implementing optimization methods, the characteristics of the existing systems are improved, thereby increasing their precision or reducing the algorithmic complexity and time. However, their limited implementation stems from the technique's current maturity level; therefore, an increase in its implementation is expected in the future. Currently, the focus is on developing new proposals with diverse characteristics, rather than implementing optimization mechanisms in existing structures.

Within the application field of data flow forecasting, several research opportunities related to the forecasted variables or phenomena are observed:

- Variables that exhibit non-stationary and dynamic characteristics over time—for example, climatic or meteorological variables, such as those in the heat exchanger system [30];
- Phenomena that demonstrate strong fluctuations over time and are interdependent on unknown variables—for example, in economic phenomena, such as stock market volatility [60,76];
- Variables in continuous data streams, such as real-time gear system faults [31] and network traffic [55];
- Time series with missing data [36] within dynamic and non-stationary data streams, such as in cryptocurrency [48].

5. Conclusions

Applications of EFSs are found in fields such as the climate, economics, energy, health, manufacturing, the social sciences, and transportation, with high-performance values in forecasting dynamic, non-stationary variables, even with continuous data flows and missing data. It is evident that it is feasible to obtain high accuracy values when forecasting phenomena with these properties, particularly in conjunction with optimization methods, which are currently of limited use, mainly due to the technique's current maturity level.

It is found that there is significant variety in the configuration and assembly proposals for EFSs for data flow forecasting. It is suggested that if higher accuracy is desired, Sugeno-type structures are the most common structures. On the other hand, if we focus on higher interpretability, only a small proportion, approximately 8.6957%, of the proposals are based on Mamdani-type fuzzy inference systems, and only two on missing data management.

6. Future Directions

In this section, we explore the possible future research avenues and advances in the field of EFSs in forecasting problems based on the review carried out. Emerging trends and unresolved problems that can advance the field are highlighted.

- Emerging Trends
 - The forecasting of dynamic, non-stationary data flows and with missing data such as economic indices, real-time health statuses, environmental data, consumer behavior, interactions on social networks, data from interactions on educational platforms, mobility data, and sensor data, among others.
 - The implementation of optimization algorithms to increase the precision of the models, seeking to maintain the levels of interpretability obtained and without modifying the semantics of the model.
 - The development of new proposals that focus on guaranteeing the forecasting of data flows with high interpretability and the handling of missing data or continuous data flows.
- Unresolved Issues
 - Regarding the classification of existing structures, more than 100 proposals for evolutionary fuzzy inference systems were found, so it is necessary to develop a classification system to make it easier to implement and compare them. For example, some proposals identified in this review are grouped by their approach, i.e., precision or interpretability, or focused on modifying rules or inputs and outputs.
 - It is necessary to efficiently combine proposals that combine high levels of interpretability and high precision in a way that is simple to implement.

Author Contributions: Conceptualization, S.-C.V.-A. and J.B.-V.; methodology, E.R.-R.; software, S.-C.V.-A.; validation, S.-C.V.-A., J.B.-V., and E.R.-R.; investigation, S.-C.V.-A.; resources, E.R.-R.; data curation, S.-C.V.-A.; writing—original draft preparation, S.-C.V.-A.; writing—review and editing, S.-C.V.-A.; visualization, E.R.-R.; supervision, J.B.-V. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Alcalá, E.; Bessa, I.; Puig, V.; Sename, O.; Palhares, R. MPC using an on-line TS fuzzy learning approach with application to autonomous driving. *Appl. Soft Comput.* **2022**, *130*, 109698. [\[CrossRef\]](#)
- Andonovski, G.; Costa, B.S.J.; Blažič, S.; Škrjanc, I. Robust evolving controller for simulated surge tank and for real two-tank plant. *Automatisierungstechnik* **2018**, *66*, 725–734. [\[CrossRef\]](#)
- Gu, X.; Angelov, P.; Shen, Q. Self-Organizing Fuzzy Belief Inference System for Classification. *IEEE Trans. Fuzzy Syst.* **2022**, *30*, 5473–5483. [\[CrossRef\]](#)
- Lughofer, E.; Angelov, P. Handling drifts and shifts in on-line data streams with evolving fuzzy systems. *Appl. Soft Comput.* **2011**, *11*, 2057–2068. [\[CrossRef\]](#)
- Gu, X. A dual-model semi-supervised self-organizing fuzzy inference system for data stream classification. *Appl. Soft Comput.* **2023**, *136*, 110053. [\[CrossRef\]](#)
- Sun, J.; Gan, X.; Gong, D.; Tang, X.; Dai, H.; Zhong, Z. A self-evolving fuzzy system online prediction-based dynamic multi-objective evolutionary algorithm. *Inf. Sci.* **2022**, *612*, 638–654. [\[CrossRef\]](#)
- Rodrigues, F.; Silva, A.; Lemos, A. Evolving fuzzy predictor with multivariable Gaussian participatory learning and multi-innovations recursive weighted least squares: eFMI. *Evol. Syst.* **2022**, *13*, 667–686. [\[CrossRef\]](#)
- Lughofer, E.; Škrjanc, I. Evolving Error Feedback Fuzzy Model for Improved Robustness under Measurement Noise. *IEEE Trans. Fuzzy Syst.* **2022**, *31*, 997–1008. [\[CrossRef\]](#)
- Škrjanc, I.; Iglesias, J.A.; Sanchis, A.; Leite, D.; Lughofer, E.; Gomide, F. Evolving fuzzy and neuro-fuzzy approaches in clustering, regression, identification, and classification: A Survey. *Inf. Sci.* **2019**, *490*, 344–368. [\[CrossRef\]](#)
- Ferdaus, M.; Pratama, M.; Anavatti, S.; Garratt, M. PALM: An Incremental Construction of Hyperplanes for Data Stream Regression. *IEEE Trans. Fuzzy Syst.* **2019**, *27*, 2115–2129. [\[CrossRef\]](#)
- Yang, Z.X.; Rong, H.J.; Angelov, P.; Yang, Z.X. Statistically Evolving Fuzzy Inference System for Non-Gaussian Noises. *IEEE Trans. Fuzzy Syst.* **2021**, *30*, 2649–2664. [\[CrossRef\]](#)
- Angelov, P.; Filev, D. An approach to online identification of Takagi-Sugeno fuzzy models. *IEEE Trans. Syst. Man Cybern. Part B Cybern.* **2004**, *34*, 484–498. [\[CrossRef\]](#) [\[PubMed\]](#)
- Dell’Anna, D.; Jamshidnejad, A. Evolving Fuzzy logic Systems for creative personalized Socially Assistive Robots. *Eng. Appl. Artif. Intell.* **2022**, *114*, 105064. [\[CrossRef\]](#)
- Ojha, V.; Abraham, A.; Snášel, V. Heuristic design of fuzzy inference systems: A review of three decades of research. *Eng. Appl. Artif. Intell.* **2019**, *85*, 845–864. [\[CrossRef\]](#)
- Huang, H.; Rong, H.J.; Yang, Z.X.; Vong, C.M. Recursive least mean dual p-power solution to the generalization of evolving fuzzy system under multiple noises. *Inf. Sci.* **2022**, *609*, 228–247. [\[CrossRef\]](#)
- Radu-Emil Precup, A.T.N.; Blažič, S. A survey on fuzzy control for mechatronics applications. *Int. J. Syst. Sci.* **2024**, *55*, 771–813. [\[CrossRef\]](#)
- Lughofer, E. On-line assurance of interpretability criteria in evolving fuzzy systems—Achievements, new concepts and open issues. *Inf. Sci.* **2013**, *251*, 22–46. [\[CrossRef\]](#)
- Baruah, R.D.; Angelov, P. Evolving fuzzy systems for data streams: A survey. *WIREs Data Min. Knowl. Discov.* **2011**, *1*, 461–476. [\[CrossRef\]](#)
- Pedrycz, W. Evolvable fuzzy systems: Some insights and challenges. *Evol. Syst.* **2010**, *1*, 73–82. [\[CrossRef\]](#)
- Page, M. PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *BMJ* **2021**, *372*, 160. [\[CrossRef\]](#)
- Haddaway, N.; Page, M.; Pritchard, C.; McGuinness, L. PRISMA2020: An R package and Shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and Open Synthesis. *Campbell Syst. Rev.* **2022**, *18*, 1230. [\[CrossRef\]](#) [\[PubMed\]](#)
- Charles, D.M. Method and System for Discovering Ancestors Using Genomic and Genealogic Data. US 2017/0213127 A1, 27 July 2017.
- Edmund, R.A.; Erwin, F.M.; James, S.M.; Kirilov, K.N.; John, G.P. Medical Applications of Adaptive Learning Systems Using Gene Expression Data. US 7370021 B2, 6 May 2008.
- Edmund, R.A.; Erwin, F.M.; James, S.M.; Kirilov, K.N.; John, G.P. Medical Applications of Adaptive Learning Systems and Diagnostic Kits. NZ 547894 A, 6 May 2008.
- Kirilov, K.N.; Erwin, F.M.; James, S.M.; Edmund, R.A. Medical Decision Support Systems Utilizing Gene Expression and Clinical Information and Method for Use. EP 1534122 B1, 20 July 2016.
- Flinn, S.D.; Moneypenny, N.F. Adaptive Recombinant Systems. WO 2005/054982 A2, 16 June 2005.

27. Xiao, F.; Liu, H.; Lu, J. A new approach based on a 1D + 2D convolutional neural network and evolving fuzzy system for the diagnosis of cardiovascular disease from heart sound signals. *Appl. Acoust.* **2024**, *216*, 109723. [\[CrossRef\]](#)
28. Li, H.; Zhao, T. A dynamic similarity weighted evolving fuzzy system for concept drift of data streams. *Inf. Sci.* **2024**, *659*, 120062. [\[CrossRef\]](#)
29. Gu, X.; Angelov, P.; Shen, Q. Semi-Supervised Fuzzily Weighted Adaptive Boosting for Classification. *IEEE Trans. Fuzzy Syst.* **2024**, *32*, 2318–2330. [\[CrossRef\]](#)
30. Cartagena, O.; Ožbot, M.; Sáez, D.; Škrjanc, I. Evolving fuzzy prediction interval for fault detection in a heat exchanger. *Appl. Soft Comput.* **2023**, *145*, 110625. [\[CrossRef\]](#)
31. Shah, J.; Wang, W. An evolving neuro-fuzzy classifier for fault diagnosis of gear systems. *ISA Trans.* **2022**, *123*, 372–380. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Ye, F.; Perrett, J.; Zhang, L.; Laili, Y.; Wang, Y. A self-evolving system for robotic disassembly sequence planning under uncertain interference conditions. *Robot. Comput. Integr. Manuf.* **2022**, *78*, 102392. [\[CrossRef\]](#)
33. Ahwiadi, M.; Wang, W. An Adaptive Evolving Fuzzy Technique for Prognosis of Dynamic Systems. *IEEE Trans. Fuzzy Syst.* **2021**, *30*, 841–849. [\[CrossRef\]](#)
34. Cordovil, L.; Coutinho, P.; Bessa, I.; Peixoto, M.; Palhares, R. Learning event-triggered control based on evolving data-driven fuzzy granular models. *Int. J. Robust Nonlinear Control* **2022**, *32*, 2805–2827. [\[CrossRef\]](#)
35. Rocha, M.; Alves, K.; Queiroz, E.; Oliveira, F.; Hell, M.; Aguiar, E. Power Transformers Thermal Modeling Based on the Modified Set-Membership Evolving Multivariable Gaussian and Variable Step-Size Evolving Multivariable Gaussian. *J. Control. Autom. Electr. Syst.* **2022**, *33*, 1044–1055. [\[CrossRef\]](#)
36. da Silva Júnior, G.A.; Silva, A. A simple and efficient incremental missing data imputation method for evolving neo-fuzzy network. *Evol. Syst.* **2022**, *13*, 201–220. [\[CrossRef\]](#)
37. Lughofer, E.; Pratama, M. Online sequential ensembling of predictive fuzzy systems. *Evol. Syst.* **2022**, *13*, 361–386. [\[CrossRef\]](#)
38. Gu, X.; Shen, Q. A self-adaptive fuzzy learning system for streaming data prediction. *Inf. Sci.* **2021**, *579*, 623–647. [\[CrossRef\]](#)
39. Pratama, M.; Za'in, C.; Lughofer, E.; Pardede, E.; Rahayu, D.A. Scalable teacher forcing network for semi-supervised large scale data streams. *Inf. Sci.* **2021**, *576*, 407–431. [\[CrossRef\]](#)
40. Gu, X. Multilayer Ensemble Evolving Fuzzy Inference System. *IEEE Trans. Fuzzy Syst.* **2020**, *29*, 2425–2431. [\[CrossRef\]](#)
41. Shaker, A.; Hüllermeier, E. TSK-Streams: Learning TSK fuzzy systems for regression on data streams. *Data Min. Knowl. Discov.* **2021**, *35*, 1941–1971. [\[CrossRef\]](#)
42. Ge, D.; Zeng, X.J. Learning data streams online—An evolving fuzzy system approach with self-learning/adaptive thresholds. *Inf. Sci.* **2020**, *507*, 172–184. [\[CrossRef\]](#)
43. Júnior, S.; Oliveira Serra, G. Intelligent forecasting of time series based on evolving distributed Neuro-Fuzzy network. *Comput. Intell.* **2020**, *36*, 1394–1413. [\[CrossRef\]](#)
44. Alves, K.; Hell, M.; Oliveira, F.; Aguiar, E. An enhanced set-membership evolving participatory learning with kernel recursive least squares applied to thermal modeling of power transformers. *Electr. Power Syst. Res.* **2020**, *184*, 106334. [\[CrossRef\]](#)
45. Soares, E.; Angelov, P.; Gu, X. Autonomous Learning Multiple-Model zero-order classifier for heart sound classification. *Appl. Soft Comput. J.* **2020**, *94*, 106449. [\[CrossRef\]](#)
46. Blazic, S.; Škrjanc, I. Incremental Fuzzy C-Regression Clustering from Streaming Data for Local-Model-Network Identification. *IEEE Trans. Fuzzy Syst.* **2019**, *28*, 758–767. [\[CrossRef\]](#)
47. Zhao, J.; Yang, L.; Wang, L.; Wang, W. A Generalized Heterogeneous Type-2 Fuzzy Classifier and Its Industrial Application. *IEEE Trans. Fuzzy Syst.* **2019**, *28*, 2287–2301. [\[CrossRef\]](#)
48. Garcia, C.; Esmin, A.; Leite, D.; Škrjanc, I. Evolvable fuzzy systems from data streams with missing values: With application to temporal pattern recognition and cryptocurrency prediction. *Pattern Recognit. Lett.* **2019**, *128*, 278–282. [\[CrossRef\]](#)
49. Pires, D.; Serra, G.O. Methodology for Evolving Fuzzy Kalman Filter Identification. *Int. J. Control Autom. Syst.* **2019**, *17*, 793–800. [\[CrossRef\]](#)
50. Shahparast, H.; Mansoori, E.; Jahromi, M. AFCGD: An adaptive fuzzy classifier based on gradient descent. *Soft Comput.* **2019**, *23*, 4557–4571. [\[CrossRef\]](#)
51. Maciel, L.; Ballini, R.; Gomide, F. Evolving fuzzy modelling for yield curve forecasting. *Int. J. Econ. Bus. Res.* **2018**, *15*, 290–311. [\[CrossRef\]](#)
52. Sa'ad, H.; Isa, N.; Ahmed, M.M.; Sa'd, A. A robust structure identification method for evolving fuzzy system. *Expert Syst. Appl.* **2018**, *93*, 267–282. [\[CrossRef\]](#)
53. Za'in, C.; Pratama, M.; Lughofer, E.; Ferdaus, M.; Cai, Q.; Prasad, M. Big Data Analytics based on PANFIS MapReduce. *Procedia Comput. Sci.* **2018**, *144*, 140–152. [\[CrossRef\]](#)
54. Andonovski, G.; Mušič, G.; Blažič, S.; Škrjanc, I. Evolving model identification for process monitoring and prediction of non-linear systems. *Eng. Appl. Artif. Intell.* **2018**, *68*, 214–221. [\[CrossRef\]](#)
55. Abdullah, S.; AL-Hashmi, A. TiSEFE: Time series evolving fuzzy engine for network traffic classification. *Int. J. Commun. Netw. Inf. Secur.* **2018**, *10*, 116–124. [\[CrossRef\]](#)
56. Hu, Z.; Bodyanskiy, Y.; Tyshchenko, O.; Boiko, O. Adaptive Forecasting of Non-Stationary Nonlinear Time Series Based on the Evolving Weighted Neuro-Neo-Fuzzy-ANARX-Model. *arXiv* **2016**, arXiv:1610.06486.

57. Pratama, M.; Lu, J.; Anavatti, S.; Lughofer, E.; Lim, C.P. An incremental meta-cognitive-based scaffolding fuzzy neural network. *Neurocomputing* **2016**, *171*, 89–105. [\[CrossRef\]](#)
58. Pratama, M.; Lu, J.; Lughofer, E.; Zhang, G.; Anavatti, S. Scaffolding type-2 classifier for incremental learning under concept drifts. *Neurocomputing* **2016**, *191*, 304–329. [\[CrossRef\]](#)
59. Alizadeh, S.; Kalhor, A.; Jamalabadi, H.; Araabi, B.; Ahmadabadi, M. Online local input selection through evolving heterogeneous fuzzy inference system. *IEEE Trans. Fuzzy Syst.* **2016**, *24*, 1364–1377. [\[CrossRef\]](#)
60. Maciel, L.; Gomide, F.; Ballini, R. Evolving Fuzzy-GARCH Approach for Financial Volatility Modeling and Forecasting. *Comput. Econ.* **2016**, *48*, 379–398. [\[CrossRef\]](#)
61. Iglesias, J.; Tiemblo, A.; Ledezma, A.; Sanchis, A. Web news mining in an evolving framework. *Inf. Fusion* **2016**, *28*, 90–98. [\[CrossRef\]](#)
62. Ahmed, M.M.; Isa, N. Evolving output-context fuzzy system for effective rule base. *Expert Syst. Appl.* **2015**, *42*, 1972–1986. [\[CrossRef\]](#)
63. Ahmed, M.M.; Isa, N. Information granularity model for evolving context-based fuzzy system. *Appl. Soft Comput. J.* **2015**, *33*, 183–196. [\[CrossRef\]](#)
64. Lughofer, E.; Cernuda, C.; Kindermann, S.; Pratama, M. Generalized smart evolving fuzzy systems. *Evol. Syst.* **2015**, *6*, 269–292. [\[CrossRef\]](#)
65. Moshtaghi, M.; Bezdek, J.; Leckie, C.; Karunasekera, S.; Palaniswami, M. Evolving fuzzy rules for anomaly detection in data streams. *IEEE Trans. Fuzzy Syst.* **2014**, *23*, 688–700. [\[CrossRef\]](#)
66. Ye, Bodyanskiy, A.; Vynokurova, O. Evolving spiking wavelet-neuro-fuzzy self-learning system. *Appl. Soft Comput.* **2014**, *14*, 252–258. [\[CrossRef\]](#)
67. Jesús Rubio, J.; Pérez-Cruz, J. Evolving intelligent system for the modelling of nonlinear systems with dead-zone input. *Appl. Soft Comput. J.* **2014**, *14*, 289–304. [\[CrossRef\]](#)
68. Iglesias, J.; Ledezma, A.; Sanchis, A. Evolving classification of UNIX users' behaviors. *Evol. Syst.* **2014**, *5*, 231–238. [\[CrossRef\]](#)
69. Pratama, M.; Anavatti, S.; Lughofer, E. Genefis: Toward an effective localist network. *IEEE Trans. Fuzzy Syst.* **2013**, *22*, 547–562. [\[CrossRef\]](#)
70. Ramasso, E.; Gouriveau, R. Remaining useful life estimation by classification of predictions based on a neuro-fuzzy system and theory of belief functions. *IEEE Trans. Reliab.* **2014**, *63*, 555–566. [\[CrossRef\]](#)
71. Maciel, L.; Gomide, F.; Ballini, R. Enhanced evolving participatory learning fuzzy modeling: An application for asset returns volatility forecasting. *Evol. Syst.* **2014**, *5*, 75–88. [\[CrossRef\]](#)
72. Lemos, A.; Caminhas, W.; Gomide, F. Adaptive fault detection and diagnosis using an evolving fuzzy classifier. *Inf. Sci.* **2013**, *220*, 64–85. [\[CrossRef\]](#)
73. Ordóñez, F.J.; Iglesias, J.; Toledo, P.; Ledezma, A.; Sanchis, A. Online activity recognition using evolving classifiers. *Expert Syst. Appl.* **2013**, *40*, 1248–1255. [\[CrossRef\]](#)
74. Tung, S.; Quek, C.; Guan, C. eT2FIS: An Evolving Type-2 Neural Fuzzy Inference System. *Inf. Sci.* **2013**, *220*, 124–148. [\[CrossRef\]](#)
75. Kalhor, A.; Araabi, B.; Lucas, C. Evolving Takagi-Sugeno fuzzy model based on switching to neighboring models. *Appl. Soft Comput. J.* **2013**, *13*, 939–946. [\[CrossRef\]](#)
76. Zarandi, M.; Hadavandi, E.; Turksen, I. A hybrid fuzzy intelligent agent-based system for stock price prediction. *Int. J. Intell. Syst.* **2012**, *27*, 947–969. [\[CrossRef\]](#)
77. Leng, G.; Zeng, X.J.; Keane, J. An improved approach of self-organising fuzzy neural network based on similarity measures. *Evol. Syst.* **2012**, *3*, 19–30. [\[CrossRef\]](#)
78. Lasota, T.; Telec, Z.; Trawinski, B.; Trawinski, K. Investigation of the eTS evolving fuzzy systems applied to real estate appraisal. *J. Mult.-Valued Log. Soft Comput.* **2011**, *17*, 229–253.
79. Iglesias, J.; Angelov, P.; Ledezma, A.; Sanchis, A. Evolving classification of agents' behaviors: A general approach. *Evol. Syst.* **2010**, *1*, 161–171. [\[CrossRef\]](#)
80. Angelov, P.; Kordon, A. Adaptive inferential sensors based on evolving fuzzy models. *IEEE Trans. Syst. Man, Cybern. Part B Cybern.* **2009**, *40*, 529–539. [\[CrossRef\]](#)
81. Ferdaus, M.M.; Dam, T.; Alam, S.; Pham, D.T. X-Fuzz: An Evolving and Interpretable Neurofuzzy Learner for Data Streams. *IEEE Trans. Artif. Intell.* **2024**, 1–12. [\[CrossRef\]](#)
82. Yan, L.; Zhao, T.; Xie, X.; Precup, R.E. OSSEFS: An online semi-supervised ensemble fuzzy system for data streams learning with missing values. *Expert Syst. Appl.* **2024**, *255*, 124695. [\[CrossRef\]](#)
83. Hu, L.; Xu, X.; Ren, W.; Han, M. Hierarchical Evolving Fuzzy System: A Method for Multidimensional Chaotic Time Series Online Prediction. *IEEE Trans. Fuzzy Syst.* **2024**, *32*, 3329–3341. [\[CrossRef\]](#)
84. Mei, Z.; Zhao, T.; Gu, X. A Dynamic Evolving Fuzzy System for Streaming Data Prediction. *IEEE Trans. Fuzzy Syst.* **2024**, 1–14. [\[CrossRef\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Evaluation of the Dynamics of Psychological Panic Factor, Glucose Risk and Estrogen Effects on Breast Cancer Model

Zahraa Aamer ¹, Shireen Jawad ¹, Belal Batiha ², Ali Hasan Ali ^{3,4,*}, Firas Ghanim ⁵ and Alina Alb Lupas ^{6,*}

¹ Department of Mathematics, College of Science, University of Baghdad, Baghdad 10071, Iraq; zahraa.aamer1703a@sc.uobaghdad.edu.iq (Z.A.); shireen.jawad@sc.uobaghdad.edu.iq (S.J.)

² Faculty of Science and Information Technology, Mathematics Department, Jadara University, Irbid 21110, Jordan; b.bateha@jadara.edu.jo

³ Institute of Mathematics, University of Debrecen, Pf. 400, H-4002 Debrecen, Hungary

⁴ Jadara University Research Center, Jadara University, Irbid 21110, Jordan

⁵ Department of Mathematics, College of Sciences, University of Sharjah, Sharjah 27272, United Arab Emirates; fgahmed@sharjah.ac.ae

⁶ Department of Mathematics and Computer Science, University of Oradea, 1 Universitatii Street, 410087 Oradea, Romania

* Correspondence: ali.hasan@science.unideb.hu (A.H.A.); alblupas@gmail.com (A.A.L.)

Abstract: Contracting cancer typically induces a state of terror among the individuals who are affected. Exploring how glucose excess, estrogen excess, and anxiety work together to affect the speed at which breast cancer cells multiply and the immune system's response model is necessary to conceive of ways to stop the spread of cancer. This paper proposes a mathematical model to investigate the impact of psychological panic, glucose excess, and estrogen excess on the interaction of cancer and immunity. The proposed model is precisely described. The focus of the model's dynamic analysis is to identify the potential equilibrium locations. According to the analysis, it is possible to establish four equilibrium positions. The stability analysis reveals that all equilibrium points consistently exhibit stability under the defined conditions. The transcritical bifurcation occurs when the glucose excess is taken as a bifurcation point. Numerical simulations are employed to validate the theoretical study, which shows that psychological panic, glucose excess, and estrogen excess could be significant contributors to the spread of tumors and weakness of immune function.

Keywords: psychological panic; glucose risk; estrogen effect; breast cancer model; stability analysis

Citation: Aamer, Z.; Jawad, S.; Batiha, B.; Ali, A.H.; Ghanim, F.; Lupas, A.A. Evaluation of the Dynamics of Psychological Panic Factor, Glucose Risk and Estrogen Effects on Breast Cancer Model.

Computation **2024**, *12*, 160.

<https://doi.org/10.3390/computation12080160>

Academic Editor: Simeone Marino

Received: 23 May 2024

Revised: 12 July 2024

Accepted: 2 August 2024

Published: 8 August 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

A group of disorders collectively referred to as cancer are those in which cells divide uncontrollably at a faster rate than healthy cells and have the potential to metastasize to other parts of the body [1]. The data indicate that the number of fatalities caused by cancer was approximately 8.2 million in the year 2010. It is anticipated that the mortality rate attributed to cancer will persistently increase, resulting in approximately 13 million fatalities by the year 2030 [2]. Breast cancer is the second most frequent type of cancer in women, second only to skin cancers, among many others. A little under 12% of women will develop invasive breast cancer in their lifetime [3]. There are three primary risk factors for breast cancer: genetic factors (such as family history), hormonal imbalance (namely estrogen), and environmental factors (including alcohol intake, poor diet, exposure to toxins, smoking, etc.) [4]. In order to study the dynamic behavior of cancer cells in the presence of different stimuli, numerous mathematical models have been developed throughout the years. For instance, Dehingia et al. developed a mathematical model to assess the impact of obesity on tumor growth and immune response, considering it as a risk factor in a study demonstrating that the accumulated fat in the body of an obese individual significantly contributes to the formation of tumors, hence elevating the risk

factors associated with survival [5]. An elevation in estrogen levels could potentially expedite the rate of tumor formation [6,7]. Sorofa et al. [8] evaluated how estrogen affects the dynamics of normal immunological cells and breast cancer. The results showed that elevated estrogen levels can slow immune system development and enhance tumor initiation.

It has been demonstrated that the majority of the energy for breast cancer cells comes from glycolysis wherein glucose is transformed to lactate for energy throughout the glycolysis process [9,10]. Multiple studies have verified that cancer cells have a high need for glucose as a source of nourishment. Therefore, the effect of glucose is incorporated into the tumor model to assess its impact on breast cancer cells. For instance, Sun et al. investigated the impact of glucose deprivation on the apoptosis of breast cancer cells in a study demonstrating that consuming an excessive amount of glucose can promote the development and proliferation of cancer cells, whereas restricting glucose consumption can decrease and suppress the growth of cancer cells [11]. A pathological state and impaired functionality of the immune system may result from a high glucose intake. However, maintaining an adequate amount of glucose is critical for the functioning of the immune system. As a result, the infiltration of a substantial amount of glucose into immune cells may have detrimental effects on the immune system. Consequently, this results in compromised immune system functionality, which initiates pathological conditions [12]. Alharbi et al. created a model that offers a fundamental understanding of the process of breast cancer in patients who already have risk factors related to hyperglycemia. They determined that their findings have the potential to be advantageous for the future delivery and treatment of breast cancer drugs [13].

On the other hand, a growing body of literature highlights the use of mathematical models for investigating the effects of anxiety on disease transmission [14,15]. Research in the medical field has shown that emotional distress might hasten the metastasis of cancer cells in a patient. Stress on a psychological level widens and thickens blood vessels, which in turn speeds up the migration of cancer cells and makes disease spreading easier [16]. According to recent research, hormones produced by stress accelerate cancer cell growth in the “lymphatic system”, which in turn makes it easier for these cells to spread to other parts of the body, a process known as metastasis [17–19].

To sum up, most of the prior research has shown that a high-glucose diet and high estrogen levels might promote the growth and proliferation of breast cancer cells, making it more likely that the disease would progress rapidly. The impact of high estrogen levels and glucose on the human immune system, on the other hand, has been the subject of very few investigations. Experimental evidence suggests that immune cells are vulnerable to the harmful effects of either high or low blood glucose levels as well as high estrogen levels [8,20]. Moreover, there is a lack of research regarding the impact of anxiety on the immune cells–breast cancer model. Therefore, we must investigate this phenomenon, as it plays a role in minimizing the occurrence of disastrous situations. Hence, this investigation is focused on examining the influence of anxiety on immune-compromised cancer patients, which may serve as a substantial factor in the proliferation of tumors and the impairment of immune system efficacy.

This research aims to investigate the dynamics of the immune cells–breast cancer model as a result of the combined influence of the psychological panic factor, increased glucose consumption, and high estrogen levels and their harmful effects on normal, cancer, and immune cells. Considering these effects, we propose a PPIGCNE model of psychological panic–immune–glucose–cancer–normal estrogen interaction. This paper’s findings provide additional context for [8] by

- (1) Analyzing the impact of glucose as a risk factor on the progression of breast cancer.
- (2) Examining the influence of psychological panic on the immune system of breast cancer patients, which may play a key role in the progression of malignancies and the impairment of immune function.

Following a description of the model's construction, we intend to determine how the above effects impact the dynamics of a PPIGCNE model. Furthermore, this paper will explore the understanding of the nonlinear dynamics exhibited by our model through the utilization of various methodologies, including stability and bifurcation analysis techniques. In conclusion, we will validate the precision of our analytical findings through a numerical simulation of the proposed system.

2. Assumptions of the Model

In this part, we will construct a model by integrating many experimental investigations and mathematical models [21–23] that elucidate the dynamics of tumor progression. The psychological panic, increased glucose consumption, and high estrogen levels interaction model are formulated. The dynamics of the PPIGCNE model significantly influence the development of breast cancer, making it a substantial risk factor. Furthermore, we analyze the interactions among the above factors with breast tumor cells. The relationships between immune cells $I(t)$, cancer cells $C(t)$, normal (breast) cells $N(t)$, and the effect of estrogen $E(t)$, are depicted in Figure 1.

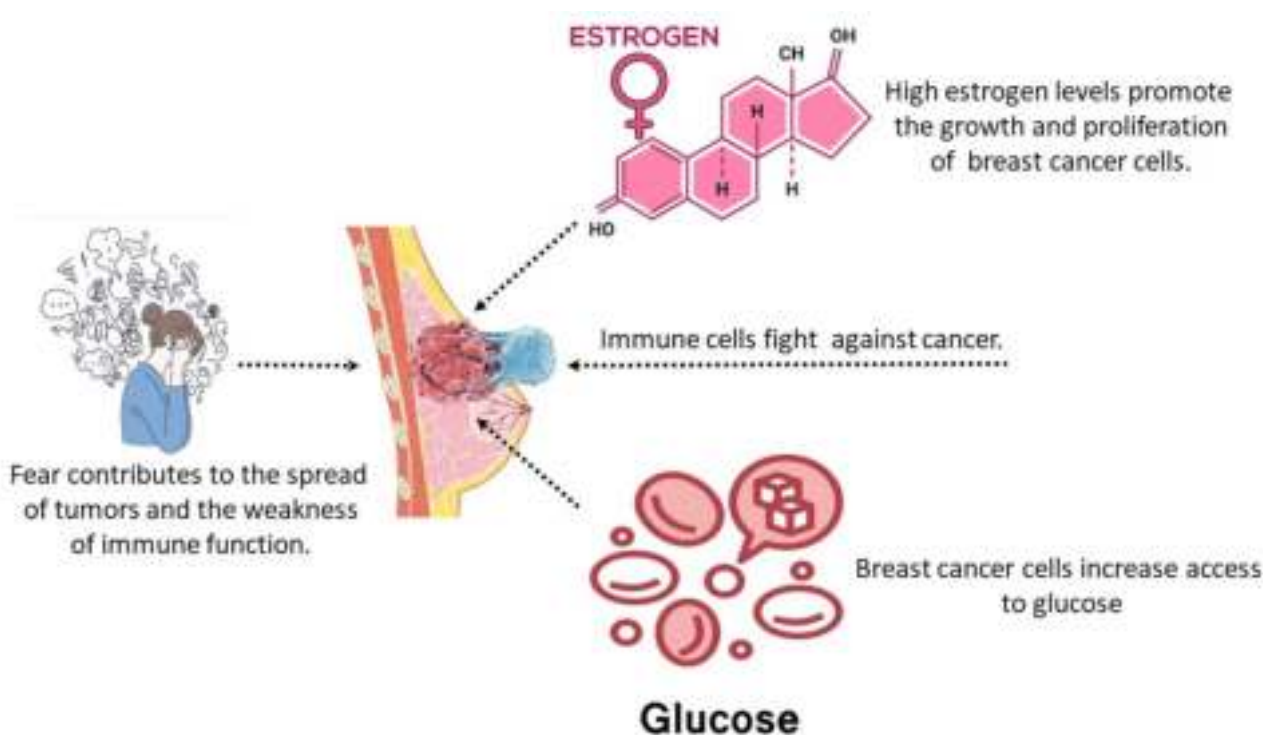


Figure 1. Schematic sketch of the PPIGCNE model.

The response of the immune system is critical for safeguarding breast tissue against malignant tumors and ensuring an appropriate response upon tumor cell recognition. Tumor cell presence stimulates the activation of the immune system. The dynamic representation of immune cells is

$$\frac{dI}{dt} = \frac{\Lambda}{1 + pC} + \frac{e_1 IC}{p_1 + C} - e_2 IC - e_3 IE - (g + \mu_1)I. \quad (1)$$

The term $\Lambda/(1 + pC)$ denotes the consistent generation of immune cells within the organism, which is negatively influenced by the psychological fear factor p in the presence of cancer. The fear function is integrated using the decreasing function $\sigma(p, C) = 1/(1 + pC)$, a concept that Wang et al. [24] first proposed. Given its biological implications, $\sigma(p, C)$ is deemed suitable since

$$\sigma(0, C) = 1, \sigma(p, 0) = 1, \lim_{p \rightarrow \infty} \sigma(p, C) = 0, \lim_{C \rightarrow \infty} \sigma(p, C) = 0, \frac{\partial \sigma(p, C)}{\partial p} < 0, \frac{\partial \sigma(p, C)}{\partial C} < 0.$$

In (1), $e_1IC/(p_1 + C)$, denoted by the Michaelis–Menten term [25], indicates the presence of tumor cells that elicit an immune response [26]. e_2IC denotes the measure of the impact of tumor cells on immune cells by suppressing the immune response with cancer cells [26]. e_3IE refers to the rate of immune suppression induced by estrogen [8]. $(g + \mu_1)I$ refers to the decline of immune cells due to excess glucose and natural death rate [13].

The breast cancer cells $T(t)$ consist of aberrant cells characterized by inflammation and fast proliferation of breast cells, leading to the formation of tumors. Therefore, we define the equation representing the tumor as:

$$\frac{dC}{dt} = \alpha_1 C(1 - k_1 C) - e_4 IC + e_5 NE - \mu_2 C + gC. \quad (2)$$

The term $\alpha_1 C(1 - k_1 C)$ represents the logistic tumor cell's growth function, where α_1 is the breast cancer intrinsic growth rate and k_1 stands for the carrying capacity of breast cancer cells. The term $e_4 IC$ means eradicating cancerous cells by the body's immune system. The outcome of DNA damage caused by estrogen is that normal cells that have been harmed will transform into tumor cells, causing the tumor cell population to increase at a rate of e_5 . This leads to a growth factor of $e_5 NE$ on the tumor cell population [8]. $\mu_2 C$ designates the death rate of cancer breast cells. gC indicates the formation of new cancer cells due to elevated glucose levels within the patient's body [13].

Healthy breast cells possess unmodified DNA that governs the activities of all breast cells during their growth, division, and demise [27]. Healthy breast cells and malignant breast cells compete for nutrition and other resources. Healthy breast cells are depicted by

$$\frac{dN}{dt} = \alpha_2 N(1 - k_2 N) - e_6 NC - e_7 NE. \quad (3)$$

The term $\alpha_2 N(1 - k_2 N)$ represents the logistic healthy breast cell growth function. The term $e_6 NC$ represents the degradation rate of healthy breast cells caused by tumor cells. The decrease in healthy breast cells due to a higher level of estrogen causes the conversion of normal cells into tumor cells, which occurs at a rate of $e_7 NE$ [8].

Estrogen, a female steroid hormone, is produced at low levels by the placenta in women and the testes in men. It serves several important purposes, including facilitating sexual development, controlling a woman's menstrual cycle, and managing the physical changes associated with puberty. Therefore, it is regarded as a crucial hormone for the optimal transformation of the breast tissues. Estrogen is recognized as a risk factor that is linked to the development of breast cancer. Estrogen has a role in activating other hormones, such as relaxin, which in turn stimulates the proliferation of breast cells [28]. Estrogen has promoted the proliferation of tumor cells when tumor cells are present. It functions as a carcinogen by damaging DNA and converting normal epithelial cells into cancerous tumor cells [29,30]. The dynamics of estrogen can be described as follows:

$$\frac{dE}{dt} = s - \mu_3 E. \quad (4)$$

where s represents the rate at which increased estrogen is produced, and μ_3 represents the rate at which estrogen is being washed out from the body system.

Upon careful evaluation of these factors, we propose the subsequent set of differential equations to determine the behavior of breast cancer cells:

$$\begin{aligned}
 \frac{dI}{dt} &= \frac{\Lambda}{1+pC} + \frac{e_1 IC}{p_1 + C} - e_2 IC - e_3 IE - (g + \mu_1)I, \\
 \frac{dC}{dt} &= \alpha_1 C(1 - k_1 C) - e_4 IC + e_5 NE - \mu_2 C + gC, \\
 \frac{dN}{dt} &= \alpha_2 N(1 - k_2 N) - e_6 NC - e_7 NE, \\
 \frac{dE}{dt} &= s - \mu_3 E.
 \end{aligned}
 \tag{5}$$

The model's parameters and their interpretations are clarified in Table 1.

Table 1. The interpretations of the PPIGCNE model's parameters.

Parameter	Interpretation	Values	Unit	Source
Λ	Body's immune cell rate production.	0.5	(cells/day)	[30]
p	Psychological panic rate from cancer.	0.1	dimensionless	Estimated
e_1	Elicitation rate of immune cells by cancer cells.	0.1	1/(cells·day)	[30]
p_1	Half-life of effector cells	0.4	(day)	[30]
e_2	Inactivation rate of immune cells due to the effect of tumor cells.	0.2	1/(cells·day)	[30]
e_3	Inhibition rate of the immune cells due to high levels of estrogen.	0.09	1/(mg/dL·day)	[8]
g	Immune cell suppression rate by high blood glucose.	0.2	1/(mg/dL·day)	[30]
μ_1	Death rate of effector cells.	0.2	(cells/day)	[30]
α_1	Breast cancer intrinsic growth rate.	0.4	(cells/day)	[8]
k_1	Carrying capacity of breast cancer cells.	1.5	(cell)	[8]
e_4	The rate at which effector cells eliminate tumor cells.	0.2	1/(cells·day)	[30]
e_5	The transformation rate of damaged normal cells into tumor cells caused by estrogen.	0.2	1/(ng/mL·day)	[8]
μ_2	Death rate of breast cancer cells.	0.05	(cells/day)	[8]
α_2	Intrinsic growth rate of healthy breast cells.	0.35	(cells/day)	[8]
k_2	Carrying capacity of healthy breast cells.	1	(cell)	[8]
e_6	The degradation rate of healthy breast cells caused by tumor cells.	0.25	1/(cells·day)	[30]
e_7	The decrease in healthy breast cells caused by a higher estrogen level.	0.1	1/(ng/mL·day)	[29]
s	The rate of higher estrogen production.	0.19	ng/mL/day	[8]
μ_3	The wash-out rate of estrogen from the body.	0.05	ng/mL/day	[8]

3. Dynamical Evaluation Results

3.1. Positivity and Boundedness of the Solution

In the following theorems, the positivity and uniformly bounded of all the solutions of the PPIGCNE model in R_+^4 are confirmed.

Theorem 1. All of the solutions of the PPIGCNE model $I(t), C(t), N(t)$ and $E(t)$ with the initial conditions $(I(0), C(0), N(0), E(0)) \in R_+^4$ are positively invariant.

Proof of Theorem 1. Let $L = (I, C, N, E)^T \in R_+^4$ and $F(L) = [f_1(L), f_2(L), f_3(L), f_4(L)]^T$, where $F(L): R_+^4 \rightarrow R^4$ and $f_i \in C_+^\infty(R_+^4), i = 1, 2, 3, 4$. Then the PPIGCNE system becomes:

$$\dot{L} = F(L),$$

with $L(0) = (I(0), C(0), N(0), E(0)) = L_0$. It is clear for any $L(0) \in R_+^4$, such that $L_i = 0$, then $[f_i(L)]_{L_i=0} \geq 0$ (for $i = 1, 2, 3, 4$). Now for any solution of $\dot{L}$ with $L_0 \in R_+^4$, say $L(t) = L(t; L_0)$, is such that $L(t) \in R_+^4$, for all $t > 0$. Thus, the PPIGCNE system is positively invariant [31]. $\square$

Theorem 2. All the solutions of the PPIGCNE model are uniformly bounded.

Proof of Theorem 2. Let $(I(0), C(0), N(0), E(0)) \in R_+^4$ be an initial condition for the PPIGCNE, then, by using the Bernoulli method, we obtain

$$\frac{dN}{dt} = \alpha_2 N(1 - k_2 N) - e_6 NC - e_7 NE \leq \alpha_2 N(1 - k_2 N) \Rightarrow N(t) \leq \frac{1}{k_2 + N(0)e^{-\alpha_2 t}}$$

Thus,

$$\limsup_{t \rightarrow \infty} [N(t)] \leq \frac{1}{k_2}.$$

Using the standard comparison theory [32] on the last equation of the PPIGCNE, we obtain

$$\limsup_{t \rightarrow \infty} [E(t)] \leq \frac{s}{\mu_3}$$

From the second equation of the PPIGCNE, we have

$$\frac{dC}{dt} = \alpha_1 C(1 - k_1 C) - e_4 IC + e_5 NE - \mu_2 C + gC \leq \alpha_1 C - \alpha_1 k_1 C^2 + gC.$$

Using the Bernoulli method, we obtain

$$C(t) \leq \frac{\alpha_1 + g}{\alpha_1 k_1 + (\alpha_1 + g)N(0)e^{-(\alpha_1 + g)t}}$$

Thus,

$$\limsup_{t \rightarrow \infty} [C(t)] \leq \frac{\alpha_1 + g}{\alpha_1 k_1}.$$

Now, by using the standard comparison theory [32], we obtain

$$\frac{dI}{dt} = \frac{\Lambda}{1 + pC} + \frac{e_1 IC}{p_1 + C} - e_2 IC - e_3 IE - (g + \mu_1)I \leq \Lambda - (g + \mu_1)I$$

$$\limsup_{t \rightarrow \infty} [I(t)] \leq \frac{\Lambda}{(g + \mu_1)}.$$

Therefore, the corresponding domain region for the PPIGCNE model is

$$\varphi = \left\{ (I, C, N, E) \in R_+^4 : I(t) \leq \frac{\Lambda}{(g + \mu_1)}, C(t) \leq \frac{\alpha_1 + g}{\alpha_1 k_1}, N(t) \leq \frac{1}{k_2}, E(t) \leq \frac{s}{\mu_3} \right\}.$$

□

3.2. Existence of Equilibria

Resetting the left-hand side of the PPIGCNE model to zero will bring the solutions of the model equations to equilibrium:

1. Tumor and breast cells-free equilibrium point $L_1 = (\bar{I}, 0, 0, E^*)$, where $\bar{I} = \frac{\Lambda \mu_3}{e_3 s + (g + \mu_1) \mu_3}$ and $E^* = \frac{s}{\mu_3}$. It is important to mention that the expression E^* remains constant for all equilibrium points that have E^* in their composition.
2. Tumor-free equilibrium point $L_2 = (\hat{I}, 0, \hat{N}, E^*)$, where $\hat{I} = \frac{\Lambda s}{e_3 \mu_3 + (g + \mu_1) s}$, and $\hat{N} =$

$\frac{\alpha_2 s - e_7 \mu_3}{\alpha_2 k_2}$. For $\hat{N} > 0$, we must have

$$\alpha_2 > \frac{e_7 \mu_3}{s} \quad (6)$$

In biological terms, this point refers to a healthy case because the tumor vanishes while the immune system and normal cells remain. The following justifies this point. The low estrogen level causes the tumor environment to be improper for growth, hence promoting the eradication of cancer cells by immune cells.

3. Breast cells-free equilibrium point $L_3 = (\check{I}, \check{C}, 0, E^*)$, where $\check{I} = \frac{\alpha_1 - \alpha_1 k_1 \check{C} - \mu_2 + g}{e_4}$ and $\check{C}$ is the root of the following equation:

$$f(C) = N_1 C^4 + N_2 C^3 + N_3 C^2 + N_4 C + N_5,$$

where

$$N_1 = \alpha_1 k_1 e_2 p,$$

$$N_2 = -\alpha_1 e_2 p s + \alpha_1 k_1 e_2 s + \alpha_1 k_1 e_2 p_1 p s + \alpha_1 k_1 e_3 \mu_3 p + \mu_2 e_2 p - g e_2 p,$$

$$N_3 = \alpha_1 e_1 p s - \alpha_1 e_2 s - \alpha_1 e_2 p_1 p s - \alpha_1 e_3 \mu_3 p - \alpha_1 k_1 e_1 s + \alpha_1 e_2 k_1 p_1 s - \alpha_1 k_1 e_3 \mu_3$$

$$+ \alpha_1 k_1 \mu_3 p_1 p - \mu_2 e_1 p s + \mu_2 e_2 s + \mu_2 e_2 p_1 p s + \mu_2 e_3 \mu_3 p + e_1 g p s - e_2 g s$$

$$- e_2 p_1 g p s - e_3 \mu_3 g p,$$

$$N_4 = \Lambda e_4 s + \alpha_1 e_1 s - \alpha_1 e_2 p_1 s - \alpha_1 e_3 \mu_3 - \alpha_1 e_3 \mu_3 p_1 p + \alpha_1 k_1 e_3 p_1 \mu_3 - \mu_2 e_1 s + \mu_2 e_2 p_1 s$$

$$+ \mu_2 e_3 \mu_3 + \mu_2 e_3 \mu_3 p_1 p - e_1 g s - e_2 p_1 g s - e_3 \mu_3 g - e_3 \mu_3 p_1 g p,$$

$$N_5 = [\Lambda e_4 s - e_3 \mu_3 (\alpha_1 + \mu_2 + g)] p_1. \text{ Clearly,}$$

$$\begin{aligned} f(0) &= N_5, \\ f(k_1) &= N_1 k_1^4 + N_2 k_1^3 + N_3 k_1^2 + N_4 k_1 + N_5, \\ f'(C) &= 4N_1 C^3 + 3N_2 C^2 + 2N_3 C + N_4. \end{aligned}$$

Thus, if any of the following criteria holds, then $f(C)$ has a unique positive root, say $C = \check{C}$, in the interval $(0, k_1)$ according to the intermediate value theorem

$$f(0) > 0, f(k_1) < 0 \text{ and } f'(C) < 0, \quad (7)$$

$$f(0) < 0, f(k_1) > 0 \text{ and } f'(C) > 0 \quad (8)$$

For $\check{I} > 0$, we must have

$$g > g_1, \quad (9)$$

where $g_1 = \alpha_1(k_1 \check{C} - 1) + \mu_2$. Biologically, condition (9) means this point is classed as a glucose risk. This is because the level of glucose (g) is higher than the rate of tumor growth, which is a factor that leads to the growth of cancer cells at a rate that is significantly higher than that of normal cells. As a result of the spread of the tumor, the patient is placed in a severe situation, which ultimately results in either a mastectomy or the patient's death.

4. The coexisting point $L_4 = (I^*, C^*, N^*, E^*)$, where $N^* = \frac{\alpha_2 \mu_3 - e_6 \mu_3 C^* - e_7 s}{\alpha_2 k_2}$, $I^* =$

$$\frac{b_1 C^{*2} + b_2 C^* + b_3}{e_4 C^*}, \text{ and}$$

$$b_1 = -\alpha_1 \alpha_2 k_1 k_2,$$

$$b_2 = \alpha_2 k_2 (\alpha_1 + g - \mu_2) - e_5 e_6,$$

$$b_3 = -e_5 \mu_3 (e_6 + e_7), \text{ and } C^* \text{ is a root of the following equation}$$

$$f(C) = l_1 C^5 + (l_1 + l_2 p) C^4 + (l_2 + l_3 p) C^3 + (l_3 + l_5 p) C^2 + (\Lambda e_4 + p l_4 + l_5) C + l_4$$

where

$$l_1 = -p e_2 b_1,$$

$$l_2 = -(p_1 e_2 + g) b_1 - \frac{s e_3}{\mu_3} - (\mu_1 + e_2) b_2,$$

$$l_3 = b_1 (e_1 - p_1 e_3 \frac{s}{\mu_3} - g p_1 - p_1 \mu_1) + b_2 (e_1 - p_1 e_2 - e_3 \frac{s}{\mu_3} - g - \mu_1) - e_2 b_3,$$

$$l_4 = -b_3 (p_1 e_3 \frac{s}{\mu_3} + g p_1 + p_1 \mu_1),$$

$$l_5 = -b_2 (p_1 e_3 \frac{s}{\mu_3} + g p_1 + p_1 \mu_1) + b_3 (e_1 - p_1 e_2 - e_3 \frac{s}{\mu_3} - g - \mu_1).$$

Clearly,

$$f(0) = l_4 > 0$$

$$f(k_1) = l_1 k_1^5 + (l_1 + l_2 p) k_1^4 + (l_2 + l_3 p) k_1^3 + (l_3 + l_5 p) k_1^2 + (\Lambda e_4 + p l_4 + l_5) k_1 + l_4$$

$$f'(C) = 5 l_1 C^4 + 4 (l_1 + l_2 p) C^3 + 3 (l_2 + l_3 p) C^2 + 2 (l_3 + l_5 p) C + (\Lambda e_4 + p l_4 + l_5).$$

Thus, if the following criteria hold, then $f(C)$ has a unique positive root, say $C = C^*$, in the interval $(0, k_1)$ according to the intermediate value theorem

$$f(k_1) < 0 \text{ and } f'(C) < 0, \quad (10)$$

For $N^* > 0$ and $I^* > 0$, we must have

$$\alpha_2 \mu_3 > e_6 \mu_3 C^* + e_7 s, \quad (11)$$

$$b_1 C^{*2} + b_2 C^* + b_3 > 0. \quad (12)$$

The coexisting point signifies the stage of interaction among immune cells, malignant, normal, and estrogen. Every cell engages in a fierce struggle for survival during this phase. Tumor appearance triggers the activation of immune cells.

3.3. Stability Analysis

The local stability behavior around the previous equilibrium points is examined in this section. The Jacobian matrix of PPIGCNE model can be expressed as follows:

$$J = \begin{bmatrix} \frac{\partial f_1}{\partial I} & \frac{\partial f_1}{\partial C} & \frac{\partial f_1}{\partial N} & \frac{\partial f_1}{\partial E} \\ \frac{\partial f_2}{\partial I} & \frac{\partial f_2}{\partial C} & \frac{\partial f_2}{\partial N} & \frac{\partial f_2}{\partial E} \\ \frac{\partial f_3}{\partial I} & \frac{\partial f_3}{\partial C} & \frac{\partial f_3}{\partial N} & \frac{\partial f_3}{\partial E} \\ \frac{\partial f_4}{\partial I} & \frac{\partial f_4}{\partial C} & \frac{\partial f_4}{\partial N} & \frac{\partial f_4}{\partial E} \end{bmatrix} = (a_{ij})_{4 \times 4},$$

where $a_{11} = \frac{e_1 C}{(p_1 + C)} - e_2 C - e_3 E - (g + \mu_1)$, $a_{12} = \frac{-\Lambda p}{(1 + pC)^2} + \frac{p_1 e_1 C}{(p_1 + C)} - e_2 I$,
 $a_{13} = 0$, $a_{14} = -e_3 I$, $a_{21} = -e_4 C$, $a_{22} = \alpha_1 - 2\alpha_1 k_1 C - e_4 I - \mu_2 + g$,
 $a_{23} = e_5 E$, $a_{24} = e_5 N$, $a_{31} = 0$, $a_{32} = e_6 N$, $a_{33} = \alpha_2 - 2\alpha_2 k_2 N - e_6 C - e_7 E$,
 $a_{34} = -e_7 N$, $a_{41} = a_{42} = a_{43} = 0$, $a_{44} = -\mu_3$.

The local analysis of the PPIGCNE system at the three equilibrium points mentioned above is computed as

(1) The Jacobian matrix at $L_1 = (\bar{I}, 0, 0, E^*)$ is given as:

$$J(L_1) = \begin{bmatrix} -e_3 E^* - (g + \mu_1) & \frac{e_1 \bar{I}}{p_1} - e_2 \bar{I} & 0 & e_3 I \\ 0 & \alpha_1 - e_4 I - \mu_2 + g & e_5 E^* & 0 \\ 0 & 0 & \alpha_2 - e_7 E^* & 0 \\ 0 & 0 & 0 & -\mu_3 \end{bmatrix}$$

The eigenvalues of $J(L_1)$ are $\lambda_{11} = -e_3 E^* - (g + \mu_1) < 0$, $\lambda_{12} = \alpha_1 - e_4 I - \mu_2 + g$,
 $\lambda_{13} = \alpha_2 - e_7 E^*$, and $\lambda_{14} = -\mu_3 < 0$. Then, L_1 is a locally asymptotic stable if

$$\alpha_2 < e_7 E^*, \quad (13)$$

$$g < g_2 \quad (14)$$

where $g_2 = e_4 \bar{I} + \mu_2 + \alpha_1$. Condition (13) shows a mastectomy could happen when the rate of decrease in healthy breast cells caused by a higher estrogen level is greater than the healthy breast cell's intrinsic growth rate. Further, from a biological standpoint, condition (14) means that cancer cells are dependent on their capacity to absorb glucose and reproduce. Consequently, a glucose rate that is lower than g_2 will substantially inhibit cancer cells, thereby reducing their strength and ability to reproduce and spread. Conversely, for $g > g_2$, L_1 is a saddle point. For $g = g_2$, then $J(L_1)$ has zero eigenvalue, making L_1 a nonhyperbolic point.

(2) The Jacobian matrix at $L_2 = (\hat{I}, 0, \hat{N}, E^*)$ is given as:

$$J(L_2) = \begin{bmatrix} -e_3 E^* - (g + \mu_1) & -\Lambda p + \frac{e_1 \hat{I}}{p_1} - e_2 \hat{I} & 0 & -e_3 \hat{I} \\ 0 & \alpha_1 - e_4 \hat{I} - \mu_2 + g & e_5 E^* & e_5 \hat{N} \\ 0 & e_6 \hat{N} & \alpha_2 - 2\alpha_2 k_2 \hat{N} - e_7 E^* & -e_7 \hat{N} \\ 0 & 0 & 0 & -\mu_3 \end{bmatrix}$$

The eigenvalues of $J(L_2)$ are

$$\lambda_{21} = -e_3 E^* - (g + \mu_1) < 0,$$

$$\lambda_{22} + \lambda_{23} = (\alpha_1 - e_4 \hat{I} - \mu_2 + g + \alpha_2 - 2\alpha_2 k_2 \hat{N} - e_7 E^*),$$

$$\lambda_{22} \cdot \lambda_{23} = \alpha_1 \alpha_2 - 2\alpha_1 \alpha_2 k_2 \hat{N} - \alpha_2 e_7 E^* - \alpha_2 e_4 \hat{I} - 2\alpha_2 e_4 \hat{I} \hat{N} + e_4 e_7 \hat{I} E^* - \mu_2 \alpha_2 + 2\alpha_2 \mu_2 k_2 \hat{N} + e_7 \mu_2 E^* + \alpha_2 g - 2\alpha_2 k_2 g \hat{N} - e_7 g E^* - e_5 e_6 \hat{N} E^*,$$

$$\lambda_{24} = -\mu_3 < 0.$$

That means L_2 is a locally asymptotical stable point if, and only if, the following conditions are satisfied:

$$\left. \begin{array}{l} g < g_3 \\ \lambda_{22} \cdot \lambda_{23} > 0 \end{array} \right\} \quad (15)$$

where $g_3 = 2\alpha_2 k_2 \hat{N} + e_7 E^* + e_4 \hat{I} + \mu_2 - (\alpha_1 + \alpha_2)$.

(3) The Jacobian matrix at $L_3 = (\check{I}, \check{C}, 0, E^*)$ is given as:

$$J(L_3) = \begin{bmatrix} \frac{e_1 \check{C}}{p_1 + \check{C}} - e_2 \check{C} - e_3 E^* - (g + \mu_1) & \frac{\Lambda p}{(1 + p \check{C})^2} + \frac{p_1 e_1 \check{I}}{(p_1 + \check{C})^2} - e_2 \check{I} & 0 & -e_3 \check{I} \\ e_4 \check{C} & \alpha_1 - 2\alpha_1 k_1 \check{C} - e_4 \check{I} - \mu_2 + g & e_3 E^* & 0 \\ 0 & 0 & \alpha_2 - e_6 \check{C} - e_7 E^* & 0 \\ 0 & 0 & 0 & -\mu_3 \end{bmatrix}$$

Then, $J(L_3)$ has the following eigenvalues

$$\lambda_{31} + \lambda_{32} = \left(\frac{e_1}{p_1 + \check{C}} + e_2 + 2\alpha_1 k_1 \right) \check{C} + 2g + \mu_1 + \alpha_1 - (e_3 E^* + e_4 \check{I} + \mu_2),$$

$$\begin{aligned} \lambda_{31} \cdot \lambda_{32} = & \left(\frac{e_1 \check{C}}{p_1 + \check{C}} - e_2 \check{C} - e_3 E^* - g - \mu_1 \right) (\alpha_1 - 2\alpha_1 k_1 \check{C} - e_4 \check{I} - \mu_2 + g) - \frac{\Lambda p e_4 \check{C}}{(1 + p \check{C})^2} \\ & + \frac{p_1 e_1 e_4 \check{I} \check{C}}{(p_1 + \check{C})^2} + e_2 e_4 \check{I} \check{C}, \end{aligned}$$

$$\lambda_{33} = \alpha_2 - e_6 \check{C} - e_7 E^*,$$

$$\lambda_{43} = -\mu_3 < 0.$$

That means L_3 is a locally asymptotical stable point if, and only if, the following conditions hold:

$$\lambda_{31} + \lambda_{32} < 0, \text{ and } \lambda_{31} \cdot \lambda_{32} > 0, \quad (16)$$

$$\alpha_2 > e_6 \check{C} + e_7 E^*. \quad (17)$$

(4) The Jacobian matrix at $L_4 = (I^*, C^*, N^*, E^*)$ can be written as:

$$J(L_4) = \begin{bmatrix} b_{11} & b_{12} & 0 & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ 0 & b_{32} & b_{33} & b_{34} \\ 0 & 0 & 0 & -\mu_3 \end{bmatrix},$$

where

$$b_{11} = \frac{e_1 C^*}{p_1 + C^*} - e_2 C^* - e_3 E^* - g - \mu_1, b_{12} = \frac{-\Lambda p}{(1 + p C^*)^2} + \frac{p_1 e_1 I^*}{(p_1 + C^*)^2} - e_2 I^*,$$

$$b_{14} = -e_3 I^*, b_{21} = -e_4 C^*, b_{22} = \alpha_1 - 2\alpha_1 k_1 C^* - e_4 I^* - \mu_2 + g, b_{23} = e_5 E^*,$$

$$b_{24} = e_5 N^*, b_{32} = -e_6 N^*, b_{33} = \alpha_2 - 2\alpha_2 k_2 N^* - e_6 C^* - e_7 E^*, b_{34} = -e_7 N^*.$$

So, the characteristic equation of $J(L_4)$ can be written as:

$$(-\mu_3 - \lambda)(\lambda^3 + B_1 \lambda^2 + B_2 \lambda + B_3) = 0, \quad (18)$$

where

$$B_1 = -(b_{11} + b_{22} + b_{33}),$$

$$B_2 = b_{11}(b_{23} + b_{32} - b_{22} - b_{11}b_{33}) - b_{22}b_{33} + b_{23}b_{32},$$

$$B_3 = b_{11}(b_{22}b_{33} - b_{23}b_{32}),$$

$$\begin{aligned} \Delta = B_1 B_2 - B_3 = & -(b_{11} + b_{22} + b_{33})[b_{11}(b_{23} + b_{32} - b_{22} - b_{11}b_{33}) - b_{22}b_{33} + b_{23}b_{32}] \\ & - (b_{11}(b_{22}b_{33} - b_{23}b_{32})). \end{aligned}$$

Thus, according to the Routh–Hurwitz rule, L_4 will be asymptotically stable if

$$B_1 > 0, B_2 > 0, \text{ and } \Delta > 0. \quad (19)$$

To reach a healthy state, we will examine the global stability surrounding L_2 to explore the dynamics of the PPIGCNE system at regions far from the equilibrium point L_2 .

Theorem 3. L_2 is a global asymptotic stable provided the following conditions hold:

$$\alpha_2 k_2 \geq \max \left\{ \frac{4e_6^2}{\alpha_1 k_1}, \frac{4e_7^2}{\mu_3} \right\}, \quad g \geq g_4 \quad (20)$$

$$\text{where } g_4 = \max \left\{ \frac{4 \left(\frac{e_1 \hat{I}}{p_1 + C} - \frac{\Lambda p}{1 + pC} - e_2 \hat{I} \right)^2}{\alpha_1 k_1} - e_3 E^* - \mu_1, \frac{4e_3^2}{\mu_3} - e_3 E^* - \mu_1 \right\}.$$

Proof of Theorem 3. Let us contemplate the positive definite function given below:

$$Y = \frac{(I - \hat{I})^2}{2} + C + \left(N - \hat{N} - \hat{N}_1 \ln \frac{N}{\hat{N}} \right) + \frac{(E - E^*)^2}{2}.$$

Thus,

$$\begin{aligned} \frac{dY}{dt} = & \left(\frac{-\Lambda p}{1 + pC} + \frac{e_1 I}{p_1 + C} - e_2 I \right) C(I - \hat{I}) - e_3(E - E^*)(I - \hat{I}) - (I - \hat{I})^2(e_3 E^* + g + \mu_1) \\ & + \alpha_1 k_1 C^2 - e_4 IC + e_5 N(E - E^*) + e_5 E^*(N - \hat{N}) - \mu_2 + g \\ & - \alpha_2 k_2 (N - \hat{N})^2 - e_6 C(N - \hat{N}) - e_7(E - E^*)(N - \hat{N}) - \mu_3(E - E^*)^2 \end{aligned}$$

Therefore,

$$\begin{aligned} \frac{dY}{dt} \leq & - \left[\sqrt{\frac{e_3 E^* + g + \mu_1}{2}} (I - \hat{I}) + \sqrt{\frac{\mu_3}{2}} (E - E^*) \right]^2 \\ & - \left[\sqrt{\frac{\alpha_2 k_2}{2}} (N - \hat{N}) + \sqrt{\frac{\mu_3}{2}} (E - E^*) \right]^2 \\ & - \left[\sqrt{\frac{e_3 E^* + g + \mu_1}{2}} (I - \hat{I}) + \sqrt{\frac{\alpha_1 k_1}{2}} C \right]^2 \\ & - \left[\sqrt{\frac{\alpha_2 k_2}{2}} (N - \hat{N}) + \sqrt{\frac{\mu_3}{2}} (E - E^*) \right]^2 + e_5 N(E - E^*) \\ & + e_5 E^*(N - \hat{N}) + g - e_4 IC - \mu_2. \end{aligned}$$

Then, $\frac{dY}{dt} < 0$ can be transformed into a negative definite form under conditions (20). Hence, Y is a Lyapunov function and L_2 is a GAS. $\square$

Thus, the cancer-free steady state L_2 fulfills the requirements for local stability, rendering the point globally stable. From a biological perspective, the immune system refers to the process of selectively eliminating tumor cells if conditions (20) are met.

3.4. Transcritical Bifurcation

A transcritical bifurcation occurs when two equilibrium points collide and exchange their stability. In the following theorems, we will investigate the possibility of a transcritical split occurring. Many scholars use Sotomayor's theorem to determine the existence of transcritical bifurcation TB, for instance, see [31–38]. For this determination, the PPIGCNE model can be rephrased in the following vector forms:

$$\frac{dL}{dt} = F(L) \text{ with } L = \begin{pmatrix} I \\ C \\ N \\ E \end{pmatrix}, \text{ and } F = \begin{pmatrix} f_1(I, C, N, E) \\ f_2(I, C, N, E) \\ f_3(I, C, N, E) \\ f_4(I, C, N, E) \end{pmatrix},$$

where $f_i, i = 1, 2, 3, 4$ are the equations on the right-hand side of the PPIGCNE system. We possess the subsequent outcomes concerning the local bifurcation around the breast cells-free equilibrium point L_3 . Now, for non-zero vector $W = (w_1, w_2, w_3, w_4)^T$:

$$D^2(W, W) = \begin{bmatrix} 2 \left[\frac{p_1 e_1}{(p_1 + C)^2} - e_2 \right] w_1 w_2 - e_3 w_1 w_4 + \left[\frac{2\Lambda p^2}{(1 + pC)^3} - \frac{2p_1 e_1 I}{(p_1 + C)^3} \right] w_2^2 \\ - 2e_4 w_1 w_2 - 2\alpha_1 k_1 w_2^2 + 2e_5 w_3 w_4 \\ - e_6 w_2 w_3 - 2\alpha_2 k_2 w_3^2 - 2e_7 w_3 w_4 \\ 0 \end{bmatrix},$$

The following theorem determines the TB of the PPIGCNE model near L_3 .

Theorem 4. For $g = g^*$, the PPIGCNE model, at L_3 has a transcritical bifurcation if the following conditions are met

$$\left. \begin{aligned} e_4 \check{I}^2 &= x_{11} \check{C} \\ w_1^{[1]} &\neq \frac{x_{11}}{e_4 \check{I}} \\ (B)^T [D^2 F_g(L_3, g^*)(W, W)] &\neq 0 \end{aligned} \right\}. \quad (21)$$

Proof of Theorem 4. The PPIGCNE model, at L_3 , has a zero eigenvalue, say λ_{31} at $g = g^*$, where g^* can be calculated using the subsequent equation:

$$\lambda_{31} \cdot \lambda_{32} = \left(\frac{e_1 \check{C}}{p_1 + \check{C}} - e_2 \check{C} - e_3 E^* - g - \mu_1 \right) (\alpha_1 - 2\alpha_1 k_1 \check{C} - e_4 \check{I} - \mu_2 + g) - \frac{\Lambda p e_4 \check{C}}{(1 + p\check{C})^2} + \frac{p_1 e_1 e_4 \check{I} \check{C}}{(p_1 + \check{C})^2} + e_2 e_4 \check{I} \check{C} = 0.$$

Here, $\lambda_{31} \cdot \lambda_{32}$ is given in the local stability of L_3 . The Jacobian matrix $J^*(L_3) = J(L_3, g^*)$, becomes:

$$J^*(L_3) = \begin{bmatrix} x_{11} & \frac{-\Lambda p}{(1 + p\check{C})^2} + \frac{p_1 e_1 \check{I}}{(p_1 + \check{C})^2} - e_2 \check{I} & 0 & -e_3 \check{I} \\ -e_4 \check{C} & x_{22} & e_5 E^* & 0 \\ 0 & 0 & \alpha_2 - e_6 \check{C} - e_7 E^* & 0 \\ 0 & 0 & 0 & -\mu_3 \end{bmatrix},$$

where $x_{11} = \frac{e_1 \check{C}}{p_1 + \check{C}} - e_2 \check{C} - e_3 E^* - (g^* + \mu_1)$ and $x_{22} = \alpha_1 - 2\alpha_1 k_1 \check{C} - e_4 \check{I} + \mu_2 - g^*$. Now, let $W = (w_1^{[1]}, w_2^{[1]}, w_3^{[1]}, w_4^{[1]})^T$ and $B = (b_1^{[1]}, b_2^{[1]}, b_3^{[1]}, b_4^{[1]})^T$ represent the eigenvectors corresponding to the zero eigenvalue of $J^*(L_3)$ and $J^{*T}(L_3)$, respectively. Direct computation gives $W = \left(\frac{\alpha_1 - 2\alpha_1 k_1 \check{C} - e_4 \check{I} + \mu_2 - g^*}{e_4 \check{C}}, 1, 0, 0 \right)^T$ and $B = \left(1, \frac{x_{11}}{e_4 \check{I}}, \frac{-e_5 E^* x_{11}}{(\alpha_2 - e_6 \check{C} - e_7 E^*) e_4 \check{I}}, \frac{-e_3 \check{I}}{\mu_3} \right)^T$, where $\alpha_2 - e_6 \check{C} - e_7 E^* \neq 0$. Then,

$$\frac{\partial F}{\partial g} = F_g(L, g) = \left(\frac{\partial f_1}{\partial g}, \frac{\partial f_2}{\partial g}, \frac{\partial f_3}{\partial g}, \frac{\partial f_4}{\partial g} \right)^T = (-I, C, 0, 0)^T.$$

So, $F_g(L_3, g^*) = (-I, \check{C}, 0, 0)^T$ and hence,

$$(B)^T F_g(L_3, g^*) = \left(1, \frac{x_{11}}{e_4 \check{I}}, \frac{-e_5 E^* x_{11}}{(\alpha_2 - \check{e}_6 \check{C} - e_7 E^*) e_4 \check{I}}, \frac{-e_3 \check{I}}{\mu_3} \right)^T (-I, \check{C}, 0, 0)^T = -\check{I} + \frac{x_{11} \check{C}}{e_4 \check{I}}.$$

$$(B)^T [DF_g(L_3, g^*)W] = \left(1, \frac{x_{11}}{e_4 \check{I}}, \frac{-e_5 E^* x_{11}}{(\alpha_2 - \check{e}_6 \check{C} - e_7 E^*) e_4 \check{I}}, \frac{-e_3 \check{I}}{\mu_3} \right)^T (-w_1^{[1]}, 1, 0, 0)^T = -w_1^{[1]} + \frac{x_{11}}{e_4 \check{I}}.$$

$$(B)^T [D^2 F_g(L_3, g^*)(W, W)] = 2 \left(w_1^{[3]} \left(\frac{p_1 e_1}{(p_1 + \check{C})^2} - e_2 \right) - b_2^{[3]} (e_4 + \alpha_1 k_1) \right) + \left(\frac{2\Lambda p^2}{(1+p\check{C})^3} - \frac{2p_1 e_1 \check{I}}{(p_1 + \check{C})^3} \right).$$

Therefore, transcritical bifurcation requirements are met under condition 21. $\square$

3.5. Numerical Simulation and Discussions

Numerical verification is essential for completing analytical studies. In this section, we visually confirmed the accuracy of our analytical findings for the PPIGCNE system using the software MATLAB [39–44]. The simulations were conducted using the parameter values specified in Table 1. The equilibrium values corresponding to the data given in Table 1 are $I^* = 1.74, C^* = 0.41, N^* = 0.59$ and $E^* = 0.37$. See Figure 2.

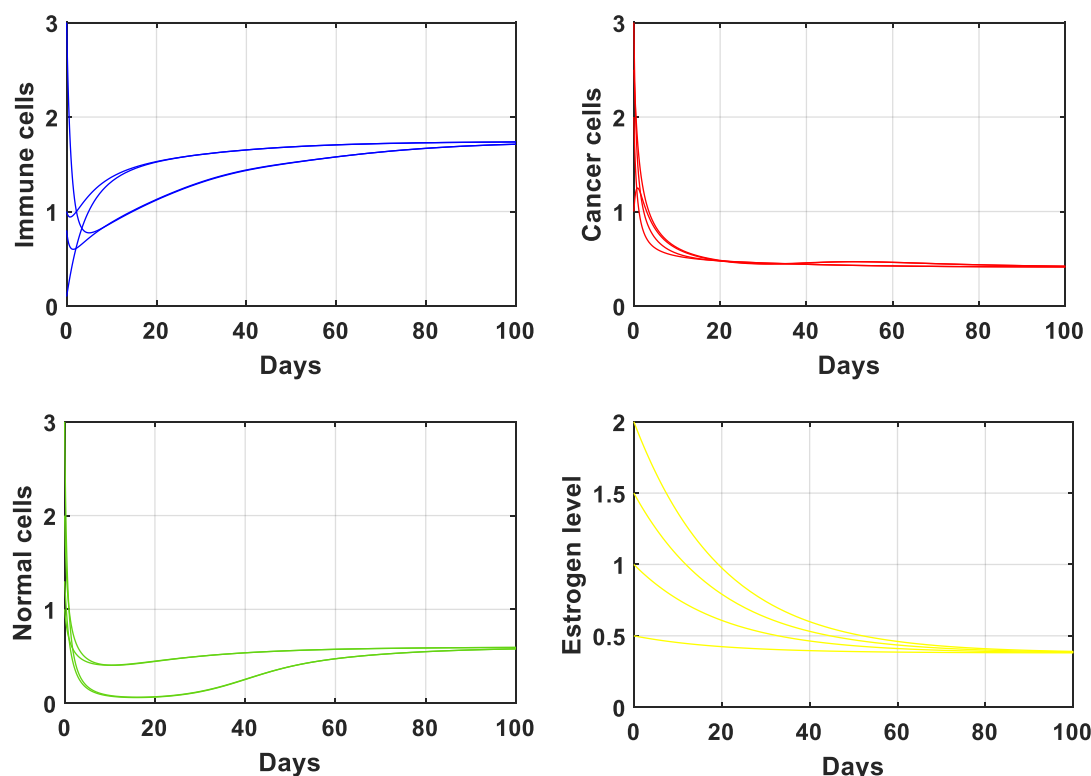


Figure 2. The solution of the PPIGCNE system with the data is given in Table 1. The initial conditions of immune cells, breast cancer cells, normal cells, and estrogen level are $I(0) = 0.25, 0.75, 1.3, 2.7$ (cells/ng), $C(0) = 0.8, 1.4, 1.6, 2.4$ (cells/ng), $N(0) = 0.94, 0.84$ (cells/ng), $0.74, 1.4$, and $E(0) = 0.5, 1.1, 1.5, 2$ ng/mL.

We will now examine three scenarios in order to understand the dynamic behavior of the PPIGCNE model and evaluate the impact of the psychological scare, high levels of estrogen, and high levels of glucose on tumor progression. The results of the three cases will then be compared. The following three cases are:

(a) The effect of a psychological panic

The objective of this case is to demonstrate the impact of psychological scare levels on the interaction between immune cells and breast cancer cells. Figure 3 explains the performance of the PPIGCNE model with various values of the psychological panic rate from cancer, i.e., p . Psychological panic has been discovered to have a negative impact on the immune system's efficiency. Increased stress leads to a significant decline in immune cells. As a result, the tumor cells grow faster. Notwithstanding the notable decline in immune system efficacy, the PPIGCNE model, despite the excess in psychological scare levels, did not attain chaos, but it came close asymptotically to the coexisting point L_4 . Furthermore, Figure 4 illustrates the impact of psychological distress on the functionality of immune cells and the rate of malignant cell growth and division with more precisely rendered. To mitigate the health risks faced by breast cancer patients, psychiatric treatment is needed. Further, some external strategies are needed to improve the immune system's performance.

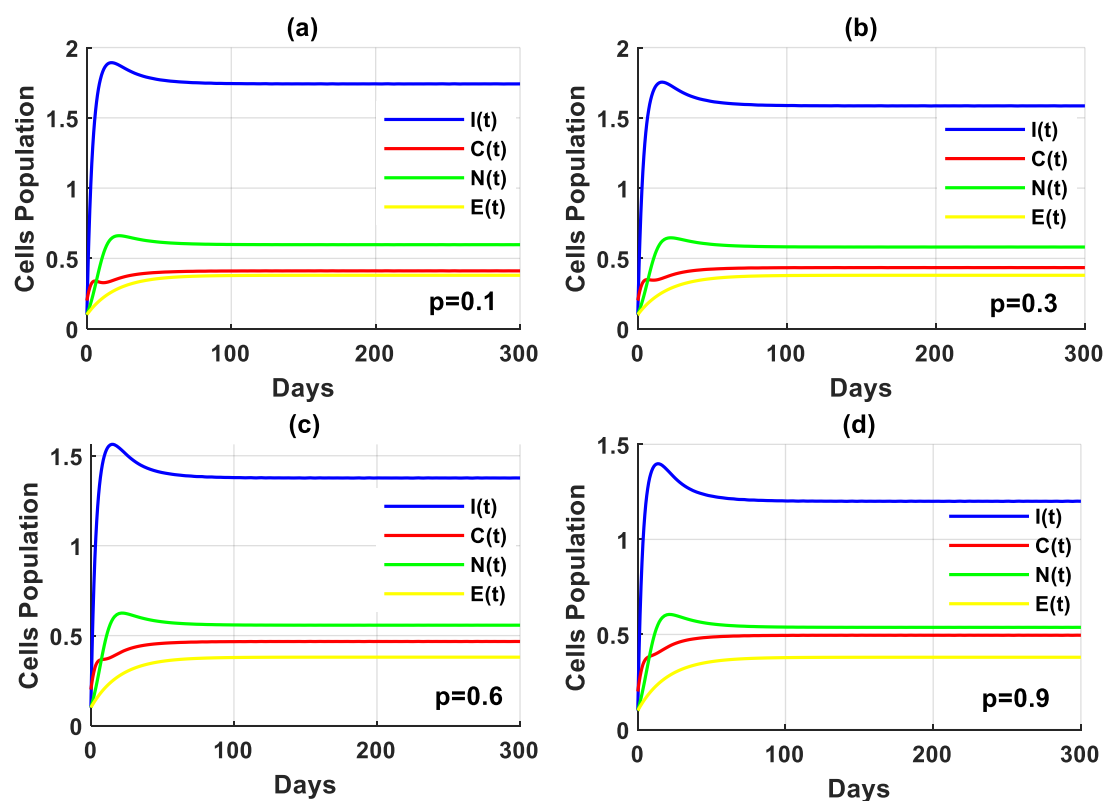


Figure 3. The solution of the PPIGCNE system with various values of psychological scare rate p and the initial conditions of immune cells, breast cancer cells, normal cells, and estrogen level are $I(0) = 0.25$ (cells/ng), $C(0) = 0.3$ (cells/ng), $N(0) = 0.3$ (cells/ng), and $E(0) = 0.12$ ng/mL.

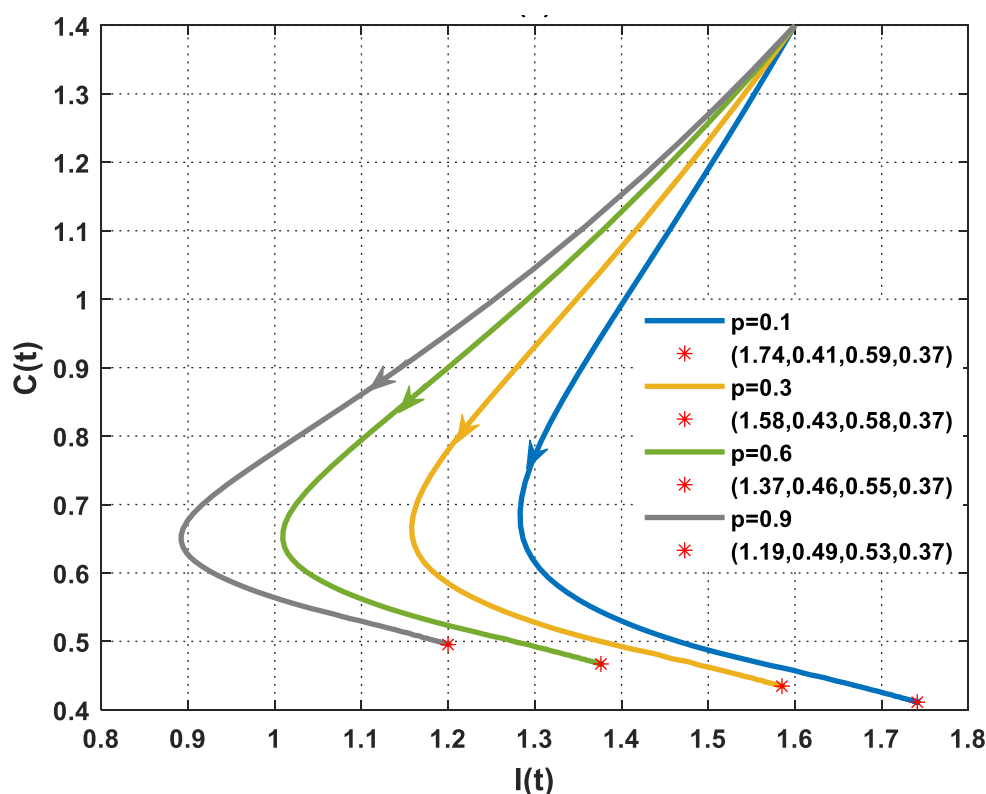


Figure 4. The dynamics of immune and cancer cells with various values of p and the initial conditions of immune cells, breast cancer cells, normal cells, and estrogen level are $I(0) = 1.6$ (cells/ng), $C(0) = 1.4$ (cells/ng), $N(0) = 0.8$ (cells/ng), and $E(0) = 0.8$ ng/mL. (*) represents the final state of the solution and the number in the bracket denotes the equilibrium point with different values of p .

(b) The impact of glucose excess

In the absence of tumor cells, the impact of excess glucose was observed to impair the functionality of the immune system (Figure 5). Furthermore, the solution of the PPIGCNE system, for various values of g , initially experiences growth before converging asymptotically to the free cancer point L_2 . Furthermore, there was a significant drop in immune cells in the presence of tumor cells, as illustrated in Figure 6. In addition, the solution of the PPIGCNE system, as a consequence of excess glucose, drops off into the breast cells-free equilibrium point $L_3 = (\tilde{I}, \tilde{C}, 0, E^*) = (0.51, 1.24, 0, 0.37)$ for $g \geq 0.46$. This result confirms the outcome of Theorem 4 which states that the PPIGCNE system faces a transcritical bifurcation at the glucose parameter $g^* = 0.46$ (please see Figure 6). The above ultimately results in either the patient's death or a mastectomy. Based on the findings, it can be inferred that there exists a direct correlation between elevated glucose levels and the proliferation and rate of division of cancer cells. Conversely, an inverse relationship can be observed between increased glucose levels and the immune system's response to combat malignant cells. Please see Figure 7 for a better idea of what happens when glucose levels are high, both when there are no cancer cells and when there are.

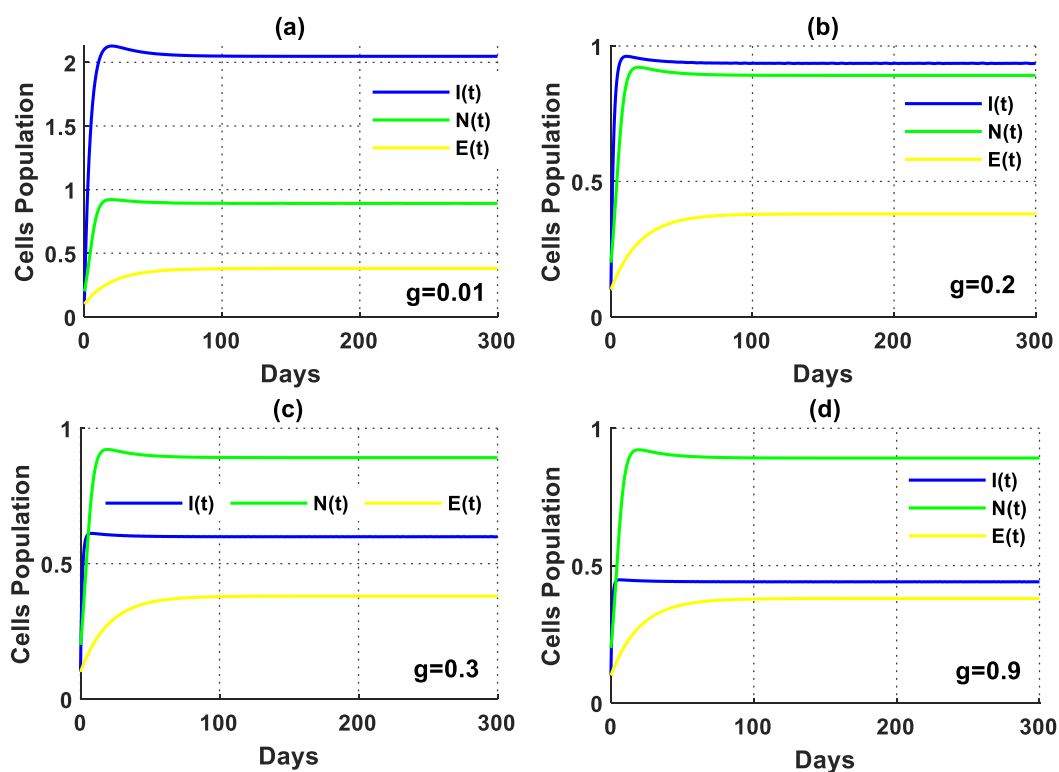


Figure 5. The solution of the PPIGCNE system in the absence of tumor cells with various values of g and the initial conditions of immune cells, breast cancer cells, normal cells, and estrogen level are $I(0) = 0.25$ (cells/ng), $N(0) = 0.3$ (cells/ng), and $E(0) = 0.12$ ng/mL.

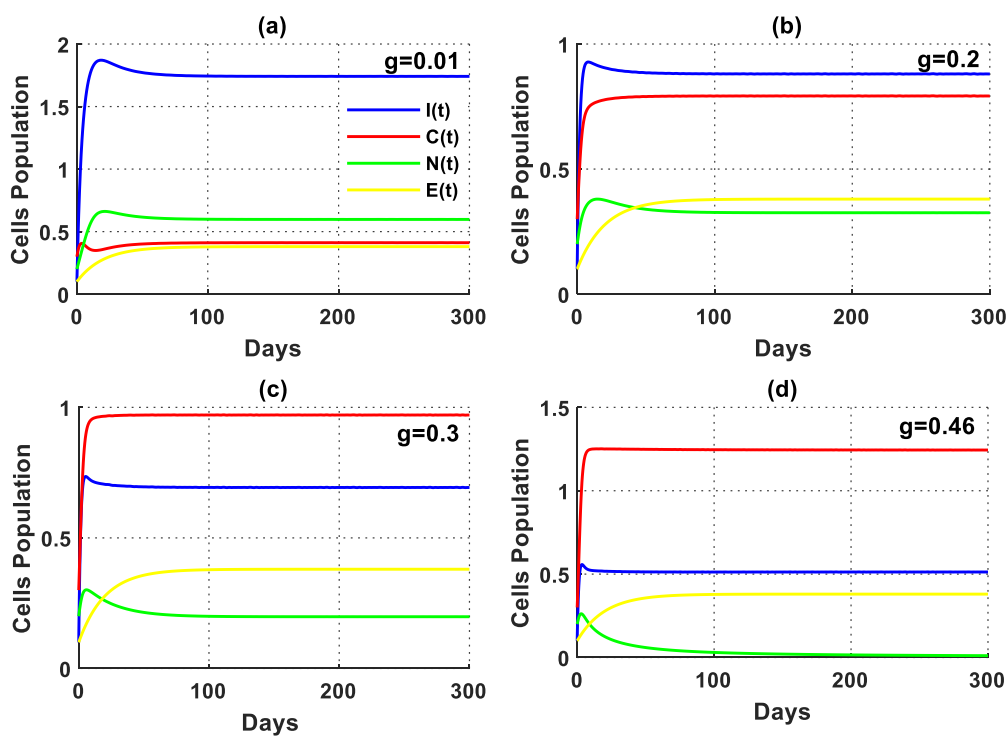


Figure 6. The solution of the PPIGCNE system with various values of g and the initial conditions of immune cells, breast cancer cells, normal cells, and estrogen level are $I(0) = 0.25$ (cells/ng), $C(0) = 0.3$ (cells/ng), $N(0) = 0.3$ (cells/ng), and $E(0) = 0.12$ ng/mL.

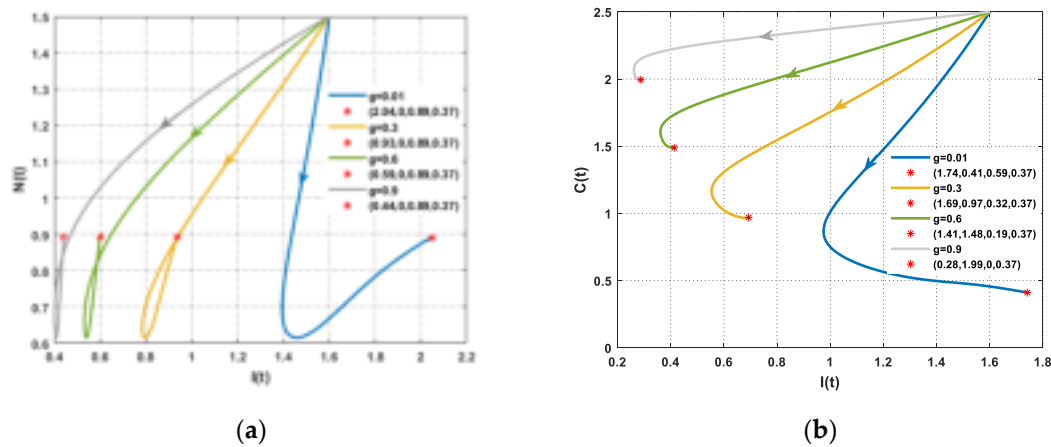


Figure 7. The performance of the PPIGCNE system with various values of g . (a) The dynamics of immune and normal cells in the absence of tumor cells with various values of g and the initial conditions are $I(0) = 1.6$ (cells/ng), $C(0) = 1.4$ (cells/ng), $N(0) = 1.5$ (cells/ng), and $E(0) = 0.8$ ng/mL. (b) The dynamics of immune and cancer cells with various values of g and the initial conditions are $I(0) = 1.6$ (cells/ng), $C(0) = 2.4$ (cells/ng), $N(0) = 1.5$ (cells/ng), and $E(0) = 0.8$ ng/mL.

(c) The impact of estrogen excess

In this scenario, the simulation is conducted by varying the values of the rate of higher estrogen production s both in the absence and presence of tumor cells. In the absence of tumor cells, the solution of the PPIGCNE system approaches asymptotically to the free tumor equilibrium point L_3 with various values s (see Figure 8). While in the presence of tumor cells, the solution of the PPIGCNE system converges to the coexisting point L_4 when $s < 0.12$. For $s \geq 0.12$, the solution of the PPIGCNE system settles down asymptotically to the breast cells-free equilibrium point $L_3 = (\tilde{I}, \tilde{C}, 0, E^*) = (1.01, 0.41, 0, 2.59)$. Please see Figure 9. The numerical solutions generally showed that in the presence of excess estrogen, immune cells, and normal cells decrease, with normal cells being the most affected (see Figure 10a), while tumor cells grow (see Figure 10b).

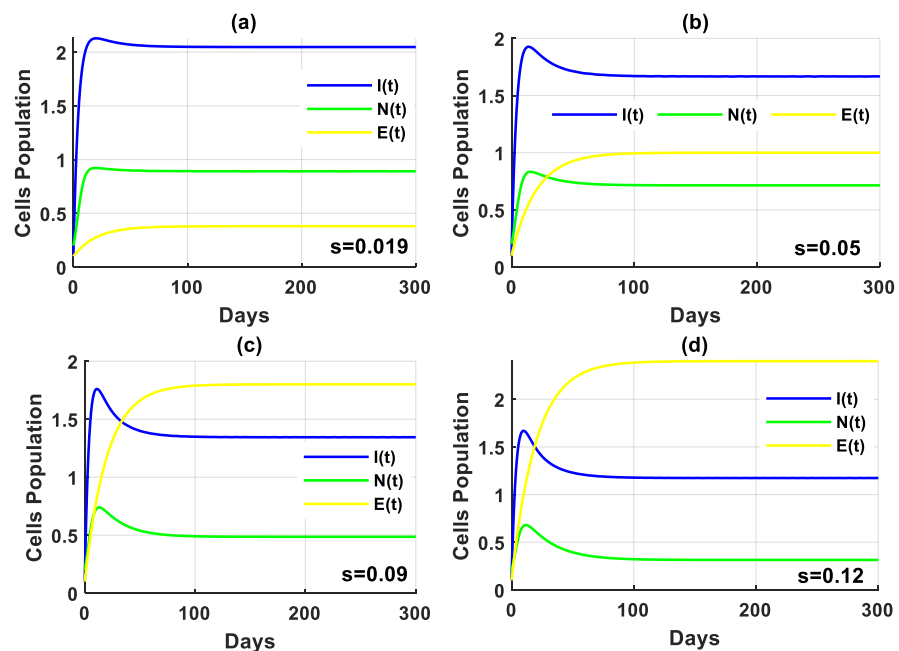


Figure 8. The solution of the PPIGCNE system in the absence of tumor cells with various values of s and the initial conditions of immune cells, breast cancer cells, normal cells, and estrogen level are $I(0) = 0.25$ (cells/ng), $N(0) = 0.3$ (cells/ng), and $E(0) = 0.12$ ng/mL.

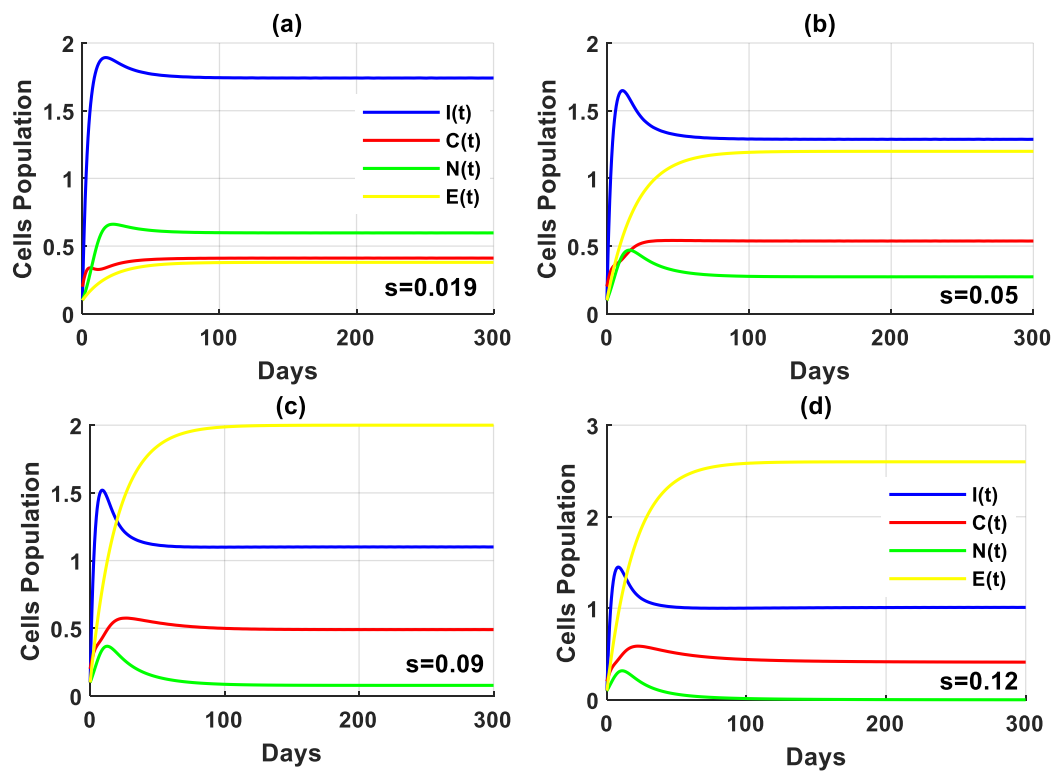


Figure 9. The solution of the PPIGCNE system with various values of s and the initial conditions of immune cells, breast cancer cells, normal cells, and estrogen level are $I(0) = 0.25$ (cells/ng), $C(0) = 0.3$ (cells/ng), $N(0) = 0.3$ (cells/ng), and $E(0) = 0.12$ ng/mL.

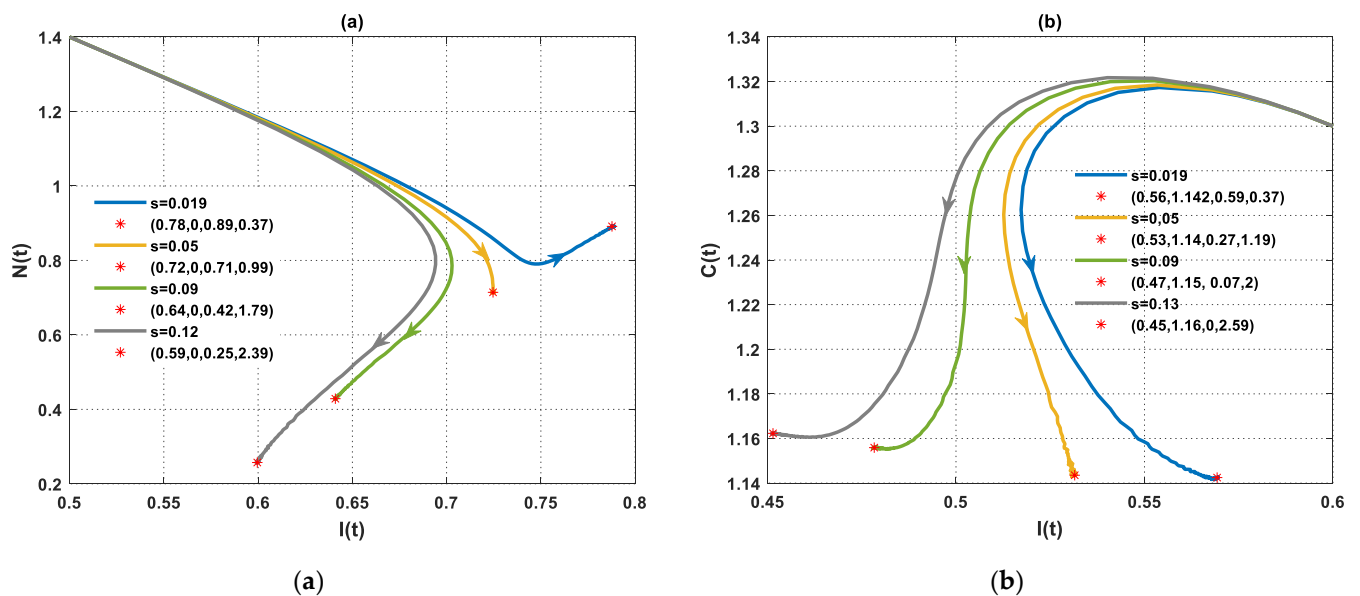


Figure 10. The performance of the PPIGCNE system with various values of s . (a) The dynamics of immune and normal cells in the absence of tumor cells with various values of s and the initial conditions are $I(0) = 0.5$ (cells/ng), $C(0) = 1.4$ (cells/ng), $N(0) = 1.4$ (cells/ng), and $E(0) = 0.8$ ng/mL. (b) The dynamics of immune and cancer cells with various values of s and the initial conditions are $I(0) = 0.6$ (cells/ng), $C(0) = 1.3$ (cells/ng), $N(0) = 1.4$ (cells/ng), and $E(0) = 0.8$ ng/mL.

Remark 1. The biological interpretation of the initial values that are smaller than one in a cancer system that have been used in Figures 2–10 refer to a parameter associated with cancer cell development or spread, such as the proliferation rate or growth factor in a mathematical model. If this initial value is less than one, cell division is slower than cancer cell death. This may imply a less aggressive cancer initially (and vice versa if the initial value is greater than one). The same interpretation is for the rest of the variables.

4. Conclusions

The dynamics of breast cancer in the presence of psychological scare, glucose excess, and estrogen excess have been presented in the form of a system of differential equations. The positivity and boundedness of the proposed system were established. Conditions of existence and local stability of the possible equilibria were illustrated. The global stability around the free breast cancer equilibrium point has been studied. The necessary threshold of glucose level to achieve a state of health has been determined. In terms of mathematics, this threshold is equivalent to a transcritical bifurcation. The numerical simulations validated the analytical findings. Precisely, the threshold value for the transcritical bifurcation was calculated, indicating the point at which normal cells (breast cells) transition from persisting to eradicating. The development of breast cancer in an individual relies on the immune system's capacity to fight against tumor cells. It has been determined that an abundance of estrogen and glucose in the body leads to instability. This suggests that an increase in the amount of estrogen and glucose leads to a higher rate of tumor growth, thereby contributing to the development of breast cancer. It can be inferred that unregulated levels of estrogen and glucose can result in an immune system that is unable to effectively combat cancer cells, leading to a failure to control the disease. The results of the findings in this research could potentially result in the development of an optimal protocol for cancer therapy, which would greatly assist oncologists in their practice.

Our upcoming research will focus on augmenting the immune system through regular vitamin intake.

Author Contributions: Conceptualization, Z.A.; Data curation, S.J.; Formal analysis, S.J.; Funding acquisition, A.A.L.; Investigation, Z.A. and B.B.; Methodology, Z.A. and S.J.; Project administration, A.H.A.; Resources, B.B. and A.A.L.; Software, A.H.A.; Supervision, F.G.; Validation, B.B. and F.G.; Visualization, A.H.A.; Writing—original draft, Z.A. and S.J.; Writing—review & editing, A.H.A., F.G. and A.A.L. All authors have read and agreed to the published version of the manuscript.

Funding: The publication of this research was supported by the University of Oradea.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Abernathy, K.; Abernathy, Z.; Baxter, A.; Stevens, M. Global dynamics of a breast cancer competition model. *Differ. Equ. Dyn. Syst.* **2020**, *28*, 791–805.
2. World Health Organization. *Global Action Plan for the Prevention and Control of Noncommunicable Diseases 2013–2020*; World Health Organization: Geneva, Switzerland, 2013.
3. Tarver, T. American cancer society. cancer facts and figures 2014. *J. Consum. Health Internet* **2012**, *16*, 366–367.
4. Patel, M.; Nagl, S. *The Role of Model Integration in Complex Systems Modelling: An Example from Cancer Biology*; Springer: Berlin/Heidelberg, Germany, 2010; ISBN 3642156037.
5. Dehingia, K.; Yao, S.-W.; Sadri, K.; Das, A.; Sarmah, H.K.; Zeb, A.; Inc, M. A study on cancer-obesity-treatment model with quadratic optimal control approach for better outcomes. *Results Phys.* **2022**, *42*, 105963.
6. Alnahdi, A.S.; Idrees, M. Nonlinear dynamics of estrogen receptor-positive breast cancer integrating experimental data: A novel spatial modeling approach. *Math. Biosci. Eng.* **2023**, *20*, 21163–21186.
7. Miziak, P.; Baran, M.; Błaszczak, E.; Przybyszewska-Podstawka, A.; Kałafut, J.; Smok-Kalwat, J.; Dmoszyńska-Graniczka, M.; Kielbus, M.; Stepulak, A. Estrogen Receptor Signaling in Breast Cancer. *Cancers* **2023**, *15*, 4689.

8. Mufudza, C.; Sorofa, W.; Chiyaka, E.T. Assessing the effects of estrogen on the dynamics of breast cancer. *Comput. Math. Methods Med.* **2012**, *2012*, 473572.
9. Almeida, L.; Denis, J.A.; Ferrand, N.; Lorenzi, T.; Prunet, A.; Sabbah, M.; Villa, C. Evolutionary dynamics of glucose-deprived cancer cells: Insights from experimentally informed mathematical modelling. *J. R. Soc. Interface* **2024**, *21*, 20230587.
10. Barbosa, A.M.; Martel, F. Targeting glucose transporters for breast cancer therapy: The effect of natural and synthetic compounds. *Cancers* **2020**, *12*, 154.
11. Sun, S.; Sun, Y.; Rong, X.; Bai, L. High glucose promotes breast cancer proliferation and metastasis by impairing angiotensinogen expression. *Biosci. Rep.* **2019**, *39*, BSR20190436.
12. Shomali, N.; Mahmoudi, J.; Mahmoodpoor, A.; Zamiri, R.E.; Akbari, M.; Xu, H.; Shotorbani, S.S. Harmful effects of high amounts of glucose on the immune system: An updated review. *Biotechnol. Appl. Biochem.* **2021**, *68*, 404–410.
13. Alblowy, A.H.; Maan, N.; Alharbi, S.A. Role of glucose risk factors on human breast cancer: A nonlinear dynamical model evaluation. *Mathematics* **2022**, *10*, 3640.
14. Thirthar, A.A.; Abboubakar, H.; Khan, A.; Abdeljawad, T. Mathematical modeling of the COVID-19 epidemic with fear impact. *AIMS Math.* **2023**, *8*, 6447–6465.
15. Doshi, D.; Karunakar, P.; Sukhabogi, J.R.; Prasanna, J.S.; Mahajan, S.V. Assessing coronavirus fear in Indian population using the fear of COVID-19 scale. *Int. J. Ment. Health Addict.* **2021**, *19*, 2383–2391.
16. Vrinten, C.; McGregor, L.M.; Heinrich, M.; von Wagner, C.; Waller, J.; Wardle, J.; Black, G.B. What do people fear about cancer? A systematic review and meta-synthesis of cancer fears in the general population. *Psycho-Oncology* **2017**, *26*, 1070–1079.
17. Niknamian, S. The Impact of Stress, Anxiety, Fear and Depression in The Cause of Cancer in Humans. *Am. J. Biomed. Sci. Res.* **2019**, *3*, 363–370.
18. Gormley, M.; Knobf, M.T.; Vorderstrasse, A.; Aouizerat, B.; Hammer, M.; Fletcher, J.; D'Eramo Melkus, G. Exploring the effects of genomic testing on fear of cancer recurrence among breast cancer survivors. *Psycho-Oncology* **2021**, *30*, 1322–1331.
19. Lebel, S.; Tomei, C.; Feldstain, A.; Beattie, S.; McCallum, M. Does fear of cancer recurrence predict cancer survivors' health care use? *Support. Care Cancer* **2013**, *21*, 901–906.
20. Von Ah Morano, A.E.; Dorneles, G.P.; Peres, A.; Lira, F.S. The role of glucose homeostasis on immune function in response to exercise: The impact of low or higher energetic conditions. *J. Cell Physiol.* **2020**, *235*, 3169–3188.
21. Jawad, S.; Winter, M.; Rahman, Z.-A.S.A.; Al-Yasir, Y.I.A.; Zeb, A. Dynamical Behavior of a Cancer Growth Model with Chemotherapy and Boosting of the Immune System. *Mathematics* **2023**, *11*, 406.
22. Tang, T.-Q.; Shah, Z.; Bonyah, E.; Jan, R.; Shutaywi, M.; Alreshidi, N. Modeling and Analysis of Breast Cancer with Adverse Reactions of Chemotherapy Treatment through Fractional Derivative. *Comput. Math. Methods Med.* **2022**, *2022*, 5636844.
23. Alharbi, S.A.; Rambely, A.S. A New ODE-Based Model for Tumor Cells and Immune System Competition. *Mathematics* **2020**, *8*, 1285.
24. Wang, X.; Zanette, L.; Zou, X. Modelling the fear effect in predator–prey interactions. *J. Math. Biol.* **2016**, *73*, 1179–1204.
25. Kuznetsov, V.A.; Makalkin, I.A.; Taylor, M.A.; Perelson, A.S. Nonlinear dynamics of immunogenic tumors: Parameter estimation and global bifurcation analysis. *Bull. Math. Biol.* **1994**, *56*, 295–321.
26. De Pillis, L.G.; Radunskaya, A. A mathematical tumor model with immune resistance and drug therapy: An optimal control approach. *Comput. Math. Methods Med.* **2001**, *3*, 79–100.
27. Gatenby, R.A. The potential role of transformation-induced metabolic changes in tumor-host interaction. *Cancer Res.* **1995**, *55*, 4151–4156.
28. Abdulkream Alharbi, S.; Dehingia, K.; Jahman Alqarni, A.; Alsulami, M.; Al Qarni, A.A.; Das, A.; Hinçal, E. A study on ODE-based model of risk breast cancer with body mass. *Appl. Math. Sci. Eng.* **2023**, *31*, 2259059.
29. Ouifki, R.; Oke, S.I. Mathematical model for the estrogen paradox in breast cancer treatment. *J. Math. Biol.* **2022**, *84*, 28.
30. Das, A.; Dehingia, K.; Ray, N.; Sarmah, H.K. Stability analysis of a targeted chemotherapy-cancer model. *Math. Model. Control* **2023**, *3*, 116–126.
31. Sahoo, B.; Poria, S. Diseased prey predator model with general Holling type interactions. *Appl. Math. Comput.* **2014**, *226*, 83–100.
32. Hubbard, J.H.; West, B.H. *Differential Equations: A Dynamical Systems Approach: Ordinary Differential Equations*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 5.
33. Alebraheem, J.; Ibrahim, T.Q.; Arif, G.E.; Hamdi, A.A.; Bazighifan, O.; Ali, A.H. The stabilizing effect of small prey immigration on competitive predator-prey dynamics. *Math. Comput. Model. Dyn. Syst.* **2024**, *30*, 605–625.
34. Ahmed, M.; Jawad, S. Bifurcation Analysis of the Role of Good and Bad Bacteria in the Decomposing Toxins in the Intestine with the Impact of Antibiotic and Probiotics Supplement. In *AIP Conference Proceedings*; AIP Publishing: Melville, NY, USA, 2024; Volume 3097.
35. Bassim, W.B.; Salem, A.J.; Ali, A.H. A new analytical study of prey-predator dynamical systems involving the effects of Hide-and-Escape and predation skill augmentation. *Results Control Optim.* **2024**, *16*, 100449.
36. Thirthar, A.A.; Jawad, S.; Majeed, S.J.; Nisar, K.S. Impact of wind flow and global warming in the dynamics of prey-predator model. *Results Control Optim.* **2024**, *15*, 100424.
37. Jawad, S.R.; Al Nuaimi, M. Persistence and bifurcation analysis among four species interactions with the influence of competition, predation and harvesting. *Iraqi J. Sci.* **2023**, *64*, 1369–1390.
38. Hassan, S.K.; Jawad, S.R. The Effect of Mutual Interaction and Harvesting on Food Chain Model. *Iraqi J. Sci.* **2022**, *63*, 2641–2649.

39. Thirthar, A.A.; Jawad, S.; Shah, K.; Abdeljawad, T. How does media coverage affect a COVID-19 pandemic model with direct and indirect transmission? *J. Math. Comput. Sci.* **2024**, *35*, 169–181.
40. Ali, A.; Jawad, S. Stability analysis of the depletion of dissolved oxygen for the Phytoplankton-Zooplankton model in an aquatic environment. *Iraqi J. Sci.* **2024**, *65*, 2736–2748.
41. Mondal, B.; Thirthar, A.A.; Sk, N.; Alqudah, M.A.; Abdeljawad, T. Complex dynamics in a two species system with Crowley–Martin response function: Role of cooperation, additional food and seasonal perturbations. *Math. Comput. Simul.* **2024**, *221*, 415–434.
42. Pakhira, R.; Mondal, B.; Thirthar, A.A.; Alqudah, M.A.; Abdeljawad, T. Developing a fuzzy logic-based carbon emission cost-incorporated inventory model with memory effects. *Ain Shams Eng. J.* **2024**, *15*, 102746.
43. Thirthar, A.A. A mathematical modelling of a plant-herbivore community with additional effects of food on the environment. *Iraqi J. Sci.* **2023**, *64*, 3551–3566.
44. Sk, N.; Mondal, B.; Thirthar, A.A.; Alqudah, M.A.; Abdeljawad, T. Bistability and tristability in a deterministic prey–predator model: Transitions and emergent patterns in its stochastic counterpart. *Chaos Solitons Fractals* **2023**, *176*, 114073.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.