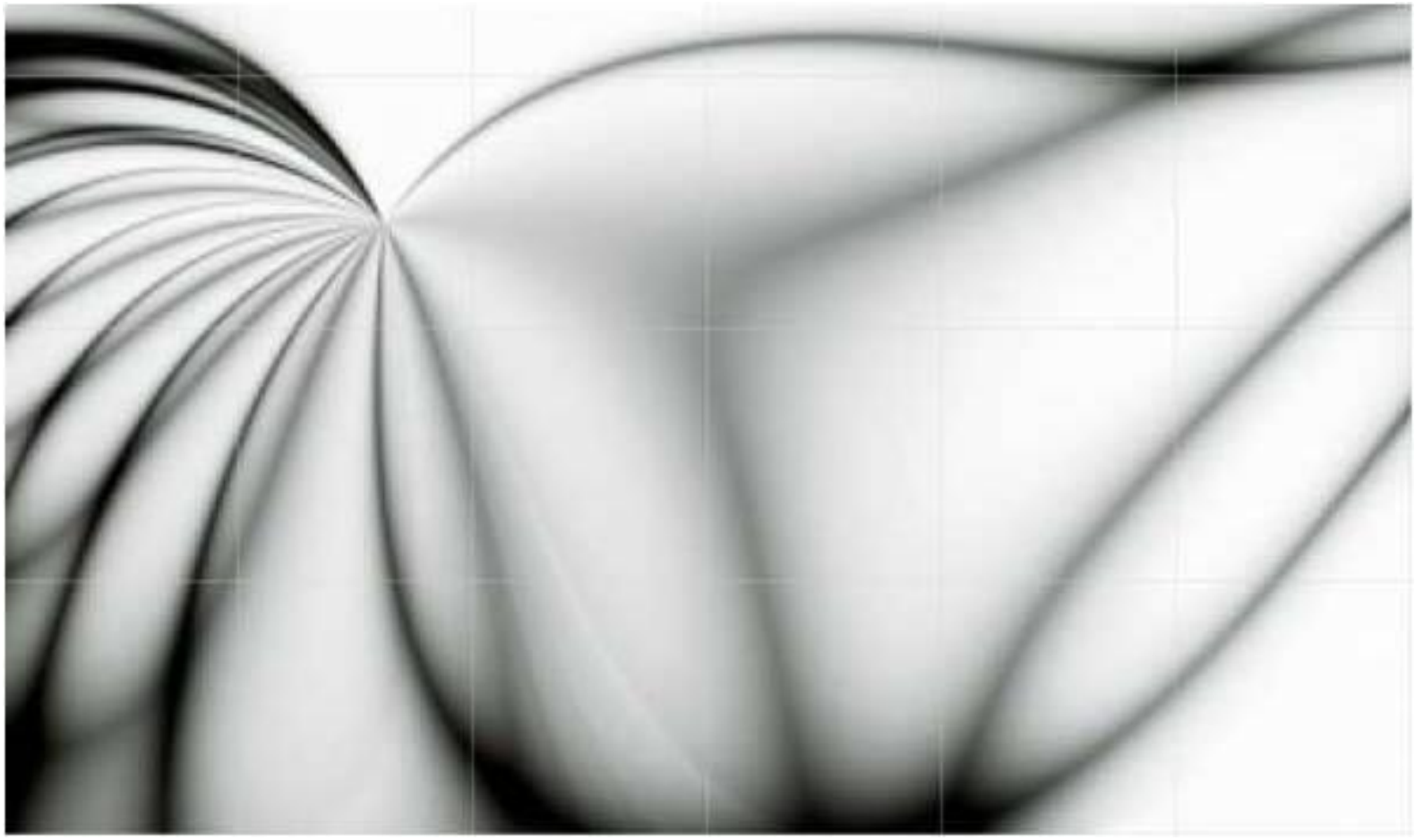
IJOCTA
An International Journal of Optimization and Control: Theories & Applications
2010

ISSN:2146-0957
eISSN:2146-5703

Volume:14 Number:4
October 2024

An International Journal of Optimization and Control:

Theories & Applications



www.ijocta.org

IJOCTA
An International Journal of Optimization and Control: Theories & Applications
2010

http://www.ijocta.org
editor@ijocta.org

An International Journal of Optimization and Control: Theories & Applications

Volume: 14, Number: 4

October 2024

Publisher & Owner *(Yayımcı & Sahibi)*:
Prof. Dr. Ramazan YAMAN
Atlas Vadi Campus, Anadolu St.
No. 40, Kagithane, Istanbul, Türkiye
Atlas Vadi Kampüsü, Anadolu Cad. No. 40, Kağıthane, İstanbul, Türkiye

ISSN: 2146-0957
eISSN: 2146-5703

Press *(Basımevi)*:
İdeal Kültür, Topkapı Mahallesi, Kahalbaşı Sokak, No: 31/1, 34093, Fatih, İstanbul, Türkiye
Ideal Kultur, Topkapi District, Kahalbasi Street, No: 31/1, 34093, Fatih, Istanbul, Türkiye

Date Printed *(Basım Tarihi)*:
October 2024
Ekim 2024

Responsible Director (Sorumlu Müdür):
Prof. Dr. Ramazan YAMAN

IJOCTA is an international, quarterly, and peer-reviewed journal indexed/ abstracted by *(IJOCTA, yılda dört kez yayımlanan ve aşağıdaki indekslerce taranan/dizinlenen uluslararası hakemli bir dergidir)*:

- Emerging Sources Citation Index (ESCI)
- Scopus - Applied Mathematics
- Scopus - Control and Optimization
- Zentralblatt MATH
- ProQuest - Electronics & Communications Abstracts
- Proquest - Computer & Information Systems Abstracts
- Proquest - Mechanical & Transportation Eng. Abstracts
- EBSCOHost
- MathSciNet
- TR Dizin - Ulakbim (Science Database)
- Directory of Open Access Journals (DOAJ)

crossref

We Use Plagiarism Detection iThenticate Learn More »

Editor in Chief
YAMAN, Ramazan - Istanbul Atlas University / Türkiye

Area Editors (Applied Mathematics & Control)
OZDEMIR, Necati - Balikesir University / Türkiye
JAJARMI, Amin - University of Bojnord / Iran
EVIRGEN, Firat - Balikesir University, Türkiye

Area Editors (Engineering Applications)
JANARDHANAN, Mukund N. - University of Warwick / UK
MANDZUKA, Sadko - University of Zagreb / Crotia

Area Editors (Fractional Calculus & Applications)
BALEANU, Dumitru - Lebanese American University / Lebanon
POVSTENKO, Yuriy - Jan Dlugosz University / Poland

Area Editors (Optimization & Applications)
WEBER, Gerhard Wilhelm - Poznan University of Technology / Poland
KUCUKKOC, Ibrahim - Balikesir University / Türkiye

Area Editor (Chaos Theory)
HAMMOUCH, Zakia - Moulay Ismail University / Morocco

Editorial Board
AGARWAL, Ravi P. - Texas A&M University Kingsville / USA
AGHABABA, Mohammad P. - Urmai University of Tech. / Iran
ATANGANA, A. - University of the Free State / South Africa
AVCI, Derya - Balikesir University / Türkiye
AYAZ, Fatma - Gazi University / Türkiye
BAGIROV, Adil - University of Ballarat / Australia
BATTINI, Daria - Universita degli Studi di Padova / Italy
BOHNER, Martin - Missouri University of Science and Technology / USA
CAKICI, Eray – IBM Deutschland GmbH / Germany
CARVALHO, Maria Adelaide P. d. Santos - Institute of Miguel Torga / Portugal
CHEN, YangQuan - University of California Merced / USA
DAGLI, Cihan H. - Missouri University of Science and Technology / USA
DAI, Liming - University of Regina / Canada
DEMIRTAS, M. - Balikesir University / Türkiye
GURBUZ, Burcu - Johannes Gutenberg-University Mainz / Germany
HRISTOV, Jordan - University of Chemical Technology and Metallurgy / Bulgaria
ISKENDER, Beyza B. - Balikesir University / Türkiye
JONRINALDI, J. - Universitas Andalas, Padang / Indonesia
KARAOGLAN, Aslan Deniz - Balikesir University / Türkiye
KATALINIC, Branko - Vienna University of Technology / Austria
MARTINEZ, Antonio J. Torija - University of Salford / UK
NANE, Erkan - Auburn University / USA
PAKSOY, Turan – Necmettin Erbakan University / Türkiye
SULAIMAN, Shamsuddin - Universiti Putra Malaysia / Malaysia
SUTIKNO, Tole - Universitas Ahmad Dahlan / Indonesia
TABUCANON, Mario T. - Asian Institute of Technology / Thailand
TEO, Kok Lay - Curtin University / Australia
TRUJILLO, Juan J. - Universidad de La Laguna / Spain
WANG, Qing - Durham University / UK
XU, Hong-Kun - National Sun Yat-sen University / Taiwan
YAMAN, Gulsen - Balikesir University / Türkiye
ZAKRZHEVSKY, Mikhail V. - Riga Technical University / Latvia
ZHANG, David - University of Exeter / UK

Production Editor
EVIRGEN, Firat - Balikesir University, Türkiye

English Editors
INAN, Dilek - Izmir Democracy University / Türkiye

An International Journal of Optimization and Control:
Theories & Applications
Volume: 14 Number: 4
October 2024

CONTENTS

RESEARCH ARTICLES

294 List coloring based algorithm for the Futoshiki puzzle
Banu Baklan Şen, Oznur Yasar Diner

308 Early prediction of fabric quality using machine learning to reduce rework in manufacturing processes
Sema Aydın, Koray Altun

322 Witte's conditions for uniqueness of solutions to a class of Fractal-Fractional ordinary differential equations
Abdon Atangana, Ilknur Koca

336 Influence of rotation on peristaltic flow for pseudoplastic fluid: a wavy channel
Hayat Adel Ali, Mohammed R. Salman

346 A comparative view to H_infinity-norm of transfer functions of linear DAEs
Hasan Gündüz, Ercan Çelik, Mesut Karabacak

355 Fuzzy-PID and interpolation: a novel synergetic approach to process control
Devashish Jha, Arifa Ahmed, Sanatan Kumar, Debanjan Roy

365 Global mathematical analysis of a patchy epidemic model
Lahcen Boulaasair, Hassane Bouzahir, Mehmet Yavuz

378 An Inverse recursive algorithm to retrieve the shape of the inaccessible dielectric objects
Ahmet Sefer

394 A local differential quadrature method for the generalized nonlinear Schrödinger (GNLS) equation
Meirikim Panmei, Roshan Thoudam

404 Modeling the dependency structure between quality characteristics in multi-stage manufacturing processes with copula functions
Pelin Toktaş, Ömer Lütfi Gebizlioğlu

An International Journal of Optimization and Control: Theories & Applications
ISSN:2146-0957 eISSN:2146-5703
Vol.14, No.4, pp.294-307 (2024)
http://doi.org/10.11121/ijocta.1432

RESEARCH ARTICLE

List coloring based algorithm for the Futoshiki puzzle

Banu Baklan Şen *, Öznur Yaşar Diner

Computer Engineering Department, Kadir Has University, Turkey
banu.baklan@stu.khas.edu.tr, oznur.yasar@khas.edu.tr

ARTICLE INFO

Article History:
Received 18 July 2023
Accepted 11 September 2024
Available Online 9 October 2024

Keywords:
List coloring
Precoloring extension
Latin square completion puzzle
Futoshiki puzzle
Personnel scheduling

AMS Classification 2010:
90C27; 05C85; 68Q25

ABSTRACT

Given a graph $G = (V, E)$ and a list of available colors $L(v)$ for each vertex $v \in V$, where $L(v) \subseteq \{1, 2, \ldots, k\}$, List k-Coloring refers to the problem of assigning colors to the vertices of G so that each vertex receives a color from its own list and no two neighboring vertices receive the same color. The decision version of the problem, List k-Coloring, is NP-complete even for bipartite graphs. As an application of list coloring problem we are interested in the Futoshiki Problem. Futoshiki is an NP-complete Latin Square Completion Type Puzzle. Considering Futoshiki puzzle as a constraint satisfaction problem, we first give a list coloring based algorithm for it which is efficient for small boards of fixed size. To thoroughly investigate the efficiency of our algorithm in comparison with a proposed backtracking-based algorithm, we conducted a substantial number of computational experiments at different difficulty levels, considering varying numbers of inequality constraints and given values. Our results from the extensive range of experiments indicate that the list coloring-based algorithm is much more efficient.

CC BY

1. Introduction

Since the 1980s, there has been significant theoretical analysis and exploration of applications for pencil puzzle games. In recent decades, research has focused on algorithmic solutions and the computational complexity of pencil puzzle games, including optimization versions of various puzzle types. Latin Square Completion-Type Puzzles (LSCP) are among the most common types of these games.

A Latin Square Completion Puzzle (LSCP) is a partial Latin square with empty cells. A Partial Latin Square (PLS) is an $n \times n$ grid that is partially filled with some numbers from $[n] = \{1, \ldots, n\}$. The goal is to fill in all the blank cells with numbers in such a way that the numbers are distinct in each row and each column. The objective of LSCP is to complete the grid by filling the remaining cells with numbers such that each number appears exactly once in each row and each column. Two notable puzzles in this category are Sudoku and Futoshiki.

The Futoshiki puzzle, also known as Unequal, is a popular Japanese board-based puzzle played on an $n \times n$ square board with additional inequality constraints between certain cells. The objective is to fill the cells with numbers, satisfying the Latin square property while respecting the specified inequalities. Inequalities can occur between horizontally or vertically neighboring cells, indicating that a number in a particular cell must be greater or smaller than the number in the adjacent cell. Let S denote the set of inequality constraints and T the set of pre-assigned cells.

The decision version of the Futoshiki game, known as the Futoshiki Problem, is defined as follows:

Futoshiki Problem (Futoshiki)
Instance: $\mathcal{F}_n(T, S)$, an $n \times n$ board, a set T of

*Corresponding Author

pre-assigned cells, and a set S of inequality constraints.
Question: Is the Futoshiki puzzle solvable on $\mathcal{F}_n(T, S)$?

The solvability of a partial Latin square is closely related to Hall's condition. However, Bobga et al. [1] demonstrated that satisfying Hall's condition is insufficient. They provided necessary and sufficient conditions on the configuration of the prescribed cells to ensure the solvability of LSCP. Further results on related topics, such as partial Latinized rectangles, can be found in [2] and [3].

Both the decision version of the Latin Square completion problem and the Futoshiki Problem have been proven to be NP-Complete [4, 5].
Let us define the optimization version of Futoshiki Problem.

Maximum Futoshiki (MaxFutoshiki)
Input: $\mathcal{F}_n(T, S)$ and sign set $S \subseteq S_L$.
Output: A Futoshiki board $\mathcal{F}_n(T, S)$ filled with maximum number of valid entries.

Various studies examine the computational complexity of problems defined on partial latin squares. The Latin Square completion problem is NP-Complete by reduction from 3-SAT [4]. In particular the Futoshiki problem is also known to be NP-Complete, as proved by Haraguchi et al. [5]. As for the optimization version of the Futoshiki, Haraguchi and Ono [5] examined the approximability of LSCPs and formulated three LSCP puzzles as maximization problems, presenting polynomial-time approximation algorithms. These maximization problems aim to fill as many cells as possible, instead of determining whether it is possible to complete the entire board. MaxFutoshiki was shown to be NP-Hard by Haraguchi [5], and related work on optimization versions of LSCP problems is reviewed by Donovan [6].

Properties of Latin squares and improvements to Galvin's solution [7] have been explored by Ivanyi and Nemeth [8]. Yato and Seta [9] investigated the computational complexity and completeness of finding alternative solutions to LSCP problems and proposed two algorithms.

The solvability of LSCP puzzles has been extensively studied in terms of time complexity, and numerous algorithmic solutions have been proposed. Sudoku, a well-known puzzle, has been approached using various algorithmic techniques for both deterministic and metaheuristic approaches. A deterministic algorithm does not contain any randomness or probabilistic elements. It always produces the same output and follows a fixed sequence of steps. Some major deterministic approaches to solve LSCP type problems are the exact cover problem with, Norvig's work with constraint propagation [10] and constraint programming that Crawford gave [11]. As for nondeterministic approaches, one can refer to the "Dancing Links" algorithm that Knuth presented [12].

The term "Metaheuristic" was first used in the study of Glover [13]. Metaheuristics are known as one of the best methods for finding sufficiently good solutions to NP-Hard problems. Traveling salesman problems, scheduling problems, and assignment problems are some of the examples that metaheuristics are used. Sudoku has been solved with one of the metaheuristic methods that are artificial bee colony algorithm [14], particle swarm optimisation [15], and ant colony optimisation algorithm [16]. Moreover, as a heuristics, we can show the study of Musliu [17] that proposes a hybrid method for solving Sudoku.

In this work, we concentrate on deterministic approaches rather than solving the puzzle with metaheuristic methods. One of the most common techniques to solve PLSs as a combinatorial optimization problem is coloring [4]. Furthermore, many graph coloring variants have been utilized to solve LSCP puzzles. For instance, in [18], the Sudoku puzzle is shown to illustrate the precoloring extension problem [19]. Precoloring extension is a variant of the precoloring problem in which some vertices are precolored and others are assigned lists of allowed colors. Notice that the NP-Completeness of the list coloring problem for general graphs [20] and bipartite graphs [21] has been proven.

Our motivation for studying the list coloring approach for the Futoshiki puzzle game stems from the fact that while Sudoku has been extensively studied as a graph coloring problem, Futoshiki has not been analyzed in the same context. In this paper, we adapt the Futoshiki puzzle game to a new variant of the list coloring problem, which we refer to as the *list precoloring extension problem* (formally defined in Section 2). We propose a list precoloring extension algorithm and discuss its complexity.

The rest of the paper is organized as follows. Section 2 provides problem definitions. Section 3 explores applications related to the Futoshiki game. Section 4 establishes the equivalence between the

list precoloring extension instance and the Futoshiki problem instance. Section 5 presents an algorithm to solve the Futoshiki problem when the board size is fixed. Section 6 analyzes the experimental results. Finally, Section 7 concludes the paper.

2. Problem definitions

In this section, we provide formal notation and terminology related to graph coloring and introduce a new graph coloring problem that models the Futoshiki problem. The notation and basic terminology used in this section follow Diestel [22].

We consider simple, finite, undirected graphs $G = (V, E)$ with a vertex set V and an edge set E. A *coloring* of a graph G is a labeling of its vertices. A *k-coloring* is a coloring that uses at most k colors from the set $[k] = 1, 2, \ldots, k$. A coloring is *proper* if no two adjacent vertices have the same color. The decision version of the graph coloring problem is defined as follows:

Coloring (Col)
Instance: A graph $G = (V, E)$ and an integer $k \geq 1$.
Question: Does G have a k-coloring?

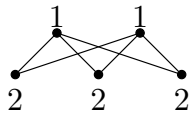


Figure 1. A graph G and a valid coloring for it.

In coloring, k is a part of the input. On the other hand, when k is fixed, i.e., when k is not a part of the input, we have the k-coloring problem. As an example, a 2-coloring is given in Figure 1: vertices in one part receive one color, while vertices in the other part receive a different color. It is worth noting that every bipartite graph can be colored using only two colors.

k-Coloring (k-Col)
Instance: A graph $G = (V, E)$.
Question: Does G have a k-coloring? k-Col is NP-Complete for $k \geq 3$ [23] and polynomial time solvable when $k = 1$ or 2 [24]. *List coloring* is a generalization of graph coloring. It is a proper coloring in which each vertex v receives a color from its own list of allowed colors. The list coloring problem is defined by Vizing [23] and Erdös, Rubin and Taylor [25] independently.

List-Coloring (LiCol)
Instance: A graph $G = (V, E)$ and a list assignment L for G.
Question: Does G have a coloring where each vertex v receives a color from its list $L(v)$?

A *list assignment* of a graph $G = (V, E)$ is a mapping L that assigns each vertex $v \in V$ a List $L(v) \subseteq \{1, 2, \ldots\}$ of admissible colors for v. When $L(v) \subseteq [k] = \{1, 2, \ldots k\}$ for every $v \in V$ we say that L is a k-list assignment of G. Thus, the total number of available colors is bounded by k in a k-list assignment. On the other hand, when $|L(v)| \leq k$ for every $v \in V$, then we say that L is a list k-assignment of G. Thus, the size of each list is bounded by k in a list k-assignment.

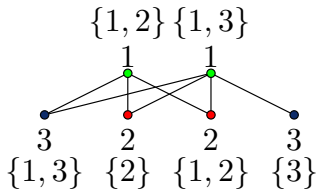


Figure 2. A list assignment L for the vertices of G, and a coloring that respects L.

The *List k-Coloring* problem is to decide whether a graph $G = (V, E)$ with a list $L(u) \subseteq \{1, \ldots, k\}$ for each $u \in V$ has a coloring c such that $c(u) \in L(u)$ for every $u \in V$. It is clearly a generalization of k-coloring, and hence it is NP-Complete for $k \geq 3$. Refer to Figure 2 for an example. It is important to note that, despite the graph being bipartite in Figure 2, two colors were not enough to color it while satisfying the constraints imposed by the assigned lists L.

List k-Coloring (Li k-Col)
Instance: A graph $G = (V, E)$ and a k-list assignment L.
Question: Does G have a coloring where each vertex v receives a color from its list $L(v)$?

A k-precoloring of a graph $G = (V, E)$ is a mapping $c_W : W \to \{1, 2, \ldots k\}$ for some subset $W \subseteq V$. We say that a k-coloring c of G is an *extension* or a *k-extension* of a k-precoloring c_W of G if $c(v) = c_W(v)$ for each $v \in W$. For a given graph G, a positive integer k and a k-precoloring c_W of G, the Precoloring Extension problem (PrExt) asks whether c_W can be extended to a k-coloring of G. If k is fixed we denote this problem as the k-Precoloring Extension problem (k-PrExt). Let us define the latter problem formally.

k-Precoloring Extension (k-PrExt)
Instance: A graph $G = (V, E)$ and a k-precoloring c_W.
Question: Is there a k-extension for c_W?

For general graphs PrExt is NP-Complete [26]. In fact, the NP-Completeness of the LSCP problem is shown via its equivalence to the k-PrExt when it is restricted to the cartesian product of K_n with itself [4].
We define a new coloring problem called the List k-Precoloring Extension problem (Li k-PrExt).

List k-Precoloring Extension (Li k-PrExt)
Instance: A graph $G = (V, E), W \subseteq V$, a k-precoloring c_W, and a list k-assignment L for each $v \in V/W$
Question: Is there a k-extension for c_W that obeys the list L?

Notice that when the list L is not assigned to the vertices in V/W, then Li k-PrExt reduces to k-PrExt. Let us denote an instance of Li k-PrExt with $\mathcal{L}_G(c_W, L)$.

3. Applications

In this section, we will first provide a brief overview of some notable applications related to the problems under consideration, namely the Futoshiki problem, list coloring, and its variants. Subsequently, we will introduce a novel application of the Futoshiki problem in the field of scheduling, specifically to optimize the efficiency of the job assignment problem.

Applications of the Futoshiki problem: The Futoshiki problem has found applications in various domains. Mahmood [27] proposed a random number generator that utilizes the Futoshiki problem to generate numbers satisfying given conditions. This generator, with good linear complexity, has potential application as an encryption key in mathematical analysis, security systems, and simulations. Additionally, Haraguchi [28] explored the evaluation values achievable in a Futoshiki puzzle with a high number of inequality signs.

The Futoshiki configuration technique, considering partial shading conditions in photovoltaic (PV) systems, has been proposed by Sahu et al. [29]. They observed that incorporating the Futoshiki structure increases the power generation of PV arrays, leading to improved energy efficiency. The technique avoids the need for changing the electrical connection of modules by rearranging them, and it effectively reduces mismatch loss under different shading models.

Applications of List Coloring and its Variants: The list coloring problem has been widely applied to solve optimization and scheduling problems [30]. In Orden and Moreira's work [31], the problem of minimizing interference threshold and the number of colors respecting that threshold was modeled as list coloring. They demonstrated that the problems are NP-Hard and proposed DSATUR, a graph coloring algorithm, to tackle them.

Garg et al. [32] tackled the channel frequency allocation problem in mobile communication networks by modeling it as a generalized list coloring problem. Their solutions prevented signal interference by selecting channels for neighboring base stations in a way that they did not overlap. This approach effectively addressed the crash failures caused by distance limitations.

In the domain of register assignment, Zeitlhofer et al. [33] presented a list-coloring algorithm that optimally assigns a large number of target variables to a small number of CPU registers. This algorithm preserves the structure of the interference graph, ensuring the retention of interval graph properties.

Sudoku puzzles can also be formulated as list coloring problems. Each cell corresponds to a vertex, and the relationships between cells are represented as edges in rows and columns. The numbers used in Sudoku can only appear once in each row and column, making it an instance of the list coloring problem. Additionally, Lastrina et al. [18] demonstrated how the precoloring extension problem can be used to illustrate the Sudoku puzzle.

4. Li k-PrExt and Futoshiki

In this section, we show that the Futoshiki problem can be reduced to the list precoloring extension problem for the Futoshiki graph G. In the reduction the revealed cells given in the Futoshiki problem are used to construct the pre-coloring for G. In addition, the list assignment is obtained using the inequality constraints.

Let $n \geq 2$ be a positive integer. A Partial Latin Square is an $n \times n$ grid that is partially filled with some numbers from $[k] = \{1, \ldots, n\}$. Let us denote a cell that is in the i'th row and the j'th column of a grid as (i, j). Each cell (i, j) is represented in the graph with a vertex v_{ij}. Two cells (i, j) and (i', j') are adjacent whenever they are in the same column or in the same row.

Thus, a Latin square is represented as a graph $G = (V, E)$ such that $V = \{v_{ij} : 1 \leq i, j \leq n\}$, and $E = \{(v_{ij}, v_{i'j'}) : ((i = i') \wedge (j \neq j')) \vee ((j = j') \wedge (i \neq i')\}$ [34]. This graph is called the *Futoshiki graph of size n*. Notice that the Futoshiki graph G has $n^2(n-1)$ edges. The graph G is isomorphic to the graph $K_n \boxtimes K_n$, which is the strong product of K_n with itself [18] and it is $(2n-2)$-regular.

Recall that the Futoshiki problem of size n, $\mathcal{F}_n(T, S)$, is defined on the Futoshiki graph of size n where S is the set of inequality constraints. Thus there are at most $2n(n-1)$ inequality signs. In Figure 3, for $n = 4$, there are $n^2 = 16$ vertices and $n^2(n-1) = 48$ edges. An instance of the problem on this graph can take up to 24 inequality constraints in total, yet in this instance, there are only 5 inequality constraints.

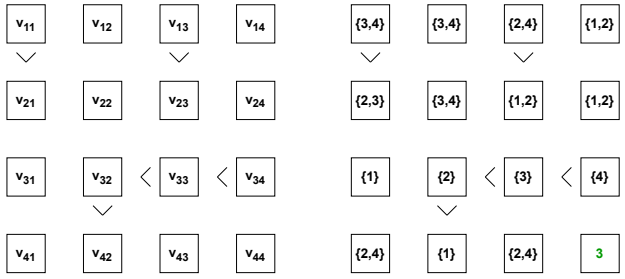


Figure 3. Vertices of the Futoshiki Puzzle.

Theorem 1. *For every Futoshiki problem $\mathcal{F}_n(T, S)$ there exists an equivalent instance of the* List k-Precoloring Extension *problem $\mathcal{L}_G(c_W, L)$ on the Futoshiki graph where $k = n$.*

Proof. We give a polynomial time reduction that converts a Futoshiki problem $\mathcal{F}_n(T, S)$ of size $n \times n$ to a Futoshiki graph G and a list-assignment L that corresponds to the list k-precoloring extension problem $\mathcal{L}_G(c_W, L)$ for some precoloring c_W. We also show that $\mathcal{F}_n(T, S)$ is solvable whenever $\mathcal{L}_G(c_W, L)$ is solvable on G. We do the latter by converting each solution of $\mathcal{F}_n(T, S)$ to a solution of $\mathcal{L}_G(c_W, L)$ and vice versa.

Given an instance $\mathcal{F}_n(T, S)$ of the Futoshiki problem the entries of the $n \times n$ board cells correspond to vertices of G, occurring exactly once in each row and column. Thus, the graph G will have n^2 vertices. The cells represent vertices and adjacent cells in each row and column represent the edges. At the beginning of the problem, if a number l is revealed in a cell that is represented with a vertex v, then we say v is precolored with color l. This gives a one-to-one correspondence between the revealed cells T and the precoloring c_W. If no number is revealed in the cell, then the corresponding vertex will be assigned the list $\{1, 2, \ldots, k\}$. If, in addition, there is an inequality sign $>$ located between some adjacent cells that are represented with vertices u and v in G, then for every color $i \in L(u)$, there must be at least one color $j \in L(v)$ such that $i > j$. For the inequality sign $<$, the construction of the list assignments of the related vertices are done similarly. This gives the construction of the list assignment L, thereby completing the reduction of $\mathcal{F}_n(T, S)$ to $\mathcal{L}_G(c_W, L)$.

Each solution of $\mathcal{F}_n(T, S)$ will naturally give a proper coloring for G which is an extension of the precoloring c_W and it will obey the list L. On the other hand, a list coloring solution of $\mathcal{L}_G(c_W, L)$ yields a solution to the given Futoshiki problem $\mathcal{F}_n(T, S)$. Notice that the precoloring and the list assignment L are constructed so that the solution to the list precoloring extension problem gives a number assignment that satisfies the inequalities located between adjacent cells.

□

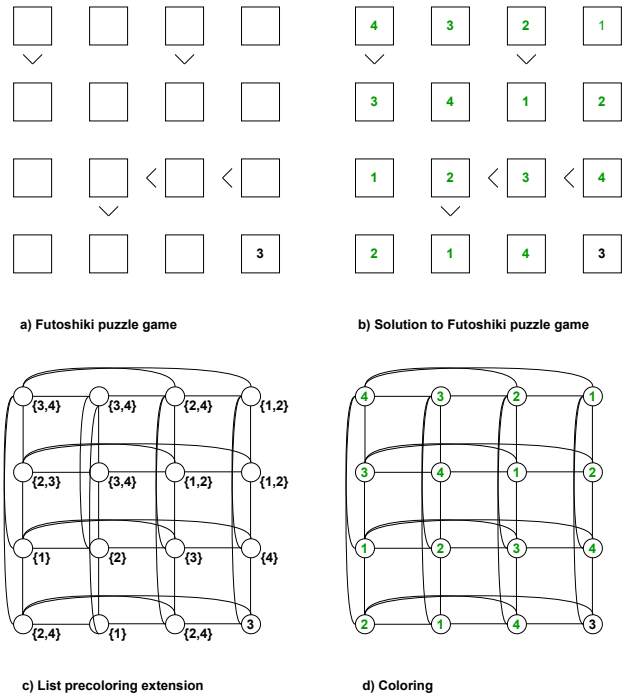


Figure 4. A Futoshiki puzzle instance and the cells that correspond to the vertices in the related Futoshiki Graph.

Figure 4 illustrates an instance of a 4×4 Futoshiki problem and the corresponding Futoshiki graph construction along with the list assignment: Figure 4.a shows the initial board of the game, Figure 4.b shows the solution of the game, Figure 4.c

shows the list precoloring extension instance and, finally, Figure 4.d shows the corresponding coloring as a solution to the list precoloring extension problem.

5. List coloring-based algorithm for the Futoshiki Problem

There are various algorithmic approaches to solve pencil puzzles. The backtracking algorithm is used for the class of Constraint Satisfaction Problems (CSPs). These problems are defined as a set of variables, a set of their respective domains of values, and a set of constraints [10]. The goal of a CSP is to find a consistent assignment of values to variables that satisfies all constraints, subject to certain conditions. CSPs have applications in various domains, including scheduling, planning, configuration, and puzzle-solving. LSCP-type puzzles are seen as constraint satisfaction problem CSPs that find a solution that satisfies all the constraints considering assignment of variables. These problems include Sudoku, Futoshiki, and Kakuro. Sudoku is the most popular one that is solved as CSP by the researchers. Norvig [10] describes two methods to solve Sudoku, namely, constraint propogation (CP) and Local Search. Constraint propagation is a technique that is commonly used in CSPs to efficiently update and reduce the domain of variables based on the constraints imposed by the problem.

In this section, we present two different deterministic algorithms for solving the Futoshiki Puzzle, called FutoshikiBT and ColorFutoshik, each incorporating backtracking, and filtering methods respectively.

5.1. BackTracking algorithm

The backtracking algorithm can be seen as the simplest solution for Sudoku puzzles which are the most commonly studied problem. Backtracking uses a recursive approach in which each cell is assigned a number from $1 \dots n$ when the board size is $n \times n$. The backtracking algorithm systematically explores the solution space by iteratively assigning values to empty cells in the puzzle and backtracking when a contradiction or violation of constraints is encountered.

Here, we give a backtracking algorithm that solves the Futoshiki Puzzle to compare its efficiency with the proposed method called ColorFutoshiki which we give in Section 5.2. The backtracking algorithm starts by selecting an empty cell in the puzzle and attempts to assign a value that satisfies the row and column constraints, as well as the inequality constraints associated with neighboring cells. It then moves on to the next empty cell and repeats the process. If a contradiction arises, such as a repeated number in a row or column, or a violation of an inequality constraint, the algorithm backtracks to the previous cell and explores alternative value assignments. Although this method guarantees a solution, it is not efficient in terms of time complexity. Let us present our FutoshikiBT Algorithm.

Algorithm 1 FutoshikiBT Algorithm

```
1:  Input: Futoshiki board with constraints.
2:  Output: Solution of the puzzle.
3:  if colorG(n, list, v = 1, given) == False
4:     print ("No solution")
5:  else
6:     print (list)
7:  colorG(n, list, v, given):
8:     if(v == V + 1)
9:        return True
10:    for c in range(1, n)
11:       if safe(v, list, c, given) == True
12:          list[v] = c
13:          if colorG(n, list, v + 1, given)
14:             return True
15:          end if
16:          if v not in given
17:             list[v] = 0
18:          end if
19:          return False
20:       end if
21:    end for
22:    return False
23: safe(v, list, c, given):
24:    if v in given and list[v] == c
25:       return True
26:    else if v in given
27:       return False
28:    end if
29:    for (i in range(1, V))
30:       if list[i] == c and neighbour(v, i)
31:          return False
32:       end if
33:       if constraints are not satisfied
34:          return False
35:       end if
36:    end for
37:    return True
```

5.2. ColorFutoshiki algorithm

In this section, using the equivalence between the Futoshiki Problem and the Li k-PrExt problem for the Futoshiki graph, as assured by Theorem 1, we construct the ColorFutoshiki algorithm to solve the Futoshiki Problem. We will observe that, this approach is equivalent to the backtracking algorithm with forward checking for the Futoshiki Problem.

The backtracking algorithm considers every solution by iterating every possible number in each cell under all satisfying conditions, assigns the first available option, backtracks when a solution is not possible for the next cell under consideration, and tries the next possible option for the previous cell. These methods guarantee the solution, but they do not give the solution in optimal time. Here we aim to improve the backtracking algorithm. This is why we need a problem space that helps our solver save us more time. The backtracking method of solving the Futoshiki problem fills each cells from left to right and top to bottom with considering inequality constraints.

The ColorFutoshiki algorithm is an improved version of the FutoshikiBT algorithm. Our aim is to reduce the number of colors in each list by eliminating inconsistent ones. This reduces the search space to be explored. At the beginning of the ColorFutoshiki algorithm, we use a filtering technique. In this filtering step we do forward checking in order to create color lists. Forward checking keeps track of the remaining possible values for unassigned variables after a variable is assigned a value. It propagates constraints by eliminating values from the domains of other variables that conflict with the newly assigned value. This technique is applied to create lists for each cell that they can use. Thus, it will begin coloring the puzzle with the minimum number of colors in the list for each cell.

5.3. Analysis of the algorithms

The input parameters for the algorithm are the Futoshiki graph of size n, the inequality constraints, and the pre-assigned entries. The algorithm begins by examining the constraints and pre-assigned numbers of the Futoshiki instance $\mathcal{F}_n(T, S)$, which then produces a k-precoloring instance $\mathcal{L}G(c_W, L)$. Next, it determines how to color the given Futoshiki graph using the list of colors assigned to each vertex through the reduction process described above. If $\mathcal{L}G(c_W, L)$ is a YES instance, the algorithm outputs a matrix $M_G = [mij]n \times n$ indicating the colors of the vertices of the graph G. Otherwise, it concludes that no solution exists.

First, let us analyze the FutoshikiBT algorithm, which builds candidates for the solutions incrementally and abandons candidates when it determines that they cannot possibly be solved with a valid solution.

The function colorG(n, list, v, given) is a recursive function that ensures the coloring process is completed by checking all vertices. It attempts to use the colors in the list for the corresponding vertex in order. Here, n represents the puzzle dimension, which is equal to the size of the color list. list is the list of colors assigned to vertices (solution). v is the vertex number. V represents the total number of vertices. given denotes the values given before the game starts.

The function safe(v, list, c, given) checks whether the given vertex can be colored with the chosen color by verifying the constraints. In this process, neighbour(v, i) checks the adjacency of the two relevant vertices, ensuring that adjacent vertices are not colored with the same color.

It is worth noting that the FutoshikiBT algorithm does not include a process for creating preassigned color lists. On the other hand, ColorFutoshiki first traverses the graph and creates a list of colors that minimizes the number of candidate colors for each cell. It then attempts to color empty cells, starting from the first vertex v_{11}.

The ColorFutoshiki algorithm is an improved version of the FutoshikiBT algorithm. Its aim is to reduce the number of colors in each list by eliminating inconsistent ones. This reduction effectively reduces the search space that needs to be explored.

The pseudocode of the algorithm is provided below.

Algorithm 2 ColorFutoshiki Algorithm

```
1: Input: Futoshiki board with constraints.
2: Output: Solution of the puzzle.

3: if colorG(pcList[1].length, list, v = 1, given)
4:    print (list)
5: else
6:    print ("No solution")
7: end if

8: colorG(n, list, v, given):
9:    if (v == V + 1)
10:      return True
11:   end if
12:   for c in range(1, n)
13:    if safe(v, list, pcList[v][c], given)
14:      list[i] = pcList[v][c]
15:      nextN = pcList[v + 1].length
16:      if colorG(nextN, list, v + 1, given)
17:         return True
18:      end if
19:      if v not in given
20:         list[v] = 0
21:      end if
22:      return False
23:    end if
24:   end for
25:   return False

26: safe(v, list, c, given):
27:   if v in given and list[v] == c
28:      return True
29:   else if v in given
30:      return False
31:   end if
32:   if constraints are not satisfied
33:      return False
34:   end if
35:   for i in range(1,n)
36:      if (hNeighbor == c & hNeighbor! = v)
37:         return False
38:      end if
39:      if (vNeighbor == c & vNeighbor! = v)
40:        return False
41:      end if
42:   end for
43:   return True
```

In the ColorFutoshiki algorithm, first, we perform filtering and create color lists for each cell based on their admissible colors and constraints. Consequently, the algorithm begins coloring the puzzle using the minimum number of colors available in the lists.

Unlike the FutoshikiBT algorithm, ColorFutoshiki uses a recursive structure that does not traverse the entire graph to color the related vertex. Instead, it only checks the horizontal and vertical neighbors of the vertex being colored. It is observed that ColorFutoshiki outperforms the FutoshikiBT algorithm in all instances of varying difficulty levels.

Now, let us explain the ColorFutoshiki algorithm.

In lines 3-7, we call the colorG function, which displays the solution if found. Here pcList is the list that each cell takes.

In lines 8-25, the recursive colorG function checks the termination condition. If this condition is not met, it checks whether the related vertex v can be colored based on the possible color list of v. If v cannot be colored, the algorithm moves on to the next color in its list. If it cannot be colored with any color in the list, the recursive function returns to previous vertex and the color of the previous vertex is updated. If it can be colored, the algorithm moves on to the next vertex, and the admissible color is added to the color list.

In lines 26-43, if the vertex to be colored has a preassigned (given) value, the algorithm proceeds to the next vertex. Then, it checks whether there is a constraint in front of or above the vertex to be colored or not. If there is, it verifies whether the constraint conditions are satisfied. After this step, we check the colors of the adjacent vertices if they have the same color. Instead of traversing all nodes, it only checks the horizontal and vertical neighbors of the relevant vertex.

Now let us analyze the time complexities of the algorithms. The FutoshikiBT algorithm traverses all vertices to check the neighborhood of the vertex being colored and to decide whether the colors are the same or not. For each empty cell, there are n possible options, where n is the total number of colors. As a result, the time complexity becomes $O(n^{n^2})$. In the ColorFutoshiki algorithm some values are removed from some domains. Since there will be some early pruning the time taken will be much less than the backtracking algorithm. However, the upper bound time complexity remains the same. The reason is that we don't know how many values are removed. As a result, the time complexity becomes $O(n^{n^2})$.

Although the worst case time complexity of the ColorFutoshiki algorithm is only a slight improvement over the FutoshikiBT algorithm, as we will observe below, its speed is remarkably faster in all of the computational experiments that we have done.

6. Results and discussion

We aim to solve the FUTOSHIKI as an instance of the LICOL problem. For this reason we first improved the **ColorFutoshiki** algorithm. It is an enumeration algorithm that incorporates additional search space reductions and bounding elements, making it an enhanced version compared to FutoshikiBT. Due to the fact that FUTOSHIKI is classified as an NP-Complete problem, ColorFutoshiki proves to be highly effective and suitable for numerous applications and purposes. Its capabilities make it a valuable tool in addressing the complexity of FUTOSHIKI.

Here, we present the computational experiments conducted to assess the efficiency of the ColorFutoshiki algorithm in solving instances of the Futoshiki problem. All the codes were implemented in the Python programming language, and the experiments were executed on a system with an Intel Core i7-6700HQ CPU operating at 2.60 GHz, with 16GB RAM, running Windows 10 (64-bit). Different instances have been generated and tested on nxn boards for $n = 6, 7, 8, 9, 15, 20, 30, 40, 50$. We ran the related code 50 times per instance and took their average. In total, 1800 Futoshiki puzzles have been used for each algorithms. Due to space limitations, we are unable to provide a table containing the running times for each individual test. This is why, we present the average experimental results for each algorithm in Table 1 and Table 2.

Solutions to the FUTOSHIKI Problem are of particular interest due to the given application in scheduling on $n \times n$ boards when the size n is large. As our experimental results demonstrate, the ColorFutoshiki algorithm works much better even on boards of a larger size where the number of inequality constraints is not necessarily restricted to be less than n for an $n \times n$ board.

A standard deviation denotes the spread of data concerning its mean. When the standard deviation is small, it signifies that the data is tightly clustered around the mean. Conversely, a high or large standard deviation suggests that the data is more widely spread.

All the results and standard deviations can be seen in Table 1. Different numbers of constraints and different numbers of givens are examined for each algorithm.

The performance of the proposed algorithms for standard search algorithms is illustrated in Figure 5 and Figure 6 with standard deviations. In this figures, we maintain a constant number of constraints while showcasing the increasing number of given values. Notably, ColorFutoshiki consistently outperforms FutoshikiBT in solving puzzles, even when the number of constraints is held constant. These results underscore the efficiency of the ColorFutoshiki algorithm, particularly on larger-sized boards, providing a performance comparison with the FutoshikiBT algorithm.

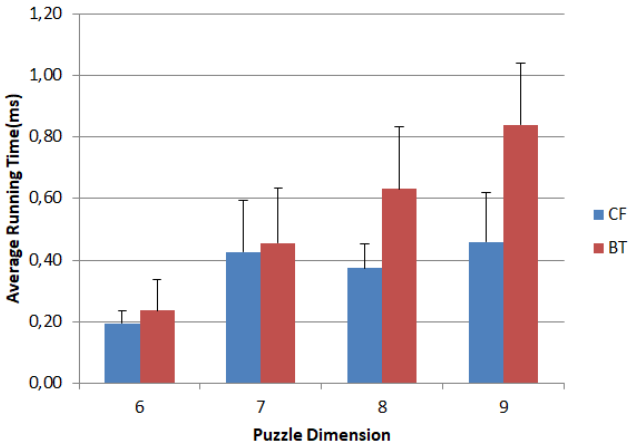


Figure 5. Comparison of **ColorFutoshiki**(CF) and **FutoshikiBT**(BT) on smaller board sizes.

In addition to running times, we also measure the number of operations for both ColorFutoshiki and FutoshikiBT. Since the number of explored nodes provides insights into the efficiency of the algorithm, we show both the number of explored nodes and the number of removed values of each algorithm in Table 2. A lower number of explored nodes generally indicates a more efficient algorithm, as it suggests that the algorithm is able to reach a solution without exhaustively searching through a large portion of the puzzle's solution space. The number explored nodes allows for comparison with other algorithms or approaches for solving the same puzzle. For this reason, we use this parameter to compare the ColorFutoshiki algorithm that we present with FutoshikiBT.

We observe that the number of explored nodes varies significantly across different instances of the puzzle, it may indicate that the algorithm's performance is sensitive to certain characteristics of the puzzle. Similar to the number of explored nodes, the number of removed nodes provides insight into the efficiency of the algorithm. In certain search algorithms, such as backtracking or constraint satisfaction algorithms, removed nodes typically refer to nodes that are pruned from the search space because they are deemed unnecessary or invalid. A lower number of removed nodes indicates that the algorithm is effectively pruning the search space, which can lead to improved

Table 1. Run time for all algorithms reported in milliseconds.

size	inequalities	givens	CF	BT
6	6	1	1.54 ± 0.11	1.61 ± 0.44
6	6	20	0.28 ± 0.05	0.29 ± 0.1
6	6	30	0.19 ± 0.04	0.24 ± 0.1
6	1	6	0.59 ± 0.08	0.66 ± 0.22
6	20	6	0.5 ± 0.08	0.54 ± 0.27
6	30	6	0.53 ± 0.17	0.4323 ± 0.1073
7	7	1	1.40 ± 0.16	1.47 ± 0.52
7	7	30	0.43 ± 0.17	0.46 ± 0.18
7	7	40	0.29 ± 0.04	0.35 ± 0.08
7	1	7	1.37 ± 0.3	1.58 ± 0.5
7	30	7	2.07 ± 0.46	2.33 ± 0.58
7	40	7	1.19 ± 0.37	1.21 ± 0.52
8	8	1	3.53 ± 0.80	3.64 ± 1.33
8	8	40	0.5 ± 0.04	0.77 ± 0.29
8	8	50	0.37 ± 0.08	0.63 ± 0.2
8	1	8	2.01 ± 0.33	2.29 ± 0.73
8	40	8	6.72 ± 1.71	9.33 ± 2.32
8	50	8	5.34 ± 1.55	10.04 ± 2.3
9	9	1	8.08 ± 3.11	8.11 ± 3.06
9	9	70	0.46 ± 0.16	0.84 ± 0.2
9	18	35	1.28 ± 0.32	1.92 ± 0.63
9	1	9	3.78 ± 1.33	4.7 ± 1.69
9	70	9	8.86 ± 0.3	10.71 ± 0.92
9	35	18	2.27 ± 3.27	2.95 ± 3.22
15	15	150	2.77 ± 0.64	11.65 ± 2.54
15	15	170	2.25 ± 0.86	11.75 ± 2.86
15	15	200	1.44 ± 0.45	10.52 ± 3.94
15	1	150	11.82 ± 0.7	12.04 ± 3.26
15	90	130	3.47 ± 1.10	12.61 ± 3.33
15	150	15	2987.7 ± 336.8	3064.9 ± 103.2
20	20	300	4.98 ± 1.48	38.54 ± 8.12
20	30	320	11.58 ± 1.27	62.2 ± 3.43
20	20	370	1.575 ± 1.62	31.57 ± 6.07
20	30	300	5.21 ± 3.48	35.98 ± 8.52
20	80	250	349.97 ± 1.42	644.16 ± 7.26
20	50	350	3.55 ± 27.31	33.27 ± 29.87
30	30	500	11.000 ± 37.377	2585.5 ± 133.84
30	30	600	612.73 ± 34.274	1360.2 ± 441.69
30	30	750	41.852 ± 4.9579	1192.3 ± 78.302
30	50	500	1103.7 ± 135.63	2412.5 ± 102.47
30	100	600	608.17 ± 29.557	1274.6 ± 71.195
30	300	750	40.34 ± 5.8942	993.95 ± 34.361
40	40	750	2827.7 ± 97.233	12159 ± 402.91
40	40	950	702.37 ± 33.654	9869.8 ± 234.02
40	40	1300	91.642 ± 8.8446	1257.7 ± 51.235
40	100	750	2913.0 ± 599.59	9359.7 ± 255.39
40	300	950	437.39 ± 23.055	4845.2 ± 130.88
40	400	950	438.99 ± 27.05	4995.3 ± 142.66
50	50	1200	4443.1 ± 151.71	37727 ± 579.75
50	100	1500	2158.0 ± 93.02	21803.0 ± 358.64
50	250	2000	1817.7 ± 75.545	20849 ± 538.47
50	400	1600	1092.8 ± 54.712	13279 ± 162.22
50	400	2000	1067.5 ± 39.437	13145 ± 263.9
50	500	2000	835.95 ± 40.199	11559 ± 284.97

efficiency as it can be seen in Table 2. As the FutoshikiBT algorithm lacks a filtering step, no values are removed from its domains, unlike the filtering steps in the ColorFutoshiki algorithm.

The running time of the algorithm varies depending on the puzzle's difficulty level. We observe that as the number of inequality signs approaches the maximum limit, the computation time significantly decreases. This behavior can be attributed to the utilization of the backtracking method in the ColorFutoshiki algorithm. Typically, Futoshiki puzzles are played on 5×5 to 9×9 boards (occasionally on 15×15 boards), and existing algorithmic solutions are primarily tested on boards with dimensions up to $n = 9$.

In Figure 5, we compare the performance of our method with previous solutions employing the FutoshikiBT algorithm on smaller-sized boards.

Motivated by the lack of performance analysis for larger-sized boards, we conducted experiments using the ColorFutoshiki algorithm on larger board sizes. Additionally, we wanted to assess the algorithm's effectiveness on larger-sized boards due to the relationship between the Futoshiki puzzle game and larger scheduling problems, as discussed in Section 3. In Figure 6, we compare the performance of our method with previous solutions employing the FutoshikiBT algorithm on larger-sized boards. We even conducted experiments for 50×50 boards. The results demonstrate that the ColorFutoshiki algorithm efficiently solves even 30×30, 40×40 and 50×50 board games as shown in Figure 7.

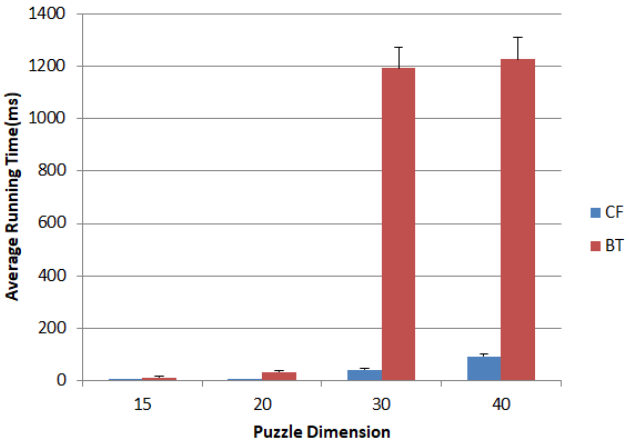


Figure 6. Comparison of **ColorFutoshiki**(CF) and **FutoshikiBT**(BT) on larger board sizes.

Overall, these findings highlight the efficiency of the ColorFutoshiki algorithm, especially on larger-sized boards, and provide a performance comparison with the FutoshikiBT algorithm.

7. Conclusion

The Futoshiki problem is aimed to be solved as a list coloring problem in ColorFutoshiki. It is an enumeration algorithm that incorporates additional search space reductions and bounding elements, making it an enhanced version compared to FutoshikiBT. Due to the fact that FUTOSHIKI is classified as an NP-Complete problem, ColorFutoshiki proves to be highly effective and suitable for numerous applications and purposes. Its capabilities make it a valuable tool in addressing the complexity of FUTOSHIKI.

A considerable number of experiments were conducted to test ColorFutoshiki and FutoshikiBT, providing a robust foundation for drawing meaningful conclusions. The extensive set of experiments carried out ensures that the findings are sufficiently supported and reliable. Observing Table 1 and Table 2, we see that ColorFutoshiki is much more efficient than FutoshikiBT.

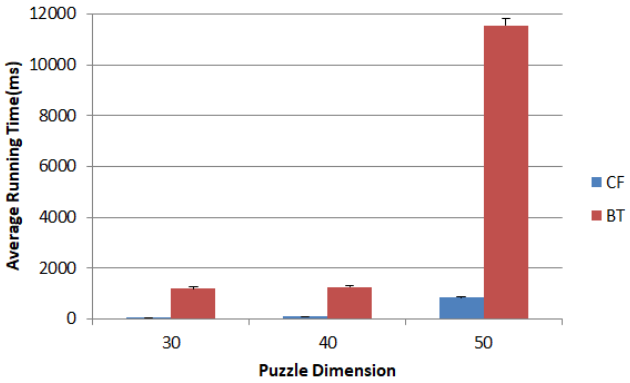


Figure 7. Comparison of **ColorFutoshiki**(CF) and **FutoshikiBT**(BT) on larger board sizes.

We incorporate a short discussion on these metaheuristic methods and their applications to Sudoku. This will enhance the comprehensiveness of our paper and provide a broader perspective on the algorithmic techniques used to solve LSCP type puzzles.

We are also interested in studying the Futoshiki problem as an application of the LI k-COL problem. To find approximation algorithms for the Futoshiki puzzle, we would like to use metaheuristics.

As for future work related with nature based algorithms, one can see whether an Ant Colony optimization (ACO) algorithm gives a more efficient algorithm to solve the Futoshiki problem. ACO is a Swarm intelligence algorithm which is one of the artificial intelligence techniques. A solution

Table 2. The count of removed nodes and explored nodes. NRN stands for "No Removed Nodes".

size	inequalities	givens	FC-EN	FC-RN	BT-EN	BT-RN
6	6	1	528	15	540	NRN
6	6	20	40	173	126	NRN
6	6	30	36	180	126	NRN
6	1	6	151	78	198	NRN
6	30	6	111	81	156	NRN
6	20	6	100	81	144	NRN
7	7	1	456	18	469	NRN
7	7	30	59	276	196	NRN
7	7	40	50	291	196	NRN
7	1	7	298	113	406	NRN
7	40	7	268	125	392	NRN
7	30	7	474	124	392	NRN
8	8	1	1145	29	1160	NRN
8	8	50	65	446	288	NRN
8	8	40	78	423	288	NRN
8	1	8	580	156	736	NRN
8	40	8	1851	161	2744	NRN
8	50	8	1293	174	2920	NRN
9	9	1	2279	24	2313	NRN
9	9	70	81	647	405	NRN
9	18	35	187	546	603	NRN
9	1	9	988	194	1357	NRN
9	70	9	1629	213	2070	NRN
9	35	18	489	385	855	NRN
15	15	150	286	3040	2010	NRN
15	15	170	239	3102	1800	NRN
15	15	200	229	3139	1800	NRN
15	1	150	286	3028	2010	NRN
15	90	130	331	2975	2025	NRN
15	150	15	674261	7472	240	NRN
20	20	300	472	7589	4200	NRN
20	30	320	760	7484	5940	NRN
20	20	370	408	7588	4220	NRN
20	30	300	472	7484	4200	NRN
20	80	250	21057	6922	49580	NRN
20	50	350	417	7578	4220	NRN
30	30	500	44063	22402	116550	NRN
30	30	600	24682	23577	56400	NRN
30	30	750	1722	25450	48060	NRN
30	50	500	41511	22403	111840	NRN
30	100	600	22654	23580	52560	NRN
30	300	750	1684	25510	41220	NRN
40	40	750	69967	54498	229320	NRN
40	40	950	17714	56006	537760	NRN
40	40	1300	2542	61272	43160	NRN
40	100	750	33348	51448	219360	NRN
40	300	950	11154	56182	220240	NRN
40	400	950	11124	56224	220240	NRN
50	50	1400	59494	119420	600200	NRN
50	100	1600	29404	120771	462900	NRN
50	250	2000	25673	120862	427900	NRN
50	400	1600	16255	120912	302150	NRN
50	400	2000	16255	120912	302150	NRN
50	500	2000	12864	120979	260000	NRN

for Sudoku is given using ACO in the study of Huw Lloyd [16]. Another solution which is the first nature-based algorithm for the NP-Complete Nurikabe problem is presented by Amos et al. [35]. This algorithm was developed based on ACO. For future work, it would be interesting to solve the Futoshiki problem using Ant Colony Optimization (ACO) and an Artificial Bee Colony (ABC)

algorithm [36]. The performance of these approaches could then be compared with existing solutions, such as the improved constraint programming method developed by Kostyukova and Tchemisova [37].

References

[1] Bobga, B., Goldwasser, J.L., Hilton A.J.W. & Johnson P.D. (2011). Completing partial latin squares: Cropper's question. *Australasian Journal of Combinatorics*, 49, 127-152.

[2] Goldwasser, J., Hilton A., Hoffman D.G. & Ozkan, S. (2015). Hall's theorem and extending partial latinized rectangles. *Journal of Combinatorial Theory Series A*, 130, 26-41. https://doi.org/10.1016/j.jcta.2014.10.007

[3] Euler, R. (2010). On the completability of incomplete Latin squares. *European Journal of Combinatorics*, 31, 535-552. https://doi.org/10.1016/j.ejc.2009.03.036

[4] Colbourn, C.J. (1984). The complexity of completing partial latin squares. *Discrete Applied Mathematics*, 8, 25-30. https://doi.org/10.1016/0166-218X(84)90075-1

[5] Haraguchi, K. & Ono, H. (2014). Approximability of Latin square completion type puzzles. *International Conference on Fun with Algorithms*, 218-229. https://doi.org/10.1007/978-3-319-07890-8_19

[6] Donovan, D. (2010). The completion of partial latin squares. *Australasian Journal of Combinatorics*, 22, 247-264.

[7] Galvin, F., Stephen, C.L., Kim, S.S. & Callan, D. (2001). A Generalization of Hall's Theorem: 10701. *The American Mathematical Monthly*, 108, 79-80. https://doi.org/10.2307/2695691

[8] Ivanyi, A. & Nemeth, Z. (2011). List coloring of Latin and Sudoku graphs. *8th Joint Conf. on Math. and Comp. Sci.*

[9] Yato, T. & Seta, T. (2003). Complexity and completeness of finding another solution and its application to puzzles. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, E86-A, 5, 1052-1060.

[10] Norvig, P. (2018). Solving every Sudoku puzzle. http://norvig.com/sudoku.html.

[11] Crawford, B., Castro, C. & Monfroy, E. (2009). Solving sudoku with constraint programming. *MCDM, CCIS Communications in Computer and Information Science*, 35, 345-348. https://doi.org/10.1007/978-3-642-02298-2_52

[12] Knuth, D.E. Dancing links. (2000). Millennial Perspectives in Computer Science. *Proceedings of the 1999 Oxford-Microsoft Symposium*, 187-214.

[13] Glover, F.W. (1986). Future paths for integer programming and links to artificial intelligence. *Computers & Operations Research*, 13(5), 533-549. https://doi.org/10.1016/0305-0548(86)90048-1

[14] Pacurib, J.A., Seno, G.M.M. & Yusiong, J.P.T. (2009). Solving Sudoku puzzles using improved artificial bee colony algorithm. *In Fourth International Conference on Innovative Computing, Information and Control (ICICIC)*, 885-888. https://doi.org/10.1109/ICICIC.2009.334

[15] Moraglio, A., & Togelius, J. (2007). Geometric particle swarm optimization for the Sudoku puzzle. *In Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation (GECCO)*, 118-125. https://doi.org/10.1145/1276958.1276975

[16] Lloyd, H. & Amos, M. (2020). Solving sudoku with ant colony optimization, *IEEE Transactions on Games*, 12, 302-311. https://doi.org/10.1109/TG.2019.2942773

[17] Musliu, N., & Winter, F. (2017). A Hybrid Approach for the Sudoku Problem: Using Constraint Programming in Iterated Local Search, *IEEE Intelligent Systems*, 32 (2), 52-62. https://doi.org/10.1109/MIS.2017.29

[18] Lastrina, M.A. (2012). *List-coloring and sum-list coloring problems on graphs*. PhD Thesis, Iawo University.

[19] Tuza, Z. (1997). Graph colorings with local constraints - a survey. *Discussiones Mathematicae Graph Theory*, 17, 161-228. https://doi.org/10.7151/dmgt.1049

[20] Karp, R.M. (1972). Reducibility among Combinatorial Problems. In: Complexity of Computer Computations. *The IBM Research Symposia Series*, Boston, MA, Springer. https://doi.org/10.1007/978-1-4684-2001-2_9

[21] Kratochvil, J., & Tuza, Z. (1994). Algorithmic complexity of list coloring. *Discrete Applied Mathematics*, 50(3), 297-302. https://doi.org/10.1016/0166-218X(94)90150-3

[22] Diestel, R. (2017). *Graph Theory*. Graduate Texts in Mathematics, Heidelberg: Springer-Verlag. https://doi.org/10.1007/978-3-662-53622-3

[23] Vizing, V.G. (1976). Coloring the vertices of a graph in prescribed colors. *Diskret. Analiz., Metody Diskret. Anal. v. Teorii Kodov i Shem*, 101, 3-10.

[24] Lovász, L. (1973). Coverings and coloring of hypergraphs. *Proc. 4th Southeastern Conf. on Combinatorics, Graph Theory and Computing*, 3-12.

[25] Erdos, P. & Rubin, A.L., Taylor. (1979). Choosability in graphs. *Proceedings of the West Coast Conference on Combinatorics, Graph Theory and Computing*, 26, 125-157.

[26] Garey, M.R. & Johnson, D.S. *Computers and Intractability, A guide to the theory of NP-Completeness*. W. H. Freeman and Co., 1979.

[27] Mahmood, A.S. (2019). Design random number generator utilizing the Futoshiki puzzle. *Journal of Information Hiding and Multimedia Signal Processing*, 10, 178-186.

[28] Haraguchi, K. (2013). The number of inequality signs in the design of Futoshiki puzzle. *Journal of Information Processing*, 21, 26-32. https://doi.org/10.2197/ipsjjip.21.26

[29] Sahu, H.S., Nayak, S.K. & Mishra, S. (2016). Maximizing the power generation of a partially shaded PV array. *IEEE Journal of Emerging and Selected Topics in Power Electronics*, 4, 626-637. https://doi.org/10.1109/JESTPE.2015.2498282

[30] Bondy, J.A. & Murty, U.S.R. (2008). *Graph Theory, Graduate Texts in Mathematics*. Springer, New York. https://doi.org/10.1007/978-1-84628-970-5

[31] Orden, D., Marsa, M.I., Gimenez, G.J.M. & Hoz, E. (2017). Spectrum graph coloring and applications to Wi-Fi channel assignment. *Symmetry*, 10(3), 65. https://doi.org/10.3390/sym10030065

[32] Garg, N., Papatriantafilou M. & Tsigas, P. (1996). Distributed list coloring: how to dynamically allocate frequencies to mobile base stations. *Eighth IEEE Symposium on Parallel and Distributed Processing*, 18-25. https://doi.org/10.1109/SPDP.1996.570312

[33] Zeitlhofer, T. & Wess, B. (2003). List-coloring of interval graphs with application to register assignment for heterogeneous register-set architectures. *Signal Processing*, 83 (7), 1411-1425. https://doi.org/10.1016/S0165-1684(03)00089-6

[34] Denes, J. & Keedwell, AD. (1991). *Latin Squares. New Developments in the Theory and Applications*, North Holland.

[35] Amos, M., Crossley, M. & Lloyd, H. (2019). Solving nurikabe with ant colony optimization. *GECCO '19: Proceedings of the Genetic and Evolutionary Computation Conference Companion*, 129-130. https://doi.org/10.1145/3319619.3338470

[36] Bektur, G. & Aslan, H.K. (2024). Artificial bee colony algorithm for operating room scheduling problem with dedicated/flexible resources and cooperative operations. *An International Journal of Optimization and Control: Theories & Applications (IJOCTA)*, 14(3), 193-207. https://doi.org/10.11121/ijocta.1466

[37] Kostyukova, O. & Tchemisova T. (2024). Exploring constraint qualification-free optimality conditions for linear second-order cone programming. *An International Journal of Optimization and Control: Theories & Applications (IJOCTA)*, 14(3), 168-182. https://doi.org/10.11121/ijocta.1421

Banu Baklan Şen *completed her BSc in Mathematics and Computer Science and her MSc in Information Tehcnologies at Bahçeşehir University, Turkey. She defended her PhD in Computer Engineering at Kadir Has University in Turkey. Her academic work includes graph theory algorithms and theoretical computer science. She has participated in international academic collaborations. Her research has led to publications in prominent journals and presentations at international conferences with the expertise in combinatorial optimization and algorithms. She is deeply committed to both teaching and research.*

https://orcid.org/0000-0003-4545-5044

Öznur Yaşar Diner *obtained her BSc in Mathematics from the Middle East Technical University in Turkey. She completed her MSc in Mathematics at the University of Göttingen in Germany and defended her Ph.D. in Mathematics at Memorial University of Newfoundland in Canada. Her primary research interests lie in structural graph theory and theoretical computer science, with a focus on combinatorial problems. She is passionate about teaching and conducting collaborative research.*

https://orcid.org/0000-0002-9271-2691

This work is licensed under a Creative Commons Attribution 4.0 International License. The authors retain ownership of the copyright for their article, but they allow anyone to download, reuse, reprint, modify, distribute, and/or copy articles in IJOCTA, so long as the original authors and source are credited. To see the complete license contents, please visit http://creativecommons.org/licenses/by/4.0/.

An International Journal of Optimization and Control: Theories & Applications
ISSN: 2146-0957 eISSN: 2146-5703
Vol.14, No.4, pp.308-321 (2024)
https://doi.org/10.11121/ijocta.1462

RESEARCH ARTICLE

Early prediction of fabric quality using machine learning to reduce rework in manufacturing processes

Sema Aydın, Koray Altun*

Department of Industrial Engineering, Bursa Technical University, Bursa, Turkey
semaydinn.16@gmail.com, koray.altun@btu.edu.tr

ARTICLE INFO

Article history:
Received: 1 October 2023
Accepted: 11 September 2024
Available Online: 9 October 2024

Keywords:
Fabric quality
Rework reduction
Machine learning
Artificial neural networks

AMS Classification 2010:
97R40, 60G25

ABSTRACT

The increasing competition and rapid technological advancements in today's business world have raised customer expectations. People now expect quick delivery, low prices, and high-quality products. As a result, companies must adapt to this competitive environment to survive. Rework, which is a significant cost in production, increases expenses, reduces production efficiency, and can lead to customer attrition. Research shows various efforts across different sectors to reduce rework, although there is still a gap in the textile sector's fabric dyeing units. Common problems in these units include non-retentive colors, customer dissatisfaction with shades, and repeated dyeing due to environmental factors or dye vat issues. This study uses logistic regression and artificial neural networks models from machine learning to predict which fabrics will need rework, using data from a textile company in Bursa. The analysis indicates that artificial neural networks models perform better.

(cc) BY

1. Introduction

In the textile industry, optimizing fabric dyeing processes is a pivotal challenge. Rework processes within fabric dyeing units are among the most critical factors contributing to cost escalation, low-quality production, and customer dissatisfaction.

“Rework” can be defined as the need for additional processing or corrective measures due to various quality issues. Rework represents a form of waste, driving research into the concept of Zero Defect Manufacturing [1,2]. This concept aims to eliminate defects before they necessitate rework. However, achieving this goal requires proactive measures to anticipate and prevent potential quality issues before they escalate into rework processes.

Machine learning techniques offer promising approaches to identifying patterns in historical data and predicting defects in manufacturing processes. Although there is growing interest in using machine learning for quality assurance and defect detection in manufacturing, a significant gap remains in research on applying these techniques specifically for "rework prediction" in the textile industry.

Correspondingly, this research aims to address this gap by evaluating the effectiveness of logistic regression and artificial neural networks (ANNs) in predicting rework instances in fabric dyeing processes.

In line with the study's objectives and the research landscape, the following contributions are emphasized:

- This study addresses a critical research gap in the textile industry by exploring the application of machine learning for predicting errors in fabric dyeing processes. While machine learning has been extensively applied in various industries, its use in the textile sector for rework prediction remains relatively underexplored.
- This study focuses on applying machine learning, specifically logistic regression and ANNs, to develop models that predict rework in fabric dyeing units within the textile industry. By leveraging data-driven methods, our goal is to enhance the early detection of potential quality issues and reduce the need for rework before it escalates a bigger problem.

The aim is to proactively predict and mitigate potential quality issues, thereby optimizing production processes and minimizing instances of rework. This study not only contributes to the expanding field of predictive

*Corresponding author

analytics and manufacturing optimization but also seeks to advance the application of machine learning within the textile sector.

2. Literature review

A review of the literature on early quality prediction and rework reveals numerous studies across various industries. However, the application of these techniques specifically within the textile sector appears to be less common.

A systematic effort was undertaken to conduct a comprehensive literature review. Prominent databases such as Science Direct, Web of Science, and Google Scholar were utilized to identify relevant materials. Two specific search strings were employed to reflect the core focus of the investigation. The first search string, "quality" AND "defect detection" AND "prediction" AND "manufacturing," aimed to cover the scholarly work on predictive methods for quality assurance and defect detection in manufacturing contexts. The second search string, "rework" AND "defect" AND "machine learning," was designed to explore research on the application of machine learning approaches to predict defects and prevent rework processes. By systematically using these search strings and examining the results from the selected databases, the literature review aimed to extract and synthesize relevant insights from the existing body of scholarly work.

A case study [1] was conducted within an automotive company, employing an Early Quality Prediction system grounded in a data-driven approach. In this study, the focus was on applying Convolutional Neural Network (CNN) techniques to time-series data to proactively predict and prevent defects. This approach aimed to minimize rework costs and optimize product quality through predictive insights. A data-driven mathematical model has been developed [3] for a dynamic manufacturing process with multiple rework lines, focusing on calculating production rates and machine efficiency for each machine. This model is designed to address scenarios involving preventive maintenance activities for each machine, with machine efficiencies computed based on real performance data for analysis. Taking into consideration the dynamic states of the production system, including rework productions, a mathematical model has been developed [4] to calculate efficiency of the process. With this model, rework strategies are predicted to ensure the production of products with the desired quality. In the context of assembly lines, rework stations are typically set up at the end of assembly lines for reprocessing faulty products. These rework stations operate within standard production processes when error rates are low. In cases where rework stations operate in dynamic conditions, a nonlinear mixed-integer programming model [5] has been proposed to enhance station efficiency. This model aims to increase the efficiency of rework stations in dynamic situations, thus improving the overall performance of the assembly line. In the automotive industry, various Machine Learning methods have been employed to predict errors in assembly lines. The results obtained from different methods have been compared [6]. To facilitate this comparison, specific metrics were established, and these metrics were contrasted among the six algorithms employed.

To provide a comprehensive overview of the current state of research on early quality prediction and rework processes, a detailed literature summary table (Table-1) has been compiled. This table highlights the key methodologies, applications, and findings from various studies, illustrating the broader landscape of machine learning applications in different sectors.

A review of the literature on the application of artificial neural networks in the textile sector reveals that they have been employed in various studies to predict yarn parameters, optimize weaving processes, enhance finishing stages, and assess fabric comfort parameters [7]. Artificial Neural Networks (ANN) has also been employed to detect fabric defects in weaving [8]. To predict the characteristics of woven fabric (width, weight, weft and warp tensile values), an advanced feedforward recurrent artificial neural network (ANN) model [9] was employed and compared with a linear regression model. Artificial Neural Network (ANN) methods have been employed [10] to predict weft errors that emerge during fabric production in a textile company. ANN methods have been employed [11] to predict the impact of chemical finishing processes on the CIELab value of the fabric color.

This study aims to develop alert systems for the identification of products with a high probability of rework and the implementation of specific actions for their mitigation. Early Quality Prediction systems employ deep learning methods as well as various techniques for this purpose. However, within the textile industry, specifically in a textile company with a fabric dyeing production process, there is a lack of research on using Machine Learning for the prediction of errors. This particular gap forms the distinct aspect of this study.

The remainder of this paper is organized as follows: Section 3 presents the problem statement, detailing the specific challenges addressed in this study. Section 4 provides the methodology. Section 5 discusses the results. Section 6 explores managerial insights and practical implementations, offering actionable recommendations based on the study's outcomes. Finally, Section 7 concludes the paper with a summary of key findings and an outlook on future research directions.

Table 1. Summary table of the literature.

Ref.	Methods used	Application area	Key findings
[1]	Deep learning, Time series data	Manufacturing	Successful use of deep learning techniques for early quality prediction.
[2]	Various machine learning methods	Zero defect manufacturing	Comprehensive review of current methods to achieve zero defect manufacturing.
[3]	Data-driven modeling and analysis	Multi-stage manufacturing systems	Used data-driven modeling to analyze quality rework cycles.
[4]	Product traceability and rework analysis	Manufacturing systems	Analyzed quality performance considering product traceability and rework.
[5]	Mixed-integer programming	Assembly lines	Developed a model for positioning rework stations to improve efficiency.
[6]	Machine learning	Assembly environment	Applied machine learning for error detection in low-automation assembly environments.
[7]	Artificial neural networks	Textile industry	Reviewed applications of artificial neural networks in the textile industry.
[8]	Artificial neural networks	Weaving technology	Comprehensive review of ANN applications in weaving technology.
[9]	Artificial neural networks	Woven fabric	Used advanced feedforward recurrent neural networks to predict woven fabric properties.
[10]	Artificial neural networks, Multiple linear regression	Fabric defects	Compared ANN and multiple linear regression models for predicting fabric defects.
[11]	Artificial neural networks	Chemical finishing processes	Predicted the impact of chemical finishing on fabric color using ANNs.
[12]	Robust mathematical model	Epidemic modeling	Developed a mathematical model to predict the course of the COVID-19 epidemic.
[13]	Machine learning	Agri-food production forecasting	Used robust and resilient machine learning methods to forecast agri-food production.
[14]	Machine learning, Deep learning	Healthcare	Applied machine learning techniques to improve early disease detection.
[15]	Support vector machines, Neural networks	Manufacturing	Compared the efficiency of SVM and ANN in predicting manufacturing defects.
[16]	Deep learning	Electronics manufacturing	Used deep learning to predict defects in electronics manufacturing.
[17]	Logistic regression, Decision trees	Automotive	Developed predictive models to reduce rework in automotive manufacturing.
[18]	Ensemble learning	Aerospace	Applied ensemble learning techniques for quality prediction in aerospace components.
[19]	Neural networks, Fuzzy logic	Food processing	Combined neural networks and fuzzy logic for defect prediction in food processing.
[20]	Gradient boosting machines	Pharmaceutical	Used gradient boosting machines to predict quality issues in pharmaceutical manufacturing.
[21]	Intelligent quality control system	Surface roughness	Enhanced surface quality control.
[22]	Artificial intelligence techniques	Production cycle	Rework control optimization.
[23]	Machine learning for quality prediction	Injection molding process	Prediction of defects in injection molding.
[24]	Real-time quality prediction	Serial-parallel manufacturing processes	Real-time quality identification and prediction.
[25]	Integration of multisource information	Manufacturing processes	Enhanced quality prediction using multisource information.
[26]	Soft computing techniques	Machining process	Intelligent quality prediction in machining.
[27]	BP neural network algorithm	Supply chain quality prediction	Optimized method for supply chain quality prediction

3. Problem statement

Fabric dyeing involves several key steps: preparing the fabric, mixing the dyes, dyeing the fabric, washing and rinsing, and drying. First, the fabric is cleaned and treated to ensure it absorbs the dye evenly. Next, the dyes are mixed to achieve the desired color. The fabric is then immersed in the dye mixture, ensuring it is thoroughly soaked. After dyeing, the fabric is washed and rinsed to remove any excess dye, and finally, it is dried.

Fabric dyeing involves several methods, each suited to different types of fabrics and production scales. Common methods include batch dyeing, where fabric is dyed in a single batch, and continuous dyeing, which is efficient for large quantities as fabric moves continuously through the dyeing process. Pad-dyeing uses rollers to ensure even dye absorption, while jet dyeing employs high-pressure jets for delicate fabrics. Beam dyeing immerses fabric wound on a beam into the dye bath, and tie-dyeing creates unique patterns by tying fabric before dyeing. Solution dyeing integrates dye into the polymer solution for synthetic fibers, and dip dyeing achieves gradient effects by submerging

fabric to different levels. Each method offers distinct advantages depending on the desired outcome and fabric type.

Rework processes are often needed in fabric dyeing due to various issues such as uneven dye distribution, incorrect color shades, or dye spots. These problems can arise from improper preparation, inaccurate dye mixing, or inconsistencies in the dyeing process. Reworking involves correcting these defects to meet the required quality standards, ensuring the final product is uniform and meets customer expectations.

The research was conducted within a textile company located in Bursa with a Fabric Dyeing Department. It was found that dyed fabrics frequently required re-dyeing or additional finishing processes due to various quality issues. These corrective processes are referred to as "Rework" within the operation. To address this issue, historical fabric dyeing job order data from the MRP software, which the company uses, was utilized. A dataset comprising 4,855 entries from the past year was collected for this purpose. The goal is to predict and minimize these rework instances to enhance efficiency and quality in the fabric dyeing process.

4. Methodology

As a result of the literature review, the decision of whether a fabric should undergo dyeing was framed as a "Classification" problem. To address this, an algorithm was developed using Logistic Regression and Artificial Neural Networks to determine whether rework is needed or not. Among the Artificial Neural Networks algorithms, the Multilayer Perceptron (MLP) algorithm was selected, specifically the MLPClassifier from the sklearn.neural_network library in Python. The parameters chosen during the construction of the Artificial Neural Networks model directly impact its accuracy. The steps involved in the study are illustrated in Figure 1.

Logistic regression

Regression is a statistical method used to determine the relationship between two variables, where one is dependent (y) and the other is independent (x). In this relationship, y is expressed as a function of x. Given the x attribute values, the continuous variable y is calculated. Regression is a supervised learning technique. Regression analysis helps identify the cause-and-effect relationship between variables.

Logistic regression is a statistical method used to predict binary outcomes. It predicts the probability of a result that can have only two values. The prediction is based on the use of one or several predictors (numerical and categorical). Linear regression is not suitable for values that can be expressed in a binary system such as yes/no or presence/absence because it can predict values outside the range of 0 and 1. Logistic regression produces a logistic curve that is limited to values between 0 and 1.

The logistic regression model aims to minimize the cost function by updating the parameters and learning the parameters that provide the best classification results [28]. In logistic regression analysis, the ratio of the probability of an event occurring to the probability of it not occurring is called the odds ratio, and the probability of not occurring is calculated as follows:

$$1 - P_i = 1 - 1/(1 + e^{-Z_i}) = (1 + e^{-Z_i} - 1)/(1 + e^{-Z_i})$$
$$= e^{-Z_i}/(1 + e^{-Z_i})$$

The odds ratio is obtained by dividing the probability of occurrence by the probability of non-occurrence. The explicit expression of the odds ratio is:

$$P_i/(1 - P_i) = 1/(1 + e^{-Z_i}) * ((1 + e^{-Z_i})/(e^{-Z_i}) = e^{Z_i}$$

By taking the natural logarithm of both sides of the logistic function, which becomes usable in linear regression analysis, a linear structure is obtained:

$$g(x) = \ln(P_i/(1 - P_i)) = \ln e^{Z_i} = Z_i = b_1 + b_2 X_i$$
$$P_i/(1 - P_i) = 1/(1 + e^{-Z_i}) * ((1 + e^{-Z_i})/(e^{-Z_i}) = e^{Z_i}$$

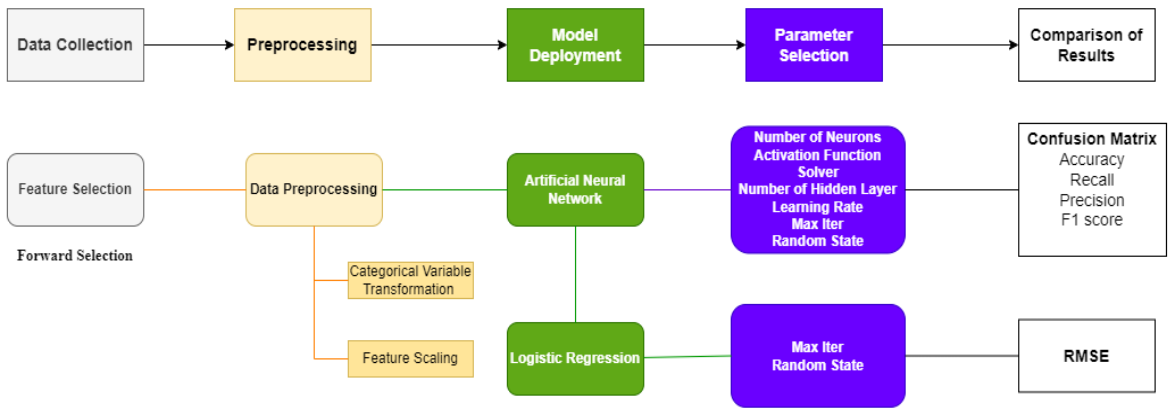


Figure 1. An overview of the methodology.

By taking the natural logarithm of both sides of the logistic function, which becomes usable in linear regression analysis, a linear structure is obtained:

$$g(x) = \ln (P_i /(1 - P_i)) = \ln e^{Z_i} = Z_i = b_1 + b_2 X_i$$

Artificial Neural Networks (ANNs)

ANNs are computer systems designed to automatically perform tasks such as generating and discovering new knowledge through learning. ANNs can be used for tasks like prediction, classification, data association, data interpretation, and data filtering. ANNs are nonlinear information and data processing systems. They consist of processing units called neurons and the connections between these neurons.

The three main components of ANNs are neurons, the connections between these neurons, and functions. ANNs are made up of layers; the input layer is where data enters the network, and hidden layers process the data received in the input layer. There can be multiple hidden layers. The output layer is where the processed data is expressed as output. When a structure has multiple hidden layers, deep neural networks are used [29]. Figure 2 shows the working principle of the ANN model.

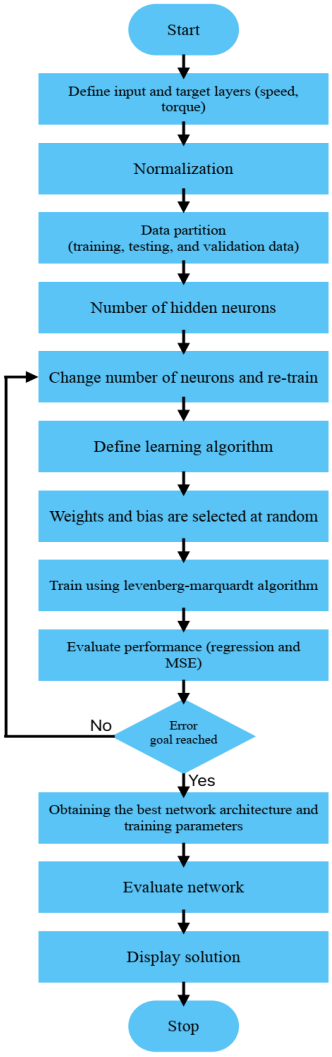


Figure 2. A typical ANN modeling workflow.

An MLP model has three layers: an input layer, hidden layers, and an output layer. Each layer can have one or more neurons, and all neurons in one layer can influence all neurons in the next layer. This relationship can be expressed as:

$$y_k = f (\sum_i w_{ki} \cdot x_i)$$

Where:

y_k is the output value of neuron k

w_{ki} is the weight value between input x_i and output neuron k

x_i is the input value i

In ANN, various transfer functions such as linear, log-sigmoid, and tan-sigmoid can be used to convert input values to output values. In our study, we selected the ReLU (Rectified Linear Unit) and logistic transfer functions to transform the weighted input values into output values.

ReLU activation function

$$f(x) = \max(0,x)$$

This function returns 0 if x is less than 0, otherwise, it returns x.

Logistic activation function

$$f(x) = 1 / 1 + e^{-x}$$

Here, e is the base of the natural logarithm, and x is the input value. This function converts the input value into an output between 0 and 1, allowing the output to be interpreted as a probability.

In this implementation, the behaviors of the parameters listed below have been examined, and based on the performance of the proposed model, suitable values for these parameters have been selected. Some of these parameters were considered together, leading to the execution of 61 different experiments. The findings of these experiments are detailed in the Results section (in Table 3).

The parameters studied

Batch size and iteration count (max_iter), Learning rate (learning_rate), Weight initialization (random_state), Neuron activation function and solver, Neuron count in the hidden layer.

When constructing the artificial neural networks model, there is no fixed formula for determining the number of hidden layers, neuron count, or activation function. Instead, a trial-and-error approach using test data is employed to determine the values that yield the best results for the model.

Test models have been created using the Relu and Logistic activation functions. Additionally, models have been built using the Adam and lbfgs optimizers. It is important to avoid setting a high learning rate, as it may lead to overfitting in the model. Therefore, through necessary testing, a learning rate of 0.001 has been selected for the model.

Generally, lbfgs tends to give optimal results on smaller datasets, while Adam is more suitable for larger datasets [30].

ReLU is a commonly used activation function, especially in deep neural networks. When used with lbfgs, ReLU often performs well because its derivative is non-zero, which can speed up training and reduce the vanishing gradient problem [31].

The logistic activation function is commonly used for binary classification problems. Using lbfgs with the logistic activation function is a good choice for small or medium-sized datasets. However, for deep neural networks or large datasets, activation functions like ReLU might be preferred due to their better performance [32, 33].

One of the important factors affecting the success of artificial neural networks is the selection of the number of neurons in each layer. There is no precise mathematical formula to determine this number, and it is often found through trial and error. Various factors should be considered when determining the number of neurons in a layer. Increasing or decreasing the number of neurons in an artificial neural network can affect the model's performance, error rate, and generalization ability. For instance, using too many neurons can lead to overfitting, where the model fits the training data too closely and fails to generalize well to real-world data [34]. Conversely, using too few neurons may cause the model to underfit, failing to capture the complexity of the dataset, which can decrease performance and increase computational costs [35]. The number of hidden layers and the number of neurons per layer have been determined through trial and error.

The Long Short Term Memory (LSTM) method, a type of deep learning technique, has been employed for real-time defect detection and texture classification on fabrics. This method aims to identify defects on fabrics using digital images.

Through a comprehensive review of literature and expert opinions, a set of 13 candidate features has been identified. The "IsTamir" outcome data serves as the target variable. Among the candidate attributes, a feature selection process has been conducted.

The Forward Selection method was employed for feature selection. This approach was chosen due to its remarkable success in classification tasks. The model commences with the most crucial variable concerning the dependent variable. Initially, the model incorporates solely one variable. Subsequent variables are incrementally added to the model. If the inclusion of a variable enhances the model's performance, it is retained within the model. Employing this process, all variables are examined, ultimately shaping the final model. List of the features are presented in Table 1.

Normalization is a technique commonly applied as part of data preprocessing in machine learning. Its purpose is to transform the values of numerical columns in a dataset onto a common scale without distorting the inherent differences in value ranges. Not every dataset requires normalization for machine learning purposes. The normalization process can be used to reduce data dimensionality, perform operations at appropriately scaled intervals with normalized values, and attain more meaningful and interpretable results.

In literature, various forms of data normalization exist. These include but are not limited to methods like minimum-maximum (min-max) scaling, decimal scaling, z-score normalization, and sigmoid normalization [36].

Table 2. Identified features.

Rework (Response variable)
Fabric quality name (Categorical)
Raw fabric pattern name (Categorical)
Color code (Categorical)
Finished fabric formation type code (Categorical)
Planned length (Numerical)
Raw fabric code (Categorical)
Master recipe name (Categorical)
Process name (Categorical)
Machine name (Categorical)
Operation flow code (Categorical)

Scaling is used to address the magnitude differences between various features in a dataset. When features have different scales, some may exert a stronger influence than others. This imbalance in feature scales can distort their equal contribution and potentially hinder algorithm performance.

The scaling process is implemented to compress feature values within a specific range. In the case of Logistic Regression and Artificial Neural Networks models, scaling has been deemed necessary. The StandardScaler scaling method from the Python sklearn.preprocessing library has been utilized to scale the features.

There are two crucial steps in prediction: first is the preparation of the data for prediction, and second is the comparison of different predictive models. The criteria for comparing models include; accuracy, speed, robustness, scalability, and interpretability.

Fundamental performance indicators employed in assessing the performance of Artificial Neural Networks and machine learning methods include R2, MSE, RMSE, and MAE [37].

To compare the models, Accuracy and F1 Score metrics were used.

Accuracy: The ratio of correctly predicted instances to the total instances.

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN)$$

TP:True Positive

TN:True Negative

FP:False Positive

FN:False Negative

Precision: Indicates how many of the instances predicted as positive are actually positive. It measures the model's ability to correctly classify the positive class.

$$\text{Precision} = TP / (TP + FP)$$

Recall: Indicates how many of the actual positive instances are correctly classified. It is the ratio of correctly predicted positive instances to all actual positive instances.

Recall = $TP / (TP + FN)$

F1 Score: The harmonic mean of Precision and Recall.

F1 Score=$2*((Precision*Recall) / (Precision+ Recall))$

Visualization of Class Distribution: To visualize the distribution of the dataset (Figure 3) in Python, the code block was used:

```
import matplotlib.pyplot as plt
import numpy as np

# Visualize class distribution
plt.bar(np.unique(y), np.bincount(y))
plt.xlabel('Class')
plt.ylabel('Number of Samples')
plt.title('Class Distribution')
plt.show()
```

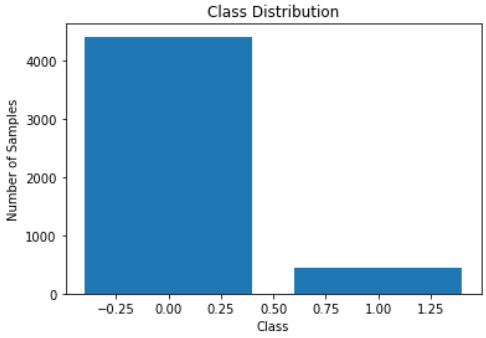


Figure 3. Class distribution.

Given that the dataset is imbalanced, the F1 Score is a more appropriate performance metric to use, as it provides a better measure of the model's performance on datasets with uneven class distributions.

5. Results

In this study, Python programming is used to implement machine learning algorithms. The process begins with handling missing and outlier data. Confirmation from the LEO MRP program ensures that no missing data is anticipated for the relevant attributes.

The get_dummy function is used for categorical data transformation. The get_dummies function takes each category of a categorical variable as a separate column and assigns a value of 1 to the rows corresponding to that category, and 0 to other rows. This way, each category becomes a distinct feature that can be utilized by machine learning models.

Logistic regression implementation

The logistic regression model was implemented using Python, with the following steps:

Data Loading and Preprocessing: The data was loaded from a CSV file and categorical variables were converted into dummy variables. The dependent variable (IsTamir) was encoded using label encoding.

Feature Engineering: The independent variables were selected and the dataset was split into training and testing sets. The data was scaled using StandardScaler.

Model Training and Prediction: A logistic regression model was trained and used to make predictions on the test set.

Evaluation Metrics: Accuracy and F1 Score metrics were used to evaluate the model's performance. Additionally, a confusion matrix and root mean squared error (RMSE) were computed.

The complete Python code is given in Appendix.

The accuracy of the logistic regression model is 92.64%. This means that the model correctly predicts the class of 92.64% of the samples in the test dataset. While accuracy is a commonly used metric, it may not be sufficient for evaluating the performance of models on imbalanced datasets.

The F1 score, which considers both precision and recall, is 0.4779. The F1 score is a better metric for imbalanced datasets because it takes into account false positives and false negatives. A higher F1 score indicates better overall performance of the model.

The confusion matrix provides insight into the model's performance across different classes. It reveals that the model correctly predicted 1431 samples as true positives and 54 samples as true negatives. However, it misclassified 36 samples as false positives and 82 samples as false negatives.

The RMSE value is 0.2713, reflecting the average difference between the actual and predicted values. Lower RMSE values indicate better model performance. However, in the context of classification, RMSE might not be the most informative metric.

While the accuracy of the logistic regression model is relatively high, the F1 score and confusion matrix reveal that the model's performance may be impacted by the imbalanced nature of the dataset. It's important to consider these results in the context of the dataset characteristics and the specific goals of the classification task.

ANNs implementation

The Python code (in Appendix) implements an Artificial Neural Network (ANN) model, starting from data preprocessing steps and extending to training the model and evaluating its performance.

The values obtained from the results based on the parameters selected during the creation of the Artificial Neural Networks models are presented in Table 3 (in appendix).

Table 3 presents the evaluation results of 61 different models based on various configurations, including the activation function, solver, number of hidden layers, and number of neurons.

Upon examining the results, it's evident that there is a wide range of performance across different configurations. Some models achieve high accuracy and F1 Score values, indicating robust classification

performance, while others exhibit lower values, suggesting potential areas for improvement.

These results provide valuable insights into the effectiveness of different configurations in training artificial neural network models for the given task. Further analysis and experimentation may help identify optimal configurations for maximizing classification performance.

Comparison of Logistic Regression and Artificial Neural Network Results

Logistic Regression and Artificial Neural Networks methods were analyzed in Python. The results were compared using the Accuracy metric. Accuracy values for multiple Artificial Neural Network models were calculated and are presented in Table 3. The accuracy and F1 Score obtained from the Logistic Regression model were compared with those from the Artificial Neural Network model with the highest accuracy.

For Logistic Regression, the Accuracy is calculated at 0.90, while for the Artificial Neural Networks (Model 1), the Accuracy is found to be 0.92. This indicates that the Artificial Neural Networks model achieves higher accuracy compared to the Logistic Regression model.

Similarly, the F1 Score performance metric was evaluated. For Logistic Regression, the F1 Score is 0.48, while for the Artificial Neural Networks (Model 1), it is 0.47. Although there is only a slight difference in the F1 Score between the two methods, the Artificial Neural Networks model has a small advantage over Logistic Regression.

In conclusion, the Artificial Neural Networks (Model 1) method achieves higher accuracy compared to Logistic Regression, while the F1 Score values are similar. These results suggest that, for this specific classification problem, the Artificial Neural Networks method may be more effective.

6. Managerial insights and practical implications

The findings of this study can inform strategic planning initiatives aimed at reducing rework costs in fabric dyeing processes. By leveraging predictive analytics, textile companies can proactively identify potential rework needs at the planning stage, enabling preemptive measures to optimize production processes and reduce rework expenses.

Implementing AI-based algorithms for predicting rework needs allows textile companies to optimize operational efficiency by streamlining production processes and minimizing downtime associated with rework activities. This, in turn, enhances overall productivity and cost-effectiveness.

Insights derived from predictive models can facilitate informed resource allocation and risk management strategies. By identifying high-risk production batches or processes prone to rework, companies can allocate resources more efficiently and implement targeted interventions to mitigate risks and minimize rework occurrences.

Table 3. Different configurations and obtained results.

Model No	Activation function	Solver	Number of hidden layers	Number of neurons	Accuracy	F1 Score
1	Relu	Adam	2	3,2	0.92	0.39
2	Relu	Adam	2	3,3	0.92	0.24
3	Relu	Adam	2	3,6	0.91	0.45
4	Relu	Adam	2	6,6	0.91	0.38
5	Relu	Adam	1	6	0.91	0.36
6	Relu	Adam	2	6,5	0.91	0.33
7	Relu	Adam	2	3,4	0.91	0.26
8	Relu	Adam	2	4,7	0.91	0.25
9	Relu	Adam	2	3,5	0.91	0.07
10	Logistic	Adam	2	7,7	0.9	0.47
11	Logistic	Adam	2	6,6	0.9	0.45
12	Logistic	Adam	2	8,8	0.9	0.45
13	Logistic	lbfgs	1	2	0.9	0.45
14	Logistic	Adam	2	5,5	0.9	0.44
15	Logistic	lbfgs	1	7	0.9	0.44
16	Relu	lbfgs	1	4	0.9	0.44
17	Relu	Adam	1	3	0.9	0.43
18	Logistic	Adam	2	7,8	0.9	0.43
19	Relu	lbfgs	2	4,4	0.9	0.42
20	Relu	Adam	4	4	0.9	0.4
21	Relu	Adam	2	4,6	0.9	0.37
22	Logistic	lbfgs	2	5,6	0.9	0.37
23	Relu	Adam	2	4,4	0.9	0.34

Table 3. Different configurations and obtained results. (continued)

Model No	Activation function	Solver	Number of hidden layers	Number of neurons	Accuracy	F1 Score
24	Relu	Adam	2	5,6	0.9	0.34
25	Logistic	lbfgs	2	4,5	0.9	0.33
26	Relu	Adam	2	4,3	0.9	0.3
27	Logistic	lbfgs	1	5	0.89	0.47
28	Logistic	Adam	3	8,8,8	0.89	0.46
29	Logistic	Adam	2	4,4	0.89	0.44
30	Logistic	Adam	1	6	0.89	0.44
31	Logistic	Adam	1	9	0.89	0.44
32	Relu	Adam	2	2,3	0.89	0.43
33	Logistic	Adam	2	8,9	0.89	0.43
34	Logistic	Adam	2	7,5	0.89	0.42
35	Logistic	lbfgs	2	5,5	0.89	0.39
36	Logistic	Adam	2	8,7	0.88	0.44
37	Logistic	lbfgs	1	9	0.87	0.42
38	Logistic	Adam	2	6,7	0.87	0.41
39	Logistic	lbfgs	2	7,7	0.87	0.37
40	Logistic	lbfgs	2	6,7	0.86	0.34
41	Logistic	lbfgs	2	2,2	0.86	0.31
42	Logistic	lbfgs	1	6	0.85	0.43
43	Logistic	lbfgs	1	11	0.85	0.42
44	Logistic	lbfgs	1	10	0.85	0.41
45	Logistic	lbfgs	1	8	0.85	0.4
46	Relu	Adam	2	6,7	0.85	0.39
47	Relu	Adam	2	6,7	0.85	0.39
48	Relu	lbfgs	2	4,3	0.85	0.39
49	Logistic	Adam	1	1	0.85	0.37
50	Relu	Adam	2	3,7	0.85	0.25
51	Logistic	lbfgs	1	12	0.84	0.4
52	Relu	Adam	2	5,5	0.84	0.38
53	Logistic	lbfgs	1	4	0.84	0.36
54	Relu	lbfgs	1	3	0.83	0.4
55	Relu	lbfgs	2	3,3	0.82	0.37
56	Relu	lbfgs	1	5	0.82	0.37
57	Relu	lbfgs	2	3,3	0.82	0.37
58	Relu	lbfgs	1	2	0.8	0.38
59	Logistic	lbfgs	1	3	0.79	0.33
60	Relu	lbfgs	2	6,6	0.77	0.35
61	Relu	Adam	2	7,7	0.73	0.25

7. Conclusions and outlook

The primary objective of this study was to develop an algorithm capable of predicting rework needs at the planning stage for textile companies with fabric dyeing processes, with the goal of minimizing rework costs due to faulty productions. Machine Learning methods, including Logistic Regression and Artificial Neural Networks, were employed to tackle this issue as a classification problem. Both Logistic Regression and Artificial Neural Networks achieved successful outcomes.

Future studies could develop a more effective method for textile companies to predict rework needs. For instance, an algorithm could be designed to create alternative production routes before engaging in re-dyeing or repair processes. Additionally, developing a model that suggests the use of different chemical compositions could significantly reduce rework costs.

In conclusion, this study takes a pivotal step towards solving a significant issue in the textile sector by providing a potential solution for predicting the need for reprocessing in fabric dyeing processes using AI-based algorithms. Future efforts could further enhance this algorithm, ultimately optimizing production processes for textile companies in terms of efficiency and cost-effectiveness.

Acknowledgments

The authors would like to thank the anonymous reviewers for their valuable comments. This paper is an outcome derived from the master's thesis of the first author.

References

[1] Saadallah A, Abdulaaty O, Büscher J, Panusch T, Morik K, Deuse J. (2022). Early quality prediction using deep learning on time series sensor data. *Procedia CIRP*, 107, 611-616.

[2] Caiazzo B, Di Nardo M, Murino T, Petrillo A, Piccirillo G, Santini S. (2022). Towards zero defect manufacturing paradigm: A review of the state-of-the-art methods and open challenges. *Computers in Industry*, 134, 103548.

[3] Zhu C, Chang Q, Arinez J. (2020). Data-enabled modelling and analysis of multistage manufacturing systems with quality reqork loops. *Journal of Manufacturing Systems*, 56, 573-584.

[4] Colledani M, Angius A. (2020). Production quality performance of manufacturing systems with in-line product traceability and rework. *CIRP Annals*, 69, 365-368.

[5] Cavdur F, Kaymaz E, Sebatli A. (2018). A mixed-integer programming model for optimizing rework station position in assembly line balancing. *Uludag University Journal of The Faculty of Engineering*, 23 (3), 273-287.

[6] Schuh G, Gützlaff A, Thomas K, Welsing M. (2021). Machine learning based defect detection in a low automated assembly environment. *Procedia CIRP*, 104, 265-270.

[7] Corekcioglu M, Ercan E, Aras Elibüyük S. (2021). Usage applications of artificial neural network methods in textile industry. *Journal of Technical Sciences*, 11(2), 14-20.

[8] Ozdemir H. (2013). Artificial neural networks and their usage in weaving technology. *Electronic Journal of Vehicle Technologies*, 7 (1), 51-68.

[9] Turker E. (2017). A research on estimation of the weave fabric properties with the artificial neural networks. *Textile and Apparel*, 27(1), 10-21.

[10] Arikan Kargi S. (2014). A comparison of artificial neural networks and multiple linear regression models as in predictors of fabric weft defects. *Textile and Apparel*, 24 (3), 309-316.

[11] Balci O, Ogulata RT. (2009). Prediction of CIELab values and color changing occurred after chemical finishing applications by artificial neural networks on dyed fabrics. *Textile and Apparel*, 19(1), 61-69.

[12] Lotfi, R., Kheiri, K., Sadeghi, A., & Babaee Tirkolaee, E. (2022). An extended robust mathematical model to project the course of COVID-19 epidemic in Iran. *Annals of Operations Research*, 1-25.

[13] Lotfi, R., Gholamrezaei, A., Kadlubek, M., Afshar, M., Ali, S. S., & Kheiri, K. (2022). A robust and resilience machine learning for forecasting agri-food production. *Scientific Reports*, 12 (1), 21787.

[14] Sadeghi, A., Kheiri, K., Lotfi, R., & Babaee Tirkolaee, E. (2023). Machine Learning and Deep Learning Techniques for Early Disease Detection. *Journal of Medical Systems*, 47(1), 25-36.

[15] Kumar, P., Mehta, M., & Singh, R. (2019). Comparative Analysis of Support Vector Machines and Neural Networks for Manufacturing Defect Prediction. *International Journal of Advanced Manufacturing Technology*, 104, 1589-1599.

[16] Wang, L., Li, J., & Zhou, Y. (2020). Application of Deep Learning in Predicting Defects in Electronics Manufacturing. *IEEE Transactions on Industrial Informatics*, 16(5), 3178-3185.

[17] Patel, S., & Gupta, R. (2018). Predictive Models for Reducing Rework in Automotive Manufacturing Using Logistic Regression and Decision Trees. *Journal of Manufacturing Science and Engineering*, 140(8), 081015.

[18] Lu, F., Zhou, G., Liu, Y., & Zhang, C. (2022). Ensemble transfer learning for cutting energy consumption prediction of aviation parts towards green manufacturing. *Journal of Cleaner Production*, *331*, 129920.

[19] Deisingh, A. K., Stone, D. C., & Thompson, M. (2004). Applications of electronic noses and tongues in food analysis. *International journal of food science & technology*, *39*(6), 587-604.

[20] Serna-Carrizales, J. C., Zárate-Guzmán, A. I., Flores-Ramírez, R., de León-Martínez, L. D., Aguilar-Aguilar, A., Warren-Vega, W. M., ... & Ocampo-Pérez, R. (2024). Application of artificial intelligence for the optimization of advanced oxidation processes to improve the water quality polluted with pharmaceutical compounds. *Chemosphere*, *351*, 141216.

[21] Gola, A., Świć, A., & Kozłowski, E. (2020). Intelligent quality control system for surface roughness. *Applied Sciences*, 10(12), 4255.

[22] Lam, S. Y., & Ip, W. H. (2011). An application of artificial intelligence techniques for the control of rework in the production cycle. *Expert Systems with Applications*, 38(7), 8314-8322.

[23] Lee, S. W., & Park, J. S. (2001). Intelligent quality prediction system using machine learning for injection moulding process. *International Journal of Advanced Manufacturing Technology*, 18(5), 329-334.

[24] Chen, D., & Hu, S. J. (2001). Real-time identification and prediction of quality for serial-parallel manufacturing processes. *International Journal of Production Research*, 39(18), 4147-4166.

[25] Liu, S., Zhang, L., Ma, J., Sun, X., Li, H., & Li, Y. (2016). Integration of multisource information for quality prediction in manufacturing. *IEEE Transactions on Industrial Informatics*, 13(4), 1925-1935.

[26] Sujatha, C., & Kumanan, S. (2016). An intelligent approach for quality prediction in machining process using soft computing techniques. *Procedia*

Technology, 25, 546-553.

[27] He, P., Gao, L., Zhao, Y., Liu, Y., & Li, W. (2019). A new method of supply chain quality prediction based on optimized BP neural network algorithm. *Computers & Industrial Engineering*, 135, 1067-1080.

[28] Kleinbaum, D. G., & Klein, M. (2020). *Logistic regression: A self-learning text (Statistics for biology and health)*. Springer.

[29] Kumar, A., Kumar, M., & Goswami, P. (2024). Numerical Solution of Coupled System of Emden-Fowler Equations using artificial neural network technique, *An International Journal of Optimization and Control: Theories & Applications*, 14(1), 62-73.

[30] Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

[31] Nair, V., & Hinton, G. E. (2010). Rectified linear units improve restricted boltzmann machines. In *Proceedings of the 27th International Conference on Machine Learning (ICML-10)* (pp. 807-814).

[32] Glorot, X., Bordes, A., & Bengio, Y. (2011). Deep sparse rectifier neural networks. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics* (pp. 315-323). JMLR Workshop and Conference Proceedings.

[33] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.

[34] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*, 15(1), 1929-1958.

[35] Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer.

[36] Jayalakshmi T, Santhakumaran A. (2011). Statistical normalization and back propagation for classification. *International Journal of Computer Theory and Engineering*, 3 (1), 1793-8201.

[37] Karasu S, Altan A, Sarac Z, Hacioglu R. (2018). Prediction of bitcoin prices with machine learning methods using time series data. *The 26th Signal Processing and Communications Applications Conference (SIU)*, Izmir, Turkey.

Sema Aydın *holds a BSc degree in Industrial Engineering from Hacettepe University, and a MSc degree in Industrial Engineering from TU Bursa. She has a keen interest in AI applications in industry.*

http://orcid.org/0009-0001-3948-0529

Koray Altun *is an Assistant Professor of Industrial Engineering at TU Bursa and a Research Collaborator at GLORAD. He is also the founder and manager of a startup focused on digital innovation and software consultancy. He holds a Ph.D. and a BSc degree in Industrial Engineering from Gaziantep University and Erciyes University, respectively. He has published papers in well-respected journals and has led consulting engagements in reputable companies with a focus on R&D, technology, and innovation management. His recent research interests include the Innovation Excellence Model, Systematic Innovation, Technology and Innovation Management, and R&D Management.*

http://orcid.org/0000-0003-0357-9495

Appendices

1. Logistic Regression in Python Code

```
import pandas as pd
import numpy as np

# Data loading
veriler = pd.read_csv('RW_Data_1.csv')
print(veriler)

# Converting categorical variables to dummy variables
df_KumasKaliteAdi = pd.get_dummies(veriler["KumasKaliteAdi"], prefix="Kalite",
drop_first=True)
df_KumasDesen = pd.get_dummies(veriler["HamKumasDesenAdi"], prefix="KumasDesen",
drop_first=True)
df_UretimPartiNo = pd.get_dummies(veriler["UretimPartiNo"], prefix="PartiNo",
drop_first=True)
df_RenkKodu = pd.get_dummies(veriler["RenkKodu"], prefix="RenkKodu",
drop_first=True)
df_LotNo = pd.get_dummies(veriler["LotNo"], prefix="LotNo", drop_first=True)
df_MamulOlusumTipKodu = pd.get_dummies(veriler["MamulOlusumTipKodu"],
prefix="MamulOlusumTipKodu", drop_first=True)
df_HamUrunKodu = pd.get_dummies(veriler["HamUrunKodu"], prefix="HamUrunKodu",
drop_first=True)
```

```
df_MasterReceteAdi = pd.get_dummies(veriler["MasterReceteAdi"],
prefix="MasterReceteAdi", drop_first=True)
df_ProsesAdi = pd.get_dummies(veriler["ProsesAdi"], prefix="ProsesAdi",
drop_first=True)
df_IslemAkisKodu = pd.get_dummies(veriler["IslemAkisKodu"], prefix="IslemAkisKodu",
drop_first=True)
df_MakinaAdi = pd.get_dummies(veriler["MakinaAdi"], prefix="MakinaAdi",
drop_first=True)

# Creating a DataFrame for PlanMt
PlanMt = veriler.iloc[:, 8:9].values
mt = pd.DataFrame(data=PlanMt, index=range(4855), columns=['PlanMt'])

# Encoding the dependent variable
from sklearn.preprocessing import LabelEncoder
le = LabelEncoder()
IsTamir = le.fit_transform(veriler.iloc[:, 0])
tamir = pd.DataFrame(data=IsTamir, index=range(4855), columns=['IsTamir'])

# Merging DataFrames
sonuc = pd.concat([df_KumasKaliteAdi, df_KumasDesen, df_UretimPartiNo, df_RenkKodu,
df_LotNo, df_MamulOlusumTipKodu, df_HamUrunKodu, df_MasterReceteAdi, df_ProsesAdi,
df_IslemAkisKodu, df_MakinaAdi, mt, tamir], axis=1)
print(sonuc)

# Splitting into independent and dependent variables
x = sonuc.iloc[:, 0:-1].values  # independent variables
y = sonuc.iloc[:, -1]           # dependent variable

# Splitting the data into training and testing sets
from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.33,
random_state=0)

# Scaling the data
from sklearn.preprocessing import StandardScaler
sc = StandardScaler()
X_train = sc.fit_transform(x_train)
X_test = sc.transform(x_test)

# Logistic regression
from sklearn.linear_model import LogisticRegression
logr = LogisticRegression(random_state=0, max_iter=1000)
logr.fit(X_train, y_train)  # training the model

# Predictions
y_pred = logr.predict(X_test)
print(y_pred)
print(y_test)

# Evaluation
from sklearn.metrics import confusion_matrix, mean_squared_error, accuracy_score,
f1_score

# Confusion matrix
cm = confusion_matrix(y_test, y_pred)
print(cm)

# Root mean squared error
RMSE = np.sqrt(mean_squared_error(y_test, y_pred))
print("RMSE:", RMSE)

# Accuracy
accuracy = accuracy_score(y_test, y_pred)
print("Accuracy:", accuracy)

# F1 Score
f1 = f1_score(y_test, y_pred)
```

```
print("F1 Score:", f1)

# Visualization of class distribution
import matplotlib.pyplot as plt
plt.bar(np.unique(y), np.bincount(y))
plt.xlabel('Class')
plt.ylabel('Number of Samples')
plt.title('Class Distribution')
plt.show()
```

2. ANN in Python Code

```
# -*- coding: utf-8 -*-
"""
import pandas as pd
import numpy as np
#The CSV file should be located in the same directory as the Python code.
veriler=pd.read_csv('RW_Data_1.csv')
print(veriler)

df_KumasKaliteAdi=pd.get_dummies(veriler["KumasKaliteAdi"],prefix="Kalite",drop_fir
st=True)
#print(df_KumasKaliteAdi)

df_KumasDesen=pd.get_dummies(veriler["HamKumasDesenAdi"],prefix="KumasDesen",drop_f
irst=True)
#print(df_KumasDesen)

df_UretimPartiNo=pd.get_dummies(veriler["UretimPartiNo"],prefix="PartiNo",drop_firs
t=True)
#print(df_UretimPartiNo)

df_RenkKodu=pd.get_dummies(veriler["RenkKodu"],prefix="RenkKodu",drop_first=True)
#print(df_RenkKodu)

df_LotNo=pd.get_dummies(veriler["LotNo"],prefix="LotNo",drop_first=True)
#print(df_LotNo)

df_MamulOlusumTipKodu=pd.get_dummies(veriler["MamulOlusumTipKodu"],prefix="MamulOlu
sumTipKodu",drop_first=True)
#print(df_MamulOlusumTipKodu)

df_HamUrunKodu=pd.get_dummies(veriler["HamUrunKodu"],prefix="HamUrunKodu",drop_firs
t=True)
#print(df_HamUrunKodu)

df_MasterReceteAdi=pd.get_dummies(veriler["MasterReceteAdi"],prefix="MasterReceteAd
i",drop_first=True)
#print(df_MasterReceteAdi)

df_ProsesAdi=pd.get_dummies(veriler["ProsesAdi"],prefix="ProsesAdi",drop_first=True
)
#print(df_ProsesAdi)

df_IslemAkisKodu=pd.get_dummies(veriler["IslemAkisKodu"],prefix="IslemAkisKodu",dro
p_first=True)
#print(df_IslemAkisKodu)

df_MakinaAdi=pd.get_dummies(veriler["MakinaAdi"],prefix="MakinaAdi",drop_first=True
)
#print(df_MakinaAdi)

PlanMt=veriler.iloc[:,8:9].values
mt=pd.DataFrame(data=PlanMt,index=range(4855),columns=['PlanMt'])

IsTamir=veriler.iloc[:,0:1].values
tamir=pd.DataFrame(data=IsTamir,index=range(4855),columns=['IsTamir'])
```

```
from sklearn import preprocessing

# DataFrame merging is used to create a new DataFrame by combining multiple
DataFrames.
sonuc=pd.concat([df_KumasKaliteAdi,df_KumasDesen,df_UretimPartiNo,df_RenkKodu,df_Lo
tNo, df_MamulOlusumTipKodu,df_HamUrunKodu,df_MasterReceteAdi,df_ProsesAdi,
df_IslemAkisKodu, df_MakinaAdi,mt,tamir],axis=1)
#print(sonuc)
x=sonuc.iloc[:,0:-1].values #independent variables
y=sonuc.iloc[:,-1] #dependent variables
#print(x)
#print(y)

# splitting the data into training and testing sets
from sklearn.model_selection import train_test_split

x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.33,random_state=0)

# scaling the data
from sklearn.preprocessing import StandardScaler
sc=StandardScaler()
X_train=sc.fit_transform(x_train) #eğitim uyguluyor
X_test=sc.transform(x_test)

#Let's create our artificial neural network model and configure our hidden layer.
from sklearn.neural_network import MLPClassifier
#iki katman ve her katman 6 nörondan oluşacak şekilde model kurulmuştur
mlpcl = MLPClassifier(hidden_layer_sizes=(3,3),activation='relu', solver='lbfgs',
max_iter=1000,random_state=42,learning_rate='constant', learning_rate_init=0.001)

#model = MLPClassifier(hidden_layer_sizes=(64, 64), activation='relu',
solver='adam', max_iter=1000, random_state=42)
mlpcl.fit(X_train, y_train.values.ravel())

# Let's make predictions on our test data.
predictions = mlpcl.predict(X_test)
print(predictions)

# Let's evaluate the performance of our predictions - our algorithm.
from sklearn.metrics import classification_report, confusion_matrix
print(confusion_matrix(y_test,predictions))
print(classification_report(y_test,predictions))

from sklearn.metrics import accuracy_score
# Calculate accuracy using your predictions and the actual labels.
accuracy = accuracy_score(y_test, predictions)

# print accuracy
print("Accuracy:", accuracy)

from sklearn.metrics import f1_score
# calculate accuracy
f1 = f1_score(y_test, predictions, pos_label='T')
print("F1 Score:", f1)
```

This work is licensed under a Creative Commons Attribution 4.0 International License. The authors retain ownership of the copyright for their article, but they allow anyone to download, reuse, reprint, modify, distribute, and/or copy articles in IJOCTA, so long as the original authors and source are credited. To see the complete license contents, please visit http://creativecommons.org/licenses/by/4.0/.

An International Journal of Optimization and Control: Theories & Applications
ISSN:2146-0957 eISSN:2146-5703
Vol.14, No.4, pp.322-335 (2024)
http://doi.org/10.11121/ijocta.1639

RESEARCH ARTICLE

Witte's conditions for uniqueness of solutions to a class of Fractal-Fractional ordinary differential equations

Abdon Atangana [a,b], Ilknur Koca [c*]

[a] *Institute for Groundwater Studies, Faculty of Natural and Agricultural Sciences, University of the Free State, Bloemfontein, 9301, South Africa*
[b] *Department of Medical Research, China Medical University Hospital, China Medical University, Taichung, Taiwan*
[c] *Department of Economics and Finance, Fethiye Business Faculty, Mugla Sıtkı Kocman University, 48300 , Mugla, Türkiye*
AtanganaA@ufs.ac.za, ilknurkoca@mu.edu.tr

ARTICLE INFO

Article History:
Received 1 July 2024
Accepted 1 September 2024
Available Online 9 October 2024

Keywords:
Fractal calculus
Witte's conditions
Uniqueness
Nonlocal operators

AMS Classification 2010:
26A33; 34A12; 31EXX

ABSTRACT

In this paper, Witte's conditions for the uniqueness solution of nonlinear differential equations with integer and non-integer order derivatives are investigated. We present a detailed analysis of the uniqueness solutions of four classes of nonlinear differential equations with nonlocal operators. These classes include classical and fractional ordinary differential equations in fractal calculus. For each case, theorems and lemmas and their proofs are presented in detail.

CC BY

1. Introduction

Nonlinear differential equations are powerful mathematical tools used to model real-world problems arising in several fields of study [1, 2]. The analysis of their solutions is of great importance, as they are for comparison with the collected data [3]. It is worth noting that, most of the time, obtaining their exact solutions is sometimes impossible. Researchers have therefore developed different approaches to help guarantee the existence and uniqueness of these solutions [4–7]. We note that several researchers have provided different conditions in the case of uniqueness in the last decades. For existence, many iterative approaches have been suggested, for example, Picard, Toneli, and others. For uniqueness, Witte provided several conditions that can be tested to conclude that a given nonlinear ordinary differential equation with a classical derivative has a unique solution. Several other researchers, like Caratheordory, Nagumo, and others, have also provided some important conditions [8, 9]. While several works have been published for ordinary differential equations with integer-order derivatives, much attention has not been devoted to classical and fractional nonlinear ordinary differential equations in fractal calculus [10, 11]. Fractional calculus and fractal calculus are interconnected fields, primarily through their shared focus on non-integer dimensions and scales. Fractional calculus extends the concept of differentiation and integration to non-integer orders, allowing for more flexible mathematical modeling of complex systems. A key connection is that fractional calculus provides the mathematical tools needed to describe the dynamics of processes on fractal structures. For example, the study by Metzler and Klafter [12] titled "The random walk's guide

*Corresponding Author

to anomalous diffusion: a fractional dynamics approach" discusses how fractional calculus can be applied to model diffusion processes on fractal media . Whereas these equations are suitable for the depiction of several complex real-world problems that cannot be modeled using classical ordinary differential equations. In this paper, we shall consider four classes of nonlinear ordinary differential equations, including those with classical differentiation in fractal calculus, those with power law, exponential decay, and generalized Mittag-Leffler kernels in fractal calculus. For each case, we will find conditions of uniqueness based on the framework of Witte [9].

2. Preliminaries

We shall provide some definitions that will be used in this paper.

$$\frac{df(t)}{dt^{\beta}} = \lim_{t_1 \to t} \frac{f(t_1) - f(t)}{t_1^{\beta} - t^{\beta}},\ \beta > 0, \tag{1}$$

which the fractal derivative of the function f with respect to a fractal measure t with scaling indice β [11]. We note that if f is differentiable then,

$$\frac{df(t)}{dt^{\beta}} = \frac{f'(t)}{\beta t^{\beta-1}}. \tag{2}$$

Fractal-fractional derivatives of the function f with power law, exponential decay and Mittag-Leffler kernel are given below respectively [10].

$${}^{FFP}_{t_0}D_t^{\alpha,\beta} f(t) = \frac{d}{dt^{\beta}} \frac{1}{\Gamma(1-\alpha)} \int_{t_0}^{t} f(\tau)(t-\tau)^{-\alpha} d\tau, \tag{3}$$

$${}^{FFE}_{t_0}D_t^{\alpha,\beta} f(t) = \frac{d}{dt^{\beta}} \frac{1}{(1-\alpha)} \int_{t_0}^{t} f(\tau) \exp\left(\frac{-\alpha}{1-\alpha}(t-\tau)\right) d\tau, \tag{4}$$

$${}^{FFM}_{t_0}D_t^{\alpha,\beta} f(t) = \frac{d}{dt^{\beta}} \frac{1}{(1-\alpha)} \int_{t_0}^{t} f(\tau) E_{\alpha}\left(\frac{-\alpha}{1-\alpha}(t-\tau)^{\alpha}\right) d\tau, \tag{5}$$

where $(\alpha, \beta) \in (0, 1]$.

Their respective integrals are given as below:

$${}^{FFP}_{t_0}J_t^{\alpha,\beta} f(t) = \frac{\beta}{\Gamma(\alpha)} \int_{t_0}^{t} (t-\tau)^{\alpha-1} \tau^{\beta-1} f(\tau) d\tau, \tag{6}$$

$${}^{FFE}_{t_0}J_t^{\alpha,\beta} f(t) = (1-\alpha)\beta t^{\beta-1} f(t) + \alpha\beta \int_{t_0}^{t} \tau^{\beta-1} f(\tau) d\tau, \tag{7}$$

$${}^{FFM}_{t_0}D_t^{\alpha,\beta} f(t) = (1-\alpha)\beta t^{\beta-1} f(t) + \frac{\alpha\beta}{\Gamma(\alpha)} \int_{t_0}^{t} (t-\tau)^{\alpha-1} \tau^{\beta-1} f(\tau) d\tau. \tag{8}$$

We note that, when $\beta = 1$, we recover all the fractional differential and integral operators.

3. The Witte's uniqueness conditions for classical fractal ordinary differential equations

In this section, we are interested in the following general fractal differential equation.

$$\begin{cases} {}^{F}_{t_0}D_t^{\alpha} y(t) = f(t, y(t)), \quad t > t_0. \\ y(t_0) = y_0, \end{cases} \tag{9}$$

The aim is to establish uniqueness conditions based on the Witte's uniqueness.

Theorem 1. *Let assume that $f(t, y)$ is continuous in $S_+ = \{(t, y) \mid t_0 < t \leq a,\ |y| < \infty\}$ and satisfies*

i) $\forall (t, y), (t, \overline{y}) \in S_+$

$$|f(t, y) - f(t, \overline{y})| \leq h(t) |y - \overline{y}|, \tag{10}$$

ii) $|f(t, y)| \leq \varphi(t) h(t) \exp\left(\int_a^t h(\tau) d\tau\right)$ *in* S_+,

where $h(t) > 0$ is continuous in $[t_0, a]$ and $\varphi(t)$ is continuous in $[t_0, a]$ and $\varphi(t_0) = 0$.

Then the considered equation has almost one solution.

Proof. To proof the above, we shall first provide the proof of the following Lemma. □

Lemma 1. *Let $\Omega(t)$ be a nonnegative continuous function on $[t_0, a]$ and let*

i) $h(t) > 0$ be continuous functions in $[t_0, a]$,

ii) There exists a function $H(t)$ in $[t_0, a]$ such that $H'(t) = h(t)$ for almost all $t \in [t_0, a]$ and $\lim_{t \to t_0^+} H(t)$ exists, it can be finite,

iii) $\Omega(t) \leq \int_{t_0}^{t} h(\tau)\Omega(\tau) d\tau$, $t \in [t_0, a]$,

iv) $\Omega(t) = o\left(\exp\left(t^{\alpha} H(t)\right)\right)$ *as* $t \to t_0^+$. *Then* $\Omega(t) = 0$.

Proof. Let the mentioned conditions hold, then

$$\Psi(t) = \alpha \int_{t_0}^{t} \tau^{\alpha-1} h(\tau)\Omega(\tau)d\tau. \tag{11}$$

Thank to the hypothesis of the Lemma $\Psi(t)$ exists and is continuous on $[t_0, a]$. Then

$$\begin{aligned} {}^{F}_{t_0}D^{\alpha}_{t}\Psi(t) &= \frac{1}{\alpha t^{\alpha-1}}\frac{d}{dt}\left[\alpha\int_{t_0}^{t}\tau^{\alpha-1}h(\tau)\Omega(\tau)d\tau\right], \\ &= \frac{1}{\alpha t^{\alpha-1}}\left[\alpha t^{\alpha-1}h(t)\Omega(t)\right], \\ &= h(t)\Omega(t) \leq h(t)\Psi(t). \end{aligned} \tag{12}$$

We define

$$F(t) = \exp\left(-t^{\alpha}H(t)\right)\Psi(t). \tag{13}$$

$$\begin{aligned} &{}^{F}_{t_0}D^{\alpha}_{t}F(t) \\ &= \frac{1}{\alpha t^{\alpha-1}}\frac{d}{dt}\int_{t_0}^{t}F(\tau)d\tau = \frac{1}{\alpha t^{\alpha-1}}F'(t), \\ &= \frac{1}{\alpha t^{\alpha-1}} \\ &\times \begin{bmatrix} \Psi'(t)\exp\left(-t^{\alpha}H(t)\right) \\ -\Psi(t)\begin{pmatrix} -\alpha t^{\alpha-1}H(t) \\ -t^{\alpha}h(t) \end{pmatrix}\exp(-t^{\alpha}H(t)) \end{bmatrix}, \\ &= \frac{1}{\alpha t^{\alpha-1}}\exp(-t^{\alpha}H(t)) \\ &\times \begin{bmatrix} \Psi'(t) \\ -\Psi(t)\begin{bmatrix} \alpha t^{\alpha-1}H(t) \\ -t^{\alpha}h(t) \end{bmatrix} \end{bmatrix}, \\ &\leq \frac{1}{\alpha t^{\alpha-1}}\exp(-t^{\alpha}H(t)) \\ &\begin{bmatrix} h(t)\Psi(t)\alpha t^{\alpha-1} \\ -\Psi(t)\begin{bmatrix} \alpha t^{\alpha-1}H(t) \\ -t^{\alpha}h(t) \end{bmatrix} \end{bmatrix}, \\ &\leq \frac{\Psi(t)\exp(-t^{\alpha}H(t))}{\alpha t^{\alpha-1}}\begin{bmatrix} \alpha t^{\alpha-1}h(t) \\ -\alpha t^{\alpha-1}H(t) - t^{\alpha}h(t) \end{bmatrix}, \\ &\leq \frac{\Psi(t)\exp(-t^{\alpha}H(t))}{\alpha t^{\alpha-1}}\begin{bmatrix} t^{\alpha}h(t) \\ -\alpha t^{\alpha-1}H(t) - t^{\alpha}h(t) \end{bmatrix}, \\ &\leq -H(t)\Psi(t)\exp(-t^{\alpha}H(t)), \\ &\leq -H(t)F(t) \leq 0. \end{aligned} \tag{14}$$

We can say that $\forall t \in [t_0, a]$, $\Psi(t)\exp(-t^{\alpha}H(t))$ is decreasing. We now choose $\varepsilon > 0$ with t small enough

$$\begin{aligned} &\Psi(t)\exp(-t^{\alpha}H(t)) \\ &= \exp(-t^{\alpha}H(t))\int_{t_0}^{t}\alpha\tau^{\alpha-1}h(\tau)\Omega(\tau)d\tau, \\ &\leq \varepsilon\exp(-t^{\alpha}H(t))\alpha\int_{t_0}^{t}\tau^{\alpha-1}h(\tau)\exp(\tau^{\alpha}H(\tau))d\tau, \\ &\leq \varepsilon\exp(-t^{\alpha}H(t))\alpha\int_{t_0}^{t}\tau^{\alpha}h(\tau)\exp(\tau^{\alpha}H(\tau))d\tau, \\ &\leq \varepsilon\exp(-t^{\alpha}H(t))\alpha\int_{t_0}^{t}\begin{pmatrix} \tau^{\alpha}h(\tau) \\ +\alpha\tau^{\alpha-1}H(\tau) \end{pmatrix} \\ &\times \exp(\tau^{\alpha}H(\tau))d\tau, \\ &= \varepsilon\alpha\exp(-t^{\alpha}H(t))\exp(t^{\alpha}H(t)), \\ &= \varepsilon\alpha. \end{aligned} \tag{15}$$

$$\lim_{t\to t_0^{+}}\exp(-t^{\alpha}H(t))\Psi(t) = 0, \tag{16}$$

thus

$$\exp(-t^{\alpha}H(t))\Psi(t) \leq 0 \text{ for } t > 0, \tag{17}$$

this also implies that

$$\alpha\int_{t_0}^{t}\tau^{\alpha-1}h(\tau)\Omega(\tau)d\tau \leq 0. \tag{18}$$

Therefore we should have

$$\Omega(t) = 0. \tag{19}$$

□

The new uniqueness criteria will be presented below. This is more general that the previous condition of the theorem.

Theorem 2. *Let $f(t, y)$ be continuous in $\overline{S}_{+}$ in addition to the hypothesis in theorem 1, we have*

$$|f(t, y) - f(t, \overline{y})| = o\left(\exp(t^{\alpha}H(t))\right), \tag{20}$$

as $t \to t_0^{+}$ uniformly with respect to $y, \overline{y} \in [-\lambda, \lambda]$, $\lambda > 0$ arbitrary with $h(t)$ and $H(t)$ the same like in Lemma 1. Then the considered equation has almost one solution in $[t_0, a]$.

Proof. Let $y(t)$ and $\overline{y}(t)$ be two different solutions of one equation

$$y(t) = y(t_0) + \alpha\int_{t_0}^{t}\tau^{\alpha-1}f(\tau, y(\tau))d\tau, \tag{21}$$

$$\begin{aligned} |y(t) - \overline{y}(t)| &\leq \alpha\int_{t_0}^{t}\tau^{\alpha-1}\left|f(\tau, y(\tau)) - f(\tau, \overline{y}(\tau))\right|d\tau, \\ &\leq \alpha\int_{t_0}^{t}\tau^{\alpha-1}h(\tau)\left|y - \overline{y}\right|d\tau, \end{aligned}$$

$$\leq \alpha \int_{t_0}^{t} \left(\tau^{\alpha-1} H(\tau) + \tau^{\alpha} h(\tau)\right) \exp\left(\tau^{\alpha} H(\tau)\right) d\tau,$$
$$\leq \varepsilon \exp\left(t^{\alpha} H(t)\right).$$

From the Lemma, the result is obtained. □

Corollary 1. *Let f satisfies the following conditions; $\forall (t,\overline{y}),(t,y) \in S_+$, $\beta \in (1,2]$ and $\alpha \in (0,1]$:*
i) $\left(f(t,\overline{y}) - f(t,y)\right)(\overline{y}-y)^{\overline{\beta}-1} \leq \frac{\overline{\beta}}{\alpha} t h(t) (\overline{y}-y)^{\overline{\beta}}$,
ii) $f(t,\overline{y}) - f(t,y) = o\left(\exp(t^{\alpha} H(t))\right)$,
as $t \to t_0^+$ uniformly with respect to $y, \overline{y} \in [-\delta,\delta], \delta > 0$ arbitrary. Then the considered equation has almost one solution.

Proof. Let $\overline{y}$ and y be different solutions in $\overline{S}_+$. Let put $\Phi(t) = (\overline{y}(t) - y(t))^{\overline{\beta}}$ then we have that,

$$\begin{aligned} {}^{F}_{t_0}D^{\alpha}_{t}\Phi(t) &= \frac{1}{\alpha t^{\alpha-1}} \frac{d}{dt}\left[\Phi(t)\right], \\ &= \frac{1}{\alpha t^{\alpha-1}} \left(\overline{\beta}\left(\overline{y}(t) - y(t)\right)\right)' (\overline{y}(t) - y(t))^{\overline{\beta}-1}, \\ &= \overline{\beta}\left({}^{F}_{t_0}D^{\alpha}_{t}\overline{y}(t) - {}^{F}_{t_0}D^{\alpha}_{t} y(t)\right)(\overline{y}(t) - y(t))^{\overline{\beta}-1}, \\ &= \overline{\beta}\left(f\left(t,\overline{y}(t)\right) - f(t,y(t))\right)(\overline{y}(t) - y(t))^{\overline{\beta}-1}. \end{aligned} \tag{22}$$

By the hypothesis (i), we have that

$$\begin{aligned} {}^{F}_{t_0}D^{\alpha}_{t}\Phi(t) &\leq \overline{\beta} h(t) (\overline{y}(t) - y(t))^{\overline{\beta}}, \\ &= \frac{\overline{\beta}}{\alpha} t h(t) \Phi(t). \end{aligned} \tag{23}$$

Therefore

$${}^{F}_{t_0}D^{\alpha}_{t}\Phi(t) \leq \frac{\overline{\beta}}{\alpha} h(t)\Phi(t) t. \tag{24}$$

Note that

$$\begin{aligned} &{}^{F}_{t_0}D^{\alpha}_{t}\left(\Phi(t)\exp\left(-\overline{\beta}t^{\alpha}H(t)\right)\right) \\ &= \frac{1}{\alpha t^{\alpha-1}} \begin{bmatrix} \Phi'(t)\exp\left(-\overline{\beta}t^{\alpha}H(t)\right) \\ +\Phi(t)\begin{bmatrix} -\overline{\beta}t^{\alpha}h(t) \\ -\overline{\beta}\alpha t^{\alpha-1}H(t) \end{bmatrix}\exp\left(-\overline{\beta}t^{\alpha}H(t)\right) \end{bmatrix}, \\ &= \exp\left(-\overline{\beta}t^{\alpha}H(t)\right)\begin{bmatrix} {}^{F}_{t_0}D^{\alpha}_{t}\Phi(t) \\ -\Phi(t)\begin{pmatrix} \frac{\overline{\beta}}{\alpha} t h(t) \\ +\overline{\beta}H(t) \end{pmatrix} \end{bmatrix}, \\ &\leq \exp\left(-\overline{\beta}t^{\alpha}h(t)\right)\begin{bmatrix} {}^{F}_{t_0}D^{\alpha}_{t}\Phi(t) \\ -\overline{\beta}\frac{t}{\alpha}h(t)\Phi(t) \end{bmatrix}, \\ &\leq 0. \end{aligned} \tag{25}$$

Since

$${}^{F}_{t_0}D^{\alpha}_{t}\Phi(t) - \overline{\beta}\frac{t}{\alpha}h(t)\Phi(t) \leq 0, \tag{26}$$

$${}^{F}_{t_0}D^{\alpha}_{t}\left(\exp\left(-\overline{\beta}t^{\alpha}H(t)\right)\Phi(t)\right) \leq 0. \tag{27}$$

The conclusion is that the function $\exp\left(-\overline{\beta}t^{\alpha}\right)\Phi(t)$ is non increasing for almost $\forall t \in [t_0, a]$. On the other hand we have that

$$\begin{aligned} &\exp\left(-\overline{\beta}t^{\alpha}H(t)\right)\Phi(t) \\ &= \exp\left(-\overline{\beta}t^{\alpha}H(t)\right)(\overline{y}(t) - y(t))^{\overline{\beta}}, \\ &= \exp\left(-\overline{\beta}t^{\alpha}H(t)\right) \\ &\times \left(\alpha \int_{t_0}^{t} \tau^{\alpha-1}\left(f\left(\tau,\overline{y}\right) - f\left(\tau,y\right)\right)d\tau\right)^{\overline{\beta}}. \end{aligned} \tag{28}$$

However by hypothesis (ii), we can find $\varepsilon > 0$ small enough such that

$$\begin{aligned} &\exp\left(-\overline{\beta}t^{\alpha}H(t)\right)\Phi(t) \\ &\leq \exp\left(-\overline{\beta}t^{\alpha}H(t)\right)\alpha^{\overline{\beta}} \\ &\times \left(\int_{t_0}^{t} \varepsilon\beta\begin{pmatrix} \alpha\tau^{\alpha-1}H(\tau) \\ +\tau^{\alpha}h\left(\tau\right) \end{pmatrix}\exp\left(\tau^{\alpha}\overline{\beta}H(\tau)\right)d\tau\right)^{\overline{\beta}}, \\ &\leq \exp\left(-\overline{\beta}t^{\alpha}H(t)\right)\alpha^{\overline{\beta}}\varepsilon^{\overline{\beta}} \\ &\times \left(\int_{t_0}^{t} \exp\left(\tau^{\alpha}\overline{\beta}H(\tau)\right)' d\tau\right), \\ &\leq \exp\left(-\overline{\beta}t^{\alpha}H(t)\right)\alpha^{\overline{\beta}}\varepsilon^{\overline{\beta}}\exp\left(t^{\alpha}\overline{\beta}H(t)\right), \\ &= \alpha^{\overline{\beta}}\varepsilon^{\overline{\beta}} = (\alpha\varepsilon)^{\overline{\beta}}, \end{aligned} \tag{29}$$

and then

$$\lim_{t\to t_0^+} \exp\left(-\overline{\beta}t^{\alpha}H(t)\right)\Phi(t) = 0. \tag{30}$$

Therefore $\Phi(t) = 0$ so we get

$$\overline{y}(t) = y(t), \tag{31}$$

which completes the proof. □

We shall now evaluation the above condition in the case of the fractal fractional with power law.This will be acheived in the next section

4. The Witte's uniqueness conditions for Fractal-Fractional ordinary differential equations with exponential kernel

We shall consider in this section, the following fractal-fractional differential equation

$$\begin{cases} {}_{t_0}^{FFE}D_t^{\alpha,\beta}y(t) = f(t,y(t)), & \text{if } t > t_0, \\ y(t_0) = y_0, & \text{if } t = t_0. \end{cases} \tag{32}$$

that under the witte's condition $\alpha, \beta \in (0,1]$. The aim of this section is to show that under the Witte's condition equation has a unique solution if such solution exists in $[t_0, a]$. We will start our investigation on with the following lemma.

Lemma 2. *Let $f(t,y(t))$, $h(t)$ and $H(t)$ satisfy the properties presented before*

i) $\Phi(t) \leq (1-\alpha)\beta t^{\beta-1}h(t)\Phi(t) + \alpha\beta\int_{t_0}^{t}\tau^{\beta-1}h(\tau)\Phi(\tau)d\tau,$

ii) $\Phi(t) = o(\exp(H(t)))$ *as* $t \to t_0^+$, *then* $\Phi(t) = 0$ *in* $[t_0, a]$.

Proof. Let set

$$\Omega(t) = (1-\alpha)\beta h(t)t^{\beta-1}\Phi(t) + \alpha\beta\int_{t_0}^{t}\tau^{\beta-1}h(\tau)\Phi(\tau)d\tau. \tag{33}$$

From the hypothesis, we have that $\Omega(t)$ exists and is continuous in $[t_0, a]$. We recall that

$${}_{t_0}^{FFE}D_t^{\alpha,\beta}\left({}_{t_0}^{FFE}J_t^{\alpha}u(t)\right) = u(t). \tag{34}$$

Thus applying ${}_{t_0}^{FFE}D_t^{\alpha,\beta}$ on both sides yields

$${}_{t_0}^{FFE}D_t^{\alpha,\beta}\Omega(t) = h(t)\Phi(t) \leq h(t)\Omega(t). \tag{35}$$

Now, we shall find the sign of the

$$\begin{aligned} &{}_{t_0}^{FFE}D_t^{\alpha,\beta}\left[\Omega(t)\exp(-H(t))\right] \\ &= \frac{1}{\beta t^{\beta-1}}\,{}_{t_0}^{CFR}D_t^{\alpha}\left[\Omega(t)\exp(-H(t))\right], \\ &= \frac{1}{\beta t^{\beta-1}}\,{}_{t_0}^{CF}D_t^{\alpha}\left[\Omega(t)\exp(-H(t))\right]. \end{aligned} \tag{36}$$

Since $\Omega(t_0) = 0$, therefore, we have that

$${}_{t_0}^{CFR}D_t^{\alpha}\Omega(t) = {}_{t_0}^{CF}D_t^{\alpha}\Omega(t). \tag{37}$$

Therefore

$$\begin{aligned} &{}_{t_0}^{FFE}D_t^{\alpha,\beta}\left[\Omega(t)\exp(-H(t))\right] \\ &= \frac{1}{\beta t^{\beta-1}}\left[\begin{array}{c} \frac{1}{1-\alpha}\int_{t_0}^{t}\exp\left(\frac{-\alpha}{1-\alpha}(t-\tau)\right) \\ \times\left[\begin{array}{c}\Omega'(\tau)\exp(-H(\tau)) \\ -h(\tau)\Omega(\tau)\exp(-H(\tau))\end{array}\right]\end{array}\right]d\tau, \\ &= \frac{1}{\beta t^{\beta-1}}\left[\begin{array}{c} \frac{1}{1-\alpha}\int_{t_0}^{t}\exp\left(\frac{-\alpha}{1-\alpha}(t-\tau)\right) \\ \times\left[\Omega'(\tau) - h(\tau)\Omega(\tau)\right]\exp(-H(\tau))\end{array}\right]d\tau, \\ &\leq 0. \end{aligned} \tag{38}$$

In reference [9] it was shown that under the condition prescribed here

$$\Omega'(t) - h(t)\Omega(t) \leq 0, \tag{39}$$

therefore

$${}_{t_0}^{FFE}D_t^{\alpha,\beta}\left[\exp(-H(t))\Omega(t)\right] \leq 0. \tag{40}$$

Since by the hypothesis the integral is positive therefore

$${}_{t_0}^{FFE}D_t^{\alpha,\beta}\left[\exp(-H(t))\Omega(t)\right] \leq 0, \tag{41}$$

almost every where in $[t_0, a]$.

$$\begin{aligned} &\exp(-H(t))\Omega(t) \\ &= \exp(-H(t))\left[\begin{array}{c}(1-\alpha)\beta h(t)\Phi(t)t^{\beta-1} \\ +\beta\alpha\int_{t_0}^{t}\tau^{\beta-1}h(\tau)\Phi(\tau)d\tau\end{array}\right]. \end{aligned}$$

For a sufficient small t, we choose $\varepsilon > 0$ such that in the view of (iv), we get

$$\begin{aligned} \Omega(t)\exp(-H(t)) &\leq \exp(-H(t)) \\ &\times\left[\begin{array}{c}(1-\alpha)\beta h(t)t^{\beta-1}\exp(H(t))\varepsilon' \\ +\beta\alpha\varepsilon'\int_{t_0}^{t}\tau^{\beta-1}h(\tau)\exp(H(\tau))d\tau\end{array}\right], \\ &\leq \exp(-H(t)) \\ &\times\left[\begin{array}{c}(1-\alpha)\beta h(t)(\bar{t}_0)^{\beta-1}\exp(H(t))\varepsilon' \\ +(\bar{t}_0)^{\beta-1}\beta\alpha\varepsilon'\exp(H(t))\end{array}\right], \\ &\leq (1-\alpha)\beta h(t)(\bar{t}_0)^{\beta-1}\varepsilon' + (\bar{t}_0)^{\beta-1}\beta\alpha\varepsilon'. \end{aligned}$$

Using the continuity of $h(t)$ in $[t_0, a]$. $\exists t_1 \in [t_0, a]$ such that $\forall t \in [t_0, a]$

$$h(t_1) \geq h(t), \tag{42}$$

therefore

$$\begin{aligned}\exp\left(-H(t)\right)\Omega(t) &\leq \left(\bar{t}_0\right)^{\beta-1}\beta\left(h\left(t_1\right)\varepsilon'\left(1-\alpha\right)+\alpha\varepsilon'\right),\\ &\leq \mu\varepsilon' = \frac{\mu}{\mu}\varepsilon = \varepsilon.\end{aligned} \tag{43}$$

where

$$\varepsilon' = \frac{\varepsilon}{\mu} = \frac{\varepsilon}{\left(\bar{t}_0\right)^{\beta-1}\beta\left(h\left(t_1\right)\left(1-\alpha\right)+\alpha\beta\right)}. \tag{44}$$

Therefore

$$\Phi(t) = 0. \tag{45}$$

□

Theorem 3. *Let $f(t,y)$ be continuous in $\overline{S}_+$ in addition to Theorem 2 and Lemma 2 we have*

$\forall \varepsilon' > 0$,

$$\varepsilon' = \frac{\varepsilon}{\left(1-\alpha\right)\beta h\left(t\right)\left(\bar{t}_0\right)^{\beta-1}+\left(\bar{t}_0\right)^{\beta-1}\beta\alpha}. \tag{46}$$

Then the initial value problem (32) has almost one solution.

Proof. Let $y(t)$ and $\overline{y}(t)$ be two different solutions of our equation, then

$$\begin{aligned}|\Phi(t)| = |\overline{y}(t)-y(t)| &\leq (1-\alpha)\beta t^{\beta-1}\left|f(t,\overline{y}(t))-f(t,y(t))\right|\\ &+\alpha\beta\int_{t_0}^{t}\tau^{\beta-1}\left|f(\tau,\overline{y}(\tau))-f(\tau,y(\tau))\right|d\tau,\\ &\leq (1-\alpha)\beta t^{\beta-1}h\left(t\right)\Phi\left(t\right)\\ &+\alpha\beta\int_{t_0}^{t}\tau^{\beta-1}h\left(\tau\right)\Phi(\tau)d\tau,\\ &\leq \left(1-\alpha\right)\beta h\left(t\right)\left(\bar{t}_0\right)^{\beta-1}\varepsilon'\exp\left(H(t)\right)\\ &+\left(\bar{t}_0\right)^{\beta-1}\beta\alpha\varepsilon'\exp\left(H(t)\right),\\ &\leq \left(\begin{array}{c}\left(1-\alpha\right)\beta h\left(t\right)\left(\bar{t}_0\right)^{\beta-1}\\ +\left(\bar{t}_0\right)^{\beta-1}\beta\alpha\end{array}\right)\varepsilon'\exp\left(H(t)\right),\\ &\leq \mu\varepsilon'\exp\left(H(t)\right) = \varepsilon\exp\left(H(t)\right).\end{aligned} \tag{47}$$

□

Theorem 4. *Let $f(t,y)$ satisfies all the condition described in Theorem 3.*

Proof. Let $y(t)$ and $\overline{y}(t)$ be two different solution of equation (32). We set as before

$$\Psi(t) = \left(\overline{y}-y\right)^{\overline{\beta}}. \tag{48}$$

We have that $\Psi(t_0) = 0$, thus

$$\begin{aligned}&{}^{FFE}_{t_0}D_t^{\alpha,\beta}\Psi(t)\\ &= \frac{1}{\beta t^{\beta-1}}\,{}^{CFR}_{t_0}D_t^{\alpha}\Psi(t) = \frac{1}{\beta t^{\beta-1}}\,{}^{CF}_{t_0}D_t^{\alpha}\Psi(t),\\ &= \frac{1}{\beta t^{\beta-1}}\frac{1}{1-\alpha}\int_{t_0}^{t}\Psi'(\tau)\exp\left(\frac{-\alpha}{1-\alpha}(t-\tau)\right)d\tau,\\ &= \frac{1}{\beta t^{\beta-1}}\frac{1}{1-\alpha}\int_{t_0}^{t}\left[\overline{\beta}\left(\overline{y}-y\right)'\left(\overline{y}-y\right)^{\overline{\beta}-1}\right]\\ &\times\exp\left(\frac{-\alpha}{1-\alpha}(t-\tau)\right)d\tau,\\ &\leq \overline{\beta}\delta\left[{}^{FFE}_{t_0}D_t^{\alpha,\beta}\overline{y}-{}^{FFE}_{t_0}D_t^{\alpha,\beta}y\right],\\ &\leq \overline{\beta}\delta\left|f(t,\overline{y}(t))-f(t,y(t))\right|,\\ &\leq \overline{\beta}\delta h(t)\Psi(t),\end{aligned} \tag{49}$$

here

$$\delta = \begin{cases}\max\limits_{t\in[t_0,a]}\left|\overline{y}-y\right|^{\overline{\beta}-1}, & \text{if} \quad y'-\overline{y}' > 0,\\ \min\limits_{t\in[t_0,a]}\left|\overline{y}-y\right|^{\overline{\beta}-1}, & \text{if} \quad y'-\overline{y}' < 0.\end{cases}$$

In the view of the first hypothesis. Thus

$${}^{FFE}_{t_0}D_t^{\alpha,\beta}\Psi(t) \leq \overline{\Delta}\Psi(t), \tag{50}$$

almost every where in $[t_0,a]$.

$$\begin{aligned}&{}^{FFE}_{t_0}D_t^{\alpha,\beta}\left[\exp\left(-\overline{\beta}H(t)\right)\Psi(t)\right]\\ &= \frac{1}{(1-\alpha)\beta t^{\beta-1}}\frac{d}{dt}\int_{t_0}^{t}\exp\left(-\frac{\alpha}{1-\alpha}\left(t-\tau\right)\right)\\ &\times\exp\left(-\overline{\beta}H(\tau)\right)\Psi(\tau)d\tau,\\ &= \frac{1}{(1-\alpha)\beta t^{\beta-1}}\int_{t_0}^{t}\exp\left(-\frac{\alpha}{1-\alpha}\left(t-\tau\right)\right)\\ &\times\left(\Psi(\tau)\exp\left(-\overline{\beta}H(\tau)\right)\right)'d\tau\\ &-\frac{1}{\beta t^{\beta-1}}\frac{1}{1-\alpha}\Psi(t_0)\exp\left(-\overline{\beta}H(t_0)\right)\exp\left(\frac{-\alpha}{1-\alpha}t\right).\end{aligned} \tag{51}$$

But

$$\Psi(t_0) = 0, \tag{52}$$

therefore

$$
\begin{aligned}
&{}_{t_0}^{FFE}D_t^{\alpha,\beta}\left[\exp\left(-\overline{\beta}H(t)\right)\Psi(t)\right] \qquad (53)\\
&=\frac{1}{(1-\alpha)\beta t^{\beta-1}}\int_{t_0}^{t}\exp\left(-\frac{\alpha}{1-\alpha}\left(t-\tau\right)\right)\\
&\times\left(\exp\left(-\overline{\beta}H(\tau)\right)\Psi(\tau)\right)'d\tau,\\
&=\frac{1}{(1-\alpha)\beta t^{\beta-1}}\int_{t_0}^{t}\exp\left(-\frac{\alpha}{1-\alpha}\left(t-\tau\right)\right)\\
&\times\begin{pmatrix}\Psi'(\tau)\exp\left(-\overline{\beta}H(\tau)\right)\\ -\overline{\beta}H(\tau)\exp\left(-\overline{\beta}H(\tau)\right)\Psi(\tau)\end{pmatrix}d\tau.
\end{aligned}
$$

In reference [9], it was shown that

$$
\Psi'(t)\exp\left(-\overline{\beta}H(t)\right)-\overline{\beta}H(t)\exp\left(-\overline{\beta}H(t)\right)\Psi(t)<0. \qquad (54)
$$

Therefore

$$
{}_{t_0}^{FFE}D_t^{\alpha,\beta}\left[\Psi(t)\exp\left(-\overline{\beta}H(t)\right)\right]\leq 0.
$$

$$
\begin{aligned}
&{}_{t_0}^{FFE}D_t^{\alpha,\beta}\left[\exp\left(-\overline{\beta}H(t)\right)\Psi(t)\right] \qquad (55)\\
&=\frac{1}{\beta t^{\beta-1}}\,{}_{t_0}^{CFR}D_t^{\alpha}\left[\exp\left(-\overline{\beta}H(t)\right)\Psi(t)\right],\\
&=\frac{1}{\beta t^{\beta-1}}\,{}_{t_0}^{CF}D_t^{\alpha}\left[\exp\left(-\overline{\beta}H(t)\right)\Psi(t)\right],\\
&=\frac{1}{\beta t^{\beta-1}}\frac{1}{1-\alpha}\int_{t_0}^{t}\exp\left(\frac{-\alpha}{1-\alpha}(t-\tau)\right)\\
&\times\begin{bmatrix}-\Psi'(\tau)\overline{\beta}\exp\left(-\overline{\beta}H(\tau)\right)\\ -\overline{\beta}h(\tau)\exp\left(-\overline{\beta}H(\tau)\right)\Psi(\tau)\end{bmatrix}d\tau,\\
&=\frac{1}{\beta t^{\beta-1}}\frac{1}{1-\alpha}\int_{t_0}^{t}\exp\left(\frac{-\alpha}{1-\alpha}(t-\tau)\right)\\
&\times\left[\left[\Psi'(\tau)+h(\tau)\Psi(\tau)\right]\exp\left(-\overline{\beta}H(\tau)\right)\right]d\tau,\\
&=\frac{1}{\beta t^{\beta}}\frac{1}{1-\alpha}\int_{t_0}^{t}\exp\left(\frac{-\alpha}{1-\alpha}(t-\tau)\right)\\
&\times\left[\left[\Psi'(\tau)-h(\tau)\Psi(\tau)\right]\exp\left(-\overline{\beta}H(\tau)\right)\right]d\tau,\\
&\leq 0.
\end{aligned}
$$

Therefore, we have

$$
{}_{t_0}^{FFE}D_t^{\alpha,\beta}\left[\exp\left(-\overline{\beta}H(t)\right)\Psi(t)\right]<0. \qquad (56)
$$

Following the routine presented earlier we shall have for ε'

$$
\begin{aligned}
&\exp\left(-\overline{\beta}H(t)\right)\Psi(t)\\
&=\exp\left(-\overline{\beta}H(t)\right)\left(\overline{y}(t)-y(t)\right)^{\overline{\beta}},\\
&=\exp\left(-\overline{\beta}H(t)\right)\begin{pmatrix}(1-\alpha)\beta t^{\beta-1}\left(f(t,\overline{y}(t))-f(t,y(t))\right)\\ +\alpha\beta\int_{t_0}^{t}\tau^{\beta-1}\left(f(\tau,\overline{y}(\tau))-f(\tau,y(\tau))\right)d\tau\end{pmatrix}^{\overline{\beta}},\\
&\leq\exp\left(-\overline{\beta}H(t)\right)\begin{pmatrix}(1-\alpha)\beta t^{\beta-1}\varepsilon'\exp\left(H(t)\right)h(t)\\ +\alpha\beta\int_{t_0}^{t}h(\tau)\tau^{\beta-1}\varepsilon'\exp\left(H(\tau)\right)d\tau\end{pmatrix}^{\overline{\beta}},\\
&\leq\exp\left(-\overline{\beta}H(t)\right)\begin{pmatrix}(1-\alpha)\beta\left(\bar{t}_0\right)^{\beta-1}\varepsilon'\exp\left(H(t)\right)h(t)\\ +\left(\bar{t}_0\right)^{\beta-1}\alpha\beta\varepsilon'\int_{t_0}^{t}h(\tau)\exp\left(H(\tau)\right)d\tau\end{pmatrix}^{\overline{\beta}},\\
&\leq\exp\left(-\overline{\beta}H(t)\right)\begin{pmatrix}(1-\alpha)\overline{\beta}\left(\bar{t}_0\right)^{\beta-1}\varepsilon'\exp\left(H(t)\right)h(t)\\ +\left(\bar{t}_0\right)^{\beta-1}\alpha\overline{\beta}\varepsilon'\exp\left(H(t)\right)\end{pmatrix}^{\overline{\beta}},\\
&\leq\exp\left(-\overline{\beta}H(t)\right)\begin{pmatrix}(1-\alpha)\beta\left(\bar{t}_0\right)^{\beta-1}\varepsilon'\exp\left(H(t)\right)h(t_1)\\ +\left(\bar{t}_0\right)^{\beta-1}\alpha\beta\varepsilon'\exp\left(H(t)\right)\end{pmatrix}^{\overline{\beta}},\\
&\leq\exp\left(-\overline{\beta}H(t)\right)\exp\left(\beta H(t)\right)\left(\varepsilon'\right)^{\overline{\beta}}\left((1-\alpha)\beta\left(\bar{t}_0\right)^{\beta-1}+\left(\bar{t}_0\right)^{\beta-1}\alpha\beta\right)^{\overline{\beta}},\\
&\leq\left(\varepsilon'\right)^{\overline{\beta}}\mu^{\overline{\beta}},\ \mu=\left((1-\alpha)\beta+\alpha\beta\right)\left(\bar{t}_0\right)^{\beta-1}.
\end{aligned} \qquad (57)
$$

We choose

$$
\varepsilon'=\frac{\varepsilon}{\mu}, \qquad (58)
$$

such that

$$
\exp\left(-\overline{\beta}H(t)\right)\Psi(t)\leq\varepsilon^2. \qquad (59)
$$

Therefore

$$
\lim_{t\to 0^+}\exp\left(-\overline{\beta}H(t)\right)\Psi(t)=0. \qquad (60)
$$

So we conclude that

$$
\begin{aligned}
&\Psi(t)=0 \qquad (61)\\
&\Rightarrow\overline{y}(t)=y(t),
\end{aligned}
$$

which concludes the proof. □

5. The Witte's uniqueness conditions for Fractal-Fractional ordinary differential equations with power-law kernel

In this section, we shall consider the following differential equation

$$
\begin{cases}{}_{t_0}^{FFP}D_t^{\alpha,\beta}y(t)=f(t,y(t)), & \text{if } t>t_0,\\ y(t_0)=y_0, & \text{if } t=t_0.\end{cases} \qquad (62)
$$

We aim to show that if the solution of the above equation exists in $S_+=\{(t,y)\mid t_0<t\leq a,\ |y|<\infty\}$, $\alpha\in(0,1]$, $\beta\in(0,1]$ then it is unique.

Lemma 3. *Let $\Phi(t)$ be a non negative continuous in $(t_0, a]$ such that $\Phi(t_0) = 0$. Let*

i) $h(t) > 0$ be continuous function in $(t_0, a]$,

ii) We can find a function $H(t)$ in $(t_0, a]$ such that $H'(t) = h(t)$ for almost all $t \in (t_0, a]$ and $\lim_{t \to t_0^+} H(t)$ exists,

iii) $$\Phi(t) \leq \frac{\beta}{\Gamma(\alpha)} \int_{t_0}^{t} \tau^{\beta-1} (t-\tau)^{\alpha-1} h(\tau)\Phi(\tau) d\tau,$$

$\forall t \in (t_0, a]$ and

iv) $\Phi(t) = o(\exp(H(t)))$ as $t \to t_0^+$. Then

$$\Phi(t) = 0, \tag{63}$$

in $(t_0, a]$.

Proof. Let

$$\Omega(t) = \frac{\beta}{\Gamma(\alpha)} \int_{t_0}^{t} \tau^{\beta-1} (t-\tau)^{\alpha-1} h(\tau)\Phi(\tau) d\tau. \tag{64}$$

The existence and the continuty of the function $\Omega(t)$ is assumed since the hypothesis of the Lemma. Therefore we have that

$$ {}^{FFP}_{t_0}D_t^{\alpha,\beta}\Omega(t) = h(t)\Phi(t) \leq h(t)\Phi(t). \tag{65}$$

We note that

$$\begin{aligned}
&{}^{FFP}_{t_0}D_t^{\alpha,\beta}\Omega(t) \\
&= \frac{\beta}{\Gamma(\alpha)} {}^{FFP}_{t_0}D_t^{\alpha,\beta} \left[\int_{t_0}^{t} \tau^{\beta-1} (t-\tau)^{\alpha-1} h(\tau)\Phi(\tau) d\tau \right], \\
&= \frac{\beta}{\Gamma(\alpha)\beta t^{\beta-1}} \frac{d}{dt} \frac{1}{\Gamma(1-\alpha)} \\
&\times \left(\int_{t_0}^{t} \tau^{\beta-1} (t-\tau)^{-\alpha} \left[\int_{t_0}^{\tau} l^{\beta-1} (\tau-l)^{\alpha-1} h(l)\Phi(l) dl \right] d\tau \right), \\
&= \frac{\beta}{\beta t^{\beta-1}} {}^{RL}_{t_0}D_t^{\alpha} \left[{}^{RL}_{t_0}J_t^{\alpha} \left(t^{\beta-1} h(t)\Phi(t) \right) \right], \\
&= \frac{\beta t^{\beta-1} h(t)\Phi(t)}{\beta t^{\beta-1}}, \\
&= h(t)\Phi(t).
\end{aligned} \tag{66}$$

We recall that $\Omega(t_0) = 0$, then

$$ {}^{RL}_{t_0}D_t^{\alpha}\Omega(t) = {}^{C}_{t_0}D_t^{\alpha}\Omega(t). \tag{67}$$

$$\begin{aligned}
&{}^{FFP}_{t_0}D_t^{\alpha,\beta} [\exp(-H(t))\Omega(t)] \\
&= \frac{1}{\beta t^{\beta-1}} {}^{RL}_{t_0}D_t^{\alpha} [\exp(-H(t))\Omega(t)], \\
&= \frac{1}{\beta t^{\beta-1}} {}^{C}_{t_0}D_t^{\alpha} [\exp(-H(t))\Omega(t)], \\
&= \frac{1}{\beta t^{\beta}} \frac{1}{\Gamma(1-\alpha)} \int_{t_0}^{t} (t-\tau)^{-\alpha} \begin{bmatrix} \Omega'(\tau)\exp(-H(\tau)) \\ -h(\tau)\Omega(\tau)\exp(-H(\tau)) \end{bmatrix} d\tau, \\
&= \frac{1}{\beta t^{\beta}} \frac{1}{\Gamma(1-\alpha)} \int_{t_0}^{t} (t-\tau)^{-\alpha} \exp(-H(\tau)) \left[\Omega'(\tau) - h(\tau)\Omega(\tau) \right] d\tau,
\end{aligned} \tag{68}$$

We have due to reference [9] that

$$\Omega'(\tau) - h(\tau)\Omega(\tau) \leq 0. \tag{69}$$

Therefore

$$ {}^{FFP}_{t_0}D_t^{\alpha,\beta} [\exp(-H(t))\Omega(t)] \leq 0. \tag{70}$$

We can now have for a small t

$$\begin{aligned}
&\exp(-H(t))\Omega(t) \\
&= \exp(-H(t)) \frac{\beta}{\Gamma(\alpha)} \int_{t_0}^{t} \tau^{\beta-1} (t-\tau)^{\alpha-1} h(\tau)\Phi(\tau) d\tau, \\
&\leq \frac{\exp(-H(t))\beta}{\Gamma(\alpha)} \int_{t_0}^{t} \tau^{\beta} (t-\tau)^{\alpha} h(\tau)\Phi(\tau) d\tau, \\
&\leq \frac{\exp(-H(t))\beta a^{\beta+\alpha}}{\Gamma(\alpha)} \int_{t_0}^{t} h(\tau)\Phi(\tau) d\tau.
\end{aligned} \tag{71}$$

By hypothesis (iv), we have

$$\begin{aligned}
\exp(-H(t))\Omega(t) &\leq \frac{\exp(-H(\tau))\beta a^{\beta+\alpha}}{\Gamma(\alpha)} \varepsilon \exp\left(H(\tau) \times \frac{\Gamma(\alpha)}{\beta a^{\beta+\alpha}} \right), \\
&\leq \varepsilon.
\end{aligned} \tag{72}$$

$$\lim_{t \to 0^+} \exp(-H(t))\Omega(t) = 0. \tag{73}$$

This leads to

$$\exp(-H(t))\Omega(t) \leq 0, \ \forall t > t_0, \tag{74}$$

which implies

$$\frac{\beta}{\Gamma(\alpha)} \int_{t_0}^{t} \tau^{\beta-1} (t-\tau)^{\alpha-1} h(\tau)\Phi(\tau) d\tau \leq 0, \tag{75}$$

which is a contradiction therefore

$$\Phi(t) = 0. \tag{76}$$

□

Theorem 5. *Let f be continuous in $S_+ = \{(t,y) \mid t_0 < t \leq a,\ |y| < \infty\}$ such that $\forall (t,y),(t,\overline{y}) \in S_+$*

i) $|f(t,y) - f(t,\overline{y})| \leq h(t)\,|y-\overline{y}|\,,$

ii) $f(t,y) - f(t,\overline{y}) = o\left(\exp(H(t))\right),$

as $t \to t_0^+$ uniformly with respect to $y,\overline{y} \in [-\delta,\delta]$, $\delta > 0$ arbitrary, where $h(t) = H'(t)$ are the same as in above. Then the considered equation has almost one solution.

Proof. Let $\overline{y}(t)$ and $y(t)$ be two different solutions, we have that

$$\begin{aligned}
|\overline{y}(t) - y(t)| &\leq \frac{\beta}{\Gamma(\alpha)} \int_{t_0}^{t} \tau^{\beta-1} (t-\tau)^{\alpha-1} \left| \begin{array}{c} f(\tau,\overline{y}(\tau)) \\ -f(\tau,y(\tau)) \end{array} \right| d\tau, \\
&\leq \frac{\beta}{\Gamma(\alpha)} \int_{t_0}^{t} \tau^{\beta-1} (t-\tau)^{\alpha-1} h(\tau)\,|y-\overline{y}|\,d\tau, \\
&\leq \frac{\beta}{\Gamma(\alpha)} \int_{t_0}^{t} \tau^{\beta-1} (t-\tau)^{\alpha} h(\tau)\,|y-\overline{y}|\,d\tau.
\end{aligned} \tag{77}$$

In the view of (ii), we have

$$\begin{aligned}
|\overline{y}(t) - y(t)| &\leq \varepsilon \frac{\beta\Gamma(\alpha)}{\beta a^{\beta+\alpha}\Gamma(\alpha)} \int_{t_0}^{t} a^{\beta+\alpha} h(\tau) \exp(H(\tau)) d\tau, \\
&\leq \varepsilon \exp(H(t)).
\end{aligned} \tag{78}$$

The result of the previous lemma leads to

$$\overline{y}(t) = y(t). \tag{79}$$

□

Theorem 6. *Let f be continuous in $\overline{S}_+ = \{(t,y) \mid t_0 < t \leq a,\ |y| < \infty\}$ such that $\overline{\beta} \in (1,2]$, $\alpha,\beta \in (0,1]$,$\forall (t,y),(t,\overline{y}) \in \overline{S}_+$, we have*

i) $\left(f(t,\overline{y}) - f(t,y)\right)(\overline{y}-y) \leq h(t)\,(\overline{y}-y)^{\overline{\beta}}\,,$

ii) $f(t,\overline{y}) - f(t,y) = o\left(h(t)\exp\left(H(t)\right)\right),$

uniformly with respect to $y,\overline{y} \in [-\delta,\delta]$, $\delta > 0$ arbitrary then

$$\overline{y}(t) = y(t). \tag{80}$$

Proof. Let $\overline{y}$ and y be two solutions, we put $\Phi(t) = (\overline{y}(t) - y(t))^{\overline{\beta}}$. We have that at $t = t_0$, $\Phi(t_0) = 0$ initial condition then we will have that

$${}_{t_0}^{C}D_t^{\alpha}\Phi(t) = {}_{t_0}^{RL}D_t^{\alpha}\Phi(t). \tag{81}$$

However,

$$\begin{aligned}
&{}_{t_0}^{FFP}D_t^{\alpha,\beta}\Phi(t) \\
&= \frac{1}{\beta t^{\beta-1}}\,{}_{t_0}^{RL}D_t^{\alpha}\Phi(t) = \frac{1}{\beta t^{\beta-1}}\,{}_{t_0}^{C}D_t^{\alpha}\Phi(t), \\
&= \frac{1}{\beta t^{\beta-1}} \frac{1}{\Gamma(1-\alpha)} \int_{t_0}^{t} (t-\tau)^{-\alpha}\,\Phi'(\tau)d\tau, \\
&= \frac{1}{\beta t^{\beta-1}} \frac{1}{\Gamma(1-\alpha)} \int_{t_0}^{t} (t-\tau)^{-\alpha} \left[\overline{\beta}\,(\overline{y}-y)'\,(\overline{y}-y)^{\beta-1}\right] d\tau, \\
&= \frac{1}{\beta t^{\beta-1}} \left[\begin{array}{c} \frac{1}{\Gamma(1-\alpha)} \int_{t_0}^{t} \overline{\beta}\,(\overline{y})'\,(\overline{y}-y)^{\beta-1}\,(t-\tau)^{-\alpha}\,d\tau \\ -\frac{1}{\Gamma(1-\alpha)} \int_{t_0}^{t} \overline{\beta} y'\,(\overline{y}-y)^{\beta-1}\,(t-\tau)^{-\alpha}\,d\tau \end{array} \right], \\
&\leq \overline{\beta}\Lambda \left({}_{t_0}^{FFP}D_t^{\alpha,\beta}\overline{y}' - {}_{t_0}^{FFP}D_t^{\alpha,\beta}y\right), \\
&\leq \overline{\beta}\Lambda\,|f(t,\overline{y}(t)) - f(t,y(t))|\,, \\
&\leq \overline{\beta}\Lambda h(t)\,(\overline{y}(t) - y(t))^{\overline{\beta}}\,.
\end{aligned} \tag{82}$$

here

$$\Lambda = \begin{cases} \max\limits_{t\in[t_0,a]} |\overline{y}-y|^{\overline{\beta}-1}\,, & \text{if} \quad y' - \overline{y}' > 0, \\ \min\limits_{t\in[t_0,a]} |\overline{y}-y|^{\overline{\beta}-1}\,, & \text{if} \quad y' - \overline{y}' < 0. \end{cases}$$

By the hypothesis (i), thus

$$\begin{aligned}
{}_{t_0}^{FFP}D_t^{\alpha,\beta}\Phi(t) &\leq \overline{\beta}\Lambda h(t)\Phi(t), \\
&\leq \overline{\beta\Lambda} h(t)\Phi(t).
\end{aligned} \tag{83}$$

$${}_{t_0}^{FFP}D_t^{\alpha,\beta}\Phi(t) - \overline{\beta\Lambda} h(t)\Phi(t) \leq 0. \tag{84}$$

$$\begin{aligned}
&{}_{t_0}^{FFP}D_t^{\alpha,\beta}\left[\exp(-\overline{\beta}H(t))\Phi(t)\right] \\
&= \frac{1}{\beta t^{\beta-1}}\,{}_{t_0}^{RL}D_t^{\alpha}\left[\exp(-\overline{\beta}H(t))\Phi(t)\right], \\
&= \frac{1}{\beta t^{\beta-1}}\,{}_{t_0}^{C}D_t^{\alpha}\left[\exp(-\overline{\beta}H(t))\Phi(t)\right], \\
&= \frac{1}{\beta t^{\beta-1}} \frac{1}{\Gamma(1-\alpha)} \int_{t_0}^{t} (t-\tau)^{-\alpha} \left(\exp(-\overline{\beta}H(\tau))\Phi(\tau)\right)' d\tau, \\
&= \frac{1}{\beta t^{\beta-1}} \frac{1}{\Gamma(1-\alpha)} \int_{t_0}^{t} (t-\tau)^{-\alpha} \left[\begin{array}{c} -\overline{\beta}h(\tau)\Phi(\tau) \\ -\beta\,(\overline{y}-y)'\,(\overline{y}-y)^{\beta-1}\,H(\tau) \end{array} \right] d\tau, \\
&= \frac{-1}{\beta t^{\beta-1}} \frac{1}{\Gamma(1-\alpha)} \int_{t_0}^{t} (t-\tau)^{-\alpha} \left[\begin{array}{c} \overline{\beta}h(\tau)\Phi(\tau) \\ +\beta\,(\overline{y}-y)'\,(\overline{y}-y)^{\beta-1}\,H(\tau) \end{array} \right] d\tau, \\
&\leq 0,
\end{aligned} \tag{85}$$

for almost all $t \in [0,a]$. This shows that $\exp(-\overline{\beta}H(t))\Phi(t)$ is non increasing for a small t.

$$\begin{aligned}
\exp(-\overline{\beta}H(t))\Phi(t) &= \exp(-\beta H(t))\,(\overline{y}-y)^{\overline{\beta}}\,, \\
&= \exp(-\beta H(t)) \\
&\quad \times \left(\frac{\beta}{\Gamma(\alpha)} \int_{t_0}^{t} \tau^{\beta-1} (t-\tau)^{\alpha-1} \left| \begin{array}{c} f(\tau,\overline{y}(\tau)) \\ -f(\tau,y(\tau)) \end{array} \right| d\tau \right)^{\overline{\beta}}.
\end{aligned} \tag{86}$$

In the view of the second hypothesis, we will have that

$$\begin{aligned}&\exp(-\overline{\beta}H(t))\Phi(t)\\&\leq\left(\frac{\beta}{\Gamma(\alpha)}\frac{\Gamma(\alpha)\,a^{\beta+\alpha}}{\beta a^{\beta+\alpha}}\right)^{\overline{\beta}}\varepsilon^{\overline{\beta}}\exp(\overline{\beta}H(t))\exp(-\overline{\beta}H(t)),\\&=\varepsilon^{\overline{\beta}}.\end{aligned}\tag{87}$$

Therefore

$$\exp(-\overline{\beta}H(t))\Phi(t)\leq\varepsilon^{\overline{\beta}},\tag{88}$$

$$\lim_{t\to0^{+}}\exp(-\beta H(t))\Phi(t)=0.$$

Therefore

$$\begin{aligned}\Phi(t)&=0,\\&\Longrightarrow\overline{y}(t)=y(t)\text{ in }[t_0,a].\end{aligned}\tag{89}$$

□

6. The Witte's uniqueness conditions for Fractal-Fractional ordinary differential equations with the Mittag Leffler kernel

In this section, we will consider the following fractal-fractional differential equation

$$\begin{cases}{}^{FFM}_{t_0}D_t^{\alpha,\beta}y(t)=f(t,y(t)), & \text{if } t>t_0,\\ y(t_0)=y_0, & \text{if } t=t_0.\end{cases}\tag{90}$$

Assuming the existence of the solution $y(t)$, we shall show that $y(t)$ is unique.

Lemma 4. *Let $\Phi(t)$ be a nonnegative continuous in $[t_0,a]$ and*
i) Let $h(t)>0$ be a continuous function in $(t_0,a]$ such that $1-zh(t_1)>0$,
ii) $\Phi(t)\leq(1-\alpha)t^{\beta-1}\beta h(t)\Phi(t)+\frac{\alpha\beta}{\Gamma(\alpha)}\int_{t_0}^{t}\tau^{\beta-1}(t-\tau)^{\alpha-1}h(\tau)\Phi(\tau)d\tau$,

iii) and $\frac{\beta}{\Gamma(\alpha)}\int_{t_0}^{t}\tau^{\beta-1}(t-\tau)^{\alpha-1}h(\tau)d\tau$, *exists Then*

$$\Phi(t)=0.\tag{91}$$

in $[t_0,a]$.

Proof. Let $\Phi(t)$ and $h(t)$ satisfy the condition of the theorem, then, we set

$$\Omega(t)=(1-\alpha)t^{\beta-1}\beta h(t)\Phi(t)+\frac{\alpha\beta}{\Gamma(\alpha)}\int_{t_0}^{t}\tau^{\beta-1}(t-\tau)^{\alpha-1}h(\tau)\Phi(\tau)d\tau.\tag{92}$$

We have from the fundamental theorem of fractal-fractional calculus that

$${}^{FFM}_{t_0}D_t^{\alpha,\beta}\left({}^{FFM}_{t_0}J_t^{\alpha,\beta}f(t)\right)=f(t).\tag{93}$$

Therefore

$${}^{FFM}_{t_0}D_t^{\alpha,\beta}\Omega(t)={}^{FFM}_{t_0}D_t^{\alpha,\beta}\left({}^{FFM}_{t_0}J_t^{\alpha,\beta}\left(h(t)\Phi(t)\right)\right)=h(t)\Phi(t),\tag{94}$$

which produces

$${}^{FFM}_{t_0}D_t^{\alpha,\beta}\Omega(t)\leq h(t)\Omega(t).\tag{95}$$

Then, we obtain $\Omega(t)$ as

$$\begin{aligned}\Omega(t)&\leq(1-\alpha)\beta t^{\beta-1}h(t)\Omega(t)\\&\quad+\frac{\alpha\beta}{\Gamma(\alpha)}\int_{t_0}^{t}\tau^{\beta-1}(t-\tau)^{\alpha-1}h(\tau)\Omega(\tau)d\tau,\\&\leq(1-\alpha)\beta(\bar{t}_0)^{\beta-1}h(t_1)\Omega(t)\\&\quad+\frac{\alpha\beta}{\Gamma(\alpha)}\int_{t_0}^{t}\tau^{\beta-1}(t-\tau)^{\alpha-1}h(\tau)\Omega(\tau)d\tau,\end{aligned}$$

$$\Omega(t)\leq\frac{\alpha\beta}{\Gamma(\alpha)(1-zh(t_1))}\int_{t_0}^{t}\tau^{\beta-1}(t-\tau)^{\alpha-1}h(\tau)\Omega(\tau)d\tau.\tag{96}$$

We put

$$\Delta=\frac{\alpha\beta}{\Gamma(\alpha)(1-zh(t_1))}.\tag{97}$$

By the Gronwall inequality

$$\begin{aligned}\Omega(t)&\leq o\exp\left(\Delta\int_{t_0}^{t}\tau^{\beta-1}(t-\tau)^{\alpha-1}h(\tau)d\tau\right),\\&=o\exp\left(\frac{\alpha}{(1-zh(t_1))}\,{}^{FFM}_{t_0}J_t^{\alpha,\beta}h(t)\right),\\&=0.\end{aligned}\tag{98}$$

$z=(1-\alpha)\beta(\bar{t}_0)^{\beta-1}$, which is contraction. Therefore

$$\Omega(t)=0\Rightarrow\Phi(t)=0,\tag{99}$$

in $[t_0,a]$. □

Lemma 5. *Let $\Phi(t)$, $h(t)$ and $H(t)$ be the same like before and $\Phi(t_0)=0$.*
i) $\Phi(t)=o\left(\exp\left(t^{\beta}H(t)\right)\right)$ as $t\to t_0^{+}$ then

$$\Phi(t)=0,\ \forall t\in[t_0,a].\tag{100}$$

Proof. Let $\Phi(t)$ and $h(t)$ satisfy the condition above, then

$$\begin{aligned}\Omega(t)&=(1-\alpha)\beta t^{\beta-1}h(t)\Phi(t)\\&\quad+\frac{\alpha\beta}{\Gamma(\alpha)}\int_{t_0}^{t}\tau^{\beta-1}(t-\tau)^{\alpha-1}h(\tau)\Phi(\tau)d\tau,\end{aligned}$$

exists

$$ {}_{t_0}^{FFM}D_t^{\alpha,\beta}\Omega(t) \leq h(t)\Omega(t). \tag{101}$$

$$\begin{aligned}
&{}_{t_0}^{FFM}D_t^{\alpha,\beta}\left(\Omega(t)\exp\left(-H(t)\right)\right)\\
&=\frac{1}{(1-\alpha)\beta t^{\beta-1}}\frac{d}{dt}\int_{t_0}^{t}E_\alpha\left(-\frac{\alpha}{1-\alpha}\left(t-\tau\right)^\alpha\right)\Omega(\tau)\exp\left(-H(\tau)\right)d\tau.
\end{aligned} \tag{102}$$

Since $\Phi\left(t_0\right)=0$, we will have $\Omega\left(t_0\right)$ therefore,

$$\begin{aligned}
&{}_{t_0}^{FFM}D_t^{\alpha,\beta}\left(\Omega(t)\exp\left(-H(t)\right)\right)\\
&=\frac{1}{(1-\alpha)\beta t^{\beta-1}}\frac{d}{dt}\int_{t_0}^{t}E_\alpha\left(-\frac{\alpha}{1-\alpha}\left(t-\tau\right)^\alpha\right)\Omega(\tau)\exp\left(-H(\tau)\right)d\tau,\\
&=\frac{t^{1-\beta}}{(1-\alpha)\beta}\int_{t_0}^{t}E_\alpha\left(-\frac{\alpha}{1-\alpha}\left(t-\tau\right)^\alpha\right)\left[\begin{array}{c}\Omega'(\tau)\exp\left(-H(\tau)\right)\\-h(\tau)\Omega(\tau)\exp\left(-H(\tau)\right)\end{array}\right]d\tau,\\
&=\frac{t^{1-\beta}}{(1-\alpha)\beta}\int_{t_0}^{t}E_\alpha\left(-\frac{\alpha}{1-\alpha}\left(t-\tau\right)^\alpha\right)\exp\left(-H(\tau)\right)\left[\begin{array}{c}\Omega'(\tau)\\-h(\tau)\Omega(\tau)\end{array}\right]d\tau,
\end{aligned} \tag{103}$$

whereas from [9], we have that

$$\Omega(\tau)-h(\tau)\Omega(\tau)\leq 0. \tag{104}$$

Therefore since $E_\alpha\left(-\frac{\alpha}{1-\alpha}\left(t-\tau\right)^\alpha\right)>0$, we concluded that

$$ {}_{t_0}^{FFM}D_t^{\alpha,\beta}\left(\Omega(t)\exp\left(-H(t)\right)\right)<0. \tag{105}$$

$$\begin{aligned}
&\exp\left(-t^\beta H(t)\right)\Omega(t)\\
&=\exp\left(-t^\beta H(t)\right)\left[\begin{array}{c}(1-\alpha)\beta t^{\beta-1}h(t)\Phi(t)\\+\frac{\alpha\beta}{\Gamma(\alpha)}\int_{t_0}^{t}\tau^{\beta-1}\left(t-\tau\right)^{\alpha-1}h(\tau)\Phi(\tau)d\tau\end{array}\right],\\
&=\exp\left(-t^\beta H(t)\right)(1-\alpha)\beta t^{\beta-1}h(t)\Phi(t)\\
&+\frac{\alpha\beta\exp\left(-t^\beta H(t)\right)}{\Gamma(\alpha)}\int_{t_0}^{t}\tau^{\beta-1}\left(t-\tau\right)^{\alpha-1}h(\tau)\Phi(\tau)d\tau,\\
&\leq\exp\left(-t^\beta H(t)\right)(1-\alpha)\beta t^{\beta}h(t)\Phi(t)\\
&+\frac{\alpha\beta\exp\left(-t^\beta H(t)\right)}{\Gamma(\alpha)}\int_{t_0}^{t}\tau^{\beta}\left(t-\tau\right)^{\alpha-1}h(\tau)\Phi(\tau)d\tau,\\
&\leq\exp\left(-t^\beta H(t)\right)(1-\alpha)\beta t^{\beta}h(t)\Phi(t)\\
&+\frac{\exp\left(-t^\beta H(t)\right)\alpha\beta\varepsilon}{\Gamma(\alpha)}\int_{t_0}^{t}\left(t-\tau\right)^{\alpha-1}\left[\begin{array}{c}\beta\tau^{\beta-1}H(\tau)\\+\tau^\beta h(\tau)\end{array}\right]\\
&\qquad\times\exp\left(-\tau^\beta H(\tau)\right)d\tau.
\end{aligned} \tag{106}$$

In the view (i)

$$\begin{aligned}
&\leq\exp\left(-t^\beta H(t)\right)(1-\alpha)\varepsilon'\beta t^\beta h(t)\exp\left(t^\beta H(t)\right)\\
&+\frac{\alpha\beta\varepsilon' a^\alpha}{\Gamma(\alpha)}\exp\left(t^\beta H(t)\right)\exp\left(-t^\beta H(t)\right),\\
&\leq\varepsilon'\left(\begin{array}{c}(1-\alpha)\beta a^\alpha h\left(t_1\right)\\+\frac{\alpha\beta a^\alpha}{\Gamma(\alpha)}\end{array}\right),\\
&\leq\frac{\varepsilon}{\left((1-\alpha)\beta a^\alpha h\left(t_1\right)+\frac{\alpha\beta a^\alpha}{\Gamma(\alpha)}\right)}\left(\begin{array}{c}(1-\alpha)\beta a^\alpha h\left(t_1\right)\\+\frac{\alpha\beta a^\alpha}{\Gamma(\alpha)}\end{array}\right),\\
&\leq\varepsilon.
\end{aligned} \tag{107}$$

Therefore

$$\lim_{t\to t_0^+}\exp\left(-t^\beta H(t)\right)\Omega(t)=0. \tag{108}$$

Thus

$$\exp\left(-t^\beta H(t)\right)\Omega(t)\leq 0 \text{ for } t>t_0, \tag{109}$$

which implies

$$\begin{aligned}
&(1-\alpha)\beta t^{\beta-1}h(t)\Phi(t)\\
&+\frac{\alpha\beta}{\Gamma(\alpha)}\int_{t_0}^{t}\tau^{\beta-1}\left(t-\tau\right)^{\alpha-1}h(\tau)\Phi(\tau)d\tau,\\
&\leq 0.
\end{aligned} \tag{110}$$

contradiction, thus

$$\Phi(t)=0. \tag{111}$$

□

Theorem 7. *Let $f(t,y(t))$ be as presented before and $\exp\left(t^\beta H(t)\right)$ as $t\to t_0^+$ uniformly with respect to y, $\overline{y}\in[-\delta,\delta]$, $\delta>0$ arbitrary $h(t)$ and $H(t)$ are the same as previously. Then equation (92) has a unique solution.*

Proof. Let $y(t)$ and $\overline{y}(t)$ be solutions of equation (92).

$$\begin{aligned}
&|y(t)-\overline{y}(t)|\leq(1-\alpha)\beta t^{\beta-1}\left|f\left(t,y(t)\right)-f\left(t,\overline{y}(t)\right)\right|\\
&+\frac{\alpha\beta}{\Gamma\left(\alpha\right)}\int_{t_0}^{t}\tau^{\beta-1}\left(t-\tau\right)^{\alpha-1}\left|f\left(\tau,y(\tau)\right)-f\left(\tau,\overline{y}(\tau)\right)\right|d\tau,\\
&\leq(1-\alpha)\beta t^{\beta-1}\left|f\left(t,y(t)\right)-f\left(t,\overline{y}(t)\right)\right|\\
&+\frac{\beta}{\Gamma\left(\alpha\right)}\int_{t_0}^{t}\tau^{\beta-1}\left(t-\tau\right)^{\alpha-1}h\left(\tau\right)\left|y-\overline{y}\right|d\tau,\\
&\leq(1-\alpha)\beta t^{\beta-1}\left|f\left(t,y(t)\right)-f\left(t,\overline{y}(t)\right)\right|\\
&+\frac{\beta}{\Gamma\left(\alpha\right)}\int_{t_0}^{t}\tau^{\beta-1}\left(t-\tau\right)^{\alpha}h\left(\tau\right)\left|y-\overline{y}\right|d\tau.
\end{aligned} \tag{112}$$

In the view of (i), we get

$$\begin{aligned}
&|y(t)-\overline{y}(t)| \leq (1-\alpha)\beta t^{\beta-1}h(t)\,|y(t)-\overline{y}(t)| \\
&+\frac{\alpha\beta}{\Gamma(\alpha)}\int_{t_0}^{t}\tau^{\beta-1}(t-\tau)^{\alpha-1}h(\tau)\,|y(\tau)-\overline{y}(\tau)|\,d\tau, \\
&\leq (1-\alpha)\beta t^{\beta-1}h(t)\,|y(t)-\overline{y}(t)| \\
&+\frac{\alpha\beta}{\Gamma(\alpha)}\int_{t_0}^{t}\tau^{\beta}(t-\tau)^{\alpha}h(\tau)\,|y(\tau)-\overline{y}(\tau)|\,d\tau, \\
&\leq (1-\alpha)\beta\varepsilon' h(t)t^{\beta}\exp\left(t^{\beta}H(t)\right) \\
&+\frac{\alpha\beta\varepsilon'}{\Gamma(\alpha)}\int_{t_0}^{t}a^{\alpha}\tau^{\beta}h(\tau)\exp\left(\tau^{\beta}H(\tau)\right)d\tau.
\end{aligned} \tag{113}$$

$$\begin{aligned}
&|y(t)-\overline{y}(t)| \leq (1-\alpha)\beta\varepsilon' h(t)a^{\beta}\exp\left(t^{\beta}H(t)\right) \\
&+\frac{\alpha\beta\varepsilon' a^{\alpha}}{\Gamma(\alpha)}\int_{t_0}^{t}\left(\tau^{\beta}h(\tau)+\beta\tau^{\beta-1}H(\tau)\right)\exp\left(\tau^{\beta}H(\tau)\right)d\tau, \\
&\leq \left((1-\alpha)\beta\varepsilon' h(t)a^{\beta}+\frac{\alpha\beta\varepsilon' a^{\alpha}}{\Gamma(\alpha)}\right)\exp\left(t^{\beta}H(t)\right), \\
&\leq \left((1-\alpha)\beta\varepsilon' h(t_1)a^{\beta}+\frac{\alpha\beta\varepsilon' a^{\alpha}}{\Gamma(\alpha)}\right)\exp\left(t^{\beta}H(t)\right), \\
&\leq \frac{\varepsilon\left((1-\alpha)\beta h(t_1)a^{\beta}+\frac{\alpha\beta a^{\alpha}}{\Gamma(\alpha)}\right)}{\left((1-\alpha)\beta h(t_1)a^{\beta}+\frac{\alpha\beta a^{\alpha}}{\Gamma(\alpha)}\right)}\exp\left(t^{\beta}H(t)\right), \\
&\leq \varepsilon\exp\left(t^{\beta}H(t)\right).
\end{aligned} \tag{114}$$

From the above Lemma

$$y(t)-\overline{y}(t)=0. \tag{115}$$

□

Corollary 2. *Let the condition in above theorem hold, then*

i) $(f(t,y)-f(t,\overline{y}))(y-\overline{y}) \leq h(t)(y-\overline{y})^{\overline{\beta}}$, $\forall(t,\overline{y}),(t,y)\in S_{+}$, $\beta\in(1,2]$.
ii) $h(t_1)=\max_{t\in(t_0,a)}h(t)$,

$$1-\overline{\beta}(1-\alpha)\beta(\overline{t}_0)^{\beta-1}\Lambda>0,$$

iii) ${}_{t_0}^{FFM}J_t^{\alpha,\beta}h(t)$ *exists.*

Proof. Let $y(t)$ and $\overline{y}(t)$ be two solutions of equation (92) then, let set $\overline{\Omega}(t_0)=0$ then we get

$$\begin{aligned}
{}_{t_0}^{FFM}D_t^{\alpha,\beta}\overline{\Omega}(t) &= \frac{1}{\beta t^{\beta-1}}\,{}_{t_0}^{ABC}D_t^{\alpha}\overline{\Omega}(t), \\
&= \frac{1}{\beta t^{\beta-1}(1-\alpha)}\int_{t_0}^{t}E_{\alpha}\left(-\frac{\alpha}{1-\alpha}(t-\tau)^{\alpha}\right)\overline{\Omega}'d\tau, \\
&= \frac{1}{\beta t^{\beta-1}(1-\alpha)}\int_{t_0}^{t}E_{\alpha}\left(-\frac{\alpha}{1-\alpha}(t-\tau)^{\alpha}\right) \\
&\times\left[\overline{\beta}(y-\overline{y})^{\overline{\beta}}(y-\overline{y})'\right]d\tau.
\end{aligned} \tag{116}$$

Let $m=\min_{\overline{t}\in[t_0,t]}|y-\overline{y}|^{\beta-1}$, $M=\max_{\overline{t}\in[t_0,t]}|y(\overline{t})-\overline{y}(\overline{t})|$. We define

$$\Lambda=\begin{cases} m, & \text{if } y'-\overline{y}'<0, \\ M, & \text{if } y'-\overline{y}'>0. \end{cases} \tag{117}$$

Therefore

$$\begin{aligned}
{}_{t_0}^{FFM}D_t^{\alpha,\beta}\overline{\Omega}(t) &\leq \frac{\overline{\beta}\Lambda}{\beta t^{\beta-1}(1-\alpha)}\int_{t_0}^{t}E_{\alpha}\left(-\frac{\alpha}{1-\alpha}(t-\tau)^{\alpha}\right)(y-\overline{y})'\,d\tau, \\
&\leq \overline{\beta}\Lambda\left({}_{t_0}^{FFM}D_t^{\alpha,\beta}y(t)-{}_{t_0}^{FFM}D_t^{\alpha,\beta}\overline{y}(t)\right), \\
&\leq \overline{\beta}\Lambda\left(f(t,y)-f(t,\overline{y})\right), \\
&\leq \overline{\beta}\Lambda h(t)(y-\overline{y})^{\overline{\beta}}.
\end{aligned} \tag{118}$$

$$\begin{aligned}
{}_{t_0}^{RL}D_t^{\alpha}\overline{\Omega}(t) &= \overline{\beta}\Lambda\beta t^{\beta-1}, \\
&\leq \overline{\beta}\Lambda h(t)\overline{\Omega}(t), \\
{}_{t_0}^{ABR}D_t^{\alpha}\overline{\Omega}(t) &\leq \beta t^{\beta-1}\Lambda h(t)\overline{\Omega}(t)\overline{\beta}.
\end{aligned} \tag{119}$$

$$\begin{aligned}
\overline{\Omega}(t) &\leq \overline{\beta}(1-\alpha)\beta t^{\beta-1}\Lambda h(t)\overline{\Omega}(t) \\
&+\frac{\alpha\beta\overline{\beta}}{\Gamma(\alpha)}\int_{t_0}^{t}\tau^{\beta-1}(t-\tau)^{\alpha-1}\Lambda h(\tau)\overline{\Omega}(\tau)d\tau, \\
&\leq \overline{\beta}(1-\alpha)\beta(\overline{t}_0)^{\beta-1}\Lambda h(t_1)\overline{\Omega}(t) \\
&+\frac{\Lambda\alpha\beta\overline{\beta}}{\Gamma(\alpha)}\int_{t_0}^{t}\tau^{\beta-1}(t-\tau)^{\alpha-1}h(\tau)\overline{\Omega}(\tau)d\tau.
\end{aligned} \tag{120}$$

If we take as

$$A=\frac{\Lambda\overline{\beta}}{1-\overline{\beta}(1-\alpha)\beta(\overline{t}_0)^{\beta-1}\Lambda}, \tag{121}$$

$$\overline{\Omega}(t) \leq \frac{A\alpha\beta}{\Gamma(\alpha)}\int_{t_0}^{t}\tau^{\beta-1}(t-\tau)^{\alpha-1}h(\tau)\overline{\Omega}(\tau)d\tau. \tag{122}$$

By the Gronwall inequality we set

$$\overline{\Omega}(t) \leq o \exp\left(A_{t_0}^{FFP} J_t^{\alpha,\beta} h(t)\right), \qquad (123)$$
$$\overline{\Omega}(t) \leq 0.$$

Therefore we have $\overline{\Omega}(t) \leq 0$ which is a contraction therefore

$$\overline{\Omega}(t) = 0 \Rightarrow y(t) = \overline{y}(t), \ \forall t \in [t_0, a]. \qquad (124)$$

□

7. Conclusion

Witte provided a set of conditions under which a given nonlinear ordinary differential equation admits unique solutions. This was established when the differential operator was in integer order. Based on the framework of Witte, we have presented a detailed analysis of the uniqueness of nonlinear ordinary differential equations with fractal-fractional derivatives.

References

[1] Zill, D.G. (2012). *A First Course in Differential Equations with Modeling Applications*. Cengage Learning. ISBN 978-1-285-40110-2.

[2] Boyce, W. E., DiPrima, R. C. & Meade, D. B. (2017). *Elementary Differential Equations*. John Wiley & Sons.

[3] Zaitsev, V. F. & Polyanin, A. D. (2002). *Handbook of Exact Solutions for Ordinary Differential Equations*. Chapman and Hall/CRC. https://doi.org/10.1201/9781420035339

[4] Sher, M., Khan, A., Shah, K. & Abdeljawad, T. (2023). Existence and stability theory of pantograph conformable fractional differential problem. *Thermal Science*, 27(Spec. issue 1), 237-244. https://doi.org/10.2298/TSCI23S1237S

[5] Shah, K., Abdeljawad, T. & Abdalla, B. (2023). On a coupled system under coupled integral boundary conditions involving non-singular differential operator. *AIMS Mathematics*, 8(4), 9890-9910. https://doi.org/10.3934/math.2023500

[6] Koca, İ. & Atangana, A. (2023). Theoretical and numerical analysis of a chaotic model with nonlocal and stochastic differential operators. *An International Journal of Optimization and Control: Theories & Applications (IJOCTA)*, 13(2), 181–192. https://doi.org/10.11121/ijocta.2023.1398

[7] Sene, N. & Ndiaye, A. . (2024). Existence and uniqueness study for partial neutral functional fractional differential equation under Caputo derivative. *An International Journal of Optimization and Control: Theories and Applications (IJOCTA)*, 14(3), 208–219. https://doi.org/10.11121/ijocta.1464

[8] Teschl, G. (2024). *Ordinary Differential Equations and Dynamical Systems* (Vol. 140). American Mathematical Society.

[9] Agarwal, R. P., Agarwal, R. P., & Lakshmikantham, V. (1993). *Uniqueness and Nonuniqueness Criteria for Ordinary Differential Equations*. (Vol. 6). World Scientific. https://doi.org/10.1142/1988

[10] Atangana, A. (2017). Fractal-fractional differentiation and integration: connecting fractal calculus and fractional calculus to predict complex system. *Chaos, Solitons & Fractals*, 102, 396-406. https://doi.org/10.1016/j.chaos.2017.04.027

[11] Chen, W. (2006). Time-space fabric underlying anomalous diffusion. *Chaos, Solitons & Fractals*, 28(4), 923-929. https://doi.org/10.1016/j.chaos.2005.08.199

[12] Metzler, R. & Klafter, J. (2000). The random walk's guide to anomalous diffusion: a fractional dynamics approach. *Physics Reports*, 339(1), 1-77. https://doi.org/10.1016/S0370-1573(00)00070-3

Abdon Atangana *is a professor at the University of the Free State, Bloemfontein, South Africa.*
https://orcid.org/0000-0002-1886-3125

Ilknur Koca *received the B.Sc. and M.Sc. degrees in 2007 and 2009 from the Department of Mathematics from Ankara University, Turkey respectively. She received her Ph.D. degree from the same university in 2013. She is working as an Professor at Department of Economics and Finance, Fethiye Business Faculty, Mugla Sıtkı Kocman University, Mugla, Türkiye. Her research interests are methods and applications of partial and ordinary differential equations,fractional differential equations, iterative methods.*
https://orcid.org/0000-0003-4393-1588

This work is licensed under a Creative Commons Attribution 4.0 International License. The authors retain ownership of the copyright for their article, but they allow anyone to download, reuse, reprint, modify, distribute, and/or copy articles in IJOCTA, so long as the original authors and source are credited. To see the complete license contents, please visit http://creativecommons.org/licenses/by/4.0/.

An International Journal of Optimization and Control: Theories & Applications
ISSN:2146-0957 eISSN:2146-5703
Vol.14, No.4, pp.336-345 (2024)
http://doi.org/10.11121/ijocta.1521

RESEARCH ARTICLE

Influence of rotation on peristaltic flow for pseudoplastic fluid: a wavy channel

Hayat Adel Ali [a*], Mohammed R. Salman [b]

[a] *Department of Applied Science, University of Technology-Iraq, Baghdad, Iraq*
[b] *Department of Mathematics, University of Kerbala, Kerbala, Iraq*
hayat.A.Ali@uotechnology.edu.iq, mohammed.razzaq@uokerbala.edu.iq

ARTICLE INFO

Article History:
Received 4 January 2024
Accepted 9 June 2024
Available Online 10 October 2024

Keywords:
Peristaltic transport
Pseudoplastic fluid
Rotation force
Wavy channel

AMS Classification 2010:
76D05; 76D07; 76N10

ABSTRACT

The phenomenon of rotation serves multiple purposes in cosmic and geophysical phenomena. It offers insights into the formation of galaxies and the circulation patterns of oceans. Moreover, rotational diffusion elucidates the orientation of nanoparticles within fluid mediums. Investigating the dynamics of fluid peristalsis under the influence of rotational forces holds significant relevance in addressing challenges associated with the transportation of conductive physiological fluids such as blood, polymeric materials, and saline water. This study focused on studying the impact of rotation on the peristaltic transport of non-Newtonian pseudoplastic fluids through a wavy channel. The complexity of flow equations, including the continuity and motion equations, is mathematically formulated and transformed into dimensionless nonlinear ordinary differential equations depending on the assumption of low Reynolds number and long wavelength approximation. Perturbation technique is employed to solve the problem for the stream function and the resulted system is implemented and plotted using MATHEMATICA software along with the boundary conditions. Graphical discussion is involved to utilize the impact of the emerging parameters in the flow characteristic, encompassing the velocity profile, pressure gradient, pressure rise, and trapping phenomenon. The research revealed that rotation significantly influences the fluid flow within the channel, diminishing the regressive and inhibitory impact of the fluid parameter, consequently enhancing the fluid flow within the channel.

CC BY

1. Introduction

Extensive research has been conducted on peristaltic motion in recent years due to the fact that it involves the study of wave-like motion in physiological fluids resulting from interaction with surrounding boundaries. Such phenomena are evident during the process of food ingestion through the esophagus, the propagation of lymphocytes within the lymphatic system, the circulation of blood through vessels, the movement of urine toward the bladder, and numerous other instances that collectively contribute to our understanding of peristalsis. Moreover, peristaltic transport has wide applications in medical engineering, science, and modern industry, such as aggressive chemicals, high solid slurries, noxious fluids (nuclear industries), heart-lung machines, blood pump machines, and dialysis machines [1–4]. The initial effort to elucidate this phenomenon was attributed to [5]. Subsequent to this progress, numerous studies, delving into the exploration of peristaltic flow of various fluid types under diverse influencing factors, were illustrated by many researchers. [6] discussed the impact of long wavelengths and the low Reynolds number assumption on peristaltic pumping. [7] determined the impact

*Corresponding Author

of the elastic wall of a hollow cylinder's channel of Jeffrey's fluid by peristaltic flow. [8] studied the heat transfer analysis of magnetohydrodynamic (MHD) peristaltic transport of Jeffrey fluid in an inclined tapered asymmetric channel. For more information, see [9–13].

Non-Newtonian fluids, including molten plastics, artificial fibers, polymeric materials, foodstuffs, blood, slurries, and synovial liquids, exhibit shear-stress-strain relationships that diverge significantly from the traditional viscous model, finding numerous applications in manufacturing and commerce [14–17]. Significant literature exists on the study of peristaltic motion in the presence of non-Newtonian fluids. Many of these types of fluids exhibit characteristics of shear-thinning yield stress materials [18], such as pseudoplastic fluid which is found in blood plasma, latex paint, polymer solutions, and similar solutions of high molecular weight substances. At low shear rates, these fluids experience the formation of shear stress that results in the reordering of the molecules to reduce the overall stress. [9] analyzed the impact of Soret and Dufour on the peristaltic flow of magnetohydrodynamic (MHD) pseudoplastic nanofluid in a tapered asymmetric channel. The impact of pseudoplasticity and dilatancy of fluid on the peristaltic flow of non-Newtonian fluid in a non-uniform asymmetric channel was investigated [19]. In 2014, [20] studied the impact of wall properties and slip conditions on the peristaltic flow of pseudoplastic fluid in a curved channel. An effect of magnetohydrodynamic (MHD) and thermal radiation on the peristaltic flow of a pseudoplastic nanofluid through a porous medium asymmetric canal with convection boundary conditions was depicted by [21].

A rotational phenomenon plays a pivotal role in various cosmic and geophysical phenomena. It aids in comprehending the emergence of galaxies and the circulation of oceans. Nanoparticles' orientation in fluids is attributed to rotational diffusion. Moreover, rotation is noteworthy in specific flow scenarios within physiological fluids like saline water and blood. This synergy of rotation facilitates the movement of biological fluids within the intestines, ureters, and arterioles. Several studies focused on the rotational system's impact on the peristaltic flow of various fluids. [22] analyzed the flow of non-Newtonian fluid with a porous medium under the effect of rotation and magnetic force. [23] was concerned with the peristaltic flow of a Jeffrey fluid in an asymmetric rotating channel. [24] illustrated the influence of magnetic force, rotation, and nonlinear heat radiation on the peristaltic transport of hybrid bio-nanofluids through a symmetric channel. [25] investigated the peristaltic flow of Bingham plastic fluid under the effect of rotation and induced magnetic field. For more information, see Refs. [26–30].

Lately, there has been a lack of attention given to studying peristaltic flows under conditions where both the fluid and the channel experience solid body rotation, i.e., the entire system, comprising both the fluid and the channel, is situated within a rotating frame characterized by a consistent angular velocity. The aforementioned studies have primarily focused on peristaltic flows involving different non-Newtonian fluids, addressing rotational effects. Nevertheless, there remains a gap in the previous literature concerning the impact of the rotation frame on the peristaltic transport of pseudoplastic fluids. In this study, We extended the inquiry delineated in [19] by elucidating the physical alterations observed in our fluid during the flow, which experiences a reduction in viscosity as the shear rate increases during rotation. As a consequence, its velocity increases, which finds application in various contexts such as in blood apheresis machine and water treatment. In this article, physical modeling governing the equation of peristaltic flow of pseudoplastic under the effect of rotation is described and reduced to the differential equation by using long wavelength and low Reynolds number assumptions. The closed-form analytic solution for the stream function and fluid velocity is provided. Subsequently, a graphical analysis is conducted using codes from the Mathematica package to illustrate the impact of key parameters on flow characteristics.

2. Mathematical model of the problem

Assuming the incompressible peristaltic motion of a non-Newtonian, electrically conducting, pseudoplastic fluid through a two-dimensional asymmetric a wavy channel of width 2 d in which the channel rotates about the horizontal with uniform angular velocity Ω see Fig. 1. The induced sinusoidal propagation waves of wavelength λ are advancing with a constant speed c in the $\overline{X}$ - axis, and the $\overline{Y}$-axis is normal to it.

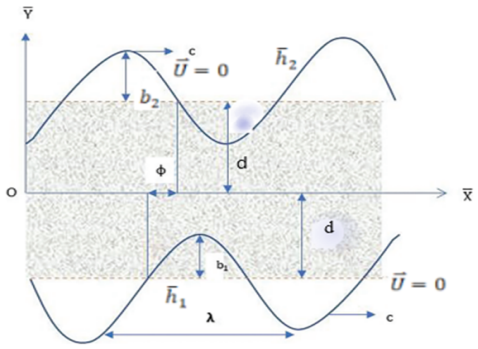


Figure 1. Geometry of problem

The mathematical equations for the channel walls are provided as follows [25]:

$$\bar{Y} = \bar{h}_1(\bar{X},\bar{t}) = d + b_1 \,\mathrm{Sin}\left(\frac{2\pi}{\lambda}(\bar{X} - c\bar{t})\right), \quad (1)$$

$$\bar{Y} = \bar{h}_2(\bar{X},\bar{t}) = -d - b_2 \,\mathrm{Sin}\left(\frac{2\pi}{\lambda}(\bar{X} - c\bar{t}) + \phi\right). \quad (2)$$

where $\bar{h}_1$ and $\bar{h}_2$ are the lower and upper walls respectively, b_1 and b_2 denote the waves' amplitudes, $\bar{t}$ stands for time, $\phi \in [0,\pi]$ and represents the phase difference. When $\phi = 0$, it indicates waves out of phase for a symmetric channel, and when $\phi = \pi$ the waves are in phase. Additionally, the values of d, b_1, b_2 and ϕ satisfy the inequity.

$$b_1{}^2 + b_2{}^2 + b_1 b_2 \,\mathrm{Cos}\,\phi \le (d)^2, \quad (3)$$

The governing equations for an incompressible fluid in the fixed frame are formulated as follows: The continuity equation is:

$$\nabla.\vec{V} = 0, \quad (4)$$

The motion equation is:

$$\rho\left(\frac{\partial \vec{V}}{\partial \bar{t}} + (\vec{V}\cdot\nabla)\vec{V}\right) + \rho[\Omega \times (\Omega \times \vec{V}) + 2\Omega \times \vec{V}] = -\nabla\bar{P} + \nabla\cdot\bar{S}, \quad (5)$$

Associated with the no- slip boundary condition bellows:

$$\vec{U} = 0 \text{ at } \bar{h}_1 \text{ and } \bar{h}_2. \quad (6)$$

In which $\vec{V} = (\bar{U},\bar{V})$ is the fluid velocity vector in $\bar{X}$ and $\bar{Y}$ coordinates respectively, $\rho, \bar{P}$ are the fluid density and the pressure, $\vec{\nabla} = \left(\frac{\partial}{\partial \bar{X}}, \frac{\partial}{\partial \bar{Y}}\right)$ is the gradient vector, $\rho(\Omega \times (\Omega \times \vec{V}))$ denotes the centrifugal force while the term $\rho(2\Omega \times \vec{V})$ refers to the Coriolis force, $\bar{S}$ is the Cauchy stress tensor for pseudoplastic fluid which defined as [9,19]:

$$\bar{S} + \lambda^*{}_1\left(\frac{\partial \bar{S}}{\partial \bar{t}} - \nabla\vec{V}\cdot\bar{S} - \bar{S}\cdot(\nabla\vec{V})^T\right) + \frac{1}{2}(\lambda^*{}_1 - \mu^*{}_1)\left(\breve{A}_1\bar{S} + \bar{S}\breve{A}_1\right) = \mu\breve{A}_1, \quad (7)$$

$$\breve{A}_1 = \nabla\vec{V} + (\nabla\vec{V})^T. \quad (8)$$

where μ is the fluid viscosity, $\lambda_1^*, \mu^*{}_1$ are the relaxation times, $\breve{A}_1$ is the first Rivlin- Ericksen tensor. Consider the wave frame $(\bar{x},\bar{y})$ traveling with speed c away from the laboratory frame. The transformation of coordinates and flow properties between fixed and wave frame is given by:

$$\bar{x} = \bar{X} - c\bar{t}, \bar{y} = \bar{Y}, \bar{u} = \bar{U} - c, \bar{v} = \bar{V}, \bar{p}(\bar{x}) = \bar{p}(\bar{X},\bar{t}). \quad (9)$$

Now, defining the dimensionless parameters and variables in the following manner.

$$x = \frac{\bar{x}}{\lambda}, y = \frac{\bar{y}}{d}, u = \frac{\bar{u}}{c}, v = \frac{\bar{v}}{c}, h_1 = \frac{\bar{H}_1(X)}{d},$$
$$h_2 = \frac{\bar{H}_2(X)}{d}, d = \frac{d_2}{d_1}, \delta = \frac{d}{\lambda}, a = \frac{b_1}{d}, b = \frac{b_2}{d},$$
$$\lambda_1 = \frac{c\lambda_1^*}{d}, p = \frac{d^2\bar{p}}{\lambda\mu c}, Re = \frac{\rho c d}{\mu}, s_{ij} = \frac{d}{\mu c}\bar{S}_{ij},$$
$$\mu_1 = \frac{c\mu_1^*}{d}, \Omega = \frac{{b_1}^2\bar{\Omega}}{\mu}, Ta = \frac{\Omega \mathrm{d} Re}{c}. \quad (10)$$

where Re and δ are the Reynolds number and the dimensionless number of waves respectively. Introducing the dimensionless stream function $\psi(x,y)$, in which,

$$u = \frac{\partial\psi}{\partial y}, v = -\delta\frac{\partial\psi}{\partial x}. \quad (11)$$

The continuity equation is identically achieved. By substituting Eq. (9) and Eq. (10) into Eqs. (2 - 7), we get:

$$\mathrm{Re}\,\delta\left((u+1)\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y}\right) - 2\,\mathrm{Ta}(u+1) = -\frac{\partial p}{\partial x} + \delta\frac{\partial s_{xx}}{\partial x} + \frac{\partial s_{xy}}{\partial y}, \quad (12)$$

$$\mathrm{Re}\,\delta^2\left((u+1)\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y}\right) - 2\,\mathrm{Ta}\,\delta v = -\frac{\partial p}{\partial y} + \delta\frac{\partial s_{yy}}{\partial y} + \delta^2\frac{\partial s_{yx}}{\partial x} - \frac{1}{\kappa}\delta^2 v, \quad (13)$$

$$s_{xx}+\lambda_1\left(\delta\left(u\frac{\partial s_{xx}}{\partial x}+v\frac{\partial s_{xx}}{\partial y}\right)-2\delta s_{xx}\frac{\partial u}{\partial x}-2s_{xy}\frac{\partial u}{\partial y}\right)+\frac{1}{2}\left(\lambda_1-\mu_1\right)\left(2s_{xy}\left(\frac{\partial u}{\partial y}+\delta^2\frac{\partial v}{\partial x}\right)+4\delta s_{xx}\frac{\partial u}{\partial x}\right)=2\delta\frac{\partial u}{\partial y}, \tag{14}$$

$$s_{xy}+\lambda_1(\delta\left(u\frac{\partial s_{xy}}{\partial x}+v\frac{\partial s_{xy}}{\partial y}\right)-\delta^2 s_{xx}\frac{\partial v}{\partial x}-s_{yy}\frac{\partial u}{\partial y}\right)+\frac{1}{2}\left(\lambda_1-\mu_1\right)\left(s_{xx}+s_{yy}\right)\left(\frac{\partial u}{\partial y}+\delta^2\frac{\partial v}{\partial x}\right)=\frac{\partial u}{\partial y}+\delta^2\frac{\partial v}{\partial x^2}, \tag{15}$$

$$s_{yy}+\lambda_1\left(\delta\left(u\frac{\partial s_{yy}}{\partial x}+v\frac{\partial s_{yy}}{\partial y}\right)-2\delta s_{xx}\frac{\partial v}{\partial y}+2\delta^2 s_{xy}\frac{\partial v}{\partial x}\right)+\frac{1}{2}\left(\lambda_1-\mu_1\right)\left(2s_{xy}\left(\frac{\partial u}{\partial y}+\delta^2\frac{\partial v}{\partial y}\right)+4\delta s_{yy}\frac{\partial v}{\partial y}\right)=2\delta\frac{\partial v}{\partial y}. \tag{16}$$

Employing Eq. (11) and assuming a low Reynolds number and a large wavelength approximation ($\delta \ll 1$), Eqs. (12 - 16) are reduced as follows:

$$\frac{\partial p}{\partial x}=\frac{\psi_{yy}}{\left(1+\xi\left(\psi_{yy}\right)^2\right)}+2Ta\left(\frac{\partial \psi}{\partial y}+1\right), \tag{17}$$

$$\frac{\partial p}{\partial y}=0. \tag{18}$$

Through Eq. (18), we conclude that pressure is not a function of y, i.e., $P \neq P(y)$. Neglecting pressure from Eq. (17), we get:

$$\psi_{yyyy}+3\xi\left(\psi_{yy}\right)^2\psi_{yyyy}+6\xi\left(\psi_{yyy}\right)^2\psi_{yy}+2\,\mathrm{Ta}\,\psi_{yy}=0, \tag{19}$$

$$s_{xx}=\left(\lambda_1+\mu_1\right)\frac{\left(\psi_{yy}\right)^2}{\left(1+\xi\left(\psi_{yy}\right)^2\right)^2}, \tag{20}$$

$$s_{xy}=\frac{\psi_{yy}}{\left(1+\xi\left(\psi_{yy}\right)^2\right)}, \tag{21}$$

$$s_{yy}=-\left(\lambda_1-\mu_1\right)\frac{\left(\psi_{yy}\right)^2}{\left(1+\xi\left(\psi_{yy}\right)^2\right)^2}. \tag{22}$$

Linked to the subsequent dimensionless boundary condition.

$$\psi=-\frac{F}{2},\psi_y=0 \text{ at } y=h_1(x), \tag{23}$$

$$\psi=\frac{F}{2},\psi_y=0 \text{ at } y=h_2(x). \tag{24}$$

where $h_1(x)=1+a\,\mathrm{Sin}(2\pi x)$ and $h_2(x)=1+b\,\mathrm{Sin}(2\pi x+\Phi)$.
where $\xi=(\lambda_1^2-\mu_1^2)$ is the Pseudoplastic fluid parameter.
The parameter F refers to the dimensionless mean flows and it is given by:

$$F=\int_{h_1}^{h_2}u(x,y)dy=\int_{h_1}^{h_2}\frac{\partial \psi}{\partial y}dy=\psi\left(h_2\right)-\psi\left(h_1\right), \tag{25}$$

Additionally, the connection between F and the nondimensional mean flows in the moving frame, θ, can be derived as:

$$F=\theta+a\,\mathrm{Sin}(2\pi x+\Phi)+b\,\mathrm{Sin}(2\pi x). \tag{26}$$

The pressure rise per unit wavelength is:

$$\Delta p=\int_0^1\frac{dp}{dx}dx. \tag{27}$$

3. Solution of the problem

Eq. (19) poses a complex nonlinear differential challenge, rendering an exact solution unattainable. In the field of fluid science, various techniques [31–37] are available to find the desired solution. Therefore, we turn to perturbation technique to calculate series solutions when dealing with a small parameter. Consequently, we applied perturbation to stream function ψ and dimensionless mean flow F, considering them up to the first order with respect to the Pseudoplastic fluid parameter ξ as:

$$\psi=\psi_0+\xi\psi_1+O\left(\xi^2\right), \tag{28}$$

$$F=F_0+\xi F_1+O\left(\xi^2\right). \tag{29}$$

3.1. Zeroth order system

$$\psi_{0yyyy}+2Ta\psi_{0yy}=0, \tag{30}$$

With the boundary conditions.

$$\psi_0=-\frac{F_0}{2},\psi_{0y}=0, \text{ at } y=h_1(x), \tag{31}$$

$$\psi_0=\frac{F_0}{2},\psi_{0y}=0 \text{ at } y=h_2(x). \tag{32}$$

where

$$F_0 = \theta + a\,\mathrm{Sin}(2\pi x + \Phi) + b\,\mathrm{Sin}(2\pi x). \quad (33)$$

3.2. First order system

$$\psi_{1yyyy} + 3\left(\psi_{0yy}\right)^2 \psi_{0yyyy} + 6\left(\psi_{0yyy}\right)^2 \psi_{0yy} + 2Ta\psi_{1yy} = 0, \quad (34)$$

With the following boundary conditions.

$$\psi_1 = -\frac{F_1}{2}, \psi_{1y} = 0, \text{ at } y = h_1(x), \quad (35)$$

$$\psi_1 = \frac{F_1}{2}, \psi_{1y} = 0, \text{ at } y = h_2(x). \quad (36)$$

Solving the two resulting systems by writing suitable codes in Mathematica software, the explicit expression of stream functions ψ_0 and ψ_1 obtained.

$$\psi_0(x,y) = \frac{e^{-\sqrt{A}y}\left(e^{2\sqrt{A}y}\mathrm{c}_1 + \mathrm{c}_2\right)}{A} + \mathrm{c}_3 + y\mathrm{c}_4, \quad (37)$$

$$\psi_1(x,y) = \mathrm{c}_7 + y\mathrm{c}_8 + \frac{1}{8A}e^{-3\sqrt{A}y}\left(-{\mathrm{c}_2}^3 + 30\mathrm{c}_1{\mathrm{c}_2}^2 e^{2\sqrt{A}y} + 30{\mathrm{c}_1}^2\mathrm{c}_2 e^{4\sqrt{A}y} - {\mathrm{c}_1}^3 e^{6\sqrt{A}y} + 8e^{4\sqrt{A}y}\mathrm{c}_5 + 8e^{2\sqrt{A}y}\mathrm{c}_6 - 6\mathrm{c}_1\mathrm{c}_2 e^{2\sqrt{A}y}\left(-\mathrm{c}_2 + \mathrm{c}_1 e^{2\sqrt{A}y}\right)\log\left(e^{2\sqrt{A}y}\right)\right). \quad (38)$$

where $A = -2\mathrm{Ta}$, and the coefficients $c_1, c_2, c_3, c_4, c_5, c_6, c_7$ and c_8 consist of complex expressions that will not be detailed here.

4. Results analysis and discussion

In this section, we examine the outcomes of different physical parameters by utilizing the visual representations provided. The analysis includes the variation in the velocity profile, the gradient of pressure, the pressure rise and the trapping phenomenon as a result of the increase in the values of the rotation parameter Ta, the pseudoplastic fluid parameter ξ, the non- dimensional mean flows θ, the phase difference parameter Φ, the lower wall amplitude parameter a, and upper wall amplitude parameter b.

4.1. Velocity profile

Figs. 2(a)-(d) illustrate a consistent pattern where the maximum velocity is consistently near the center of the channel, and all velocity profiles exhibit a parabolic shape. Figs. 2(a) and 2(b) depict a decrease in velocity profile at the central part of the channel whereas an increasing effect is noticed toward the boundaries, as well as inflection points appearing via ascending values of Ta and ξ. This outcome arises from the physical phenomenon wherein viscosity decreases as shear rate escalates during rotational motion. As a consequence, its velocity increases. It's worth noting from Fig. 2(c) that as θ enlarges, the fluid velocity reduces. However, an opposite reaction on velocity profile is observed from Fig. 2(d), that means as the phase difference between waves increases (Φ), the axial velocity increases across the entire range of the y-axis.

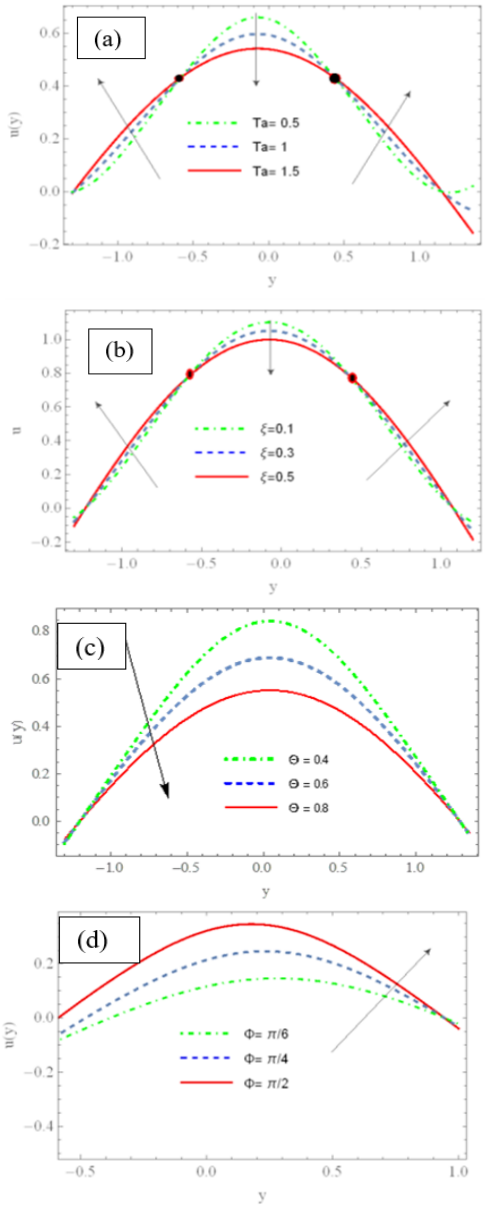
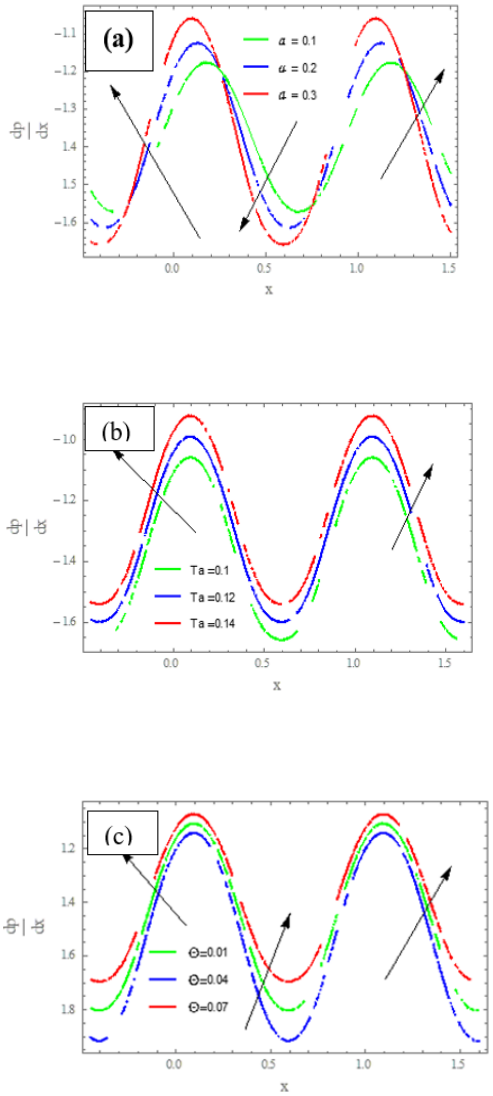


Figure 2. Velocity profile for ascending values of (a) rotation parameter (b) Pseudoplastic parameter (c) non- dimensional mean flows parameter (d) phase difference parameter and fixed $b = 0.3, a = 0.6, x = 0.5, t = 0.2$.

4.2. Gradient of pressure

Figs. 3(a)-(d) record the fluctuations in the axial pressure gradient (dp/dx), exhibiting sinusoidal behavior across the entire x-axis range when it is analyzed under the impact of increasing values of the lower wall wave amplitude a, the rotation parameter Ta, non- dimensional mean flows θ , and the Pseudoplastic fluid parameter ξ. It is noticed from Fig. 3(a) The observed elevation in the magnitude of the pressure gradient towards the central region of the channel compared to its boundaries is attributed to the augmentation of a, leading to an expansion in the dimensions of the channel wall. While from Figs. 3(b) and 3(c), we conclude that the rate of change for (dp/dx) with respect to Ta and θ means the flow can smoothly pass without requiring a significant pressure gradient. Fig. 3(d) illustrated that as ξ increases, a reversal in the situation is depicted, as this parameter inversely correlates with the velocity of the fluid. Consequently, a notable pressure gradient is necessitated to ensure fluid flow remains smooth.



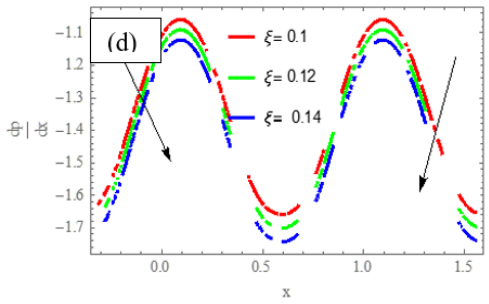
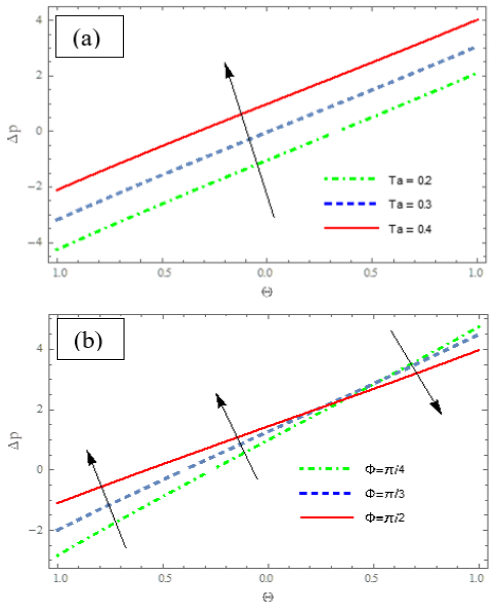


Figure 3. Pressure gradient profile for ascending values of (a) low wall amplitude parameter (b) rotation parameter (c) non- dimensional mean flows parameter (d) Pseudoplastic parameter and fixed $b = 0.3, a = 0.6, x = 0.5, t = 0.2$.

4.3. Pressure rise profile

Figs. 4(a)-(d) elucidate the behavior of nondimensional pressure rise Δp versus the dimensionless mean flow rate θ through plotting Δp profile with various values of the rotation parameter Ta, the phase difference parameter Φ, the lower wall amplitude a, and the upper wall amplitude b. It is clear from Fig. 4(a) that the pumping rate Δp is enlarged in the whole region as the value of Ta is increased. Whereas the impact of enhancing Φ and a on pumping rate anticipated in Figs. 4(b) and 4(c). The plots recorded a decay in peristaltic pumping region with $(\theta > 0, \Delta p > 0)$ while an increment impact is noticed in the retrograde pumping area $(\theta < 0, \Delta p > 0)$ and no flow area $(\theta < 0, \Delta p < 0)$. The opposite scene is depicted with an increasing b, that means the peristaltic pumping region is enhanced while the retrograde pumping and no flow areas are dampened, see Fig. 4(d).



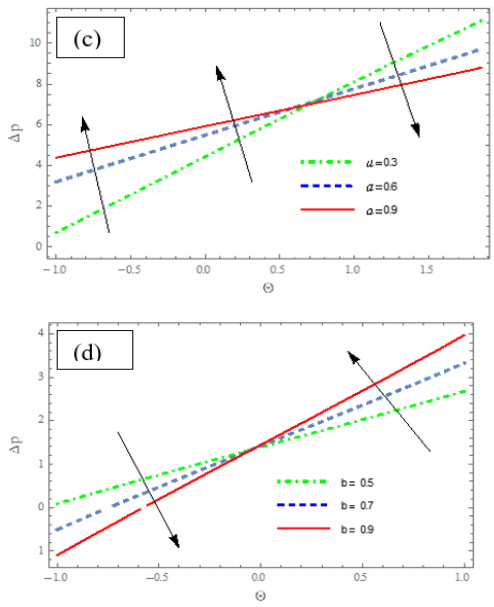


Figure 4. Pressure rise profile for ascending values of (a) rotation parameter (b) phase difference parameter (c) lower wall amplitude parameter (d) upper wall amplitude parameter (d) and fixed $\{\xi = 0.3, \theta = 0.6, x = 0.5, t = 0.2\}$.

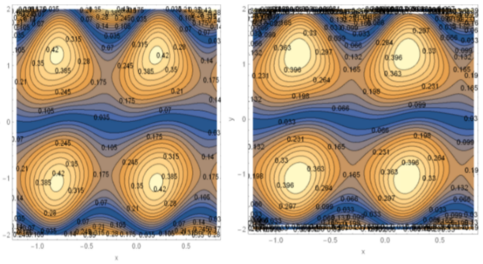


Figure 5. Stream lines for ascending values of rotation parameter $\{Ta = 0.1, Ta = 0.3\}$.

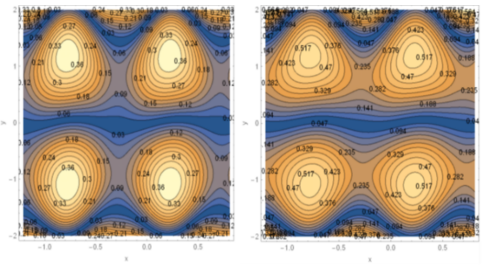


Figure 6. Stream lines for ascending values of rotation parameter $\{\theta = 0.2, \theta = 0.5\}$.

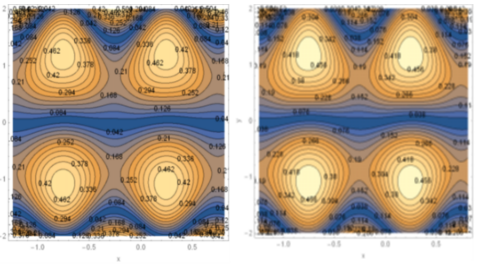


Figure 7. Stream lines for ascending values of pseudoplastic fluid parameter $\{\xi = 0.1, \xi = 0.3\}$.

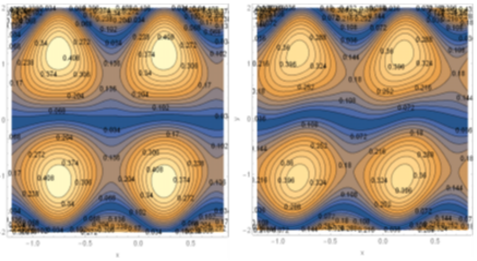


Figure 8. Stream lines for ascending values phase difference parameter $\{\Phi = \pi/6, \Phi = \pi/4\}$.

4.4. Trapping phenomenon

Streamlines depict the paths followed by fluid particles within a flow. The creation of an enclosed, circulating mass of fluid due to the closed streamlines is referred to as trapping phenomenon. Figs. 5-8 are sketched to elucidate the influence of the rotation parameter Ta, the non-dimensional mean flow parameter θ, the Pseudoplastic fluid parameter ξ, and the phase difference parameter Φ on the absolute value of stream function $|\psi|$. Moreover, we noticed from the graphs that the trapped bolus is composed and focused near the channel's walls. Figs. 5 and 6 illustrate that as the Ta and θ values increase, the trapped bolus size increases. This outcome correlates with the observation that an increase in these parameters leads to a rise in the flow rate, thereby resulting in the generation of more streamlines and boluses. Figs. 7 and 8 reveal a decrease in the size and number of the trapped bolus as the magnitude of ξ and Φ are enlarged. This result aligns well with findings from previous studies conducted by [9] and [19].

5. Conclusions

The influence of rotation on peristaltic motion for non-Newtonian pseudoplastic fluid in a wavy channel has yielded significant insights. By modeling and transforming the basic governing equations and employing perturbation method, we

have obtained analytical expressions for key parameters such as stream function, velocity, pressure gradient, and pressure rise. The study has led to several noteworthy findings:

1. The velocity profile exhibits a parabolic shape, with a decrease in velocity profile attributed to an increase in the pseudoplastic fluid parameter ξ. This decrease is due to the shear-thickening effect or reduction in fluid viscosity as the rate of deformation increases, thereby opposing the flow.
2. The velocity profile exhibits a reduction in the central region with increasing rotation parameter Ta and phase difference Φ, while it increases towards the boundaries. This phenomenon arises from the enhancement of kinematic forces induced by rotation, and wave phase difference particularly at the boundaries, thereby accelerating fluid flow.
3. Because of the direct effect of Ta parameter on the fluid velocity, a smooth flow can occur without requiring a significant pressure gradient with increasing Ta and mean flow rate θ, while an opposite trend is observed as ξ increases, necessitating a substantial pressure gradient.
4. The peristaltic pumping region contracts with an increase in the lower wall amplitude a, while it strengthens with an increase in the upper wall amplitude b.
5. The size of the trapped bolus increases with increasing values of Ta and θ, while its volume and number decrease with the magnitude of ξ and Φ.
6. These findings offer valuable insights into the behavior of peristaltic motion in complex fluid systems, with potential applications in various fields such as biomedical engineering, microfluidics, and industrial processes. Understanding these phenomena can aid in the design and optimization of systems involving fluid transport, leading to improved efficiency and performance.

Acknowledgments

The authors would like to thank the editor and anonymous referees for helpful comments and suggestions.

References

[1] Ali, H. A. (2021). Radiative peristaltic transport of Ree-Eyring fluid through porous medium in asymmetric channel subjected to combined effect of inclined MHD and convective conditions. *In Journal of Physics: Conference Series* (Vol. 1879, No. 2, p. 022101). IOP Publishing. https://doi.org/10.1088/1742-6596/1879/2/022101

[2] Ali, H. A. (2022). Impact of Varying Viscosity with Hall Current on Peristaltic Flow of Viscoelastic Fluid Through Porous Medium in Irregular Microchannel. *Iraqi Journal of Science,* 1265-1276. https://doi.org/10.24996/ijs.2022.63.3.31

[3] Waseem, F., Sohail, M., Lone, S. A., & Chambashi, G. (2023). Numerical simulations of heat generation, thermal radiation and thermal transport in water-based nanoparticles: OHAM study. *Scientific Reports*, 13(1), 15650. https://doi.org/10.1038/s41598-023-42582-4

[4] Li, S., Akbar, S., Sohail, M., Nazir, U., Singh, A., Alanazi, M., & Hassan, A. M. (2023). Influence of buoyancy and viscous dissipation effects on 3D magneto hydrodynamic viscous hybrid nano fluid (MgO-TiO2) under slip conditions. *Case Studies in Thermal Engineering*, 49, 103281. https://doi.org/10.1016/j.csite.2023.103281

[5] Latham, T. W. (1966). *Fluid motions in a peristaltic pump (Doctoral dissertation*, Massachusetts Institute of Technology).

[6] Shapiro, A. H., Jaffrin, M. Y., & Weinberg, S. L. (1969). Peristaltic pumping with long wavelengths at low Reynolds number. *Journal of fluid mechanics*, 37(4), 799-825. https://doi.org/10.1017/S0022112069000899

[7] Al-Khafajy, D. G. S., & Noor, A. L. (2022). The peristaltic flow of jeffrey fluid through a flexible channel. *Iraqi Journal of Science*, 5476-5486. https://doi.org/10.24996/ijs.2022.63.12.33

[8] Salman, M. R., & Ali, H. A. (2020). Approximate treatment for the MHD peristaltic transport of jeffrey fluid in inclined tapered asymmetric channel with effects of heat transfer and porous medium. *Iraqi Journal of Science*, 3342-3354. https://doi.org/10.24996/ijs.2020.61.12.22

[9] Hayat, T., Iqbal, R., Tanveer, A., & Alsaedi, A. (2018). Variable viscosity effect on MHD peristaltic flow of pseudoplastic fluid in a tapered asymmetric channel. *Journal of Mechanics*, 34(3), 363-374. https://doi.org/10.1017/jmech.2016.111

[10] Nazir, U., Sohail, M., Mukdasai, K., Singh, A., Alahmadi, R. A., Galal, A. M., & Eldin, S. M. (2022). Applications of variable thermal properties in Carreau material with ion slip and Hall forces towards cone using a non-Fourier approach via FE-method and mesh-free study. *Frontiers in Materials*, 9, 1054138. https://doi.org/10.3389/fmats.2022.1054138

[11] Hayat, T., Iqbal, R., Tanveer, A., & Alsaedi, A. (2016). Soret and Dufour effects in MHD peristalsis of pseudoplastic nanofluid with chemical reaction. *Journal of Molecular Liquids*, 220, 693-706.

https://doi.org/10.1016/j.molliq.2016.04.123

[12] Imran, N., Javed, M., Sohail, M., Qayyum, M., & Mehmood Khan, R. (2023). Multi-objective study using entropy generation for Ellis fluid with slip conditions in a flexible channel. *International Journal of Modern Physics* B, 37(27), 2350316. https://doi.org/10.1142/S0217979223503162

[13] Liu, J., Nazir, U., Sohail, M., Mukdasai, K., Singh, A., Alanazi, M., & Chambashi, G. (2023). Numerical investigation of thermal enhancement using MoS2-Ag/C2H6O2 in Prandtl fluid with Soret and Dufour effects across a vertical sheet. *AIP Advances*, 13(7). https://doi.org/10.1063/5.0152262

[14] Khan, R. M., Ashraf, W., Sohail, M., Yao, S. W., & Al-Kouz, W. (2020). On behavioral response of microstructural slip on the development of magnetohydrodynamic micropolar boundary layer flow. *Complexity*, 2020, 1-12. https://doi.org/10.1155/2020/8885749

[15] Sohail, M., Alyas, N., & Saqib, M. (2023). Contribution of double diffusion theories and thermal radiation on three dimensional nanofluid flow via optimal homotopy analysis procedure. *Int J Emerg Multidis Math*, 2(1), 2790-3257. https://doi.org/10.59790/2790-3257.1038

[16] Sohail, M., Nazir, U., Naz, S., Singh, A., Mukdasai, K., Ali, M. R., Khan, M. J. & Galal, A. M. (2022). Utilization of Galerkin finite element strategy to investigate comparison performance among two hybrid nanofluid models. *Scientific Reports*, 12(1), 18970. https://doi.org/10.1038/s41598-022-22571-9

[17] Hou, E., Jabbar, N., Nazir, U., Sohail, M., Javed, M. B., Shah, N. A., & Chung, J. D. (2022). *Significant mechanism of Lorentz force in energy transfer phenomena involving viscous dissipation via numerical strategy.*

[18] Sohail, M., & Nazir, U. (2023). Numerical computation of thermal and mass transportation in Williamson material utilizing modified fluxes via optimal homotopy analysis procedure. *Waves in Random and Complex Media*, 1-22. https://doi.org/10.1080/17455030.2023.2226230

[19] Tahir, M., & Ahmad, A. (2020). Impact of pseudoplaticity and dilatancy of fluid on peristaltic flow and heat transfer: Reiner-Philippoff fluid model. *Advances in Mechanical Engineering*, 12(12), 1687814020981184. https://doi.org/10.1177/1687814020981184

[20] Hina, S., Hayat, T., Mustafa, M., & Alsaedi, A. (2014). Peristaltic transport of pseudoplastic fluid in a curved channel with wall properties and slip conditions. *International Journal of Biomathematics*, 7(02), 1450015. https://doi.org/10.1142/S1793524514500156

[21] Salman, M. R. (2023, February). Effect of convective conditions in a radiative peristaltic flow of pseudoplastic nanofluid through a porous medium in a tapered an inclined asymmetric channel. *In AIP Conference Proceedings* (Vol. 2414, No. 1). AIP Publishing. https://doi.org/10.1063/5.0117556

[22] Alshareef, T. S. (2020). Impress of rotation and an inclined MHD on waveform motion of the non-Newtonian fluid through porous canal. *In Journal of Physics: Conference Series*, 1591(1), 012061. IOP Publishing. https://doi.org/10.1088/1742-6596/1591/1/012061

[23] Abd-Alla, A. M., & Abo-Dahab, S. M. (2016). Rotation effect on peristaltic transport of a Jeffrey fluid in an asymmetric channel with gravity field. *Alexandria Engineering Journal*, 55(2), 1725-1735. https://doi.org/10.1016/j.aej.2016.03.018

[24] Nassief, A. M., & Abdulhadi, A. M. (2023). Rotation and magnetic force effects on peristaltic transport of non-newtonian fluid in a symmetric channel. *Ibn AL-Haitham Journal For Pure and Applied Sciences*, 36(2), 436-453. https://doi.org/10.30526/36.2.3066

[25] Mohaisen, H. N., & Abdalhadi, A. M. (2022). Influence of the induced magnetic and rotation on mixed convection heat transfer for the peristaltic transport of bingham plastic fluid in an asymmetric channel. *Iraqi Journal of Science*, 1770-1785. https://doi.org/10.24996/ijs.2022.63.4.35

[26] Abd-Alla, A. M., & Abdallah, M. (2022). Effect of heat transfer and rotation on the peristaltic flow of a micropolar fluid in a vertical symmetric channel. *Sohag Journal of Sciences*, 7(3), 117-122.

[27] Salim, S. H., Al-Anbari, R. H., & Haider, A. J. (2021). Polymeric membrane with nanomaterial's for water purification: a review. *In IOP Conference Series: Earth and Environmental Science*, 779(1), 012103. IOP Publishing. https://doi.org/10.1088/1755-1315/779/1/012103

[28] Radhi, A. B., Khashan, K. S., & Sulaiman, G. M. (2023). Antibacterial Activity of Gold Nanoparticles Produced by One-Step Pulsed Laser Ablation in Liquid. *Plasmonics*, 1-13. https://doi.org/10.1007/s11468-023-02081-8

[29] Salman, M. R., & Abdulhadi, A. M. (2018, May). Influence of heat and mass transfer on inclined (MHD) peristaltic of pseudoplastic nanofluid through the porous medium with couple stress in an inclined asymmetric channel. *In Journal of Physics: Conference Series* (Vol. 1032, No. 1, p. 012043). IOP Publishing. https://doi.org/10.1088/1742-6596/1032/1/012043

[30] Khan, R. M., Imran, N., Mehmood, Z., & Sohail, M. (2022). A Petrov-Galerkin finite element approach for the unsteady boundary layer upper-convected rotating Maxwell fluid flow and heat transfer analysis. *Waves in Random and Complex Media*, 1-18. https://doi.org/10.1080/17455030.2022.2055201

[31] Ali, U., Alqahtani, A. S., Rehman, K. U., & Malik, M. Y. (2019). On Cattaneo-Christov heat flux

analysis with magneto-hydrodynamic and heat generation effects in a Carreau nano-fluid over a stretching sheet. *Revista mexicana de física*, 65(5), 479-488. https://doi.org/10.31349/RevMexFis.65.479

[32] Bibi, M., Zeeshan, A., Malik, M. Y., & Rehman, K. U. (2019). Numerical investigation of the unsteady solid-particle flow of a tangent hyperbolic fluid with variable thermal conductivity and convective boundary. *The European Physical Journal Plus*, 134, 1-10. https://doi.org/10.1140/epjp/i2019-12651-9

[33] Rehman, K. U., Malik, M. Y., Makinde, O. D., & Malik, A. A. (2017). A comparative study of nanofluids flow yields by an inclined cylindrical surface in a double stratified medium. *The European Physical Journal Plus*, 132(10), 427. https://doi.org/10.1140/epjp/i2017-11679-1

[34] Rehman, K. U., Al-Mdallal, Q. M., Tlili, I., & Malik, M. Y. (2020). Impact of heated triangular ribs on hydrodynamic forces in a rectangular domain with heated elliptic cylinder: finite element analysis. *International Communications in Heat and Mass Transfer*, 112, 104501. https://doi.org/10.1016/j.icheatmasstransfer.2020.104501

[35] Rehman, K. U., Shatanawi, W., & Al-Mdallal, Q. M. (2022). A comparative remark on heat transfer in thermally stratified MHD Jeffrey fluid flow with thermal radiations subject to cylindrical/plane surfaces. *Case Studies in Thermal Engineering*, 32, 101913. https://doi.org/10.1016/j.csite.2022.101913

[36] Abass, G., & Shihab, S. (2023). Operational Matrix of New Shifted Wavelet Functions for Solving Optimal Control Problem. *Mathematics*, 11(14), 3040. https://doi.org/10.3390/math11143040

[37] Aydin, C., & Tezer-Sezgin, M. (2019). The DRBEM solution of Cauchy MHD duct flow with a slipping and variably conducting wall using the well-posed iterations. *An International Journal of Optimization and Control: Theories & Applications (IJOCTA)*, 9(3), 76-85. https://doi.org/10.11121/ijocta.01.2019.00677

Hayat Adel Ali *is full professor in applied mathematics at the Department of Mathematics and Computer Application, University of Technology-Iraq. My research interests include numerical analysis and fluid mechanic.*

https://orcid.org/0000-0002-0368-0764

Mohammed R. Salman *is an assistance professor in fluid mechanics at the Department of Mathematics, College of Education for Pure Sciences, University of Karbala, Iraq. His research interest includes fluid mechanics.*

https://orcid.org/0000-0001-9611-4544

This work is licensed under a Creative Commons Attribution 4.0 International License. The authors retain ownership of the copyright for their article, but they allow anyone to download, reuse, reprint, modify, distribute, and/or copy articles in IJOCTA, so long as the original authors and source are credited. To see the complete license contents, please visit http://creativecommons.org/licenses/by/4.0/.

An International Journal of Optimization and Control: Theories & Applications
ISSN:2146-0957 eISSN:2146-5703
Vol.14, No.4, pp.346-354 (2024)
http://doi.org/10.11121/ijocta.1575

RESEARCH ARTICLE

A comparative view to $\mathcal{H}_\infty$-norm of transfer functions of linear DAEs

Hasan Gündüz [a*], Ercan Çelik [b], Mesut Karabacak[c]

[a] *Department of Mathematics, Bingol University, Türkiye*
[b] *Department of Applied Mathematics and Informatics, Kyrgyz-Turkish Manas University, Kyrgyzstan*
[c] *Department of Mathematics, Ataturk University, Türkiye*
hgunduz@bingol.edu.tr, ercan.celik@manas.edu.kg, mkarabacak@atauni.edu.tr

ARTICLE INFO

Article History:
Received 31 March 2024
Accepted 9 July 2024
Available Online 10 October 2024

Keywords:
DAEs systems
$\mathcal{H}_\infty$-norm
Bisection method
Extended-balanced singular perturbation method

AMS Classification 2010:
90B23; 90B56

ABSTRACT

In this paper, bisection and extended-balanced singular perturbation methods are used to calculate the $\mathcal{H}_\infty$-norm of the transfer function of a linear DAEs system for the particular case $D = 0$. In the beginning the approaches' algorithms and error analysis are provided separately. Next, the methods are employed to calculate the $\mathcal{H}_\infty$-norms of a numerical example pertaining to an automotive gas turbine model, and the error limits are used to check the norms in the suitable range, respectively. Ultimately, every solution is compared individually with the problem's $\mathcal{H}_\infty$-norm values, which are retrieved from MATLAB.

(cc) BY

1. Introduction

The phenomenon of control means "making a system capable of acting as desired". It dates back to Ancient Egypt. The first known control tools are water clocks, Although it is not known when and by whom water clocks were first invented, the first example was found in tomb of Pharaoh Amenhotep I, in 1500s BC. These mechanisms, known to have been designed by Vitrivius and Ktesibos in 325 BC and called clepsydra (water thief) were used by the Greeks to adjust speaking times in assemblies and courts.

Control, in the modern sense, begins with Watt's steam engine, in 1789. Until the 1870s, hundreds of regulators (governors) were patented worldwide using Watt's principles. From then until today, major seminal works in the field of control were carried out by many famous scientists from various areas such that; Maxwell, Vyshnegradski, Routh, Lyapunov, Hurwitz, Sickels, McFarlane, Farcot, Minorsky, Nyquist, Bode, Bellman, Pontryagyn, Kalman etc [1]. For more detail about historical development of control theory see [2].

One of the most powerful techniques of modern control is $\mathcal{H}_\infty$ control. $\mathcal{H}_\infty$ control is a very useful tool for large-scale multivariable problems to numerically measure the performance, sensitivity and durability of closed loop (feedback) system controllers. Its primary aim is to reduce modeling inaccuracies and account for unquantified disturbances, such as environmental factors, inner uncertainties, and noise, by transforming an optimization problem into a sensitivity problem involving the $\mathcal{H}_\infty$-norm. Here, $\mathcal{H}_\infty$ refers to "the space encompassing all bounded analytic matrix-valued functions within the open right-half complex plane." This concept was initially introduced by Zames in 1981 [3] and has since found applications across numerous works utilizing various control theory techniques [4–9]. For more comprehensive information, please refer to [10].

*Corresponding Author

Boyd, Balakrishnan and Kamamba (1988) presented the Bisection algorithms, which is an important tool in control theory [11, 12]. In this field, the energy associated with each state of the system is characterized by Hankel singular values, whereas eigenvalues indicate the stability of a system. Making a link between the stability of the system and the energy of its states is the basic idea underlying the bisection method. This entails connecting the imaginary eigenvalues of the associated Hamiltonian matrix, designated as M_γ in Eq.(12), with the singular values of the transfer matrix evaluated along the imaginary axis. The technique is applied in numerous works [13, 14].

In high-level control problems which contain large number of variables and parameters, researches confront many difficulties and complexity. To cope with these adverse conditions, researchers try to create some alternative methods to convert these high-level problems into far smaller dimensional models which can be solved more easily, without losing structural characteristic of the original problems. These kinds of methods are called model order reductions [15–18]. One of the methods is balanced truncation approach. Balanced truncation approach means, to find appropriate balanced realization and truncate this realization preserving the structural characteristic of the original problems.

Let $\mu > 0$ be a parameter, a dynamical system which contains some state component derivatives with μ coefficients is called a singular perturbation model. Singular perturbation models are represented by following set of equations,

$$\begin{aligned} \dot{x}_1 &= A_{11}x_1 + A_{12}x_2 + B_1u \quad (1)\\ \mu\dot{x}_2 &= A_{21}x_1 + A_{22}x_2 + B_2u \quad (2)\\ y &= C_1x_1 + C_2x_2 + Du \quad (3)\end{aligned}$$

here x_1, x_2 are called slow and fast variables, respectively, Eq.(1), Eq.(2) are called slow (powerful) and fast(weak) subsystems, respectively and μ is called perturbation parameter.

Analysis of these system types is done by singular perturbation theory. Singular perturbation theory means to investigate behavior of solutions of the system Eq.(2) for an interval $0 \le t \le T$ (or $0 \le t < +\infty$). The basic idea of singular perturbation method is to protect the slow(low-frequency) part (Eq.(1)) while neglecting the fast(high-frequency) (Eq.(1)). When considered from this point of view the method can be associated with a dominant mode state. In other words, it is process of examining solutions of the given system for $\mu = 0$ [18, 19]. μ-parameter may correspond to different concepts depending on the structure of the system. For example, it represents machine reactance or transients in voltage regulators in power systems, actuators in industrial control, enzymes in biochemical models and fast neutrons in nuclear reactor models.

The extended-balanced singular perturbation method represents a hybrid approach that combines the principles of both balanced truncation and singular perturbation methods. It begins by reducing the model order through the application of balanced truncation. Subsequently, the norm of the transfer function for the reduced model is determined using the singular perturbation method.

This paper organized into six sections. A number of basic definitions and notations which will be used next chapters are given in Section 2. In Section 3, general information about bisection method is told and algorithm of the method is given. In Section 4, extended-balanced singular perturbation method is told and its algorithm summarize as a table with the error bounds. A numerical example is solved by both methods and tolerances are computed in Section 5. Finally, in Section 6, the results are compared and discussed.

2. Preliminaries

Let's examine the linear dynamic system;

$$\begin{aligned} \dot{x} &= Ax + Bu\\ y &= Cx + Du \quad (4)\end{aligned}$$

where $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}, C \in \mathbb{R}^{p\times n}, D \in \mathbb{R}^{p\times m}$. Transfer matrix (or function) of the system Eq. (4) is defined as;

$$G(s) = C(sI - A)^{-1}B + D \quad (5)$$

Let $\lambda_j(M), \sigma_j(M)$ denote the j^{th} eigenvalue and j^{th} singular value of a matrix M respectively, where $\sigma_j(M) = \sqrt{\lambda_j\,(MM^T)}$.$A$ is stable if $Re\,(\lambda_j(A)) < 0$ for all j. If A is stable H_∞-norm of the transfer matrix $G(s)$ is given as follows;

$$\|G\|_\infty = \sup_{\text{Re } s>0} \sigma_{\max}(G(s)) = \sup_{\omega\in\mathbb{R}} \sigma_{\max}(G(i\omega)) \quad (6)$$

where $\sup_{\omega\in\mathbb{R}}$ denotes least upper bound for all real frequencies ω.

Let $J_{2n\times 2n} = \begin{bmatrix} 0_n & I_n \\ -I_n & 0_n \end{bmatrix}$ be a skew-symmetric matrix where $0_n, I_n$ are n-dimensional zero and

identity matrices, respectively. $H_{2n\times 2n}$ is called a Hamiltonian matrix, if HJ is symmetric, such that $(HJ)^T = HJ$. The definition confirms that the distinctive block structure form of Hamiltonian matrices is as follows;

$H = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & -H_{11}^T \end{bmatrix}$, where H_{12} and H_{21} are symmetric. For the system Eq.(4) the matrices $W_{\mathcal{C}}(t)$ and $W_{\mathcal{O}}(t)$ are called controllable and observable Grammians, respectively, defined as follows;

$$\begin{aligned} W_{\mathcal{C}}(t) &= \int_0^t e^{A\tau} B B^T e^{A^T\tau} d\tau \\ W_{\mathcal{O}}(t) &= \int_0^t e^{A^T\tau} C^T C e^{A\tau} d\tau \end{aligned} \tag{7}$$

which satisfy the Lyapunov equations as follows;

$$\begin{aligned} A^T W_{\mathcal{O}} + W_{\mathcal{O}} A + C^T C &= 0 \\ A W_{\mathcal{C}} + W_{\mathcal{C}} A^T + B B^T &= 0 \end{aligned} \tag{8}$$

and singular values of $W_{\mathcal{C}}(t)W_{\mathcal{O}}(t)$ are called Hankel singular values of the system Eq.(4) which describes the energy of each state of the system Eq.(4) and are denoted as σH_j for $j = 1, 2, \ldots$

Any positive definite matrix M can be expressed in the form of

$$M = LL^T \tag{9}$$

where L is a lower triangular matrix. The expression Eq.(9) and the matrix L are called Cholesky factorization and Cholesky factor of M, respectively. Let $M \in \mathbb{R}^{m\times n}$ and $rank(M) = r = \min(m, n)$, the expression

$$M = U\Sigma V^T \tag{10}$$

is called singular value decomposition of the matrix M. Here U and V are orthogonal matrices of type of $m \times m$ and $n \times n$, respectively, that is, $U^T U = I_m, V^T V = I_n$ and Σ is a half-diagonal matrix which contains singular values $(\sigma_1, \ldots, \sigma_r)$ of the matrix M. Singular value decomposition can be formulated clearly as follows for a matrix M,

$$M = U\Sigma V^T = \underbrace{\begin{bmatrix} u_1 & | & u_2 & | & \cdots & | & u_m \end{bmatrix}}_{u(m\times m)} \times \underbrace{\begin{bmatrix} \sigma_1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & \ddots & 0 & 0 & 0 & \ldots & 0 \\ \vdots & 0 & \sigma_r & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & \ldots & 0 & \ldots & 0 \end{bmatrix}}_{\Sigma(m\times n)} \underbrace{\begin{bmatrix} v_1^T \\ \hline v_2^T \\ \hline \vdots \\ \hline v_n^T \end{bmatrix}}_{V^T(n\times n)}. \tag{11}$$

3. Bisection method

Let $\gamma > 0$ related Hamiltonian matrix M_γ for system Eq. (4) is given as follows;

$$\begin{aligned} M_\gamma &= \begin{bmatrix} A & 0 \\ 0 & -A^T \end{bmatrix} + \begin{bmatrix} B & 0 \\ 0 & -C^T \end{bmatrix} \times \\ &\quad \begin{bmatrix} -D & \gamma I \\ \gamma I & -D^T \end{bmatrix}^{-1} \begin{bmatrix} C & 0 \\ 0 & B^T \end{bmatrix} \\ &= \begin{bmatrix} A - BR^{-1}D^T C & -\gamma B R^{-1} B^T \\ \gamma C^T S^{-1} C & -A^T + C^T D R^{-1} B^T \end{bmatrix} \end{aligned} \tag{12}$$

where $R = D^T D - \gamma^2 I$ and $S = DD^T - \gamma^2 I$.

For special case

$$D = 0, M_\gamma = \begin{bmatrix} A & \frac{1}{\gamma} B B^T \\ -\frac{1}{\gamma} C^T C & -A^T \end{bmatrix}.$$

Prior to initiating the bisection algorithm, it is essential to establish clear lower (γ_{lb}) and upper (γ_b) bounds. While one option is to set $\gamma_{lb} = 0$ and γ_{ub} to a sufficiently large value before proceeding with the bisection protocol, this approach can be time-consuming and inefficient. To streamline this process and determine suitable bounds, we can leverage Hankel singular values, as derived by Enns [20] and Glover [21], which are outlined below:

$$\begin{aligned} \gamma_{lb} &= \max\left\{\sigma_{\max}(D), \sqrt{Tr\left(W_{\mathcal{C}} W_{\mathcal{O}}\right)/n}\right\} \\ \gamma_{ub} &= \sigma_{\max}(D) + 2\sqrt{nTr\left(W_{\mathcal{C}} W_{\mathcal{O}}\right)} \end{aligned} \tag{13}$$

or alternative formulas;

$$\begin{aligned} \gamma_{lb} &= \max\left\{\sigma_{\max}(D), \sigma H_1\right\} \\ \gamma_{ub} &= \sigma_{\max}(D) + 2\sum_{j=1}^{n} \sigma H_i \end{aligned} \tag{14}$$

here, σH_i s represents the Hankel singular values, while $W_{\mathcal{O}}$ and $W_{\mathcal{C}}$ stand for the observability and controllability Grammians of the system Eq.(4)

Assuming A is stable and $\varepsilon > 0$ is the error margin for system Eq.(4), the bisection algorithm is outlined as follows:

Step 1. Determine the lower and upper bounds for the bisection algorithm, where

$$\gamma_{lb} = \max\{\sigma_{\max}(D), \sigma H_1\}$$
$$\gamma_{ub} = \sigma_{\max}(D) + 2\sum_{j=1}^{n} \sigma H_i$$

Step 2. Set $\gamma = (\gamma_{lb} + \gamma_{ub})/2$
If $\gamma_{ub} - \gamma_{lb} < \frac{\varepsilon}{2}$, end.

Step 3. Calculate M_γ.

Step 4. Check eigenvalues of M_γ. If there exists a purely imaginary eigenvalue set $\gamma_{lb} = \gamma$. Else set $\gamma_{ub} = \gamma$.

4. Extended balanced singular perturbation method

The extended balanced singular perturbation method, as introduced in the Introduction, combines the principles of both balanced truncation and singular perturbation methods, as described below.

Suppose we have an asymptotically stable, minimal realization of the system Eq.(4) as defined in equation Eq.(5). The algorithm for the balanced truncation approach is implemented using the following MATLAB commands:

Step 1. Find controllable and observable Grammians $W_\mathcal{C}$ and $W_\mathcal{O}$ of the given system through the Lyapunov eqtions with the MATLAB commands

Wc=gram(sys,'c')
Wo=gram(sys, 'o').

Step 2. Find the Cholesky factors $L_\mathcal{C}$ and $L_\mathcal{O}$ of $W_\mathcal{C}$ and $W_\mathcal{O}$, respectively, such that

$$W_\mathcal{C} = L_\mathcal{C} L_\mathcal{C}^T$$
$$W_\mathcal{O} = L_\mathcal{O} L_\mathcal{O}^T$$

with the MATLAB commands

$$Lc = chol(Wc,' lower')$$
$$Lo = chol(Wo,' lower').$$

Step 3. Find the singular value decomposition of $L_\mathcal{O}^T L_\mathcal{C}$ such that

$$L_\mathcal{O}^T L_\mathcal{C} = U\Sigma V^T$$

with the MATLAB command

$$[U, S, V] = svd(Lo\ '\ *\mathrm{Lc}).$$

Step 4. Make the transformation $T = L_\mathcal{C} V \Sigma^{-1/2}$ and obtain coefficient matrices of balanced system by similarity transformation as follows,

$$\tilde{A} = T^{-1}AT, \quad \tilde{B} = T^{-1}B, \quad \tilde{C} = CT, \quad \tilde{D} = D$$

where $\tilde{G}(s) = \left[\begin{array}{c|c} \tilde{A} & \tilde{B} \\ \hline -\tilde{C} & \tilde{D} \end{array}\right]$ and find controllable and observable Grammians of the balanced system $\widetilde{W}_C$ and $\widetilde{W}_0$ respectively which are given as below,

$$\widetilde{W}_\mathcal{C} = T^{-1} W_\mathcal{C} T^{-T}$$
$$\widetilde{W}_\mathcal{O} = T^T W_\mathcal{O} T$$

here $\widetilde{W}_\mathcal{C} = \widetilde{W}_\mathcal{O} = \Sigma = diag\,(\sigma_1, \sigma_1, \ldots, \sigma_n)$.

Let $\tilde{G}(s) = \left[\begin{array}{c|c} \tilde{A} & \tilde{B} \\ \hline \tilde{C} & \tilde{D} \end{array}\right]$ be the balanced system obtained by balanced truncation approach, the algorithm of singular perturbation method is given as follows;

Step 1. Separate the balanced system $\tilde{G}(s) = \left[\begin{array}{c|c} \tilde{A} & \tilde{B} \\ \hline \tilde{C} & \tilde{D} \end{array}\right] \Leftrightarrow \left[\begin{array}{cc} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{array}\right]$ into two subsystem as slow(powerful) and fast(weak). Choose A_{11} as coefficient matrix of the slow part where $A_{11}, \Sigma_1 \in \mathbb{R}^{r\times r}$, for $r \ll n$. Rearrange the matrices $\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D}$ in block matrix form as seen below,

$$\tilde{A} = \left[\begin{array}{cc} A_{11} & A_{12} \\ A_{21} & A_{22} \end{array}\right], \tilde{B} = \left[\begin{array}{c} B_1 \\ B_2 \end{array}\right]$$
$$\tilde{C} = \left[\begin{array}{cc} C_1 & C_2 \end{array}\right], \tilde{D} = D$$

add perturbation parameter μ and rewrite $\tilde{G}(s)$ as the followings,

$$\left[\begin{array}{c} \dot{x}_1 \\ \mu\dot{x}_2 \end{array}\right] = \left[\begin{array}{cc} A_{11} & A_{12} \\ A_{21} & A_{22} \end{array}\right]\left[\begin{array}{c} x_1 \\ x_2 \end{array}\right] + \left[\begin{array}{c} B_1 \\ B_2 \end{array}\right] u$$
$$y = \left[\begin{array}{cc} C_1 & C_2 \end{array}\right]\left[\begin{array}{c} x_1 \\ x_2 \end{array}\right] + Du.$$

Step 2. Eliminate the fast(weak) part $\mu = 0$ and find the system as;

$$\dot{x}_1 = A_{11}x_1 + A_{12}x_2 + B_1 u$$
$$0 = A_{21}x_1 + A_{22}x_2 + B_2 u$$
$$y = C_1 x_1 + C_2 x_2 + Du$$

and weak variable as,

$x_2 = -A_{22}{}^{-1}A_{21}x_1 - A_{22}{}^{-1}B_2 u.$

Step 3. Substitute x_2 to the other equations to get the final version of the system which denoted by $G_f(s)$ as is below

$$G_f(s) = \left[\begin{array}{c|c} A_f & B_f \\ \hline C_f & D_f \end{array}\right] =$$

Table 1. Algorithm of extended balanced singular perturbation method step by step.

	Balanced Truncation Approach	Singular Perturbation Method
Step 1.	Find Grammians of the original system $(W_{\mathcal{C}}, W_O)$	Separate the balanced system $\tilde{G}(s)$ into two parts as; strong and weak
Step 2.	Find Cholesky factors of Grammians $(L_{\mathcal{C}}, L_{\mathcal{O}})$	Eliminate the weak part taking $\mu = 0$ and find weak variable x_2
Step 3.	Find singular value decomposition of $L_O^T L_{\mathcal{C}} = U\Sigma V^T$	Substitute x_2 in other equations, get the final version of the system $G_f(s)$
Step 4.	Make the transformation $T = L_{\mathcal{C}} V \Sigma^{-1/2}$ and find the balanced system $\tilde{G}(s)$	Obtain the H_∞-norm of $\|G_f(s)\|_\infty$
Error Analysis	Compute actual and theoretical infinity error bounds and apply the error tolerance criterion which says actual bound must be less than or equal to theoretical bound	

$$\left[\begin{array}{c|c} A_{11} - A_{12}A_{22}{}^{-1}A_{21} & B_1 - A_{12}A_{22}{}^{-1}B_2 \\ \hline C_1 - C_2 A_{22}^{-1} A_{21} & D - C_2 A_{22}{}^{-1} B_2 \end{array}\right]$$

Step 4. Obtain the H_∞-norm of $\|G_f(s)\|_\infty$ in MATLAB.

The algorithm of extended balanced singular perturbation method is summarized in Table 1 as follows.

To analyze the error tolerance first we define modelling error transfer function as follows;

$$E_r = [G(s) - G_f(s)] \tag{15}$$

then, we have a criterion about sufficiency of error tolerance which is based on comparison of two error bounds called actual infinity error bound and theoretical infinity error bound defined in [22, 23] given as below;

- Actual infinity error bound: $\|E_r\|_\infty = \|[G(s) - G_f(s)]\|_\infty$
- Theoretical infinity error bound: $2\sum_{i=r+1}^{n} \sigma_i$
- The criterion:

$$\|E_r\|_\infty \leq 2 \sum_{i=r+1}^{n} \sigma_i$$

5. Application to a numerical example

Example 1. *The system two-input, twelve-state, two-output model of an automobile gas turbine [24].*

For more detail and examples see [25]. Consider the system Eq.(4) with the coefficient matrices given as follows:

When employing the bisection method for this problem, we obtain the values presented in Table 2. The first and the last columns in the table pertain to number of iteration that denoted as Itr briefly and verifying the presence of purely imaginary eigenvalues that denoted as Eig briefly, respectively.

Table 2. Related values of Example 2.

Itr	γ_{lb}	γ_{ub}	γ	Eig
1	3.0368	36.4417	19.7397	no
2	3.0368	19.7397	11.3881	yes
3	11.3881	19.7397	15.5637	no
4	11.3881	15.5637	13.4759	yes
5	13.4759	15.5637	14.5198	no
6	13.4759	14.5198	13.9979	no
7	13.4759	13.9979	13.7369	yes
8	13.7369	13.9979	13.8674	no
9	13.7369	13.8674	13.8022	no
10	13.7369	13.8022	13.7695	no
11	13.7369	13.7695	13.7532	no
12	13.7369	13.7532	13.7450	no
13	13.7369	13.7450	13.7410	no
14	13.7369	13.7410	13.7389	no
15	13.7369	13.7389	13.7379	yes
16	13.7379	13.7389	13.7384	yes
17	13.7384	13.7389	13.7387	yes
18	13.7387	13.7389	13.7388	no
19	13.7387	13.7388	13.7388	yes
20	13.7388	13.7388	13.7388	

After 20 iterations γ_{ub} and γ_{lb} are so close, the all next iterations will be automatically assigned the same value by MATLAB and error margin ε is also will be satisfied. Thus, $\mathcal{H}_\infty$ norm of transfer function of the given problem found as $\|G(s)\|_\infty \approx 13.7388$. Now, if we apply balanced truncation approach algorithm step by step finally we get

$$\tilde{G}(s) = \left[\begin{array}{c|c} \tilde{A} & \tilde{B} \\ \hline \tilde{C} & \tilde{D} \end{array}\right] \text{ where;}$$

and Hankel singular values of the original system as,

$$\sigma(G) = (7.1833 \quad 1.4904 \quad 0.9279 \quad 0.5876 \quad 0.4633 \quad 0.2368 \quad 0.1613 \quad 0.0936 \quad 0.0006 \quad 0 \quad 0 \quad 0).$$

$$A = \begin{bmatrix}
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-0.202 & -1.15 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -2.36 & -13.6 & -12.8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -1.62 & -9.4 & -9.15 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -188 & -111.6 & -116.4 & -20.8
\end{bmatrix}$$

$$B = \begin{bmatrix}
0 & 1.0439 & 0 & 0 & -1.794 & 0 & 0 & 1.0439 & 0 & 0 & 0 & -1.794 \\
0 & 4.1486 & 0 & 0 & 2.6775 & 0 & 0 & 4.1486 & 0 & 0 & 0 & 2.6775
\end{bmatrix}^T$$

$$C = \begin{bmatrix}
0.264 & 0.806 & -1.42 & -15 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 4.9 & 2.12 & 1.95 & 9.35 & 25.8 & 7.14 & 0
\end{bmatrix}$$

$$D = 0$$

$$\tilde{A} = \begin{bmatrix}
-0.1647 & 0.0185 & 0.1116 & -0.0528 & -0.4698 & 0.2334 & 0.0559 & -0.0336 & 0.0179 & -0.0049 & -0.0001 & 0 \\
-0.0071 & -0.8293 & -0.2913 & 0.3813 & 0.2261 & -0.0513 & -0.7095 & -1.6433 & 0.0079 & -0.0065 & 0.0008 & -0.0003 \\
-0.0605 & 0.6599 & -0.1368 & 0.1909 & 1.9093 & -0.6958 & 0.0074 & 0.3709 & -0.0453 & 0.0129 & 0.0001 & 0 \\
-0.0676 & -0.2951 & -0.1328 & -0.1111 & 0.0156 & 0.2976 & 0.3495 & 0.4641 & 0.0339 & -0.0075 & -0.0005 & 0.0001 \\
-0.4673 & -0.3296 & -1.3820 & -1.3904 & -5.8806 & 3.7574 & 1.4509 & -1.1650 & 0.4207 & -0.1141 & -0.0018 & 0.0004 \\
0.2332 & 0.0786 & 0.4107 & 0.5352 & 3.7468 & -2.6771 & -2.2760 & 1.6069 & -0.3983 & 0.1083 & 0.0016 & -0.0004 \\
-0.0103 & 0.7485 & -0.3240 & 0.0251 & -0.4105 & 1.6010 & -1.2640 & -2.6186 & -0.0926 & 0.0100 & 0.0033 & -0.0010 \\
0.0416 & 1.6502 & -0.7122 & 0.3698 & 1.9050 & -1.8933 & -4.2162 & -11.6604 & 0.3158 & -0.1526 & 0.0113 & -0.0035 \\
-0.0144 & -0.0545 & 0.0096 & -0.0499 & -0.3612 & 0.3313 & 0.1843 & 0.9343 & -6.2677 & 3.6203 & 0.0367 & -0.0080 \\
-0.0032 & -0.0197 & 0.0056 & -0.0128 & -0.0838 & 0.0753 & 0.0693 & 0.3210 & -3.5719 & -13.7547 & -0.0809 & 0.0128 \\
-0.0002 & 0 & -0.0002 & -0.0005 & -0.0042 & 0.0040 & -0.0004 & 0.0010 & -0.0948 & -0.5578 & -0.2477 & 0.1440 \\
0 & -0.0001 & 0.0001 & 0.0001 & 0.0011 & -0.0011 & 0.0005 & 0.0011 & 0.0177 & 0.0725 & 0.1343 & -0.9059
\end{bmatrix}$$

$$\tilde{B} = \begin{bmatrix}
-0.4823 & 1.4882 & -0.5001 & 0.0220 & -0.5164 & 0.2995 & -0.6184 & -1.3307 & 0.0274 & 0.0129 & -0.0003 & 0.0002 \\
-1.4609 & -0.5073 & -0.0613 & -0.3607 & -2.2765 & 1.0855 & 0.1595 & 0.6417 & -0.0796 & -0.0200 & -0.0008 & 0.0002
\end{bmatrix}^T$$

$$\tilde{C} = \begin{bmatrix}
-0.5368 & 1.4993 & 0.0399 & -0.3178 & -0.8139 & 0.3740 & 0.6374 & 1.3230 & 0.0181 & -0.0010 & -0.0008 & 0.0002 \\
-1.4417 & -0.4735 & 0.5023 & -0.1721 & -2.1879 & 1.0621 & 0.0399 & -0.6574 & 0.0823 & -0.0238 & -0.0001 & 0
\end{bmatrix}$$

$$\tilde{D} = 0$$

It is seen clearly in the Figure 1 the first three Hankel singular values are much greater than the others so we choose $r = 6$ and apply extended balanced singular perturbation method. First separate the balanced system $\tilde{G}(s)$ into two parts as slow(powerful) and fast(weak) and rewrite the system for perturbation parameter $\mu = 0$ as is given below;

$$\begin{aligned}
\dot{x}_1 &= A_{11}x_1 + A_{12}x_2 + B_1u \\
0 &= A_{21}x_1 + A_{22}x_2 + B_2 \\
y &= C_1x_1 + C_2x_2 + Du
\end{aligned}$$

where;

$$A_{11} = \begin{bmatrix}
-0.1647 & 0.0185 & 0.1116 & -0.0528 & -0.4698 & 0.2334 \\
-0.0071 & -0.8293 & -0.2913 & 0.3813 & 0.2261 & -0.0513 \\
-0.0605 & 0.6599 & -0.1368 & 0.1909 & 1.9093 & -0.6958 \\
-0.0676 & -0.2951 & -0.1328 & -0.1111 & 0.0156 & 0.2976 \\
-0.4673 & -0.3296 & -1.3820 & -1.3904 & -5.8806 & 3.7574 \\
0.2332 & 0.0786 & 0.4107 & 0.5352 & 3.7468 & -2.6771
\end{bmatrix}$$

$$A_{12} = \begin{bmatrix}
0.0559 & -0.0336 & 0.0179 & -0.0049 & -0.0001 & 0 \\
-0.7095 & -1.6433 & 0.0079 & -0.0065 & 0.0008 & -0.0003 \\
0.0074 & 0.3709 & -0.0453 & 0.0129 & 0.0001 & 0 \\
0.3495 & 0.4641 & 0.0339 & -0.0075 & -0.0005 & 0.0001 \\
1.4509 & -1.1650 & 0.4207 & -0.1141 & -0.0018 & 0.0004 \\
-2.2760 & 1.6069 & -0.3983 & 0.1083 & 0.0016 & -0.0004
\end{bmatrix}$$

$$A_{21} = \begin{bmatrix}
-0.0103 & 0.7485 & -0.3240 & 0.0251 & -0.4105 & 1.6010 \\
0.0416 & 1.6502 & -0.7122 & 0.3698 & 1.9050 & -1.8933 \\
-0.0144 & -0.0545 & 0.0096 & -0.0499 & -0.3612 & 0.3313 \\
-0.0032 & -0.0197 & 0.0056 & -0.0128 & -0.0838 & 0.0753 \\
-0.0002 & 0 & -0.0002 & -0.0005 & -0.0042 & 0.0040 \\
0 & -0.0001 & 0.0001 & 0.0001 & 0.0011 & -0.0011
\end{bmatrix}$$

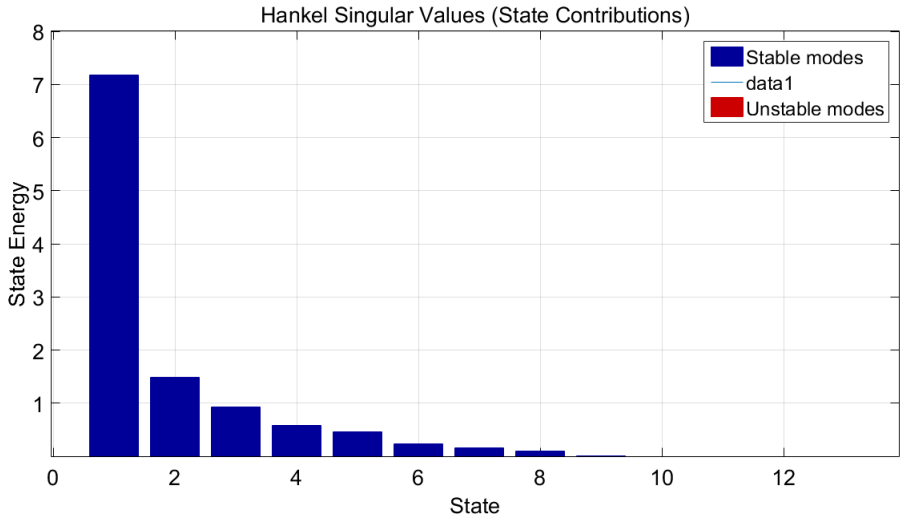


Figure 1. Hankel singular values of the original system.

$$A_{22} = \begin{bmatrix} -1.2640 & -2.6186 & -0.0926 & 0.0100 & 0.0033 & -0.0010 \\ -4.2162 & -11.6604 & 0.3158 & -0.1526 & 0.0113 & -0.0035 \\ 0.1843 & 0.9343 & -6.2677 & 3.6203 & 0.0367 & -0.0080 \\ 0.0693 & 0.3210 & -3.5719 & -13.7547 & -0.0809 & 0.0128 \\ -0.0004 & 0.0010 & -0.0948 & -0.5578 & -0.2477 & 0.1440 \\ 0.0005 & 0.0011 & 0.0177 & 0.0725 & 0.1343 & -0.9059 \end{bmatrix}$$

$$B_1 = \begin{bmatrix} -0.4823 & -1.4609 \\ 1.4882 & -0.5073 \\ -0.5001 & -0.0613 \\ 0.0220 & -0.3607 \\ -0.5164 & -2.2765 \\ 0.2995 & 1.0855 \end{bmatrix} \quad B_2 = \begin{bmatrix} -0.6184 & 0.1595 \\ -1.3307 & 0.6417 \\ 0.0274 & -0.0796 \\ 0.0129 & -0.0200 \\ -0.0003 & -0.0008 \\ 0.0002 & 0.0002 \end{bmatrix}$$

$$C_1 = \begin{bmatrix} -0.5368 & 1.4993 & 0.0399 & -0.3178 & -0.8139 & 0.3740 \\ -1.4417 & -0.4735 & 0.5023 & -0.1721 & -2.1879 & 1.0621 \end{bmatrix}$$

$$C_2 = \begin{bmatrix} 0.6374 & 1.3230 & 0.0181 & -0.0010 & -0.0008 & 0.0002 \\ 0.0399 & -0.6574 & 0.0823 & -0.0238 & -0.0001 & 0 \end{bmatrix} \quad D = 0$$

and from the second equation find weak variable as, $x_2 = -{A_{22}}^{-1}A_{21}x_1 - {A_{22}}^{-1}B_2u$. Continue from Step3 make necessary algebraic matrix operations and finally get, $G_f(s) = \left[\begin{array}{c|c} A_f & B_f \\ \hline C_f & D_f \end{array}\right]$ where;

$$A_f = \begin{bmatrix} -0.1690 & 0.0951 & 0.0784 & -0.0663 & -0.6557 & 0.6759 \\ -0.0058 & -1.1996 & -0.1312 & 0.3503 & 0.2631 & -0.5240 \\ -0.0513 & 0.5609 & -0.0938 & 0.2259 & 2.3058 & -1.5693 \\ -0.0772 & -0.0122 & -0.2554 & -0.1296 & -0.3903 & 1.3889 \\ -0.5873 & 1.7458 & -2.2822 & -1.7701 & -11.0375 & 15.9797 \\ 0.4156 & -3.1167 & 1.7961 & 1.1096 & 11.5963 & -21.3316 \end{bmatrix}$$

$$B_f = \begin{bmatrix} -0.5472 & -1.4593 \\ 1.7922 & -0.6034 \\ -0.4144 & -0.0472 \\ -0.2146 & -0.3261 \\ -2.2766 & -2.2466 \\ 3.0081 & 1.0291 \end{bmatrix} \quad C_f = \begin{bmatrix} -0.5419 & -1.4612 \\ 1.8765 & -0.2345 \\ -0.1233 & 0.3985 \\ -0.3049 & -0.2438 \\ -1.0190 & -3.0315 \\ 1.1786 & 2.9523 \end{bmatrix}^T$$

$$D_f = \begin{bmatrix} -0.3115 & 0.0807 \\ -0.2056 & -0.0223 \end{bmatrix}$$

Obtain the $\mathcal{H}_\infty$-norm in MATLAB as $\|G_f(s)\|_\infty = 13.7413$ which is so close to the $\mathcal{H}_\infty-$ norm of the original system $\|G(s)\|_\infty = 13.7388$. Let's now assess the error tolerance between the original system and the reduced-order balanced model using both actual and theoretical infinity error bounds, as outlined below.

$$\|E_r\|_\infty = \|[G(s) - G_f(s)]\|_\infty = 0.3774$$

and for $r = 6$ and $n = 12$,

$$2\sum_{i=r+1}^{n} \sigma_i = 2(0.1613 + 0.0936 + 0.0006 + 0 + 0 + 0) = 0.5110$$

It is obvious that $\|E_r\|_\infty \leq 2\sum_{i=r+1}^{n} \sigma_i$ thus we can say that error tolerance is in a satisfied level.

6. Conclusion

In this research, we applied both the bisection method and the extended-balanced singular perturbation method to analyze a linear dynamic system with the parameter D set to 0 . Our goal was to compute the $\mathcal{H}_\infty-$ norm of its transfer function. We conducted a numerical experiment using both methods and performed a detailed error analysis. The outcomes of our investigation revealed that the bisection method performed satisfactorily, with error tolerances falling within an acceptable range after a certain number of iterations. Similarly, the extended-balanced singular perturbation method demonstrated satisfactory performance, as the error tolerances met the criteria for investigating the accuracy of the reduced-order models. According to the $\mathcal{H}_\infty$-norms computed by methods, we conclude that bisection

method is a slightly accurate than extended-balanced singular perturbation method. Utilizing bisection and extended balanced singular perturbation methods, the research not only provides detailed algorithms and error analysis but also demonstrates practical application through a numerical example involving an automotive gas turbine model, enhancing the precision and reliability of $\mathcal{H}_\infty$-norm computations in real-world systems.

References

[1] Gunduz, H. (2023). $\mathcal{H}_\infty$ control of linear differential algebraic equation (DAEs) systems. PhD Thesis. Ataturk University.

[2] Bennett, S. (1996). A brief history of automatic control. *IEEE Control Systems Magazine*, 16(3), 17-25. https://doi.org/10.1109/37.506394

[3] Zames, G. (1981). Feedback and optimal sensitivity: Model reference transformations, multiplicative seminorms and approximate inverses. *IEEE Transactions on Automatic Control*, 26(2), 301-320. https://doi.org/10.1109/TAC.1981.1102603

[4] Zhou, K. & Doyle, J. C. (1998). Essential of Robust Control. Prentice Hall, Upper Saddle River, NJ, USA.

[5] Petersen, I. (1987). Disturbance attenuation and $\mathcal{H}_\infty$ optimization: a design method based on the algebraic Riccati equation. *IEEE Transactions on Automatic Control*, 32,(5), 427-429. https://doi.org/10.1109/TAC.1987.1104609

[6] Khargonekar, P. P., Petersen, I. R. & Rotea, M. A. (1988). $\mathcal{H}_\infty$ optimal control with state feedback. *IEEE Transactions on Automatic Control*, 33(8), 786-788 https://doi.org/10.1109/9.1301

[7] Doyle, J., Glover, K. & Khargonekar, P. P. (1988). State space solutions to standard H_2 and $\mathcal{H}_\infty$ control problems, proceedings, *IEEE American Control Conference*, 1691-1696.

[8] Khargonekar, P. P., Petersen, I. R. & Zhou K. (1990). Robust stabilization of uncertain linear systems:Quadratic stabilization and $\mathcal{H}_\infty$ control theory. *IEEE Transactions on Automatic Control*, 33(3), 356-361. https://doi.org/10.1109/9.50357

[9] Jiang, B., Wu, Z. & Karimi, H. R. (2022). A distiributed dynamic event-triggered mechanism to HMM-based observer design for $\mathcal{H}_\infty$ sliding mode control of markov jump systems. *Automatica*, 142, 110357. https://doi.org/10.1016/j.automatica.2022.110357

[10] Tacx, P. & Oomen, T. (2021). Accurate $\mathcal{H}_\infty$-norm estimation via finite frequency norms of local parametric models. *American Control Conference (ACC)*, 321-326.

[11] Boyd, S., Balakrishnan, V. & Kamamba, P. (1988). On computing the $\mathcal{H}_\infty$-norm of a transfer matrix. *Proceedings of 1988 American Control Conference*, 369-397.

[12] Boyd, S., Balakrishnan, V. & Kamamba, P. (1989). A bisection method for computing $\mathcal{H}_\infty$-norm of a transfer function matrix and related problems. Springer-Verlag, New York.

[13] Kuster, G. E. (2012). H infinity norm calculation via a state-space formulation. PhD Thesis. Faculty of Polytechique Institute, Blacksburg, Virginia.

[14] Gunduz, H. & Celik, E. (2022). $\mathcal{H}_\infty$-norm evaluation for a transfer matrix via bisection algorithm. *Thermal Science*, 26(2), 745-751. https://doi.org/10.2298/TSCI22S2745G

[15] Moore, B. (1981). Principal component analysis in linea systems:Controllability, observability and model reduction. *IEEE Transactions on automatic Control*, 26(1), 17-32. https://doi.org/10.1109/TAC.1981.1102568

[16] Pernebo, L. & Silverman, L. (1989). Model reduction via balanced state-space representations. *IEEE Transactions on Automatic Control*, 27(2), 382-387. https://doi.org/10.1109/TAC.1982.1102945

[17] Imran, M., Ghafoor, A. & Sreeram, V. (2014). A frequency weighted model order reduction technique and error bound. *Automatica*, 50(12), 3304-3309. https://doi.org/10.1016/j.automatica.2014.10.062

[18] Kokotovich, P. V., O'Malley Jr, R. E. & Sannuti, P. (1976). Singular perturbation and order reduction in control theory: an overview. *Automatica*, 12(2), 123-132. https://doi.org/10.1016/0005-1098(76)90076-5

[19] N'Diaye, M., Hussain, S., Suliman, I. M. A. & Toure, L. (2023). Robust uncertainity alleviation by H infinity analysis and control for singularity perturbed systems with disturbances. Journal of Xi'an Shioyu University, Natural Science Edition, Volume 19, Issue 01, 728-737.

[20] Enns, D. F. (1984). Model reduction with balanced realization: an error bound and a frequency weighted generalization. The 23rd IEEE Conference on Decision and Control, 127-132. https://doi.org/10.1109/CDC.1984.272286

[21] Glover, K. (1984). All optimal Hankel-norm approximations of linear multivarible systems and their $\mathcal{L}_\infty$-error bounds. *Internationl Journal of Control*, 39(6), 1115-1193. https://doi.org/10.1080/00207178408933239

[22] Datta, B. N. (2004). Numerical Methods for Linear Control Systems (Vol. 1). Academic Press, London, New York.

[23] Antoulas, A. C. & Benner P., Feng, L. (2018). Model reduction by iterative error system approximation. *Mathematical and Computer Modelling of Dynamical Systems*, 24(2), 103-118. https://doi.org/10.1080/13873954.2018.1427116

[24] Hung, Y. S. & MacFarlane, A. G. S. (1982). Multivariable feedback: a quasi-classical approach lecture notes in control and information in sciences. Springer-Verlag, Berlin-Heidelberg, New York.

[25] Leibfritz, P. & Lipinski, W. (2003). Description of benchmark examples in COMPl_eib 1.0, Tech Report.

Hasan Gunduz *is an assistant professor at Bingol University. He received his PhD in applied mathematics from Ataturk University in 2023. His research interests include differential algebraic equations (DAEs), control systems, optimization, numerical analysis.*
https://orcid.org/0000-0001-5851-3819

Ercan Çelik *obtained his PhD degree in Mathematics, Ataturk University, Erzurum, Turkey in 2002. He is currently working as a professor at the Department of Mathematics, Kyrgyz-Turkish Manas University, Kyrgyzstan. His fields of research are optimization, numerical analysis, applied mathematics.*
https://orcid.org/0000-0002-1402-1457

Mesut Karabacak *is an associate professor in the Department of Mathematics at Ataturk University, Erzurum, Turkey. He holds a PhD in Applied Mathematics (2015) from the same university. His research interests are in numerical analysis, wavelet theory, fractional partial differential equations (PDEs), and information theory.*
https://orcid.org/0000-0002-0057-8115

This work is licensed under a Creative Commons Attribution 4.0 International License. The authors retain ownership of the copyright for their article, but they allow anyone to download, reuse, reprint, modify, distribute, and/or copy articles in IJOCTA, so long as the original authors and source are credited. To see the complete license contents, please visit http://creativecommons.org/licenses/by/4.0/.

An International Journal of Optimization and Control: Theories & Applications
ISSN:2146-0957 eISSN:2146-5703
Vol.14, No.4, pp.355-364 (2024)
http://doi.org/10.11121/ijocta.1483

RESEARCH ARTICLE

Fuzzy-PID and interpolation: a novel synergetic approach to process control

Devashish Jha [a*], Arifa Ahmed [a], Sanatan Kumar [b], Debanjan Roy [c]

[a] *Department of Electronics & Communication Engineering, Madanapalle Institute of Technology & Science, India*
[b] *Department of Electrical and Electronics Engineering, Technocrat Institute of Technology, Bhopal, India*
[c] *Department of Electrical Engineering, Teerthanker Mahaveer University, Moradabad, India*
devashish.sit@gmail.com, arifanerist@gmail.com, skumar0185@gmail.com, debanjanroy88@gmail.com

ARTICLE INFO

Article History:
Received 4 November 2023
Accepted 25 July 2024
Available Online 10 October 2024

Keywords:
PID
Fuzzy
Linguistic variables
Interpolation
FIE

AMS Classification 2010:
93B52; 93C10; 93C42

ABSTRACT

This paper presents a novel approach for tuning a fuzzy-based proportional-integral-derivative (PID) controller to enhance the control performance of a chemical process control system. The proposed approach combines the advantages of fuzzy- PID and interpolation to achieve improved control performance. Properly tuned PID controllers can help maintain process stability, minimize deviations from setpoints, and ensure efficient operation in industrial systems. Fuzzy logic allows for the incorporation of expert knowledge and linguistic rules, enabling the controller to handle uncertain and imprecise process information. Fuzzy PID controllers combine fuzzy logic and conventional PID control to enhance control performance, particularly in systems with complex or nonlinear dynamic such as chemical plant. It dynamically adjusts the PID parameters—proportional gain (Kp), integral gain (Ki), and derivative gain (Kd)—based on error $e(t)$ and change of error $\Delta e(t)$. Interpolation plays a crucial role in this context by filling in the gaps or handling situations not explicitly covered by the fuzzy rules. Comparative studies are conducted to evaluate the performance of the fuzzy PID controller against conventional PID controllers and other advanced control techniques. It is demonstrated that the synergy between fuzzy logic and interpolation not only enhances control performance but also offers a more intuitive and adaptable solution for addressing the complexities of modern chemical process control systems.

CC BY

1. Introduction

Process control systems should be designed to ensure the efficient and safe operation of chemical processes in industries such as oil refining, petrochemicals, pharmaceuticals, and food processing. These systems monitor and regulate process variables such as temperature, pressure, flow rate, level and concentration with the aim of achieving optimal process performance, product quality, and resource utilization. The complexity of chemical processes, often characterized by nonlinear dynamics, time delays, and interactions between various process variables, poses significant challenges. The primary objective of a chemical process control system is to maintain these variables within desired operating ranges, despite disturbances and uncertainties in the process environment. Traditionally, control systems in chemical processes have relied on classical control techniques such as Proportional-Integral-Derivative (PID) control. PID controllers are widely used owing to their inherent simplicity, ease of implementation, and familiarity among control engineers. A

*Corresponding Author

PID controller calculates a control action based on the error between the desired setpoint and the actual process variable, taking into account the proportional, integral, and derivative components. However, traditional PID control approaches have limitations when applied to complex chemical processes. These limitations include difficulty in handling nonlinearities, interactions, and time-varying dynamics, as well as challenges in tuning the controller parameters for optimal performance. Consequently, there is a need for advanced control strategies that can address these limitations and enhance the control performance of chemical process systems.

Several classical methods pertaining to PID control have been proposed to achieve desired controller performance. Ziegler and Nichols in 1942 [1], first introduced the method of PID tuning. The tuning method described by them is a heuristic approach that provides simple and practical technique for initial tuning but had limitations in terms of control performance and robustness. To overcome the limitations of the Ziegler-Nichols (Z-N) method, researchers focused on developing more sophisticated techniques. Passivity-based control strategy [2] offers a promising alternative for controlling generalized passive systems, providing stability, robustness, and adaptability to different system dynamics. Cohen and Coon's method [3] focuses on achieving better control performance by accounting for the time delay, which can significantly impact the system's response. Despite of some limitations, Z-N is one of the most widely used PID tuning method because of its simplicity. A modified Ziegler-Nichols method is proposed in [4] by refining the calculation of ultimate gain, ultimate period and tuning rules. An Internal Mode Control (IMC) based design was proposed in [5].

This seminal work introduced the concept of Internal Model Control (IMC) and provided a comprehensive framework for designing PID controllers based on this innovative control strategy. However, the paper mainly focused on linear systems, and the robustness analysis was limited to linear uncertainties. Skogestad and Postlethwait [6] proposed the modified IMC method, which aimed to achieve the desired closed-loop response by designing an internal model that mimics the process dynamics. The IMC method showed improved disturbance rejection and robustness compared to the Ziegler-Nichols method. A comprehensive reference guide that focuses on providing practical insights and guidelines for tuning proportional-integral-derivative (PID) controllers has been presented in [7]. Detailed review of various PID controller tuning method has been discussed in [8]. In order to reduce the time and knowledge of tuning process, Åström and Hägglund [9] proposed method that focuses on the automatic tuning of simple regulators, likely referring to control systems that use basic feedback control techniques like proportional, integral, and derivative (PID) control. The main emphasis of the paper is on achieving desired phase and amplitude margins for stability and performance of control systems. Author went further to explore more of autotuning method and presented an experimental comparison of various PID autotuners. The objective of this study is likely to evaluate the performance and effectiveness of different autotuning methods in real-world scenarios [10].

In addition to aforementioned methods for tuning PID controllers, there are numerous other. Some of the popular tuning methods are Model-based Tuning, 'Trial and error', Optimization Techniques etc. However, in many real-world scenarios, the system's behavior may be complex, uncertain, or difficult to model accurately. In such cases, expert knowledge from experienced operators or domain experts can be valuable in improving the control performance.

Fuzzy logic has been used to incorporate the expert knowledge for the PID controller tuning. Uçak. K introduced the concept of fuzzy and proposed a novel adaptive multi-input multi-output (MIMO) fuzzy PID controller for time delay systems, building upon prior work on single-input single-output (SISO) system. The study evaluates the controller's performance in stabilization, tracking, and disturbance rejection against classical PID controllers. Results demonstrate the effectiveness of the proposed adaptation mechanism, suggesting its successful application in delay systems [11]. An another article showing the usefulness of integrated fuzzy-PI/PID control has been authored by Demirtas and Papanikolopoulos [12]. In this, AC voltage controller capable of operating at different power factors presents a power factor correction (PFC) scheme using various controllers (PI, fuzzy logic PI, and fractional order PI).

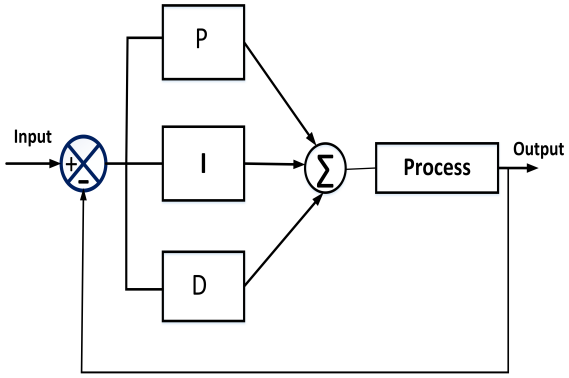


Figure 1. PID control system.

A single-phase boost converter is modeled in MATLAB/Simulink, and a filter is designed to minimize THD. The proposed model demonstrate the combination of fuzzy and PI controllers achieves the best power factor control. Tzafestas and Papanikopoulos [13] were among the first who introduced concept of Fuzzy-tuned PID controller design.

They suggested enhancing the performance of a closed-loop system by making adjustments to PID parameters. This was accomplished through a fuzzy matrix, encapsulating the operator's experiential insights within a concise rule base. In [14], [15] researchers introduced an auto-tuning algorithm for PID controller using Fuzzy logic. This algorithm aimed to dynamically adjust PID parameters in real-time, utilizing the generated error signal from the closed-loop system as input. An observer and error based adaptive proportional-integral-derivative (PID) controller has been introduced for type-2 fuzzy based system [16].

Numerous other scholars have also adopted a similar type of approach, introducing error to implement expert insights within the PID tuning process. The fuzzy tuned PID controller has some advantages over other tuning methods. Fuzzy logic-based tuning methods provide an effective means of adjusting the gains of a PID controller while minimizing overshoot, settling time, and steady-state error, particularly in nonlinear and complex systems. Contrary to other tuning methods, the fuzzy logic approach does not need an exact mathematical model of the system and can handle non-literariness and uncertainties in the process. The paper introduces a novel approach to controller tuning by combining both fuzzy logic and computational techniques to optimize the PID controller parameters. Fuzzy logic for rule-based decision-making, and interpolation techniques are combined in this work. This represents a unique approach, which has not been encountered in research arena.

A PID controller is a type of feedback control system commonly used in engineering and industrial processes (Figure 1). It continuously measures the difference between a desired set-point and the current value of a controlled variable, and adjusts an output signal to bring the two values closer together. The first term of PID provides an output proportional to the error signal, the integral term sums up the past errors to correct for any steady-state errors, and the derivative term predicts the future error based on the current rate of change. By merging all three parameters, a PID controller can achieve stable and accurate control of a wide range of systems, from simple temperature control to complex manufacturing processes.

In recent years, various tuning methods based upon fuzzy logic have been proposed for optimizing the performance of fuzzy logic controllers. Traditional approaches, such as heuristic tuning and optimization algorithms, often rely on iterative procedures and expert knowledge to define the fuzzy ranges. These methods, while effective, can be time-consuming and may not generalize well to different problem domains. In contrast, the proposed method leverages interpolation to set the fuzzy ranges, providing a unique and efficient alternative to conventional tuning techniques. This approach simplifies the tuning process by reducing the dependency on expert intervention. Moreover, it ensures controller parameters remain within specified bounds, a crucial consideration for safety and practicality in real-world applications. A third-order reactor plant has been taken into consideration as the system. Aim of this work is to maintain the concentration of effluent in each at a certain desired level. Simulation has been done using MATLAB software to get the desired response of the system subjected to step input. The results were observed and compared with few existing relevant literature.

1.1. Tuning of PID controller

PID controller tuning refers to the process of adjusting the parameters of the controller (proportional gain, integral time constant, and derivative time constant) to achieve the desired performance of the control system. The process of tuning involves adjusting the parameters based on the response of the system to different inputs, and is done through many methods.

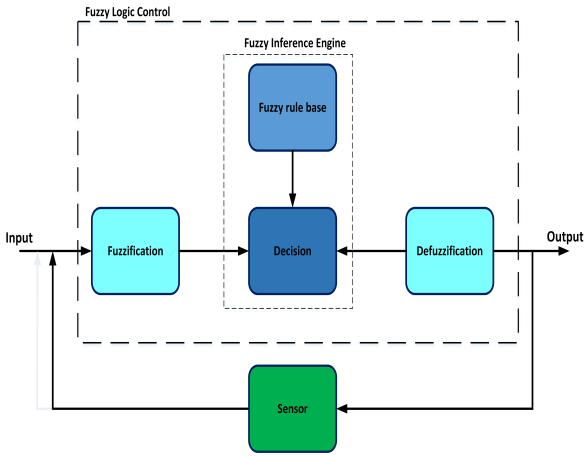


Figure 2. General Structure of Fuzzy Logic Control System

The importance of tuning a PID controller is rooted in its crucial role in ensuring the effective and efficient functioning of a control system. If the controller is not tuned correctly, the system may be unstable or oscillate, leading to inefficient operation or even damage to the system. On the other hand, a well-tuned PID controller can help maintain stable control of the process variable, reduce overshoot, improve settling time, and improve the overall performance of the system.

1.2. Fuzzy-tuned PID controller

Fuzzy logic is mathematical tool to deal with uncertainty in the system. In traditional binary logic, propositions are either true or false, but in fuzzy logic, propositions can be partially true or partially false, and the degree of truth or falsity is expressed using a range of values between 0 and 1. Primary benefit of fuzzy logic is its knowledge, that is efficient to handle imperfect and ambiguous data. Fuzzy logic is appropriate for modeling complex systems that are difficult to describe using traditional mathematical models. It allows for the use of linguistic variables, which can be more intuitive and easier to understand than traditional mathematical models. A typical Fuzzy logic control structure is depicted in Figure 2.

It takes crisp value as inputs variables. After fuzzification, processing with Fuzzy Inference Engine (FIE) and defuzzification, again desired crisp output is obtained. Fuzzy knowledge base together with decision form FIE in the control structure. A fuzzy tuned PID controller is a type of proportional-integral-derivative (PID) controller that uses fuzzy logic to tune its parameters. PID controllers are widely used in control systems to maintain a desired set-point by adjusting the output as per the difference between the set-point and the measured process variable. However, the performance of a PID controller is highly dependent on its tuning parameters, which can be difficult to determine for complex systems. Fuzzy tuned PID controllers use fuzzy logic to determine the optimal tuning parameters for the PID controller based on the current state of the system. This approach can be more effective than traditional PID tuning methods because it takes into account the complexity and uncertainty of the system being controlled. Additionally, fuzzy tuned PID controllers can adapt to changing conditions in the system, making them more robust and effective in real-world applications.

2. System description

A chemical process control system typically involves monitoring and adjusting the variables in a chemical reaction to optimize the output. In such a control industry it is required commonly to maintain the concentration of solute in a solvent. In the proposed work, the n-numbers of cascaded tanks are arranged in series, with the first tank receiving the solvent and the solute as shown in Figure 3. The output of the first tank is fed into the second tank, where further mixing occurs. The output of second tank is then fed into the third tank, where further mixing occurs. This process continues up to nth tank of the system where final mixing occurs. The concentration of solute in the solvent in final tank is measured and the signal is sent to controller, which adjusts the flow rate of the solute feed to the fist tank in order to maintain the desired concentration.

2.1. Modeling of the system

In order to obtain simple mathematical model of the aforementioned system, volume-flow rate and corresponding volumes are assumed to be constant. Writing mass-balance equation for first two tank.

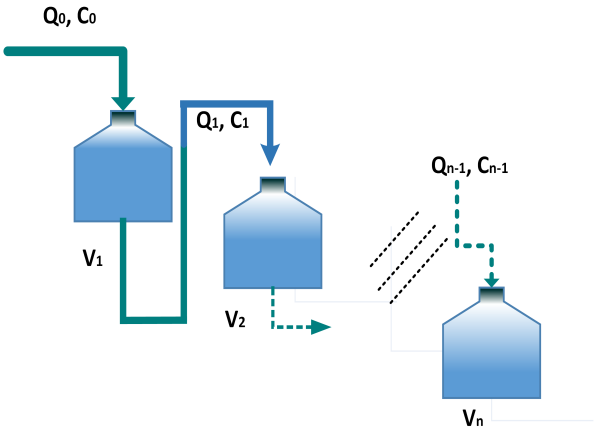


Figure 3. Cascaded 'n' number of tanks.

$$T_1 \frac{dC_1}{dt} = C_0 - C_1, \tag{1}$$

$$T_2 \frac{dC_2}{dt} = C_1 - C_2, \tag{2}$$

where, $T_1 = V_1/q0$ and $T_2 = V_2/q0$. In the above equation V, C and $q0$ represents volume, concentration and inlet flow rate of the both the tanks in proper unit with subscript number denoting the tank number. Transforming the above equations in Laplace domain yields.

$$\frac{C_1}{C_0} = \frac{1}{T_1 s + 1}, \tag{3}$$

$$\frac{C_2}{C_1} = \frac{1}{T_2 s + 1}. \tag{4}$$

The effect of C_0 on C_2 can be computed as:

$$\frac{C_2}{C_0} = \frac{1}{T_1 s + 1} \times \frac{1}{T_2 s + 1}. \tag{5}$$

For 'n' number of tanks in series, the generalized equation can be expressed as;

$$\frac{C_n}{C_0} = \frac{1}{(T_1 s + 1)(T_2 s + 1).....(T_n s + 1)}. \tag{6}$$

Order of the system described by the equation (6) will vary with the value of 'n' or number of tanks. for the proposed work a 3-tank process control system has been considered, where, values of T1, T2 and T3 is assumed to be 1,3 and 5 respectively. Transfer function of the aforementioned system can be written as:

$$G(s) = \frac{1}{(S+1)} \times \frac{1}{(S+3)} \times \frac{1}{(S+5)}. \tag{7}$$

Subsequently, rewritten as:

$$G(s) = \frac{1}{(S+1)(0.33S+1)(0.2S+1)}. \tag{8}$$

2.2. Design of fuzzy tuned PID controller

The solute concentration in the solvent is measured continuously and the signal is sent back through feedback to a fuzzy PID (Proportional-Integral-Derivative) controller. The fuzzy PID controller uses a set of rules and linguistic variables to adjust the flow rate of the solute feed to the first tank, in order to maintain the desired solute concentration in the solvent. The controller takes into account factors the current concentration error the rate of change of the error. The proposed structure of Fuzzy-PID control has been shown in Figure 4.

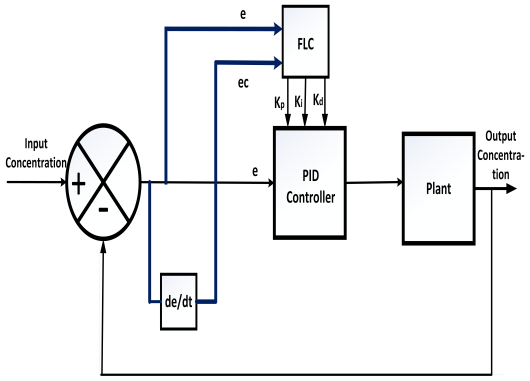


Figure 4. Fuzzy-PID controller for a chemical mixing plant

The design process involves tuning of PID controller parameters with the help of Fuzzy logic controller, which takes 'error(e)' and 'rate of change of error(ec)' as inputs and provide Kp, Ki and Kd as output. Hence, there are total of five linguistic variables. For each of these variables total of seven fuzzy values (NL,NA,NT,ZE,PT,PA,PL) have been chosen. The range of the 'e' and 'ec' are between -3 and 3, whereas, by formulating a kind of interpolation, fuzzy values for Kp,Ki and Kd are kept between 0 and 1. Aim of performing interpolation is to keep the values of the variable within some range. In the following equations, interpolation has been applied to keep the value of parameters (Kp, Ki and Kd) between 0 to 1. As the result, parameters are scaled to a common range, enabling a consistent, normalized and controlled analysis.

The determination of these ranges involves the following two steps:

- **Identification of Extreme Values:** The maximum and minimum values of all the three parameters of controller are identified.
- **Normalization Process** Once the minimum and maximum values are identified, the values of controller parameter are normalized to the range [0, 1] using the following equations:

$$K_p = \frac{K_{p,0} - K_{p,\min}}{K_{p,\max} - K_{p,\min}} \tag{9}$$

$$K_i = \frac{K_{i,0} - K_{i,\min}}{K_{i,\max} - K_{i,\min}} \tag{10}$$

$$K_d = \frac{K_{d,0} - K_{d,\min}}{K_{d,\max} - K_{d,\min}} \tag{11}$$

$K_{p,0}$, $K_{i,0}$ and $K_{d,0}$ are initially estimated values of K_p, K_i and K_d respectively. Fuzzy rule base for the three PID parameters are shown in Table 1, 2 and 3.

Table 1. Rule base for proportional gain, K_p

e/ec	NL	NA	NT	ZE	PT	PA	PL
NL	PL	PL	PA	PA	PT	ZE	ZE
NA	PL	PL	PA	PT	PT	NT	NT
NT	PA	PA	PA	PT	ZE	NT	NT
ZE	PA	PA	PT	ZE	NT	NA	NA
PT	PT	PT	ZE	NT	NT	NA	NA
PA	PT	ZE	NT	NA	NA	NL	NL
PL	PT	ZE	NA	NA	NA	NL	NL

Table 2. Rule base for proportional gain, K_i

e/ec	NL	NA	NT	ZE	PT	PA	PL
NL	NL	NL	NA	NA	NT	ZE	ZE
NA	NL	NL	NA	NT	NT	ZE	ZE
NT	NL	NA	NT	NT	ZE	PT	PT
ZE	NA	NA	NT	ZE	PT	PA	PA
PT	NA	NT	ZE	PT	PT	PA	PL
PA	ZE	ZE	PT	PT	PA	PL	PL
PL	ZE	ZE	PT	PA	PA	PL	PL

Table 3. Rule base for proportional gain, K_d

e/ec	NL	NA	NT	ZE	PT	PA	PL
NL	PT	NT	NL	NL	NL	NA	PT
NA	PT	NT	NL	NA	NA	NT	ZE
NT	ZE	NT	NA	NA	NT	NT	ZE
ZE	ZE	NT	NT	NT	NT	NT	ZE
PT	ZE	ZE	ZE	ZE	ZE	ZE	ZE
PA	PL	NT	PT	PT	PT	PT	PL
PL	PL	PA	PA	PA	PT	PT	PL

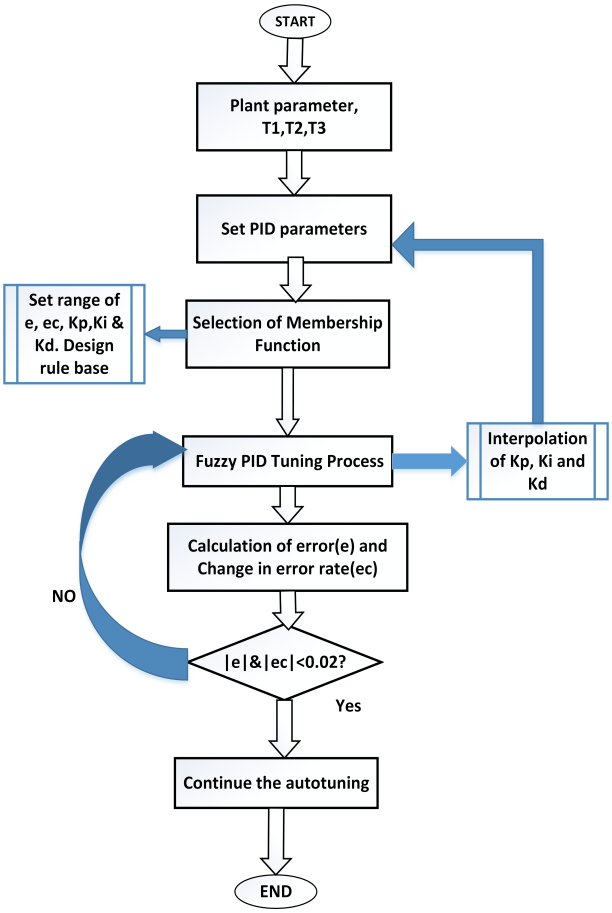
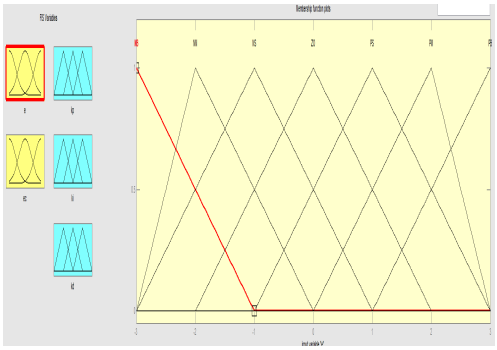


Figure 5. Flow chart of the control action.

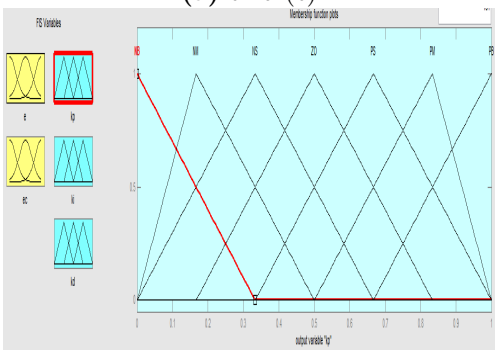
2.3. FIS computational model and control action

Entire scheme of control action has been shown in the form of flow chart in the Figure 5. A fuzzy inference engine (FIE) employs fuzzy logic to approximate human reasoning/experience. Two input "error(e)" and "rate of change of error(ec))" have been taken as input variables as well as K_p, K_i and K_d as output. Triangular membership function has been chosen for both input and output variables. For the illustration purpose

range of the membership function for 'e' and 'K_p' have been shown in the Figure 6.



(a) error(e)

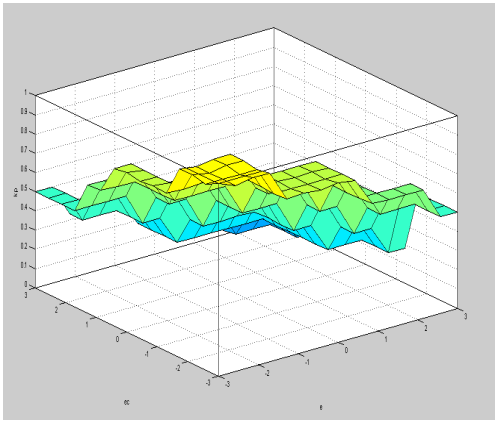


(b) K_p

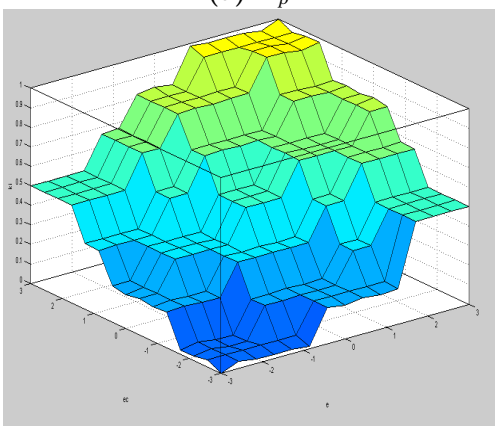
Figure 6. Membership function plots with its range for (a)error(e) and (b) proportionality constant (K_p)

Fuzzy logic uses linguistic variables, which are variable that take on values in a fuzzy set. On the basis of these variables linguistic rules (rule base) are made that plays a crucial role in the interpretation of fuzzy logic systems by defining the meaning of linguistic variables and interpreting the output in a way that is understandable to humans.

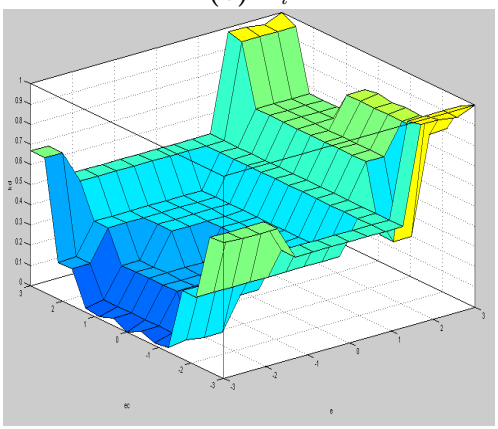
A number of such rules are made to provide the direction to controller for appropriate action. With seven assigned membership values to each variable, a total of 49 rules base have been proposed to completely describes the control action. These rules for PID coefficients (K_p, K_i and K_d) are summarized in the form



(a) K_p



(b) K_i



(c) K_d

Figure 7. Surface view of (a) K_p (b)K_i and (c) K_d

of three tables. The variations of these with respect to inputs (e and ec) are shown as the form of surface view in Figure 7.

Table 4. Comparison in terms of few performance indicators of different methods.

Method	T_r	T_s	$\%M_p$	$ITAE$
Salem	9	10	0.2	0.9405
Z-N	0.5	5	60	7.905
CA	0.8	1.6	2	50
Proposed	1.7	2	0.8	0.63

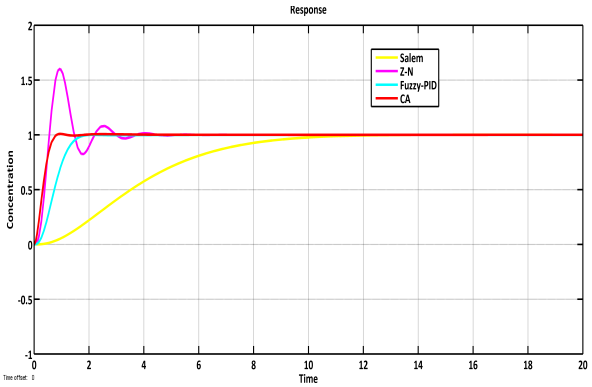


Figure 8. Response of system without any disturbance.

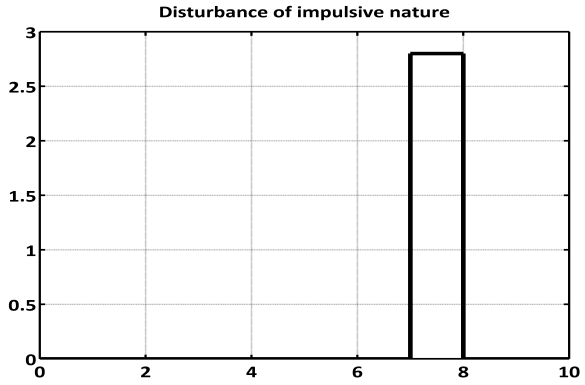


Figure 9. Disturbance to the system

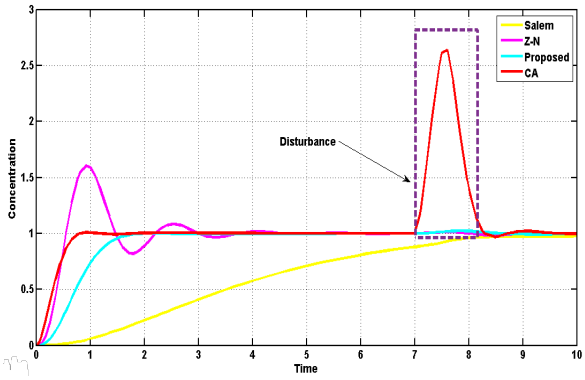


Figure 10. Response of system with some disturbance.

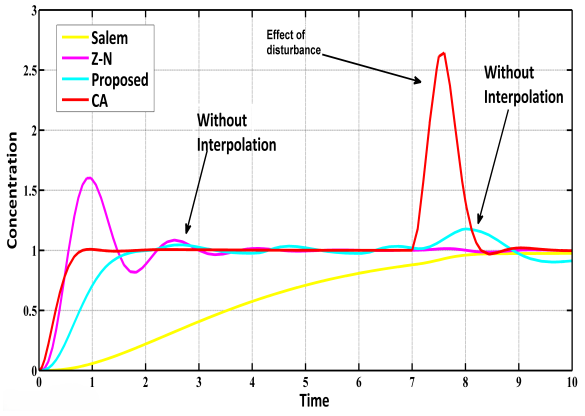


Figure 11. Response of system without using interpolation.

3. Results and discussions

The proposed Fuzzy-PID process control system is designed to maintain the concentration of solute in the tanks. With the step input(concentration of solute), proposed system is analyzed considering two cases, first assuming no disturbance on the system, secondly taking disturbances into the account at seventh second in given test time. The function of the system is to track this step input. Again, in chemical process control systems, disturbances can occur due to fluctuations in feed composition, temperature variations, or unexpected reactions. These disturbance are usually of impulsive nature . In the proposed work, disturbance that is taken for account, is of periodic impulsive nature, whose amplitude and time period have been considered of 2.8 unit and 10 seconds respectively, while pulse width is 10 percent of time period as shown in Fig. 9 . The fuzzy logic system continuously monitors the system's performance under such disturbances and adjusts the PID parameters accordingly. The interpolation mechanism allows for smooth adjustments, preventing abrupt changes that could destabilize the system. A total of four performance specification parameters have been considered analyzing the working of the system. A step input is applied to the proposed system and compared with Ziegler-Nicolas(Z-N) method [1], model-based design in [7] and computational approach (CA). response of the proposed system is shown in Figure 8. It can be observed that PID-tuning by Z-N method has highest overshoot and settle to steady state in approximately 5 second, when system is subjected to step input. Model based approach proposed by *Salem* [14] is quite sluggish as rise time of the response is 9 seconds. Computational approach(CA) [9] has somewhat satisfactory results in terms of rise time, settling time and maximum overshoot. However, CA has poor "disturbance rejection" capacity having *Integral time absolute error(ITAE)* nearly equal to 50, that result in spike(shown by the arrow) when subjected to disturbance of impulsive nature at 7th second as shown in Figure 10. Proposed system qualify in all the four performance specification parameter and shows satisfactory performance even at the onset of sudden disturbance. It has fast response, low settling time, low overshoot and good disturbance rejection capacity. Values of performance parameters of different methods are tabulated in Table 4. It is also worth mentioning the importance of interpolation technique in settling the system response specially under the

effect of disturbance which is occurring at the 7th second of operation. Without interpolation, disturbance try to hinder the tracking of desired value in the proposed method as shown in Fig. 11. Effect can also be seen at the 2.5 sec, when there is slight overshoot causing unsettling of system-response. Thus interpolation mechanism used in the self-tuning Fuzzy PID controller is achieving smooth the PID parameters. By interpolating between predefined range, the controller can fine-tune its responses to varying process conditions more effectively. Therefore, synergetic approach introduced in this work enhances the self-tuning Fuzzy PID controller by combining the strengths of fuzzy logic and traditional PID control. This combination provides a dual benefit: fuzzy logic handles the nonlinearity and uncertainty in the process, while the PID control ensures precise regulation. The Synergetic approach's particularly useful scenarios with high process variability, where it maintains control performance and robustness.

4. Conclusions

In this study, a interpolation-enhanced Fuzzy-PID controller has been designed to control the concentration of a solute in a process control system. The proposed synergistic approach has been applied to a chemical process control system consisting of three tanks, aimed at maintaining a constant effluent concentration, has proven to be successful. This method has demonstrated the ability to attain the desired performance in the closed-loop system by dynamically adjusting the controller parameters, K_p, K_i and K_d. Performance of Fuzzy-PID controller has been assessed for third order chemical process system. Aim of the controller function is to regulate the level of concentration of solute in the solution in three tanks. A comprehensive set of simulation and case studies have been conducted to illustrate the versatility and robustness of the fuzzy PID controller across a range of chemical processes. It has been observed that proposed system seems is working satisfactorily even in the case of sudden disturbance and it tries to maintain the level of solute at pre-defined values in all the tanks. The simulation results indicate that the suggested approach outperforms the conventional PID controller when it comes to handling disturbances and accurately following set-points. The fuzzy PID controller display improved performance in the presence of uncertainties and disturbances. The result was compared with few existing literature, proposed system seems to be working better among all. The Synergetic methods developed in this work have potential applications in various industrial domains where precise and adaptable process control is critical.

References

[1] Ziegler, J. G & Nichols, N.B. (1942). Optimum settings for automatic controllers: The new global context. *Transactions of the American society of mechanical engineers*, 64(8), 759-765. https://doi.org/10.1115/1.4019264

[2] Chien, K. L., Hrones, J. A. & Reswick, J. B. (1952). Optimum settings for automatic controllers: On the automatic control of generalized passive systems. *Transactions of the American Society of Mechanical Engineers*, 74(2), 175-183. https://doi.org/10.1115/1.4015724

[3] Cohen, G.H. & Coon, G. A. (1953). Optimum settings for automatic controllers: On the automatic control of generalized passive systems. *Transactions of the American Society of Mechanical Engineers*, 75(5), 827-834. https://doi.org/10.1115/1.4015452

[4] Åström, K. J. & Hägglund, T. (2004). Revisiting the Ziegler-Nichols step response method for PID control. *Journal of process control*, 14(6), 635-650. https://doi.org/10.1016/j.jprocont.2004.01.002

[5] Rivera, D. E., Morari, M., & Skogestad, S. (1986). Internal model control: PID controller design. *Industrial & Engineering Chemistry Process Design and Development*, 25(1), 252-265. https://doi.org/10.1021/i200032a041

[6] Skogestad, S. & Postlethwaite, I. (2005). *Multivariable Feedback Control: Analysis and Design*. John Wiley & sons.

[7] O'dwyer, A. (2009). *Handbook of PI and PID Controller Tuning Rules*. World Scientific. https://doi.org/10.1142/9781848162433

[8] Borase, R. P., Maghade, D. K., Sondkar, S. Y. & Pawar, S. N. (2021). A review of PID control, tuning methods and applications. *International Journal of Dynamics and Control*, 9, 818-827. https://doi.org/10.1007/s40435-020-00665-4

[9] Åström, K. J., & Hägglund, T. (1984). Automatic tuning of simple regulators with specifications on phase and amplitude margins. *Automatica*, 20(5), 645-651. https://doi.org/10.1016/0005-1098(84)90014-1

[10] Berner, J., Soltesz, K., Hägglund, T., & Åström, K. J. (2018). An experimental comparison of PID autotuners. *Control Engineering Practice*, 73, 124-133. https://doi.org/10.1016/j.conengprac.2018.01.006

[11] Uçak, K., & Arslantürk, B. N. (2023). Adaptive MIMO fuzzy PID controller based on peak observer. *International Journal of Optimization & Control: Theories & Applications (IJOCTA)*, 13(2), 139-150. https://doi.org/10.11121/ijocta.2023.1247

[12] Demirtas, M., & Ahmad, F. (2023). Fractional fuzzy PI controller using particle swarm optimization to improve power factor by boost converter. *An International Journal of Optimization and Control: Theories& Applications (IJOCTA)*, 13(2), 205-213. https://doi.org/10.11121/ijocta.2023.1260

[13] Tzafestas, S. & Papanikolopoulos, N. P. (1990). Incremental fuzzy expert PID control. *IEEE Transactions on Industrial Electronics*, 37(5), 365-371. https://doi.org/10.1109/41.103431

[14] Salem, F. A. (2013). New efficient model-based PID design method. *European Scientific Journal*, 9(15), 181-199.

[15] Mirrashid, N., Alibeiki, E., & Rakhtala, S. M. (2022). Development and control of an upper limb rehabilitation robot via ant colony optimization-PID and fuzzy-PID controllers. *International Journal of Engineering*, 35(8), 1488-1493. https://doi.org/10.5829/IJE.2022.35.08B.04

[16] Yesil, E., Kumbasar, T., Dodurka, M. & Sakalli, A. (2014). Peak observer based self-tuning of type- 2 fuzzy PID controller. 10th IFIP International Conference on Artificial Intelligence Applications and Innovations (AIAI), Rhodes, Greece, 487- 497. https://doi.org/10.1007/978-3-662-44654-6_48

Devashish Jha *received his B.Tech degree from silicon Institute of Technology Bhubaneswar, M.Tech degree from the National Institute of Technology Patna and Ph.D. from the National Institute of Technology Jamshedpur. He is currently working as an Assistant Professor in the Department of Electronics and Communication Engineering at the Madanapalle Institute of Technology and Science, Andhra Pradesh, India. His research interests primarily focus on control applications in renewable energy. He has contributed significantly to interdisciplinary area of research and has authored several research papers in reputed journals and conferences.*

https://orcid.org/0000-0002-5616-750X

Arifa Ahmed *received the B.Tech. degree in Electronics and Communication Engineering from North-Eastern Hill University (NEHU), Shillong, in 2011 and the M.Tech. degree from the same department, North Eastern Regional Institute of Science and Technology (NERIST), Itanagar, in 2014. She completed her Ph.D. degree in Electronics and Communication Engineering at National Institute of Technology (NIT) Silchar, India in 2022. She is currently working as an Assistant Professor in the Department of Electronics and Communication Engineering, Madanapalle Institute of Technology and Science, Andhra Pradesh, India. Her research interests include wireless communications, wireless networking, cognitive radio, and optimization.*

https://orcid.org/0000-0002-6695-9973

Sanatan Kumar *received his M.Tech degree from the National Institute of Technical Teachers' Training & Research, Bhopal and earned his Ph.D. from the National Institute of Technology Jamshedpur. He is currently working as an Assistant Professor in the Department of Electrical and Electronics Engineering, Technocrat Institute of Technology, Bhopal, India India. His research interests primarily focus on Power Electronics and Control.*

https://orcid.org/0000-0002-5054-9654

Debanjan Roy *graduated in Electrical Engineering from Future Institute of Engineering and Management, Sonarpur, under West Bengal University of Technology (WBUT) in the year 2010. He subsequently pursued his post-graduation in Power Electronics and Drives from Kalinga Institute of Industrial Technology Deemed to be University, Bhubaneswar, graduating in 2014. He received Ph.D. degree from National Institute of Technology Jamshedpur, India. Currently, he holds the position of an Assistant Professor in the Department of Electrical Engineering at Teerthanker Mahaveer University, Moradabad, India. Dr. Debanjan Roy has contributed significantly to his field and has authored over 8 research papers in areas such as Multilevel inverters, Space vector pulse width modulation, Advanced PWM techniques, and their applications in motor drive systems. He is a dedicated member of the Institute of Electrical and Electronics Engineers and also holds an associate (Life) membership with the Indian Society of Lighting Engineers.*

https://orcid.org/0000-0002-6886-0598

This work is licensed under a Creative Commons Attribution 4.0 International License. The authors retain ownership of the copyright for their article, but they allow anyone to download, reuse, reprint, modify, distribute, and/or copy articles in IJOCTA, so long as the original authors and source are credited. To see the complete license contents, please visit http://creativecommons.org/licenses/by/4.0/.

An International Journal of Optimization and Control: Theories & Applications
ISSN:2146-0957 eISSN:2146-5703
Vol.14, No.4, pp.365-377 (2024)
http://doi.org/10.11121/ijocta.1558

RESEARCH ARTICLE

Global mathematical analysis of a patchy epidemic model

Lahcen Boulaasair,[a] Hassane Bouzahir,[a] Mehmet Yavuz[b,c,*]

[a] *ISTI Lab, Ibn Zohr University, ENSA, PO Box 1136, Agadir, Morocco*
[b] *Necmettin Erbakan University, Department of Mathematics and Computer Sciences, Meram Yeniyol, 42090 Meram, Konya, Türkiye*
[c] *Centre for Environmental Mathematics, Faculty of Environment, Science and Economy, University of Exeter, Cornwall TR10 9FE, UK*
lahcen.boulaasair@edu.uiz.ac.ma, hbouzahir@yahoo.fr, mehmetyavuz@erbakan.edu.tr

ARTICLE INFO

Article History:
Received 6 March 2024
Accepted 10 July 2024
Available Online 14 October 2024

Keywords:
Disease dynamics
Two-patch epidemic model
Stochastic analysis
Deterministic stability
Exponential disease extinction
Persistence of the disease
Numerical simulation

AMS Classification 2010:
34K50; 60H30; 65C30; 92B05; 92D25

ABSTRACT

The dissemination of a disease within a homogeneous population can typically be modeled and managed in a uniform fashion. Conversely, in non-homogeneous populations, it is essential to account for variations among subpopulations to achieve more precise predictive modeling and efficacious intervention strategies. In this study, we introduce and examine the comprehensive behavior of a deterministic two-patch epidemic model alongside its stochastic counterpart to assess disease dynamics between two heterogeneous populations inhabiting distinct regions. First, utilizing a specific Lyapunov function, we demonstrate that the disease-free equilibrium of the deterministic model is globally asymptotically stable. For the stochastic model, we establish that it is well-posed, meaning it possesses a unique positive solution with probability one. Subsequently, we ascertain the conditions necessary to ensure the total extinction of the disease across both regions. Furthermore, we explicitly determine a threshold condition under which the disease persists in both areas. Additionally, we discuss a scenario wherein the disease persists in one region while simultaneously becoming extinct in the other. The article concludes with a series of numerical simulations that corroborate the theoretical findings.

(cc) BY

1. Introduction

Infectious diseases are defined as illnesses caused by pathogenic agents, transmitted from an infected person, animal, or contaminated inanimate object to a susceptible host [1]. They are the main cause death worldwide killing more people than all wars and natural disasters combined [2,3]. For instance, during the past three years, the world has been under enormous threat from the highly contagious coronavirus which first emerged in China and has spread rapidly to cover almost the entire globe leading to the death of more than six million people, according to the statistics of the World Health Organization [4]. In addition to the human casualties from the coronavirus, the economic and social disruption caused by this pandemic is devastating. Millions of people at risk of crossing the poverty line, thousands of companies face an existential threat and almost 50% of the global workforce, comprising 3.3 billion individuals, faces the threat of unemployment [5].

To understand how infectious diseases spread, the mathematical models are useful tools to describe and simulate concrete situations for anticipating their future behaviour. Most models for the transmission of infectious diseases descend from the classical SIR model of Kermack and McKendrick established in 1927 [6]. Such model is called a compartmental model where the population is divided into compartments of susceptible, infected, and recovered. A non-linear ordinary differential

*Corresponding Author

equations are used to model the dynamics between these compartments.

A major criticism for this model and the models that followed (for example [7–15]), is that the total population is assumed to be entirely homogenous and all individuals behave the same. Therefore, the model may not represent complex mobility and contact patterns for many real-world diseases. To overcome this inconvenience, Calvo et al. [16] have incorporated population heterogeneity to examine interactions between urban and rural populations on the dynamics of disease spreading by using a compartmental framework of susceptible–infected–susceptible dynamics with some level of immunity. The proposed model is as follows:

$$\begin{aligned}
\frac{dS_1}{dt} =& \mu_1 N_1 + \beta_1 \frac{I_1}{N_1} S_1 - \mu_1 S_1 + \delta_{21} S_2 - \delta_{12} S_1,\\
\frac{dI_1}{dt} =& \beta_1 \frac{I_1}{N_1} S_1 - (\mu_1 + \gamma_1) I_1 + \rho_1 \frac{I_1}{N_1} R_1 + \delta_{21} I_2 - \delta_{12} I_1,\\
\frac{dR_1}{dt} =& \gamma_1 I_1 - \rho_1 \frac{I_1}{N_1} R_1 - \mu_1 R_1 + \delta_{21} R_2 - \delta_{12} R_1,\\
\frac{dS_2}{dt} =& \mu_2 N_2 + \beta_2 \frac{I_2}{N_2} S_2 - \mu_2 S_2 + \delta_{12} S_1 - \delta_{21} S_2,\\
\frac{dI_2}{dt} =& \beta_2 \frac{I_2}{N_2} S_2 - (\mu_2 + \gamma_2) I_2 + \rho_2 \frac{I_2}{N_2} R_2 + \delta_{12} I_1 - \delta_{21} I_2,\\
\frac{dR_2}{dt} =& \gamma_2 I_2 - \rho_2 \frac{I_2}{N_2} R_2 - \mu_2 R_2 + \delta_{12} R_1 - \delta_{21} R_2.
\end{aligned} \tag{1}$$

The subscript 1 is used to denote the urban parameters and variables, and the subscript 2 for the rural parameters and variables. For $i \in \{1, 2\}$, S_i, I_i, R_i and N_i denote the numbers of susceptible, infected, post-recovery susceptible individuals and the total population, respectively. The parameter μ_i is the rate of birth and death. β_i is the infection transmission coefficient between susceptible and infected individuals. The post-recovery susceptible individuals are infected at rate ρ_i, while infected individuals become post-recovery susceptibles at rate γ_i. The motion between urban and rural populations is modeled by the function $\delta_{ij}(t)$ which denotes the fraction of individuals who travel from patch $i \in \{1, 2\}$ to patch $j \in \{1, 2\}$ (with $i \neq j$) at time t. To study the dynamics of system (1), the authors of [16] compute steady states, showing the local stability of the disease-free steady state, and identify conditions for the existence of the endemic steady states.

In the model above, infectious diseases can spread through interactions between the urban and rural populations. Therefore, infected individuals in urban area can infect rural population and the rural infected can transmit the disease to the urban dwellers. In this paper, we assume that there is no immunity. From this perspective, we propose another version of model (1) by introducing four infection transmission coefficients β_i $(i = 1, 2, 3, 4)$, presented as follows:

$$\begin{cases}
dx_1 = & (A_1 - \beta_1 x_1 y_1 - \beta_3 x_1 y_2 - \mu_1 x_1)dt,\\
dy_1 = & (\beta_1 x_1 y_1 + \beta_4 x_2 y_1 - \mu_1 y_1)dt,\\
dx_2 = & (A_2 - \beta_2 x_2 y_2 - \beta_4 x_2 y_1 - \mu_2 x_2)dt,\\
dy_2 = & (\beta_2 x_2 y_2 + \beta_3 x_1 y_2 - \mu_2 y_2)dt.
\end{cases} \tag{2}$$

For the variables x_i and y_i $(i = 1, 2)$, we use the subscript 1 to denote the urban variable and the subscript 2 for the rural one. All the other parameters appearing in model (2) are assumed to be constant positives. The symbols involved in the model are described in Table 1.

On the other hand, the spread of diseases is characterized by randomness due to the unpredictability of the natural behavior [17]. A lot of scholars have introduced the white noise into the deterministic models to reveal the effect of the evironmental fluctuations on the spread of diseases. For example, Cao et al. [18] considered a stochastic SEI epidemic model with saturation incidence and logistic growth. By constructing a suitable Lyapunov function, they established sufficient conditions for the existence and uniqueness of an ergodic stationary distribution of the solutions to the model. They also established sufficient conditions for the extinction of the disease. In [19], Pang et al. discussed the dynamics of a stochastic SIQS epidemic model and investigated the boundedness, extinction and the persistence of the stochastic system. Khan et al. [20] proposed a stochastic model to analyze the dynamics of the novel coronavirus disease. They studied the extinction and the persistence of the disease. For more details on the impact of environmental fluctuations on the spread of diseases and population dynamics, we refer the readers to [21–39].

Based on the aforementioned facts, we substitute $\beta_i dt$ in model (2) by $\beta_i dt + \sigma_i dB_i(t)$, where $B_i(t)$ are mutually independent standard Brownian motions and $\sigma_i > 0$ are the intensities of their corresponding white noises, $i = 1, 2, 3, 4$. All these Brownian Motions are defined on a filtered probability space $\left(\Omega, \mathcal{F}_\Omega, (\mathcal{F}_{\{t\}})_{t\geq 0}, \mathbb{P}\right)$ endowed with a filtration that meets the usual criteria. Thus, we get a stochastic version of the deterministic model (2), defined as follows:

$$\begin{aligned}
dx_1(t) =&(A_1 - \beta_1 x_1(t)y_1(t) - \beta_3 x_1(t)y_2(t) - \mu_1 x_1(t))dt - \sigma_1 x_1(t)y_1(t)dB_1(t) - \sigma_3 x_1(t)y_2(t)dB_3(t),\\
dy_1(t) =&(\beta_1 x_1(t)y_1(t) + \beta_4 x_2(t)y_1(t) - \mu_1 y_1(t))dt + \sigma_1 x_1(t)y_1(t)dB_1(t) + \sigma_4 x_2(t)y_1(t)dB_4(t),\\
dx_2(t) =&(A_2 - \beta_2 x_2(t)y_2(t) - \beta_4 x_2(t)y_1(t) - \mu_2 x_2(t))dt - \sigma_2 x_2(t)y_2(t)dB_2(t) - \sigma_4 x_2(t)y_1(t)dB_4(t),\\
dy_2(t) =&(\beta_2 x_2(t)y_2(t) + \beta_3 x_1(t)y_2(t) - \mu_2 y_2(t))dt + \sigma_2 x_2(t)y_2(t)dB_2(t) + \sigma_3 x_1(t)y_2(t)dB_3(t).
\end{aligned} \tag{3}$$

Here, we assume that the urban susceptibles contamined by the rural infected individuals stay in the rural infectd compartment and rural susceptibles contamined by urban infected people stay in the urban infected class.

For convenience, the abbreviation "a.s." means "almost surely", while $\langle f(t)\rangle = t^{-1}\int_0^t f(r)dr$ is the time average of a continuous function f. For two numbers a and b, the symbols $a \wedge b$ and $a \vee b$ stand for the minimum and the maximum of a and b, respectively.

The rest of the paper proceeds as follows. In the next section, we study the stability of the equilibrium state $E = (\frac{A_1}{\mu_1}, 0, \frac{A_2}{\mu_2}, 0)$ for the deterministic model (2). Section 3 is devoted to verify if the stochastic model (3) has a unique positive solution with probability one. In section 4, the conditions ensuring the exponential extinction of the disease in both patches are established. Afterwards, we will carry out an analysis leading to defining a threshold for the disease to persist completely. There remains a case where the disease persists in one patch, and disappears in the other, which is the main theme of the section 6. The paper ends with the realization of numerical simulations using the software Matlab 2015b.

2. Stability of the deterministic model

The aim of mathematical modeling of the spread of epidemics is to know the conditions under which the epidemic dies out. The deterministic model (2) has one free-disease equilibrium $E = (\frac{A_1}{\mu_1}, 0, \frac{A_2}{\mu_2}, 0)$.

The following theorem gives sufficient conditions for local and global asymptotic stability of the free-disease equilibrium E.

Theorem 1. *If* $\Big(\big(\beta_1\frac{A_1}{\mu_1} + \beta_4\frac{A_2}{\mu_2} - \mu_1\big) \vee \big(\beta_3\frac{A_1}{\mu_1} + \beta_2\frac{A_2}{\mu_2} - \mu_2\big)\Big) < 0$, *then the equilibrium E is locally asymptotically stable.*

Proof. The Jacobian matrix related to model (2) at the equilibrium E is

$$J(E) = \begin{pmatrix} -\mu_1 & -\beta_1\frac{A_1}{\mu_1} & 0 & -\beta_3\frac{A_1}{\mu_1} \\ 0 & \beta_1\frac{A_1}{\mu_1} - \mu_1 + \beta_4\frac{A_2}{\mu_2} & 0 & 0 \\ 0 & -\beta_4\frac{A_2}{\mu_2} & -\mu_2 & -\beta_2\frac{A_2}{\mu_2} \\ 0 & 0 & 0 & \beta_2\frac{A_2}{\mu_2} - \mu_2 + \beta_3\frac{A_1}{\mu_1} \end{pmatrix}.$$

According to the Hurwitz criterion, if $\Big(\big(\beta_1\frac{A_1}{\mu_1} + \beta_4\frac{A_2}{\mu_2} - \mu_1\big) \vee \big(\beta_3\frac{A_1}{\mu_1} + \beta_2\frac{A_2}{\mu_2} - \mu_2\big)\Big) < 0$, then the eigenvalues of matrix $J(E)$ are all negatives.

Thus, the equilibrium state E is locally asymptotically stable. □

Theorem 2. *If* $\Big(\big((\beta_1 + \beta_4)\frac{A_1+A_2}{\mu_1\wedge\mu_2} - \mu_1\big) \vee \big((\beta_2 + \beta_3)\frac{A_1+A_2}{\mu_1\wedge\mu_2} - \mu_2\big)\Big) < 0$, *then the equilibrium E is globally asymptotically stable.*

Proof. Consider the Lyapunov function $\mathcal{V}$ defined by

$$\mathcal{V}(x_1(t), y_1(t), x_2(t), y_2(t)) = \frac{1}{2}(y_1^2 + y_2^2).$$

The derivative of $\mathcal{V}$ along the trajectories of solution of model (2) is as follows

$$\begin{aligned}
&\frac{d\mathcal{V}(x_1(t), y_1(t), x_2(t), y_2(t))}{dt}\\
&= y_1(t)(\beta_1 x_1(t)y_1(t) + \beta_4 x_2(t)y_1(t) - \mu_1 y_1(t))\\
&+ y_2(t)(\beta_2 x_2(t)y_2(t) + \beta_3 x_1(t)y_2(t)\\
&- \mu_2 y_2(t))\\
&= y_1^2(t)(\beta_1 x_1(t) + \beta_4 x_2(t) - \mu_1)\\
&+ y_2^2(t)(\beta_2 x_2(t) + \beta_3 x_1(t) - \mu_2)\\
&\leq y_1^2(t)\Big((\beta_1 + \beta_4)\frac{A_1 + A_2}{\mu_1 \wedge \mu_2} - \mu_1\Big)\\
&+ y_2^2(t)\Big((\beta_2 + \beta_3)\frac{A_1 + A_2}{\mu_1 \wedge \mu_2} - \mu_2\Big).
\end{aligned}$$

Assuming that $\Big(\big((\beta_1 + \beta_4)\frac{A_1+A_2}{\mu_1\wedge\mu_2} - \mu_1\big) \vee \big((\beta_2 + \beta_3)\frac{A_1+A_2}{\mu_1\wedge\mu_2} - \mu_2\big)\Big) < 0$, we get

$$\frac{d\mathcal{V}(x_1(t), y_1(t), x_2(t), y_2(t))}{dt} < 0 \quad \text{for any } t \geq 0,$$

which means that E is globally asymptotically stable. □

Table 1. Description of symbols in model (2).

Parameter	Meaning
x_i	The number of susceptible individuals to the disease, where $i = 1, 2$.
y_i	The number of infective members.
A_i	A constant input of new members into the population i per unit time.
μ_i	Natural death rate of x_i and y_i.
β_1	Transmission coefficient between x_1 and y_1.
β_2	Transmission coefficient between x_2 and y_2.
β_3	Transmission coefficient between x_1 and y_2.
β_4	Transmission coefficient between x_2 and y_1.

3. Well-posedeness of the stochastic model

Lemma 1. *The set* $\Gamma = \Big\{(x_1(t), y_1(t), x_2(t), y_2(t)) \in \mathbb{R}^4_+ : N(t) = x_1(t)+y_1(t)+x_2(t)+y_2(t) \leq \frac{A_1+A_2}{\mu_1\wedge\mu_2}\Big\}$ *is positively invariant for the stochastic model (3).*

Proof. From system (3), we have

$$\begin{aligned} dN(t) =& \big(A_1 + A_2 - \mu_1(x_1(t) + y_1(t)) \\ & - \mu_2(x_2(t) + y_2(t))\big)dt \\ \leq& (A_1 + A_2 - (\mu_1 \wedge \mu_2)N(t))dt. \end{aligned} \tag{4}$$

Thus

$$N(t) \leq \frac{A_1 + A_2}{\mu_1 \wedge \mu_2} + \Big(N(0) - \frac{A_1 + A_2}{\mu_1 \wedge \mu_2}\Big)e^{-(\mu_1\wedge\mu_2)t}.$$

If $N(0) \leq \frac{A_1+A_2}{\mu_1\wedge\mu_2}$, then $N(t) \leq \frac{A_1+A_2}{\mu_1\wedge\mu_2}$ for all $t > 0$. □

Theorem 3. *For any* $(x_1(0), y_1(0), x_2(0), y_2(0)) \in \Gamma$*, the stochastic system (3) is mathematically well-posed in the sense that it has a unique solution* $(x_1(t), y_1(t), x_2(t), y_2(t)) \in \Gamma$ *with probability one.*

Proof. The coefficients of system (3) are locally Lipschitz continuous, for any given initial value $(x_1(0), y_1(0), x_2(0), y_2(0))$, then there is a unique local solution $(x_1(t), y_1(t), x_2(t), y_2(t))$ on $t \in [0, \tau_e)$, where τ_e is the explosion time.

Let $k_0 > 0$ such that $x_1(0), y_1(0), x_2(0), y_2(0) > k_0$. For $k \leq k_0$, we consider the stopping times

$$\begin{aligned} \tau_k =& \inf\big\{t \in [0, \tau_e) : x_1(t) \leq k \ \text{ or } \ y_1(t) \leq k \\ & \text{or } \ x_2(t) \leq k \ \text{ or } \ y_2(t) \leq k\big\}, \\ \tau =& \lim_{k\to 0}\tau_k = \inf\big\{t \in [0, \tau_e) : x_1(t) \leq 0 \ \text{ or } \ y_1(t) \leq 0 \\ & \text{or } \ x_2(t) \leq 0 \ \text{ or } \ y_2(t) \leq 0\big\}. \end{aligned}$$

Let

$$\begin{aligned} V(X(t)) &= V((x_1(t), y_1(t), x_2(t), y_2(t))) \\ &= 4\ln\frac{A_1 + A_2}{\mu_1 \wedge \mu_2} - \ln(x_1(t)y_1(t)x_2(t)y_2(t)). \end{aligned}$$

Applying Itô formula on V, we obtain

$$\begin{aligned} dV =& -\frac{dx_1}{x_1} - \frac{dx_2}{x_2} - \frac{dy_1}{y_1} - \frac{dy_2}{y_2} + \Big(\frac{1}{2}\sigma_1^2y_1^2 + \frac{1}{2}\sigma_3^2y_2^2 \\ & + \frac{1}{2}\sigma_2^2y_2^2 + \frac{1}{2}\sigma_4^2y_1^2 + \frac{1}{2}\sigma_1^2x_1^2 + \frac{1}{2}\sigma_4^2x_2^2 + \frac{1}{2}\sigma_2^2x_2^2 + \frac{1}{2}\sigma_3^2x_1^2\Big)dt \\ \leq& \Big[-\frac{A_1}{x_1} + \beta_1y_1 + \beta_3y_2 + \mu_1 - \beta_1x_1 - \beta_4x_2 + \mu_1 - \frac{A_2}{x_2} \\ & + \beta_2y_2 + \beta_4y_1 + \mu_2 - \beta_2x_2 - \beta_3x_1 + \mu_2 + \Big(\sigma_1^2 + \sigma_2^2 + \sigma_3^2 \\ & + \sigma_4^2\Big)\Big(\frac{A_1 + A_2}{\mu_1 \wedge \mu_2}\Big)^2\Big]dt + \sigma_1(y_1 - x_1)dB_1 + \sigma_2(y_2 - x_2)dB_2 \\ & + \sigma_3(y_2 - x_1)dB_3 + \sigma_4(y_1 - x_2)dB_4 \\ \leq& \Big[2\mu_1 + 2\mu_2 + \Big(\beta_1 + \beta_2 + \beta_3 + \beta_4\Big)\frac{A_1 + A_2}{\mu_1 \wedge \mu_2} + \Big(\sigma_1^2 + \sigma_2^2 \\ & + \sigma_3^2 + \sigma_4^2\Big)\Big(\frac{A_1 + A_2}{\mu_1 \wedge \mu_2}\Big)^2\Big]dt + \sigma_1(y_1 - x_1)dB_1 \\ & + \sigma_2(y_2 - x_2)dB_2 + \sigma_3(y_2 - x_1)dB_3 + \sigma_4(y_1 - x_2)dB_4. \end{aligned}$$

Integrating the both sides of the inequality above and then taking the expectation give

$$\mathbb{E}[V(X(t))] \leq \lambda t + V(X(0)), \tag{5}$$

where

$$\begin{aligned} \lambda =& 2\mu_1 + 2\mu_2 + \Big(\beta_1 + \beta_2 + \beta_3 + \beta_4\Big)\frac{A_1 + A_2}{\mu_1 \wedge \mu_2} \\ & + \Big(\sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \sigma_4^2\Big)\Big(\frac{A_1 + A_2}{\mu_1 \wedge \mu_2}\Big)^2. \end{aligned}$$

Using the stopping time τ_k, one has

$$\begin{aligned}\mathbb{E}[V(X(t\wedge\tau_k))] =&\mathbb{E}[V(X(t\wedge\tau_k))\,\mathbb{I}_{(\tau_k\le t)}]\\ &+\mathbb{E}[V(X(t\wedge\tau_k))\,\mathbb{I}_{(\tau_k>t)}]\\ \ge&\mathbb{E}[V(X(\tau_k))\,\mathbb{I}_{(\tau_k\le t)}],\end{aligned}$$

where $\mathbb{I}_A$ is the characteristic function of A.
Notice that there is some component of $X(\tau_k)$ equals to k. Therefore

$$V(X(\tau_k)) \ge \ln\left(\frac{A_1+A_2}{\mu_1\wedge\mu_2}\,\frac{1}{k}\right).$$

As a result, we have

$$\mathbb{E}[V(X(t\wedge\tau_k))] \ge \ln\left(\frac{A_1+A_2}{\mu_1\wedge\mu_2}\,\frac{1}{k}\right)\times\mathbb{P}\Big(\tau_k\le t\Big).$$

Together with (5), we get

$$\mathbb{P}\Big(\tau_k\le t\Big) \le \frac{\lambda t+V(X(0))}{\ln\left(\frac{A_1+A_2}{\mu_1\wedge\mu_2}\,\frac{1}{k}\right)}.$$

If we let $k\to 0$, we obtain for all $t\ge 0$: $\mathbb{P}\Big(\tau\le t\Big)=0$.
Hence

$$\mathbb{P}\Big(\tau=\infty\Big)=1.$$

As $\tau_e\ge\tau$, then $\tau_e=\infty$ a.s.
Finally, the solution is global. □

4. Decline of the disease

Theorem 4. *Let $(x_1(t),y_1(t),x_2(t),y_2(t))$ be the solution of system (3) with any initial value $(x_1(0),y_1(0),x_2(0),y_2(0))\in\Gamma$.*

(1) *If* $\frac{\beta_1^2}{2\sigma_1^2}+\frac{\beta_4^2}{2\sigma_4^2}<\mu_1$, *then* $\lim\limits_{t\to\infty}y_1(t)=0$ *a.s.*
(2) *If* $\frac{\beta_2^2}{2\sigma_2^2}+\frac{\beta_3^2}{2\sigma_3^2}<\mu_2$, *then* $\lim\limits_{t\to\infty}y_2(t)=0$ *a.s.*

Proof. 1. Applying Itô formula to system (3), we get

$$\begin{aligned}d\ln y_1(t) =&\frac{1}{y_1(t)}dy_1(t)-\frac{1}{2}\,\frac{1}{y_1^2}(dy_1(t))^2\\ =&\Big(\beta_1x_1(t)+\beta_4x_2(t)-\mu_1-\frac{\sigma_1^2}{2}x_1^2(t)\\ &-\frac{\sigma_4^2}{2}x_2^2(t)\Big)dt+\sigma_1x_1(t)dB_1(t)\\ &+\sigma_4x_2(t)dB_2(t).\end{aligned}$$

It follows that

$$\begin{aligned}\frac{1}{t}\ln\frac{y_1(t)}{y_1(0)} =&\beta_1\langle x_1(t)\rangle+\beta_4\langle x_2(t)\rangle-\mu_1-\frac{\sigma_1^2}{2}\langle x_1^2(t)\rangle\\ &-\frac{\sigma_4^2}{2}\langle x_2^2(t)\rangle+\frac{M_1(t)}{t}\\ \le&\beta_1\langle x_1(t)\rangle+\beta_4\langle x_2(t)\rangle-\mu_1-\frac{\sigma_1^2}{2}\langle x_1(t)\rangle^2\\ &-\frac{\sigma_4^2}{2}\langle x_2(t)\rangle^2+\frac{M_1(t)}{t}\\ =&-\frac{\sigma_1^2}{2}\langle x_1(t)\rangle^2+\beta_1\langle x_1(t)\rangle-\frac{\sigma_4^2}{2}\langle x_2(t)\rangle^2\\ &+\beta_4\langle x_2(t)\rangle-\mu_1+\frac{M_1(t)}{t}\\ =&-\frac{\sigma_1^2}{2}\Big(\langle x_1(t)\rangle^2-2\frac{\beta_1}{\sigma_1^2}\langle x_1(t)\rangle\Big)\\ &-\frac{\sigma_4^2}{2}\Big(\langle x_2(t)\rangle^2-2\frac{\beta_4}{\sigma_4^2}\langle x_2(t)\rangle\Big)-\mu_1\\ &+\frac{M_1(t)}{t}\\ =&-\frac{\sigma_1^2}{2}\Big(\langle x_1(t)\rangle-\frac{\beta_1}{\sigma_1^2}\Big)^2+\frac{\beta_1^2}{2\sigma_1^2}\\ &-\frac{\sigma_4^2}{2}\Big(\langle x_2(t)\rangle-\frac{\beta_4}{\sigma_4^2}\Big)^2+\frac{\beta_4^2}{2\sigma_4^2}-\mu_1\\ &+\frac{M_1(t)}{t}\qquad(6)\\ \le&\frac{\beta_1^2}{2\sigma_1^2}+\frac{\beta_4^2}{2\sigma_4^2}-\mu_1+\frac{M_1(t)}{t},\end{aligned}$$

where

$$M_1(t)=\sigma_1\int_0^t x_1(r)dB_1(r)+\sigma_4\int_0^t x_2(r)dB_2(r).$$

Bearing in mind the strong law of large numbers for martingales, we obtain

$$\limsup_{t\to\infty}\frac{\ln y_1(t)}{t}\le\frac{\beta_1^2}{2\sigma_1^2}+\frac{\beta_4^2}{2\sigma_4^2}-\mu_1\quad\text{a.s.}$$

Since $\frac{\beta_1^2}{2\sigma_1^2}+\frac{\beta_4^2}{2\sigma_4^2}<\mu_1$, then $\limsup\limits_{t\to\infty}\frac{\ln y_1(t)}{t}<0$ a.s.,
which implies

$$\lim_{t\to\infty}y_1(t)=0\quad\text{a.s.}$$

2. Similarly, we get : $\lim\limits_{t\to\infty}y_2(t)=0$ a.s., under the condition $\frac{\beta_2^2}{2\sigma_2^2}+\frac{\beta_3^2}{2\sigma_3^2}<\mu_2$.

This completes the proof of Theorem 4. □

Theorem 5. *Let $(x_1(t), y_1(t), x_2(t), y_2(t))$ be the solution of system (3) with any initial value $(x_1(0), y_1(0), x_2(0), y_2(0)) \in \Gamma$.*

(1) *If* $\frac{A_1+A_2}{\mu_1} \leq \frac{\beta_1}{\sigma_1^2}$, $\frac{A_1+A_2}{\mu_2} \leq \frac{\beta_4}{\sigma_4^2}$ *and* $\beta_1\frac{A_1+A_2}{\mu_1} + \beta_4\frac{A_1+A_2}{\mu_2} - \frac{\sigma_1^2}{2}\Big(\frac{A_1+A_2}{\mu_1}\Big)^2 - \frac{\sigma_4^2}{2}\Big(\frac{A_1+A_2}{\mu_2}\Big)^2 < \mu_1$, *then*

$$\lim_{t\to\infty} y_1(t) = 0 \quad a.s.$$

(2) *If* $\frac{A_1+A_2}{\mu_1} \leq \frac{\beta_3}{\sigma_3^2}$, $\frac{A_1+A_2}{\mu_2} \leq \frac{\beta_2}{\sigma_2^2}$ *and* $\beta_3\frac{A_1+A_2}{\mu_1} + \beta_2\frac{A_1+A_2}{\mu_2} - \frac{\sigma_3^2}{2}\Big(\frac{A_1+A_2}{\mu_1}\Big)^2 - \frac{\sigma_2^2}{2}\Big(\frac{A_1+A_2}{\mu_2}\Big)^2 < \mu_2$, *then*

$$\lim_{t\to\infty} y_2(t) = 0 \quad a.s.$$

Before giving the proof of Theorem 5, we will present the two following Lemmas.

Lemma 2. *Let $(x_1(t), y_1(t), x_2(t), y_2(t))$ be the solution of system (3). We have*

$$\lim_{t\to\infty} \frac{x_1(t)+y_1(t)+x_2(t)+y_2(t)}{t} = 0 \quad a.s.$$

Proof. Let $N(t) = x_1(t) + y_1(t) + x_2(t) + y_2(t)$.

From system (3), one has

$$\begin{aligned} dN(t) = &\left(A_1 + A_2 - \mu_1\left(x_1(t) + y_1(t)\right)\right) \\ &- \left(\mu_2\left(x_2(t) - y_2(t)\right)\right)dt. \end{aligned} \tag{7}$$

Then

$$\begin{aligned} &\Big(A_1 + A_2 - (\mu_1 \vee \mu_2)N(t)\Big)dt \leq dN(t) \\ &\leq \Big(A_1 + A_2 - (\mu_1 \wedge \mu_2)N(t)\Big)dt. \end{aligned} \tag{8}$$

Thus

$$\frac{A_1+A_2}{\mu_1 \vee \mu_2} + \Big(N(0) - \frac{A_1+A_2}{\mu_1 \vee \mu_2}\Big)e^{-(\mu_1\vee\mu_2)t} \leq N(t),$$

and

$$N(t) \leq \frac{A_1+A_2}{\mu_1 \wedge \mu_2} + \Big(N(0) - \frac{A_1+A_2}{\mu_1 \wedge \mu_2}\Big)e^{-(\mu_1\wedge\mu_2)t}.$$

Hence

$$\lim_{t\to\infty} \frac{x_1(t)+y_1(t)+x_2(t)+y_2(t)}{t} = 0 \quad a.s.$$

The proof is complete. □

Lemma 3. *Let $(x_1(t), y_1(t), x_2(t), y_2(t))$ be the solution of system (3). Then*

$$\begin{aligned} \limsup_{t\to\infty} \langle x_1(t)\rangle &\leq \frac{A_1+A_2}{\mu_1} \quad a.s., \\ \limsup_{t\to\infty} \langle x_2(t)\rangle &\leq \frac{A_1+A_2}{\mu_2} \quad a.s. \end{aligned}$$

Proof. By (7), we obtain

$$\begin{aligned} \langle x_1(t)\rangle &\leq \frac{A_1+A_2}{\mu_1} - \frac{\phi(t)}{\mu_1}, \\ \langle x_2(t)\rangle &\leq \frac{A_1+A_2}{\mu_2} - \frac{\phi(t)}{\mu_2}, \end{aligned}$$

where

$$\begin{aligned} \phi(t) =& \frac{x_1(t)+y_1(t)+x_2(t)+y_2(t)}{t} \\ &- \frac{x_1(0)+y_1(0)+x_2(0)+y_2(0)}{t}. \end{aligned}$$

Bearing in mind Lemma 2, we get the seeked results. □

Proof of Theorem 5. 1. By Lemma 3, there is $T_1 > 0$ such that, for any $t \geq T_1$,

$$\langle x_1(t)\rangle \leq \frac{A_1+A_2}{\mu_1} \quad \text{and} \quad \langle x_2(t)\rangle \leq \frac{A_1+A_2}{\mu_2}.$$

For all $t \geq T_1$, we assume that

$$\langle x_1(t)\rangle \leq \frac{A_1+A_2}{\mu_1} \leq \frac{\beta_1}{\sigma_1^2},$$

and

$$\langle x_2(t)\rangle \leq \frac{A_1+A_2}{\mu_2} \leq \frac{\beta_4}{\sigma_4^2}.$$

Together with (6), we have

$$\begin{aligned} \frac{1}{t}\ln\frac{y_1(t)}{y_1(0)} \leq& -\frac{{\sigma_1}^2}{2}\Big(\frac{A_1+A_2}{\mu_1} - \frac{\beta_1}{\sigma_1^2}\Big)^2 + \frac{\beta_1^2}{2\sigma_1^2} \\ &- \frac{\sigma_4^2}{2}\Big(\frac{A_1+A_2}{\mu_2} - \frac{\beta_4}{\sigma_4^2}\Big)^2 + \frac{\beta_4^2}{2\,\sigma_4^2} - \mu_1 \\ =& \beta_1\frac{A_1+A_2}{\mu_1} + \beta_4\frac{A_1+A_2}{\mu_2} - \frac{\sigma_1^2}{2}\Big(\frac{A_1+A_2}{\mu_1}\Big)^2 \\ &- \frac{\sigma_4^2}{2}\Big(\frac{A_1+A_2}{\mu_2}\Big)^2 - \mu_1. \end{aligned}$$

Since

$$\beta_1\frac{A_1+A_2}{\mu_1}+\beta_4\frac{A_1+A_2}{\mu_2}-\frac{\sigma_1^2}{2}\Big(\frac{A_1+A_2}{\mu_1}\Big)^2 -\frac{\sigma_4^2}{2}\Big(\frac{A_1+A_2}{\mu_2}\Big)^2-\mu_1<0,$$

then

$$\limsup_{t\to\infty}\frac{\ln y_1(t)}{t}<0 \quad \text{a.s.}$$

Consequently

$$\lim_{t\to\infty} y_1(t)=0 \quad \text{a.s.}$$

2. Following the same method above, we get

$$\lim_{t\to\infty} y_2(t)=0 \quad \text{a.s.}$$

□

5. Disease prevalence

Theorem 6. *Let $(x_1(t),y_1(t),x_2(t),y_2(t))$ be the solution of system (3) with any initial value $(x_1(0),y_1(0),x_2(0),y_2(0))\in\Gamma$.*

(1) *If* $(\beta_1\wedge\beta_4)\frac{A_1+A_2}{\mu_1\vee\mu_2}-\frac{\sigma_1^2+\sigma_4^2}{2}\Big(\frac{A_1+A_2}{\mu_1\wedge\mu_2}\Big)^2>(\beta_1\wedge\beta_4)+\mu_1$, *then:* $\liminf_{t\to\infty}\langle y_1(t)\rangle>0$ *a.s.*

(2) *If* $(\beta_2\wedge\beta_3)\frac{A_1+A_2}{\mu_1\vee\mu_2}-\frac{\sigma_2^2+\sigma_3^2}{2}\Big(\frac{A_1+A_2}{\mu_1\wedge\mu_2}\Big)^2>(\beta_2\wedge\beta_3)+\mu_2$, *then:* $\liminf_{t\to\infty}\langle y_2(t)\rangle>0$ *a.s.*

Proof. 1. From (7), we have

$$(\mu_1\vee\mu_2)\langle x_1(t)+x_2(t)\rangle\ge A_1+A_2-\mu_1\langle y_1(t)\rangle-\mu_2\frac{A_1+A_2}{\mu_1\wedge\mu_2}-\phi(t).$$

Then, one can get

$$\begin{aligned}\langle x_1(t)+x_2(t)\rangle&\ge\frac{A_1+A_2}{\mu_1\vee\mu_2}-\frac{\mu_1}{\mu_1\vee\mu_2}\langle y_1(t)\rangle-\frac{\mu_2}{\mu_1\vee\mu_2}\frac{A_1+A_2}{\mu_1\wedge\mu_2}-\frac{\phi(t)}{\mu_1\vee\mu_2}\\&\ge\frac{A_1+A_2}{\mu_1\vee\mu_2}-\langle y_1(t)\rangle-\frac{A_1+A_2}{\mu_1\wedge\mu_2}-\frac{\phi(t)}{\mu_1\vee\mu_2}.\end{aligned}\tag{9}$$

On the other hand, one can have

$$\frac{1}{t}\ln\frac{y_1(t)}{y_1(0)}\ge(\beta_1\wedge\beta_4)\langle x_1(t)+x_2(t)\rangle-\mu_1-\frac{\sigma_1^2+\sigma_4^2}{2}\Big(\frac{A_1+A_2}{\mu_1\wedge\mu_2}\Big)^2+\frac{M_1(t)}{t}.\tag{10}$$

Combining (9) and (10) yields

$$\begin{aligned}\frac{1}{t}\ln\frac{y_1(t)}{y_1(0)}&\ge(\beta_1\wedge\beta_4)\frac{A_1+A_2}{\mu_1\vee\mu_2}-(\beta_1\wedge\beta_4)\langle y_1(t)\rangle\\&\quad-(\beta_1\wedge\beta_4)\frac{A_1+A_2}{\mu_1\wedge\mu_2}-\frac{\beta_1\wedge\beta_4}{\mu_1\vee\mu_2}\phi(t)-\mu_1\\&\quad-\frac{\sigma_1^2+\sigma_4^2}{2}\Big(\frac{A_1+A_2}{\mu_1\wedge\mu_2}\Big)^2+\frac{M_1(t)}{t}.\end{aligned}$$

As a consequence, we get

$$\begin{aligned}\liminf_{t\to\infty}(\beta_1\wedge\beta_4)\langle y_1(t)\rangle\ge&(\beta_1\wedge\beta_4)\frac{A_1+A_2}{\mu_1\vee\mu_2}\\&-\frac{\sigma_1^2+\sigma_4^2}{2}\Big(\frac{A_1+A_2}{\mu_1\wedge\mu_2}\Big)^2\\&-\Big((\beta_1\wedge\beta_4)+\mu_1\Big)\quad\text{a.s.}\end{aligned}$$

Immediately, under the condition stated in the first part of Theorem 6, we deduce that

$$\liminf_{t\to\infty}\langle y_1(t)\rangle>0 \quad \text{a.s.}$$

2. Concerning the second part of Theorem 6, we get the desired result using the above method.

□

6. Simultaneous extinction and persistence

Theorem 7. *Let $(x_1(t),y_1(t),x_2(t),y_2(t))$ be the solution of system (3) with any initial value $(x_1(0),y_1(0),x_2(0),y_2(0))\in\Gamma$.*

(1) *If* $(\beta_2\vee\beta_3)\frac{A_1+A_2}{\mu_1\vee\mu_2}-\frac{\sigma_2^2+\sigma_3^2}{2}\big(\frac{A_1+A_2}{\mu_1\wedge\mu_2}\big)^2>\mu_2$ *and* $\lim_{t\to\infty}y_1(t)=0$ *a.s., then:* $\liminf_{t\to\infty}\langle y_2(t)\rangle>0$ *a.s.*

(2) *If* $(\beta_1\vee\beta_4)\frac{A_1+A_2}{\mu_1\vee\mu_2}-\frac{\sigma_1^2+\sigma_4^2}{2}\big(\frac{A_1+A_2}{\mu_1\wedge\mu_2}\big)^2>\mu_1$ *and* $\lim_{t\to\infty}y_2(t)=0$ *a.s., then:* $\liminf_{t\to\infty}\langle y_1(t)\rangle>0$ *a.s.*

Proof. 1. In the case of the extinction of urban infected, we have: $\lim_{t\to\infty} y_1(t) = 0$ a.s.

Then, for any $\epsilon > 0$, there exist$T_2 > 0$ such that: $y_1(t) \leq \epsilon$ for all $t \geq T_2$.

Together with (7), one can get

$$\begin{aligned}\phi(t) &\geq A_1 + A_2 - (\mu_1 \vee \mu_2)\langle x_1(t) + x_2(t)\rangle \\ &\quad - \mu_1\langle y_1(t)\rangle - \mu_2\langle y_2(t)\rangle \\ &= A_1 + A_2 - (\mu_1 \vee \mu_2)\langle x_1(t) + x_2(t)\rangle \\ &\quad - \mu_1\frac{1}{t}\int_0^{T_2} y_1(r)dr - \mu_1\frac{1}{t}\int_{T_2}^{t} y_1(r)dr \\ \mu_2\langle y_2(t)\rangle &\geq A_1 + A_2 - (\mu_1 \vee \mu_2)\langle x_1(t) + x_2(t)\rangle \\ &\quad - \frac{T_2}{t}\,\mu_1 \sup_{r\in[0,T_2]} y_1(r) - \mu_1(1 - \frac{T_2}{t})\epsilon - \mu_1\langle y_2(t)\rangle.\end{aligned}$$

Then

$$\begin{aligned}\langle x_1(t) + x_2(t)\rangle \geq &\frac{A_1 + A_2}{\mu_1 \vee \mu_2} - \frac{T_2}{t}\sup_{r\in[0,T_2]} y_1(r) \\ &- \epsilon - \langle y_2(t)\rangle - \frac{\phi(t)}{\mu_1 \vee \mu_2}. \quad (11)\end{aligned}$$

Now, we apply Itô formula on system (3) to obtain

$$\begin{aligned}\frac{1}{t}\ln\frac{y_2(t)}{y_2(0)} \geq &(\beta_2 \wedge \beta_3)\langle x_1(t) + x_2(t)\rangle - \mu_2 \\ &- \frac{\sigma_2^2 + \sigma_3^2}{2}\Big(\frac{A_1 + A_2}{\mu_1 \wedge \mu_2}\Big)^2 + \frac{M_2(t)}{t}, \quad (12)\end{aligned}$$

where

$$M_2(t) = \sigma_2\int_0^t x_2(r)dB_2(r) + \sigma_3\int_0^t x_1(r)dB_3(r).$$

Injecting (11) on (12) gives

$$\begin{aligned}\frac{1}{t}\ln\frac{y_2(t)}{y_2(0)} \geq &(\beta_2 \wedge \beta_3)\frac{A_1 + A_2}{\mu_1 \vee \mu_2} \\ &- (\beta_2 \wedge \beta_3)\frac{1}{t}\sup_{r\in[0,T_2]} y_1(r) \\ &- (\beta_2 \wedge \beta_3)\langle y_2(t)\rangle - (\beta_2 \wedge \beta_3)\epsilon \\ &- \frac{\beta_2 \wedge \beta_3}{\mu_1 \vee \mu_2}\phi(t) - \mu_2 \\ &- \frac{\sigma_2^2 + \sigma_3^2}{2}\Big(\frac{A_1 + A_2}{\mu_1 \wedge \mu_2}\Big)^2 + \frac{M_2(t)}{t}.\end{aligned}$$

According to Lemma 2, we can have

$$\begin{aligned}\liminf_{t\to\infty}(\beta_2 \wedge \beta_3)\langle y_2(t)\rangle \geq &(\beta_2 \wedge \beta_3)\frac{A_1 + A_2}{\mu_1 \vee \mu_2} - \mu_2 \\ &- \frac{\sigma_2^2 + \sigma_3^2}{2}\Big(\frac{A_1 + A_2}{\mu_1 \wedge \mu_2}\Big)^2 \quad \text{a.s.}\end{aligned}$$

2. Similarly, we get

$$\begin{aligned}\liminf_{t\to\infty}(\beta_1 \wedge \beta_4)\langle y_1(t)\rangle \geq &(\beta_1 \wedge \beta_4)\frac{A_1 + A_2}{\mu_1 \vee \mu_2} - \mu_1 \\ &- \frac{\sigma_1^2 + \sigma_4^2}{2}\Big(\frac{A_1 + A_2}{\mu_1 \wedge \mu_2}\Big)^2 \quad \text{a.s.}\end{aligned}$$

The proof is complete. □

7. Numerical results

The main goal of this section is to perform a numerical verification of the results obtained in the previous sections. First of all, we choose the initial value as $(x_1(0), y_1(0), x_2(0), y_2(0)) = (0.5, 0.7, 0.4, 0.9)$. The other parameters values are summarized in Table 2 split into 8 tests.

7.1. Deterministic stability

Based on the values of Test 0, we have

$$\begin{aligned}\Big(\big((\beta_1 + \beta_4)\frac{A_1 + A_2}{\mu_1 \wedge \mu_2} - \mu_1\big) \vee \big((\beta_2 + \beta_3)\frac{A_1 + A_2}{\mu_1 \wedge \mu_2} - \mu_2\big)\Big) \\ = -0.41477.\end{aligned}$$

According to Theorem 2, the equilibrium $E = (0.2308, 0, 0.1667, 0)$ is globally asymptotically stable which is depicted in Figure 1.

7.2. Stochastic extinction of the epidemic

By considering the values of Test 1, we have the following calculation

$$\frac{\beta_1^2}{2\sigma_1^2} + \frac{\beta_4^2}{2\sigma_4^2} - \mu_1 = -0.0036,$$

and

$$\frac{\beta_2^2}{2\sigma_2^2} + \frac{\beta_3^2}{2\sigma_3^2} - \mu_2 = -0.1375.$$

From Figure 2, we observe that: $\lim_{t\to\infty} y_1(t) = \lim_{t\to\infty} y_2(t) = 0$, which conform to the Theorem 4. Second, we obtain for Test 2

$$\frac{A_1 + A_2}{\mu_1} - \frac{\beta_1}{\sigma_1^2} = -1.3626, \ \frac{A_1 + A_2}{\mu_2} - \frac{\beta_4}{\sigma_4^2} = -2.441$$

and

$$\beta_1\frac{A_1+A_2}{\mu_1}+\beta_4\frac{A_1+A_2}{\mu_2}-\frac{\sigma_1^2}{2}\Big(\frac{A_1+A_2}{\mu_1}\Big)^2 - \frac{\sigma_4^2}{2}\Big(\frac{A_1+A_2}{\mu_2}\Big)^2-\mu_1=-0.017.$$

According to Theorem 5, $y_1(t)$ converges exponentially to zero (see Figure 2).
Next, based on the parameters values for Test 3, the numerical values are

$$\frac{A_1+A_2}{\mu_1}-\frac{\beta_3}{\sigma_3^2}=-0.9437,\quad \frac{A_1+A_2}{\mu_2}-\frac{\beta_2}{\sigma_2^2}=-2.2292$$

and

$$\beta_3\frac{A_1+A_2}{\mu_1}+\beta_2\frac{A_1+A_2}{\mu_2}\frac{\sigma_3^2}{2}\Big(\frac{A_1+A_2}{\mu_1}\Big)^2 - \frac{\sigma_2^2}{2}\Big(\frac{A_1+A_2}{\mu_2}\Big)^2-\mu_2=-0.0799.$$

From Figure 2, we see that $y_2(t)$ tends to zero, which agrees with Theorem 5.

7.3. Stochastic persistence of the epidemic

We choose values of Test 4 and Test 5 to get the following

$$(\beta_1\wedge\beta_4)\frac{A_1+A_2}{\mu_1\vee\mu_2}-\frac{\sigma_1^2+\sigma_4^2}{2}\Big(\frac{A_1+A_2}{\mu_1\wedge\mu_2}\Big)^2 - (\beta_1\wedge\beta_4)-\mu_1=0.1589$$

and

$$(\beta_2\wedge\beta_3)\frac{A_1+A_2}{\mu_1\vee\mu_2}-\frac{\sigma_2^2+\sigma_3^2}{2}\Big(\frac{A_1+A_2}{\mu_1\wedge\mu_2}\Big)^2 - (\beta_2\wedge\beta_3)-\mu_2=0.1489.$$

By virtue of Theorem 6, the epidemic will be persistent in both urban and rural areas (see Figure 3).

7.4. Simultaneous extinction and persistence

Case 1. We have already considered that $\lim\limits_{t\to\infty} y_1(t)=0$. From values of Test 6, we obtain

$$(\beta_2\vee\beta_3)\frac{A_1+A_2}{\mu_1\vee\mu_2}-\frac{\sigma_2^2+\sigma_3^2}{2}\Big(\frac{A_1+A_2}{\mu_1\wedge\mu_2}\Big)^2 - \mu_2=0.0631.$$

Therefore, Theorem 7 yields

$$\liminf_{t\to\infty}\langle y_2(t)\rangle>0,$$

which is well confirmed by Figure 4.

Case 2. Based on the values of Test 7 and , we get that

$$(\beta_1\vee\beta_4)\tfrac{A_1+A_2}{\mu_1\vee\mu_2}-\tfrac{\sigma_1^2+\sigma_4^2}{2}\big(\tfrac{A_1+A_2}{\mu_1\wedge\mu_2}\big)^2-\mu_1=0.0528.$$

If we consider $\lim\limits_{t\to\infty} y_2(t)=0$, then Theorem 7 implies that: $\liminf\limits_{t\to\infty}\langle y_1(t)\rangle>0$.

Therefore, Figure 4 reflects perfectly the statement of Theorem 7.

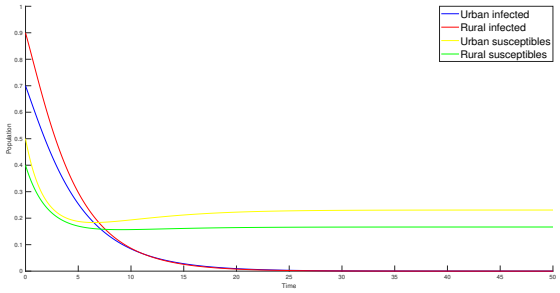


Figure 1. Computer simulation of $x_1(t)$, $x_2(t)$, $y_1(t)$ and $y_2(t)$ for model (2), corresponding to Test 0.

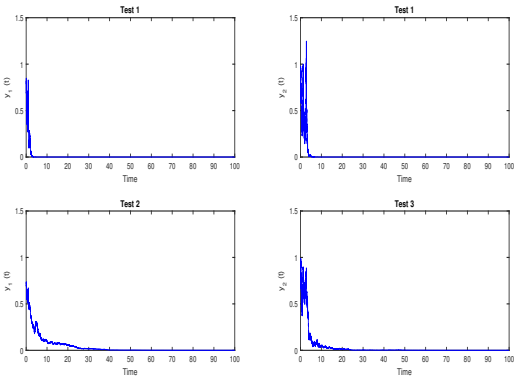


Figure 2. The paths of $y_1(t)$ and $y_2(t)$ for model (3), corresponding to Test 1, Test 2 and Test 3.

Table 2. Parameters values.

Parameters	Test 0	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6	Test 7
A_1	0.06	0.3	0.06	0.04	0.1	0.1	0.06	0.04
A_2	0.05	0.25	0.05	0.035	0.1	0.2	0.05	0.035
μ_1	0.26	0.1	0.26	0.2	0.08	0.08	0.26	0.2
μ_2	0.3	0.2	0.1	0.2	0.09	0.09	0.1	0.2
β_1	0.14	0.2	0.14	—	0.5	—	0.14	0.6
β_2	0.1	0.15	0.1	0.15	—	0.5	0.5	0.15
β_3	0.2	0.2	0.2	0.211	—	0.4	0.4	0.211
β_4	0.08	0.3	0.211	—	0.4	—	0.211	0.7
σ_1	—	0.7	0.28	—	0.2	—	0.28	0.28
σ_2	—	0.6	—	0.24	—	0.2	0.2	0.24
σ_3	—	0.8	—	0.4	—	0.2	0.2	0.4
σ_4	—	0.9	0.244	—	0.2	—	0.244	0.244

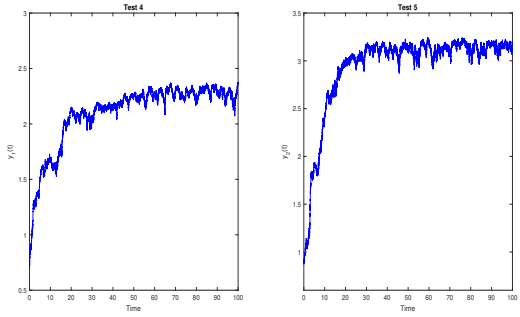


Figure 3. The paths of $y_1(t)$ and $y_2(t)$ for model (3), corresponding to Test 4 and Test 5.

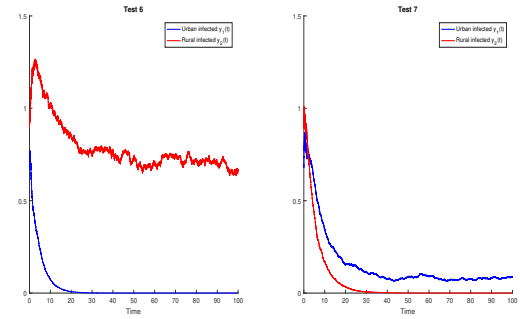


Figure 4. The paths of $y_1(t)$ and $y_2(t)$ for model (3), corresponding to Test 6 and Test 7.

8. Conclusion

In this paper, we elucidate the dynamics of disease transmission between two groups from distinct regions, operating under the assumption of comprehensive and unrestricted interaction. We consider both a deterministic two-patch epidemic model and its stochastic counterpart. For the deterministic model (2), we examine the global asymptotic stability of the equilibrium $E = (\frac{A_1}{\mu_1}, 0, \frac{A_2}{\mu_2}, 0)$. This result is illustrated in Figure 1. Regarding the stochastic version of model (2), we demonstrate the uniqueness of a positive solution for model (3). The thresholds that determine whether the disease will disappear are identified, as detailed in Theorems 4 and 5. Additionally, in Theorem 6, we establish conditions ensuring disease persistence. We also highlight a third scenario, distinct from those studied in sections 4 and 5, where the disease persists in one patch while simultaneously disappearing in the other. The accuracy of our theoretical findings is validated in the numerical simulation section.

Acknowledgments

The authors extend their deepest gratitude to the editor and the reviewers for their diligent efforts, meticulous evaluation, and invaluable feedback.

References

[1] Van Seventer, J.M., & Hochberg, N.S. (2016). Principles of Infectious Diseases: Transmission, Diagnosis, Prevention, and Control. *International Encyclopedia of Public Health*, 22-39. https://doi.org/10.1016/B978-0-12-803678-5.00516-6

[2] Braun, D. (2016). Health Security: Abwarten ist keine Option; Epidemien erfordern die Zusammenarbeit von Sicherheits-und Gesundheitsexperten. *DEU*, 1-4.

[3] Patrick, S. (2011). Weak links: fragile states, global threats, and international security. *Oxford University Press*.

[4] Wang, C., Horby, P.W., Hayden, F.G., & Gao, G.F. (2020). A novel coronavirus outbreak of global health concern. *Lancet (London, England)*, Vol.395, 470-473. https://doi.org/10.1016/S0140-6736(20)30185-9

[5] Chriscaden, K.(2020). Impact of COVID-19 on people's livelihoods, their health and our food systems. https://www.who.int/news/item/13-10-2020-impact-of-covid-19-on-people's-livelihoods-their-health-and-our-food-systems.

[6] Kermack, W.O., & Mckendrick, À.G. (1991). Contributions to the mathematical theory of epidemics I. *Bulletin of Mathematical Biology*, 53, 33-55. https://doi.org/10.1016/S0092-8240(05)80040-0

[7] Yavuz, M., ur Rahman, M., Yildiz, M., & Joshi, H. (2024). Mathematical Modeling of Middle East Respiratory Syndrome Coronavirus with Bifurcation Analysis. Contemporary Mathematics, 5(3), 3997-4012. https://doi.org/10.37256/cm.5320245004

[8] Ur Rahman, M., Yavuz, M., Arfan, M., & Sami, A. (2024). Theoretical and numerical investigation of a modified ABC fractional operator for the spread of polio under the effect of vaccination. *AIMS Biophysics*, 11(1), 97-120. https://doi.org/10.3934/biophy.2024007

[9] ur Rahman, M., Arfan, M., & Baleanu, D. (2023). Piecewise fractional analysis of the migration effect in plant-pathogen-herbivore interactions. *Bulletin of Biomathematics*, 1(1), 1-23. https://doi.org/10.59292/bulletinbiomath.2023001

[10] Li, M.Y., & Muldowney, J.S. (1995). Global stability for the SEIR model in epidemiology. *Mathematical Biosciences*, 125(2), 155-64. https://doi.org/10.1016/0025-5564(95)92756-5

[11] Smith, H.L., Wang, L., & Li, M.Y. (2001). Global Dynamics of an SEIR Epidemic Model with Vertical Transmission. *SIAM Journal on Applied Mathematics*, 62, 58-69. https://doi.org/10.1137/S0036139999359860

[12] Castillo-Chavez, C., Hethcote, H.W., Andreasen, V., Levin, S.A., & Liu, W. (1989). Epidemiological models with age structure, proportionate mixing, and cross-immunity. *Journal of Mathematical Biology*, 27, 233-258. https://doi.org/10.1007/BF00275810

[13] Shulgin, B., Stone, L., & Agur, Z. (1998). Pulse vaccination strategy in the SIR epidemic model. *Bulletin of Mathematical Biology*, 60, 1123-1148. https://doi.org/10.1016/S0092-8240(98)90005-2

[14] d'Onofrio, A. (2005). On pulse vaccination strategy in the SIR epidemic model with vertical transmission. *Applied Mathematics Letters*, 18, 729-732. https://doi.org/10.1016/j.aml.2004.05.012

[15] Shi, Z., Zhang, X., & Jiang, D. (2019). Dynamics of an avian influenza model with half-saturated incidence. *Applied Mathematics and Computation*, 355, 399-416. https://doi.org/10.1016/j.amc.2019.02.070

[16] Calvo, J.G., Hern'andez, A., Porter, M.A., & Sanchez, F. (2019). A two-patch epidemic model with nonlinear reinfection. *Revista de Matemática: Teoría y Aplicaciones*, 27(1), 23-48. https://doi.org/10.15517/rmta.v27i1.39946

[17] Moore, C. (1990). Unpredictability and undecidability in dynamical systems. *Physical Review Letters*, 64(20), 2354-2357. https://doi.org/10.1103/PhysRevLett.64.2354

[18] Cao, Z., Feng, W., Wen, X., & Zu, L. (2019). Dynamical behavior of a stochastic SEI epidemic model with saturation incidence and logistic growth. *Physica A: Statistical Mechanics and its Applications*, 523, 894-907. https://doi.org/10.1016/j.physa.2019.04.228

[19] Pang, Y., Han, Y., & Li, W. (2014). The threshold of a stochastic SIQS epidemic model. *Advances in Difference Equations*, 2014, 1-15. https://doi.org/10.1186/1687-1847-2014-320

[20] Khan, T., Zaman, G., & El-Khatib, Y. (2021). Modeling the dynamics of novel coronavirus (COVID-19) via stochastic epidemic model. *Results in Physics*, 24, 104004. https://doi.org/10.1016/j.rinp.2021.104004

[21] Joshi, H., & Yavuz, M. (2024). Numerical analysis of compound biochemical calcium oscillations process in hepatocyte cells. *Advanced Biology*, 8(4), 2300647. https://doi.org/10.1002/adbi.202300647

[22] Joshi, H., Yavuz, M., & Özdemir, N. (2024). Analysis of novel fractional order plastic waste model and its effects on air pollution with treatment mechanism. *Journal of Applied Analysis & Computation*, 14(6), 3078-3098. https://doi.org/10.11948/20230453

[23] Adel, W., Elsonbaty, A., & Mahdy, A. (2024). On some recent advances in fractional order modeling in engineering and science. *Computation and Modeling for Fractional Order Systems*, 169-197. https://doi.org/10.1016/B978-0-44-315404-1.00016-3

[24] Salih, R. I., Jawad, S., Dehingia, K., & Das, A. (2024). The effect of a psychological scare on the dynamics of the tumor-immune interaction with optimal control strategy. *An International Journal of Optimization and Control: Theories & Applications (IJOCTA)*, 14(3), 276-293. https://doi.org/10.11121/ijocta.1520

[25] Evirgen, F., Özköse, F., Yavuz, M., & Özdemir, N. (2023). Real data-based optimal control strategies for assessing the impact of the Omicron variant on heart attacks. *AIMS Bioengineering*, 10(3), 218-239. https://doi.org/10.3934/bioeng.2023015

[26] El-Mesady, A., Ahmed, N., Elsonbaty, A., & Adel, W. (2023). Transmission dynamics and control measures of reaction-diffusion pine wilt disease model. *The European Physical Journal Plus*, 138(12), 1078. https://doi.org/10.1140/epjp/s13360-023-04705-8

[27] Yavuz, M., Boulaasair, L., Bouzahir, H., Diop, M.A., & Benaid, B. (2024). The impact of two independent gaussian white noises on the behavior of a stochastic epidemic model. *Journal of Applied Mathematics and Computational Mechanics*, 23(1), 121-134. https://doi.org/10.17512/jamcm.2024.1.10

[28] Izadi, M., El-Mesady, A., & Adel, W. (2024). A novel Touchard polynomial-based spectral matrix collocation method for solving the Lotka-Volterra

competition system with diffusion. *Mathematical Modelling and Numerical Simulation with Applications*, 4(1), 37-65. https://doi.org/10.53391/mmnsa.1408997

[29] Ayaz, F., & Heredağ, K. (2024). Fractional model for blood flow under MHD influence in porous and non-porous media. *An International Journal of Optimization and Control: Theories & Applications (IJOCTA)*, 14(2), 156-167. https://doi.org/10.11121/ijocta.1497

[30] Raeisi, E., Yavuz, M., Khosravifarsani, M., & Fadaei, Y. (2024). Mathematical modeling of interactions between colon cancer and immune system with a deep learning algorithm. *The European Physical Journal Plus*, 139(4), 345. https://doi.org/10.1140/epjp/s13360-024-05111-4

[31] Boulaasair, L., Bouzahir, H., Vargas, A.N., & Diop, M.A. (2022). Existence and uniqueness of solutions for stochastic urban-population growth model. *Frontiers in Applied Mathematics and Statistics*, 8, 960399. https://doi.org/10.3389/fams.2022.960399

[32] Naik, P. A., Yavuz, M., Qureshi, S., Owolabi, K. M., Soomro, A., & Ganie, A. H. (2024). Memory impacts in hepatitis C: A global analysis of a fractional-order model with an effective treatment. *Computer Methods and Programs in Biomedicine*, 254, 108306. https://doi.org/10.1016/j.cmpb.2024.108306

[33] Evirgen, F., Uçar, E., Uçar, S., & Özdemir, N. (2023). Modelling influenza a disease dynamics under Caputo-Fabrizio fractional derivative with distinct contact rates. *Mathematical Modelling and Numerical Simulation with Applications*, 3(1), 58-73. https://doi.org/10.53391/mmnsa.1274004

[34] Sabbar, Y. (2023). Asymptotic extinction and persistence of a perturbed epidemic model with different intervention measures and standard lévy jumps. *Bulletin of Biomathematics*, 1(1), 58-77. https://doi.org/10.59292/bulletinbiomath.2023004

[35] Kiouach, D., & Boulaasair, L. (2018). Stationary distribution and dynamic behaviour of a stochastic SIVR epidemic model with imperfect vaccine. *Journal of Applied Mathematics*, 2018(1), 1291402. https://doi.org/10.1155/2018/1291402

[36] Elsonbaty, A., Alharbi, M., El-Mesady, A., & Adel, W. (2024). Dynamical analysis of a novel discrete fractional lumpy skin disease model. *Partial Differential Equations in Applied Mathematics*, 9, 100604. https://doi.org/10.1016/j.padiff.2023.100604

[37] Boulaasair, L. (2023). Threshold properties of a stochastic epidemic model with a variable vaccination rate. *Bulletin of Biomathematics*, 1(2), 177-191. https://doi.org/10.59292/bulletinbiomath.2023009

[38] Adel, W., Amer, Y.A., Youssef, E.S., & Mahdy, A.M. (2023). Mathematical analysis and simulations for a Caputo-Fabrizio fractional COVID-19 model. *Partial Differential Equations in Applied Mathematics*, 8, 100558. https://doi.org/10.1016/j.padiff.2023.100558

[39] Fatima, B., Yavuz, M., Rahman, M.U., & Al-Duais, F.S. (2023). Modeling the epidemic trend of middle eastern respiratory syndrome coronavirus with optimal control. *Mathematical Biosciences and Engineering*, 20(7), 11847-11874. https://doi.org/10.3934/mbe.2023527

Lahcen Boulaasair *earned his Ph.D. from the National School of Applied Sciences of Agadir, under the mentorship of Professor Hassane Bouzahir. He is presently engaged as a researcher at the LISTI laboratory within the same institution. His research interests lie in the field of applied mathematics, with a focus on stochastic differential equations, partial differential equations and time-series.*

https://orcid.org/0000-0001-6495-5760

Hassan Bouzahir *is a Full Professor of Mathematics/Statistics at the National School of Applied Sciences, Ibn Zohr University, Agadir, Morocco. He earned his Ph.D. in Mathematics in April 2001 and has held notable positions such as Head of Departments and Director of Research Laboratories. With numerous publications, including research books, book chapters, and articles in prestigious journals and databases, his work spans various topics in Engineering and Mathematics. He has also contributed to IEEE conferences and delivered talks globally. Known for his exceptional teaching, he has mentored many graduate students and Ph.D. candidates. His international career includes professional activities in numerous countries, making him a respected figure in the global academic community.*

https://orcid.org/0000-0002-1254-967X

Mehmet Yavuz *received a BSc degree (2009) from Zonguldak Bulent Ecevit University, Türkiye, and both MSc (2012) and PhD (2016) degrees in applied mathematics from Balıkesir University, Türkiye. He visited the Department of Mathematics, University of Exeter, U.K., in the period 2019-2020 for postdoctoral research. He is currently an Associate Professor in the Mathematics and Computer Sciences Department at Necmettin Erbakan University, Türkiye. He has published more than 100 research papers in reputed journals and conference papers as well as 10 book chapters in international books. Moreover, he edited 4 international books published by esteemed publishers. He is currently concerned with fractional calculus, and its applications to the different fields of science, mathematical biology, nonlinear dynamics, and optimal control. He has 12 years of research and teaching experience in these fields. Apart from being an Associate Editor in several esteemed journals, he is the Editor-in-Chief of the journals "Mathematical Modelling and Numerical Simulation with Applications" and "Bulletin of Biomathematics".*

https://orcid.org/0000-0002-3966-6518

An International Journal of Optimization and Control: Theories & Applications (http://www.ijocta.org)

This work is licensed under a Creative Commons Attribution 4.0 International License. The authors retain ownership of the copyright for their article, but they allow anyone to download, reuse, reprint, modify, distribute, and/or copy articles in IJOCTA, so long as the original authors and source are credited. To see the complete license contents, please visit http://creativecommons.org/licenses/by/4.0/.

An International Journal of Optimization and Control: Theories & Applications
ISSN:2146-0957 eISSN:2146-5703
Vol.14, No.4, pp.378-393 (2024)
http://doi.org/10.11121/ijocta.1608

RESEARCH ARTICLE

An Inverse recursive algorithm to retrieve the shape of the inaccessible dielectric objects

Ahmet Sefer*

Department of Electrical and Electronics Engineering, FMV Işık University, Istanbul, Türkiye
ahmet.sefer@isikun.edu.tr

ARTICLE INFO

Article History:
Received 14 May 2024
Accepted 11 September 2024
Available Online 16 October 2024

Keywords:
Inverse electromagnetic scattering
Two-patch epidemic model
Electromagnetic imaging
Shape reconstruction
Surface integral equations
Newton's algorithm

AMS Classification 2010:
47A52; 47J06; 65R30; 65R32; 78A45

ABSTRACT

A regularized electromagnetic iterative inverse algorithm is formulated and implemented to reconstruct the shape of 2D dielectric objects using the far-field pattern of the scattered field data. To achieve this, an integral operator that maps the unknown boundary of the object onto the far-field pattern of the scattered field is defined and solved for the unknown boundary. The addressed inverse problem has an ill-posed nature and inherits nonlinearity. To overcome these, the proposed solution is linearized via Newton and regularized by Tikhonov in the sense of least squares. Besides, the dominance of the shadow region in the inverse-imaging process is exceeded by considering the superposition of multi-incoming plane waves, leading to less computational cost and a very fast inversion process. Comprehensive numerical analyses are carried out to ascertain the algorithm's feasibility, revealing that it is very efficient and promising.

(cc) BY

1. Introduction

Inverse electromagnetic (EM) imaging methods that utilize scattered field data to reconstruct the shape of an unknown scatterer precisely are in significant demand across a wide variety of engineering fields, such as non-destructive testing, microwave imaging, and geophysical exploration, and so on [1–8]. Apart from its diverse applications, recovering objects from scattered fields poses an immense challenge due to its inherently nonlinear and ill-posed nature [9]. Recently, deep learning schemes and the corresponding applications have been of great interest among many engineering fields [10–12]. In addition to these, significant advancements in deep learning (DL) have led to substantial research investments in the field of inverse electromagnetic imaging problems [13–19]. In [14] and [15], the ill-posed problem is regularized considering Landweber iterations that are implemented in the regularized DL framework. [16] proposes two-step DL framework. In the first step, the dielectric properties of the inaccessible object are recovered. The object's shape is then reconstructed using the outcomes of the first step. In the sense of rough surface imaging, [18] recovers the statistical parameters of the randomly formed rough surfaces. The shape of random rough surfaces are directly recovered in [17] and [19] for different scattering scenarios.

In addition to DL applications and with the notable exception of certain non-iterative inversion techniques such as Fourier method [20], reverse time migration (RTM)-based [21] approach, and equivalent source model [22], the vast majority of algorithms developed to address these issues are based on recursive applications of regularization and linearization techniques [23–30]. Generally, these are constructed with consideration for multiple incident illuminations to enhance the precision of reconstructions [24, 25, 29, 30]. Nevertheless, these undertakings incur additional

*Corresponding Author

computational expenses. Furthermore, many of these solutions address the inverse problem associated with perfectly electric conducting (PEC) and sound-soft boundary conditions acoustically. This is because conceptually and physically, the recovery of a penetrable scatterer presents a more difficult inverse problem than the inverse problems associated with impenetrable obstacles [31]. In this regard, [29] proposes a solution to recover the 2D profile of an acoustically sound-soft scatterer by using the far-field pattern. The method utilizes multi-incidence monochromatic incident fields for illumination and applies multi-frequency measurement for higher accuracy. It solves a linearized system with a huge number of unknowns. The same consequence is valid for the linearized iterative methodologies presented in [24] and [25], which offers multi-incidence illumination for rigorous reconstructions. The principal reason for illumination is to reduce (or eliminate) the effect of shadow regions. Namely, the information in the far-field data becomes blurred for a limited amount of illumination, so the lack of information becomes dominant and yields unsuccessful reconstructions as the problem is inherently ill-posed. Alternatively, a recursive linearized method that uses only single incident illumination is proposed in [32]. The method is applied to recover unknown non-penetrable acoustically sound-soft obstacles using the far-field measured field pattern. Later, it is expanded to reconstruct the shape of penetrable objects in [26]. Within this context, for penetrable and non-penetrable cases, the far-field measured field pattern is represented by the single layer potential form [9].

This paper proposes a regularized and linearized recursive inverse algorithm to recover unknown penetrable objects using the scattered field data measured in the far-field region. Unlike other multi-illuminated inversion algorithms, the proposed algorithm applies the superposition of the multi-incident illuminations. Thus, the unknown 2D scatterer is illuminated by multiple sources simultaneously, and the scattered field is collected only once due to these simultaneous illuminations. This allows a fast inverse algorithm to obtain robust and successful reconstructions with a reduced computational cost. The superposition of incident fields is first considered for reconstructing sound-soft obstacles in [33]. Furthermore, unlike the open literature, the inverse algorithm uses a combination of double and single-layer potentials to represent the far-field measured scattered field data.

The paper's outline is provided as follows: Section 2 presents the considered EM scattering scenario, and the following Section 3 briefly summarizes the direct problem. Section 4 presents the regularized recursive inverse-imaging solution, which utilizes the far-field measurements to recover the unknown surface profile in detail. In Section 5, an extensive numerical study was conducted using various scattering scenarios to illustrate the algorithm's efficiency and examine its validation restrictions. Final remarks are given in the Section 6.

2. Geometry of the Problem

Fig. 1 represents the considered 2D scattering geometry. The unknown dielectric body is denoted as Ω embedded in infinite free space medium with permittivity ε_0 and permeability μ_0. The body is a simple non-magnetic lossy object defined in terms of constitutive electromagnetic parameters, where its permittivity and conductivity are denoted with $\varepsilon_1 = \varepsilon_0\varepsilon_r$, and σ (S/m), respectively. Accordingly, the body has a constant complex wave number k_1 with $Re\{k_1\} > 0$ and $Im\{k_1\} \geq 0$, precisely its square equals to $k_1^2 = \omega^2\varepsilon_1\mu_0 + i\omega\mu_0\sigma$, where ω is the radial operating frequency. The cross-section of the body constitutes the principle unknown of the problem, which is denoted with $\partial\Omega$ and defined as:

$$\partial\Omega := \Big\{(\rho,\varphi)\,|\,\rho = r(\varphi)\Big\}, \tag{1}$$

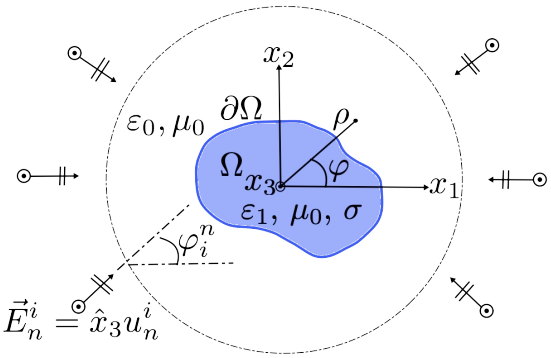


Figure 1. Geometry of the problem.

where (ρ,φ) are the cylindrical polar coordinates, such that $\rho > 0$ and $\varphi \in (0, 2\pi)$. Within this context, Ω has a star-like shape. As depicted, the 2D region is illuminated by a superposition of several monochromatic incident electric fields. The incident fields are TE polarized (transverse to wave propagation direction). The n^{th} incident is defined with a function u_n^i

$$\vec{E}_n^i = \hat{x}_3 u_n^i, \tag{2}$$

where u_n^i is the monochromatic plane wave, precisely equals

$$u_n^i = u_n^i(r(\varphi), \varphi_i^n) := e^{ik_0 r(\varphi)\cos\left(\varphi-\varphi_i^n\right)}. \quad (3)$$

Here, $k_0 > 0$ is the wavenumber of the free space. The superposition of the incident fields enables to consider the summation of incident illuminations as a single unique field, *i.e.,*

$$u^i = \sum_{n=1}^{N} u_n^i(r(\varphi), \varphi_i^n). \quad (4)$$

Regarding the incident illumination, both within the body and in the surrounding free space, the electric field vectors are in the $\hat{x}_3$ direction. Consequently, the entire problem may be simplified to a scalar one. To do so, let u_0 and u_1 denote the total fields in free space and the dielectric body, respectively. Then, both of them satisfy the scalar wave equation

$$\begin{aligned}\left(\Delta + k_0^2\right)u_0 &= 0 \text{ in } \mathbb{R}^2 \setminus \Omega,\\ \left(\Delta + k_1^2\right)u_1 &= 0 \text{ in } \Omega. \end{aligned} \quad (5)$$

Denote the derivative $r'(\varphi) = \frac{dr(\varphi)}{d\varphi}$ and the outer normal vector of $\partial\Omega$ as $\hat{\nu}$, which precisely equals

$$\hat{\nu} = \frac{\hat{\rho} r(\varphi) - \hat{\varphi} r'(\varphi)}{\sqrt{r(\varphi) + r'^2(\varphi)}}. \quad (6)$$

It follows that the fields and their derivatives with respect to the outward surface normal exhibit continuity on $\partial\Omega$. Namely, the boundary conditions imply the following:

$$u_0 = u_1, \quad (7a)$$
$$\psi_0 = \psi_1, \text{ on } \partial\Omega. \quad (7b)$$

Noting that the fields $\psi_m = \hat{\nu} \cdot \nabla u_m$ ($m = \{1, 2\}$), where "∇" denotes the gradient operator. Namely, ψ_0 and ψ_1 describe derivatives of u_0 and u_1 with respect to ν, respectively. The scattered field, in this regard, is defined as the difference

$$u^s = u_0 - u^i, \quad (8)$$

which is an outgoing wave and fulfills the Sommerfield radiation condition

$$\lim_{\rho\to\infty} \sqrt{\rho}\left(\frac{\partial u^s}{\partial \rho} - ik_0 u^s\right) = 0, \ r \text{ in } \mathbb{R}^2 \setminus \Omega \quad (9)$$

in a uniform way in all directions. Furthermore, it is straightforward to demonstrate that u^s exhibits the subsequent asymptotic behavior:

$$u^s(\rho, \varphi) = \frac{e^{ik\rho}}{\sqrt{\rho}} u_\infty(\varphi) + \mathcal{O}\left(\frac{1}{\rho^{3/2}}\right), \ \rho \to \infty. \quad (10)$$

Here, u_∞ represents the scattered field measured far away from the source, namely the far-field pattern. It is worth noting that the field also depends on k_0 and the incoming direction. However, the assumption is made by taking these quantities fixed so that u_∞ has only φ dependence. The fields defined in (7) represent the surface currents on the cross-section $\partial\Omega$. These are the unknowns of the direct scattering problem for which $\partial\Omega$ and the constitutive parameters are known. Once the currents are obtained, one can take an opportunity to obtain the fields scattered anywhere in the first medium (in our particular case, in free space). The whole procedure constitutes the "direct EM scattering problem". In the inverse problem, conversely, the inputs and the outputs are reversed. That is, the main concern is to recover the unknown cross-section, $\partial\Omega$, utilizing the measured far-field pattern of the scattered field data, *i.e.,* u_∞. To this aim, the integral representation of the scattered field data, described in the following subsection, is taken as a mapping operator into account, which maps $\partial\Omega$ onto u_∞. Hence, the problem turns into taking the inverse of the mapping operator. The following subsection describes the direct EM scattering problem applied to acquire the synthetic scattered field data utilized in the inverse problem.

3. Direct EM Problem

As mentioned above, the forward scattering problem mainly considers obtaining the surface currents given in (7) within the knowledge of the 2D cross-section $\partial\Omega$. Using Green's theorem both in the free space and Ω, one can easily obtain the integral representations of the surface currents u_0 and ψ_0 [34]

$$\begin{aligned} u_0(\boldsymbol{r}) = u^i(\boldsymbol{r}) &+ \int_{\partial\Omega} u_0(\boldsymbol{r'}) K_0(\boldsymbol{r};\boldsymbol{r'})\, ds(\varphi')\\ &- \int_{\partial\Omega} \psi_0(\boldsymbol{r'}) G_0(\boldsymbol{r};\boldsymbol{r'})\, ds(\varphi')\\ u_1(\boldsymbol{r}) = &- \int_{\partial\Omega} u_1(\boldsymbol{r'}) K_1(\boldsymbol{r};\boldsymbol{r'})\, ds(\varphi')\\ &+ \int_{\partial\Omega} \psi_1(\boldsymbol{r'}) G_1(\boldsymbol{r};\boldsymbol{r'})\, ds(\varphi') \end{aligned} \quad (11)$$

Here, $G_m(\boldsymbol{r};\boldsymbol{r'})$ ($m = \{0, 1\}$) is the fundamental solution of the scalar wave equation in 2D, *i.e.,*

$$G_m(\boldsymbol{r};\boldsymbol{r'}) = \frac{i}{4}H_0^{(1)}(k_m|\boldsymbol{r}-\boldsymbol{r'}|), \qquad (12)$$

where $\boldsymbol{r} = \hat{x}_1\rho\cos(\varphi) + \hat{x}_2\rho\sin(\varphi)$ and $\boldsymbol{r'} = \hat{x}_1 r(\varphi)\cos(\varphi') + \hat{x}_2 r(\varphi)\sin(\varphi')$ so that the argument of the Hankel-type function precisely

$$|\boldsymbol{r}-\boldsymbol{r'}| = \sqrt{\rho^2 + r^2 - 2\rho r\cos(\varphi-\varphi')}, \qquad (13)$$

and the integrand is also

$$ds(\varphi) = \sqrt{r^2 + r'^2}d\varphi. \qquad (14)$$

Noting that $r := r(\varphi)$ and $r' := r'(\varphi)$. Moreover, $K_m = \partial G_m/\partial\nu$. In regards to (8) and (11), the scattered field has an integral representation as a combination of single and double potential integral operators [9] as:

$$u^s(\boldsymbol{r}) = \int_{\partial\Omega}\Big(u_0(\boldsymbol{r'})K_0(\boldsymbol{r};\boldsymbol{r'}) - \psi_0(\boldsymbol{r'})G_0(\boldsymbol{r};\boldsymbol{r'})\Big)\,ds(\varphi'). \qquad (15)$$

Now, by substituting (7) into (11) and considering the jump relations, the subsequent classical set of the boundary integral equations are obtained [35]:

$$\begin{aligned} u^i(\boldsymbol{r}) &= \frac{1}{2}u_0(\boldsymbol{r}) - \fint_{\partial\Omega} K_0(\boldsymbol{r};\boldsymbol{r'})u_0(\boldsymbol{r'})\,ds(\varphi') \\ &\quad + \int_{\partial\Omega} G_0(\boldsymbol{r};\boldsymbol{r'})\psi_0(\boldsymbol{r'})\,ds(\varphi') \\ 0 &= -\frac{1}{2}u_0(\boldsymbol{r}) - \fint_{\partial\Omega} K_1(\boldsymbol{r};\boldsymbol{r'})u_0(\boldsymbol{r'})\,ds(\varphi') \\ &\quad + \int_{\partial\Omega} G_1(\boldsymbol{r};\boldsymbol{r'})\psi_0(\boldsymbol{r'})\,ds(\varphi'). \end{aligned} \qquad (16)$$

Accordingly, one can find the unknown surface currents by the numerical solution of the integral equations such as the method of moments (MoM) [35] and then obtain the scattered field using the integral representation given in (15). It is worth noting that the recursive inverse algorithm described in Section 4 also needs to solve the direct problem for the reconstructed shape at each iteration step. Thus, an accurate solution for the direct scattering case, for which a numerical MoM solution is applied in this study, is essential. To verify the numerical solution of the direct problem, the following subsection is designed by considering scattering from an infinite-length cylinder, which has an analytical expression in terms of the Mie series.

3.1. Validation: Scattering by dielectric cylinder

Consider an infinitely long cylinder with a cross-section radius $r = a$ located at the origin. Suppose it is a non-magnetic lossy dielectric with complex wavenumber k_1 and embedded in free space with the wavenumber k_0. An incident plane wave with the angle of incidence φ_i illuminates the cylinder. The plane wave can be expressed with the infinite Bessel series by Jacobi-Anger identity [36]:

$$\begin{aligned} u^i(r,\varphi) &= e^{ik_0 r\cos(\varphi-\varphi_i)} \\ &= \sum_{n=-\infty}^{n=+\infty} i^n J_n(k_0 r)e^{-in(\varphi-\varphi_i)} \end{aligned} \qquad (17)$$

The fields inside and outside the cylinder are denoted u_0 and u_1, respectively. They both satisfy the scalar Helmholtz equation regarding (5) and continuous on the boundary ($r = a$) as in (7). It is worth reminding that $\hat{\nu} = \hat{\rho}$ for the circular cylinder. They have series representations

$$\begin{aligned} u_0(r,\varphi) &= \sum_{n=-\infty}^{n=+\infty}\Big(B_n J_n(k_0 r) + C_n H_n^{(1)}(k_0 r)\Big)e^{-in\varphi}, \\ u_1(r,\varphi) &= \sum_{n=-\infty}^{n=+\infty} A_n J_n(k_1 r)e^{-in\varphi}, \end{aligned} \qquad (18)$$

where $B_n = i^n e^{in\varphi_i}$ according to (17). Substituting (18) into (7) yields

$$\begin{aligned} B_n J_n(k_0 a) + C_n H_n^{(1)}(k_0 a) &= A_n J_n(k_1 a) \\ B_n J_n'(k_0 a) + C_n H_n'^{(1)}(k_0 a) &= \frac{k_1}{k_0}A_n J_n'(k_1 a). \end{aligned} \qquad (19)$$

Leading with $\varsigma = k_1/k_0$, the solution reads:

$$\begin{aligned} A_n &= B_n\frac{H_n^{(1)}(k_0 a)J_n'(k_0 a) - H_n'^{(1)}(k_0 a)J_n(k_0 a)}{\varsigma H_n^{(1)}(k_0 a)J_n'(k_1 a) - H_n'^{(1)}(k_0 a)J_n(k_1 a)}, \\ C_n &= B_n\frac{J_n(k_1 a)J_n'(k_0 a) - \varsigma J_n(k_0 a)J_n'(k_1 a)}{\varsigma H_n^{(1)}(k_0 a)J_n'(k_1 a) - H_n'^{(1)}(k_0 a)J_n(k_1 a)}. \end{aligned} \qquad (20)$$

Accordingly, one can compute the total fields inside and outside the cylinder by substituting (20) into (18). In the context of the numerical MoM-point matching solution, the whole cylinder sampled as follows: The circular cylinder with radius $r(\varphi) = a$ has a circumference $2\pi a$ with $0 \le \varphi < 2\pi$İt is divided into the number of N equally spaced segments. n^{th} segment is denoted φ_n with its width is $\Delta\varphi$, precisely

$$\varphi_n = (n-1)\Delta\varphi \text{ with } \Delta\varphi = \frac{2\pi}{N}. \tag{21}$$

Here, $n \in [1,\ N]$ and the sample number is settled as

$$N = 10\left|\frac{k_1}{k_0}\right|2\pi a \tag{22}$$

The unknown surface fields in (16) are expanded as linear combinations of the pulse-basis subdomain functions with some unknown coefficients positioned at each segment's center. Namely,

$$\begin{Bmatrix} u_0 \\ \psi_0 \end{Bmatrix} \approx \sum_{n=1}^{N} \begin{Bmatrix} u_n \\ \psi_n \end{Bmatrix} f_n(\varphi), \tag{23}$$

where the pulse basis function,

$$f_n(\varphi) = \begin{cases} 1 & (\varphi_n - \Delta\varphi/2) \le \varphi \le (\varphi_n + \Delta\varphi/2) \\ 0 & \text{elsewhere .} \end{cases} \tag{24}$$

The length of each segment is sufficiently small so that the integrand doesn't vary significantly. Regarding point matching, the whole equation is weighted by Dirac-delta functions. Accordingly, one can obtain the matrix equation system.

$$\overline{\overline{Z}}_{die}\,\overline{I}_{die} = \overline{V}_{die}, \tag{25}$$

where $\overline{I}_{die}$ and $\overline{V}_{die}$ are vectors with size $2N \times 1$. The elements of the vectors and the impedance matrices are precisely given in the appendix. The numerical comparison of the analytic and MoM solutions for a dielectric cylinder is given in the following subsection.

3.2. Numerical comparisons for a dielectric cylinder

To validate MoM-point matching with the analytic Mie series solution, a dielectric lossy cylinder with radius $r = 2$m is considered. Outside of the cylinder is free space, and the dielectric parameters are $\varepsilon_r = 4$ and $\sigma = 5 \times 10^{-5}$. In (18), $n = 64$ and in (22), $N = 252$. Operating frequency is 300 MHz, the angle of the incident plane wave illumination $\varphi = 60°$. The following figures 2 and 3 show the modulus and the phase of the surface fields acquired by MoM and analytic series, respectively.

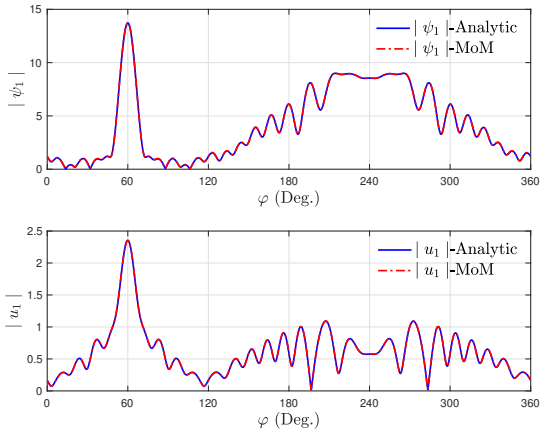


Figure 2. Modulus of the surface fields on dielectric cylinder

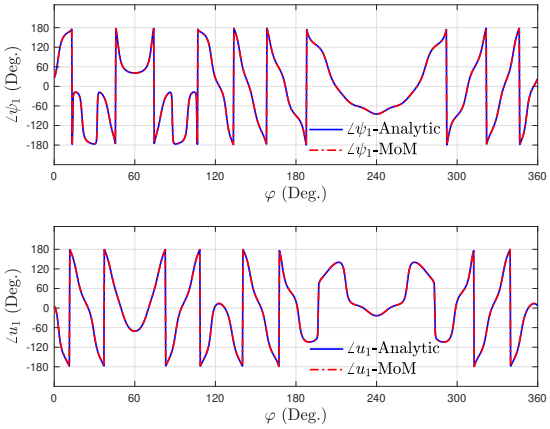


Figure 3. phase of the surface fields on dielectric cylinder

To compare the results quantitatively, an ℓ_2 norm-based error is defined between the fields obtained from analytic and MoM solutions:

$$e(\%) = \frac{\|u_{\mathrm{A}} - u_{\mathrm{MoM}}\|_2}{\|u_{\mathrm{A}}\|_2} \times 100. \tag{26}$$

Here, u_{A} and u_{MoM} denote the surface fields obtained by analytic and MoM solutions, respectively. The obtained surface fields with MoM solution in the figures 2 and 3 requires sampling $N = 252$. High agreements were achieved between the MoM and analytical solutions. The numerical errors obtained for this sampling number are below 2% for both u_0 and ψ_0. As expected, increasing the number of samples in MoM improves the agreement between the two methods. Hence, the quantitative error decreases for a higher sampling number. The obtained errors for increased N values are shown in Fig. 4

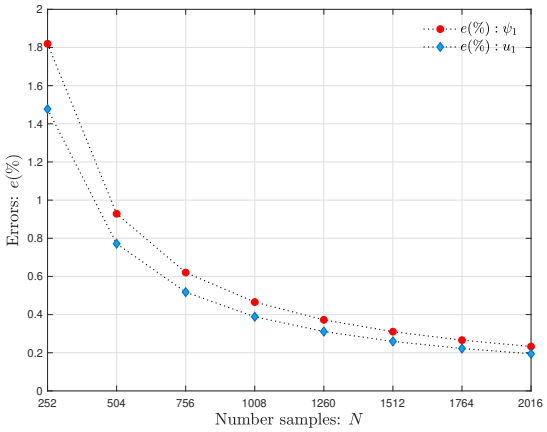


Figure 4. Errors vs. the number of samples in MoM

4. Inverse Problem

The inverse problem addressing involves determining the boundary $\partial\Omega$ of the scatterer Ω, given the far-field pattern u_∞ for the superposition of the incident plane wave illuminations u^i. To define the far-field pattern precisely, it is necessary to consider the asymptotic behavior of the Hankel type functions for large argument, which equals [36]

$$H_0^{(1)}(\omega) = \sqrt{\frac{2}{\pi\omega}} e^{i\omega - \frac{\pi}{4}} \left(1 + \mathcal{O}\left(\frac{1}{\omega}\right)\right), \omega \to \infty. \tag{27}$$

Within this context, assume that a point in the far field is described as $\boldsymbol{r_s} := r_s(\hat{x}_1 \cos(\varphi_s) + \hat{x}_2 \sin(\varphi_s))$ and the cross-section is represented as in (1). Then, in accordance with Huygens' principle [37], the standard approximation for the 2D Green's function given in (12) and (13) have the phase term and the amplitude term as following

$$\tilde{G}(\boldsymbol{r}; \boldsymbol{r_s}) = \gamma e^{ik_0 r_s} e^{-ik_0 r(\varphi)\cos(\varphi_s - \varphi)}, \tag{28}$$

where the constant

$$\gamma = \frac{i}{4}\sqrt{\frac{2}{\pi k_0 r_s}} e^{-i\frac{\pi}{4}}. \tag{29}$$

That is, the modulus term of $\tilde{G}$ is approximated as $|\boldsymbol{r} - \boldsymbol{r_s}| \approx r_s$ and the phase term is $|\boldsymbol{r} - \boldsymbol{r_s}| \approx r_s - k_0(\boldsymbol{r_s} \cdot \boldsymbol{r})$. It is naturally a good approximation for Green's functions and is conventionally applied to represent a far-field pattern of the scattered field. The reader may refer to [38] for the Huygens' principle in 3D and 2D scattering problems [39], for the details of the far-field expansion of 2D Green's function, and thus the far-field expansion of the Greens' function [34, 37]. In this context, one can easily define the derivative of the function with respect to the surface normal $\hat{\nu}$ as

$$\frac{\partial \tilde{G}(\boldsymbol{r}; \boldsymbol{r_s})}{\partial \nu} = \hat{\nu} \cdot \nabla \tilde{G}(\boldsymbol{r}; \boldsymbol{r_s}) = -i\tilde{G}(\boldsymbol{r}; \boldsymbol{r_s})\hat{k}_s \cdot \hat{\nu}. \tag{30}$$

Here, $\hat{k}_s$ is the wavenumber vector in scattering direction with the angle φ_s, particularly equals to

$$\hat{k}_s = k_0\Big(\hat{x}_1 \cos(\varphi_s) + \hat{x}_2 \sin(\varphi_s)\Big), \tag{31}$$

which can be converted into polar coordinates:

$$\hat{k}_s = k_0\Big(\hat{\rho} \cos(\varphi_s - \varphi) + \hat{\varphi} \sin(\varphi_s - \varphi)\Big). \tag{32}$$

Accordingly, the far-field scattered field has the integral representation as

$$u_\infty(\varphi) = \gamma \int_{\partial\Omega} \Big(- i\hat{k}_s \cdot \hat{\nu} u_0(\varphi') + \psi_0(\varphi') \Big) e^{-ik_0 r(\varphi')\cos(\varphi_s - \varphi')}\, ds(\varphi'). \tag{33}$$

Substituting (6) and (31) into (33) yields more precise expression for u_∞ as

$$\begin{aligned} u_\infty(\varphi) = -ik_0\gamma \int_0^{2\pi} \Big\{ \Big(& r(\varphi')\cos(\varphi_s - \varphi') \\ & - r'(\varphi')\sin(\varphi_s - \varphi') \Big) u_0(\varphi') \\ & + \psi_0(\varphi') \sqrt{r(\varphi')^2 + r'(\varphi')^2} \Big\} \\ & e^{-ik_0 r(\varphi')\cos(\varphi_s - \varphi')}\, d\varphi'. \end{aligned} \tag{34}$$

To have a much more compact form of (34), the integral equation can be defined in an operator form $\mathcal{D}$. Thus, with the knowledge of the u_∞ and the surface fields, the inverse problem consists of solving the nonlinear and ill-posed equation

$$\mathcal{D}(r, u_0, \psi_0) = u_\infty, \tag{35}$$

for the unknown boundary $\partial\Omega$ represented by $r := r(\varphi)$. To start with the inversion process, first, the operator is linearized via Newton's type iterations and then regularized by Tikhonov. Within the context of linearization, let r_0 be the initially guessed shape, for which one can solve the direct problem to obtain the surface currents (u_{1_0}, ψ_{1_0}) of the guessed shape. Accordingly, the linearization proceeds in the sense of Newton as

$$\mathcal{D}(r,u_0,\psi_0)\approx\mathcal{D}(r_0,u_{0_0},\psi_{0_0})+\mathcal{D}'(r_0;u_{0_0},\psi_{0_0})\delta r_0. \tag{36}$$

Here, $\mathcal{D}'(r_0;u_{1_0},\psi_{1_0})\delta r_0$ is the Frechet derivative of the surface with respect to r, and δr_0 is the updated correlation function for which (36) has to be solved. For the regularization procedure, let $\mathcal{D}_0'$ stand for the Frechet derivative for a short notation, and its ad-joint be denoted by $\mathcal{D}_0'^{\dagger}$. Then, by defining a regularization parameter $0<\tau<1$, δr_0 is the solution of

$$\tau\delta r_0+\mathcal{D}_0'^{\dagger}\mathcal{D}_0'=\mathcal{D}_0'^{\dagger}\Delta u_\infty, \tag{37}$$

where $\Delta u_\infty=u_\infty-\mathcal{D}(r_0,u_{1_0},\psi_{1_0})$. Furthermore, one may also consider a scaling (tuning) parameter to have a much more robust δr_0. In this sense, the solution of (37) is written as

$$\delta r_0=\alpha\big[\tau\mathbf{I}+\mathcal{D}_0'^{\dagger}\mathcal{D}_0'\big]^{-1}\mathcal{D}_0'^{\dagger}\Delta u_\infty. \tag{38}$$

Here, $\mathbf{I}$ is the identity matrix, and $0<\alpha<1$ is the scaling parameter. The reader may refer to [40] for the details of α and τ. For a predetermined threshold ξ, the procedure is repeated recursively until the stopping criteria $\|\delta r_N\|_2\leq\xi$. Accordingly, n^{th} approximated boundary is updated by setting

$$r_{n+1}=r_n+\delta r_n. \tag{39}$$

Moreover, to have a more robust reconstruction, the solution is obtained via the least squares [41]. To this aim, the update correlation is expanded by the linear combination of some basis functions $\Phi_q(\varphi)$, $q=1,\cdots,Q$ as

$$\delta r(\varphi)=\sum_{q=1}^{Q}a_q\Phi_q(\varphi). \tag{40}$$

Hence, the problem turns into finding unknown coefficients of (40). For a set of grid points $\varphi^1,\cdots,\varphi^P$, the unknown coefficients are determined by minimizing the sum of squares at n^{th} iteration, *i.e.,*

$$\sum_{p=1}^{P}\left|\mathcal{D}'(r_n,u_{0_n},\psi_{0_n})\sum_{q=1}^{Q}a_q\Phi_q(\varphi^p)-\Delta u_\infty\right|^2. \tag{41}$$

The whole procedure is summarized as follows:

(i). Choose a closed curve for the initial guess $\rho=r_0(\varphi)$
(ii). Obtain the surface currents of the closed curve and thus the far-field pattern using (16) and (34), respectively.
(iii). Solve (36) and (38) in the sense of least squares (40)-(41) for the updated correlation function δr_0
(iv). Obtain the new surface profile via (39)
(v). Repeat (ii)-(iv) for n times ($n>1$) such that $r_{n+1}=r_n+\delta r_n$.
(vi). Break the loop, if $\|\delta r_n\|\leq\xi$

Here, the crucial part of the whole framework is the Frechet derivative part, which is for a mapping from a domain of functions [42]. For the sake of simplicity, one may consider the Frechet derivative $\mathcal{D}'(r_0;u_{1_0},\psi_{1_0})\delta r_0:=\mathcal{D}_0'$ as the superposition of two operators

$$\mathcal{D}_0'=\mathcal{F}_{\mathcal{D}}'(r_0;\psi_{1_0})\delta r_0+\mathcal{F}_{\mathcal{N}}'(r_0;u_{1_0})\delta r_0 \tag{42}$$

where the Frechet operators are:

$$\begin{aligned}\mathcal{F}_{\mathcal{D}}'(r_0;\psi_{1_0})\delta r_0=&-\gamma\int_0^{2\pi}ik\cos\big(\varphi_s-\varphi'\big)\\&e^{-ikr_0(\varphi')\cos(\varphi_s-\varphi')}\psi_{1_0}(\varphi')\\&\sqrt{r(\varphi')^2+r'(\varphi')^2}\,\delta r_0 d\varphi',\end{aligned} \tag{43}$$

$$\begin{aligned}\mathcal{F}_{\mathcal{N}}'(r_0;u_{1_0})\delta r_0=&-\gamma\int_0^{2\pi}\Bigg(ik\cos\big(\varphi_s-\varphi'\big)\\&\kappa(r_0,\varphi',\varphi_s)u_{1_0}(\varphi')\\&+\frac{\partial\kappa(r_0,\varphi',\varphi_s)}{\partial r_0}u_{1_0}(\varphi')\Bigg)\\&e^{-ikr_0(\varphi')\cos(\varphi_s-\varphi')}\,\delta r_0 d\varphi',\end{aligned} \tag{44}$$

and κ function

$$\kappa(r_0,\varphi,\varphi_s)=r_0\cos(\varphi_s-\varphi)-r_0'\sin(\varphi_s-\varphi) \tag{45}$$

is basically the result of $\hat{k}_s\cdot\hat{\nu}$.

5. Numerical results and discussion

The section is reserved to demonstrate the feasibility of the proposed inverse framework. For all considered scattering scenarios, the operating frequency is 300MHz so that the wavelength in free-space $\lambda_0=1$m. For all numerical examples, the predetermined threshold is $\xi=0.07$. Except for

one example, the far-field pattern is assumed to be known at 64 points equally distributed around the unit circle. For the expansion of the unknown updated correlation function with some basis functions, $\Phi_q(\varphi) = e^{-iq\varphi}$, $q = 0, \pm 1, \cdots \pm Q$. To verify the success of the reconstructions quantitatively, an ℓ_2-norm based error is defined precisely

$$err(\%) = \frac{\|r(\varphi) - r_n(\varphi)\|}{\|r(\varphi)\|} \times 100, \qquad (46)$$

where $r(\varphi)$ and $r_n(\varphi)$ represent the actual and the reconstructed surfaces, respectively.

The first example aimed to put forth the effect of the penetrability of the object on the inverse algorithm. To this aim, a kite-like object is considered for reconstruction, considering both PEC and dielectric cases. The unknown kite-like surface is a radial function

$$r(\varphi) = 1.5\big(1 + 0.15\cos(3\varphi)\big). \qquad (47)$$

For the dielectric case, the constitutive EM parameters are $\varepsilon_1 = 4\varepsilon_0$ and $\sigma = 10^{-5}$ (S/m). For both PEC and dielectric cases, the region is illuminated by 7 incident illuminations simultaneously, for which the angles of incidence are selected in the range: $0° : \Delta\varphi_i : 330°$ with the $\Delta\varphi_i = 55°$ angular increments. The reconstructions and the actual surface for PEC and Dielectric scenarios are shown in Fig. 5 and Fig. 6, respectively.

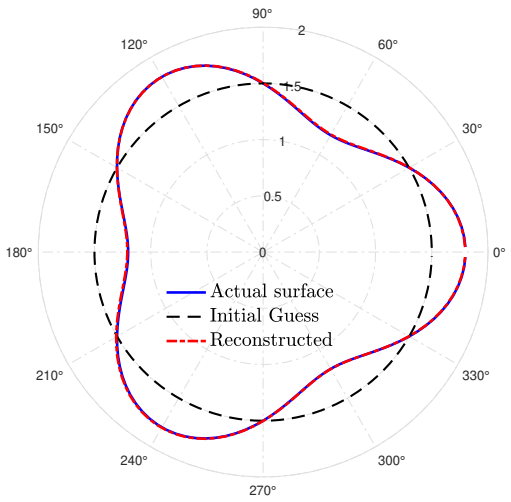


Figure 5. Reconstruction of the surface defined in (47) for PEC case.

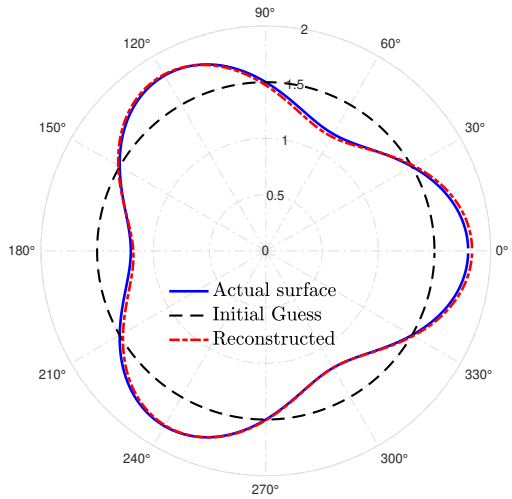


Figure 6. Reconstruction of the surface defined in (47) for Dielectric (penetrable) case.

It is worth noting that for the PEC case, the far-field pattern of the scattered field is represented with a single layer potential as shown in [32] in detail. Accordingly, for the Frechet derivative of the PEC case, (43) should be taken into consideration. As illustrated in the figure, a cross-section of an infinitely long cylinder with a radius $1.5\lambda_1$ is considered as the initial guess for both cases. Moreover, both cases' stopping criteria are $\|\delta r\| \leq \xi = 0.07$. Accordingly, the needed 16 iteration is for the PEC case, and the 28 iteration is for the dielectric case. To visualize the expected decreasing tendency of the $\|\delta r\|$ for each new iteration, Fig. 7 shows $\|\delta r\|$ versus the number of iterations for the dielectric case.

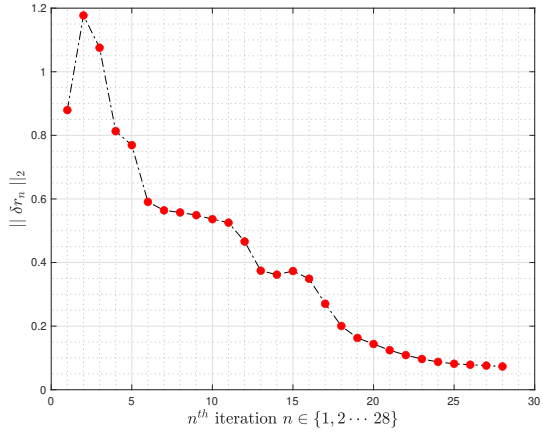


Figure 7. $\|\delta r_n\|$ vs. iterations.

In addition, the reconstruction of PEC is better than the dielectric case qualitatively and quantitatively. This is mostly because the PEC case has no penetrated field to the second region. Hence, a more powerful scattered field contains more information for the scatterer, leading to better

reconstructions. The quantitative errors for both cases are $err(\%) = 0.31\%$ and $err(\%) = 1.37\%$, respectively, for PEC and dielectric cases.

The next analysis covers the algorithm's sensitivity to the constitutive parameters of the dielectric object, *i.e.,* ε_r and σ. Accordingly, the response of the algorithm is tested for higher and lower dielectric permittivity and conductivity values. In this regard, "4− leaf" shape boundary curve is assumed to be unknown, and it is reconstructed for different ε_r and σ values. The considered "leaf-shaped" radial function is defined as

$$r(\varphi) = 1.3\big(1 + 0.15\cos(4\varphi)\big) \tag{48}$$

First, the objective is to observe the dielectric dependency of the iterative inverse reconstruction algorithm. In order to notice the sensitivity to the dielectric permittivity, the algorithm is run for different ε_r values for a fixed conductivity $\sigma = 10^{-5}$ (S/m). Within this context, the dielectric permittivity range is taken into account $\varepsilon_r \in [2, 10]$. Hence, the dielectric permittivities are defined in a wide range, from very penetrable cases to high levels. For all reconstructions, run for different ε_r, the circle with radius $1.5\lambda_1$ is considered for the initial guess. The obtained quantitative errors of the reconstructions for different dielectric permittivity values are shown in Fig. 8.

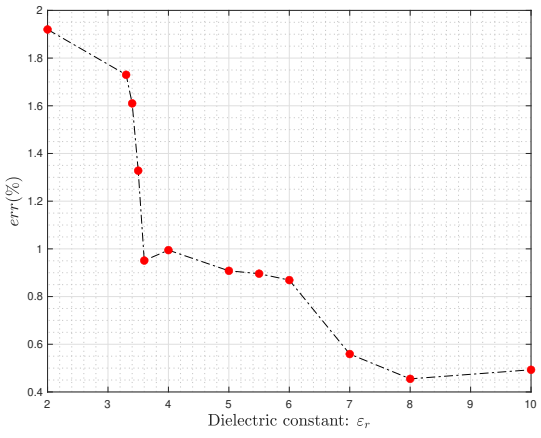


Figure 8. Error: $err(\%)$ vs. dielectric constant: ε_r

As shown, the bigger contrast between free space and the dielectric object yields better reconstructions quantitatively. It is worth noting that if $\varepsilon_r < 2$, it becomes impossible to obtain accurate reconstructions. Furthermore, $\varepsilon_r > 10$ yields higher computational cost as the inversion algorithm needs a direct solver whose unknown is directly related to the dielectric permittivity ε_r. A similar analysis was also carried out to observe the conductivity sensitivity. To this aim, $\varepsilon_r = 2$ is fixed and the conductivity varies in the range $\sigma \in [5 \times 10^{-7}, 10^{-2}]$ (S/m). The algorithm's errors for different conductivity values are shown in Fig. 9

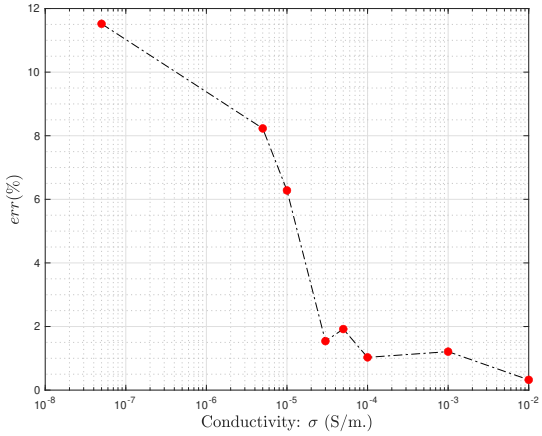


Figure 9. Error: $err(\%)$ vs. conductivity σ (S/m)

Accordingly, the higher conductivity yields better reconstructions such that for $\sigma = 10^{-2}$, the error $err < 0.4\%$. However, such a high conductivity yields a huge loss, so the unknown object can almost turn into PEC rather than a penetrable object. Fig. 10 shows the worst and the best cases together to demonstrate reconstructions visually.

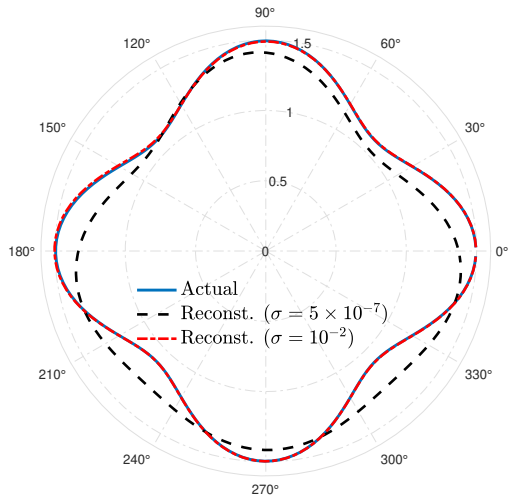


Figure 10. Reconstructions for the highest and the lowest conductivity values

Noting that the region is illuminated for 12 incident plane waves simultaneously where the angles of incidence are defined as $0° : \Delta\varphi_i : 330°$ with the $\Delta\varphi_i = 30°$ angular increments. Again, the

circle with the radius $1.5\lambda_1$ is considered the initial guess. For all analyses conducted to obtain the sensitivity to the constitutive parameters, 19 exponential-type basis functions are applied in the sense of least squares.

The next analysis covers the algorithm's sensitivity against noise. For this purpose, a synthetic noise is added to the far-field pattern. The noisy scattered field is defined as $\tilde{u}_\infty = u_\infty + n_\ell |u_\infty| e^{i2\pi r_d}$, where n_ℓ is the noise-level and r_d is the random number in the interval $0 < r_d < 1$. The bean-shaped object is considered for the noise analyses. It is defined as

$$r(\varphi) = 0.8\frac{1 + 0.85\cos\left(\varphi + \frac{\pi}{4}\right) + 0.05\sin\left(2\varphi + \frac{\pi}{4}\right)}{1 + 0.5\cos\left(\varphi + \frac{\pi}{4}\right)}. \tag{49}$$

In regards to the scattering scenario, $\varepsilon_r = 4$ and $\sigma = 10^{-5}$ (S/m), and the number of 6 incident illuminations is applied simultaneously, where $\varphi_i^n = \{0°, 60°, 120°, 180°, 240°, 300°\}$. Fig. 11 shows the obtained quantitative errors for different noise levels.

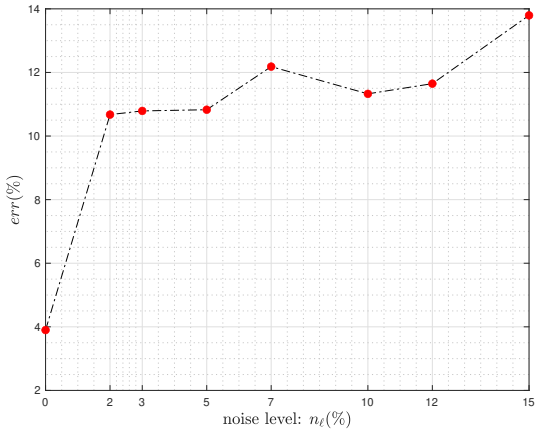


Figure 11. Error: $err(\%)$ vs. noise level: $n_\ell(\%)$

Accordingly, the algorithm is sensitive to the noise such that $n_\ell(\%) \leq 10\%$ for satisfactory reconstructions. The reconstructions for the noise-free case and with the highest noise level are shown in Fig. 12

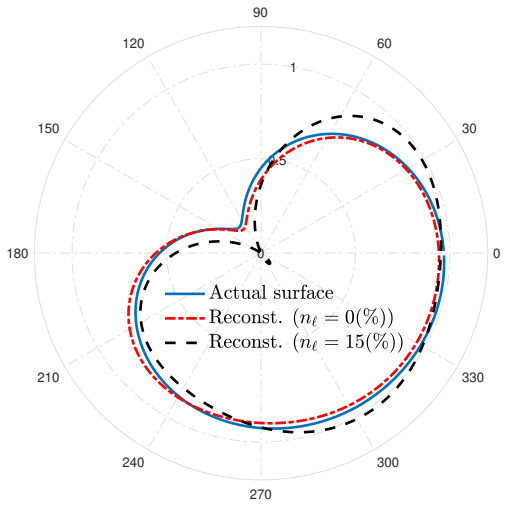


Figure 12. Reconstruction of the bean-shaped object for the noise-free and the highest noise level

The following example investigates the sensitivity of the reconstructions to the amount of scattered field data. Let #MP denote the number of measurement points. It is worth to remind that, up to this example, #MP = 64. To this aim, a 5-leaf shape is reconstructed for different numbers of scattered field data. The inaccessible 5-leaf shape is defined as

$$r(\varphi) = 1.3\big(1 + 0.15\cos(5\varphi)\big) \tag{50}$$

For the sensitivity analysis to the number of scattered field data, the remaining parameters are kept constant such as the number of incident fields is 11, precisely defined in the range $\varphi_i = 0° : 30° : 300°$, the penetrable medium parameters are $\varepsilon_r = 4$, $\sigma = 10^{-5}$ S/m and the number of applied exponential basis functions are 27. The obtained error vs the amount of the measured field data is given in Fig. 13

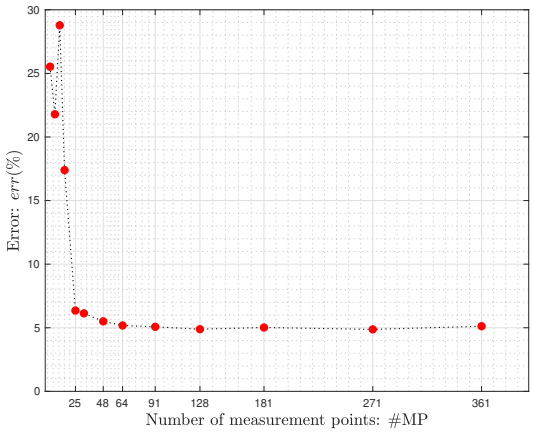


Figure 13. Error: $err(\%)$ vs. the number of measurement points: #MP

Accordingly, insufficient reconstructions observed for#MP < 25. For an accurate result, #MP $\geq$

25. As shown in Fig. 13, the error remains around 5% such that the differences between the reconstructions cannot be distinguished with the naked eye. To visualize this, the unsuccessful reconstructions obtained for #MP = 8, #MP = 16 and the satisfactory result of #MP = 32 are shown in Fig. 14

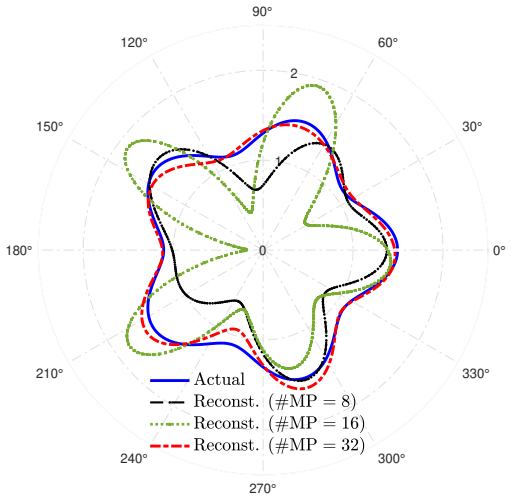


Figure 14. Reconstructions for different #MP values

The next analysis is carried out to emphasize the significance of multi-illumination. For this purpose, a potato-shaped curved object is considered. It is defined as a radial function

$$r(\varphi) = 4\sqrt{\left(\tfrac{1}{4} + \tfrac{3}{40}\cos(2\varphi)\right)^2 + \left(\tfrac{1}{4} + \tfrac{3}{100}\cos(3\varphi)\right)^2}. \tag{51}$$

In the first case, the object is illuminated with a single incident plane wave with the angle of incidence $\varphi_i = 0°$. The result is shown in Fig. 15

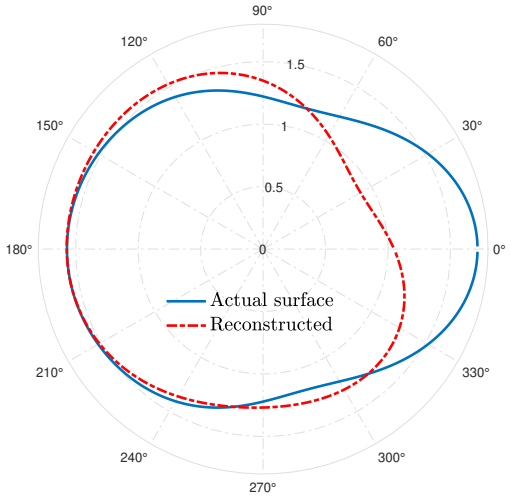


Figure 15. Reconstruction of the potato-shaped object with a single incident illumination

As illustrated, the reconstruction is unsatisfactory as the shadow region predominates inversion [24,33,43]. To overcome this, the same object is recovered for the superposition of 4 incident plane waves. The angles of incidence are $\varphi_i^n = \{-30, 30, 150, 210\}$, for $n = \{1, 2, 3, 4\}$. The satisfactory reconstruction for the multi-incidence is illustrated in Fig. 16

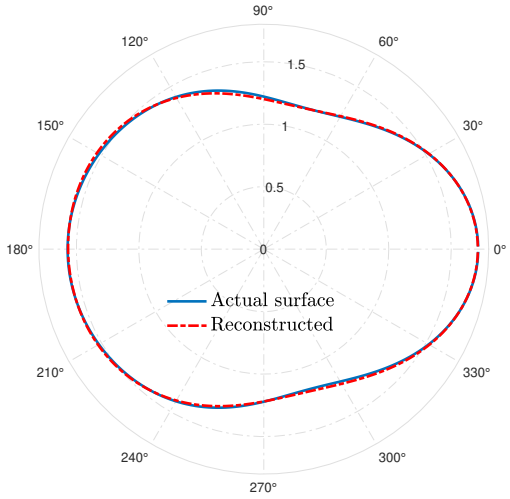


Figure 16. Reconstruction of the potato-shaped object with superposition of multi-incident illumination

Noting that, for both multi and single-illumination cases of potato shape objects, $\varepsilon_r = 4$ and $\sigma = 10^{-5}$ (S/m) and there are 13 exponential basis functions were utilized. Actually, the superposition of multi-incidence illumination is one of the essential factors for satisfactory reconstructions. To underline this, all 2D shapes reconstructed so far have been reconstructed again by considering different numbers of incident illumination in the following analysis. The error for different number of illuminations are shown in Fig. 17

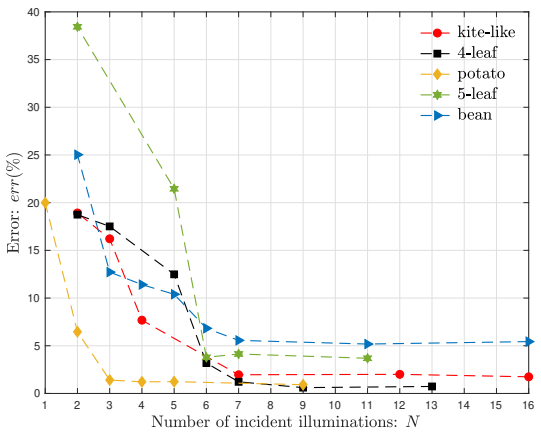


Figure 17. Error: $err(\%)$ vs. number of multi-incidence illumination: N

Table 1. Parameters for the superposition of the multi-incident illumination

$r(\varphi)$	Angels of incidence: $\varphi_i^1 : \Delta\varphi : \varphi_i^N$	increments: $\Delta\varphi$ (Deg.)	number of incident field in superposition: N
3-leaf	$0° : \Delta\varphi : 330°$	{22, 30, 55, 110, 165, 330}	{2, 3, 4, 7, 12, 16}
4-leaf	$0° : \Delta\varphi : 360°$	{30, 45, 60, 72, 90, 180, 360}	{2, 3, 5, 6, 7, 9, 13}
5-leaf	$0° : \Delta\varphi : 300°$	{20, 30, 45, 60, 75, 90, 150, 300}	{2, 3, 4, 5, 6, 7, 11, 16}
potato	$-30° : \Delta\varphi : 210°$	{30, 60, 80, 120, 240, }	{1, 2, 3, 4, 5, 9} (Single illumination at 0°)

As shown in Fig. 17, there are no satisfactory reconstructions for a single or double illumination. As expected, the error starts to decrease for increasing the number of illuminations, and after a specific number, it remains almost constant for each specific reconstruction. There is no certain value because it differs for every shapes. However, one can conclude that at least 2 incidence illumination should be considered even for a simple object (like the potato). The details of the analysis are given in Table 1. The table states the incident fields angles, given in (4), by defining $\varphi_i^1 : \Delta\varphi : \varphi_i^N$. Here, φ_i^1 and φ_i^N are the initial and the final angles of the illumination, and $\Delta\varphi$ denotes the increments.

The next example shows the reconstruction of a peanut-shaped object for the same constitutive EM parameters with different initial guessed surfaces. The object is defined with a radial function

$$r(\varphi) = 0.7\sqrt{\left(0.2\cos^2(\varphi) + \sin^2(\varphi)\right)}. \qquad (52)$$

Two incident illuminations were utilized with $\varphi_i = \{90°, 270°\}$. In this example, the reconstruction carried out for considering both a circular cylinder with a radius $0.6\lambda_1$ and an ellipse defined as:

$$r_0(\varphi) = \frac{ab}{\sqrt{\left(a\cos^2(\varphi) + b\sin^2(\varphi)\right)}}, \qquad (53)$$

where $a = 0.6\lambda_1$ and $b = 0.4\lambda_1$. Since the elliptical initial guess is more similar to the actual peanut-shaped object, one may consider that using an elliptical cylinder as the initial guess would lead to a better reconstruction. However, both qualitatively and quantitatively, the difference between reconstructions is almost negligible. It is obtained $err = 3.73\%$ with the circular cylinder initial guess and $err = 3.61\%$ for the case of the elliptical cylinder. The only difference is that it requires 7 iterations for the elliptical case, whereas it costs 11 for the circular initial guess. Thus, the method is stable and robust. Fig. 18 demonstrates the reconstruction, the elliptical initial guess, and the actual surface.

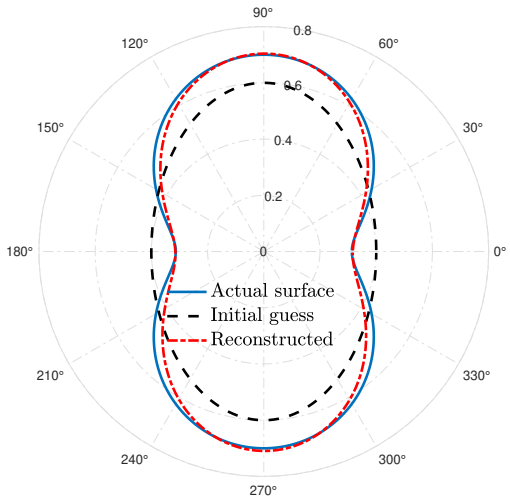


Figure 18. Reconstruction of the peanut-shaped object with the initial guess

For the peanut reconstruction cases, $n_\ell = 2\%$ noise is applied for all the simulations, and 13 exponential functions are applied in the least square sense.

6. Conclusion

A regularized and linearized iterative framework is presented to recover the shape of inaccessible 2-dimensional dielectric objects. The proposed framework utilizes the far-field pattern of the scattered field data for this aim. The inversion is done in accordance with the boundary integral representation of the far-field pattern considering the combination of the double and single-layer potentials. The problem is inherently ill-posed and nonlinear. Within this context, the Newton-type iterative linearization technique is applied, and it is regularized via Tikhonov in the sense of the least squares approach. To overcome the adverse effect of the shadow region on the imaging process, the superposition of the multi-incident wave

is taken into account. Accordingly, robust and fast inversion is achieved with a very low computational cost. The feasibility of the proposed framework and its validation limits are asserted via various scattering scenarios.

The algorithm may be extended for three-dimensional imaging problems in the acoustic case, as it still requires a scalar solution to the wave equations. However, the EM case must be reformulated, as the scattering problem needs a vectorial solution. The validation limits can be enlarged with hybrid approaches generated with deep learning techniques. All these issues are left as future works.

Appendix

$\overline{I}_{die}$ and $\overline{V}_{die}$ are vectors with size $2N \times 1$ whose elements are precisely

$$\overline{I}_{die} = \left[u_0\, u_1 \cdots u_N\, \psi_0\, \psi_1\, \cdots \psi_N \right]^T, \qquad \text{(A.1)}$$

and tested incident fields with Dirac-delta function yields:

$$\overline{V}_{die} = \left[u^i(\varphi_1)\; u^i(\varphi_2) \cdots u^i(\varphi_N) \underbrace{0\;\; 0 \cdots 0}_{N} \right]^T. \qquad \text{(A.2)}$$

Finally, the impedance matrix has a size of $2N \times 2N$, which is composed of 4 sub-matrices, each of which has a size $N \times N$:

$$\overline{Z}_{die} = \begin{bmatrix} \overline{Z^{1u}} & \overline{Z^{1\psi}} \\ \overline{Z^{2u}} & \overline{Z^{2\psi}} \end{bmatrix}. \qquad \text{(A.3)}$$

The elements of $\overline{Z^{j\psi}}$ $(j = \{1, 2\})$ are $N \times N$ are given as

$$Z^{j\psi}_{mn} = \Delta\varphi_n \begin{cases} G_j(\boldsymbol{r_m}; \boldsymbol{r_n}) & m \neq n \\ \frac{i}{4}\left[1 + i\frac{2}{\pi} \ln\left(\frac{\gamma k_0}{4e} \Delta\varphi_n \right) \right] & m = n. \end{cases} \qquad \text{(A.4)}$$

Here, $\gamma = 1.78107$. The sub-matrix $\overline{Z^{1u}}$ has the elements

$$Z^{1u}_{mn} = \Delta\varphi_n \begin{cases} -K_0(\boldsymbol{r_m}; \boldsymbol{r_n}) & \text{for } m \neq n \\ \frac{1}{2} & \text{for } m = n. \end{cases} \qquad \text{(A.5)}$$

$\overline{Z^{2u}}$ has the same format with (A.5) taking the wavenumber k_1 instead of k_0 into account. Moreover, the diagonal elements of $\overline{Z^{2u}} = -1/2$ in accordance with the (16).

References

[1] Tuz, M.(2017). Boundary values for an eigenvalue problem with a singular potential. *An International Journal of Optimization and Control: Theories & Applications (IJOCTA)*, 7(3), 293–300. https://doi.org/10.11121/ijocta.01.2017.00507

[2] Aydin, C., & Tezer, M. (2019). the DRBEM solution of cauchy MHD duct flow with a slipping and variably conducting wall using the well-posed iteration. (2019) *An International Journal of Optimization and Control: Theories & Applications (IJOCTA)*, 9(3), 76–85. https://doi.org/10.11121/ijocta.01.2019.00677

[3] Karatas Akgul, E. (2018). Reproducing kernel Hilbert space method for solutions of a coefficient inverse problem for the kinetic equation. *An International Journal of Optimization and Control: Theories & Applications (IJOCTA)*, 8(2), 141–151. https://doi.org/10.11121/ijocta.01.2018.00568

[4] Ozmen, A. (2022). Multi-objective regression modeling for natural gas prediction with ridge regression and CMARS. *An International Journal of Optimization and Control: Theories & Applications (IJOCTA)*, 12(1), 56–65. https://doi.org/10.11121/ijocta.2022.1084

[5] Acil, M., & Konuralp, A. (2021). Reconstruction of potential function in inverse Sturm-Liouville problem via partial data. *An International Journal of Optimization and Control: Theories & Applications (IJOCTA)*, 11(2), 186–198. https://doi.org/10.11121/ijocta.01.2021.001090

[6] Sefer, A. (2022). Locally perturbed inaccessible rough surface profile reconstruction via phaseless scattered field data. *IEEE Transactions on Geoscience and Remote Sensing,*, 60, 1–8. https://doi.org/10.1109/TGRS.2021.3105257

[7] Bilgin, E., Cayoren, M., Joof, S., Cansiz, G., Yilmaz, T., & Akduman, I. (2022). Single-slice microwave imaging of breast cancer by reverse time migration. *Medical Physics*, 49(10), 599–6608. https://doi.org/10.1002/mp.15917

[8] Dogu, S., Akinci, M. N., & Gose, E. (2022). Experimental moving target imaging in a nonanechoic environment with linear sampling method. *IEEE Geoscience and Remote Sensing Letters*, 18(3), 441–445. https://doi.org/10.1109/LGRS.2020.2976594

[9] Colton, D., & Kress, R. Smith, G. (2019). Ill-posed Problems. In: P. Holmes, S.S. Antman, K. Sreenivasan eds. *Inverse Acoustic and Electromagnetic Scattering Theory.* Wiley - 3rd ed. 4th ed. Springer, London. 95–118.

https://doi.org/10.1007/978-1-4614-4942-3_4

[10] Ghazi, F. F., & Tawfiq, L. N. M. (2024). Design optimal neural network based on new LM training algorithm for solving 3D - PDEs. *An International Journal of Optimization and Control: Theories & Applications (IJOCTA)*, 14(3), 249–250. https://doi.org/10.11121/ijocta.1519

[11] Ozmen, A. (2022). Multi-objective regression modeling for natural gas prediction with ridge regression and CMARS. *An International Journal of Optimization and Control: Theories & Applications (IJOCTA)*, 12(1), 56–65. https://doi.org/10.11121/ijocta.2022.1084

[12] Kumar A., Kumar, M., & Goswami, P.(2024). Numerical solution of coupled system of Emden-Fowler equations using artificial neural network technique. *An International Journal of Optimization and Control: Theories & Applications (IJOCTA)*, 14(1), 62–73. https://doi.org/10.11121/ijocta.1424

[13] Salucci, M., Arrebola, M., Shan, T., & Li, M. (2022). New frontiers in real-time inverse scattering and electromagnetic imaging. *IEEE Transactions on Antennas and Propagation*, 70(8), 6349–6364. https://doi.org/10.1109/TAP.2022.3177556

[14] Desmal, A. (2022). A trained iterative shrinkage approach based on born iterative method for electromagnetic imaging. *IEEE Transactions on Antennas and Propagation*, 70(11), 4991–4999. https://doi.org/10.1109/TMTT.2022.3205650

[15] Desmal, A., & Alsaei, J. (2024). Multifrequency trained projection nonlinear framework for electromagnetic imaging with contrast-source landweber-kaczmarz. *IEEE Geoscience and Remote Sensing Letters*, 21, 1–5, Art no. 3002505. https://doi.org/10.1109/LGRS.2024.3397694

[16] Yao, H. M., Sha, W.E.I.,& Jiang,L. (2019). Two-step enhanced deep learning approach for electromagnetic inverse scattering problems. *IEEE Antennas and Wireless Propagation Letters*, 18(11), 2254–2258. https://doi.org/10.1109/LAWP.2019.2925578

[17] Aydin, I., Guven, B., Sefer, A., & Yapar, A. (2022). CNN-based deep learning architecture for electromagnetic imaging of rough surface profiles. *IEEE Transactions on Antennas and Propagation*, 70(10), 9752–9763. https://doi.org/10.1109/TAP.2022.3177493

[18] Song, L., Kuang, L., Han, Y., Wang, Y., & Liu, Q. H. (2018). Inversion of rough surface parameters from SAR images using simulation-trained convolutional neural networks. *IEEE Geoscience and Remote Sensing Letters*, 15(7), 1130–1134. https://doi.org/10.1109/LGRS.2018.2822821

[19] Aydin, I., Guven, B., Sefer, A., & Yapar, A. (2022). Recovery of impenetrable rough surface profiles via CNN-based deep learning architecture. *International Journal of Remote Sensing*, 43(15-16), 5658–5685. https://doi.org/10.1080/01431161.2022.2105177

[20] Wang, X., Zhu, J., Song, M., & Wu, W. (2022). Fourier method for reconstructing elastic body force from the coupled-wave field. *Inverse Problems and Imaging*, 16(2), 325–340. https://doi.org/10.3934/ipi.2021052

[21] Sefer, A., Yapar, A., & Yelkenci, T. (2024). Imaging of rough surfaces by RTM method. *IEEE Transactions on Geoscience and Remote Sensing*, 62, 1–12, Art no. 2003312, doi: 10.1109/TGRS.2024.3374972 https://doi.org/10.1109/TGRS.2024.3374972

[22] Sefer, A., & Yapar, A., (2022). Inverse scattering by perfectly electric conducting (PEC) rough surfaces: An equivalent model with line sources. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1–9, Art no. 2007109. https://doi.org/10.1109/TGRS.2022.3210657

[23] Blasten, E. L.K., & Liu, H. (2021). Scattering by curvatures, radiationless sources, transmission eigenfunctions, and inverse scattering problems. *SIAM Journal on Mathematical Analysis*, 53(4), 3801–3837. https://doi.org/10.1137/20M1384002

[24] Guo, Y., & Zhang, D. (2011). An optimization method for acoustic inverse obstacle scattering problems with multiple incident waves. *Inverse Problems in Science and Engineering*, 19(4), 461–484. https://doi.org/10.1080/17415977.2010.518286

[25] Bao, G., Li, P., Lin, J., & Triki, F. (2015). Inverse scattering problems with multi-frequencies. *Inverse Problems*, 31, 093001. https://doi.org/10.1088/0266-5611/31/9/093001

[26] Altundag, A., & Kress, R. (2012). On a two dimensional inverse scattering problem for a dielectric. *Applicable Analysis*, 91(4), 757-771. https://doi.org/10.1080/00036811.2011.619981

[27] Qu, F., Yang, J., & Zhang, H. (2019). shape reconstruction in inverse scattering by an inhomogeneous cavity with internal measurement. *SIAM Journal on Imaging Sciences*, 12(2), 788–808. https://doi.org/10.1137/18M1232401

[28] Cheng, Z., & Dong, H. (2023). Uniqueness and reconstruction method for inverse elastic wave scattering with phaseless data. *Inverse Problems and Imaging*, 18(12), 406–433.

[29] Borges, C., Gillman, A., & Greengard, L. (2017). High resolution inverse scattering in two dimensions using recursive linearization. *SIAM Journal on Imaging Sciences*, 10(2), 641–664. https://doi.org/10.1137/16M1093562

[30] Borges, C., Rachh, M., & Greengard, L. (2023). On the robustness of inverse scattering for penetrable, homogeneous objects with complicated boundary. *Inverse Problems*, 39(3), 035004. https://doi.org/10.1088/1361-6420/acb2ec

[31] Sefer, A., & Yapar, A. (2021). An iterative algorithm for imaging of rough surfaces separating two dielectric media. *IEEE Transactions on Geoscience and Remote Sensing*, 59(2), 1041–1051. https://doi.org/10.1109/TGRS.2020.2997637

[32] Johansson, T., & Sleeman, B. D. (2007). Reconstruction of an acoustically sound-soft obstacle from one incident field and the far-field pattern. *IMA Journal of Applied Mathematics*, 72(1), 96–112. https://doi.org/10.1093/imamat/hxl026

[33] Zinn, A. (1989). On an optimization method for the full and the limited-aperture problem in inverse acoustic scattering for a sound-soft obstacle. *Inverse Problems*, 5(2), 239–253. https://doi.org/10.1088/0266-5611/5/2/009

[34] Tsang, L., Kong, J. A., Ding, K.-H., & Ao, C. O. (2001). *Scattering of Electromagnetic Waves, Numerical Simulations*, 1st ed. Wiley, NY USA. https://doi.org/10.1002/0471224308

[35] Bourlier, C., Pinel, N., & Kubicke, G. (2013). Validation of the Method of Moments for a Single Scatterer. In: Joseph Saillard, ed. *Method of Moments for 2D Scattering Problems*. Wiley - ISTE, NY USA, 31–72. https://doi.org/10.1002/9781118648674.ch2

[36] Abramowitz, M., & Stegun, A. (1964). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 1st ed. Dover, NY USA.

[37] Bourlier, C., Pinel, N., & Kubicke, G. (2013). Integral Equations for a Single Scatterer: Method of Moments and Rough Surfaces. In: Joseph Saillard, ed. *Method of Moments for 2D Scattering Problems*. Wiley - ISTE, NY USA, 1–30. https://doi.org/10.1002/9781118648674.ch1

[38] Ishimaru, A. (2017). Radiation from Apertures and Beam Waves. In: Tariq Samad, ed. *Electromagnetic Wave Propagation Radiation and Scattering*. 2nd ed. IEEE Press, Piscataway, NJ. 169–199. https://doi.org/10.1002/9781119079699.ch6

[39] Altundag, A (2012). *On a Two-Dimensional Inverse Scattering Problem for a Dielectric*. PhD Thesis. Georg-August-Universität zu Göttingen.

[40] Fan, J., & Pan, J. (2004). Inexact Levenberg-Marquardt method for nonlinear equations. *Discrete and Continuous Dynamical Systems Series B*, 4(4), 1223–1232. https://doi.org/10.3934/dcdsb.2004.4.1223

[41] Sefer, A. (2022). Optimization of inverse problems involving surface reconstruction: Least-square application. *Proceedings of the 3rd AT-AP-RASC*. 1–4. https://doi.org/10.23919/AT-AP-RASC54737.2022.9814221

[42] Potthast, R. (1994). Frechet differentiability of boundary integral operators in inverse acoustic scattering. *Inverse Problems*, 10(2), 431–447. https://doi.org/10.1088/0266-5611/10/2/016

[43] Kirsch, A., & Kress, R. (1988). Two methods for solving the inverse acoustic scattering problem. *Inverse Problems*, 4(3), 749–770. https://doi.org/10.1088/0266-5611/4/3/013

Ahmet Sefer *received a B.Sc. in electrical and electronics engineering from Bilkent University, Ankara, Turkey, in 2010 and a Ph.D. from the Graduate School of Istanbul Technical University, Istanbul, Turkey, in 2021. Since September 2021, he has been an Assistant Professor with the Department of Electrical and Electronics Engineering at FMV Isik University, Istanbul, Turkey. Since March 2023, he has been a postdoctoral fellow at King Abdullah University of Science and Technology (KAUST), Saudi Arabia. Dr. Sefer received the Leopold B. Felsen Excellence in Electrodynamics Award from Leopold B. Felsen Fund, in 2020. He is currently a member of the IEEE Antennas and Propagation Society and IEEE Geoscience and Remote Sensing Society. He has been Chair of IEEE Antennas and Propagation Society-Istanbul Chapter. His research interests include electromagnetic theory, direct and inverse scattering problems, integral equations, and numerical techniques.*

https://orcid.org/0000-0001-5168-4367

An International Journal of Optimization and Control: Theories & Applications (http://www.ijocta.org)

This work is licensed under a Creative Commons Attribution 4.0 International License. The authors retain ownership of the copyright for their article, but they allow anyone to download, reuse, reprint, modify, distribute, and/or copy articles in IJOCTA, so long as the original authors and source are credited. To see the complete license contents, please visit http://creativecommons.org/licenses/by/4.0/.

An International Journal of Optimization and Control: Theories & Applications
ISSN:2146-0957 eISSN:2146-5703
Vol.14, No.4, pp.394-403 (2024)
http://doi.org/10.11121/ijocta.1546

RESEARCH ARTICLE

A local differential quadrature method for the generalized nonlinear Schrödinger (GNLS) equation

Meirikim Panmei[a,*], Roshan Thoudam[a]

[a] *Department of Mathematics, Manipur University, Canchipur - 795003, Manipur, India*
meirikimp@gmail.com, roshandmc@gmail.com

ARTICLE INFO

Article History:
Received 16 February 2024
Accepted 23 July 2024
Available Online 16 October 2024

Keywords:
Differential quadrature method
Fourier series expansion
Generalized nonlinear Schrödinger equation
Solitary waves

AMS Classification 2010:
65M99; 42A38; 35Q55

ABSTRACT

A local differential quadrature method based on Fourier series expansion numerically solves the generalized nonlinear Schrödinger equation. For time integration, a Runge-Kutta fourth-order method is used. Matrix stability analysis is used to examine the method's stability. Three test problems involving the motion of a single solitary wave, the interaction of two solitary waves, and a solution that blows up in finite time, respectively, demonstrate the accuracy and efficiency of the provided method. Finally, the numerical results obtained from the presented method are compared with the exact solution and those obtained in earlier works available in the literature.

(cc) BY

1. Introduction

We consider the generalized nonlinear Schrödinger (GNLS) equation, given by

$$i\omega_t + \omega_{xx} + q_1|\omega|^2\omega + q_2|\omega|^4\omega + iq_3(|\omega|^2)_x\omega + iq_4|\omega|^2\omega_x = 0, \quad (1)$$

where $i = \sqrt{-1}$, ω is a complex-valued function of the spatial coordinate x and time t. The subscripts t and x denote differentiation with respect to time, t and space, x and q_1, q_2, q_3, q_4 are real parameters. Eq.(1) describes the modulation of a quasi-monochromatic wave train in a weakly nonlinear dispersive medium [1]. It also describes the behaviour of the Stokes wave near the state of modulation instability, which was independently proposed by Johnson [2], Kakutani and Michihiro [3]. The GNLS Eq.(1) takes some special forms [1] and these forms have found many applications [1, 4]. One of the special forms of Eq.(1) is the well known cubic nonlinear Schrödinger (CNLS) equation:

$$i\omega_t + \omega_{xx} + q_1|\omega|^2\omega = 0, \quad (2)$$

which has found applications in nonlinear optics [5], plasma physics [6] and fluid dynamics [7]. Other special forms of Eq.(1) have applications including propagation of nonlinear Alfvên waves [8] and the self-modulation of the complex amplitude of the solution to the Benjamin-Ono equation [9]. Under the condition that the initial condition $\omega(x, 0)$ vanishes for sufficiently large x, the CNLS Eq.(2) has analytic solution given by [6,10]. There are many papers about the numerical and analytical solutions of the CNLS equation. However, a few papers can be found in the past about the numerical and analytical solutions of the GNLS equation. Exact solution of the GNLS equation was obtained by using Gauss transformation by Pathria and Morris [1]. They also obtained the numerical solution of the GNLS equation using the pseudo-spectral split-step method. Different

*Corresponding Author

split-step pseudo-spectral methods were implemented for the numerical solution of the GNLS equation by Pathria and Morris [4]. Muslu and Erbay [11] used first, second and fourth-order versions of the split-step Fourier method to solve the GNLS equation numerically. The quintic B-spline collocation method was used to solve the GNLS equation by Irk and Dağ [12]. A meshfree method based on RBFs has been used to solve the GNLS equation by Uddin and Haq [13]. Bashan, Ali, et al. [14–16] used various methods based on the differential quadrature method to find the solution of the nonlinear Schödinger Equation. In literature, [17–19], many researchers have developed various types of differential quadrature methods (DQM) using different base functions.

Assuming that ω and all its derivatives tend to zero rapidly as $x \to \pm\infty$, the solutions of the GNLS equation possess the following conservation laws [1, 4]:

$$I_1 = \int_{-\infty}^{\infty} |\omega|^2 dx, \tag{3}$$

$$I_2 = \int_{-\infty}^{\infty} \left[|\omega_x|^2 - \frac{1}{2}(2q_3 + q_4)|\omega|^2 Im(\omega\bar{\omega}_x) - \frac{1}{2}q_1|\omega|^4 + \frac{1}{6}\{q_3(2q_3 + q_4) - 2q_2\}|\omega|^6 \right] dx, \tag{4}$$

and

$$I_3 = \int_{-\infty}^{\infty} \left[2Im(\omega\bar{\omega}_x) - q_3|\omega|^4 \right] dx, \tag{5}$$

which are the conservation of mass, energy and impulse respectively.

The differential quadrature method (DQM) was first introduced by Bellman et al. [20] in 1972 as a simple and versatile numerical technique for solving complex differential equations. This method approximates a function's derivative at a certain point through a weighted linear sum of the functional values at specific collocation points, whereby a key aspect is related to the computation of weighting coefficients. Many authors have used various test functions to formulate various DQ methods, like Legendre polynomials, Lagrange interpolating polynomials, spline functions, radial basis functions (RBF), Chebychev polynomials, etc. [20–26]. Shu proposed a better method for computing weighting coefficients [27]. Shu and Richards [28] used Lagrange interpolating polynomials, which have no limitation on the choice of grid points. This leads to the polynomial-based differential quadrature (PDQ) method. They also obtained a recurrence formula to compute the weighting coefficients for higher-order derivatives. When using Fourier series expansion, we call it the Fourier-based differential quadrature (FDQ) method.

The main advantage of DQM is their high accuracy. In general, DQMs are global in nature [29], which means that they approximate a function and its derivative at a point by using the functional values at all collocation points in the domain. The number of collocation points in the given domain must be large enough to achieve high accuracy approximation. However, it was found that DQM is inefficient when the number of collocation points is larger [30] because of instability. In this regard, Zong and Lam [31] introduced a localized DQ method to keep a balance between stability and accuracy. It has been demonstrated that accuracy and stability can be balanced by approximating the derivative of a function at a position using a weighted sum of functional values at the points in its neighbourhood rather than all collocation points. Therefore, we proposed an efficient numerical approach based on the local differential quadrature method using Fourier series expansion to solve the GNLS Eq.(1).

The paper is organized as follows: Section 2 briefly introduces DQM and the local Fourier-based differential quadrature (L-FDQ) method. The L-FDQ method is implemented in section 3 to solve the GNLS equation. In section 4, we discuss the matrix stability analysis of the proposed method. Section 5 reports the numerical results of the GNLS equation for some test problems. In section 6, we draw a brief conclusion about the presented method.

2. Differential quadrature method (DQM)

DQM is an approximation to the derivative of a function at any grid point using the weighted linear sum of all functional values at certain collocation points in the given domain of definition. We consider an arbitrarily distributed N grid points $x_1 < x_2 < \cdots < x_N$ on the real axis. Then, according to DQ discretization, the n^{th} order derivatives of $U(x,t)$ w.r.t. the spatial coordinate x at a point x_i is given by

$$U^{(n)}(x_i, t) = \sum_{j=1}^{N} w_{i,j}^{(n)} U(x_j, t), \tag{6}$$

where $w_{i,j}^{(n)}$ represents the weighting coefficients, $i, j = 1, \ldots, N$ and $n = 1, \ldots, N-1$.

2.1. Fourier-based differential quadrature (FDQ)

For Fourier-based differential quadrature (FDQ), we consider an arbitrary function defined on the interval $[a, b]$. Two typical sets of base functions are used to compute the weighting coefficients. These sets of base functions are:

$$1, \cos\left(\frac{\pi x}{b-a}\right), \sin\left(\frac{\pi x}{b-a}\right), \ldots, \cos\left(\frac{(N-1)\pi x}{2(b-a)}\right), \sin\left(\frac{(N-1)\pi x}{2(b-a)}\right), \tag{7}$$

and

$$g_j(x) = \frac{G(x)}{\sin\left(\frac{\pi(x-x_j)}{2(b-a)}\right) G^{(1)}(x_j)}, \quad x \in [a,b], \quad j = 1, 2, \ldots, N \tag{8}$$

where

$$G(x) = \prod_{k=1}^{N} \sin\left(\frac{\pi(x-x_k)}{2(b-a)}\right),$$

$$G^{(1)}(x_j) = \prod_{k=1, k\neq j}^{N} \sin\left(\frac{\pi(x_j-x_k)}{2(b-a)}\right), \ x \in [a,b], \quad j = 1, \ldots, N.$$

Using these sets of base functions given in Eq.(7) and Eq.(8), the weighting coefficients for the first and second-order derivatives as evaluated by Shu [29] are as follows:

$$\left.\begin{aligned} w_{i,j}^{(1)} &= \frac{\pi}{2(b-a)} \frac{G^{(1)}(x_i)}{\sin\left(\frac{\pi(x_i-x_j)}{2(b-a)}\right) G^{(1)}(x_j)}, \ i \neq j, \\ w_{i,j}^{(2)} &= w_{i,j}^{(1)} \left(2w_{i,i}^{(1)} - \frac{\pi}{(b-a)} \cot\left(\frac{\pi(x_i-x_j)}{2(b-a)}\right)\right), \\ & i \neq j, \\ w_{i,i}^{(n)} &= -\sum_{j=1, j\neq i}^{N} w_{i,j}^{(n)}, \ n = 1, 2 \end{aligned}\right\} \tag{9}$$

where $i, j = 1, 2, \ldots, N$. We used equally spaced grid points in the space direction to approximate the derivative of the unknown function.

2.2. Local Fourier-based differential quadrature (L-FDQ)

We consider a partition $a = x_1 < x_2 < \cdots < x_i < \cdots < x_N = b$ of the domain [a,b]. Following the method adopted by Shu [27], we consider a location x_i $(i = 1, 2, \ldots, N)$ and for each i, consider a stencil $S_i = \{x_{i-K_1}, x_{i-K_1+1}, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots x_{i+K_2}\}$ containing $K+1$ $(K = K_1 + K_2)$ grid points. For the left boundary point $x_1, K_1 = 0, K_2 = K$, while for the right boundary point x_N, $K_1 = K, K_2 = 0$. Then using the $K+1$ grid points $x_{i-K_1}, x_{i-K_1+1}, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots x_{i+K_2}$ the n^{th} order partial derivative of the function $U(x,t)$ with respect to x at x_i is given by

$$U_x^{(n)}(x_i, t) = \sum_{j=-K_1}^{K_2} w_{i,i+j}^{(n)} U(x_{i+j}, t) \tag{10}$$

where , the L-FDQ weighting coefficients for the first and second-order derivatives in Eq.(9) are given by

$$w_{i,i+j}^{(1)} = \frac{\pi}{2(b-a)} \frac{G^{(1)}(x_i)}{\sin\left(\frac{\pi(x_i-x_{i+j})}{2(b-a)}\right) G^{(1)}(x_{i+j})}, \quad \text{for } j \neq 0, j = -K_1, \ldots, K_2 \tag{11}$$

where

$$G^{(1)}(x_i) = \prod_{\substack{k=K_1 \\ k\neq 0}}^{K_2} \sin\left(\frac{\pi(x_i-x_{i+k})}{2(b-a)}\right) \text{ and}$$

$$G^{(1)}(x_{i+j}) = \prod_{\substack{k=-K_1 \\ k\neq j}}^{K_2} \sin\left(\frac{\pi(x_{i+j}-x_{i+k})}{2(b-a)}\right),$$

and

$$w_{i,i+j}^{(2)} = w_{i,i+j}^{(1)} \left(2w_{i,i}^{(1)} - \frac{\pi}{b-a} \cot\left(\frac{\pi(x_i-x_{i+j})}{2(b-a)}\right)\right), \quad j \neq 0, j = -K_1, \ldots, K_2. \tag{12}$$

For the diagonal coefficients $w_{i,i}^{(n)}$, we have

$$w_{i,i}^{(n)} = -\sum_{\substack{k=-K_1, \\ k\neq 0}}^{K_2} w_{i,i+k}^{(n)} \quad ; \ n = 1, 2. \tag{13}$$

Once the weighting coefficients are computed, then we make the following differentiation matrices, $\mathbf{W}^{(1)} = \left(w_{i,j}^{(1)}\right)_{N\times N}$ and $\mathbf{W}^{(2)} = \left(w_{i,j}^{(2)}\right)_{N\times N}$ to approximate the first and second-order spatial

derivatives of $U(x,t)$ in the domain $[a,b]$. These differentiation matrices are banded.

3. Implementation of L-FDQ

The GNLS Eq.(1) is examined in this part, with the following initial and boundary conditions applied across the interval $[a,b]$:

$$\omega(x,0) = f(x), x \in [a,b] \tag{14}$$
$$\omega(a,t) = \omega(b,t) = 0, t \in (0,T]. \tag{15}$$

Taking $\omega(x,t)= u(x,t) + iv(x,t)$, where $i = \sqrt{-1}$ the GNLS Eq.(1) with the initial and boundary conditions (14) and (15) are transformed into the following coupled initial-boundary value problem (IBVP):

PDEs:

$$\left.\begin{aligned} u_t =& -v_{xx} - [q_1(u^2+v^2) + q_2(u^2+v^2)^2]v \\ & - [2q_3u^2 + q_4(u^2+v^2)]u_x - 2q_3uvv_x \\ v_t =& \quad u_{xx} + [q_1(u^2+v^2) + q_2(u^2+v^2)^2]u \\ & - [2q_3v^2 + q_4(u^2+v^2)]v_x - 2q_3uvu_x \end{aligned}\right\} \tag{16}$$

$$\text{ICs} : u(x,0) = f_u(x), v(x,0) = f_v(x), x \in [a,b] \tag{17}$$

$$\text{BCs} : u(a,t) = u(b,t) = 0, v(a,t) = v(b,t) = 0, \quad t \in (0,T]. \tag{18}$$

To solve the system (16) with ICs (17) and BCs (18) at the collocation points $\{x_1, x_2, \ldots, x_N\}$ with uniform step size $h = x_{i+1} - x_i$, for $i = 1, 2, \ldots, N-1$ we define the following:

$$\mathbf{U}(t) = [u_1(t), u_2(t), \ldots, u_N(t)]^T,$$

$$\mathbf{V}(t) = [v_1(t), v_2(t), \ldots, v_N(t)]^T,$$

where $u_i(t) = u(x_i,t), v_i(t) = v(x_i,t)$ for all $i = 1, 2, \ldots, N$.

Using these definitions and the differentiation matrices $\mathbf{W}^{(1)}$ and $\mathbf{W}^{(2)}$ as defined in section 3, the system of PDEs (16) reduces to the following system of ordinary differential equations (ODEs), which can be written in the following matrix form:

$$\left.\begin{aligned} \mathbf{U}'(t) =& -\mathbf{W}^{(2)} \cdot \mathbf{V}(t) - \left[q_1\left(\mathbf{U}^2(t) + \mathbf{V}^2(t)\right) + q_2\left(\mathbf{U}^2(t) + \mathbf{V}^2(t)\right)^2\right] * \mathbf{V}(t) \\ & - \left[2q_3\mathbf{U}^2(t) + q_4\left(\mathbf{U}^2(t) + \mathbf{V}^2(t)\right)\right] * \left(\mathbf{W}^{(1)} \cdot \mathbf{U}(t)\right) - 2q_3\mathbf{U}(t) * \mathbf{V}(t) * \left(\mathbf{W}^{(1)} \cdot \mathbf{V}(t)\right) \\ \mathbf{V}'(t) =& \ \mathbf{W}^{(2)} \cdot \mathbf{U}(t) + \left[q_1\left(\mathbf{U}^2(t) + \mathbf{V}^2(t)\right) + q_2\left(\mathbf{U}^2(t) + \mathbf{V}^2(t)\right)^2\right] * \mathbf{U}(t) \\ & - \left[2q_3\mathbf{V}^2(t) + q_4\left(\mathbf{U}^2(t) + \mathbf{V}^2(t)\right)\right] * \left(\mathbf{W}^{(1)} \cdot \mathbf{V}(t)\right) - 2q_3\mathbf{U}(t) * \mathbf{V}(t) * \left(\mathbf{W}^{(1)} \cdot \mathbf{U}(t)\right) \end{aligned}\right\} \tag{19}$$

where " · " indicates the multiplication of two matrices and $\mathbf{U}(t) * \mathbf{V}(t)$, $\mathbf{U}^2(t) = \mathbf{U}(t) * \mathbf{U}(t)$ denote the component by component multiplication of two matrices.

Using the corresponding ICs and BCs (17) and (18), we solve the above system of ODEs (19) by the usual RK4 method.

4. Stability analysis

The method's stability is analyzed using the matrix stability analysis as suggested in literature [30,32]. After linearization of the system of ODEs (19), the resulting system can be written in the following matrix form:

$$\mathbf{X}'(t) = \mathbf{A} \cdot \mathbf{X}(t) \tag{20}$$

(since, both the system of ODEs and the BCs are homogeneous) or

$$\begin{bmatrix} \mathbf{U}'(t) \\ \mathbf{V}'(t) \end{bmatrix} = \begin{bmatrix} -\alpha\mathbf{W}^{(1)} & -\beta\mathbf{I} - \gamma\mathbf{W}^{(1)} - \mathbf{W}^{(2)} \\ \beta\mathbf{I} - \gamma\mathbf{W}^{(1)} + \mathbf{W}^{(2)} & -\alpha\mathbf{W}^{(1)} \end{bmatrix} \begin{bmatrix} \mathbf{U}(t) \\ \mathbf{V}(t) \end{bmatrix}$$

where $\mathbf{I}$ is the identity matrix of order $N \times N$, $\mathbf{W}^{(1)}$ and $\mathbf{W}^{(2)}$ are the weighting coefficient matrices for first and second order derivatives respectively, as defined in section-3. Also, we have taken $\alpha = 2q_3\bar{u}^2 + q_4(\bar{u}^2 + \bar{v}^2)$, $\beta = q_1(\bar{u}^2 + \bar{v}^2) + q_2(\bar{u}^2 + \bar{v}^2)^2$ and $\gamma = 2q_3\bar{u}\bar{v}$, where $\bar{u} = ||\mathbf{U}||_\infty$ and $\bar{v} = ||\mathbf{V}||_\infty$.
The stability region for the complex eigenvalues is shown in Figure 1 [33].

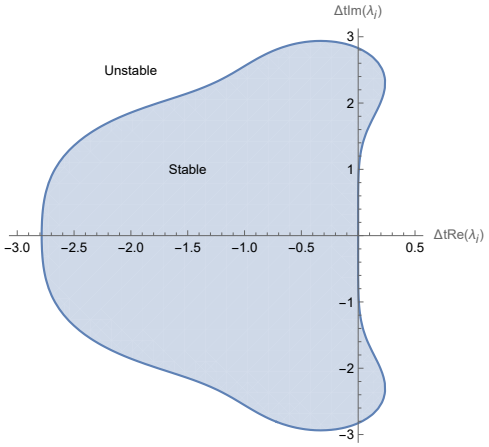


Figure 1. Stability region for complex eigenvalues.

The stability of the numerical integration of the system (20) is related to the stability of the numerical scheme for solving it. If the system of ODEs in (20) is not stable, then the stable numerical scheme for temporal discretization may not generate the converged solution.

The stability of (20) depends on the eigenvalues of the coefficient matrix $\mathbf{A}$ since its exact solution can be found using the eigenvalues. Let λ_i be the eigenvalues of the coefficient matrix $\mathbf{A}$, then the stable solution of $\mathbf{X}(t)$ as $t \to \infty$ requires:

(1) if all the eigenvalues are real $-2.78 < \Delta t\ \lambda_i < 0$
(2) if eigenvalues are imaginary, $-2\sqrt{2} < \Delta t\ \lambda_i < 2\sqrt{2}$
(3) if eigenvalues are complex $\Delta t\ \lambda_i$ should be in the region shown in Figure 1

At the end of section 5, we will calculate the eigenvalues of the coefficient matrix $\mathbf{A}$ and we will see that our scheme is stable with the proper choice of the time step Δt.

5. Numerical experiment

The accuracy and effectiveness of the present method are demonstrated by taking three test problems. The accuracy of the method is measured by using $L_\infty-$ error norm, which is defined as

$$L_\infty = \left\| \, |\omega|_{exct} - |\omega|_{approx} \right\|_\infty = \max_{1\le i\le N} \left| \, |\omega(i)|_{exct} - |\omega(i)|_{approx} \right|$$

5.1. Single solitary wave solution

The exact solitary wave solution of the GNLS Eq.(1) for the parameters $q_1 = 0.5$, $q_2 = -1.75$, $q_3 = -1.0$ and $q_4 = -2.0$ is given by [1,4]:

$$\omega(x,t) = \frac{2e^{i\phi(x,t)}}{\sqrt{4 + 3\sinh^2(x - 2t - x_0)}}, \tag{21}$$

where $\phi(x,t) = 2\tanh^{-1}\left[\frac{1}{2}\tanh(x - 2t - x_0)\right] + x - x_0$.

The modulus of the above solution represents a single solitary wave initially located at x_0, moving to the right with constant speed 2. The exact values of the three conserved quantities I_1, I_2 and I_3 as given in Eq.(3)-Eq.(5), for this problem can be found as:

$$\left.\begin{aligned} I_1 &= 2\log 3 \approx 2.19722, \\ I_2 &= -1.5 + 3.875\log 3 \approx 2.75712, \\ I_3 &= 4 - 9\log 3 \approx -5.88751. \end{aligned}\right\} \tag{22}$$

Table 1. $L_\infty-$ errors and conserved quantities for different stencil sizes K, for a single solitary wave motion, when $h = 0.1$, $\Delta t = 0.001$ over the domain $-20 \le x \le 30$.

K	I_1	I_2	I_3	L_∞
2	2.19724	2.75715	-5.88743	2.40377×10^{-2}
4	2.19712	2.75701	-5.88717	1.26695×10^{-3}
6	2.19712	2.75701	-5.88717	1.07408×10^{-4}
8	2.19712	2.75701	-5.88717	9.39029×10^{-6}
10	2.19712	2.75701	-5.88717	4.23275×10^{-6}
12	2.19712	2.75701	-5.88717	9.42336×10^{-6}

Using the initial condition obtained from (21) and imposing the boundary conditions (18), the GNLS equation is simulated by the proposed method for different stencils over the solution domain $[-20, 30]$. The $L_\infty-$ error and the three conserved quantities for different stencil sizes K are reported in Table 1. From Table 1, we observe that $L_\infty-$ error decreases when K, the size of the stencil increases from $K = 2$ to $K = 10$, however $L_\infty-$ increase when K reached 12 .

Table 2. Comparison of $L_\infty-$ error at $T = 3$, for a single solitary wave motion, with $x_0 = 15$, $N = 513$, $K = 18$ and $-20 \le x \le 60$..

Δt	Present Method	Collocation [12]	First Order [11]	Second Order [11]
0.010	2.75032×10^{-5}	3.0×10^{-4}	3.1×10^{-3}	3.0×10^{-5}
0.005	5.65732×10^{-6}	3.1×10^{-5}	1.6×10^{-3}	2.0×10^{-5}
0.001	3.86334×10^{-7}	2.1×10^{-6}	3.1×10^{-4}	8.0×10^{-7}

Table 3. $L_\infty-$ error norms and Rate of Convergence (ROC) for various numbers of grid points for $K = 6$, $K = 8$ and $K = 10$ with $\Delta t = 0.001$ at $T = 5$.

N	$K = 6$		$K = 8$		$K = 10$	
	Error	*ROC*	*Error*	*ROC*	*Error*	*ROC*
201	1.83495×10^{-2}	–	1.01788×10^{-2}	–	6.13463×10^{-3}	–
301	1.98217×10^{-3}	5.5111	4.52132×10^{-4}	7.71185	1.13554×10^{-4}	9.87958
401	3.91842×10^{-4}	5.65132	5.23908×10^{-5}	7.51345	7.34062×10^{-6}	9.548

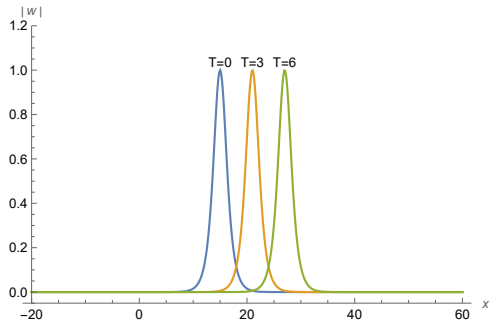


Figure 2. Motion of a single solitary wave at different time levels.

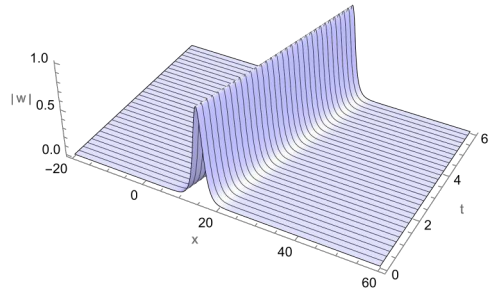


Figure 3. Space time graph of a single solitary wave for GNLS equation ($N = 513, K = 18, \Delta t = 0.001$)

In order to compare our result with [11, 12], we choose $x_0 = 15$, $N = 513$ over the space interval $[-20, 60]$ with time step size, $\Delta t = 0.001$. This comparison is reported in Table 2. Figure 2 and 3 represent the space-time graph of the numerical solution of single solitary wave up to time $T = 6$.

The absolute error distribution at time $T = 6$ for this case is shown in Figure 4.

The numerical rate of convergence (ROC) is calculated by using the formula [33],

$$ROC \approx \frac{\ln\left(\frac{E(N_2)}{E(N_1)}\right)}{\ln\left(\frac{N_1}{N_2}\right)}$$

where $E(N_i)$ is the $L_\infty-$ error norm when using N_i grid points.

The $L_\infty-$ error norm and numerical rate of convergence analysis for various number of grid points are shown in Table 3. From the table it is evident that the rate of convergence (ROC) depends on the value of K.

5.2. Interaction of two solitons

In this test problem, we consider the interaction of two solitons for the GNLS equation, in which the coefficients are taken as $q_1 = 1$, $q_2 = 1$, $q_3 = -2$ and $q_4 = 0$. With these coefficients, we take the initial conditions as given by [1, 4] :

$$\omega(x, 0) = \omega_1(x, 0) + \omega_2(x, 0), \tag{23}$$

where

$$\omega_1(x, 0) = \frac{1}{\sqrt{2}} \operatorname{sech}\left[\frac{1}{2}(x - 15)\right] e^{i\left[\frac{1}{4}(x-15)+\tanh\{\frac{1}{2}(x-15)\}\right]},$$

and

$$\omega_2(x, 0) = \frac{1}{2\sqrt{2}} \operatorname{sech}\left[\frac{1}{4}(x - 35)\right] e^{i\left[-\frac{1}{2}(x-35)+\frac{1}{2}\tanh\{\frac{1}{4}(x-35)\}\right]}.$$

Table 4. Conserved quantities at different time levels, for interaction of two solitons over the space interval $[-50, 100]$ with $N = 501$ and $\Delta t = 0.001$ with stencil size $K = 14$.

T	I_1	I_2	I_3
0	3.00145	0.18974	-1.02223×10^{-4}
5	3.00139	0.18977	-1.02687×10^{-4}
10	3.00141	0.18986	-1.02482×10^{-4}
15	3.00154	0.18968	-1.02161×10^{-4}
20	3.00142	0.18975	-1.02403×10^{-4}

The exact values of the conserved quantities for this problem are $I_1 = 3.0$, $I_2 = \frac{3}{16}$ and $I_3 = 0.0$. The initial condition defined in Eq.(23), represents two solitons, one initially located at $x_1 = 15$, moving to the right with speed $\frac{1}{2}$ and having amplitude $\frac{1}{\sqrt{2}}$ and another initially located at $x_2 = 35$ moving to the left with unit speed and having amplitude $\frac{1}{2\sqrt{2}}$. We have simulated this problem with the present method. These two solitons interact, and after the interaction, they retain their shapes and speeds, which has been shown in Figure 5. In Table 4, the conserved quantities I_1, I_2 and I_3 at different time levels are reported. From the table, we see that the variations of these conserved quantities from the exact values are negligible.

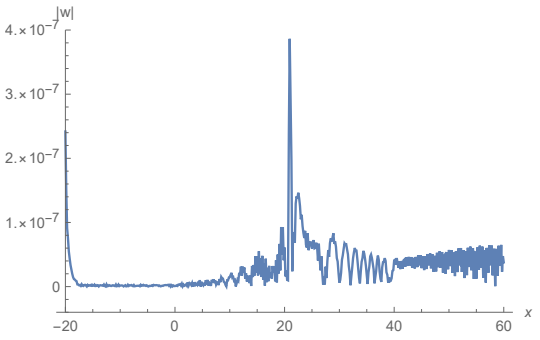


Figure 4. Error distribution of a single solitary wave for GNLS equation at t=6, $(N = 513, K = 18, \Delta t = 0.001)$.

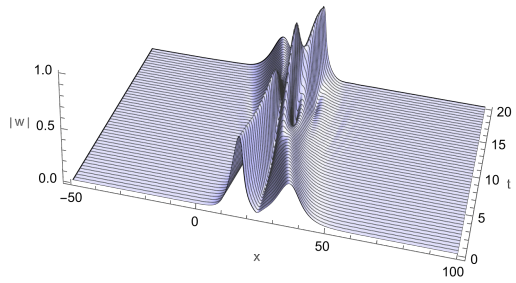


Figure 5. Interaction of two solitons $(N = 501, K = 14, \Delta t = 0.001)$.

5.3. Blow-up

In [1], it has been reported that for specific values of the coefficients and for certain initial conditions, the solutions of the GNLS equation have finite time blow-up. To see this experience, we take $q_1 = -2$, $q_2 = 20$ and $q_3 = q_4 = 0$ and the Gaussian function $\omega(x, 0) = e^{-x^2}$ as the initial condition, the numerical simulation has been conducted by our method. The exact values of the conserved quantities for this problem are found to be $I_1 = \sqrt{\pi/2} \approx 1.5331$, $I_2 = \sqrt{\pi}(9\sqrt{2} + 9 - 20\sqrt{6})/18 \approx -2.68447$ and $I_3 = 0$.

Table 5. Conserved quantities at different time levels for case of finite time blow-up ($N = 151$, $-7.5 \leq x \leq 7.5$, $\Delta t = 10^{-4}$ and $K = 6$).

T	I_1	I_2	I_3
0.00	1.25330	−2.68419	0.0
0.02	1.25329	−2.68414	-4.40186×10^{-17}
0.06	1.25107	−2.34227	-1.13798×10^{-15}
0.07	1.15379	−2.07551	-6.16625×10^{-14}
0.08	1.24722	−2.63257	3.98570×10^{-14}
0.10	1.25129	−2.88368	5.09393×10^{-14}
0.15	1.10587	−2.23570	-6.12399×10^{-12}
0.20	1.24751	−3.39736	1.97811×10^{-11}

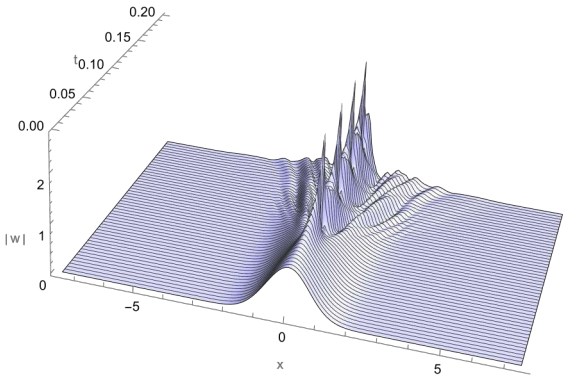


Figure 6. Finite time blow-up, with initial condition $\omega(x, 0) = e^{-x^2} (-7.5 \leq x \leq 7.5, N = 151, \Delta t = 0.0001, K = 12)$.

The conserved quantities for this problem have been reported in Table 5, and from the table, we see that the variation of the conserved quantity is more in I_2. Figure 6 shows the space-time graph of this test problem. According to [1], it has been shown that the exact solution $\omega(x, t)$ for this problem will blow up in finite time, and an upper bound on the blow-up time is $t \approx 1.7$. However, from Figure 6, we observed that the blow-up

is evident at $t = 0.07$, and this result is consistent with the result obtained using the quintic b-spline collocation method in [12]. Also, the graph shows that the blow-up is well occurring at about $t = 0.07, 0.11, 0.15$ and 0.20.

As a part of the stability analysis, we have calculated the eigenvalues of the coefficient matrix, $\mathbf{A}$ as defined in Eq.(20). We take $\bar{u} = \bar{v} = 1$, so that $\alpha = 2q_3 + 2q_4$, $\beta = 2q_1 + 4q_2$ and $\gamma = 2q_3$. The maximum absolute values of the eigenvalues of the coefficient matrix $\mathbf{A}$ for a single solitary wave motion is determined to be 248.273 ($N = 513$). Therefore, for maintaining stability the maximum value of Δt is given by $\Delta t < \frac{2\sqrt{2}}{248.273} = 0.0113924$. However, we take smaller values of Δt in order to get more accurate results. The distribution of eigenvalues for this case is shown in Figure 7. The figure shows that more eigenvalues are distributed near the imaginary axis.

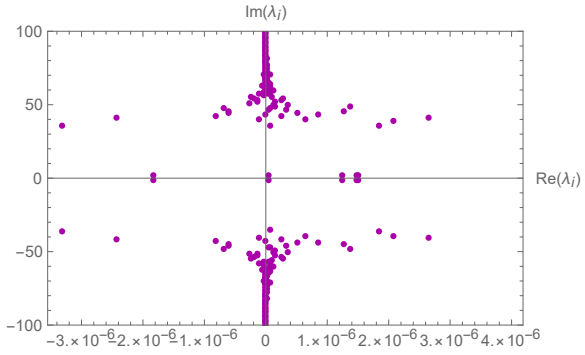


Figure 7
Distribution of eigenvalues for the coefficient matrix, $\mathbf{A}$ ($N = 513, K = 12$).

6. Conclusion

In this study, we have examined the numerical solution of the GNLS equation by means of the L-FDQ method. The GNLS equation is discretized in space using differentiation matrices obtained from the L-FDQ method, and the resulting system of ordinary differential equations in time t is solved by the usual RK4 method. By the present method, the motion of a single solitary wave has been investigated, and the results obtained are compared with the exact solution and some other results obtained in earlier works. It has also been studied how two solitons interact, and it has been found that after the encounter, the solitons maintained their identities. The finite time blow-up problem has also been tackled by the suggested approach, which is consistent with the previous findings. Further, this study found that the finite time blow-up is repeating.

References

[1] Pathria, D., & Morris, J. L. (1989). Exact solutions for a generalized nonlinear Schrödinger equation. *Physica Scripta*, 39(6), 673-679. https://doi.org/10.1088/0031-8949/39/6/001

[2] Johnson, R. S. (1977). On the modulation of water waves in the neighbourhood of kh ≈ 1.363. Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences, 357(1689), 131-141. https://doi.org/10.1098/rspa.1977.0159

[3] Kakutani, T., & Michihiro, K. (1983). Marginal State of Modulational Instability-Note on Benjamin-Feir Instability. *Journal of the Physical Society of Japan*, 52(12), 4129-4137. https://doi.org/10.1143/JPSJ.52.4129

[4] Pathria, D., & Morris, J. L. (1990). Pseudo-spectral solution of nonlinear Schrödinger equations. *Journal of Computational Physics*, 87(1), 108-125. https://doi.org/10.1016/0021-9991(90)90228-S

[5] Strauss, W. A. (1978). The nonlinear Schrödinger equation. In North-Holland Mathematics Studies (Vol. 30, pp. 452-465). North-Holland. https://doi.org/10.1016/S0304-0208(08)70877-6

[6] Shabat, A., & Zakharov, V. (1972). Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media. *Soviet Physics JETP*, 34(1), 62.

[7] Hasimoto, H., & Ono, H. (1972). Nonlinear modulation of gravity waves. *Journal of the Physical Society of Japan*, 33(3), 805-811. https://doi.org/10.1143/JPSJ.33.805

[8] Kaup, D. J., & Newell, A. C. (1978). An exact solution for a derivative nonlinear Schrödinger equation. *Journal of Mathematical Physics*, 19(4), 798-801. https://doi.org/10.1063/1.523737

[9] Tanaka, M. (1982). Nonlinear self-modulation problem of the Benjamin-Ono equation. *Journal of the Physical Society of Japan*, 51(8), 2686-2692. https://doi.org/10.1143/JPSJ.51.2686

[10] Karpman, V. I., & Krushkal, E. M. (1969). Modulated waves in nonlinear dispersive media. *Soviet Journal of Experimental and Theoretical Physics*, 28, 277.

[11] Muslu, G. M., & Erbay, H. A. (2005). Higher-order split-step Fourier schemes for

the generalized nonlinear Schrödinger equation.*Mathematics and Computers in Simulation*, 67(6), 581-595. https://doi.org/10.1016/j.matcom.2004.08.002

[12] Irk, D., & Dağ, İ. (2011). Quintic B-spline collocation method for the generalized nonlinear Schrödinger equation. *Journal of the Franklin Institute*, 348(2), 378-392. https://doi.org/10.1016/j.jfranklin.2010.12.004

[13] Uddin, M., & Haq, S. (2013). On the numerical solution of generalized nonlinear Schrodinger equation using radial basis functions. *Miskolc Mathematical Notes*, 14(3), 1067-1084. https://doi.org/10.18514/MMN.2013.486

[14] Başhan, A. (2019). A mixed methods approach to Schrödinger equation: Finite difference method and quartic B-spline based differential quadrature method. *An International Journal of Optimization and Control: Theories & Applications (IJOCTA)*, 9(2), 223-235. https://doi.org/10.11121/ijocta.01.2019.00709

[15] Bashan, A., Yagmurlu, N. M., Ucar, Y., & Esen, A. (2017). An effective approach to numerical soliton solutions for the Schrödinger equation via modified cubic B-spline differential quadrature method. *Chaos, Solitons & Fractals*, 100, 45-56. https://doi.org/10.1016/j.chaos.2017.04.038

[16] Başhan, A., Uçar, Y., Murat Yağmurlu, N., & Esen, A. (2018). A new perspective for quintic B-spline based Crank-Nicolson-differential quadrature method algorithm for numerical solutions of the nonlinear Schrödinger equation. *The European Physical Journal Plus*, 133(1), 12. https://doi.org/10.1140/epjp/i2018-11843-1

[17] Uçar, Y., Yağmurlu, M., & Başhan, A. (2019). Numerical solutions and stability analysis of modified Burgers equation via modified cubic B-spline differential quadrature methods. *Sigma Journal of Engineering and Natural Sciences*, 37(1), 129-142.

[18] Başhan, A., Yağmurlu, N. M., Uçar, Y., & Esen, A. (2021). Finite difference method combined with differential quadrature method for numerical computation of the modified equal width wave equation. *Numerical Methods for Partial Differential Equations*, 37(1), 690-706. https://doi.org/10.1002/num.22547

[19] Başhan, A., Uçar, Y., Yağmurlu, N. M., & Esen, A. (2016, October). Numerical solution of the complex modified Korteweg-de Vries equation by DQM. In Journal of Physics: Conference Series (Vol. 766, No. 1, p. 012028). IOP Publishing. https://doi.org/10.1088/1742-6596/766/1/012028

[20] Bellman, R., Kashef, B. G., & Casti, J. (1972). Differential quadrature: a technique for the rapid solution of nonlinear partial differential equations. *Journal of Computational Physics*, 10(1), 40-52. https://doi.org/10.1016/0021-9991(72)90089-7

[21] Bellman, R., Kashef, B., Lee, E. S., & Vasudevan, R. (1975). Differential quadrature and splines. *Computers & Mathematics with Applications*, 1(3-4), 371-376. https://doi.org/10.1016/0898-1221(75)90038-3

[22] Quan, J. R., & Chang, C. T. (1989). New insights in solving distributed system equations by the quadrature method-I. Analysis. *Computers & Chemical Engineering*, 13(7), 779-788. https://doi.org/10.1016/0098-1354(89)85051-3

[23] Quan, J. R., & Chang, C. T. (1989). New insights in solving distributed system equations by the quadrature method-II. Numerical experiments. *Computers & Chemical Engineering*, 13(9), 1017-1024. https://doi.org/10.1016/0098-1354(89)87043-7

[24] Shu, C., & Xue, H. (1997). Explicit computation of weighting coefficients in the harmonic differential quadrature. *Journal of Sound and Vibration*, 204(3), 549-555. https://doi.org/10.1006/jsvi.1996.0894

[25] Shu, C., & Wu, Y. L. (2007). Integrated radial basis functions-based differential quadrature method and its performance. *International Journal for Numerical Methods in Fluids*, 53(6), 969-984. https://doi.org/10.1002/fld.1315

[26] Yigit, G., & Bayram, M. (2017). Chebyshev differential quadrature for numerical solutions of higher order singular perturbation problems. arXiv preprint arXiv:1705.09484.

[27] Shu, C. (1991). Generalized differential-integral quadrature and application to the simulation of incompressible viscous flows including parallel computation. University of Glasgow (United Kingdom).

[28] Shu, C., & Richards, B. E. (1992). Application of generalized differential quadrature to solve two-dimensional incompressible Navier-Stokes equations. *International Journal for Numerical Methods in Fluids*, 15(7), 791-798.

https://doi.org/10.1002/fld.1650150704

[29] Shu, C. (2000). Differential quadrature and its application in engineering. Springer Science & Business Media.

[30] Civan, F., & Sliepcevich, C. M. (1984). Differential quadrature for multi-dimensional problems. *Journal of Mathematical Analysis and Applications*, 101(2), 423-443. https://doi.org/10.1016/0022-247X(84)90111-2

[31] Zong, Z., & Lam, K. Y. (2002). A localized differential quadrature (LDQ) method and its application to the 2D wave equation. *Computational Mechanics*, 29, 382-391. https://doi.org/10.1007/s00466-002-0349-4

[32] Tomasiello, S. (2011). Numerical stability of DQ solutions of wave problems. *Numerical Algorithms*, 57, 289-312. https://doi.org/10.1007/s11075-010-9429-2

[33] Korkmaz, A., Aksoy, A. M., & Dağ, I. (2011). Quartic B-spline differential quadrature method. *International Journal of Nonlinear Science*, 11(4), 403-411.

Meirikim Panmei *received his undergraduate degree in B.Sc. mathematics from Dhanamanjuri College, Imphal, Manipur. He has completed his M.Sc. in applied mathematics at Manipur University. He is currently a research scholar in the Department of Mathematics at Manipur University. His research interests include the differential quadrature methods and the finite difference methods.*

https://orcid.org/0009-0000-6054-1990

Roshan Thoudam *is an Assistant Professor in the Department of Mathematics at Manipur University. He completed his M.Sc. at Manipur University in 1994. He obtained his Ph.D. in numerical solutions of solitary waves based on the B-spline finite element method from the Department of Mathematics at Manipur University in 2012. His research includes differential equations, CFD, numerical methods, and scientific computing.*

https://orcid.org/0000-0001-5273-416X

This work is licensed under a Creative Commons Attribution 4.0 International License. The authors retain ownership of the copyright for their article, but they allow anyone to download, reuse, reprint, modify, distribute, and/or copy articles in IJOCTA, so long as the original authors and source are credited. To see the complete license contents, please visit http://creativecommons.org/licenses/by/4.0/.

An International Journal of Optimization and Control: Theories & Applications
ISSN: 2146-0957 eISSN: 2146-5703
Vol.14, No.4, pp.404-418 (2024)
https://doi.org/10.11121/ijocta.1594

RESEARCH ARTICLE

Modeling the dependency structure between quality characteristics in multi-stage manufacturing processes with copula functions

Pelin Toktaş [a *], Ömer Lütfi Gebizlioğlu[b]

[a] *Department of Industrial Engineering, Başkent University, Turkey*
[b] *Department of International Trade and Finance, Kadir Has University, Turkey*
ptoktas@baskent.edu.tr, omer.gebizlioglu@khas.edu.tr

ARTICLE INFO

Article history:
Received: 17 April 2024
Accepted: 22 October 2024
Available Online: 24 October 2024

Keywords:
Multi-stage manufacturing process
Statistical process control
State-space model
Kalman filter
Copula modeling

AMS Classification 2010:
93E11, 93C05, 62H05

ABSTRACT

This study is about multi-stage manufacturing processes and their control by statistical process control modeling. There are two kinds of dependence structures in a multi-stage manufacturing process: one is the dependence between the stages of the process, and the other is the dependence between the concerned quality characteristics. This study employs state-space models to demonstrate the dependency structure between the process stages and uses the Kalman filter method to estimate the states of the processes. In this setup, copula modeling is proposed to determine the dependence structure between the quality characteristics of interest. A simulation study is conducted to assess the model's accuracy. As a result, it was found that the model gives highly accurate predictions according to the mean absolute percentage error (MAPE) criteria (<10%).

(cc) BY

1. Introduction

Today, production and service processes generally consist of many serial or parallel stages in which products are completed by passing from one to the other. In a multi-stage manufacturing process, it is not clear from which stage and which variables characterizing the process arise the variability in quality characteristics. The key to reducing quality variability in a product is understanding how, much of this variability occurs at each stage of the process and how much is transmitted to other stages.

The most important problem in the multi-stage manufacturing process is how to define the process in the context of interactions within and between stages and time dynamics. In past research, multistage processes have been described with statistical models such as the linear regression model. Conversely, for more effective monitoring and control of the process, engineering knowledge must also be combined with statistics in modeling and analysis of the multi-stage process. In this context; Many articles can be found in the sources that describe the multi-stage manufacturing process in a linear state-space model structure based on production engineering knowledge. A complex system, such as a multistage manufacturing process, may have many inputs and outputs. These inputs and outputs can be complexly interrelated. The hierarchical structure of the data obtained can be explained by multi-level dynamic models. An example of this is a two-level linear state-space model.

In this study, in addition to a dynamic modeling approach such as the state-space model of the dependency between stages in multi-stage manufacturing processes, it is proposed to use copula modeling to reveal the internal dependencies of the quality characteristics of interest at each stage. In order to present the practical implications of the proposed model, the process was simulated and the applicability of the model was discussed.

The following sections of the study are organized as follows: In the second section, studies on statistical process control (SPC) methods used for modeling multi-stage manufacturing processes and monitoring these processes will be discussed. In the third section, modeling of multi-stage manufacturing processes with state-space models will be explained. Additionally, this chapter will include the proposal of the Kalman filter method for the statistical estimation of the state variables of the process equations put forward by state-space models. In the fourth chapter, the statistical dependence of quality characteristics and the explanation of dependence with copula functions will be highlighted, and multi-stage manufacturing process modeling under dependence will be presented. Multi-stage manufacturing processes under dependency The example and process simulation of SPC approaches

*Corresponding author

will be presented in the fifth chapter. In the sixth chapter, the results of the study and some determinations about future studies as extensions of these will be stated.

2. Literature review

In order to identify out-of-control situations in multi-stage manufacturing processes, SPC methods have been applied to the quality measurements of the product in the final stage of the process. Generally, Shewart, CUSUM, EWMA control charts for univariate quality measurements of the final product; It has been suggested to use Hotelling's T^2 control chart for multivariate quality measurements [1]. Since these control charts were applied to a single stage of the process, they were insufficient to determine the stage that was the source of variability. In another study, quality measurements obtained from each stage of the process were evaluated separately [2]. In this study, where simultaneous confidence intervals were established for the average of each of the quality variables, it was examined whether the quality measurements of interest were within the confidence intervals in terms of the defined quality levels, and it was stated that the explanatory power of the method decreased as the size of the problem increased.

Statistical process control tools used to monitor multi-stage manufacturing processes have a wide place in the literature. These tools can be examined under three headings: multivariate control charts, control charts based on regression modeling, and methods based on engineering-based models.

In many production processes, it may be necessary to simultaneously monitor and control one or more interrelated quality characteristics. Independent examination of quality characteristics causes loss of information to be obtained from the process. The concept of multivariate quality control originated in Hotelling's work in 1947 [3]. In this study, he applied his proposed method to bombardment viewfinder data used in World War II. The most well-known multivariate process monitoring and control method used to monitor the mean vector of the process is Hotelling's T^2 control chart, which is similar to the univariate Shewhart's $\bar{x}$ chart. Applied to multi-stage manufacturing processes, Hotelling's T^2 chart indicates when the entire process is out of control, but does not indicate which stage is out of control. Alternatively, quality metrics at each stage can be tracked with T^2 cards. In this case, the effect of the quality output of the previous stage on the quality measurements at a certain stage will be ignored. As a result, it is difficult to interpret an out-of-control situation in a multi-stage manufacturing process with a T^2 chart [1]. Following this pioneering work by Hotelling, control methods for many related variables have been proposed [4]. Nowadays, the issue of multivariate quality control (or process monitoring) has maintained its importance as many quality characteristics of products manufactured with automatic inspection methods can be measured at the same time. For example; Chemical and semiconductor manufacturers try to keep the process under control by constantly updating their databases for hundreds of important variables in their manufacturing processes.

It was thought that quality measurements in multi-stage manufacturing processes are affected by the output of the previous stage and the regression analysis technique was introduced [5]. This method is based on establishing univariate control cards for the residuals obtained from the multivariate regression line established on other variables for each quality variable [6]. Regression models can give misleading results when quality measurements from different stages are strongly correlated with each other. This problem in regression analysis can be partially reduced by the cause-selection method and is effective in identifying out-of-control stages [7]. A compilation of cause-selection method studies was compiled by Wade and Woodall [8]. Nowadays, the use and applications of cause-selection schemes for multi-stage processes are also found in Shu and Tsung's article [9].

The hierarchical structure of data obtained from the multistage manufacturing process suggests a two-level model: At the first level, quality measurements are fitted to the system input and quality information. At the second level, the change in quality measurements is modeled as a function of measurements obtained from earlier stages of the process. An example of this situation is the state-space model.

Quality measurements for the kth stage of a production process consisting of N stages are formulated as a linear state-space model as in Eq. (1) and Eq. (2) [10].

$$x_k = A_{k-1}x_{k-1} + v_k \tag{1}$$

$$y_k = C_k x_k + w_k \quad \{k\} \subset \{1,2,\dots,N\} \tag{2}$$

In Eq. (1), x_k shows unobservable product quality information such as dimensional deviations of products at the kth stage. v_k indicates the cause of variability and unmodelable errors (process noise). $A_{k-1}x_{k-1}$ shows the transformation of quality information from the $(k-1)$th stage to the kth stage. In Eq. (2), w_k is the measurement error of the product, and C_k is the matrix used to relate x_k with quality measurements (y_k).

A_{k-1} and C_k are constant matrices obtained from engineering knowledge, laws of physics and process/product design information and known at the kth stage of the process. For univariate cases, $v_k \sim N(0, \sigma_{v_k}^2)$ and $w_k \sim N(0, \sigma_{w_k}^2)$ with its variance depending on the stage index k and the initial state $x_0 \sim N(a_0, \tau^2)$. Various methods have been researched to monitor whether the process is out of control, and fixture errors, machine errors and thermal errors in the process are seen as process out of control or process errors.

A multistage manufacturing process can have many inputs and outputs. These inputs and outputs can be

intricately related to each other. There are many articles explaining multi-stage manufacturing processes with state-space models based on process management expertise. Lawless et al. [11] and Agrawal et al. [12] revealed quality variability in multi-stage manufacturing processes with AR(1) type models in the form of state-space models. Part assembly process [12] and sheet metal assembly [13] are examples of modeling proposals in the form of state-space model. Detailed descriptions of state-space models can be found in [10] and [14]. There are many studies in the literature on error detection, error prevention and corrective methods in multi-stage manufacturing operations. Tsung et al.'s study compiled past studies on multi-stage manufacturing and service operations and provided ideas for future research [15].

Today, modeling for monitoring and control of multi-stage manufacturing processes, which have become more complex with developing technology, is a complex issue that still maintains its importance. State-space models are a modeling method that has a wide place in the literature and includes the physics rules surrounding engineering and production structures suitable for the structure of multi-stage manufacturing processes. In a dynamic system represented by a state-space model, the state of the system can be predicted from the input and output information together with the previous information of the model. Estimation of the state of the system from a series of noisy measurements obtained from a dynamic system can be made with the Kalman filter.

In this study, it is suggested to model the dependency structure between quality characteristics with copula and combine it with Kalman filter. Some studies in which copulas, Kalman filter and/or state space models are used together are given in Table 1.

Table 1. Some selected studies on copulas, Kalman filter and state space models.

Authors	Methods/Models	Examples/Application Area
Lindsey [16].	Kalman filter and copulas	The application to autoimmunity in multiple sclerosis data
Junker, Szimayer and Wagner [17	Kalman filter based on copula functions	Nonlinear cross-sectional dependence in the term structure of US-Treasury yields and points out risk management implications
Hafner and Manner [18]	A multivariate stochastic volatility models with Gaussian copula	The application to two bivariate stock index series
Goto [19]	State space model to describe the target system's behaviour	A simulation study conducted to show the effectiveness of the developed controller
Creal and Tsay [20]	Gaussian, Student's t, grouped Student's t, and generalized hyperbolic copulas with time-varying correlations matrices	Modeling an unbalanced, 200-dimensional panel consisting of credit default swaps and equities for 100 US corporations
Alpay and Hayat [21]	Copula and Data Envolopment Analysis (DEA)	The application to simulated and real hospital data
Zhang and Choudhry [22]	Four generalized autoregressive conditional heteroscedasticity (GARCH) models and the Kalman filter method	Empirically forecasting the daily betas of a few European banks during the pre-global financial crisis period and the crisis period
Fernández, García and González-López [23]	Copula and the multivariate Markov chain	Spike prediction in neuronal data
Smith and Maneesoonthorn [24]	Construction of copulas from the inversion of nonlinear state space models	Forecasting of quarterly U.S. broad inflation and electricity inflation
Wang, Meng, Liui Fu and Cau [25]	The Unscented Kalman Filter (UKF), copula and the worst case analysis	A two-stage dynamic attack strategy using global network information
Xu, Liang, Li and Wang [26]	Characterization of the dependence among all components by a copula function	Investigation of the optimal condition-based maintenance policy under periodic inspection for a *K*-out-of-*N*: G system
Kreuzer, Dalla Valle and Czado [27]	Non-linear non-Gaussian state space model	Estimation of airborne pollutant concentrations
Ly, Sriboonchitta, Tang and Wong [28]	A hybrid of ARMA-GARCH, static and dynamic copulas and dynamic state space models	Investigation of dependence and integration among the European electricity markets
Wang, Xu, Trajcevski, Zhang, Zhong and Zhou [29]	A non-linear neural state space model based on copula-augmented mechanism	Electricity forecasting
Kreuzer, Dalla Valle and Czado [30]	Multivariate nonlinear non-Gaussian state space models	The application to atmospheric pollutant measurement data

The rest of the study is organized as follows. In the third section of the study, state-space models will be discussed. In the fourth section, the copulas proposed to model the dependency structure between quality characteristics will be explained in detail. Application of the proposed approach by a simulation study is given in the fifth section. The last section includes the conclusions of the study, and the future studies.

3. Multi-stage manufacturing processes and state-space models

Dynamic systems, such as multistage manufacturing processes, can be more generally represented in the form of state-space models by the equations shown in Eq. (3) and Eq. (4).

$$x_k = A_{k-1}x_{k-1} + B_k u_k + D_k \varepsilon_k \quad (3)$$

$$y_k = C_k x_k + H_k \eta_k \quad (4)$$

Similar to Eq. (1) and Eq. (2), x_k is the state and y_k is the measurement or observation vectors ($k = 1, \dots, N$). The vectors ε_k and η_k express the noise in the state and the observations, and the vector u_k represents the effects of managerial inputs at the kth stage in the process in Eq. (3) and Eq. (4).

Estimation of the state vector $x_k, k = 1, \dots, N$ in state-space models and other related analyzes can be done within the framework of three main approaches [31]. These are Bayesian, Fisher and unknown-bounded approaches. In the Bayesian approach, the error terms ε_k and η_k in the equations are stochastic, and the initial state vector x_0 is a random variable. In the Fisher approach, the measurement equation term η_k has a stochastic feature, ε_k can be stochastic or completely unknown, and x_0 can be random. Within the framework of the unknown – bounded approach, ε_k, η_k and x_0 are unknown but are limited from above to the values of the ellipsoids expressing the variance-covariance quantities [32].

When A_{k-1}, B_k and C_k matrices are accepted as known matrices in state-space models, the model estimation problem is solved by using the observation values $y_1, y_2, \dots, y_{k_1}$ obtained up to time k_1 and estimating x_{k_2} at time k_2. When $k_1 = k_2$, the estimation problem becomes a filtering process, for which Kalman filter (KF) or weighted least squares (WLS) methods can be used. Estimation equations that can be applied within the framework of the Bayesian model approach are known as Kalman filters in the literature [33].

The Bayesian model approach is the most widely used state-space modeling approach and can offer flexible perspectives on the dependence and independence of the vectors ε_k, η_k and x_0 within and among themselves in the time dimension. In this sense, the issues of determining the prior and posterior probability distributions for the random variables in the state-space model and the expected value and covariance functions are needed in estimation process.

Control effects that can be applied in a dynamic stochastic process are represented by the sequence $\{u_k\}$ in state-space models. While control effects, state vectors should be a function of x_k's, in the absence of a complete and direct observation of the situations, measurement or observation values must be considered as a function of y_k's and determined by the opinion of system experts; $u_k = \omega_0\,(y_0, y_1, \dots, y_k)$. In the literature, it is also recommended to impose a constraint such as $|u_k| \leq 1$ for u_k's [34].

The Kalman filter and its calculation equations are explained in detail in the next section. In the weighted least squares method, the aim is to estimate the state vector with the deviation of x_k, which minimizes the quantity in Eq. (5), where the covariance matrix of the variable η_k is $R_k > 0$.

$$J(x_k) = (y_k - C_k x_k)' R_k^{-1} (y_k - C_k x_k) \quad (5)$$

In Eq. (5), the R_k^{-1} matrix is a positive definite matrix and must be determined in the context of the inputs, states and outputs of the dynamic system of interest. For x_k estimation that gives the smallest value of $J(x_k)$. The solution in Eq. (6) is found for the x_k estimation that gives the smallest value of $J(x_k)$.

$$\hat{x}_k = (C_k R_k^{-1} C_k)^{-1} C_k R_k^{-1} y_k \quad (6)$$

Estimation of x_k in the context of the weighted least squares method for the state-space model in Eq. (5) and Eq. (6); P_0 is the covariance matrix for the initial state vector x_0, and Q_k is the covariance matrix for the vector ε_k, and Equation 7 is obtained by reaching its minimum value under the $x_k = A_{k-1}x_{k-1} + B_k u_k + D_k e_k$ constraint.

$$J(x_k, \varepsilon_0, \varepsilon_1, \dots, \varepsilon_{k-1}) = \sum_{k=1}^{N} (y_k - C_k x_k)' R_k^{-1} (y_k - C_k x_k) + \sum_{k=0}^{n-1} \varepsilon_k Q_k^{-1} \varepsilon_k + x_0' P_0^{-1} x_0 \quad (7)$$

In Eq. (7), R_k, Q_k and P_0 matrices are positive definite and determined based on expert knowledge about the dynamic system of interest [10].

3.1. State estimation with filtering: Kalman filter

The state of the system may not be directly measurable. In a dynamic system represented by a state-space model, the state of the system can be estimated by using the model's information obtained at previous times and its output information. Kalman filter, which was first introduced by Kalman in 1960, is an effective analysis algorithm that estimates the state of the system from a series of measurements obtained from a dynamic system that may contain error (noise), and updates the estimate as observations are made [35]. The Kalman filter combines measurement data, a priori information about the system, and indirectly measuring state values to make the desired predictions by minimizing the error

statistically. Therefore, it gives better results than most other filters for statistical estimation purposes. Within the framework of the Bayesian approach, by conditioning the real data information provided by measuring devices, the spread of conditional probability densities for the features to be estimated can be filtered. Kalman filter helps the purpose of predictive analysis of a system that can be expressed with a linear model, where measurement errors are white noise and normally distributed, by providing conditional probability distribution [36].

For the dynamic and stochastic multi-stage production system represented by the state-space model equations Eq. (3) and Eq. (4), a series of prediction and filtering processes are required in line with the estimation of the state vector x_k at stage k. The difference equations needed for this purpose within the scope of the Bayesian approach are known as Kalman or Kalman-Bucy equations. There are various approaches and generalizations in determining the equations in question, and it is possible to consider equivalent criteria that form the basis for all of them. Minimizing the expected value of prediction error squares is one of these criteria [37].

3.1.1. Minimization of expected value of squared error criteria method

In order to make state estimation with the Kalman filter, explanations about the variables and coefficients in the state-space model equations Eq. (3) and Eq. (4) are given below:

$x_k \in R^n$: System state vector.

$y_k \in R^m$: System observation vector.

A_k: $n \times n$ dimensional system transition matrix.

B_k: $n \times n$ dimensional system input matrix.

C_k: $m \times n$ dimensional observation transition matrix.

u_k: Vector expressing the effect of managerial inputs at time (stage) k.

D_k: $n \times n$ dimensional system noise matrix.

H_k: $m \times n$ dimensional observation noise matrix.

It is assumed that the matrices A_k, B_k, C_k, D_k and H_k are known at all times $k = 0,1,2,\ldots$. The zero-mean white noise processes $\varepsilon_k \in R^n$ and $\eta_k \in R^m$ are assumed to satisfy the following assumptions for each k, j value in Eqs. (8)-(17).

$$E[\varepsilon_k] = 0 \tag{8}$$

$$E[\eta_k] = 0 \tag{9}$$

$$E[\varepsilon_k \varepsilon_j'] = Q_k \delta_{kj} \tag{10}$$

$$E[\eta_k \eta_j'] = R_k \delta_{kj} \tag{11}$$

$$\delta_{kj} = \begin{cases} 1, & k = j \\ 0, & k \neq j \end{cases} \tag{12}$$

$$E[\varepsilon_k \eta_j'] = 0 \tag{13}$$

$$E[x_0] = \bar{x}_0 \tag{14}$$

$$E[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)'] = P_0 \tag{15}$$

$$E[x_0 \varepsilon_j'] = 0 \tag{16}$$

$$E[x_0 \eta_j'] = 0 \tag{17}$$

Table 2. Discrete time Kalman filter equations based on minimization of mean squared errors [38].

System dynamic model:
$x_k = A_{k-1}x_{k-1} + B_k u_k + \varepsilon_k\ ,\ \varepsilon_k \sim N(0, Q_k)$
Measurement (observation) model:
$y_k = C_k x_k + \eta_k,\ \ \eta_k \sim N(0, R_k)$
Starting conditions:
$x_0 \sim N(\bar{x}_0, P_0),\ \ \hat{x}_{0\|-1} = \bar{x}_0,\ \ \ P_{0\|-1} = P_0$
Independence conditions:
$E[\varepsilon_k \eta_j'] = 0,\ \ E[x_0 \varepsilon_j'] = 0,$
$E[x_0 \eta_j'] = 0,\ \ \forall k, j$
Estimation of prediction stage:
State estimation:
$\hat{x}_{k\|k-1} = A_{k-1}\hat{x}_{k-1\|k-1} + B_k u_k$
Measurement condition:
$\hat{y}_{k\|k-1} = C_k \hat{x}_{k\|k-1}$
$= C_k[A_{k-1}\hat{x}_{k-1\|k-1} + B_k u_k]$
Errors of predition stage:
State error:
$x_k^* = x_k - \hat{x}_{k\|k-1}$
Measurement error:
$w_k = y_k - \hat{y}_{k\|k-1} = C_k(x_k - \hat{x}_{k\|k-1}) + H_k \eta_k$
Covariance matrix of prediction stage:
$P_k(w) = E[w_k w_k']$
$= C_k P_{k\|k-1} C_k' + H_k R_k H_k'$
Update of error covariance for prediction stage:
For state:
$P_{k\|k-1} = A_{k-1} P_{k-1\|k-1} A_{k-1}' + D_k Q_k D_k'$
For measurement:
$P_k(y) = C_k[A_{k-1}P_{k-1}A_{k-1}' + D_k Q_k D_k']C_k' + H_k R_k H_k'$
Observational update of the state estimate:
$\hat{x}_{k\|k} = \hat{x}_{k\|k-1} + B_k u_k + P_{k\|k-1} C_k P_k^{-1}(w)[y_k - \hat{y}_{k\|k-1}]$
$= \hat{x}_{k\|k-1} + K_k\{y_k - C_k \hat{x}_{k\|k-1}\}$
Update of error covariance for filtering stage:
$P_{k\|k} = [I - K_k C_k] P_{k\|k-1}$
Kalman gain matrix:
$K_k = P_{k\|k-1} C_k' P_k^{-1}(w)$

In addition to all given assumptions, it is assumed that the matrices Q_k and R_k are known. It is aimed to obtain $\hat{x}_{k|m}$ by using observations $\{y_1, y_2, \ldots, y_m\}$ for the best

estimation of the x_k vector. In this direction, It is possible to use the covariance matrix $(P_{k|m})$ of the estimation error $(x_k - \hat{x}_{k|m})$. When $k = m$, the estimation is called as filtering. Considering that the observations are not error-free, the assumption of $R_k > 0$ will be a realistic and necessary assumption. Let the vector $Y_k = [y_1, \dots, y_k]'$ represent the observations obtained until time (stage) k. If $\vartheta_{k|m}$ is the estimation error of $\hat{x}_{k|m}$ using Y_k, the covariance matrix of this error is expressed as $P_{k|m} = E[\vartheta_{k|m}\vartheta'_{k|m}]$, with $E[\hat{x}_{k|m}] = E(x_k)$. The estimation of vector x_k is done in two stages with various calculation steps. In Table 2, discrete time Kalman filter equations are summarized according to the method of minimizing the expected value of error squares by showing the filter system relationship.

On the other hand, η_k in the system equation Eq. (4) may become unobtainable. The solution to this problem requires adding additional state equations to the system equation. Bryson and Johanson proposed the first general solution to the problem in question [39]. To solve the problem, Brown and Hwang suggest removing exactly known state variables from the system equations and estimating the remaining ones by filtering [34]. This recommendation requires the separation of system state variables from other exactly known system variables by linear transformation. Simon summarized adequate explanations and methods of Kalman filter application approaches by considering the dependence as linear dependence and correlation for the cases where the random vectors η_k and ε_k are dependent within and between themselves [40].

4. Modeling multi-stage manufacturing processes under the dependency between quality characteristics

In this section, a method is proposed by including copula functions in the approach of modeling and evaluating multi-stage manufacturing processes with state-space models under dependency. It has been suggested to use copula modeling to reveal the internal dependencies of the quality features within the state vector at each stage. With copula models, the stochastic relationship between quality characteristics can be determined by revealing the dependency structure without the need for common distributions of quality characteristics. In this context; Statistical properties such as marginal distributions, covariance, conditional probability distributions (and therefore regression function determination) of quality features that are random variables can be expressed.

4.1. Copula functions

Copula functions are statistical tools used to model dependency. Copulas are functions that combine multivariate distributions with their univariate marginal distributions. Let F be the m-dimensional cumulative distribution function and $F_1, F_2, \dots, F_m$ be the cumulative distribution functions of one-dimensional marginals. In this case, the m-dimensional copula function is defined as in Eq. (18).

$$F(y_1, y_2, \dots, y_m) = C(F_1(y_1), F_2(y_2), \dots, F_m(y_m); \theta) \quad (18)$$

θ in Eq. (18) is called the dependency parameter and the marginal distributions of each of the quality characteristics express the relationship. The most basic theoretical determination about copula functions is put forward by the Scalar theorem.

Theorem 1. *(Sklar's Theorem) The m-dimensional copula is a function C defined from the m-dimensional interval $[0,1]^m$ to the unit interval $[0,1]$ and satisfies the following conditions* [41].

- $C(1, \dots, 1, a_n, 1, \dots, 1) = a_n$, $\forall n \leq m$ *and* $a_n \in [0,1]$.
- If $a_n = 0$ for any $n \leq m$, $C(a_1, \dots, a_m) = 0$.
- C is m-increasing.

In other words, the m-copula is an m-dimensional distribution function with m univariate marginals, each of which is uniformly distributed in the range $(0,1)$.

There are many copula functions belonging to different copula families in the literature. When its application areas are investigated, it is seen that it has widespread use in finance, actuarial, time series and risk analysis. In this study, the focus is on the Gaussian (normal) copula, which belongs to the elliptic copula family and has many useful features.

Definition 1. *(Gaussian Copula) Consider random variables $Z_1, Z_2, \dots, Z_k$ with correlation coefficients $\rho_{ij} = \rho(Z_i, Z_j)$ with multivariate normal probability distribution. Let the joint cumulative distribution function of the random variables $Z_1, Z_2, \dots, Z_k$ be $\Phi_G(z_1, z_2, \dots, z_k)$. In this case, the multivariate Gaussian (Normal) copula is defined in Eq. (19)* [42].

$$(u_1, \dots, u_k) = \Phi_G\left(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_k)\right) \quad (19)$$

The two-variable Gaussian (Normal) copula is in the form of Eq. (20).

$$C(u_1, u_2; \theta) = \Phi_G(\Phi^{-1}(u_1), \Phi^{-1}(u_2); \theta) = \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \frac{1}{2\pi(1-\theta^2)^{1/2}} \times \left\{\frac{-(s^2 - 2\theta st + t^2)}{2(1-\theta^2)}\right\} dsdt \quad (20)$$

In Eq. (20), Φ denotes the cumulative distribution function for the standard normal random variable and $\Phi_G(u_1, u_2)$ denotes the standard bivariate normal distribution with the correlation parameter θ, which takes values in the range of $(-1,1)$. This copula function was proposed by Lee in 1983 [43]*. The density function of the two-variable Gaussian copula is also in the form in Eq. (21).*

$$c(u_1,u_2;\theta)=\frac{1}{\sqrt{1-\theta^2}}\exp\left[\frac{-(u_1^2-2\theta u_1u_2+u_2^2}{2(1-\theta^2)}\right]\times\exp\left(\frac{u_1^2+u_2^2}{2}\right) \quad (21)$$

According to the scalar theorem, the bivariate probability distribution of the random vector $X=(X_1,X_2)'$ can be determined by the non-normal (any distribution) marginal distributions of the vector and the Gaussian copula [44].

In order to determine the probability distribution of a random vector $X=(X_1,X_2)'$, it is necessary to determine the marginal distribution of each X_j and find the dependency structure between X_j. In order to determine the dependency structure between random variables, it is necessary to mention the measures and some special dependency structures included in the copula functions. There is a relationship between copula functions expressing dependence and dependence measurements, especially for two-variable cases. Dependency can be measured by many methods. The Pearson correlation coefficient is one of them; it is sensitive to outliers and does not change under strictly increasing linear transformations. The expression of the Pearson correlation coefficient in terms of copulas is shown in Eq. (22) [45].

$$\rho_P(X,Y)=\frac{1}{\sigma_X\sigma_Y}\int_0^1\int_0^1[C(u_1,u_2)-u_1u_2]dF_X^{-1}(u_1)dF_Y^{-1}(u_2),\qquad u_i\in[0,1] \quad (22)$$

4.2. Integration of state-space model with copula modeling

In this section, the state vector of quality characteristics under dependency is estimated by combining the state-space model, Kalman filtering and copula functions for multi-stage manufacturing processes. Therefore, a unique approach has been introduced to monitor quality in a multi-stage manufacturing process.

4.2.1. Prediction error

Considering the general state-space model representation of a multi-stage manufacturing process with Eq. (3) and Eq. (4), the Kalman filter method for estimating the state vector x_k is introduced in Section 3.1. In the prediction phase of the estimation, it was seen that the uncertainty in the state vector $\hat{x}_{k|k-1}$ is a function of the estimation of $\hat{x}_{k-1|k-1}$ and the covariance Q_k of ε_k. In the next step, the prediction error components for vector x_k are; The statistical inference prediction error for x_k is $x_k-\hat{x}_{k|k-1}$ and the prediction error for the observation vector y_k is η_k. Therefore, as expressed in Table 2, the conditional variance of the prediction error given in Eq. (23) should be evaluated as a function of the uncertainties or errors related to $\hat{x}_{k|k-1}$ and R_k.

$$P_k(w)=C_kP_{k|k-1}C_k'+H_kR_kH_k' \quad (23)$$

According to the information obtained up to stage or time $k-1$, based on the conditional probability distribution of x_k to $y_k-\hat{y}_{k|k-1}$, the final estimation $\hat{x}_{k|k}$ and its covariance $P_{k|k-1}$ are obtained. Assuming that the joint probability distribution $X=x_k$ and $Y=y_k-\hat{y}_{k|k-1}$ is a normal distribution given in Eq. (24), the conditional probability distribution X given that Y is $N(\mu_{X|y},\Sigma_{XX|y})$ with parameters $\mu_{X|y}=\mu_X+\Sigma_{XY}\Sigma_{YY}^{-1}(Y-\mu_Y)$ and $\Sigma_{XX|y}=\Sigma_{XX}-\Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{YX}$.

$$\left[\binom{\mu_X}{\mu_Y},\begin{pmatrix}\Sigma_{XX}&\Sigma_{XY}\\\Sigma_{YX}&\Sigma_{YY}\end{pmatrix}\right] \quad (24)$$

In Eq. (24), $\mu_X=\hat{x}_{k|k-1}$, $\Sigma_{XX}=P_{k|k-1}$, $\Sigma_{XY}=P_{k|k-1}C_k'$ and $\Sigma_{YY}=P_k(w)$ (as in Eq. (23)) are some definitions. It is seen that by using these definitions, the expressions $\hat{x}_{k|k}$ ve $P_{k|k}$ which are the final estimates in the second stage of the Kalman filter, will be reached (see Table 2). The importance of the conditional variance $P_k(w)$ of the prediction error in the estimation of x_k can be revealed from another perspective. For example; assuming that x_0,ε_k and η_k are random variables whose joint distribution is the normal distribution, the probability distribution of the random vector y_k conditional on the information $\{y_1,y_2,\dots,y_{k-1}\}$ is normal distribution $N[y_{k|k-1},P_k(w)]$, the estimates of x_k depend on the parameters of the distribution, conditional expected value and conditional variance-covariance $y_{k|k-1}$ and $P_k(w)$, respectively. The estimation of the model parameters of interest can be achieved by maximizing the function given in Eq. (25), which expresses the log-likelihood in the context of all observation values.

$$L_C^{(k)}(whole)=\ln\mathcal{L}=-\frac{1}{2}\sum_k\ln[2\pi|P_k(w)|]-\frac{1}{2}\sum_k w_k'[P_k(w)]^{-1}w_k=\sum_{k=1}^{N}L_C^{(k)} \quad (25)$$

In the filtering stage, equations that express the estimates $\hat{x}_{k|k}$ and $P_{k|k}$ emerge. This was mentioned in Section 3.1, where the error quantities $y_k-\hat{y}_{k|k-1}=w_k$, and the variance expression of these error quantities $P_k(w)$ are highlighted and discussed in state vector estimation. It has been emphasized more clearly that it constitutes the necessary essential element. Maximazing the likelihood function specified in Eq. (25) is equivalent to minimizing the quantity in Eq. (26), especially under the assumption of a normal distribution in general.

$$J=E[y_k-C_k\hat{x}_{k|k-1}][y_k-C_k\hat{x}_{k|k-1}]' \quad (26)$$

The assumptions and definitions for an illustrative example of this when the state vector x_k is a two-element random vector with two quality characteristics are as follows:

- x_k is a 2×1 dimensional state vector and y_k is a 2×1 dimensional measurement vector.
- The quality characteristics in the measurement vector have marginal distributions $y_{k1}\sim F_{k1}(.)$ and $y_{k2}\sim F_{k2}(.)$, respectively, and it is assumed that the internal dependency structure between them is modeled with an appropriate copula. The cumulative joint probability distribution of random variables y_{k1} and y_{k2} can be determined through copula functions as $F_{k12}(y_{k1}, y_{k2})$.
- $E(y_{ki}) = \mu_{ki}, \quad Var(y_{k1}) = \sigma_{ki}^2, \quad E(y_{ki}^2) = \sigma_{ki}^2 + \mu_{ki}^2, \quad k = 1,2,\dots,N; \quad i = 1,2.$
- $Cov(y_{k1}, y_{k2}) = \sigma_{k12}, \quad E(y_{k1}y_{k2}) = \sigma_{k12} + \mu_{k1}\mu_{k2}, \quad k = 1,2,\dots,N.$
- It is assumed that the 2×2 dimensional matrices A_k, B_k and C_k are known.
- $y_k = \begin{bmatrix} y_{k1} \\ y_{k2} \end{bmatrix}, \quad \hat{x}_{k|k-1} = \begin{bmatrix} \hat{x}_{k1} \\ \hat{x}_{k2} \end{bmatrix}, \quad \hat{x}_{k-1|k-1} = \begin{bmatrix} \hat{x}_{(k-1)1} \\ \hat{x}_{(k-1)2} \end{bmatrix}, \quad u_k = \begin{bmatrix} u_{k1} \\ u_{k2} \end{bmatrix} \quad k = 1,2,\dots,N.$
- $B_k = \begin{bmatrix} b_{11}^{(k)} & b_{12}^{(k)} \\ b_{21}^{(k)} & b_{22}^{(k)} \end{bmatrix}, \quad A_{k-1} = \begin{bmatrix} a_{11}^{(k-1)} & a_{12}^{(k-1)} \\ a_{21}^{(k-1)} & a_{22}^{(k-1)} \end{bmatrix}, \quad C_k = \begin{bmatrix} c_{11}^{(k)} & c_{12}^{(k)} \\ c_{21}^{(k)} & c_{22}^{(k)} \end{bmatrix}, \quad k = 1,2,\dots,N.$

If Eq. (26) is rewritten according to the definitions, the matrix in Eq. (27) is obtained.

$$J = E\left[y_k - C_k\hat{x}_{k|k-1}\right]\left[y_k - C_k\hat{x}_{k|k-1}\right]' = E\begin{bmatrix} N & L \\ L & M \end{bmatrix} \quad (27)$$

The expansion of the matrix elements in Eq. (27) is given in Eqs. (28)-(30).

$$N = \left(y_{k1} - c_{11}^{(k)}\hat{x}_{k1} - c_{12}^{(k)}\hat{x}_{k2}\right)^2 \quad (28)$$

$$M = \left(y_{k2} - c_{21}^{(k)}\hat{x}_{k1} - c_{22}^{(k)}\hat{x}_{k2}\right)^2 \quad (29)$$

$$L = \left(y_{k1} - c_{11}^{(k)}\hat{x}_{k1} - c_{12}^{(k)}\hat{x}_{k2}\right) \times \left(y_{k2} - c_{21}^{(k)}\hat{x}_{k1} - c_{22}^{(k)}\hat{x}_{k2}\right) \quad (30)$$

When the elements of matrix J are considered separately, the expected values in Eqs. (31)-(33) are obtained.

$$E(N) = \sigma_{k1}^2 + \mu_{k1}^2 - 2\left(c_{11}^{(k)}\hat{x}_{k1} + c_{12}^{(k)}\hat{x}_{k2}\right)\mu_{k1} + \left(c_{11}^{(k)}\hat{x}_{k1} + c_{12}^{(k)}\hat{x}_{k2}\right)^2 \quad (31)$$

$$E(M) = \sigma_{k2}^2 + \mu_{k2}^2 - 2\left(c_{21}^{(k)}\hat{x}_{k1} + c_{22}^{(k)}\hat{x}_{k2}\right)\mu_{k1} + \left(c_{21}^{(k)}\hat{x}_{k1} + c_{22}^{(k)}\hat{x}_{k2}\right)^2 \quad (32)$$

$$E(L) = \sigma_{k12} + \mu_{k1}\mu_{k2} - \left(c_{21}^{(k)}\hat{x}_{k1} + c_{22}^{(k)}\hat{x}_{k2}\right)\mu_{k1} - \left(c_{11}^{(k)}\hat{x}_{k1} + c_{12}^{(k)}\hat{x}_{k2}\right)\mu_{k2} + \left(c_{11}^{(k)}\hat{x}_{k1} + c_{12}^{(k)}\hat{x}_{k2}\right) \times \left(c_{21}^{(k)}\hat{x}_{k1} + c_{22}^{(k)}\hat{x}_{k2}\right) \quad (33)$$

If the system transition matrix C_k is optimized (minimum), the partial derivatives of the expected values according to the elements of the C_k matrix are equal to zero. Then, the values in Eq. (34) and Eq. (35) for $c_{11}^{(k)}$ and $c_{12}^{(k)}$ are obtained.

$$c_{11}^{(k)} = \frac{\mu_{k1} - c_{12}^{(k)}\hat{x}_{k2}}{\hat{x}_{k1}} \quad (34)$$

$$c_{12}^{(k)} = \frac{\mu_{k1} - c_{11}^{(k)}\hat{x}_{k1}}{\hat{x}_{k2}} \quad (35)$$

It is necessary to test that the expressions in Eq. (34) and Eq. (35) are the values that minimize $E(N)$. The values found for this are the values that make the second derivatives of $E(N)$ with respect to $c_{11}^{(k)}$ and $c_{12}^{(k)}$ greater than zero, ($\partial^2 E(N)/\partial\left(c_{11}^{(k)}\right)^2 = 2\hat{x}_{k1}^2 > 0$ and $\partial^2 E(N)/\partial\left(c_{12}^{(k)}\right)^2 = 2\hat{x}_{k2}^2 > 0$, will be the values that minimize $E(N)$. Similarly, if the partial derivatives according to $c_{21}^{(k)}$ and $c_{22}^{(k)}$ in $E(M)$ are taken and set equal to zero, the expressions in Eq. (36) and Eq. (37) are obtained.

$$c_{21}^{(k)} = \frac{\mu_{k2} - c_{22}^{(k)}\hat{x}_{k2}}{\hat{x}_{k1}} \quad (36)$$

$$c_{22}^{(k)} = \frac{\mu_{k2} - c_{21}^{(k)}\hat{x}_{k1}}{\hat{x}_{k2}} \quad (37)$$

Since, $\partial^2 E(M)/\partial\left(c_{21}^{(k)}\right)^2 = 2\hat{x}_{k1}^2 > 0$ and $\partial^2 E(M)/\partial\left(c_{22}^{(k)}\right)^2 = 2\hat{x}_{k2}^2 > 0$, the values in Eq. (36) and Eq. (37) are the values that minimize $E(M)$. If these values are substituted in $E(L)$, the result will be as in Eq. (38) for k=1,2,…,N.

$$E(L) = E(Y_{k1}Y_{k2}) - \mu_{k1}\mu_{k2} = Cov(Y_{k1}, Y_{k2}) = \sigma_{k12}, \quad k = 1,\dots,N \quad (38)$$

In conclusion, the dependency between quality characteristics at any stage k is a phenomenon that affects the quality values of the production process. In the derivation made above, it is seen that the variance and covariance values directly affect the values symbolizing the quality status of the system, under the assumptions about the moments of the *Y* variables, which express the observable values of the *X* variables, which are the quality characteristics. Considering that variance and covariance values are quantities that determine correlation values; The conclusion is that the dependence, which can be expressed in general and specifically in the context of Gaussian copulas, is effective in the Kalman filter state estimation equations.

To state this more clearly, let us consider the Pearson correlation measure $\rho_P = (y_{k1}, y_{k2}) = \sigma_{k12}(\sigma_{k1}\sigma_{k2})^{-1}$ in the context of observations for a two-element state vector. Pearson correlation measure can be expressed as a function of the copula function $C(.,.)$ and the marginal distributions F_1 and F_2, as shown in Eq. (22), in the form in Eq. (39).

$$\rho_P(y_{k1}, y_{k2}) = (\sigma_{k1}\sigma_{k2})^{-1} \int_0^1 \int_0^1 [C(u_1, u_2) - u_1 u_2]\, dF_1^{-1}(u_1) F_2^{-1}(u_2), \quad u_i \in [0,1], \quad i = 1,2 \qquad (39)$$

By expressing the covariance σ_{k12} given in Eq. (38) in terms of copula, using Eq. (22), the adequacy of combining copula functions in the estimation of state-space models through the Kalman filter is demonstrated with Eq. (40).

$$\sigma_{k12} = (\sigma_{k1}\sigma_{k2})\rho_P(y_{k1}, y_{k2}) = \int_0^1 \int_0^1 [C(u_1, u_2) - u_1 u_2]\, dF_1^{-1}(u_1) F_2^{-1}(u_2) \qquad (40)$$

4.2.2. Copula likelihood functions

Considering the copula functions and Sklar's Theorem, the joint probability distribution function for the elements of the observation vector y_k, which takes continuous values, will be in the form in Eq. (41) with the expression of the copula function.

$$F_{k12}(y_{k1}, y_{k2}; \gamma, \theta) = C(F_{k1}(y_{k1}; \gamma), F_{k2}(y_{k2}, \gamma); \theta) \qquad (41)$$

In Eq. (41), the vector γ represents the probability distribution parameters except the dependence parameter θ between y_{k1} and y_{k2}. It is not necessary for θ parameter to express only correlation. If the distribution function in Eq. (41) is differentiated according to (y_{k1}, y_{k2}), the joint probability density function in Eq. (42) is obtained $k = 1,2, \dots, N$.

$$f_{k12}(y_{k1}, y_{k2}; \gamma, \theta) = c(F_{k1}(y_{k1}; \gamma_1), F_{k2}(y_{k2}; \gamma_2); \theta) \times f_{k1}(y_{k1}; \gamma_1) f_{k2}(y_{k2}; \gamma_2) \qquad (42)$$

$c(.,.;\theta)$ in Eq. (42) is the copula density function corresponding to $C(.,.;\theta)$. Assuming that there are n observations that can be expressed as $(y_{k11}, \dots, y_{k1n})$ and $(y_{k21}, \dots, y_{k2n})$ for each of the y_k vector elements y_{k1} and y_{k2} at any stage or time k, the copula log-likelihood function, For k =1,2,…,N, the expression in Eq. (43) is obtained.

$$\mathcal{L}_c^{(k)} = \sum_{i=1}^{n} \ln f_{k12}(y_{1k}, y_{2k}; \gamma, \theta) = \sum_{i=1}^{n} \ln c(F_{k1}(y_{k1}; \gamma_1), F_{k2}(y_{k2}; \gamma_2); \theta) + \sum_{i=1}^{n} \ln f_{k1}(y_{k1}; \gamma_1) + \sum_{i=1}^{n} \ln f_{k2}(y_2; \gamma_2) \qquad (43)$$

The difference between the expression in Eq. (43) and the ordinary log-likelihood function is that the sum of the log copula density functions is included in the equation. In the observation vector given in Eq. (3) where $H_k = I_k$, $\eta_{k1} = y_{k1} - c_{11}^{(k)} x_{k1} - c_{12}^{(k)} x_{k2}$ and $\eta_{k2} = y_{k2} - c_{21}^{(k)} x_{k1} - c_{22}^{(k)} x_{k2}$ are defined as in the covariance matrix R_k in Eq. (44) for $y_k = (y_{k1}, y_{k2})'$ with $\eta_k = (\eta_{k1}, \eta_{k2})' \sim N(0, R_k)$, $k = 1,2, \dots, N$.

$$R_k = \begin{bmatrix} Var(\eta_{k1}) & Cov(\eta_{k1}, \eta_{k2}) \\ Cov(\eta_{k1}, \eta_{k2}) & Var(\eta_{k2}) \end{bmatrix} = \begin{bmatrix} \sigma_{\eta_{k1}}^2 & \rho_k \sigma_{\eta_{k1}} \sigma_{\eta_{k2}} \\ \rho_k \sigma_{\eta_{k1}} \sigma_{\eta_{k2}} & \sigma_{\eta_{k2}}^2 \end{bmatrix} \qquad (44)$$

When the marginal density functions and joint probability density functions are given in Eq. (45), Eq. (46) and Eq. (47), respectively, the corresponding Gaussian (normal) copula density function expression can be calculated only with the help of the marginal distribution functions as in Eq. (48).

$$f_{k1}\left(y_{k1}; x_{k1}, x_{k2}, c_{11}^{(k)}, c_{12}^{(k)}\right) = \frac{1}{\sqrt{2\pi\sigma_{\eta_{k1}}^2}} \exp\left\{-\frac{\left(y_{k1} - c_{11}^{(k)} x_{k1} - c_{12}^{(k)} x_{k2}\right)^2}{2\sigma_{\eta_{k1}}^2}\right\} \qquad (45)$$

$$f_{k2}\left(y_{k2}; x_{k1}, x_{k2}, c_{21}^{(k)}, c_{22}^{(k)}\right) = \frac{1}{\sqrt{2\pi\sigma_{\eta_{k2}}^2}} \exp\left\{-\frac{\left(y_{k2} - c_{21}^{(k)} x_{k1} - c_{22}^{(k)} x_{k2}\right)^2}{2\sigma_{\eta_{k2}}^2}\right\} \qquad (46)$$

$$f_{k12}(y_k; x_k, C_k, R_k) = \frac{1}{2\pi\sqrt{|R_k|}} \times \exp\left\{-\frac{1}{2}(y_k - C_k x_k)' R_k^{-1} (y_k - C_k x_k)\right\} \qquad (47)$$

$$c(F_{k1}(y_{k1}; x_{k1}, \gamma_{1k}), F_{k2}(y_{k2}; x_{k2}, \gamma_{2k}); \theta) = \frac{1}{\sigma_{\eta_{k1}}^2 \sigma_{\eta_{k2}}^2 - \rho_k^2} \exp\left\{-\frac{(\eta_{k1}^2 - 2\rho_k \eta_{k1}\eta_{k2} + \eta_{k2}^2)}{2(1-\rho_k^2)}\right\} \times \exp\left(\frac{\eta_{k1}^2 + \eta_{k2}^2}{2}\right) \qquad (48)$$

As seen from the copula log-likelihood function obtained in Eq. (43), the value size of the copula log-likelihood function is determined by the dependency parameter values when other parameters are given. Under normal distribution, the dependence parameter θ is the parameter expressed in terms of moment factors and corresponding to the Pearson correlation.

4.2.3. Copula functions and Kalman filter

In order to define the stochastic dependency structure between the quality characteristics of a product in multi-stage manufacturing processes with copula functions, it is sufficient to know the marginal probability distributions of the characteristics. By

analyzing the representation of multi-stage manufacturing processes with the state-space models approach under dependency, it is possible to make Kalman filter estimations better trackable and interpretable on the basis of copula likelihood functions. To better express this, it would be useful to express the combination of Kalman filter estimation steps with copula functions, as shown in Table A1 in Appendix.

For the explicit expression of the copula log-likelihood functions $L_c(prediction)$ and $L_c(whole)$ in Table A1, it is necessary to know or predict the likelihood distribution models for the state-space model state vector x_k and therefore the observation vector y_k. In the predictions made for the Kalman filter method state-space model, normal distribution is assumed for the relevant model variables, and it is stated by many researchers that the predictions are efficient under these conditions [42].

For this reason, it is necessary to determine copula density functions and copula log-likelihood functions under certain distributions by using Eq. (43) to express the $L_c(prediction)$ and $L_c(whole)$ functions in Table A1.

For example; when the joint probability distribution of x_0, e_k and η_k is a normal distribution, the $L_c(prediction)$ and $L_c(whole)$ functions for the y_k observation vector will be as in Eq. (49) and Eq. (50).

$$L_c(prediction) = \sum_{i=1}^{n} \ln f_{k12}(y_{1k}, y_{2k}; \gamma, \theta)$$

$$= \sum_{i=1}^{n} \ln c(F_{k1}(y_{k1}; \gamma_1), F_{k2}(y_{k2}; \gamma_2); \theta)$$

$$+ \sum_{i=1}^{n} \ln f_{k1}(y_{k1}; \gamma_1) + \sum_{i=1}^{n} \ln f_{k2}(y_2; \gamma_2) \quad (49)$$

$$L_c(whole) = \sum_{i=1}^{N} L_c(prediction) \quad (50)$$

The copula density function in Eq. (49) is defined in Eq. (48). On the other hand, γ_1 and γ_2 in the expressions $f_{k1}(y_{k1}; \gamma_1)$ and $f_{k2}(y_{k2}; \gamma_2)$ defined in Eq. (45) and Eq. (46) show the distribution parameters.

It has been stated in the previous sections that in determining the x_k and $P_{k|k}$ expressions in the filtering stage of the Kalman filter equations, the likelihood function should be maximized or, equivalently, the sum of squares of the errors $w_k = y_k - \hat{y}_{k|k-1}$ should be minimized. .In this regard, copula log-likelihood functions must be determined to write the $L_c(prediction)$ and $L_c(whole)$ expressions shown in Table A1, the joint probability function of the random variables x_k and $w_k = y_k - \hat{y}_{k|k-1}$ with normal distribution and the copula function were used.

Let $y_k = (y_{k1}, y_{k2})'$ be the values observed about the quality characteristics of the production process at the kth stage of the multi-stage manufacturing process. In the case of the existence of an observation set of size n, considering the equations $x_{k|k-1}, P_{k|k-1}$, $E(w_k w_k') = P_k(w)$ in Table A1, the probability density functions and the copula density function are given in Eqs.(51)-(54) where $w_{ki} = y_{ki} - y_{k|k-1}^{(i)}$, $i = 1,2$.

$$f_{k1}(w_{k1}; x_{k1}, \gamma_1)$$

$$= \frac{1}{\sqrt{2\pi\sigma_{k1}^2}} \exp\left\{ - \frac{\left(y_{k1} - y_{k|k-1}^{(1)}\right)^2}{2\sigma_{w_{k1}}^2} \right\}, \quad (51)$$

$$f_{k2}(w_{k2}; x_{k2}, \gamma_2)$$

$$= \frac{1}{\sqrt{2\pi\sigma_{k2}^2}} \exp\left\{ - \frac{\left(y_{k2} - y_{k|k-1}^{(2)}\right)^2}{2\sigma_{w_{k2}}^2} \right\}, \quad (52)$$

$$f_{k12}(w_{k1}, w_{k2}; x_{k1}, x_{k2}, \gamma_1, \gamma_2, \theta)$$

$$= \frac{1}{2\pi\sqrt{|P_k(w)|}} \exp\left\{ - \frac{1}{2} \left(y_k - y_{k|k-1}\right)' P_k(w) \left(y_k - y_{k|k-1}\right) \right\}, \quad (53)$$

$$c(F_{k1}(w_{k1}; x_{k1}, \gamma_1), F_{k2}(w_{k2}; x_{k2}, \gamma_2); \theta)$$

$$= \frac{1}{\sigma_{w_{k1}}^2 \sigma_{w_{k2}}^2 - \rho_k^2} \exp\left\{ - \frac{(w_{k1}^2 - 2\rho_k w_{k1} w_{k2} + w_{k2}^2)}{2(1 - \rho_k^2)} \right\}$$

$$\times \exp\left(\frac{w_{k1}^2 + w_{k2}^2}{2} \right) \quad (54)$$

$y_{k|k-1}^{(i)}, i = 1,2$ in Eq. (51) and Eq. (52) shows the prediction made for the ith element of vector y_k based on the values observed until the kth stage. $P_k(w)$ in Eq. (53) shows the error covariance matrix in the Kalman filter prediction stage and is in the form in Eq. (55). ρ_k, which shows the correlation between w_{k1} and w_{k2} in the sense of Pearson correlation, is a copula correlation parameter for y_{k1} and y_{k2} since $w_k = y_k - \hat{y}_{k|k-1}$ and $y_{k|k-1}$ are calculated values.

$$P_k(w) = \begin{bmatrix} \sigma_{w_{k1}}^2 & \sigma_{w_{k1}} \sigma_{w_{k2}} \rho_k \\ \sigma_{w_{k1}} \sigma_{w_{k2}} \rho_k & \sigma_{w_{k2}}^2 \end{bmatrix} \quad (55)$$

Then, the copula log-likelihood function derived by the log-likelihood expression in Eq. (43) is given Eq. (56).

$$L_c(w_k | y_1, \dots, y_{k-1})$$

$$= \sum_{j=1}^{n} \ln c\left(F_{k1}\left(w_{k1j}; x_{k1}, \gamma_1\right), F_{k2}\left(w_{k2j}; x_{k2}, \gamma_2\right); \theta\right)$$

$$+ \sum_{j=1}^{n} \ln f_{k1}\left(w_{k1j}; x_{k1}, \gamma_1\right)$$

$$+ \sum_{j=1}^{n} \ln f_{k2}\left(w_{k2j}; x_{k2}, \gamma_2\right) \quad (56)$$

As seen in Eq. (56), when the copula density function dependence parameter is different from zero, the copula log-likelihood function creates an increasing or

decreasing effect on its value. The equations $x_{k|k}$ and $P_{k|k}$, which are the expressions of the Kalman filter filtering stage estimates, emerge as a result of the optimization of the copula log-likelihood functions according to w_{k1} and w_{k2} are obtained by using maximum likelihood (MLE) or minimization of the mean squared error (MMSE) approaches, effective estimates for x_k are obtained. When the marginal probability distributions of observation vectors y_k and error vectors w_k are distributions other than the normal distribution, the effectiveness of Kalman filter estimations may decrease.

5. Application of the proposed approach

In this section, a multi-stage manufacturing process is simulated with the modeling method presented and the results are discussed. The production process in the simulation study is based on assumption. Assume that parts are processed in the wood workshop of a factory. The wooden pieces, which are processed through a three-stage manufacturing process, are expected to weigh 150 grams (g) and be 30 centimeters (cm) long at the end of the production process. When unprocessed wood pieces arrive at the factory, they are weighed and their lengths are measured in the input quality control department. Based on past measurements, it will be assumed that the lengths of untreated wood pieces have a normal distribution with a mean of 32 cm and a standard deviation of 0.5 cm. Similarly, the weights of the raw parts will be assumed to have a normal distribution with a mean of 152g and a standard deviation of 1.1g. It will be assumed that the parts entering the processing process are cut in the first stage, sanded in the second stage and polished in the last stage. Fig.(1) shows a representative version of this process.

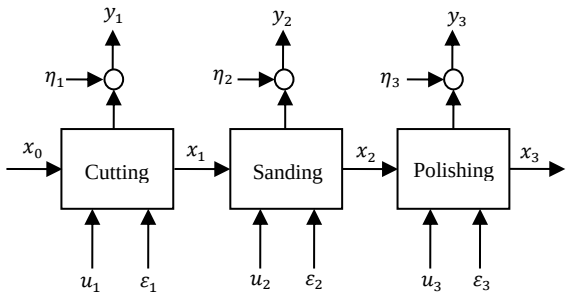


Figure 1. Three-stage processing scheme

The wooden parts, first checked in the input quality control department, are cut to the desired size during the cutting stage. After the parts are sanded, they move on to the polishing stage. Then, the products are left to dry to take their final form. It is assumed that the parts are weighed at the end of each stage and their length measured. Under these assumptions, the state equation will be as in Eq. (57), and the output (measurement) equation will be as in Eq. (58) where $D_k = I_k$ and $H_k = I_k$.

$$x_k = A_{k-1}x_{k-1} + B_k u_k + \varepsilon_k, \qquad k = 1,2,3 \quad (57)$$

$$y_k = C_k x_k + \eta_k \ , \qquad k = 1,2,3 \quad (58)$$

In a multi-stage processing process revealed by the linear state-space model, x_k is the directly unobservable quality characteristic of the product being inspected. x_k is a vector that contains all the information about the current state of the process at the kth stage. In this simulation study, the state vector x_k consists of two quality characteristics: x_{k1} represents the actual value of the part size at the kth stage; x_{k2} shows the actual value of the part weight at the kth stage. The vectors $x_k = [x_{k1} \quad x_{k2}]', k = 1,2,3$ in size (2×1) are known positive definite matrices that show the deviation during the transition from the kth stage to the $(k+1)$th stage of the process given as
$A_0 = [-0.645 \quad 0.343\,; \, -3.165 \quad 1.660]$,
$A_1 = [-0.639 \quad 0.329; \, -3.170 \quad 1.670]$ and
$A_2 = [-0.761 \quad 0.352; \, -3.660 \quad 1.730]$
with MATLAB notation.

B_k is defined as the input matrix at the kth stage and $u_k = \left[u_1^{(k)} \quad u_2^{(k)}\right]', k = 1,2,3$ is defined as a (2×1) dimensional vector showing the contribution of the kth stage in the state equation. Here $u_i^{(k)}, i = 1,2; \, k = 1,2,3$ is the contribution of the kth stage to the ith quality characteristic. This contribution is provided by the multiplication of the known matrix B_k and the vector u_k. In this application, $B_1 u_1 = [-0.01 \; -0.025]'$, $B_2\, u_2 = [-0.01 \; -0.01]'$ and $B_3 u_3 = [0.01 \;\; 0.01]'$. Let the unobservable process noise be defined as $\varepsilon_k = \left[e_1^{(k)} \quad e_2^{(k)}\right]', k = 1,2,3$. In the simulation study, it is assumed that $e_i^{(k)} \sim N(\mu = 0, \sigma^2), i = 1,2; \, k = 1,2,3$. It is also assumed that the measurements can be taken from every stage. In this case, the measurement vector $Y_k = [y_{k1} \quad y_{k2}]', k = 1,2,3$ can be observed for every value of the phase index k. Let the C_k matrices, which provide the transition between the actual values of the quality characteristics and the measurement values, be determined as $C_1 = [1.30 \;\; -0.01; 0.01 \;\; 0.99], C_2 = [1.01 \;\; -0.01; 0.01 \;\; 0.99]$, and $C_3 = [1.01 \;\; -0.001; 0.01 \;\; 0.99]$.

The elements of the vector $\eta_k = \left[\xi_1^{(k)} \xi_2^{(k)}\right]', k = 1,2,3$, which show the measurement error are distributed normally given as $\xi_i^{(k)} \sim N(\mu = 0\,, \sigma^2), i = 1,2; k = 1,2,3$. In this study, different correlation coefficient values (0.99, 0.90, 0.70, 0.50, 0.30 and 0.1) were tested for 0.1, 0.5 and 1 values of the σ^2 parameter, which indicates the noise level. Additionally, it is assumed that the dependency structure between the length and the weight measurement values can be determined with the Gaussian copula.

In this case, the Kalman filter equations in Table 2 will be taken into consideration for the observed values of quality characteristics $y_k = (y_{k1}, y_{k2})'$ in the kth stage of the multi-stage manufacturing process. Pearson correlation (ρ) between w_{k1} and w_{k2} is a copula correlation parameter for y_{k1} and y_{k2} since $w_k = y_k - \hat{y}_{k|k-1}$ and $y_{k|k-1}$ are calculated values. The values of

the Gaussian copula dependence parameter which models the dependency structure between y_{k1} and y_{k2} are calculated for each stage by a MATLAB code. As previously shown in Eq. (22) and Eq. (23), by expressing the σ_{k12} covariance in terms of copula, the adequacy of combining copula functions in the estimation of state-space models through the Kalman filter was demonstrated. The obtained Gaussian copula dependence parameter values were used instead of ρ_k in the matrix given in Eq. (55). The prediction stage covariance matrix $P_k(w)$ is included in the Kalman gain matrix as in $K_k = P_{k|k-1}C_k' P_k^{-1}(w)$. As a result, the Kalman filter equations and copula functions have been integrated.

Assume that 100 wooden parts go through this machining process under the defined conditions. Due to the structure of the multi-stage manufacturing process, the output of the previous stage will be the input for any stage. For example, the outputs from the cutting stage will be the input for the sanding stage. The outputs obtained in the sanding stage will be the input for the polishing stage (see in Figure 1).

Table 3. MAPE values (%) for the simulation study

		Cutting		Sanding		Polishing	
ρ	σ^2	L	W	L	W	L	W
0.99	0.1	4.95	9.28	0.85	0.84	8.85	1.07
0.99	0.5	4.80	9.27	2.26	1.14	8.91	1.62
0.99	1.0	5.03	9.25	2.57	1.28	8.49	1.96
0.90	0.1	4.97	9.27	1.21	0.93	8.51	1.11
0.90	0.5	4.92	9.24	2.06	1.17	8.93	1.66
0.90	1.0	4.64	9.27	3.55	1.68	7.90	2.05
0.70	0.1	4.95	9.26	1.12	0.89	8.64	1.08
0.70	0.5	4.87	9.26	2.09	1.15	8.51	1.77
0.70	1.0	4.74	9.26	3.09	1.37	8.52	2.02
0.50	0.1	4.96	9.25	1.02	0.88	8.76	1.28
0.50	0.5	4.84	9.26	2.23	1.14	8.64	1.44
0.50	1.0	4.75	9.30	2.48	1.18	9.05	2.06
0.30	0.1	4.90	9.27	1.01	0.90	8.71	1.15
0.30	0.5	4.99	9.25	2.18	1.19	8.78	1.55
0.30	1.0	4.91	9.29	4.02	2.11	8.94	3.04
0.10	0.1	4.88	9.24	1.29	0.95	8.76	1.09
0.10	0.5	4.97	9.26	2.32	1.21	8.50	1.96
0.10	1.0	4.95	9.23	2.90	1.19	8.81	1.71

A MATLAB code was written to obtain simulation values for the quality characteristics, weight (W) and length (L), for each production stage under the assumptions. The mean absolute percentage error (MAPE) criterion was used to measure the performance of the Kalman filter model under the copula dependency. Table 3 displays the MAPE values that are obtained for various noise levels (σ^2) and correlation coefficients (ρ). Since every MAPE value is less than 10%, it is evident that the proposed model, which provides remarkably accurate predictions, allows for the examination of the dependencies between quality characteristics at every stage [46].

6. Conclusions

In this study, the state-space model established for the dependency between the stages in the multi-stage manufacturing process is integrated with the copula modeling used to reveal the internal dependency structure between the quality characteristics in a stage. The importance of the conditional variance of the prediction error in the Kalman filter equations in the estimation of x_k has been revealed and it has been emphasized that it constitutes the main element that needs to be addressed. In the application part of the study, first, the dependency structure between the quality variables of interest in a hypothetical production process was expressed with copulas, system state predictions were made with the Kalman filter, and evaluated under the mean absolute percentage error (MAPE) criterion. The resulting model has shown that it is a model that allows examining the dependency between quality characteristics at every stage and gives extremely accurate predictions.

The original contribution of this study to the theory, method and practice on the subject is as follows: The Kalman filter estimation method, based on state-space models of multi-stage manufacturing processes, has been presented in a broad perspective, with a solution algorithm proposed and subject-specific comments. In order to take into account the statistical dependence between the quality characteristics of interest at any stage of the process, the dependence was expressed with copula functions and integrated with the Kalman filter method.

The innovations and improvements that the specified original contributions brought to the modeling and analysis of multi-stage manufacturing processes are as follows: The fact that quality characteristics are essentially interdependent is reflected in the models and internalized in the analyses. Model components, structure and calculation steps that are dependent on modeling and analysis are clearly stated. The internal dependency structure that can exist between the quality characteristics of interest at any production stage is integrated with the dependency structure between the stages.

In order to further the results put forward in the study, future studies that are deemed useful are as follows: Prediction methods for various copulas that can be used in modeling the internal dependency between the quality characteristics of interest at any stage of multi-stage manufacturing process structures can be investigated and implemented. State-space modeling generalizations involving dependency can be made with multivariate copula models for more than two quality characteristics. In the presence of models

containing noise terms and observation errors with distributions other than normal distribution, the robustness of Kalman filter estimates can be addressed.

Acknowledgements

The authors would like to thank the editor and anonymous referees for helpful comments and suggestions.

References

[1] Montgomery, D. C. (2019). Introduction to statistical quality control. John wiley & sons.

[2] Hayter, A. J., & Tsui, K. L. (1994). Identification and quantification in multivariate quality control problems. Journal of quality technology, 26(3), 197-208.

[3] Hotelling, H, (1947). Multivarite Quality Control, Techniques of Statistical Analysis, Einshart, Hastay and Wallis (eds.), McGraw-Hill, New York, A.B.D.

[4] Hicks, C. R. (1955). Some Applications of Hotelling's T2. Industrial Quality Control, 11(9), 23-26.

[5] Hawkins, D. M. (1991). Multivariate quality control based on regression-adiusted variables. Technometrics, 33(1), 61-75.

[6] Hawkins, D. M. (1993). Regression adjustment for variables in multivariate quality control. Journal of quality Technology, 25(3), 170-182.

[7] Zhang, G. X. (1985). Cause-selecting control charts-a new type of quality control charts. The QR Journal, 12(4), 221-225.

[8] Wade, M. R., & Woodall, W. H. (1993). A review and analysis of cause-selecting control charts. Journal of quality technology, 25(3), 161-169.

[9] Shu, L., & Tsung, F. (2003). On multistage statistical process control. Journal of the Chinese Institute of Industrial Engineers, 20(1), 1-8.

[10] Ding, Y., Ceglarek, D., & Shi, J. (2002). Fault diagnosis of multistage manufacturing processes by using state space approach. J. Manuf. Sci. Eng., 124(2), 313-322.

[11] Lawless, J. F., Mackay, R. J., & Robinson, J. A. (1999). Analysis of variation transmission in manufacturing processes—part I. Journal of Quality Technology, 31(2), 131-142.

[12] Agrawal, R., Lawless, J. F., & Mackay, R. J. (1999). Analysis of variation transmission in manufacturing processes—part II. Journal of Quality Technology, 31(2), 143-154.

[13] Jin, J., & Shi, J. (1999). State space modeling of sheet metal assembly for dimensional control, *J. Manuf. Sci. Eng*., 121(4): 756-762.

[14] Basseville, M., & Nikiforov, I. V. (1993). Detection of abrupt changes: theory and application (Vol. 104). Englewood Cliffs: prentice Hall.

[15] Tsung, F., Li, Y., & Jin, M. (2008). Statistical process control for multistage manufacturing and service operations: a review and some extensions. International Journal of Services Operations and Informatics, 3(2), 191-204.

[16] Lindsey, J. K. (2000). A family of models for uniform and serial dependence in repeated measurements studies. Journal of the Royal Statistical Society: Series C (Applied Statistics), 49(3), 343-357.

[17] Junker, M., Szimayer, A., & Wagner, N. (2006). Nonlinear term structure dependence: Copula functions, empirics, and risk implications. *Journal of Banking & Finance*, *30*(4), 1171-1199.

[18] Hafner, C. M., & Manner, H. (2012). Dynamic stochastic copula models: Estimation, inference and applications. *Journal of applied econometrics*, *27*(2), 269-295.

[19] Goto, H. (2013). Model predictive control-based scheduler for repetitive discrete event systems with capacity constraints. *An International Journal of Optimization and Control: Theories & Applications (IJOCTA)*, *3*(2), 73-83.

[20] Creal, D. D., & Tsay, R. S. (2015). High dimensional dynamic stochastic copula models. *Journal of Econometrics*, *189*(2), 335-345.

[21] Alpay, O., & Hayat, E. A. (2017). Copula approach to select input/output variables for DEA. *An International Journal of Optimization and Control: Theories & Applications (IJOCTA)*, *7*(1), 28-34.

[22] Zhang, Y., & Choudhry, T. (2017). Forecasting the daily time-varying beta of European banks during the crisis period: comparison between GARCH models and the Kalman Filter. *Journal of Forecasting*, *36*(8), 956-973.

[23] Fernández, M., García, J. E., & González-López, V. A. (2018). A copula-based partition Markov procedure. *Communications in Statistics-Theory and Methods*, *47*(14), 3408-3417.

[24] Smith, M. S., & Maneesoonthorn, W. (2018). Inversion copulas from nonlinear state space models with an application to inflation forecasting. *International Journal of Forecasting*, *34*(3), 389-407.

[25] Wang, H., Meng, A., Liu, Y., Fu, X., & Cao, G. (2019). Unscented Kalman Filter based interval state estimation of cyber physical energy system for detection of dynamic attack. *Energy*, *188*, 116036.

[26] Xu, J., Liang, Z., Li, Y. F., & Wang, K. (2021). Generalized condition-based maintenance optimization for multi-component systems considering stochastic dependency and imperfect maintenance. *Reliability Engineering & System Safety*, *211*, 107592.

[27] Kreuzer, A., Dalla Valle, L., & Czado, C. (2022). A Bayesian non-linear state space copula model for air pollution in Beijing. *Journal of the Royal Statistical*

Society Series C: Applied Statistics, *71*(3), 613-638.

[28] Ly, S., Sriboonchitta, S., Tang, J., & Wong, W. K. (2022). Exploring dependence structures among European electricity markets: Static and dynamic copula-GARCH and dynamic state-space approaches. *Energy Reports*, *8*, 3827-3846.

[29] Wang, Z., Xu, X., Trajcevski, G., Zhang, K., Zhong, T., & Zhou, F. (2022, June). PrEF: Probabilistic electricity forecasting via Copula-augmented state space model. In *Proceedings of the AAAI*.

[30] Kreuzer, A., Dalla Valle, L., & Czado, C. (2023). Bayesian multivariate nonlinear state space copula models. Computational Statistics & Data Analysis, 188, 107820.

[31] Schweppe, F.C. (1973) Uncertain Dynamic Systems. Prentice-Hall.

[32] Kim, C-J., Nelson, C.R. (1999) State-Space Models with Regime Switching. The MIT Press.

[33] Anderson, B.D.O. and Moore, J.B. (1979) Optimal Filtering, Prentice-Hall.

[34] Brown, R.G. and Hwang, P.Y.C. (1992) Introduction to Random Signals and Applied Kalman Filtering. John Wiley and Sons.

[35] Davis, M. H. A. and Vinter, R. B. (1984) Stochastic Modelling and Control, Chapman and Hall, London NewYork.

[36] Kalman, R. E. (1960). A new approach to linear filtering and prediction problems, *J. Basic Eng*. Mar 1960, 82(1): 35-45.

[37] Welch, G. and Bishop, G. An Introduction to the Kalman Filter. http://www.cs.unc.edu/~welch/media/pdf/kalman_intro.pdf. Last access date: 05.09.2024.

[38] Özbek, L. (2000) Durum-Uzay Modelleri ve Kalman Filtresi, *Gazi Ünv. Fen Bilimleri Ens. Dergisi*, 13(1), 113-126.

[39] Bryson, A., & Johansen, D. (1965). Linear filtering for time-varying systems using measurements containing colored noise. IEEE Transactions on Automatic Control, 10(1), 4-10.

[40] Simon, D. (2006). Optimal state estimation: Kalman, H infinity, and nonlinear approaches. John Wiley & Sons.

[41] Nelsen, R. B. (1999) An Introduction to Copulas. Lecture Notes in Statistics, Vol. 39, Spring-Verlag, New York.

[42] Trivedi, P. K., & Zimmer, D. M. (2007). Copula modeling: an introduction for practitioners. Foundations and Trends® in Econometrics, 1(1), 1-111.

[43] Lee, L. F. (1983). Generalized econometric models with selectivity. Econometrica: Journal of the Econometric Society, 507-512.

[44] Kolev, N., & Paiva, D. (2009). Copula-based regression models: A survey. Journal of statistical planning and inference, 139(11), 3847-3856.

[45] Drouet-Mari, D. and Kotz, S. (2001) Correlation and Dependence, First Edition, Imperial College Press, London.

[46] Lewis C. D. (1982) Industrial and Business Forecasting Methods, Butterworths, London.

[47] Toktaş, P., (2011) Çok Aşamalı İmalat Süreçleri için İstatistiksel Proses Kontrolü Modellemesi, (*Doktora Tezi, Ankara Üniversitesi, Fen Bilimleri Enstitüsü*).

Pelin Toktaş *received the B.Sc. degree in Statistics from Middle East Technical University, Ankara, Turkey, in 1997, the M.Sc. degree in industrial engineering from Bilkent University, Ankara, Turkey, in 2000, and the Ph.D. degree in statistics from Ankara University, Ankara, Turkey, in 2011. She is currently working with Engineering Faculty, Industrial Engineering Department, Başkent University, Ankara, Turkey, as an Assistant Professor. She has written various articles, and many studies presented at various national and international conferences. Her research interests include statistical analysis and its applications in industrial engineering and social sciences, multicriteria decision making methods, and statistical quality control applications.*

http://orcid.org/0000-0001-6622-4646

Ömer Lütfi Gebizlioğlu *received his B.S. degree (high honors) in Economics and Statistics from Midde East Technical University, Turkey, M.Sc. degree in Statistics and Ph.D. degree in Administrative and Engineering Systems from the Union Graduate School of Union College and University, New York, USA. Professor Gebizlioğlu has had academic positions at the Statistics Departments of Middle East Technical University and Ankara University, and at the Department of International Trade and Finance of Kadir Has University. He received several research funds from the national research institutions in the USA and Türkiye. He has numerous scientific publications in the areas of theory and applications of probability and statistics, risk measures, risk theory and loss distributions, actuarial science, insurance mathematics and economics, financial risk management and system reliability.*

http://orcid.org/0000-0002-3824-281X

Appendix

Table A1. Kalman filter and copula functions [47]

Stages	Related Equations
Initial:	
x_0, P_0	
Prediction stage:	
$\hat{x}_{k\|k-1}$	(Table 2) $\hat{x}_{k\|k-1} = A_{k-1}\hat{x}_{k-1\|k-1} + B_k u_k$
$P_k(y)$	(Table 2) $P_k(y) = C_k\left[A_{k-1}P_{k-1\|k-1}A'_{k-1} + D_k Q_k D'_k\right]C'_k + H_k R_k H'_k$
$w_k = y_k - \hat{y}_{k\|k-1}$	(Table 2) $w_k = C_k\left(x_k - \hat{x}_{k\|k-1}\right) + H_k\eta_k$
$P_k(w) = E(w_k w'_k)$	(Table 2) $P_k(w) = C_k P_{k\|k-1}C'_k + H_k R_k H'_k$
Estimation prediction stage – copula likelihood function $L_c(prediction)$	General: Eq. (43) $\mathcal{L}_c^{(k)} = \sum_{i=1}^{n} \ln f_{k12}(y_{1k}, y_{2k}; \gamma, \theta)$ $= \sum_{i=1}^{n} \ln c(F_{k1}(y_{k1};\gamma_1), F_{k2}(y_{k2};\gamma_2); \theta)$ $+ \sum_{i=1}^{n} \ln f_{k1}(y_{k1};\gamma_1) + \sum_{i=1}^{n} \ln f_{k2}(y_2;\gamma_2)$ Specific: Copula likelihood function for $\hat{x}_{k\|k-1}$ and w_k
Estimation-filtering stage	
$\hat{x}_{k\|k}$	(Table 2) $\hat{x}_{k\|k} = \hat{x}_{k\|k-1} + K_k\{y_k - C_k\hat{x}_{k\|k-1}\}$
$P_{k\|k}$	(Table 2) $P_{k\|k} = [I - K_k C_k]P_{k\|k-1}$
Copula likelihood function for all process stages $L_c(whole)$	Generalization of Eq. (43)

This work is licensed under a Creative Commons Attribution 4.0 International License. The authors retain ownership of the copyright for their article, but they allow anyone to download, reuse, reprint, modify, distribute, and/or copy articles in IJOCTA, so long as the original authors and source are credited. To see the complete license contents, please visit http://creativecommons.org/licenses/by/4.0/.

INSTRUCTIONS FOR AUTHORS

Aims and Scope

An International Journal of Optimization and Control: Theories & Applications (IJOCTA) is a scientific, peer-reviewed, open-access journal that publishes original research papers and review articles of high scientific value in all areas of applied mathematics, optimization and control. It aims to focus on multi/inter-disciplinary research into the development and analysis of new methods for the numerical solution of real-world applications in engineering and applied sciences. The basic fields of this journal cover mathematical modeling, computational methodologies and (meta)heuristic algorithms in optimization, control theory and their applications. Note that new methodologies for solving recent optimization problems in operations research must conduct a comprehensive computational study and/or case study to show their applicability and practical relevance.

Journal Topics

The topics covered in the journal are (not limited to):

Applied Mathematics, Financial Mathematics, Control Theory, Optimal Control, Fractional Calculus and Applications, Modeling of Bio-systems for Optimization and Control, Linear Programming, Nonlinear Programming, Stochastic Programming, Parametric Programming, Conic Programming, Discrete Programming, Dynamic Programming, Nonlinear Dynamics, Stochastic Differential Equations, Optimization with Artificial Intelligence, Operational Research in Life and Human Sciences, Heuristic and Metaheuristic Algorithms in Optimization, Applications Related to Optimization in Engineering.

Submission of Manuscripts

New Submissions

Solicited and contributed manuscripts should be submitted to IJOCTA via the journal's online submission system. You need to make registration prior to submitting a new manuscript (please click here to register and do not forget to define yourself as an "Author" in doing so). You may then click on the "New Submission" link on your User Home.

IMPORTANT: If you already have an account, please click here to login. It is likely that you will have created an account if you have reviewed or authored for the journal in the past.

On the submission page, enter data and answer questions as prompted. Click on the “Next” button on each screen to save your work and advance to the next screen. The names and contact details of at least four internationally recognized experts who can review your manuscript should be entered in the "Comments for the Editor" box.

You will be prompted to upload your files: Click on the “Browse” button and locate the file on your computer. Select the description of the file in the drop down next to the Browse button. When you have selected all files you wish to upload, click the “Upload” button. Review your submission before sending to the Editors. Click the “Submit” button when you are done reviewing. Authors are responsible for verifying all files have uploaded correctly.

You may stop a submission at any phase and save it to submit later. Acknowledgment of receipt of the manuscript by IJOCTA Online Submission System will be sent to the corresponding author, including an assigned manuscript number that should be included in all subsequent correspondence. You can also log-on to submission web page of IJOCTA any time to check the status of your manuscript. You will receive an e-mail once a decision has been made on your manuscript.

Each manuscript must be accompanied by a statement that it has not been published elsewhere and that it has not been submitted simultaneously for publication elsewhere.

Manuscripts can be prepared using LaTeX (.tex) or MSWord (.docx). However, manuscripts with heavy mathematical content will only be accepted as LaTeX files.

Preferred first submission format (for reviewing purpose only) is Portable Document File (.pdf). Please find below the templates for first submission.

Click here to download Word template for first submission (.docx)

Click here to download LaTeX template for first submission (.tex)

Revised Manuscripts

Revised manuscripts should be submitted via IJOCTA online system to ensure that they are linked to the original submission. It is also necessary to attach a separate file in which a point-by-point explanation is given to the specific points/questions raised by the referees and the corresponding changes made in the revised version.

To upload your revised manuscript, please go to your author page and click on the related manuscript title. Navigate to the "Review" link on the top left and scroll down the page. Click on the "Choose File" button under the "Editor Decision" title, choose the revised article (in pdf format) that you want to submit, and click on the "Upload" button to upload the author version. Repeat the same steps to upload the "Responses to Reviewers/Editor" file and make sure that you click the "Upload" button again.

To avoid any delay in making the article available freely online, the authors also need to upload the source files (Word or LaTex) when submitting revised manuscripts. Files can be compressed if necessary. The two-column final submission templates are as follows:

Click here to download Word template for final submission (.docx)

Click here to download LaTeX template for final submission (.tex)

Authors are responsible for obtaining permission to reproduce copyrighted material from other sources and are required to sign an agreement for the transfer of copyright to IJOCTA.

Article Processing Charges

There are **no charges** for submission and/or publication.

English Editing

Papers must be in English. Both British and American spelling is acceptable, provided usage is consistent within the manuscript. Manuscripts that are written in English that is ambiguous or incomprehensible, in the opinion of the Editor, will be returned to the authors with a request to resubmit once the language issues have been improved. This policy does not imply that all papers must be written in "perfect" English, whatever that may mean. Rather, the criteria require that the intended meaning of the authors must be clearly understandable, i.e., not obscured by language problems, by referees who have agreed to review the paper.

Presentation of Papers

Manuscript style

Use a standard font of the **11-point type: Times New Roman** is preferred. It is necessary to single line space your manuscript. Normally manuscripts are expected not to exceed 25 single-spaced pages including text, tables, figures and bibliography. All illustrations, figures, and tables are placed within the text at the appropriate points, rather than at the end.

During the submission process you must enter: (1) the full title, (2) names and affiliations of all authors and (3) the full address, including email, telephone and fax of the author who is to check the proofs. Supply a brief **biography** of each author at the end of the manuscript after references.

- Include the name(s) of any **sponsor(s)** of the research contained in the paper, along with **grant number(s).**
- Enter an **abstract** of no more than 250 words for all articles.

Keywords

Authors should prepare no more than 5 keywords for their manuscript.

Maximum five **AMS Classification number** (http://www.ams.org/mathscinet/msc/msc2010.html) of the study should be specified after keywords.

Writing Abstract

An abstract is a concise summary of the whole paper, not just the conclusions. The abstract should be no more than 250 words and convey the following:

1. An introduction to the work. This should be accessible by scientists in any field and express the necessity of the experiments executed.
2. Some scientific detail regarding the background to the problem.
3. A summary of the main result.
4. The implications of the result.
5. A broader perspective of the results, once again understandable across scientific disciplines.

It is crucial that the abstract conveys the importance of the work and be understandable without reference to the rest of the manuscript to a multidisciplinary audience. Abstracts should not contain any citation to other published works.

Reference Style

Reference citations in the text should be identified by numbers in square brackets "[ ]". All references must be complete and accurate. Please ensure that every reference cited in the text is also present in the reference list (and vice versa). Online citations should include date of access. References should be listed in the following style:

Journal article

Author, A.A., & Author, B. (Year). Title of article. Title of Journal, Vol(Issue), pages.

Castles, F.G., Curtin, J.C., & Vowles, J. (2006). Public policy in Australia and New Zealand: The new global context. Australian Journal of Political Science, 41(2), 131–143.

Book

Author, A. (Year). Title of book. Publisher, Place of Publication.

Mercer, P.A., & Smith, G. (1993). Private Viewdata in the UK. 2nd ed. Longman, London.

Chapter

Author, A. (Year). Title of chapter. In: A. Editor and B. Editor, eds. Title of book. Publisher, Place of publication, pages.

Bantz, C.R. (1995). Social dimensions of software development. In: J.A. Anderson, ed. Annual review of software management and development. CA: Sage, Newbury Park, 502–510.

Internet document

Author, A. (Year). Title of document [online]. Source. Available from: URL [Accessed (date)].

Holland, M. (2004). Guide to citing Internet sources [online]. Poole, Bournemouth University. Available from: http://www.bournemouth.ac.uk/library/using/guide_to_citing_internet_sourc.html [Accessed 4 November 2004].

Newspaper article

Author, A. (or Title of Newspaper) (Year). Title of article. Title of Newspaper, day Month, page, column.

Independent (1992). Picking up the bills. Independent, 4 June, p. 28a.

Thesis

Author, A. (Year). Title of thesis. Type of thesis (degree). Name of University.

Agutter, A.J. (1995). The linguistic significance of current British slang. PhD Thesis. Edinburgh University.

Illustrations

Illustrations submitted (line drawings, halftones, photos, photomicrographs, etc.) should be clean originals or digital files. Digital files are recommended for highest quality reproduction and should follow these guidelines:

- 300 dpi or higher
- Sized to fit on journal page
- TIFF or JPEG format only
- Embedded in text files and submitted as separate files (if required)

Tables and Figures

Tables and figures (illustrations) should be embedded in the text at the appropriate points, rather than at the end. A short descriptive title should appear above each table with a clear legend and any footnotes suitably identified below.

Proofs

Page proofs are sent to the designated author using IJOCTA EProof system. They must be carefully checked and returned within 48 hours of receipt.

Offprints/Reprints

Each corresponding author of an article will receive a PDF file of the article via email. This file is for personal use only and may not be copied and disseminated in any form without prior written permission from IJOCTA.

Submission Preparation Checklist

As part of the submission process, authors are required to check off their submission's compliance with all of the following items, and submissions may be returned to authors that do not adhere to these guidelines.

1. The submission has not been previously published, nor is it before another journal for consideration (or an explanation has been provided in Comments for the Editor).
2. The paper is in PDF format and prepared using the IJOCTA's two-column template.
3. All references cited in the manuscript have been listed in the References list (and vice-versa) following the referencing style of the journal.
4. There is no copyright material used in the manuscript (or all necessary permissions have been granted).
5. Details of all authors have been provided correctly.
6. ORCID profile numbers of "all" authors are mandatory, and they are provided at the end of the manuscript as in the template (visit https://orcid.org for more details on ORCID).
7. The text adheres to the stylistic and bibliographic requirements outlined in the Author Guidelines.
8. Maximum five AMS Classification number (http://www.ams.org/mathscinet/msc/msc2010.html) of the study have been provided after keywords.
9. The names and email addresses of at least FOUR (4) possible reviewers have been indicated in "Comments for the Editor" box in "Paper Submission Step 1-Start". Please note that at least two of the recommendations should be from different countries. Avoid suggesting reviewers you have a conflict of interest.

Peer Review Process

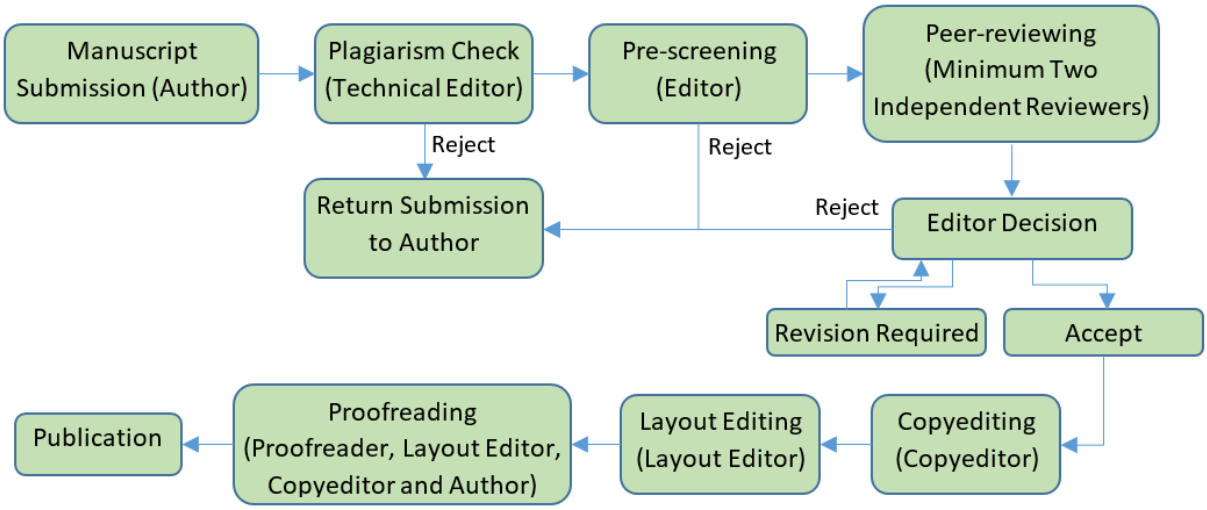
All contributions, prepared according to the author guidelines and submitted via IJOCTA online submission system are evaluated according to the criteria of originality and quality of their scientific content. The corresponding author will receive a confirmation e-mail with a reference number assigned to the paper, which he/she is asked to quote in all subsequent correspondence.

All manuscripts are first checked by the Technical Editor using plagiarism detection software (iThenticate) to verify originality and ensure the quality of the written work. If the result is not satisfactory (i.e. exceeding the limit of 30% of overlapping), the submission is rejected and the author is notified.

After the plagiarism check, the manuscripts are evaluated by the Editor-in-Chief and can be rejected without reviewing if considered not of sufficient interest or novelty, too preliminary or out of the scope of the journal. If the manuscript is considered suitable for further evaluation, it is first sent to the Area Editor. Based on his/her opinion the paper is then sent to at least two independent reviewers. Each reviewer is allowed up to four weeks to return his/her feedback but this duration may be extended based on his/her availability. IJOCTA has instituted a blind peer review process where the reviewers' identities are not known to authors. When the reviews are received, the Area Editor gives a decision and lets the author know it together with the reviewer comments and any supplementary files.

Should the reviews be positive, the authors are expected to submit the revised version usually within two months the editor decision is sent (this period can be extended when the authors contact to the editor and let him/her know that they need extra time for resubmission). If a revised paper is not resubmitted within the deadline, it is considered as a new submission after all the changes requested by reviewers have been made. Authors are required to submit a new cover letter, a response to reviewers letter and the revised manuscript (which ideally shows the revisions made in a different color or highlighted). If a change in authorship (addition or removal of author) has occurred during the revision, authors are requested to clarify the reason for change, and all authors (including the removed/added ones) need to submit a written consent for the change. The revised version is evaluated by the Area editor and/or reviewers and the Editor-in-Chief brings a decision about final acceptance based on their suggestions. If necessary, further revision can be asked for to fulfil all the requirements of the reviewers.

When a manuscript is accepted for publication, an acceptance letter is sent to the corresponding author and the authors are asked to submit the source file of the manuscript conforming to the IJOCTA two-column final submission template. After that stage, changes of authors of the manuscript are not possible. The manuscript is sent to the Copyeditor and a linguistic, metrological and technical revision is made, at which stage the authors are asked to make the final corrections in no more than a week. The layout editor prepares the galleys and the authors receive the galley proof for final check before printing. The authors are expected to correct only typographical errors on the proofs and return the proofs within 48 hours. After the final check by the layout editor and the proofreader, the manuscript is assigned a DOI number, made publicly available and listed in the forthcoming journal issue. After printing the issue, the corresponding metadata and files published in this issue are sent to the databases for indexing.



Publication Ethics and Malpractice Statement

IJOCTA is committed to ensuring ethics in publication and quality of articles. Conforming to standards of expected ethical behavior is therefore necessary for all parties (the author, the editor(s), the peer reviewer) involved in the act of publishing.

International Standards for Editors

The editors of the IJOCTA are responsible for deciding which of the articles submitted to the journal should be published considering their intellectual content without regard to race, gender, sexual orientation, religious belief, ethnic origin, citizenship, or political philosophy of the authors. The editors may be guided by the policies of the journal's editorial board and constrained by such legal requirements

as shall then be in force regarding libel, copyright infringement and plagiarism. The editors may confer with other editors or reviewers in making this decision. As guardians and stewards of the research record, editors should encourage authors to strive for, and adhere themselves to, the highest standards of publication ethics. Furthermore, editors are in a unique position to indirectly foster responsible conduct of research through their policies and processes.

To achieve the maximum effect within the research community, ideally all editors should adhere to universal standards and good practices.

- Editors are accountable and should take responsibility for everything they publish.
- Editors should make fair and unbiased decisions independent from commercial consideration and ensure a fair and appropriate peer review process.
- Editors should adopt editorial policies that encourage maximum transparency and complete, honest reporting.
- Editors should guard the integrity of the published record by issuing corrections and retractions when needed and pursuing suspected or alleged research and publication misconduct.
- Editors should pursue reviewer and editorial misconduct.
- Editors should critically assess the ethical conduct of studies in humans and animals.
- Peer reviewers and authors should be told what is expected of them.
- Editors should have appropriate policies in place for handling editorial conflicts of interest.

Reference:

Kleinert S & Wager E (2011). Responsible research publication: international standards for editors. A position statement developed at the 2nd World Conference on Research Integrity, Singapore, July 22-24, 2010. Chapter 51 in: Mayer T & Steneck N (eds) Promoting Research Integrity in a Global Environment. Imperial College Press / World Scientific Publishing, Singapore (pp 317-28). (ISBN 978-981-4340-97-7) [Link].

International Standards for Authors

Publication is the final stage of research and therefore a responsibility for all researchers. Scholarly publications are expected to provide a detailed and permanent record of research. Because publications form the basis for both new research and the application of findings, they can affect not only the research community but also, indirectly, society at large. Researchers therefore have a responsibility to ensure that their publications are honest, clear, accurate, complete and balanced, and should avoid misleading, selective or ambiguous reporting. Journal editors also have responsibilities for ensuring the integrity of the research literature and these are set out in companion guidelines.

- The research being reported should have been conducted in an ethical and responsible manner and should comply with all relevant legislation.
- Researchers should present their results clearly, honestly, and without fabrication, falsification or inappropriate data manipulation.
- Researchers should strive to describe their methods clearly and unambiguously so that their findings can be confirmed by others.
- Researchers should adhere to publication requirements that submitted work is original, is not plagiarised, and has not been published elsewhere.
- Authors should take collective responsibility for submitted and published work.
- The authorship of research publications should accurately reflect individuals' contributions to the work and its reporting.
- Funding sources and relevant conflicts of interest should be disclosed.
- When an author discovers a significant error or inaccuracy in his/her own published work, it is the author's obligation to promptly notify the journal's Editor-in-Chief and cooperate with them to either retract the paper or to publish an appropriate erratum.

Reference:

Wager E & Kleinert S (2011) Responsible research publication: international standards for authors. A position statement developed at the 2nd World Conference on Research Integrity, Singapore, July 22-24, 2010. Chapter 50 in: Mayer T & Steneck N (eds) Promoting Research Integrity in a Global Environment.

Imperial College Press / World Scientific Publishing, Singapore (pp 309-16). (ISBN 978-981-4340-97-7) [Link].

Basic principles to which peer reviewers should adhere

Peer review in all its forms plays an important role in ensuring the integrity of the scholarly record. The process depends to a large extent on trust and requires that everyone involved behaves responsibly and ethically. Peer reviewers play a central and critical part in the peer-review process as the peer review assists the Editors in making editorial decisions and, through the editorial communication with the author, may also assist the author in improving the manuscript.

Peer reviewers should:

- respect the confidentiality of peer review and not reveal any details of a manuscript or its review, during or after the peer-review process, beyond those that are released by the journal;
- not use information obtained during the peer-review process for their own or any other person's or organization's advantage, or to disadvantage or discredit others;
- only agree to review manuscripts for which they have the subject expertise required to carry out a proper assessment and which they can assess within a reasonable time-frame;
- declare all potential conflicting interests, seeking advice from the journal if they are unsure whether something constitutes a relevant conflict;
- not allow their reviews to be influenced by the origins of a manuscript, by the nationality, religion, political beliefs, gender or other characteristics of the authors, or by commercial considerations;
- be objective and constructive in their reviews, refraining from being hostile or inflammatory and from making libellous or derogatory personal comments;
- acknowledge that peer review is largely a reciprocal endeavour and undertake to carry out their fair share of reviewing, in a timely manner;
- provide personal and professional information that is accurate and a true representation of their expertise when creating or updating journal accounts.

Reference:

Homes I (2013). COPE Ethical Guidelines for Peer Reviewers, March 2013, v1 [Link].

Copyright Notice

All articles published in An International Journal of Optimization and Control: Theories & Applications (IJOCTA) are made freely available with our Open Access policy without any publication/subscription fee.

Under the CC BY license, authors retain ownership of the copyright for their article, but authors grant others permission to use the content of publications in IJOCTA in whole or in part provided that the original work is properly cited. Users (redistributors) of IJOCTA are required to cite the original source, including the author's names, IJOCTA as the initial source of publication, year of publication, volume number and DOI (if available).

Authors grant IJOCTA the right of first publication. Although authors remain the copyright owner, they grant the journal the irrevocable, nonexclusive rights to publish, reproduce, publicly distribute and display, and transmit their article or portions thereof in any manner.

Articles are published under the Creative Commons Attribution 4.0 International License (CC BY 4.0).

An International Journal of Optimization and Control: Theories & Applications

Volume: 14 Number: 4
October 2024

CONTENTS

294 List coloring based algorithm for the Futoshiki puzzle
Banu Baklan Şen, Oznur Yasar Diner

308 Early prediction of fabric quality using machine learning to reduce rework in manufacturing processes
Sema Aydın, Koray Altun

322 Witte's conditions for uniqueness of solutions to a class of Fractal-Fractional ordinary differential equations
Abdon Atangana, Ilknur Koca

336 Influence of rotation on peristaltic flow for pseudoplastic fluid: a wavy channel
Hayat Adel Ali, Mohammed R. Salman

346 A comparative view to H_infinity-norm of transfer functions of linear DAEs
Hasan Gündüz, Ercan Çelik, Mesut Karabacak

355 Fuzzy-PID and interpolation: a novel synergetic approach to process control
Devashish Jha, Arifa Ahmed, Sanatan Kumar, Debanjan Roy

365 Global mathematical analysis of a patchy epidemic model
Lahcen Boulaasair, Hassane Bouzahir, Mehmet Yavuz

378 An Inverse recursive algorithm to retrieve the shape of the inaccessible dielectric objects
Ahmet Sefer

394 A local differential quadrature method for the generalized nonlinear Schrödinger (GNLS) equation
Meirikim Panmei, Roshan Thoudam

404 Modeling the dependency structure between quality characteristics in multi-stage manufacturing processes with copula functions
Pelin Toktaş, Ömer Lütfi Gebizlioğlu

www.ijocta.org
editor@ijocta.org

ISSN 2146-0957