Rubem P. Mondaini *Editor*

Trends in Biomathematics: Stability and Oscillations in Environmental, Social, and Biological Models

Selected Works from the BIOMAT Consortium Lectures, Rio de Janeiro, Brazil, 2021

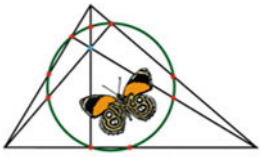
Springer

Trends in Biomathematics: Stability and Oscillations in Environmental, Social, and Biological Models

Rubem P. Mondaini
Editor

Trends in Biomathematics: Stability and Oscillations in Environmental, Social, and Biological Models

Selected Works from the BIOMAT Consortium Lectures, Rio de Janeiro, Brazil, 2021



Springer

Editor
Rubem P. Mondaini
BIOMAT Consortium
International Institute for Interdisciplinary
Mathematical and Biological Sciences
Rio de Janeiro, Brazil

Federal University of Rio de Janeiro
Rio de Janeiro, Brazil

ISBN 978-3-031-12514-0 ISBN 978-3-031-12515-7 (eBook)
https://doi.org/10.1007/978-3-031-12515-7

Mathematics Subject Classification: 92Bxx, 92-08, 92-10

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The publication of the present book follows the fine tradition of books already published with selected papers accepted at the BIOMAT International Symposia on Interdisciplinary Mathematical and Biological Sciences. All these papers were selected for publication as chapters in the book, after their presentation in the scientific sessions of the conference and after a final peer review. We hoped that the pandemic situation will improve and that we can have a face-to-face conference in 2021. Unfortunately, we again had to organize the conference in an online format. We believe that once again the tradition of BIOMAT conferences in terms of the absence of parallel sessions, as well as their fundamental mission of strengthening the interdisciplinary scientific activities of our colleagues from developing countries, has been so efficient to support this tradition.

We hope that security conditions will allow the organization of at least one hybrid conference in 2022. We are however quite sure that if occasional difficulties should again impede the hope of organizing a face-to-face conference that would be desirable for a true interdisciplinary conference series, our tradition and the sense of duty of all our colleagues in the scientific community of the BIOMAT Consortium will once again suffice to preserve the professional competence of the organization and the scientific participation that is essentially due to all of them.

For this second online version of a BIOMAT Symposium, BIOMAT 2021—21st International Symposium on Mathematical and Computational Biology, we again benefited from the cooperation of the Brazilian Academic Network—RNP, and Dr. Beatriz Zoss, to whom we now express our sincere thanks for her professional expertise. BIOMAT 2021 benefited from the cooperation of all colleagues/authors of accepted papers and abstracts to bring together professional researchers and research students from 17 countries on four continents. The Scientific Program, which was followed with all necessary rigor, had 23 contributed lectures, 16 Keynote Speaker lectures, and 2 Keynote Speaker Tutorial lectures. We chose to work with Greenwich Mean Time (GMT), and the lectures were scheduled from 10:00 to 16:45 GMT, in order to maximize the presence of participants in countries from four continents. The lectures were organized into 20 scientific sessions from Monday morning, November 1, to Friday evening, November 5.

At the end of the opening session on Monday morning, November 1, 2021, 10:00 GMT, a few words were said in honor of Professor Nick Britton, who passed away in December 2020. We transcribe here what was said as a testimony of our deep gratitude to what our friend Nick Britton did for the BIOMAT Consortium and the honor of having him as a friend.

"At this opening session of the 21st BIOMAT International Symposium on Mathematical and Computational Biology, we have the duty and honor to pay tribute to Professor Nicholas Ferris Britton, Professor Emeritus of the University of Bath, UK."

Nick Britton, as he became known throughout the world, especially in the 67 countries where he developed scientific activities, was responsible for the invitation made to the BIOMAT Consortium to organize the 2010 Annual Meeting of the Society for Mathematical Biology in Rio de Janeiro, Brazil, as a joint conference with the BIOMAT 2010 International Symposium.

Nick Britton was a keynote speaker at the BIOMAT International Symposia held in 2008, 2009, 2010, 2011, and 2014. He was also a member of the Consortium Board of Directors as one of the representatives of Western Europe. He was besides everything a wonderful human being, a very polite gentleman, and a great scientist. He had an enormous scientific dedication to interdisciplinary topics in mathematics and biology.

Nick passed away in December 2020, just before Christmas. He is deeply missed by all his colleagues, as well as former students and family members, such as his wife, mother, daughter, and three sisters, for the pain of his loss and the remembrance of what he represents to the entire scientific community. We would like to ask everyone to observe a minute's silence to pay tribute to an old and great friend.

The Editor of the book would like to thank again the professional help given by his former PhD student, Dr. Simão C. de Albuquerque Neto, with all the conveniences (and inconveniences) in the use of the electronic platform and his kind way of helping the presentations of our colleagues in the scientific sessions. He also acknowledges the patience and competence of his wife Carmem Lucia S.C. Mondaini with the administration of technical matters related to the financial support of the conference, as well as for her collaboration with the reception of papers to be submitted to the peer review procedure. The Editor also takes this opportunity to inform you that our 12-year-old son, Romolo Mondaini, after trying to attend some of the talks in the time available between his classes, confirmed what he said last year during the BIOMAT 2020 conference that he also wants to be a scientist and a teacher. After 47 years as a university professor, all we can say is: May God forgive and help him!

Rio de Janeiro, Brazil R. P. Mondaini
November 06, 2021

Editorial Board of the BIOMAT Consortium

Contents

Whole-Graph Embedding and Adversarial Attacks for Life Sciences

L. Maddalena, M. Giordano, M. Manzo, and M. R. Guarracino

1 Introduction

Networks provide a suitable model for many scientific and technological problems where it is important to represent not only complex entities but also their relations, which can be derived from structural or functional information or their integration. Networks have been successfully used in many life science applications, where data is abundant and publicly accessible. Structural information and omics data are the most abundant and publicly available resources for extracting knowledge through graph theory-based approaches. Examples can be found in brain networks [48]), molecular structures data (e.g., chemical compounds [23], proteins structure [57] networks), or physical interactions between molecules (e.g., protein-protein [42], lncRNA–protein interaction networks [63]). In other applications, functional networks are derived from the integration of available knowledge with omics data analyses (e.g., metabolic [18, 22], gene-regulatory [13], co-expression [58], drug-disease association [67] networks). In such applications, various entities are used as nodes, such as gene, disease, RNA, etc. The edges between nodes mean known associations between pairs of entities, such as miRNA–disease interaction.

L. Maddalena (✉) · M. Giordano
National Research Council, Institute for High-Performance Computing and Networking (ICAR), Naples, Italy
e-mail: lucia.maddalena@icar.cnr.it

M. Manzo
ITS, University of Naples "L'Orientale", Naples, Italy

M. R. Guarracino
University of Cassino and Southern Lazio, Cassino, Italy
National Research University Higher School of Economics, Moscow, Russia

R. P. Mondaini (ed.), *Trends in Biomathematics: Stability and Oscillations in Environmental, Social, and Biological Models*,
https://doi.org/10.1007/978-3-031-12515-7_1

Once the problem has been modeled, different analysis techniques can be used to obtain more insight into it. Statistics and machine learning have provided many algorithms to tackle three fundamental problems. The first is related to *edge prediction*, which translates into disease association prediction (discovering the associated factors with various diseases, e.g., RNA–disease and disease–gene association), drug development and discovery (e.g., protein–protein and ligand–protein interaction prediction, prediction of molecular properties, de novo molecule design, drug response, and drug–drug interaction prediction), and medical imaging (e.g., early detection and evaluation of diseases based on medical images segmentation, classification, and recognition, brain connectivity research). The second fundamental problem is related to *node classification or regression*. This means predicting a class label or a continuous value for a node. This has been applied for the analysis of the proteome and transcriptome (e.g., protein structure and function prediction, alternative splicing). The last problem is related to the analysis of whole networks. Here, the sample under investigation contains different networks, one for each observation. This is the case when studying the metabolism of cancer cells [21] or the similarity of protein structures.

Graph embedding techniques provide efficient and effective network analysis techniques, as they aim at converting networks into a vector space while their structural and topological information is preserved. Single nodes or whole networks can be represented as vectors in the embedding space. Therefore, after this unsupervised embedding step, standard algorithms such as linear regression, generalized linear models, Support Vector Machine (SVM), and random forest, can be applied to such vector representations. In this paper, we focus on applications of whole-graph embedding techniques. This means we will only describe embedding techniques where samples composed of different networks are embedded in a vector space, in which each point represents a network.

Many whole-graph embedding methods have been developed [6], aimed at transforming graph data into vector data and automatically learning a low-dimensional feature representation for each node in the graph where the similarity (or distance) between the nodes is preserved as much as possible. Embedding-based algorithms are often faster than those working on the original networks. Additionally, the learned embedding can be used for downstream analysis, either by direct interpretation of the embedding space or by machine learning techniques applicable to vector data. Indeed, preserving the graph structure, the low-dimensional representations can be used as features for applications such as link prediction, community detection, graphs/node classification, and clustering. Most of the studies evaluate the graph embedding approaches on non-biomedical networks. Few attempts have been made to compare the efficiency of these methods on biomedical datasets, mainly focusing on node embedding approaches [40, 61], but also on whole-graph embedding [33].

Despite the considerable success of whole-graph embedding techniques, robustness to noise in training data has not been extensively evaluated yet. This is in particular important in life science applications, where corrupted data, either due to acquisition noise or to intentional modifications, could lead to misleading

conclusions. The difficulty might be found in the lack of interpretability of the models, which makes it difficult to apply standard optimization and statistical inference methods.

In this paper, we use adversarial machine learning techniques for evaluating and eventually enhancing the robustness of whole-graph embedding methods for supervised classification.

The paper is structured as follows. In Sect. 2, we focus on whole-graph embedding, defining its aim and giving an overview of the classes of methods frequently adopted in the literature. In Sect. 3, we focus on adversarial attacks for graphs, introducing interesting taxonomies for graph attacks and surveying the literature related to graph-level tasks. In Sect. 4, we present our evaluation approach, describing the considered graph attacks and the adopted strategy to compare the robustness to such attacks of whole-graph embedding methods. In Sect. 5, we present the results of our evaluation, describing the adopted life sciences datasets and implementation details. Conclusions and future research directions are given in Sect. 6.

2 Whole-Graph Embedding

In a very general definition, graph embedding is a mechanism for learning a mapping from a graph to a vector space still preserving main graph properties. Let $G = (V, E)$ represent a graph, where $V = \{v_i\}_{i=1}^{N}$ is the set of nodes, and $E \subseteq V \times V$ is the set of edges, each one represented by a pair of nodes (v_i, v_j).

Definition 2.1 Given a graph $G{=}(V, E)$, a *graph embedding* (or node-level graph embedding) is a mapping ϕ: $v_i \in V \rightarrow \mathbf{y}_i \in \mathbb{R}^d$, $i = 1, \ldots, N$, $d \in \mathbb{N}$, such that the function ϕ preserves some proximity measure defined on graph G.

Thus, a graph embedding is a space reduction that maps the nodes of a graph into a d-dimensional feature vector space, also known as *latent space*, trying to preserve local and structural information. In this way, graphs can be represented as compact yet informative vectors in the *latent space*, suitable to be effectively and efficiently treated by non-network-based machine learning methods [49].

Methods for node-level graph embedding are the most widespread in the literature for applications such as link prediction and node label predictions [6, 30, 37]. However, other graph-level tasks, such as classification, similarity ranking, generation, and visualization of graphs, require information at a higher granularity [10]. In these cases, whole-graph embedding, which allows representing a whole graph as a single vector of fixed length, appears to be a more appropriate mechanism.

Definition 2.2 Given a set of graphs $\mathcal{G} = \{G_i\}_{i=1}^{M}$ with the same set of nodes V, a *whole-graph embedding* is a mapping $\psi : G_i \rightarrow \mathbf{y}_i \in \mathbb{R}^d$, $i = 1, \ldots, M$, $d \in \mathbb{N}$, such that the function ψ preserves some proximity measure defined on $\mathcal{G}$.

Thus, whole-graph embedding takes a collection of graphs and outputs a fixed-length representation for each of them, still preserving important information on the entire graph set.

2.1 Approaches to Whole-Graph Embedding

The literature provides a set of interesting approaches aimed at transforming the space from graph to vector. For this purpose, the section is divided into four categories, as suggested in [46, 54]: direct methods, graph kernel methods, neural networks methods, and statistical and spectral representations. Some of the most relevant and recent methods are summarized below.

2.1.1 Direct Methods

Direct methods work to extract features from graphs directly. They were the first to appear in the literature (e.g., Graph Edit Distance [47]). Below are some of these recently published.

In [14], Dutta et al. propose an approach called Pyramidal Stochastic Graphlet Embedding (PSGE). The goal is to combine the result of pyramidal embedding representation to obtain a global delineation of the original graph. Similarly, Dutta et al. [15] consider the hierarchical structure of a graph constructed by topologically clustering the graph nodes and considering each cluster as a node in the upper hierarchical level. Once this hierarchical structure is built, several configurations are taken into account to define the mapping into a vector space.

In [66], Zhou et al. present a data augmentation approach on graphs based on four methods: random mapping, vertex-similarity mapping, motif-random mapping, and motif-similarity mapping, to generate more weakly labeled data via heuristic transformation. Furthermore, the authors propose a generic model evolution framework, called M-Evolve, which combines graph augmentation, data filtration, and model retraining to optimize pre-trained graph classifiers.

2.1.2 Kernel Methods

One of the most popular categories of space reduction for graph classification is that of graph kernel methods. Generally, a kernel defines a distance between pairs of elementary substructures of two graphs. The resulting matrix is the representation of the distance between the two given graphs. The main limitation concerns the large computational load for the calculation, and for this reason they were not adopted in the experimental phase. Below are some of these recently published.

Huang et al. [25] describe a graph kernel using a longest common subsequence (LCS) kernel to compute a more comprehensive similarity between paths and walks,

which resolves substructure isomorphism difficulties. Furthermore, the kernel is combined with optimal transport theory to extract more in-depth features of graphs.

In [51], Sun et al. propose a feature reduction algorithm based on semantic similarity for graph classification. First, it learns vector representations of subtree patterns using neural language models and then merges semantically similar subtree patterns into a new feature.

Bai et al. [1] describe two local-global nested graph kernels, namely the nested aligned kernel and the nested reproducing kernel, drawing on depth-based complexity traces. The first is defined by measuring the global alignment kernel, based on the dynamic time warping framework, between the complexity traces. The second is defined by measuring the basic reproducing kernel between the complexity traces.

2.1.3 Neural Network Methods

Graph neural networks (GNNs) have become the standard toolkit for analyzing and learning from data on graphs. The state of the art includes a wide range of models, such as PSCN [41], DGCNN [62], UGraphEmb [2, 3], DGNN [28], CommPOOL [53], GSSNN [68], EigenPooling [31]. Here, we focus on some of the methods that have been explored in the experimental phase.

Narayanan et al. [39] provide a Skip-Gram neural network model, called Graph2vec, typically adopted in the Natural Language Processing (NLP) domain. It learns data-driven distributed representations of arbitrarily sized graphs. The resulting embeddings are learned in an unsupervised manner and are task-unaware.

In [7], Chen et al. propose an extended version of Graph2vec. The method is named GL2vec (Graph and Line graph to vector) because it concatenates the embedding of an original graph to that of the corresponding line graph. The line graph is an edge-to-vertex dual graph of the original graph. Specifically, GL2vec integrates either the edge label information or the structural information, which Graph2vec misses with the embeddings of the line graph.

Manipur et al. [35] introduce Netpro2vec, a neural-network method that produces embedding of whole-graphs that are independent of the task and nature of the data. It first represents graphs as textual documents whose words are formed by processing probability distributions of graph node paths/distances (e.g., the Transition Matrix, TM, or the Node Distance Distribution, NDD). Then, it embeds graph documents by using the Doc2Vec method [29].

2.1.4 Statistical and Spectral Representations

Methods from this class use statistical properties and graph spectrum of a graph to generate graph feature vectors. Initial works in this field adopt summary-graph-statistics and handcrafted features, such as node degrees, to extract local graph

properties. Again, here we focus on some of the methods that have been explored in our experiments.

In [17], Galland et al. present a method called IGE, which extracts handcrafted invariant features based on graph spectral decomposition. These features are easy to compute, permutation-invariant, and include sufficient information on the graph's structure.

Tsitsulin et al. [54] introduce NetLSD. It computes a compact graph signature derived from the solution of the heat equation involving the normalized Laplacian matrix. It is permutation and size-invariant, scale-adaptive, and computationally efficient.

In [56], Verma et al. provide a graph representation, named FGSD, based on a family of graph spectral distances with uniqueness, stability, sparsity, and computational efficiency properties.

Rozemberczki et al. [45] propose FeatherGraph, which adopts characteristic functions defined on graph nodes to describe the distribution of node features at multiple scales. The probability weights of these functions are defined as the transition probabilities of random walks. The node-level features are combined by mean pooling to create graph-level statistics.

3 Adversarial Attacks for Graphs

Due to the demonstrated vulnerability of deep neural networks (DNNs), and consequently, of GNNs, recent research is moving toward the investigation of graph adversarial attacks and their countermeasures, as reviewed in [9, 27, 50]. While [9, 27] mainly concentrate on GNN-based methods, [50] also covers attack and defense models for non-GNN methods, even though all the considered methods devoted to graph-level tasks are based on GNNs. However, these surveys provide a broad and updated view of the field, also thanks to their accompanying repositories. Indeed, Chen et al. [9] created and maintain the *Awesome Graph Adversarial Learning repository*[1] with links to 271 related papers of the last five years (as of October 2021). Sun et al. [50] created the *Graph Adversarial Learning Literature repository*[2] that includes a curated list of more than 110 adversarial attack and defense papers on graph-structured data, together with links to available codes. Finally, Jin et al. [27] made available and maintain the *DeepRobust adversarial learning library*[3] that includes the majority of representative attack and defense algorithms for both graph data and image data.

[1] The Awesome Graph Adversarial Learning repository is available at https://github.com/gitgiter/Graph-Adversarial-Learning.

[2] The Graph Adversarial Learning Literature repository is available at https://github.com/safe-graph/graph-adversarial-learning-literature.

[3] The DeepRobust pytorch library is available at https://github.com/DSE-MSU/DeepRobust.

3.1 Taxonomies for Adversarial Attacks

Interesting taxonomies for graph attacks have been given, based on the goal, resources, knowledge, and capacity of the attackers [9, 27, 50].

Perturbation To generate adversarial samples on graph data, we can modify the nodes or edges in the original graph. In the case of *edge-level perturbations*, the attacker is capable of adding/removing/rewiring edges in the whole original graph within a given budget. In this case, the number of modified edges is often used to evaluate the magnitude of perturbation. With *node-level perturbations*, the attacker is capable of adding/removing nodes, or manipulating the features of target nodes. The related evaluation metric can be computed based on the number of nodes modified or the distance between the benign and adversarial feature vectors.

Stage Adversarial attacks can happen at two stages: evasion attack (model testing) and poisoning attacks (model training), depending on the attacker's capacity to insert adversarial perturbations. A *poisoning attack* tries to affect the performance of the model by adding adversarial samples into the training dataset. Most existing works are poisoning attacks, and their node classification tasks are performed in the transductive learning setting. In this case, once the attacker changes the data, the model is retrained. In an *evasion attack*, the parameters of the trained model are assumed to be fixed, and the attacker tries to generate the adversarial samples of the trained model. Evasion attacks only change the testing data, which does not require retraining the model.

Knowledge The attacker could hold different information on the model to attack it. Based on this, the dangerous levels of existing attacks can be characterized as: (1) *white-box attack*, where the attacker can get all information and use it to attack the system, such as the prediction result, gradient information, etc. The attack may not work if the attacker does not fully break the system first; (2) *grey-box attack*, where the attacker gets limited information to attack the system. Comparing to white-box attacks, it is more dangerous to the system since the attacker only needs partial information; (3) *black-box attack*, where the attacker can only do black-box queries on some of the samples. Thus, the attacker generally can not make poisoning attacks on the trained model. However, if a black-box attack can work, it would be the most dangerous attack compared with the other two because the attacker can attack the model with the most limited acknowledge; (4) *restricted black-box attack*, which is one step further than black-box, as only black-box queries on some of the samples can be done, and the attacker is asked to create adversarial modifications to other samples.

Goal According to the goals, attacks for graph-level classification can be divided into *targeted attacks* (or *integrity* attacks in [50]), that aim to induce the model to give a specific label to a given graph sample, and *untargeted attacks* (or *availability* attacks in [50]), that aim to insert poison so that the model just performs incorrectly.

Task Corresponding to different tasks for graph data, attacks can be considered for (1) *node-relevant tasks*, such as node classification or node embedding, (2) *link-relevant tasks*, such as link prediction, or (3) *graph-relevant tasks*, such as graph classification or graph clustering.

3.2 Adversarial Attack Methods for Graph-Level Tasks

Most of the literature on adversarial attacks for graph data is devoted to node-relevant and link-relevant tasks [9, 27, 50]; only limited attention is given to graph-relevant tasks, mainly for graph classification.

RL-S2V,[4] by Dai et al. [11], is one of the first examples of adversarial attacks to GNN-based graph classification. Here, an attacker perturbs the structure of a testing graph such that the target model misclassifies the perturbed test graph. It adopts a reinforcement learning technique to generate adversarial attacks on graph data under the black-box setting. The attack procedure is modeled as a Markov Decision Process adapted for both node-level (targeted) and graph-level (untargeted) attacks to standard GNN models.

Chen et al. [8] propose the GraphAttacker framework that allows adjusting the structures and the attack strategies according to different graph analysis tasks. Adversarial samples are generated based on GAN (Generative Adversarial Network) through alternate training on three key components: the Multi-strategy Attack Generator, the Similarity Discriminator, and the Attack Discriminator. To achieve attacks within a perturbation budget, they also introduce a Similarity Modification Rate constraint to quantify the similarity between nodes and limit the attack budget.

Tang et al. [52] focus on adversarial attacks on hierarchical graph convolutional networks (GCNs), and specifically, the Hierarchical Graph Pooling model [65], for graph classification tasks. They propose an adversarial graph generating strategy with a surrogate model to explore their vulnerability to adversarial samples. They also perform robust training on the target models to demonstrate that the retrained graph classification models can better defend against the attack from adversarial samples.

Some literature is concerned with *backdoor attacks*, where a trojan model forces its host system to misbehave when certain predefined conditions, named "triggers", are present but function normally otherwise. Xi et al. [59] investigate the vulnerability of DNNs to backdoor attacks. They propose GTA[5] (Graph Trojaning Attack), a graph attack framework that can be applied readily without knowing data models or tuning strategies to optimize both attack effectiveness and evasiveness. Triggers are defined as specific subgraphs, including topological structures and descriptive features, entailing a large design spectrum for the adversary. Rather

[4] RL-S2V code and data are available at https://github.com/Hanjun-Dai/graph_adversarial_attack.

[5] GTA code is available at https://github.com/HarrialX/GraphBackdoor.

than defining a fixed trigger for all the graphs, it generates triggers tailored to the characteristics of individual graphs. The framework can be instantiated for both transductive and inductive tasks, such as node classification and graph classification, respectively. Another backdoor attack for the task of graph classification is proposed by Zhang et al. in [64].[6] They propose a subgraph-based backdoor attack to GNNs such that the GNN classifier, trained on attacked graphs, predicts an attacker-chosen target label for a testing graph once the same trigger is injected into it.

Since strong attacks are often intractable to compute, an alternative approach that has recently attracted considerable interest is to construct *certificates of robustness*, i.e., sufficient conditions which, once satisfied, guarantee the immunity to any admissible attack [26]. A robustness certificate for graph classification using GCNs under structural attacks is proposed by Jin et al. [26].[7] Given a trained GCN and a threat model with both local and global budgets (number of edges that can be removed for each node or across all nodes, respectively), their certificate can efficiently verify that no topological perturbation can change the graph prediction.

Rather than adding or deleting edges to graphs, the *rewiring* operation has also been considered [32]. Given three nodes, two of which are connected, it consists in removing the edge between the two connected nodes and adding an edge between one of them and the third node. The rationale is that this operation preserves some basic properties of the attacked graph (e.g., number of nodes and edges, total degree of the graph), thus resulting in less noticeable attacks as compared to adding/deleting edges. Based on rewiring, Ma et al. [32] propose ReWatt,[8] an untargeted black-box attack for the task of graph classification, where the series of rewiring operations is described through a discrete Markov decision process and reinforcement learning is adopted to learn how to make these decisions. However, as highlighted by the authors, rewiring may lead to abnormal structure of the graphs, which could make them invalid for some applications, such as modeling of chemical molecules.

All the above-described approaches propose different types of adversarial attacks for graph-level tasks, mainly in the context of GNNs. However, none of them shares our aim, i.e., to compare the robustness to adversarial attacks of different whole-graph embedding methods.

[6] The code for the backdoor attack to GNNs in [64] is available at https://github.com/zaixizhang/graphbackdoor.

[7] Code and data for [26] are available at https://github.com/RobustGraph/RoboGraph.

[8] ReWatt code is available at https://github.com/alge24/ReWatt.

4 Adversarial Attacks to Whole-Graph Embedding Methods for Graph Classification

In [36], we considered two node-based attacks, consisting in removing from each graph a certain amount of nodes, chosen according to different criteria. Since some of the considered datasets (namely, PROTEINS and MUTAG, see Sect. 5.1) have a tiny number of nodes, removing few highest centrality nodes could almost destroy the graph structures (and their meaning). Therefore, in this work, we consider two edge-based attack strategies

- *Random Edge-based Attack* (REA): a percentage p of graph edges randomly chosen is removed from each graph. It can be considered a local attack and is performed with a low computational overhead.
- *Edge Betweenness Centrality-based Attack* (EBA): a percentage p of graph edges having the highest edge betweenness centrality is removed from each graph. The betweenness centrality b_e for an edge $e \in E$ is defined as the sum of the fraction of all-pairs shortest paths that pass through it [5]

$$b_e = \sum_{i,j=1}^{N} \frac{\sigma(v_i, v_j|e)}{\sigma(v_i, v_j)},$$

  where $\sigma(v_i, v_j)$ is the number of shortest paths from node v_i to node v_j and $\sigma(v_i, v_j|e)$ is the number of those paths that pass through the target edge e. It is a global attack strategy, as the path information is aggregated from the whole network. Clearly, global information carries significant computational overhead compared to local attacks.

According to the attack taxonomies presented in Sect. 3.1, the adversarial attacks that we consider are (1) *edge-level perturbations*, consisting in the removal of edges in the input graphs within the budget p; (2) *poisoning* attacks, modifying graphs used both for training and for testing; (3) *black-box* attacks, as these attacks do not need any information on the embedding model to attack; (4) *untargeted* attacks, as no specific label is enforced to the attacked data.

The attack task that we focus on is *graph classification*, which is the less frequently considered task in the adversarial attack literature. Indeed, we aim to evaluate the robustness of whole-graph embedding methods to adversarial attacks in terms of their performance in classifying the attacked graphs. To this end, we consider an unsupervised representation setting [60], where the graph embeddings generated by the compared unsupervised methods are fed into a down-stream classifier (see Algorithm 1 in Sect. 5.2).

Table 1 Main properties of the adopted datasets

Property	Kidney	MUTAG	Proteins
# graphs	299	188	1113
# classes	3	2	2
# samples per class	159/90/50	125/63	663/450
Average # nodes	1034	17.93	39.06
Average # edges	3226.00	19.79	72.82
Average edge density	0.006	0.138	0.212
# distinct node labels	1034	7	3
Edge weights	✓	✗	✗
Minimum diameter	7	5	1
Maximum diameter	7	15	54
Average degree	6.24	2.19	3.73

5 Experimental Results

5.1 Datasets

For the experiments, we adopted different life sciences datasets, including both weighted and non-weighted graphs, with both few and many nodes as well edges (see details in Table 1).

The *Kidney* dataset describes real metabolic networks created for validating related research [18, 19, 34]. It contains networks derived from data of 299 patients divided into three classes: 159 clear cell Renal Cell Carcinoma (KIRC), 90 Papillary Renal Cell Carcinoma (KIRP), and 50 Solid Tissue Normal samples. We obtained the networks by mapping gene expression data coming from the Genomic Data Commons (GDC, https://portal.gdc.cancer.gov) portal (Projects TCGA-KIRC and TCGA-KIRP) on the biochemical reactions extracted from the kidney tissue metabolic model [55] (https://metabolicatlas.org). Graph nodes represent the metabolites, and the edges connect reagent and product metabolites in the same reaction, weighted by the average of the expression values of the genes/enzymes catalyzing that reaction [18]. The simplification procedure described in [19] is applied to reduce the complexity of the networks, leading to reduce the number of nodes from 4022 to 1034.

MUTAG [12] is a small molecule benchmark dataset and is composed of networks of 188 mutagenic aromatic and heteroaromatic nitro compounds. The nodes represent atoms, while the edges represent chemical bonds between them. The graphs contain both node and edge labels. The two classes indicate whether or not the compound has mutagenic effects on a bacterium.

PROTEINS [4] is a bioinformatics dataset consisting of 1113 graphs that represent protein molecules. The nodes represent secondary structure elements of three distinct types (helix, sheet, or turn), and there is an edge between two nodes if they are neighbors in the amino-acid sequence or in the 3D space. The two classes indicate whether the graphs represent enzyme or non-enzyme proteins.

5.2 Implementation Details

The experimental test bench we chose for investigating the robustness of whole-graph embedding methods under adversarial attacks consists of methods selected from the neural network- and statistical/spectral representation-based approaches (see Sect. 2.1). In particular, six of the benchmarking methods (namely, GL2vec, Graph2vec, IGE, NetLSD, FGSD, and FeatherGraph) are implementations provided by the Karate Club software [44], which is a Python library consisting of state-of-the-art methods to do unsupervised learning on graph structured data. A recent whole-graph embedding method, named Netpro2vec [35], is included in our experimental bench of methods. This method is a neural network-based graph embedding technique implemented in Python and publicly available.[9]

The experimental pipeline is summarized in the pseudo-code of Algorithm 1. With the exception of IGE, FeatherGraph, and NetLSD, the embedding methods

Algorithm 1 The experimental pipeline

Require: graph $dataset$, graph labels y
Ensure: embedding validation scores
$graphs, y \leftarrow Load(dataset)$
$graphs \leftarrow Attack(graphs)$ ▷ apply attack strategy on graphs
$WGEmbedder.init()$ ▷ init and set parameters for embedder
$WGEmbedder.fit(graphs)$ ▷ build the model (unsupervised)
$X \leftarrow WGEmbedder.get\_embedding()$ ▷ embed the graphs
$X \leftarrow MinMaxScaler(X)$ ▷ embedding normalization
$scores \leftarrow Cross\_Validate(X, y)$ ▷ 10-fold stratified CV

allow to select the dimension d of the produced embedding vectors for graphs, set to 512 in all the experiments. GL2Vec and Graph2Vec were applied with a recursion depth of 5, while the remaining parameters were set to the library defaults. Netpro2Vec method has been applied by jointly exploiting NDD and TM1 distribution information of graphs, with no threshold cutoff; in addition, its skipgram model was trained for one hundred epochs, while the remaining parameters were set to the default. All values for neural network-based methods have been experimentally chosen to maximize accuracy. Embedding results were normalized by min-max scaling.

The validation step was carried out through ten-fold stratified cross-validation: at each cross-validation iteration, 90% of the embedding matrix X was used to train an SVM with linear kernel, while all other parameters were set to the defaults, as specified by the chosen SVM implementation which is available in the Scikit-

[9] Netpro2vec is available at https://github.com/cds-group/Netpro2vec; the implementation also includes the code for extracting the NDD and TM distribution matrices, based on the GraphDistances R package [20], and the doc2vec embedding is performed using the gensim NLP library [43].

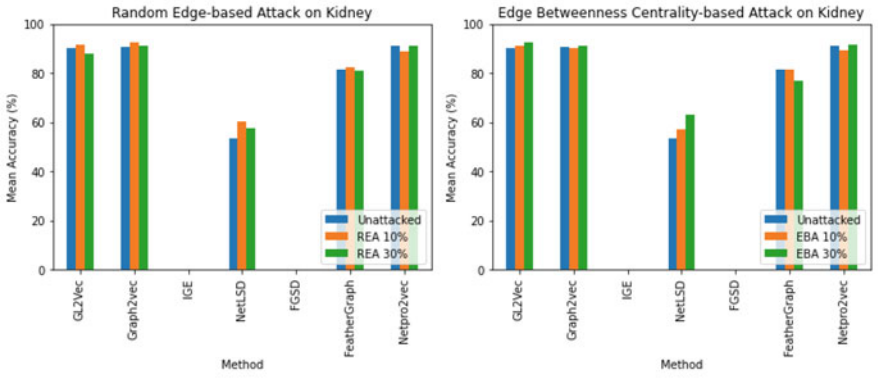
Learn Machine Learning library (https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html); the remaining 10% of X was used for testing prediction of the SVM model built on the embedding representation of graphs. Partitions of embedding matrix X are non-overlapping, thus ensuring that all the vectors of X are used (only once) for testing in one validation iteration. All the experiments were run on Google Colab hardware which provided a virtual machine based on a bi-processor with two Intel(R) Xeon(R) CPU @ 2.20 GHz with 13 GB RAM and 108 GB HDD.

5.3 *Performance Results*

Performance results obtained using the seven whole-graph embedding methods described in Sect. 5.2 on the three datasets detailed in Sect. 5.1, under the adversarial attacks described in Sect. 4, are reported in the bar plots of Figs. 1, 2, and 3, while detailed numerical results are given in the Appendix (Tables 2, 3, 4, 5, 6, and 7).

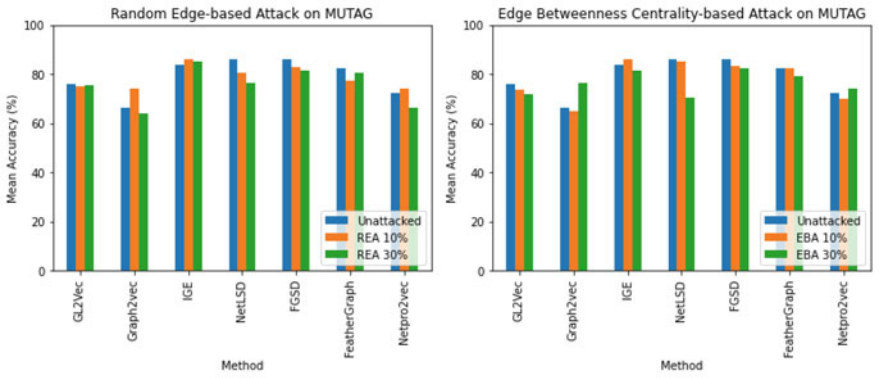
Here, we consider the results achieved using the original network data (Unattacked), as well as those using data that underwent the removal of the 10 and 30% of the edges having the highest betweenness centrality (EBA) or the randomly chosen (RAE), respectively. The choice of these percentages p of removed edges aims at investigating the effects of both moderate (10%) and strong (30%) adversarial attacks.

The performance is evaluated in terms of the mean and standard deviation across the CV iterations of Accuracy and Matthews Correlation Coefficient (MCC) [38]. As well known, Accuracy provides the percentage of correctly classified samples and assumes values in [0,1], reaching 1 for perfect classification. MCC gives the



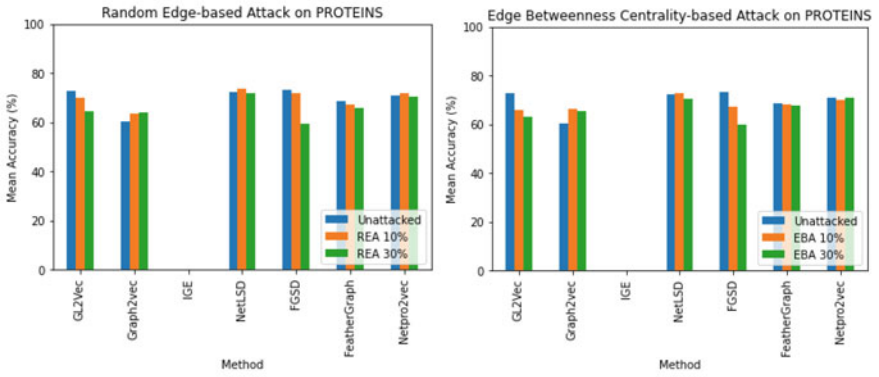
(a) (b)

Fig. 1 Plots of accuracy measures on Kidney dataset for (**a**) Random edge-based attack and (**b**) Edge betweenness centrality-based attack



(a) (b)

Fig. 2 Plots of accuracy measures on MUTAG dataset for (**a**) Random edge-based attack and (**b**) Edge betweenness centrality-based attack



(a) (b)

Fig. 3 Plots of accuracy measures on PROTEINS dataset for (**a**) Random edge-based attack and (**b**) Edge betweenness centrality-based attack

correlation coefficient between observed and predicted binary classifications, taking into account class balancing. It assumes values in $[-1, 1]$, where 1 indicates perfect classification. The value -1 is asymptotically reached in the extreme misclassification case of a confusion matrix with all zeros but in two symmetric entries, while MCC is equal to 0 when the confusion matrix contains all zeros but for one column (i.e., all samples have been classified to be of one class). Missing values in Figs. 1 and 3 (as well as in Tables 2, 3, 6, and 7) indicate non-convergence of the related method on Kidney and PROTEINS datasets.

From the analysis of the plots in Figs. 1, 2, and 3, we can observe that, as the attack strength grows (from 0 to 30%), FGSD is the only method that always shows a performance degradation regardless of the attack type (both REA

and EBA) and the dataset (even though nothing can be said on Kidney dataset, where it fails to converge). GL2Vec also degrades performance on MUTAG and PROTEINS datasets in the case of EBA, while on Kidney dataset, its performance remains substantially unaltered. Feathergraph experiences performance degradation on MUTAG and Kidney datasets under edge betweenness centrality-based attacks. Netpro2vec proved to have fairly constant performance across all datasets and attack types, thus showing significant robustness in all the case studies. The remaining methods have fluctuating trends, thus showing dropping or stable performance depending on the specific dataset and attack strategy.

It should be observed that some of the compared methods improve their performance under increasing strength of the attacks. Specifically, NetLSD shows this behavior only on Kidney dataset; the same can be said for Graph2Vec on MUTAG and PROTEINS datasets, while in the Kidney case study, its performance stays unaltered under both types of attack. This probably means that adversarial learning, i.e., the injection of adversarial examples into the training set such that the trained model can correctly classify the future adversarial examples [9, 27, 50], in these cases succeeds in enhancing the robustness of the methods.

6 Conclusions and Future Work

In this paper, we evaluated adversarial attacks to whole-graph embedding methods for the graph classification task. We considered methods from different classes, attacked by varying edge-level attacks. We mainly focused on networks of interest for life sciences, covering metabolic, mutagenic, and molecular networks. Our analysis has shown that some of the compared methods are intrinsically quite robust to the considered attacks, showing substantially unchanged performance under different strengths and types of edge-level attacks. It also highlighted that, in some cases, some of the methods improved their performance under attack, suggesting that adversarial learning could lead to higher robustness.

Even though we considered datasets showing a variety of graph properties, further work is needed to extend the analysis to medium- and large-scale datasets, such as the Benchmarking Graph Neural Networks datasets [16] and the Open Graph Benchmark [24]. Moreover, an analogous evaluation for inductive, rather than transductive, embedding methods, eventually coupled with suitable defense mechanisms, could shed more light on the issue.

Acknowledgments This work has been partially funded by the BiBiNet project (H35F21000430002) within POR-Lazio FESR 2014–2020. It was carried out also within the activities of the authors as members of the ICAR-CNR INdAM Research Unit and partially supported by the INdAM research project “Computational Intelligence methods for Digital Health”. The work of Mario R. Guarracino was conducted within the framework of the Basic Research Program at the National Research University Higher School of Economics (HSE). Mario Manzo thanks Prof. Alfredo Petrosino for the guidance and supervision during the years of working together.

Appendix

Tables 2, 3, 4, 5, 6, and 7 report detailed numerical results in terms of the mean and standard deviation across the CV iterations of Accuracy and MCC values achieved by the seven whole-graph embedding methods described in Sect. 5.2 on the three datasets detailed in Sect. 5.1, under the adversarial attacks described in Sect. 4.

Table 2 Accuracy (%) and MCC (mean ± std) of whole-graph embedding models on kidney dataset under edge betweenness centrality-based attack (EBA)

	Accuracy			MCC		
Method	Unattacked	EBA 10%	EBA 30%	Unattacked	EBA 10%	EBA 30%
GL2Vec	90.09±4.74	91.32±4.75	92.31±3.95	0.83±0.08	0.85±0.07	0.87±0.06
Graph2vec	90.79±5.11	90.32±6.89	91.32±3.70	0.83±0.08	0.84±0.10	0.85±0.06
IGE	–	–	–	–	–	–
NetLSD	53.46±7.02	57.16±9.20	63.21±3.67	0.11±0.16	0.22±0.18	0.40±0.07
FGSD	–	–	–	–	–	–
FeatherGraph	81.51±7.96	81.26±5.01	76.91±6.63	0.68±0.13	0.68±0.08	0.40±0.07
Netpro2vec	91.27±4.45	89.31±3.86	91.64±3.72	0.86±0.07	0.82±0.06	0.86±0.05

Table 3 Accuracy (%) and MCC (mean ± std) of whole-graph embedding models on kidney dataset under random edge-based attack (REA)

	Accuracy			MCC		
Method	Unattacked	REA 10%	REA 30%	Unattacked	REA 10%	REA 30%
GL2Vec	90.09±4.74	91.66±4.76	87.99±5.40	0.83±0.08	0.86±0.07	0.80±0.08
Graph2vec	90.79±5.11	92.66±3.88	91.32±4.75	0.83±0.08	0.87±0.06	0.85±0.07
IGE	–	–	–	–	–	–
NetLSD	53.46±7.02	60.48±8.72	57.55±6.35	0.11±0.16	0.31±0.15	0.26±0.13
FGSD	–	–	–	–	–	–
FeatherGraph	81.51±7.96	82.60±4.71	80.92±8.35	0.68±0.13	0.70±0.07	0.68±0.13
Netpro2vec	91.27±4.45	88.97±4.48	90.98±4.95	0.86±0.07	0.81±0.07	0.85±0.08

Table 4 Accuracy (%) and MCC (mean ± std) of whole-graph embedding models on MUTAG dataset under edge betweenness centrality-based attack (EBA)

	Accuracy			MCC		
Method	Unattacked	EBA 10%	EBA 30%	Unattacked	EBA 10%	EBA 30%
GL2Vec	76.11±8.48	73.77±11.29	71.73±9.13	0.31±0.24	0.37±0.29	0.31±0.24
Graph2vec	66.32±9.72	64.88±11.11	76.52±9.41	0.15±0.24	0.18±0.24	0.45±0.23
IGE	83.72±7.92	86.14±6.38	81.32±6.59	0.61±0.16	0.70±0.14	0.58±0.15
NetLSD	86.23±7.68	85.00±6.54	70.64±7.64	0.69±0.16	0.66±0.15	0.26±0.24
FGSD	86.01±7.77	83.54±6.79	82.43±8.55	0.70±0.16	0.65±0.15	0.65±0.13
FeatherGraph	82.40±8.24	82.37±8.94	79.24±12.38	0.60±0.17	0.61±0.23	0.53±0.25
Netpro2vec	72.06±9.64	70.20±8.57	73.98±9.86	0.60±0.17	0.30±0.19	0.43±0.21

Table 5 Accuracy (%) and MCC (mean ± std) of whole-graph embedding models on MUTAG dataset under random edge-based attack (REA)

	Accuracy			MCC		
Method	Unattacked	REA 10%	REA 30%	Unattacked	REA 10%	REA 30%
GL2Vec	76.11±8.48	75.03±5.21	75.50±8.64	0.31±0.24	0.44±0.13	0.43±0.20
Graph2vec	66.32±9.72	73.95±5.43	63.86±12.51	0.15±0.24	0.40±0.11	0.16±0.27
IGE	83.72±7.92	86.14±5.93	85.06±5.83	0.61±0.16	0.70±0.11	0.67±0.12
NetLSD	86.23±7.68	80.79±11.22	76.49±10.74	0.69±0.16	0.56±0.24	0.43±0.30
FGSD	86.01±7.77	82.87±6.65	81.29±5.83	0.70±0.16	0.63±0.16	0.59±0.13
FeatherGraph	82.40±8.24	77.54±10.04	80.35±7.04	0.60±0.17	0.53±0.23	0.56±0.14
Netpro2vec	72.06±9.64	73.92±7.65	66.5±7.63	0.60±0.17	0.39±0.19	0.22±0.19

Table 6 Accuracy (%) and MCC (mean ± std) of whole-graph embedding models on PROTEINS dataset under edge betweenness centrality-based attack (EBA)

	Accuracy			MCC		
Method	Unattacked	EBA 10%	EBA 30%	Unattacked	EBA 10%	EBA 30%
GL2Vec	72.87±3.44	65.85±3.87	62.90±3.42	0.43±0.08	0.25±0.09	0.18±0.08
Graph2vec	60.28±1.20	66.22±3.16	65.32±2.74	0.07±0.08	0.26±0.07	0.24±0.07
IGE	–	–	–	–	–	–
NetLSD	72.50±3.41	72.78±3.13	70.44±3.58	0.41±0.07	0.42±0.07	0.38±0.09
FGSD	73.23±2.79	67.12±2.57	59.66±0.28	0.44±0.06	0.33±0.05	0.01±0.03
FeatherGraph	68.73±2.08	68.10±2.94	67.83±2.10	0.33±0.05	0.31±0.07	0.30±0.04
Netpro2vec	70.89±3.17	69.90±3.39	70.98±1.90	0.40±0.06	0.38±0.06	0.40±0.03

Table 7 Accuracy (%) and MCC (mean ± std) of whole-graph embedding models on PROTEINS dataset under random edge-based attack (REA)

	Accuracy			MCC		
Method	Unattacked	REA 10%	REA 30%	Unattacked	REA 10%	REA 30%
GL2Vec	72.87±3.44	69.99±4.33	64.51±2.13	0.43±0.08	0.36±0.11	0.23±0.06
Graph2vec	60.28±1.20	63.71±3.57	63.97±5.01	0.07±0.08	0.21±0.08	0.23±0.10
IGE	–	–	–	–	–	–
NetLSD	72.50±3.41	73.67±3.47	71.61±4.14	0.41±0.07	0.44±0.07	0.39±0.09
FGSD	73.23±2.79	71.79±2.92	59.57±0.44	0.44±0.06	0.40±0.06	0.00±0.04
FeatherGraph	68.73±2.08	67.11±2.92	65.68±2.67	0.33±0.05	0.29±0.07	0.25±0.07
Netpro2vec	70.89±3.17	71.60±4.62	70.62±3.24	0.40±0.06	0.41±0.09	0.38±0.06

References

1. Lu Bai, Lixin Cui, Luca Rossi, Lixiang Xu, Xiao Bai, and Edwin Hancock. Local-global nested graph kernels using nested complexity traces. *Pattern Recognition Letters*, 134:87–95, 2020.
2. Yunsheng Bai, Hao Ding, Yang Qiao, Agustin Marinovic, Ken Gu, Ting Chen, Yizhou Sun, and Wei Wang. Unsupervised inductive graph-level representation learning via graph-graph proximity, 2019.
3. Yunsheng Bai, Hao Ding, Yang Qiao, Agustin Marinovic, Ken Gu, Ting Chen, Yizhou Sun, and Wei Wang. Unsupervised inductive graph-level representation learning via graph-graph proximity. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, IJCAI'19, page 1988–1994. AAAI Press, 2019.
4. Karsten M. Borgwardt, Cheng Soon Ong, Stefan Schönauer, S. V. N. Vishwanathan, Alex J. Smola, and Hans-Peter Kriegel. Protein function prediction via graph kernels. *Bioinformatics*, 21(suppl 1):i47–i56, 06 2005.
5. Ulrik Brandes. On variants of shortest-path betweenness centrality and their generic computation. *Social Networks*, 30(2):136–145, 2008.
6. Hongyun Cai, Vincent W Zheng, and Kevin Chen-Chuan Chang. A comprehensive survey of graph embedding: Problems, techniques, and applications. *IEEE Transactions on Knowledge and Data Engineering*, 30(9):1616–1637, 2018.
7. Hong Chen and Hisashi Koga. Gl2vec: Graph embedding enriched by line graphs with edge features. In *Int. Conf. on Neural Information Processing*, pages 3–14. Springer, 2019.
8. Jinyin Chen, Dunjie Zhang, Zhaoyan Ming, and Kejie Huang. GraphAttacker: A general multi-task graphattack framework. *arXiv preprint arXiv:2101.06855*, 2021.
9. Liang Chen, Jintang Li, Jiaying Peng, Tao Xie, Zengxu Cao, Kun Xu, Xiangnan He, and Zibin Zheng. A survey of adversarial learning on graphs. *CoRR*, abs/2003.05730, 2020.
10. Noé Cécillon, Vincent Labatut, Richard Dufour, and Georges Linarès. Graph embeddings for abusive language detection. *SN Computer Science*, 2(1), Jan 2021.
11. Hanjun Dai, Hui Li, Tian Tian, Xin Huang, Lin Wang, Jun Zhu, and Le Song. Adversarial attack on graph structured data. In *International conference on machine learning*, pages 1115–1124. PMLR, 2018.
12. AK Debnath, RL Lopez de Compadre, G Debnath, AJ Shusterman, and C. Hansch. Structure-activity relationship of mutagenic aromatic and heteroaromatic nitro compounds. correlation with molecular orbital energies and hydrophobicity. *J Med Chem*, 34, 1991.
13. Fernando M Delgado and Francisco Gómez-Vela. Computational methods for gene regulatory networks reconstruction and analysis: A review. *Artif. Intell. Med.*, 95:133–145, 2019.
14. Anjan Dutta, Pau Riba, Josep Lladós, and Alicia Fornés. Pyramidal stochastic graphlet embedding for document pattern classification. In *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, volume 1, pages 33–38. IEEE, 2017.
15. Anjan Dutta, Pau Riba, Josep Lladós, and Alicia Fornés. Hierarchical stochastic graphlet embedding for graph-based pattern recognition. *Neural Computing and Applications*, 32(15):11579–11596, 2020.
16. Vijay Prakash Dwivedi, Chaitanya K. Joshi, Thomas Laurent, Yoshua Bengio, and Xavier Bresson. Benchmarking graph neural networks, 2020.
17. Alexis Galland and Marc Lelarge. Invariant embedding for graph classification. In *ICML 2019 Workshop on Learning and Reasoning with Graph-Structured Representations*, 2019.
18. Ilaria Granata, Mario R Guarracino, Valery A Kalyagin, Lucia Maddalena, Ichcha Manipur, and Panos M Pardalos. Supervised classification of metabolic networks. In *2018 IEEE Int. Conf. on Bioinformatics and Biomedicine (BIBM)*, pages 2688–2693. IEEE, 2018.
19. Ilaria Granata, Mario R Guarracino, Valery A Kalyagin, Lucia Maddalena, Ichcha Manipur, and Panos M Pardalos. Model simplification for supervised classification of metabolic networks. *Ann. Math. Artif. Intell.*, 88(1):91–104, 2020.
20. Ilaria Granata, Mario Rosario Guarracino, Lucia Maddalena, and Ichcha Manipur. Network distances for weighted digraphs. In Yury Kochetov, Igor Bykadorov, and Tatiana Gruzdeva,

editors, *Mathematical Optimization Theory and Operations Research*, pages 389–408. Springer Int. Publishing, 2020.
21. Ilaria Granata, Mario Rosario Guarracino, Lucia Maddalena, Ichcha Manipur, and Panos M. Pardalos. On network similarities and their applications. In Rubem P. Mondaini, editor, *Trends in Biomathematics: Modeling Cells, Flows, Epidemics, and the Environment: Selected Works from the BIOMAT Consortium Lectures, Szeged, Hungary, 2019*, pages 23–41, Cham, 2020. Springer International Publishing.
22. Ilaria Granata, Mario Manzo, Ari Kusumastuti, and Mario Rosario Guarracino. Learning from metabolic networks: Current trends and future directions for precision medicine. *Current Medicinal Chemistry*, 28(32), 2021.
23. Maya Hirohara, Yutaka Saito, Yuki Koda, Kengo Sato, and Yasubumi Sakakibara. Convolutional neural network based on smiles representation of compounds for detecting chemical motif. *BMC bioinformatics*, 19(19):526, 2018.
24. Weihua Hu, Matthias Fey, Marinka Zitnik, Yuxiao Dong, Hongyu Ren, Bowen Liu, Michele Catasta, and Jure Leskovec. Open graph benchmark: Datasets for machine learning on graphs, 2021.
25. Jianming Huang, Zhongxi Fang, and Hiroyuki Kasai. LCS graph kernel based on Wasserstein distance in longest common subsequence metric space. *Signal Processing*, 189:108281, 2021.
26. Hongwei Jin, Zhan Shi, Venkata Jaya Shankar Ashish Peruri, and Xinhua Zhang. Certified robustness of graph convolution networks for graph classification under topological attacks. *Advances in Neural Information Processing Systems*, 33, 2020.
27. Wei Jin, Yaxing Li, Han Xu, Yiqi Wang, Shuiwang Ji, Charu Aggarwal, and Jiliang Tang. Adversarial attacks and defenses on graphs. *SIGKDD Explor. Newsl.*, 22(2):19–34, January 2021.
28. Nadeem Iqbal Kajla, Malik Muhammad Saad Missen, Muhammad Muzzamil Luqman, Mickael Coustaty, Arif Mehmood, and Gyu Sang Choi. Additive angular margin loss in deep graph neural network classifier for learning graph edit distance. *IEEE Access*, 8:201752–201761, 2020.
29. Quoc Le and Tomas Mikolov. Distributed representations of sentences and documents. In *Int. Conf. on machine learning*, pages 1188–1196, 2014.
30. Bentian Li and Dechang Pi. Network representation learning: a systematic literature review. *Neural Computing and Applications*, pages 1–33, 2020.
31. Yao Ma, Suhang Wang, Charu C Aggarwal, and Jiliang Tang. Graph convolutional networks with eigenpooling. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 723–731, 2019.
32. Yao Ma, Suhang Wang, Tyler Derr, Lingfei Wu, and Jiliang Tang. Graph adversarial attack via rewiring. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, KDD '21, page 1161–1169, New York, NY, USA, 2021. Association for Computing Machinery.
33. Lucia Maddalena, Ichcha Manipur, Mario Manzo, and Mario Rosario Guarracino. On whole-graph embedding techniques. In Rubem P. Mondaini, editor, *Trends in Biomathematics: Chaos and Control in Epidemics, Ecosystems, and Cells: Selected Works from the 20th BIOMAT Consortium Lectures, Rio de Janeiro, Brazil, 2020*, pages 115–131, Cham, 2021. Springer International Publishing.
34. Ichcha Manipur, Ilaria Granata, Lucia Maddalena, and Mario Rosario Guarracino. Clustering analysis of tumor metabolic networks. *BMC Bioinformatics*, 21(349), 2020.
35. Ichcha Manipur, Mario Manzo, Ilaria Granata, Maurizio Giordano, Lucia Maddalena, and Mario Rosario Guarracino. Netpro2vec: a graph embedding framework for biomedical applications. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2021.
36. Mario Manzo, Maurizio Giordano, Lucia Maddalena, and Mario Rosario Guarracino. Performance evaluation of adversarial attacks on whole-graph embedding models. In Dimitris E. Simos, Panos M. Pardalos, and Ilias S. Kotsireas Kotsireas, editors, *Learning and Intelligent Optimization*, LNCS. Springer, 2021.

37. Mario Manzo and Alessandro Rozza. Dopsie: Deep-order proximity and structural information embedding. *Machine Learning and Knowledge Extraction*, 1(2):684–697, 2019.
38. B.W. Matthews. Comparison of the predicted and observed secondary structure of t4 phage lysozyme. *Biochimica et Biophysica Acta (BBA) - Protein Structure*, 405(2):442–451, 1975.
39. Annamalai Narayanan, Mahinthan Chandramohan, Rajasekar Venkatesan, Lihui Chen, Yang Liu, and Shantanu Jaiswal. graph2vec: Learning distributed representations of graphs. In *Proceedings of the 13th International Workshop on Mining and Learning with Graphs (MLG)*, 2017.
40. Walter Nelson, Marinka Zitnik, Bo Wang, Jure Leskovec, Anna Goldenberg, and Roded Sharan. To embed or not: network embedding as a paradigm in computational biology. *Front Genet*, 10:381, 2019.
41. Mathias Niepert, Mohamed Ahmed, and Konstantin Kutzkov. Learning convolutional neural networks for graphs. In *International conference on machine learning*, pages 2014–2023, 2016.
42. Saeid Rasti and Chrysafis Vogiatzis. A survey of computational methods in protein–protein interaction networks. *Ann. Oper. Res.*, 276(1–2):35–87, 2019.
43. Radim Řehůřek and Petr Sojka. Software Framework for Topic Modelling with Large Corpora. In *Proc. of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, pages 45–50, Valletta, Malta, May 2010. ELRA.
44. Benedek Rozemberczki, Oliver Kiss, and Rik Sarkar. Karate Club: An API Oriented Open-source Python Framework for Unsupervised Learning on Graphs. In *Proc. of the 29th ACM Int. Conf. on Information and Knowledge Management (CIKM '20)*. ACM, 2020.
45. Benedek Rozemberczki and Rik Sarkar. Characteristic functions on graphs: Birds of a feather, from statistical descriptors to parametric models. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, pages 1325–1334, 2020.
46. Anwar Said, Saeed-Ul Hassan, Waseem Abbas, and Mudassir Shabbir. NetKI: A Kirchhoff index based statistical graph embedding in nearly linear time. *Neurocomputing*, 433:108–118, 2021.
47. Alberto Sanfeliu and King-Sun Fu. A distance measure between attributed relational graphs for pattern recognition. *IEEE transactions on systems, man, and cybernetics*, 3:353–362, 1983.
48. CJ Stam, ECW Van Straaten, E Van Dellen, P Tewarie, G Gong, A Hillebrand, J Meier, and P Van Mieghem. The relation between structural and functional connectivity patterns in complex brain networks. *Int. J. Psychophysiol.*, 103:149–160, 2016.
49. Chang Su, Jie Tong, Yongjun Zhu, Peng Cui, and Fei Wang. Network embedding in biomedical data science. *Briefings in Bioinformatics*, 21(1):182–197, 12 2018.
50. Lichao Sun, Ji Wang, Philip S. Yu, and Bo Li. Adversarial attack and defense on graph data: A survey. *CoRR*, abs/1812.10528, 2020.
51. Zhigang Sun, Hongwei Huo, Jun Huan, and Jeffrey Scott Vitter. Feature reduction based on semantic similarity for graph classification. *Neurocomputing*, 397:114–126, 2020.
52. Haoteng Tang, Guixiang Ma, Yurong Chen, Lei Guo, Wei Wang, Bo Zeng, and Liang Zhan. Adversarial attack on hierarchical graph pooling neural networks. *arXiv preprint arXiv:2005.11560*, 2020.
53. Haoteng Tang, Guixiang Ma, Lifang He, Heng Huang, and Liang Zhan. Commpool: An interpretable graph pooling framework for hierarchical graph representation learning. *Neural Networks*, 143:669–677, 2021.
54. Anton Tsitsulin, Davide Mottin, Panagiotis Karras, Alexander Bronstein, and Emmanuel Müller. NetLSD: hearing the shape of a graph. In *Proc. of the 24th ACM SIGKDD Int. Conf. on Knowledge Discovery & Data Mining*, pages 2347–2356, 2018.
55. Mathias Uhlén, Linn Fagerberg, Björn M Hallström, Cecilia Lindskog, Per Oksvold, Adil Mardinoglu, Åsa Sivertsson, Caroline Kampf, Evelina Sjöstedt, Anna Asplund, et al. Tissue-based map of the human proteome. *Science*, 347(6220), 2015.
56. Saurabh Verma and Zhi-Li Zhang. Hunt for the unique, stable, sparse and fast feature learning on graphs. In *Adv Neural Inform Process Syst*, pages 88–98, 2017.

57. Juan Salamanca Viloria, Maria Francesca Allega, Matteo Lambrughi, and Elena Papaleo. An optimal distance cutoff for contact-based protein structure networks using side-chain centers of mass. *Scientific reports*, 7(1):1–11, 2017.
58. Yadi Wang, Xiaoping Li, and Rubén Ruiz. Weighted general group lasso for gene selection in cancer classification. *IEEE Trans Cybern*, 49(8):2860–2873, 2018.
59. Zhaohan Xi, Ren Pang, Shouling Ji, and Ting Wang. Graph backdoor. In *30th USENIX Security Symposium (USENIX Security 21)*, 2021.
60. Yuning You, Tianlong Chen, Yongduo Sui, Ting Chen, Zhangyang Wang, and Yang Shen. Graph contrastive learning with augmentations. In H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 5812–5823. Curran Associates, Inc., 2020.
61. Xiang Yue, Zhen Wang, Jingong Huang, Srinivasan Parthasarathy, Soheil Moosavinasab, Yungui Huang, Simon M Lin, Wen Zhang, Ping Zhang, and Huan Sun. Graph embedding on biomedical networks: methods, applications and evaluations. *Bioinformatics*, 36(4):1241–1251, 2020.
62. Muhan Zhang, Zhicheng Cui, Marion Neumann, and Yixin Chen. An end-to-end deep learning architecture for graph classification. In *AAAI*, pages 4438–4445, 2018.
63. Wen Zhang, Xiang Yue, Guifeng Tang, Wenjian Wu, Feng Huang, and Xining Zhang. Sfpel-lpi: Sequence-based feature projection ensemble learning for predicting lncrna-protein interactions. *PLoS computational biology*, 14(12):e1006616, 2018.
64. Zaixi Zhang, Jinyuan Jia, Binghui Wang, and Neil Zhenqiang Gong. Backdoor attacks to graph neural networks. In *Proceedings of the 26th ACM Symposium on Access Control Models and Technologies*, pages 15–26, 2021.
65. Zhen Zhang, Jiajun Bu, Martin Ester, Jianfeng Zhang, Chengwei Yao, Zhi Yu, and Can Wang. Hierarchical graph pooling with structure learning. *CoRR*, abs/1911.05954, 2019.
66. Jiajun Zhou, Jie Shen, Shanqing Yu, Guanrong Chen, and Qi Xuan. M-evolve: Structural-mapping-based data augmentation for graph classification. *IEEE Transactions on Network Science and Engineering*, 8(1):190–200, 2020.
67. Renyi Zhou, Zhangli Lu, Huimin Luo, Ju Xiang, Min Zeng, and Min Li. Nedd: a network embedding based method for predicting drug-disease associations. *BMC bioinformatics*, 21(13):1–12, 2020.
68. Shichao Zhu, Lewei Zhou, Shirui Pan, Chuan Zhou, Guiying Yan, and Bin Wang. GSSNN: graph smoothing splines neural networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 7007–7014. AAAI Press, 2020.

Photochemical Scheme for Reversibly Photoswitchable Fluorescent Proteins Used in Biological Imaging

A. Pellissier-Tanon, R. Chouket, T. Le Saux, L. Jullien, A. Espagne, and A. Lemarchand

1 Introduction

Living matter provides a variety of challenges. Cells contain a very large number (more than 10^6) of chemical components whose concentrations vary from 10^{-12} mol.L^{-1} to 10^{-1} mol.L^{-1} and which exhibit a well-defined heterogeneous spatial distribution. As a consequence, observing living cells requires sensitive imaging of a large number of components.

A crucial issue is to image one or several targeted components of the cell without interference from the other components. It is therefore necessary that the target exhibits a specific signature. However, the elemental composition and functional groups of the cell components are essentially the same. An exogenous chemical construct with a specific signature can be introduced by genetic encoding to label a targeted component[1–3]. The properties of the label can then be harnessed to improve selective imaging.

Due to the high sensitivity of fluorescence, fluorescent labels and in particular fluorescent proteins are widely exploited. Light is used to both probe and determine the amount of a fluorescent label. Among fluorescent proteins, reversibly photoswitchable fluorescent proteins (RSFPs) are of particular interest due to their complex photochemical kinetics[4]. RSFPs have been recently popularized

A. Pellissier-Tanon (✉) · R. Chouket · T. Le Saux · L. Jullien · A. Espagne
Pasteur, Département de chimie, École normale supérieure, PSL University, Sorbonne Université, CNRS, Paris, France
e-mail: agnes.pellissier-tanon@ens.fr

A. Lemarchand
Sorbonne Université, CNRS, Laboratoire de Physique Théorique de la Matière Condensée, Paris, France
e-mail: annie.lemarchand@sorbonne-universite.fr

R. P. Mondaini (ed.), *Trends in Biomathematics: Stability and Oscillations in Environmental, Social, and Biological Models*,
https://doi.org/10.1007/978-3-031-12515-7_2

by super-resolution microscopies[5–8]. They are also involved in the tracking of protein movements and interactions[9, 10], optical control of protein activity[11, 12], photoacoustic phenomena[13], and kinetic discrimination protocols[14–18]. A better knowledge of the complex kinetics of RSFPs is therefore desirable.

The detailed photoswitching mechanisms of RFSPs are not fully established. Based on the existing literature, we propose a seven-state photoswitching mechanism which accounts for the known steps of isomerization and protonation[19–23]. We probe the photochemical mechanism of RSFPs using illuminations at the wavelengths $\lambda_1 = 480$ nm and $\lambda_2 = 405$ nm. We focus on the evolution of fluorescence emission occurring in the 1 μs–1 s time window and detected at 525 nm. Our goal is to show that, depending on the light intensities, different kinetic regimes are reached. We aim to deduce different characteristic times from the fluorescence evolution associated with well-chosen experimental conditions. Starting from a reasonable photoswitching mechanism, we establish relationships between the experimental characteristic times and the rate constants of some reaction steps. The results are applied to the RSFP Dronpa-2 (Dronpa-M159T)[24] with the aim of providing unknown kinetic information on photoswitching.

2 Seven-State Photoswitching Mechanism in the 1 μs–1 s Time Window

Figure 1a gives the seven-state mechanism we adopt to describe the kinetics of RSFPs, in the 1 μs–1 s window. The Cis-O^- state, denoted A, absorbs light mainly at the wavelength $\lambda_1 = 480$ nm and to a lesser extent at the wavelength $\lambda_2 = 405$ nm forming the excited state A^* (with the rate constant k_1). The Cis-O^- state is recovered by internal conversion and fluorescence emission (rate constant k_{-1}) or by a longer reaction path described below. Isomerization of A^* (rate constant k_B) yields the Trans-O^- state, denoted B, which presumably possesses a similar relaxation rate of its excited state B^* as the Cis-O^- state. Protonation of B (rate constant k_C) leads to the Trans-OH state, denoted C. The excited state C^* of the Trans-OH state is assumed to be only formed by absorption at the wavelength $\lambda_2 = 405$ nm. The Cis-O^- state may be directly formed from Trans-OH by thermal isomerization (rate constant k_{CA}) or obtained by light excitation of Trans-OH (rate constant k_2), then isomerization yielding the Cis-OH state denoted D (rate constant k_D), and finally deprotonation (rate constant k_A)[19–23].

According to literature, the orders of magnitude of the rate constants involved in Fig. 1a are given in Table 1. In addition, we assume that the protonation rate constant k_C is of the same order of magnitude as the deprotonation rate constant k_A.

The seven-state photoswitching mechanism displayed in Fig. 1a possesses both photochemical steps associated with the rate constants k_1 and k_2 and thermal steps associated with all other rate constants. Therefore tuning light intensities may change the rate-limiting steps. We consider two different illumination protocols,

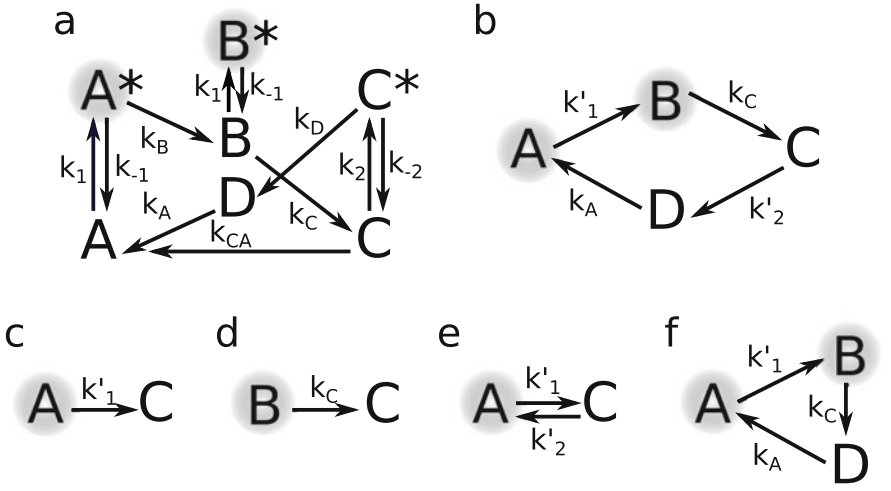


Fig. 1 Photochemical mechanisms of reversibly photoswitchable fluorescent proteins relevant in the 1 μs–1 s time window. (**a**): seven-state mechanism. (**b**): four-state mechanism obtained after elimination of the fast excited states A*, B*, and C*. (**c**): photochemical two-state mechanism in regime α_l. (**d**): Thermal two-state mechanism in regime α_h. (**e**): Photochemical two-state mechanism in regime β_l. (**f**): Photochemical and thermal three-state mechanism in regime β_h. Species with a halo are bright or acquire brightness due to mechanism reduction

α and β. The illumination protocol α is characterized by a light intensity I_1 at the wavelength $\lambda_1 = 480$ nm applied to the sample initially in the A state. The illumination protocol β is defined by a light intensity I_2 at the wavelength $\lambda_2 = 405$ nm and a small intensity I_1 at the wavelength $\lambda_1 = 480$ nm applied to the sample initially in the B state. In the protocol β, the complementary illumination at $\lambda_1 = 480$ nm is necessary to obtain a fluorescent signal since the excitation of the states A* and B* at the wavelength $\lambda_2 = 405$ nm does not generate a significant fluorescence emission.

The seven-state mechanism yields the following rate laws for the concentrations of the different species

$$\frac{\mathrm{d}A}{\mathrm{d}t} = -k_1 A + k_{-1} A^* + k_{CA} C + k_A D \tag{1}$$

$$\frac{\mathrm{d}A^*}{\mathrm{d}t} = k_1 A - (k_{-1} + k_B)\, A^* \tag{2}$$

$$\frac{\mathrm{d}B}{\mathrm{d}t} = k_B A^* - (k_1 + k_C)\, B + k_{-1} B^* \tag{3}$$

$$\frac{\mathrm{d}B^*}{\mathrm{d}t} = k_1 B - k_{-1} B^* \tag{4}$$

$$\frac{\mathrm{d}C}{\mathrm{d}t} = k_C B - (k_2 + k_{CA})\, C + k_{-2} C^* \tag{5}$$

Table 1 Order of magnitude of parameters of the photoisomerization mechanisms of RSFPs

	Parameter	Order of magnitude
Seven-state mechanism		
Absorption cross sections[19–23]	$\epsilon_{11}, \epsilon_{22}$	10^3–10^4 $m^2\ mol^{-1}$
	ϵ_{12}	10^2–10^3 $m^2\ mol^{-1}$
Relaxation[19–23]	k_{-1}	$10^9\ s^{-1}$
	k_{-2}	$10^{11}\ s^{-1}$
Isomerization[19–23]	k_B	$10^7\ s^{-1}$
	k_D	$10^{10}\ s^{-1}$
Proton exchange[19–23]	k_C, k_A	10^3–$10^6\ s^{-1}$
Thermal isomerization[19–23]	k_{CA}	10^{-4}–$10^{-2}\ s^{-1}$
Four-state mechanism		
Light intensities	I_1, I_2	$\ll 10^5$ Ein $m^{-2}\ s^{-1}$
Isomerization cross sections[15, 17]	σ_{11}, σ_{22}	10–10^2 $m^2\ mol^{-1}$
	σ_{12}	1–10 $m^2\ mol^{-1}$
Two-state mechanism α_l		
Light intensity	I_1	$\ll$ 10–10^2 Ein $m^{-2}\ s^{-1}$
Two-state mechanism α_h		
Light intensity	I_1	$\gg$ 10–10^5 Ein $m^{-2}\ s^{-1}$
Two-state mechanism β_l		
Light intensities	I_1	~ 0.1 Ein $m^{-2}\ s^{-1}$
	I_2	$\ll$ 10–10^2 Ein $m^2\ s^{-1}$
Three-state mechanism β_h		
Light intensities	I_1	~ 0.1 Ein $m^{-2}\ s^{-1}$
	I_2	$\gg$ 10–10^5 Ein $m^2\ s^{-1}$

$$\frac{dC^*}{dt} = k_2 C - (k_{-2} + k_D)\, C^* \tag{6}$$

$$\frac{dD}{dt} = k_D C^* - k_A D \tag{7}$$

The concentrations obey the law of matter conservation $C_{\text{tot}} = A + A^* + B + B^* + C + C^* + D$. The photochemical rate constants k_1 and k_2 are given by

$$k_1 = \epsilon_{11} I_1 + \epsilon_{12} I_2 \tag{8}$$

$$k_2 = \epsilon_{22} I_2 \tag{9}$$

where ϵ_{11} and ϵ_{12} are the absorption cross sections of A and B at the wavelengths $\lambda_1 = 480$ nm and $\lambda_2 = 405$ nm, respectively, and ϵ_{22} is the absorption cross section of C at the wavelength $\lambda_2 = 405$ nm. We assume that only illumination at $\lambda_1 = 480$ nm leads to fluorescence emission. The rate constant k_{-1} takes into account the radiative and non radiative de-excitation of the excited states A^* and B^*. We neglect the fluorescence of all species except A^* and B^* which leads to the following

expression of the fluorescence intensity

$$I_F = k^F_{A^*} A^* + k^F_{B^*} B^* \tag{10}$$

where $k^F_{A^*}$ and $k^F_{B^*}$ are the radiative parts of the rate constant k_{-1} associated with the de-excitation of A^* and B^*, respectively. Species B and B^* are the least known species of the mechanism. In particular the radiative de-excitation of B^* has not been evidenced. The hypothesis on the fluorescent nature of B^* needs to be validated.

We consider variable light intensities I_1 and I_2 in order to explore different kinetic regimes. We harness the different orders of magnitude of the rate constants and use the steady-state approximation to reduce the ordinary system of differential equations given in Eqs. (1)–(7) (see Appendix 1). The mechanism being composed of first-order steps, the condition of validity for the elimination of a short-lived chemical species is easy to express: The sum of the rate constants of the steps consuming the considered species must be larger than the sum of the rate constants of the steps producing the species. The fastest reactive species are first eliminated and the next timescale of interest is defined by the larger remaining characteristic times. We repeatedly eliminate the fastest species until the remaining characteristic times are close within the chosen 1 μs–1 s time range.

3 Elimination of the Excited States: The Four-State Mechanism

The rate constant $k_{CA} \leq 10^{-2}$ s^{-1} is associated with a slow kinetics inaccessible in the explored time window and therefore ignored in the following. According to Table 1, the fastest species are the excited states A^*, B^*, and C^*. Their elimination is valid if the conditions $k_1 \ll k_{-1} + k_B$, $k_1 \ll k_{-1}$, and $k_2 \ll k_{-2} + k_D$ are fulfilled, i.e. if the light intensities obey $I_1 \ll k_{-1}/\epsilon_{11}$ and $I_2 \ll k_{-2}/\epsilon_{22}$. Using Table 1, we find that light intensities smaller than 10^5 Ein.m^{-2}.s^{-1} meet these conditions. According to the steady-state approximation, we eliminate the species A^*, B^*, and C^* using $\frac{\mathrm{d}A^*}{\mathrm{d}t} = 0$, $\frac{\mathrm{d}B^*}{\mathrm{d}t} = 0$, and $\frac{\mathrm{d}C^*}{\mathrm{d}t} = 0$ where A^*, B^*, and C^* obey Eqs. (2), (4), and (6). The dynamics is then given by

$$\frac{\mathrm{d}A}{\mathrm{d}t} = -k'_1 A + k_A D \tag{11}$$

$$\frac{\mathrm{d}B}{\mathrm{d}t} = k'_1 A - k_C B \tag{12}$$

$$\frac{\mathrm{d}C}{\mathrm{d}t} = k_C B - k'_2 C \tag{13}$$

$$\frac{\mathrm{d}D}{\mathrm{d}t} = k'_2 C - k_A D \tag{14}$$

where $k'_1 = k_1 k_B/(k_{-1} + k_B)$ and $k'_2 = k_2 k_D/(k_{-2} + k_D)$. Following Appendix 1, we obtain the reduced four-state mechanism given in Fig. 1b with the apparent photochemical rate constants

$$k'_1 = \sigma_{11} I_1 + \sigma_{12} I_2 \tag{15}$$

$$k'_2 = \sigma_{22} I_2 \tag{16}$$

where $\sigma_{11} = \epsilon_{11} k_B/(k_{-1} + k_B)$ and $\sigma_{12} = \epsilon_{12} k_B/(k_{-1} + k_B)$ are the photoisomerization cross sections of A at the wavelengths $\lambda_1 = 480$ nm and $\lambda_2 = 405$ nm, respectively, and $\sigma_{22} = \epsilon_{22} k_D/(k_{-2} + k_D)$ is the photoisomerization cross section of C at the wavelength $\lambda_2 = 405$ nm.

The elimination of the fluorescent states A* and B* leads to the apparently fluorescent A and B species in the four-state mechanism. The fluorescence intensity is deduced from Eq. (10) leading to

$$I_F = (Q_A A + Q_B B)\, I_1 \tag{17}$$

where $Q_A = \epsilon_{11} k^F_{A^*}/(k_{-1} + k_B)$ and $Q_B = \epsilon_{11} k^F_{B^*}/k_{-1}$ are the apparent brightnesses of species A and B, respectively. It is to be noted that the nonvanishing brightness of species B is a direct consequence of the hypothesis on the fluorescent nature of B*.

The thermal rate constants k_A and k_C involved in the four-state mechanism have the same order of magnitude in the investigated time range. The values of the photochemical rate constants, k'_1 and k'_2, are determined by the light intensities I_1 and I_2. Consequently the light intensities control which species is the fastest and can be eliminated.

4 Illumination Protocol α

The illumination protocol α implies that $I_2 = 0$ which leads to $k'_2 = 0$. Hence species D and the rate constants k'_2 and k_A are not involved in the kinetics. The four-state mechanism reduces to a three-state mechanism. In addition, the initial condition is assumed to be $A = C_{\text{tot}}$. The three-state mechanism can be further reduced depending on the value of the light intensity I_1.

4.1 Low Intensity I_1: Regime α_l

We first consider the regime α_l of low light intensity obeying $I_1 \ll k_C/\sigma_{11}$. According to Table 1, this condition is written $I_1 \ll 10\text{–}10^2$ Ein m^{-2} s^{-1}. The condition $k'_1 \ll k_C$ is then fulfilled enabling the elimination of the fast species

B. Following Appendix 1, we write $\frac{dB}{dt} = 0$ where B is given in Eq. (12). We obtain the two-state mechanism given in Fig. 1c. Using Eq. (17), we write the fluorescence intensity as

$$I_F = \left(Q_A + Q_B \frac{k'_1}{k_A} \right) A I_1 \tag{18}$$

In this regime, the fluorescence intensity behaves as the concentration A following a monoexponential decay associated with the characteristic time (see Appendix 2 and apply the results to F ≡ A, N ≡ C, $k \equiv k'_1$, and $k' = 0$)

$$\tau = \frac{1}{k'_1} \tag{19}$$

The characteristic time is inversely proportional to I_1 and the photochemical step associated with the rate constant k'_1 controls the fluorescence evolution.

4.2 High Intensity I_1: Regime α_h

Then we consider the regime α_h of high light intensity such that the condition $I_1 \gg k_A/\sigma_{11}$, i.e. $k_A \ll k'_1$ is fulfilled. Table 1 is used to assess the interval of $I_1 \gg 10$–10^5 Ein m^{-2} s^{-1} in which the condition is fulfilled. In this condition, species A is fast and can be eliminated using the steady-state approximation. Following Appendix 1, we write $\frac{dA}{dt} = 0$ where A is given in Eq. (11). We obtain the two-state thermal mechanism shown in Fig. 1d. Using Eq. (17), we write the fluorescence intensity as

$$I_F = Q_B B I_1 \tag{20}$$

In the regime α_h, the fluorescence intensity evolves as the concentration B, which monoexponentially decreases with the characteristic time (see Appendix 2 and apply the results to F ≡ B, N ≡ C, $k \equiv k_C$, and $k' = 0$)

$$\tau = \frac{1}{k_C} \tag{21}$$

The characteristic time is independent of I_1 and a thermal step controls the fluorescence evolution. The hypothesis on the fluorescent nature of species B* and consequently B induces the existence of a fluorescent signal in the regime α_h.

5 Illumination Protocol β

In the case of illumination β, we assume that the initial condition is $C = C_{\text{tot}}$. The four-state mechanism for low and high light intensities I_2 is reduced in the following.

5.1 Low Intensity I_2: Regime β_l

We consider the regime β_l associated with a low light intensity I_2 with I_1 fixed around 0.1 Ein m^{-2} s^{-1} in order to obtain a sufficient level of fluorescence. We assume that the condition $I_2 \ll \min(k_A/\sigma_{22}, k_C/\sigma_{12})$ is met, so that $k'_1 \ll k_C$ and $k'_2 \ll k_A$. Table 1 is used to evaluate the interval of $I_2 \ll$ 10–10^2 Ein m^{-2} s^{-1} in which the condition is fulfilled. The two species B and D are fast and can be eliminated. Following Appendix 1, we write $\frac{dB}{dt} = 0$ and $\frac{dD}{dt} = 0$ where B and D are given in Eqs. (12) and (14). We obtain the two-state mechanism shown in Fig. 1e. According to Eq. (18), the fluorescence intensity behaves as the concentration A, which monoexponentially increases with the characteristic time (see Appendix 2 and apply the results to F $\equiv$ A, N $\equiv$ C, $k \equiv k'_1$, and $k' \equiv k'_2$)

$$\tau = \frac{1}{k'_1 + k'_2} \tag{22}$$

The characteristic time is a function of both I_1 and I_2. Photochemical steps control the fluorescence evolution.

5.2 High Intensity I_2: Regime β_h

Finally we consider the regime β_h associated with high light intensities I_2 and $I_1 \sim$ 0.1 Ein m^{-2} s^{-1}. We suppose that the intensity I_2 obeys $I_2 \gg k_C/\sigma_{22}$, so that $k_C \ll k'_2$. Table 1 is used to estimate the interval of $I_2 \gg$ 10–10^5 Ein m^{-2} s^{-1} in which the condition is fulfilled. The fast species C is eliminated. Following Appendix 1, we write $\frac{dC}{dt} = 0$ where C is given in Eq. (13). We obtain the three-state mechanism shown in Fig. 1f. The fluorescence intensity is given in Eq. (17).

In the regime β_h, the dynamics is described by two independent variables. Two characteristic times describe the fluorescence evolution

$$\tau_1 = \frac{2}{k'_1 + k_A + k_C + \sqrt{(k'_1 + k_A - k_C)^2 - 4k'_1 k_A}} \tag{23}$$

$$\tau_2 = \frac{2}{k'_1 + k_A + k_C - \sqrt{(k'_1 + k_A - k_C)^2 - 4k'_1 k_A}} \tag{24}$$

The characteristic time τ_1 is smaller than τ_2. Consequently fluorescence evolution is governed at short times by τ_1 and at long times by τ_2. For the initial condition $C = C_{\text{tot}}$, the amplitude of the fluorescence evolution associated with τ_1 is negative and the amplitude associated with τ_2 is positive. Consequently, fluorescence first increases and then decreases. Both photochemical and thermal steps control the dynamics.

6 Comparison of Experiments Involving the RSFP Dronpa-2 with the Predictions of the Reduced Mechanisms

The evolution of the fluorescence intensity of a solution of the RSFP Dronpa-2[24] is shown in Fig. 2 for four light intensity conditions illustrating the four illumination regimes α_l, α_h, β_l, and β_h. The evolution has been recorded during 0.4 s with a sampling interval of 10^{-6} s. As a consequence, only reaction steps associated with characteristic times larger than 1 µs and smaller than 0.1 s can be analyzed. Monoexponential functions that have the best fit to the experimental data are determined.

As shown in Fig. 2a for the regime α_l, the fluorescence intensity decays monoexponentially, in agreement with the prediction of the two-state mechanism given in Fig. 1c. We have checked that varying the light intensity I_1 leads to a linear variation of the inverse of the experimental characteristic time of fluorescence evolution. For a low enough light intensity I_1 under the illumination protocol α, the kinetics is governed by a photochemical step. The fitted monoexponential function to the fluorescence data provides the characteristic time $\tau = 67.8 \pm 0.1$ ms for $I_1 = 0.08$ Ein m^{-2} s^{-1} and $I_2 = 0$. Using the expression of the characteristic time τ given in Eq. (19) for the reduced mechanism associated with the regime α_l and the relationship between a rate constant and the light intensity in a photochemical step, we evaluate the isomerization cross-section at $\lambda_1 = 480$ nm to 184 m^2 mol^{-1}, close to the tabulated value[17] $\sigma_{11} = 196$ m^2 mol^{-1}.

The reduced scheme given in Fig. 1d is a two-state mechanism which predicts a monoexponential decay of fluorescence intensity in the case of the regime α_h. However, Fig. 2b shows at least a biexponential decay of fluorescence. The experimental results reveal that the seven-state mechanism does not capture all the complexity of RSFP kinetics. Nevertheless our model satisfactorily reproduces the thermal nature of the chemical reaction experimentally validated by the invariance of the characteristic times as the light intensity increases. The fitted monoexponential function to the fluorescence data provides the characteristic time $\tau = 240 \pm 20$ µs. This result evidences the radiative de-excitation of species B*, i.e. Trans-O^{-*}. Indeed, if A*, i.e. Cis-O^{-*}, is the only source of fluorescence, the characteristic time

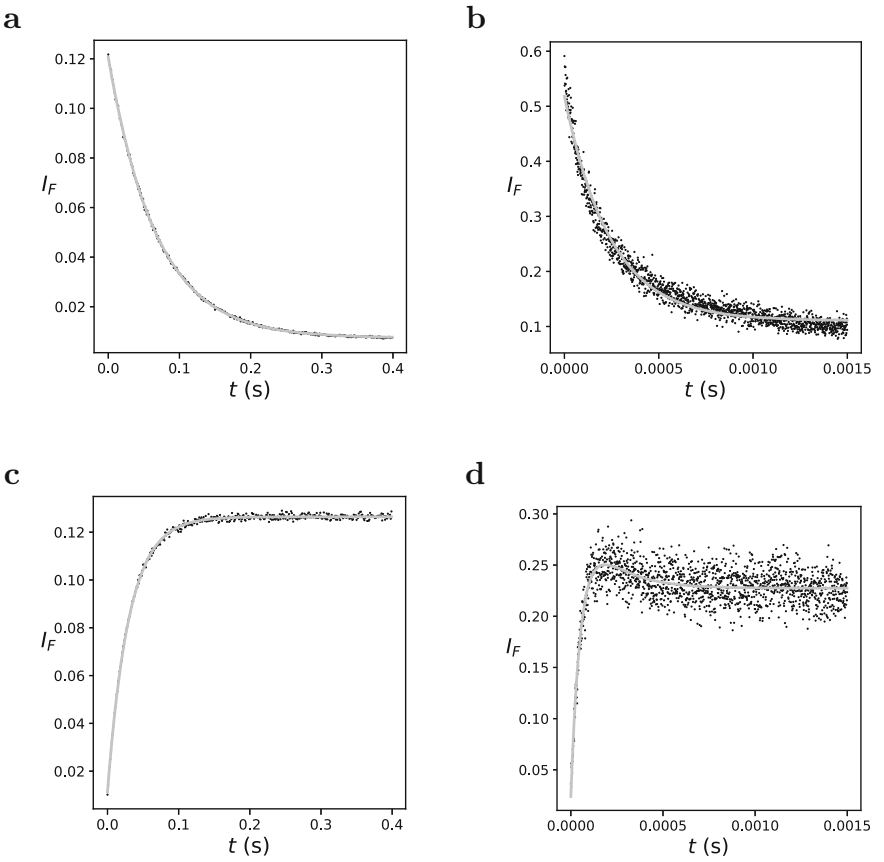


Fig. 2 Experimental evolution of the fluorescence intensity I_F (black dots) of a 20 µM solution of the RSFP Dronpa-2 in pH 7.4 PBS (50 mM sodium phosphate, 150 mM NaCl) at T = 298 K under (**a**): illumination α_l for $I_1 = 0.08$ Ein m^{-2} s^{-1}, (**b**): illumination α_h for $I_1 = 228.8$ Ein m^{-2} s^{-1}, (**c**): illumination β_l for $I_1 = 0.15$ Ein m^{-2} s^{-1} and $I_2 = 0.01$ Ein m^{-2} s^{-1}, (**d**): illumination β_h for $I_1 = 0.15$ Ein m^{-2} s^{-1} and $I_2 = 106$ Ein m^{-2} s^{-1}. The gray lines in (**a**), (**b**), and (**c**) are monoexponential fitted curves to the experimental data. The gray line in (**d**) is a biexponential fitted curve

of fluorescence would be given by $1/k_1' = 22$ µs in the experimental conditions, much smaller than the found τ value. Using the expression of the characteristic time τ given in Eq. (21) for the reduced mechanism associated with the regime α_h, we obtain the following evaluation of the unknown protonation rate constant $k_C = (4.1 \pm 0.4) \times 10^3$ s^{-1}.

As displayed in Fig. 2c for the regime β_l, the fluorescence intensity increases monoexponentially, in agreement with the prediction of the two-state mechanism given in Fig. 1e. We have checked that varying the light intensity I_2 leads to a linear variation of the inverse of the experimental characteristic time of fluorescence

evolution. For a low enough light intensity I_2 with a small light intensity I_1 under the illumination protocol β, the kinetics is governed by a photochemical step. The fitted monoexponential function to the fluorescence data provides the characteristic time $\tau = 30.3 \pm 0.1$ ms for $I_1 = 0.15$ Ein m^{-2} s^{-1} and $I_2 = 0.01$ Ein m^{-2} s^{-1}. Using the expression of the characteristic time τ given in Eq. (22) for the reduced mechanism associated with the regime β_l and the relationship between a rate constant and the light intensities in a photochemical step, we evaluate the sum of the isomerization cross-sections at $\lambda_2 = 405$ nm to 360 m^2 mol^{-1}, close to the value of the literature[17] $\sigma_{12} + \sigma_{22} = 413$ m^2 mol^{-1}.

The reduced scheme given in Fig. 1f is a three-state mechanism which predicts a fast increase followed by a slow decrease of fluorescence intensity in the case of the regime β_h. Indeed Fig. 2d shows that the fluorescence intensity reaches a maximum. However the sum of at least three exponential terms has an acceptable fit to the experimental data. We have experimentally checked that increasing the light intensity I_2 does not affect the characteristic time associated with the initial increase of fluorescence intensity but modifies the characteristic times associated with the decrease. Once again the results obtained at high light intensities which are sensitive to thermal steps reveal unknown subtleties of RSFP dynamics. According to literature[20], the deprotonation rate constant is given by $k_A = 5.2 \times 10^4$ s^{-1}. For the light intensity $I_1 = 0.15$ Ein m^{-2} s^{-1} and the isomerization cross-section $\sigma_{11} = 196$ m^2 mol^{-1}, the photochemical rate constant is $k_1' = 29.4$ s^{-1}. Assuming $k_1' \ll k_C \ll k_A$, we expand the characteristic times given in Eqs. (23) and (24) at leading order in $k_1'/(k_A - k_C)$ and find $\tau_1 \simeq 1/k_A$ and $\tau_2 \simeq 1/k_C$. The fitted biexponential function to the fluorescence data provides the characteristic times $\tau_1 = 50 \pm 3$ µs and $\tau_2 = 230 \pm 30$ µs. The results yield $k_A = (2.0 \pm 0.1) \times 10^4$ s^{-1} and $k_C = (4.3 \pm 0.5) \times 10^3$ s^{-1}. The two values of k_C derived in the regimes α_h and β_h remarkably agree. The found value of k_A is in satisfying agreement with the result obtained using UV-visible transient absorption spectroscopy[20]. The difference of less than a factor of 3 could be due to a different sensitivity of the absorption and fluorescence observables to additional microsecond protein reorganization steps[23] that were neglected in the seven-state mechanism. An effect of the solution pH (7.4 instead of 8.0 in the transient absorption study[20]) on the chromophore deprotonation kinetics is moreover not excluded.

7 Conclusion

In this work we investigate the kinetics of reversibly photoswitchable fluorescent proteins (RSFPs) used in biological imaging. Light intensity is tuned to change rate limiting steps, from photochemical steps for low light intensity to thermal steps for high enough light intensity. Fluorescence evolution is analyzed to evaluate characteristic times. Starting from a reasonable seven-state mechanism, we use the steady-state approximation to reduce the dynamics to simple schemes valid in each light intensity regime. Hence a small number of characteristic times can be

associated with the experimental fluorescence data allowing us to evaluate some rate constants. In particular we assert the fluorescent nature of species Trans-O^{-*} and determine the value of the protonation rate constant k_C unknown in the literature. Transient absorption spectroscopy commonly used to investigate the kinetics of RSFPs has not yet been conclusive due to the small isomerization cross-section σ_{11} which limits the signal. Our approach based on the analysis of fluorescence photoswitching kinetics under continuous illumination benefits from the very high sensitivity of this observable and does not require any expensive pulsed laser equipment.

Acknowledgments This work is supported by the French research agency programs ANR-19-CE29-0003 and ANR-19-CE11-0005.

Appendix 1: The Steady-State Approximation

The steady-state approximation consists in eliminating a concentration from the rate equations, provided that the time dependence of this variable contains a component relaxing faster than the other concentrations. In this work, first-order reactions are involved and we illustrate the steady-state approximation in the case of a system of linear differential equations for the two variables A and B associated with the following chemical scheme

$$\mathrm{A} \longrightarrow [\mathrm{k}_1]\mathrm{C}_1 \tag{25}$$

$$\mathrm{A} \rightleftharpoons [\mathrm{k}_2][\mathrm{k}_3]\mathrm{B} \tag{26}$$

$$\mathrm{B} \longrightarrow [\mathrm{k}_4]\mathrm{C}_2 \tag{27}$$

where the concentrations C_i for $i = 1, 2$ are supposed to be constant due to the action of appropriate reservoirs called chemostats. The rate equations are given by

$$\frac{\mathrm{d}A}{\mathrm{d}t} = -(k_1 + k_2)A + k_3 B \tag{28}$$

$$\frac{\mathrm{d}B}{\mathrm{d}t} = k_2 A - (k_3 + k_4)B \tag{29}$$

Considering the small parameter $\epsilon = k_3/(k_1 + k_2)$, we find at leading order that $\frac{\mathrm{d}A}{\mathrm{d}t} = 0$, which amounts to an instantaneous relaxation of the fast variable A. Equation (28) then leads to the relationship $A = k_3/(k_1 + k_2)B$ allowing us to eliminate A from the rate laws. The small ϵ condition is equivalent to assuming that the sum of the rate constants of the steps consuming the fast species is larger than the sum of the rate constants of the steps producing the considered species. In addition, the fast species can be eliminated from the chemical scheme leading to the reduced mechanism

$$\mathrm{B} \xrightarrow{k_4+\frac{k_1 k_3}{k_1+k_2}} \mathrm{C}_2 \tag{30}$$

Appendix 2: Fluorescence Evolution in a Two-State Mechanism

We consider the standard two-state mechanism describing the reversible exchange between a bright fluorescent state F and a dark nonfluorescent state N

$$\mathrm{F} \underset{k'}{\overset{k}{\rightleftharpoons}} \mathrm{N} \tag{31}$$

where k and k' are the rate constants. The law of conservation of matter leads to $F + N = C_{\mathrm{tot}}$. The evolution of the system is given by the single equation

$$\frac{\mathrm{d}F}{\mathrm{d}t} = -(k + k')F + k'C_{\mathrm{tot}} \tag{32}$$

Choosing as initial condition $F = C_{\mathrm{tot}}$, we find

$$F = C_{\mathrm{tot}}\left(1 - \exp\left(\frac{-t}{\tau}\right)\right) \tag{33}$$

where $\tau = 1/(k + k')$ is the characteristic time. The fluorescence evolution is proportional to the concentration of the bright state F and consequently decays monoexponentially.

References

1. K. M. Marks and G. P. Nolan, *Nat. Methods* **3**, 591 (2006).
2. K. M. Dean and A. E. Palmer, *Nat. Chem. Biol.* **10**, 512 (2014).
3. E. A. Specht, E. Braselmann, and A. E. Palmer, *Annual review of physiology* **79**, 93 (2017).
4. S. Habuchi, R. Ando, P. Dedecker, W. Verheijen, H. Mizuno, A. Miyawaki, and J. Hofkens, *Proc. Natl. Acad. Sci. USA* **102**, 9511 (2005).
5. S. W. Hell J. and Wichmann, *Opt. Lett.* **19**, 780 (1994).
6. M. Hofmann, C. Eggeling, S. Jakobs, and S. W. Hell, *Proc. Natl. Acad. Sci. USA* **102** 17565 (2005).
7. P. Dedecker, J. Hotta, C. Flors, M. Sliwa, H. Uji-i, M. B. J. Roeffaers, R. Ando, H. Mizuno, A. Miyawaki, Atsushi, and J. Hofkens, *J. Am. Chem. Soc.* **129** 16132 (2007).
8. P. Dedecker, G. H. H. Mo, Gary, T. Dertinger, and J. Zhang, *Proc. Natl. Acad. Sci. USA* **109** 10909 (2012).
9. R. Ando, H. Mizuno, and A. Miyawaki, *Science* **306**, 1370 (2004).
10. F. Ciruela, *Curr. Opin. Biotechnol.* **19**, 338 (2008).

11. X. X. Zhou, H. K. Chung, A. J. Lam, and M. Z. Lin, *Science* **338** 810 (2012).
12. X. X. Zhou and L. Z. Michael, *Curr. Opin. Chem. Biol.* **17**, 682 (2013).
13. F. Garzella, B. Storti, R. Bizzarri, A. Losi, W. Gärtner, S. Abbruzzetti, P. Bianchini, C. Viappiani, and A. Diaspro, *Biophys. J.* **120**, 363a (2021).
14. G. Marriott, S. Mao, T. Sakata, J. Ran, D. K. Jackson, C. Petchprayoon, T. J. Gomez, E. Warp, O. Tulyathan, H. L. Aaron, E. Y. Isacoff, and Y. Yan, *Proc. Natl. Acad. Sci. U. S. A.* **105**, 17789 (2008).
15. J. Quérard, T.-Z. Markus, M.-A. Plamont, C. Gauron, P. Wang and A. Espagne, M. Volovitch, S. Vriz, V. Croquette, A. Gautier, T. Le Saux, L. Jullien, *Angew. Chem. Int. Ed.* **127**, 2671 (2015).
16. G. Abbandonato, and B. Storti, G. Signore, F. Beltram, and R. Bizzarri, *Microsc. Res. Tech.* **79**, 929 (2016).
17. J. Quérard, R. Zhang, Z. Kelemen, M.-A. Plamont, X. Xie, R. Chouket, I. Roemgens, Y. Korepina, S. Albright, E. Ipendey, M. Volovitch, H. L. Sladitschek, P. Neveu, L. Gissot, A. Gautier, J.-D. Faure, V. Croquette, T. Le Saux, and L. Jullien, *Nat. Comm.* **8**, 969 (2017).
18. R. Chouket, A. Pellissier-Tanon, A. Lemarchand, A. Espagne, T. Le Saux, and L. Jullien, *Chem. Sci.* **11**, 2882 (2020).
19. M. M. Warren, M. Kaucikas, A. Fitzpatrick, P. Champion, J. T. Sage, and J. J. van Thor, *Nat. Comm.* **4**, 1461 (2013).
20. D. Yadav, F. Lacombat, N. Dozova, F. Rappaport, P. Plaza, and A. Espagne, *J. Phys. Chem. B* **119**, 2404 (2015).
21. J. P. Colletier, M. Sliwa, F. X. Gallat, M. Sugahara, V. Guillon, G. Schiro, N. Coquelle, J. Woodhouse, L. Roux, G. Gotthard, A. Royant, L. M. Uriarte, C. Ruckebusch, Y. Joti, M. Byrdin, E. Mizohata, E. Nango, T. Tanaka, K. Tono, M. Yabashi, V. Adam, M. Cammarata, I. Schlichting, D. Bourgeois, M. Weik, *J. Phys. Chem. Lett.* **7**, 882 (2016).
22. N. Coquelle, M. Sliwa, J. Woodhouse, G. Schiro, V. Adam, A. Aquila, T. R. M. Barends, S. Boutet, M. Byrdin, S. Carbajo, E. De la Mora, R. B. Doak, M. Feliks, F. Fieschi, L. Foucar, V Guillon, M. Hilpert, M. S. Hunter, S. Jakobs, J. E. Koglin, G. Kovacsova, T. J. Lane, B. Lévy , M. Liang, K. Nass, J. Ridard, J. S. Robinson, C. M. Roome, C., Ruckebusch, M. Seaberg, M. Thepaut, M. Cammarata, I. Demachy, M. Field, R. L. Shoeman, D. Bourgeois, J. P. Colletier, I. Schlichting, and M. Weik, *Nat. Chem.* **10**, 31 (2017).
23. S.P. Laptenok, A. A. Gil, C. R. Hall, *et al.*, *Nat. Chem.* **10**, 845 (2018).
24. R. Ando, C. Flors, H. Mizuno, J. Hofkens, and A. Miyawaki, *Biophys. J.* **92**, L97 (2007).

Probability Distributions of p53 Mutations and Their Corresponding Shannon Entropies in Different Cancer Cell Types

S. A. Moghadam, S. I. Omar, and J. A. Tuszynski

1 Introduction

A permanent change in the nucleic acid sequence of a gene is known as a gene mutation. A mutation stems from an error in the DNA replication process, meiosis, mitosis, or for any other DNA damage reason. The smallest mutation happens when a single base pair (in a codon) is replaced by another base pair. In synonymous mutations, replacing a base pair does not change the amino acid that codes for the corresponding protein peptide sequence [1–3]. In contrast, in nonsynonymous mutations, the change in the codon will change the protein sequence. Gene mutations can be attributed to two different origins, namely somatic (acquired) or hereditary (also called germline) mutations [1–3]. A somatic mutation might occur locally in a tissue or an organ, often due to some environmental factors such as UV radiation. Parents have a significant role in the former category since these hereditary mutations or germline mutations are present in every cell of the body [4, 5]. Hence, considering the importance of p53 in the pathogenesis of human cancer offers a great motivation to study the probability of amino acids represented

S. A. Moghadam
Department of Physics, University of Alberta, Edmonton, AB, Canada

S. I. Omar
Department of Oncology, University of Alberta, Edmonton, AB, Canada

J. A. Tuszynski (✉)
Department of Physics, University of Alberta, Edmonton, AB, Canada

Department of Oncology, University of Alberta, Edmonton, AB, Canada

DIMEAS, Politecnico di Torino, Turin, Italy
e-mail: jackt@ualberta.ca

R. P. Mondaini (ed.), *Trends in Biomathematics: Stability and Oscillations in Environmental, Social, and Biological Models*,
https://doi.org/10.1007/978-3-031-12515-7_3

by p53 mutations in different cancer types in order to find any correlation between them.

TP53 codes for the tumor suppressor protein called p53 [6, 7]. In human DNA, *TP53* is located on the 17th chromosome (17P13.1). *TP53* codes for over 15 various isoforms of its product protein denoted p53 [6–8]. The p53 protein is made of 393 amino acids, which are divided into five main domains: (1) N-terminal transactivation domain (amino acids 1–43 and 44–60), which is involved in the activation of different transcription factors, binds to transcription factors and plays the role of a mediator in some interactions [9–12]. (2) Pro-rich domain (amino acids 61–100), which is important for p53 stability and also has a function in transcription activation and induction of transcription-independent apoptosis [12–14]. (3) DNA binding domain (DBD) (amino acids 101–300), which primarily binds to DNA. It is also responsible for binding with the p53 corepressor [12, 14]. (4) Tetramerization domain (amino acids 301–323), which plays a role in the regulation of the oligomeric state of p53 [12, 15–17]. (5) Basic C-terminal domain (amino acids 360–393), which is important in the regulation of the sequence [15, 18].

The DBD of p53 is made of an immunoglobulin-like β-sandwich of two antiparallel β-sheets, providing a scaffold for a flexible DNA-binding surface. This DNA-binding surface is created by two large loops stabilized by a zinc atom and a loop-sheet-helix motif [19–23]. Zinc binding is critical for correct protein folding and requires a reduction of thiol groups on cysteines [8, 19–24]. In its role as a tumor suppressor protein, p53 binds to the DNA regulating the cell cycle [8, 12, 24]. The p53 protein controls the following cellular processes: (a) cell proliferation, (b) cell death, (c) nutrient deprivation, (d) nucleotide depletion, (e) hypoxia and oxidative stress and (f) hyperproliferative signals [12]. These and other cellular functions are performed by p53 primarily by triggering apoptosis, DNA repair, regulation of energy metabolism and anti-oxidant defense [8]. Stimuli that activate p53 include DNA damage, nutrient deprivation, nucleotide depletion, hypoxia, oxidative stress, and hyperproliferative signals [8, 12, 24]. The activated protein plays its role by virtue of being a transcription factor as it binds to the promoter region of different genes to activate their expression in order to induce the above-listed functions as well as cell cycle arrest when required [8, 12, 24].

Numerous studies show that virtually all cancer types exhibit p53 protein mutations, and several studies used computational methods, such as molecular docking, to find pharmacological compounds that are predicted to restore the function of the p53 mutant to its wild-type state [25–42]. It has been hypothesized that on their own, these mutations can lead to tumor initiation and progression [25–37]. Due to the importance of preventing cancer formation in multicellular organisms and the significant role of the p53 protein in conserving the cell's stability, p53 has been described as "the guardian of the genome" [24, 43–52]. A vast majority of the p53 mutations, approximately 95%, take place in the DNA binding domain. Interestingly, about 40% of these amino acid mutations happen in only six specific positions, known as hotspot mutations, in which the frequency of the hotspot mutations is much higher than in other regions of the DNA sequence. These hotspot mutations involve the following specific residue changes R175H, G245S, R248W,

R249S, R273H, and R282Q [4]. The most common type of mutation in cancer is mainly missense, nonsense and deletion but the pattern of mutations is different in different ethnic groups, which also depends on the geographical location [53]. Most mutations in the DBD region are missense; in contrast, outside this region, missense mutations represent only about 40%, the majority of mutations being nonsense or frameshift [54]. *TP53* mutations occur in nearly all types of cancer, such as: ovarian, esophageal, colorectal, head and neck, laryngeal and lung cancers, sarcomas, breast, brain, testicular cancer, cervical cancers malignant melanoma, and leukemia. Mutations have been found to be more frequent in advanced stages of the disease. Interestingly, it was also found that in elephants, cancer prevalence was significantly lower than expected based on extrapolation from other species, including humans, which stems partly from the number of copies of the p53 protein in these animals compared to humans, namely twenty copies in elephants and one in humans [50, 51]. The p53 gene counts as the highly frequent mutated gene in human cancers, and more than half of the human tumors include deletions or mutations of the p53 gene bases. For instance, individuals having a single p53 gene's functional copy develop Li-Fraumeni syndrome (LFS), which leads to their predisposition to developing cancer. These rare conditions create multiple autonomous tumors in different tissues. This demonstrates the importance of studying p53 mutations due to their consequences for cell division.

Using different experimental biological techniques, such as gene knockout in mice, has revealed vital information regarding the molecular mechanisms of cancer initiation and progression [55–57]. When the p53 protein binds to the promoter region of the p21 gene, it activates its transcription and hence its expression. The p21 protein interacts with a cyclin-dependent kinase2 (CDK2), which is a protein normally involved in cell division [8, 24, 58–61]. The formation of the p21-CDK2 complex inhibits the function of the latter protein and hence progression of the cell-cycle is inhibited [8, 24, 58–61]. Mutations in p53 can, therefore, inhibit its transcriptional activity and hence alter its control over the cell cycle. Thus, cell division would progress without control and consequently, a tumor can form. A recent study by Baugh et al., discussed the causes behind the hotspot mutations in p53 [4, 62], which were listed as: (1) the mutations in the gene alter the structure of the expressed protein, (2) in a specific DNA sequence, such as a methylated cytosine residue in a CpG dinucleotide, changing it to thymidine causes hotspot mutations to occur at these residues, (3) environmental mutagens create specific changes in the p53 gene and (4) the altered protein causes cancer due to an allele-specific gain of function [4]. In the present paper, we investigate the probability distribution of p53 mutations among the various amino acids and across a number of cancer types.

2 Methodology

For this study, we extracted information regarding the probability distributions of the available p53 mutations from the IARC TP53 database (http://p53.iarc.fr/). This

database has organized and gathered all the published information on the *TP53* gene variations from peer-reviewed literature on human cancers since 1989 [24]. The IARC dataset provides valuable information on *TP53* gene variations and mutations associated with each human cancer sample. This information includes *TP53* germline mutations, somatic mutations, synonymous or nonsynonymous mutations, functional classifications (based on the transcriptional activity), exon numbers, and several other details. Among these categories, *TP53* somatic mutations were mainly considered in this research. Somatic mutations refer to the mutations in sporadic (as opposed to germline) cancers reported in primary tissues, cell lines, and fluids in the body.

We are interested in finding how frequent is a specific conversion of an amino acid into another amino acid in the given gene sequence. For instance, we need to know the frequency of mutating arginine to the other 19 amino acids and compare it with other amino acids, so that this would result in a matrix of 19 by 19 possibilities or 20 by 20 including non-mutated cases. In the gene sequence, different types of mutations occur, and they are recorded in the p53 database as well. The mutation types are missense, silent, nonsense, frameshift, splice, insertions or in-frame deletions, intronic, and upstream mutations in the 5' or 3' UTR (untranslated region). In missense mutations, which are in the nonsynonymous substitution category in the genetic code, a single nucleotide is altered and the produced codon codes for a different amino acid. This type of point mutation is a missense mutation, and it changes the protein sequence encoded. Silent mutations are those types of point mutations in which the changed nucleotide still codes for the same amino acid, and the encoded protein remains the same. The other mutation type involves nonsense mutations that arise when a point mutation of a nucleotide is an introduction of a stop codon. In this case, this mutation in the DNA sequence leads to a premature termination of a protein [24, 63]. Splice mutations refer to the mutations that delete, insert or change the number of nucleotides in the specific site at which splicing occurs during the processing of precursor messenger RNA into mature messenger RNA and are located in the two first and last intron nucleotides, which remain conserved and hence, nominated for change in splicing. Also, intronic mutations happen in introns that are located outside of the splicing site. In human cancers, approximately 90% of the mutations are missense mutations, and the produced protein by these mutations is not sufficiently able to bind to the DNA sequence to regulate the transcriptional pathway of p53 [4]. Among the 189 different mutations in the trinucleotides, eight of them are referred to the codons that contain about 28% of all p53 mutations. Therefore, in our calculations, we are interested in finding the frequency of missense and silent mutations in p53 protein in different types of cancer. Using the IARC database, all wild-type to mutant changes of p53 in 75 different cancer types have been found. Table 1 shows all the human cancer types studied in the present paper. The probability distribution of each of the mutations to other amino acids has been obtained considering the somatic mutations among all the cancer diseases listed in Table 1. Each mutation of the p53 protein is associated with a number of the human samples in the IARC database. The

probability for each mutation from the wild-type sequence is obtained using the formula

$$P_{ij} = \frac{n_{ij}}{N} . \tag{1}$$

where $1 < i < 20$ refers to each amino acid for all somatic mutations available in the database, n_{ij} is the frequency of missense or silent mutations involving ij amino acid pairs, N is the total number of mutations reported in the database and $\sum_{ij} P_{ij} = 1$. Figure 7 in the Appendix shows glycine to alanine mutations of p53 protein in different human cancer types extracted from the IARC database. The total number of mutations for this amino acid is 38, and they are distributed unevenly between the 17 cancer types. For instance, liver cancer has a total number of 1198 mutations, among which 8 stem from glycine to alanine mutations, which gives us the probability distribution for this specific amino acid. Similarly, using Eq. (1), the probability distribution for other amino acid mutations were extracted.

3 Results and Discussion

Having used the IARC TP53 database, the amino acid mutations of p53 protein in different human cancer have been analyzed. Similar to the example of glycine to alanine shown in Fig. 7 in the Appendix, all of the amino acid mutations can be presented as the elements of a matrix whose size is 20 by 20. In the p53 sequence, some of the amino acids do not mutate as reported in the IARC database. From all 400 possible permutations, 189 cases were mutated and the rest (211) did not involve any mutations. The terminology "mutated" means that there is at least one mutation between two amino acids regardless of the number of repetitions in cases in the IARC TP53 database (R20, July 2019) [24]. For instance, there is some information about glycine to alanine mutations in the database and this number is 38 and it is repeated in 17 different cancer types.

Moreover, mutations for each category were extracted based on wild-type to mutant changes in amino acids, including missense and silent mutations. Figure 1a shows the number of all mutations found in different cancer types for each amino acid. Lung cancer has a total number of 3047 different reported mutations, which is the highest number of mutations compared to the other cancers. Among this number, 1880 are missense and silent mutations. Other cancers, such as bladder and breast cancer are the second and third highest mutated cancers, respectively. Summation over the elements of the 20×20 mutation matrix has been calculated using Eq. (2) as

$$P_{M,S} = \sum_{i,j=1}^{20} (p_{ij})^{\alpha} . \tag{2}$$

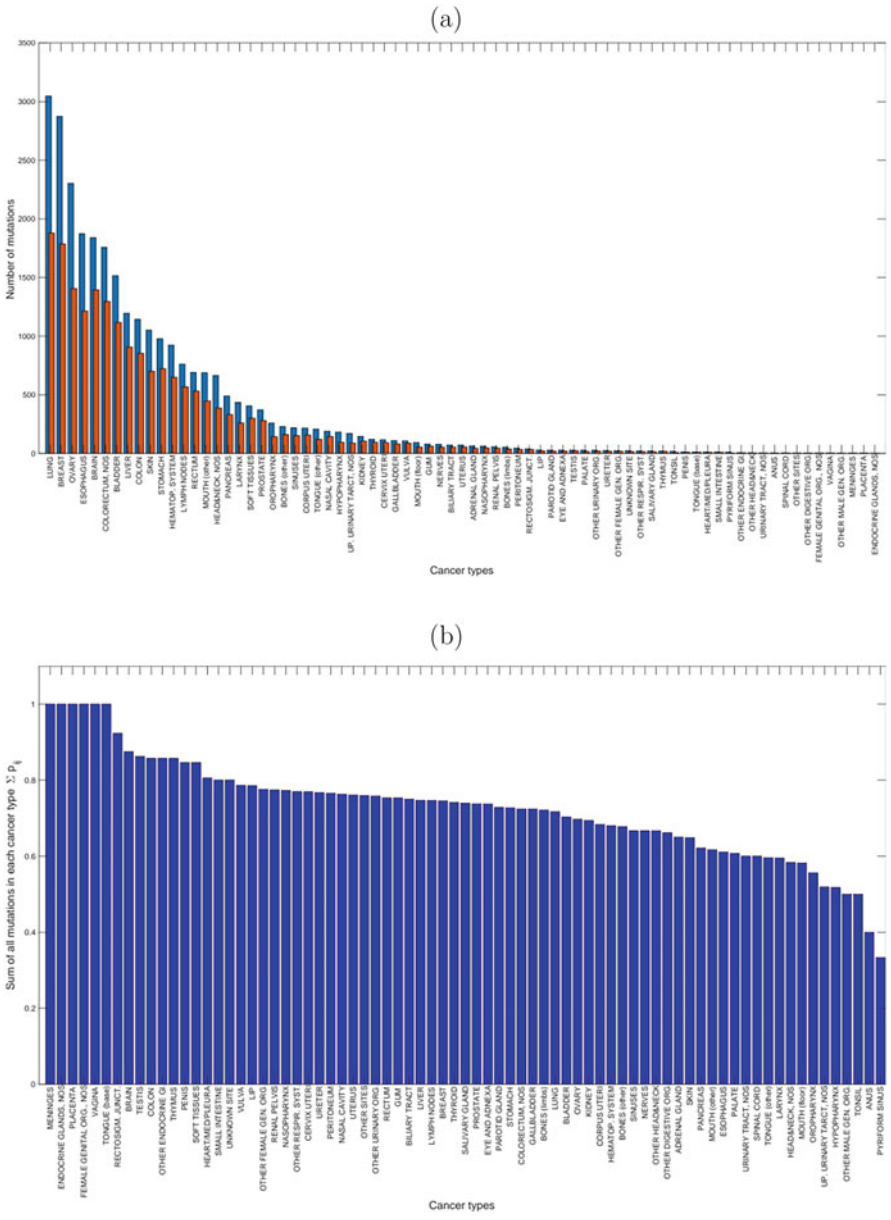


Fig. 1 **(a)** Total number of amino acid mutations for each cancer type in decreasing order, blue bar shows all types of mutations recorded for each cancer type in the IARC database, the red bar shows the missense and silent mutations for each cancer type. **(b)** Sum over all mutation occurrence frequencies in each cancer type as given in Eq. (2) (The summation reaches one for each cancer if other types of mutations are taken into account (not only missense and nonsense)

where $P_{M,S}$ refers to the sum over all missense and silent mutations, p_{ij} stands for the occurrence frequency of each mutation (i to j), and α refers to the cancer type ($1 < \alpha < 75$). The results of Eq. (2) are presented in Fig. 1b. Since we only focus on the missense and silent mutations, the summation over all the probabilities is not equal to 1. The rest of the contributions are for the other types of mutations (in order to compare Fig. 1a, b for each cancer type, the same order is chosen for the cancer type, along the x-axis, see Appendix Fig. 8). In addition, Fig. 2 shows the summation over occurrence frequency, p_{ij}, of all cancer types in one graph. The red bars demonstrate the hot spot mutations of the p53 protein, which are R175H, G245S, R248W, R249S, R273H and R282Q. Mutation of arginine to histidine, R-H, arginine-to-tryptophan R-W, arginine-to-glutamine R-Q, arginine-to-cysteine R-C, glycine-to-serine G-S and arginine-to-serine R-S, are the top-six highly mutated amino acid pairs.

Among all 75 studied cancer types, 79% have at least one arginine-to-histidine mutation, 73% have at least one arginine to glutamine, 71% have arginine to tryptophan. For the next two hotspot mutations, this number drops to 55% for glycine to serine and 48% for arginine to serine. Moreover, in $\sim$ **84**% of the cancer types at least one of the hotspot mutations has a higher frequency compared to other mutants. Figure 3 shows the mutation frequency of p53 in two of the highly mutated cancer types, which are lung (a) and breast (b) cancers. In Fig. 9 in the Appendix, a histogram of all the mutations in the different types of cancer has been plotted separately. Similarly, in most of them, the highest frequency mutations belong to one or more hotspot mutations of p53 protein.

The two-dimensional (2-D), and three-dimensional (3-D) heat map representations of the amino acid mutations' frequency have been plotted. Figure 10a in the Appendix demonstrates a 20×20 matrix with the occurrence frequency of the corresponding mutations of the p53 protein in 2-D and Fig. 10b in the Appendix is a 3-D representation, in which zero means there is no amino acid mutation in that cancer type. The results are color-coded starting from blue, which means there were no mutations, to yellow as the frequency of that p53 mutation increases. Lung cancer has been reported to have the highest number of mutations among all cancers. Arginine to histidine and arginine to tryptophan are the highest frequency mutations in lung cancer.

Next, we investigate the dissimilarity factor relative to a reference number. First, we consider lung cancer, which has the highest number of reported mutations. The dissimilarity factor $\Delta^{\alpha\beta}$ is defined as

$$\Delta^{\alpha\beta} = \sqrt{\frac{\sum_{i,j=1}^{20}(p_{ij}^{\alpha} - p_{ij}^{\beta})^2}{N}} = \sqrt{\frac{\sum_{i,j=1}^{20}(\delta_{ij}^{\alpha\beta})^2}{N}} \,. \tag{3}$$

where N is the normalization factor, α corresponds to all the other cancer types relative to cancer type β, and p_{ij} is the occurrence probability of a mutation of i to j. Equation (3), $\Delta^{\alpha\beta}$, varies from zero to one ($0 < \Delta^{\alpha\beta} < 1$), whereby 1 indicates that the two compared cancer type mutations are dissimilar (the higher the

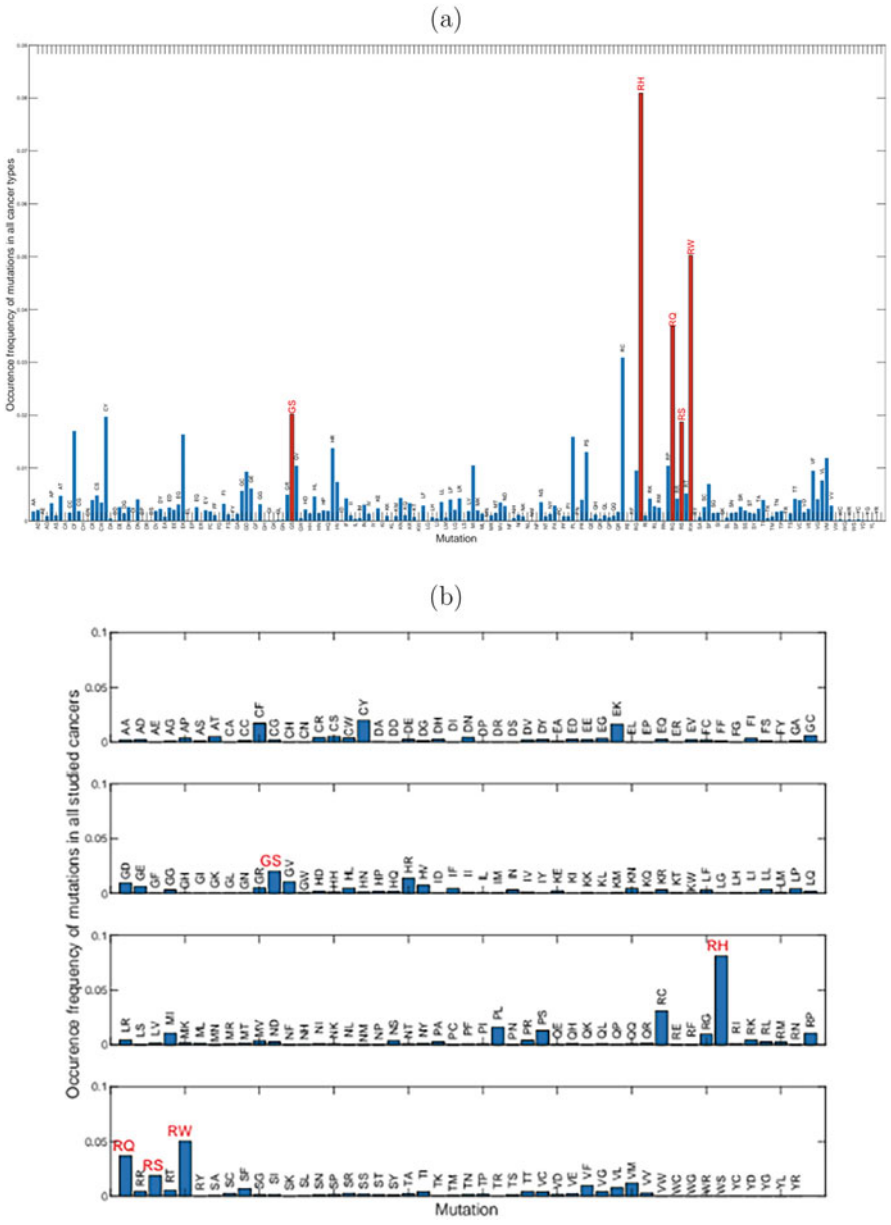


Fig. 2 **(a)** Summation over all the occurrence frequencies of p53 mutations in 75 different cancers types, red bars show the occurrence frequency of p53 hotspot mutations, **(b)** to represent these large data points better, the same data are shown in 4 subplots in the same order as plot (**a**) and the hotspot mutations are labeled in red

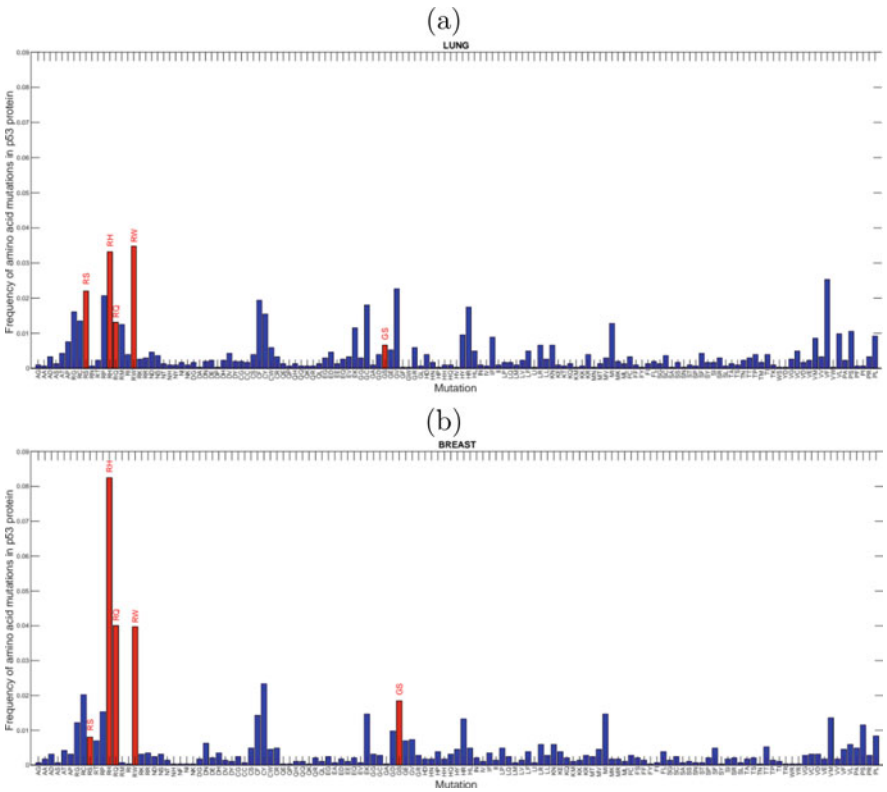


Fig. 3 Frequency of amino acid mutations in the p53 protein in **(a)** lung and **(b)** breast cancer using the IARC database (for only missense and nonsense mutations). The red bars in both plots represent the hotspot mutations of the p53 protein. They are more frequent in these cancer types as well

value, the lower the similarity). Moreover, the similarity factor can be obtained from $\Delta^{\alpha\beta'} = 1 - \Delta^{\alpha\beta}$. In Fig. 4a, b, dissimilarity and similarity coefficients are plotted for all the 75 cancer types. It should be noted that both plots in Fig. 4 are complementary to each other. One conclusion that can be readily drawn is that most of the cancers have similar mutations to those found in lung cancer. Furthermore, some of them, such as cancer of the endocrine glands, placenta and meninges, which have a small number of mutations, are less likely to have similar amino acid mutations to those in lung cancer and hence the similarity factor is low.

As a general calculation, by taking each cancer as a reference, dissimilarity and similarity factors have been obtained for all other cancer types and plotted in Fig. 11 in the Appendix, which shows a 75 by 75 symmetric matrix for dissimilarity factors. As explained in Eq. (3), $\delta_{ij}^{\alpha\beta}$ represents the value of each matrix element. The diagonal elements are zero since they represent the dissimilarity of a mutation frequency of each cancer to itself, $\delta_{ij}^{\beta\beta} = 0$ and the off-diagonal elements show the dissimilarity factor between two cancer types.

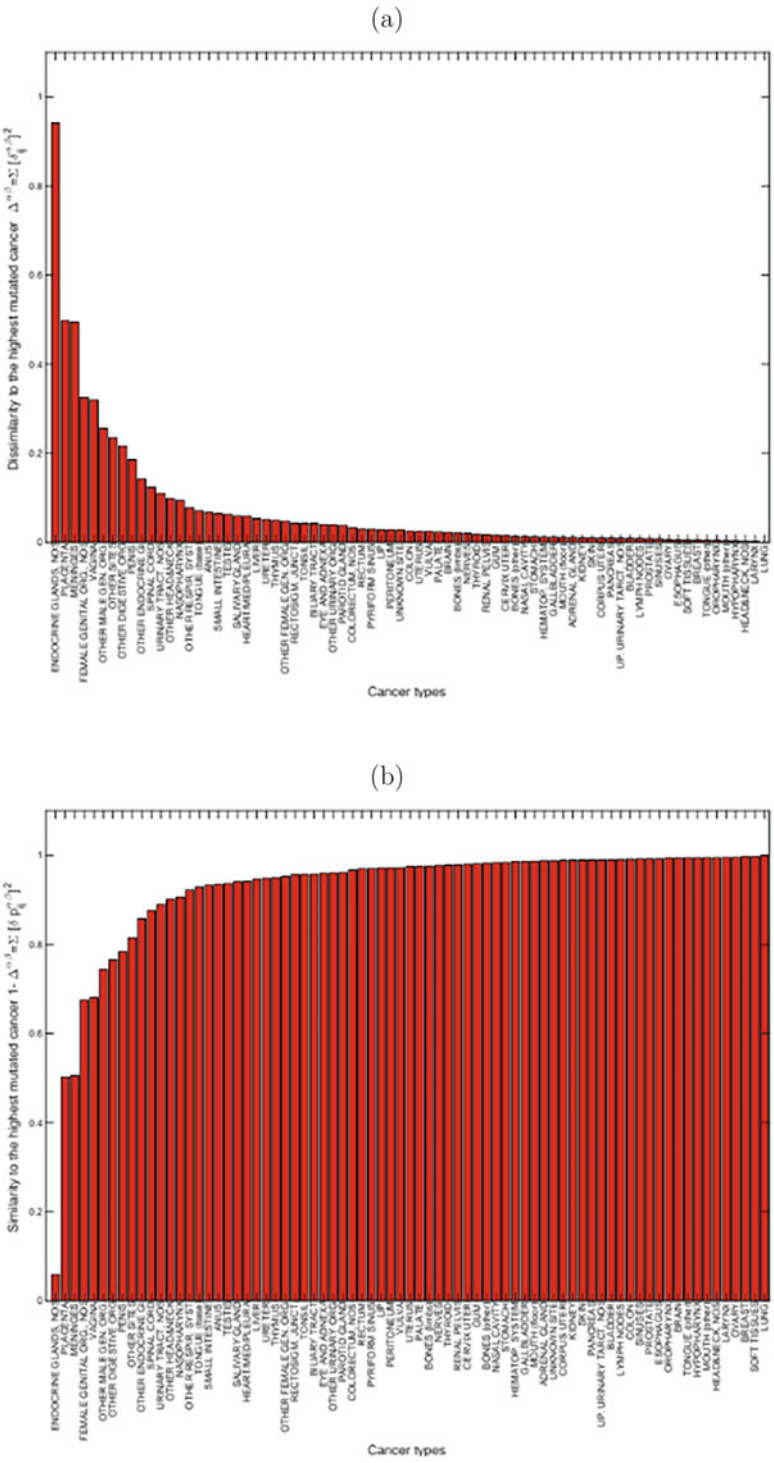


Fig. 4 Mutation dissimilarity factors **(a)** and mutation similarity factors **(b)** obtained from Eq. (3) for all cancers with respect to the lung cancer (these plots are complementary to each other). Lung cancer has been assumed to be the reference case due to its highest number of mutations. The more similar to the reference cancer type, the closer the value to zero is (i.e., similarity for lung cancer to itself is zero and it means they are identical in terms of mutation types)

The next stage to this approach is to consider the entropy values corresponding to the probability distribution of mutations in different cancers. Entropy is an important concept in studying cancer from the view point of information theory and statistical thermodynamics [64–72] Based on the second principle of thermodynamics, in an isolated system, entropy always increases. Also, at the macroscopic level, entropy is a statistical measure of disorder [73]. Entropy can be computed as a system-specific entity that allows us to predict the gap between the present and estimate the final stage of a biological system based on the statistics ofmacroscopic characteristics of the system. The dynamics of the carcinogenesis process, which, among other processes that are dysregulated, is associated with the misplacement of internal cellular information leading to pathological transformations. It can be quantified by accumulation of genomic mutations, which can be studied by using concepts from information theory [73]. Previous studies showed that Shannon entropy is a useful concept for creating a theoretical model of carcinogenesis and prognostic models for patient survival. In this study, we apply the Shannon entropy relation to obtain the entropy of p53 mutations in different cancers [72–75]. The Shannon entropy of a system, which is characterized by a probability distribution p_{ij}^{α}, can be computed using the relation

$$S^{\alpha} = -\kappa_B \sum_{i,j}^{20} p_{ij}^{\alpha} \ln(p_{ij}^{\alpha}) \tag{4}$$

where p_{ij}^{α} is the occurrence probability of an amino acid mutation and κ_B is the Boltzmann constant, $\kappa_B = 0.0083144621$ kJ/mol·K. With the help of Eq. (4) the entropy values for the studied mutations for various cancers are obtained and plotted in Fig. 5. As can be clearly seen, lung cancer has the highest entropy. Recall that lung cancer also has the highest number of p53 mutations. However, this trend is not seen in the rest of the cancer types. For example, bladder and ovarian cancers are the second and third ranked cancers, respectively, in terms of the number of mutations, although here they rank fourth and eleventh among the cancers ordered by their mutation entropy values.

Also, it is interesting to establish if there is a correlation between the entropy of p53 mutations for a given type of cancer and the corresponding 5-years survival rate. Therefore, the 5-year survival rates for the investigated cancer type can be compared using the available databases. Using the statistical information provided by the Surveillance, Epidemiology, and End Results (SEER) Program, which is an authoritative source for cancer statistics located in the United States, the statistical data on 5-years survival of cancer patients were collected. There is a dedicated website at https://seer.cancer.gov/. SEER collects and curates information on cancer cases from around the world. The 5-years survival rates were collected based on the patient's information for the period 2009–2015. These values are obtained comparing survival rates in people who are diagnosed with cancer with those who are healthy without diagnosed cancer, having the same age, sex, and race [76, 77]. Figure 6 shows a plot of entropy as a function of the survival rate for all the

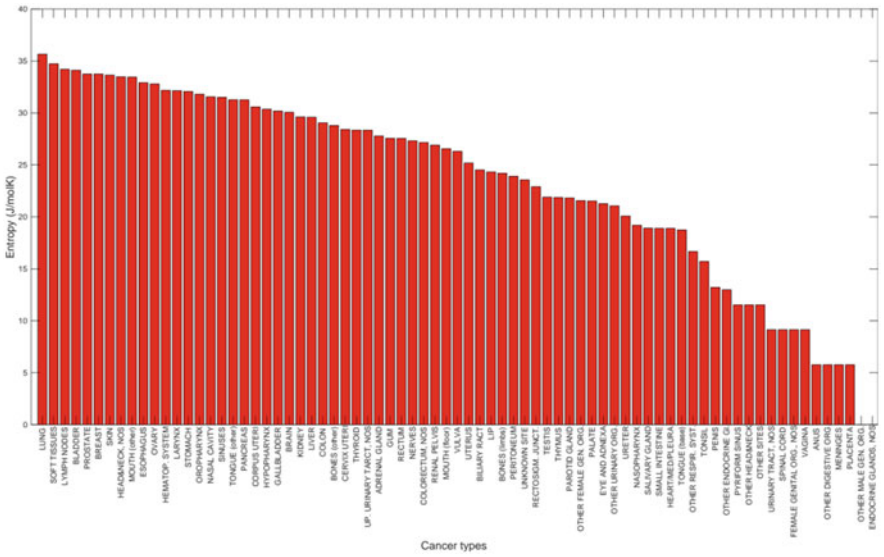


Fig. 5 The Shannon entropy for different cancer types calculated from Eq. (4) (using the occurrence frequency obtained from the IARC database)

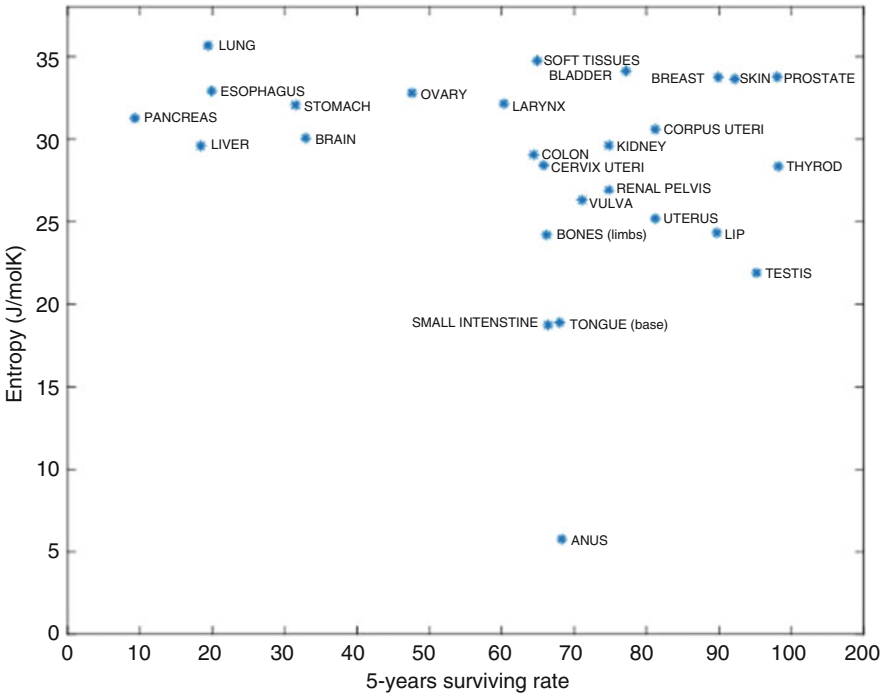


Fig. 6 Mutation entropy of the p53 protein as a function of 5-years survival rate (using SEER database). Each of the cancer types is shown with a blue star symbol and only a weak correlation can be seen

available cancer types in the SEER database. The data are scattered and only a weak correlation can be found [76, 77].

As discussed earlier, if it is equally likely for any amino acid to be mutated to any other amino acid, then it is a reasonable expectation to have the same probability distribution for any amino acid mutations in every cancer type. However, our results showed that some of the amino acids are more probable than others in general. Moreover, in recent decades studies specifically focused on the p53 protein showed that there are several hotspot mutations in well-defined locations of the p53 sequence. Our results show that most of the p53 hotspot mutations have a higher occurrence frequency as well. These observations are in contradiction to the natural selection theory and indicate that the mutation location in the genetic sequence is important for a mutation to happen and it is not a random event. Therefore, there should be a correlation between the occurrence probability of a mutation and the location in the p53 protein sequence. Despite the findings in this study, mutations in p53 could still be random. However, p53 might activate cell death in abnormal cells if the mutations do not affect the wild-type activity of p53 [6, 78]. If this is the case, then only cells with p53 mutations that alter the protein's tumor suppressor activity would be reported and hence bias the results showing that some mutations are more likely than others.

4 Conclusions

In this study, the frequency of p53 mutations of amino acid has been studied in a large number of cancer types. In terms of the number of somatic mutations, lung cancer has the highest number of such mutations. After lung cancer, breast, ovarian, esophageal, brain, and colorectal cancers have the next highest numbers of mutations. We showed that in 84% of somatic mutations, at least one of the hotspot mutations has the highest frequency. The top-five highly mutated amino acids are; arginine-to-histidine, arginine-to-tryptophan, arginine-to-glutamine glycine-to-serine, arginine-to-serine. Moreover, the Shannon entropy of the mutations was also computed and analyzed as a possible characteristic of the associated malignancy. Lung cancer has the highest entropy value of all cancer types and also the highest number of p53 mutations. However, our results indicate that there is no correlation between the entropy of p53 mutations and the number of mutations for all cancer types in general. We also examined the hypothesis that entropy may be correlated with the 5-year survival rate for the available cancers types as listed in the SEER database. Except for the lung cancer, which is the most highly mutated cancer, no obvious trend could be found between the p53 mutation entropy and either the 5-year patient survival rate or the occurrence frequency of mutations across all cancer types (Figs. 7, 8, 9, 10, and 11).

Acknowledgments This research was partially funded by a grant awarded to JAT by NSERC (Canada).

Appendix

Table 1 Cancer types studied with respect to p53 mutations [4]

Topography	Database total	Topography	Database total	Topography	Database total
Adrenal gland	65	Liver	1196	Prostate	373
Anus	5	Lung	3047	Pyriform sinus	12
Biliary tract	73	Lymph nodes	762	Rectosigm. Junct.	40
Bladder	1516	Meninges	2	Rectum	691
BONES (limbs)	53	MOUTH (floor)	94	Renal pelvis	58
BONES (other)	231	MOUTH (other)	689	Salivary gland	22
Brain	1840	Nasal cavity	190	Sinuses	219
Breast	2874	Nasopharynx	62	Skin	1052
Cervix uteri	117	Nerves	79	Small intestine	13
Colon	1144	Oropharynx	259	Soft tissues	406
Colorectum, nos	1758	Other digestive org.	3	Spinal cord	5
Corpus uteri	217	Other endocrine gl.	7	Stomach	978
Endocrine glands, nos	1	Other female gen. org.	25	Testis	29
Esophagus	1873	Other head and neck	6	Thymus	21
Eye and adnexa	29	Other male gen. org.	2	Thyroid	121
Female genital org., nos	3	Other respir. Syst.	22	TONGUE (base)	13
Gallbladder	110	Other sites	4	TONGUE (other)	208

(continued)

Table 1 (continued)

Topography	Database total	Topography	Database total	Topography	Database total
Gum	81	Other urinary org.	28	Tonsil	18
Head & neck, nos	665	Ovary	2303	Unknown site	25
Heart/med/pleura	13	Palate	28	Up. Urinary tract, nos	172
Hematop.System	925	Pancreas	490	Ureter	26
Hypopharynx	183	Parotid gland	29	Urinary tract, nos	5
Kidney	147	Penis	14	Uterus	73
Larynx	437	Peritoneum	46	Vagina	3
Lip	30	Placenta	2	Vulva	108

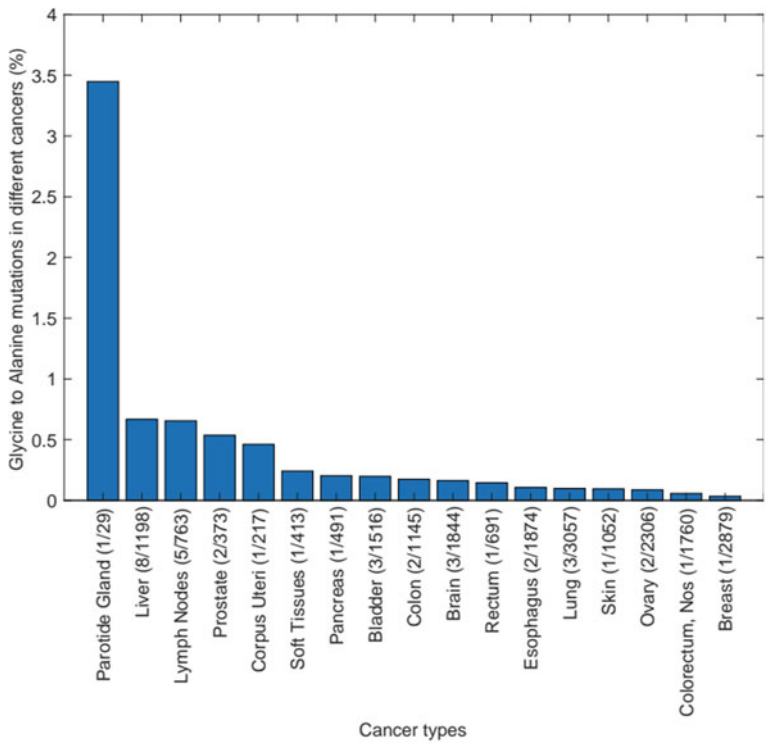


Fig. 7 Somatic mutations of glycine to alanine in different cancers, IARC TP53 database, R20, July 2019 [24]

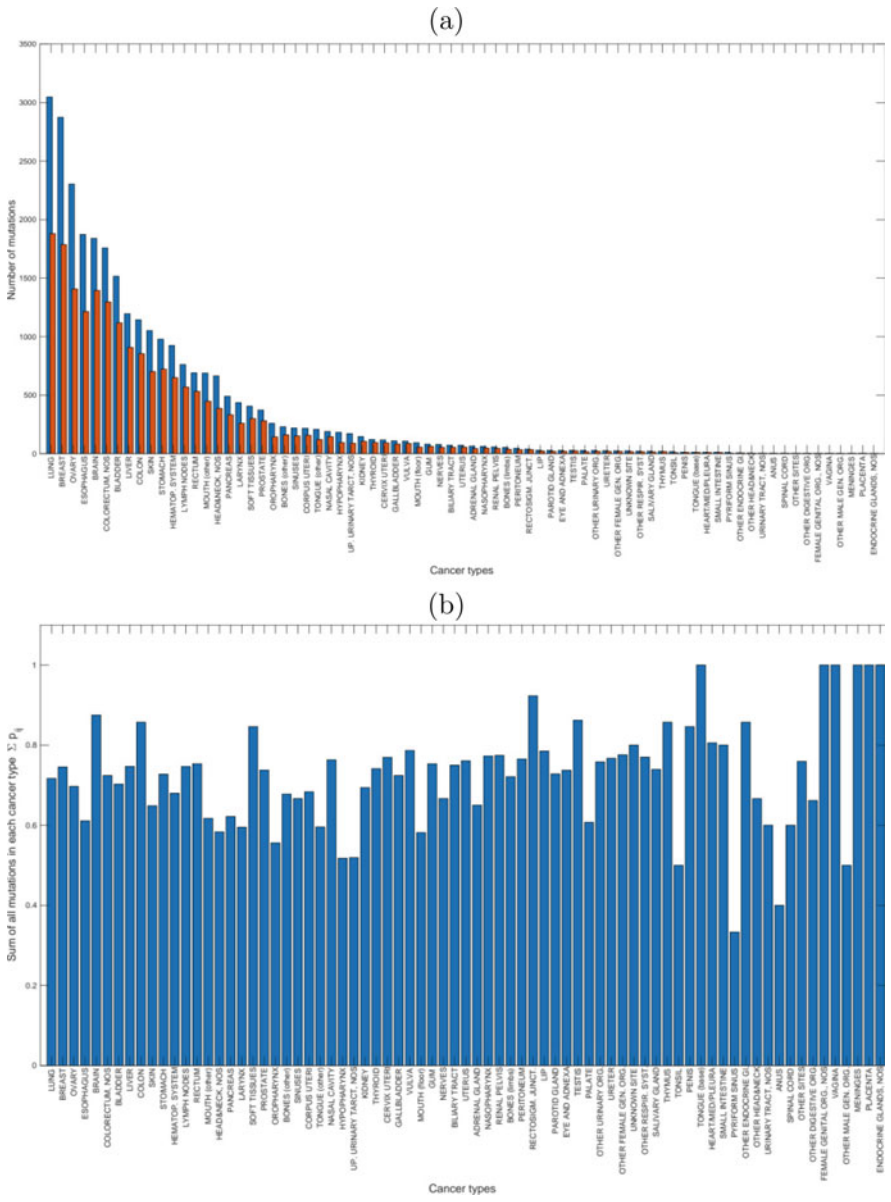
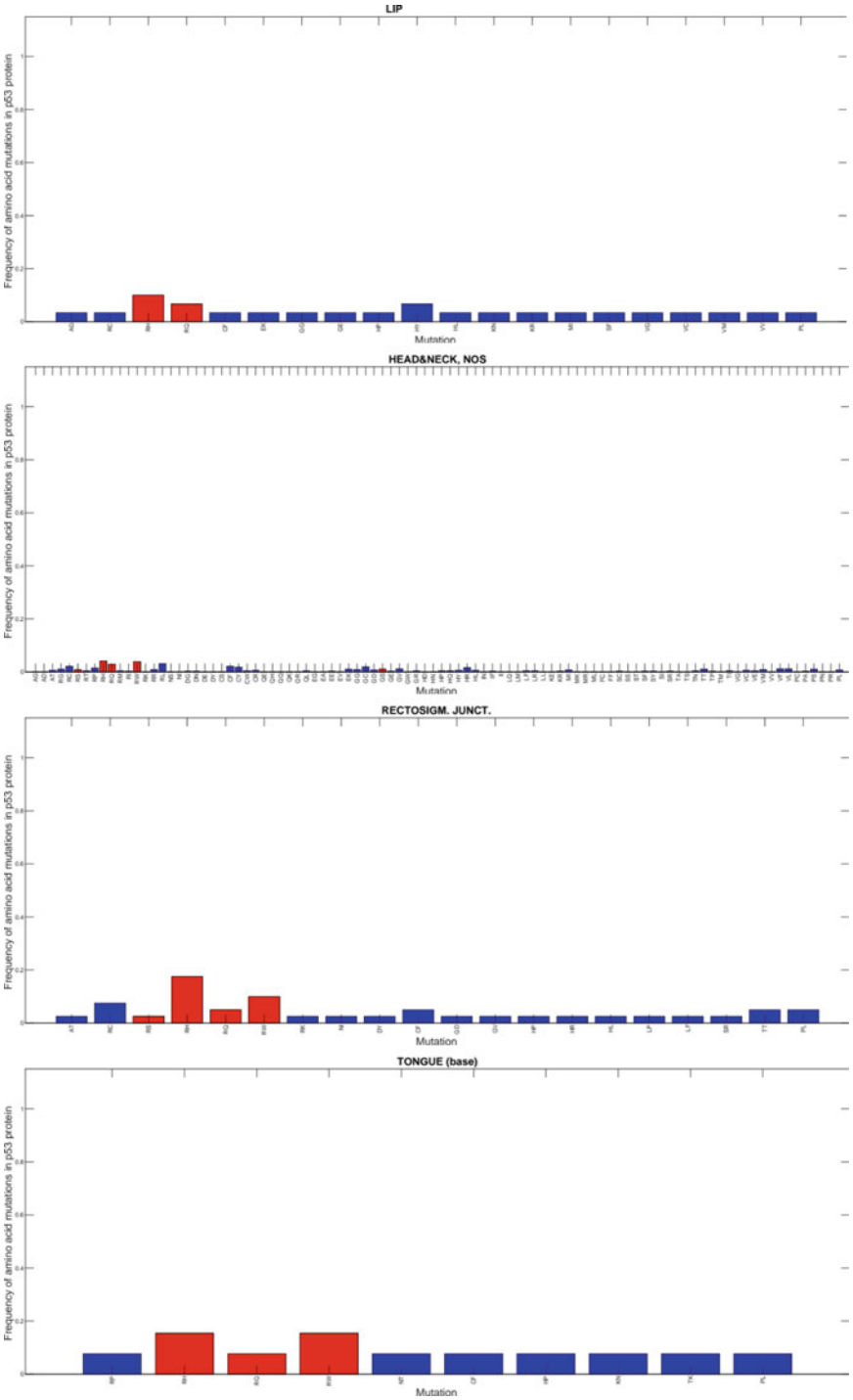
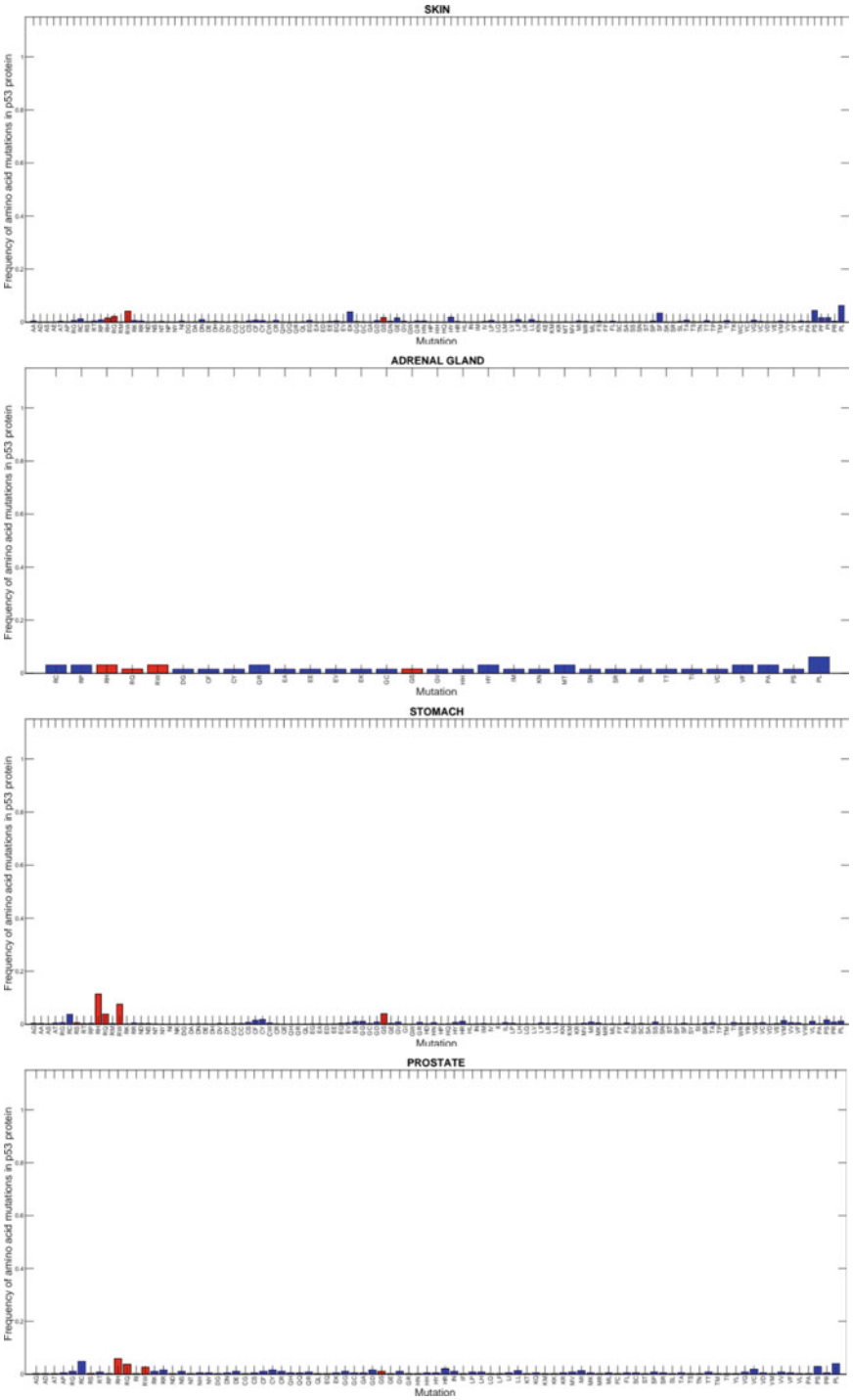
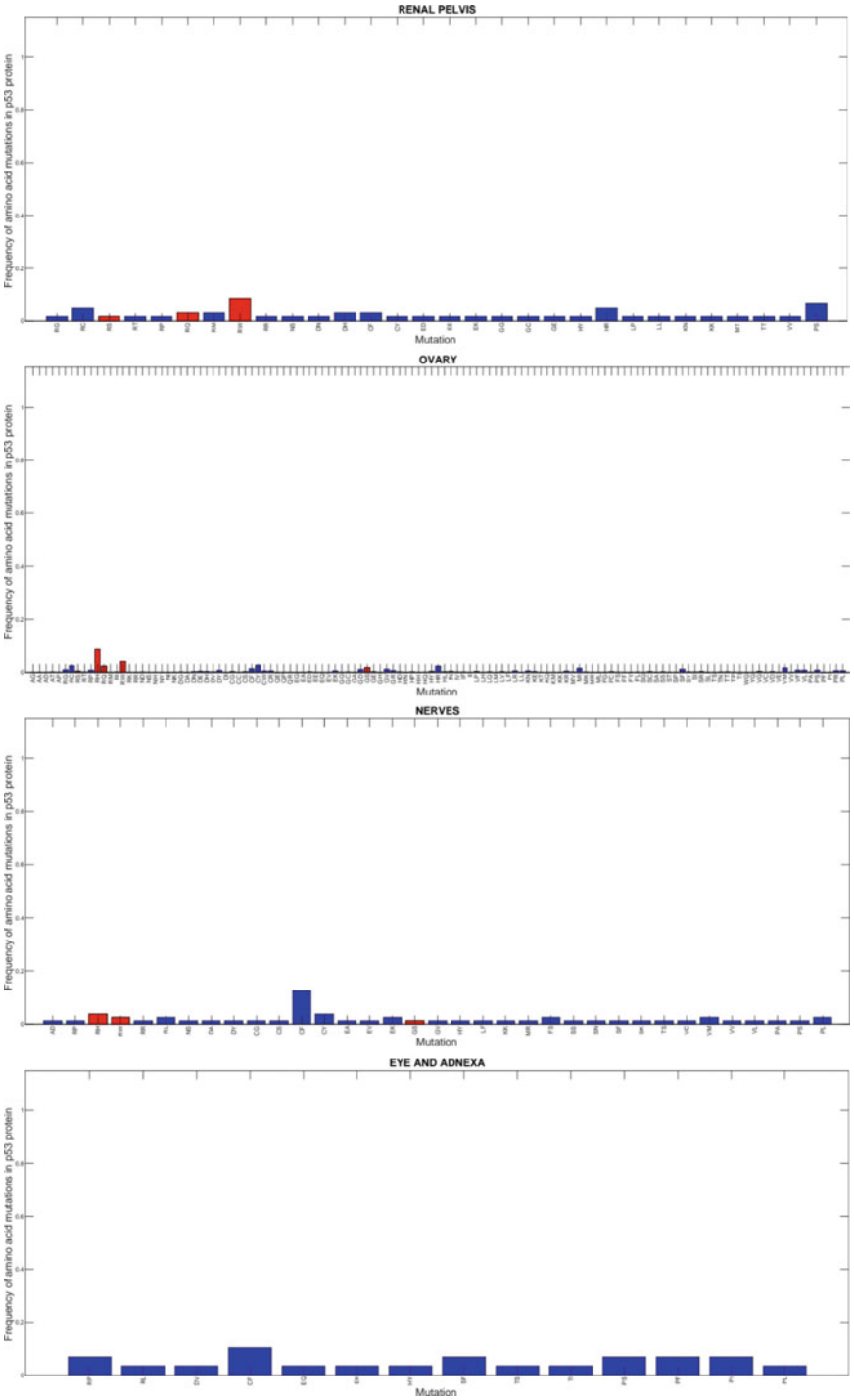
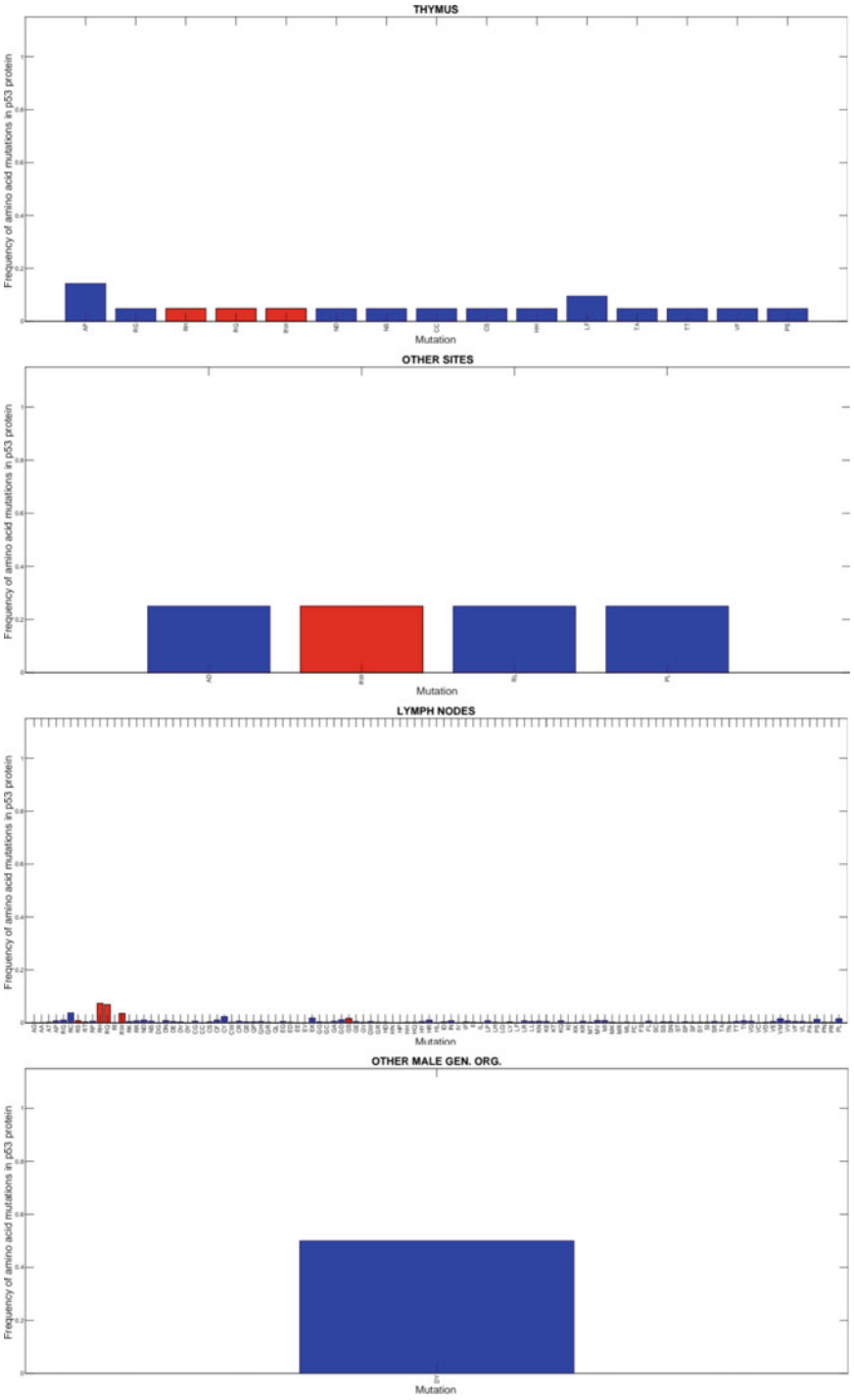


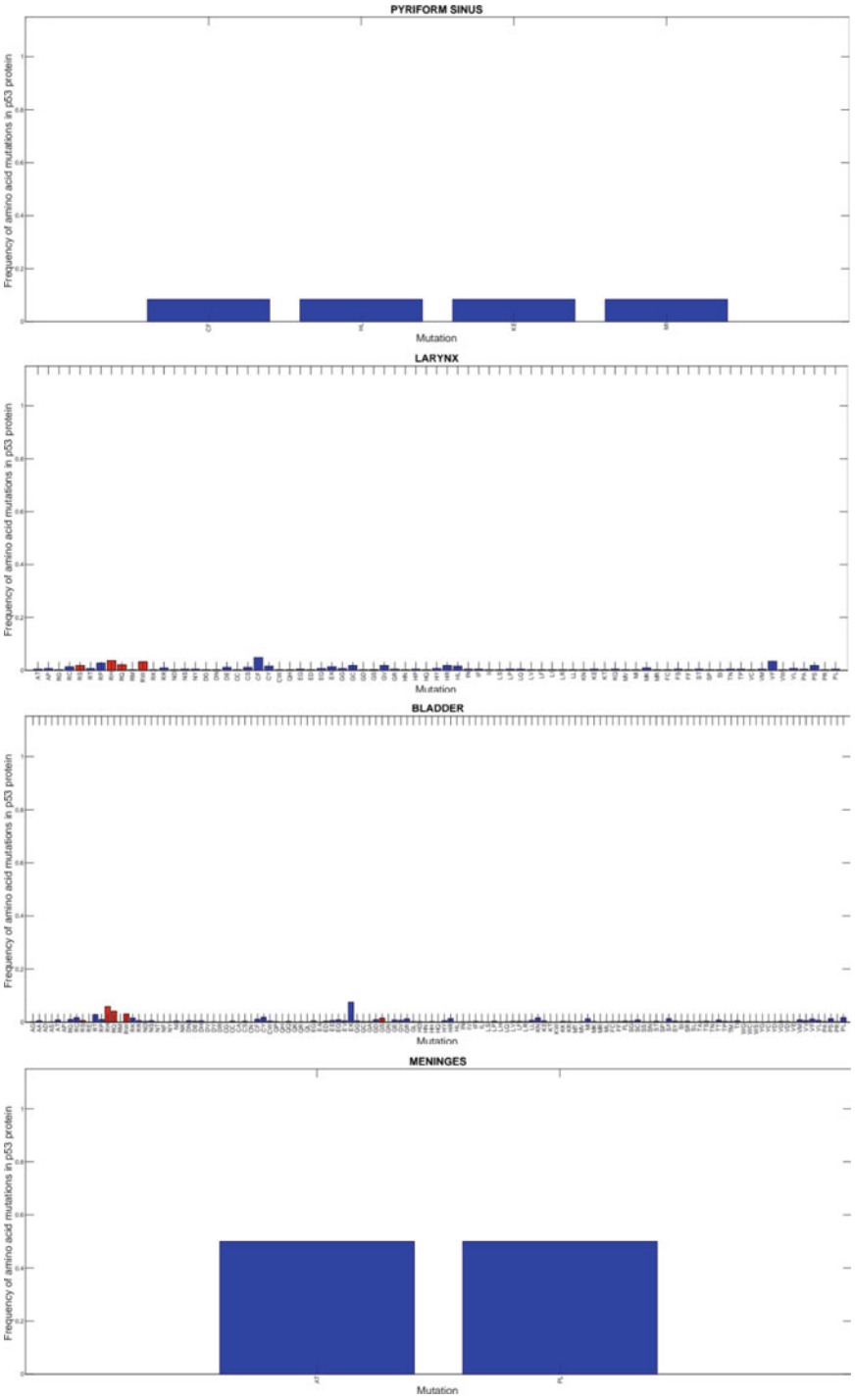
Fig. 8 **(a)** Numbers of mutations for each cancer type shown in decreasing order. **(b)** Sum over all the amino acid occurrence frequencies in each cancer type in Eq. (2) in the same order of appearing in plot (**a**)

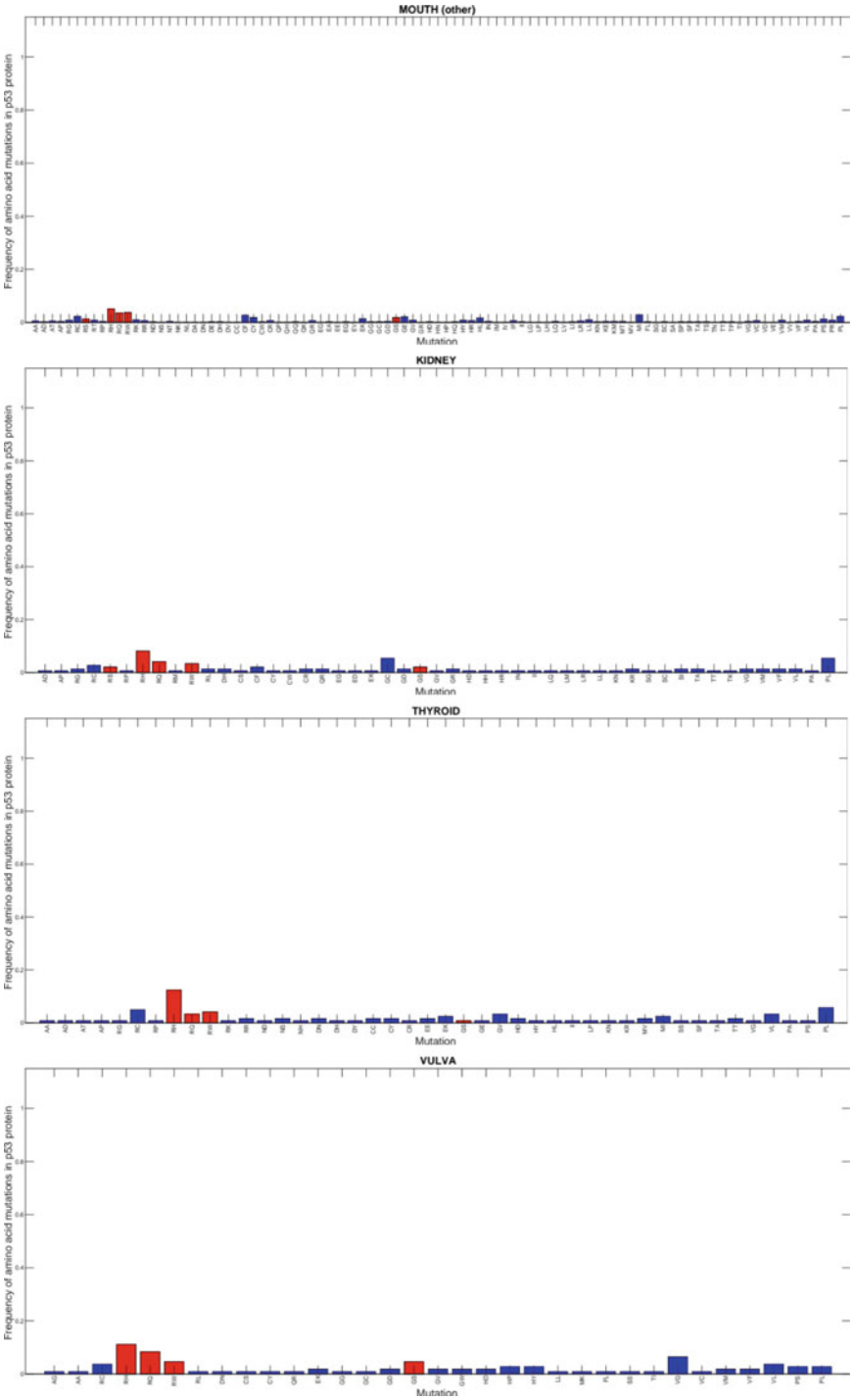

LIP
Frequency of amino acid mutations in p53 protein
Mutation
HEAD&NECK, NOS
Frequency of amino acid mutations in p53 protein
Mutation
RECTOSIGM. JUNCT.
Frequency of amino acid mutations in p53 protein
Mutation
TONGUE (base)
Frequency of amino acid mutations in p53 protein
Mutation

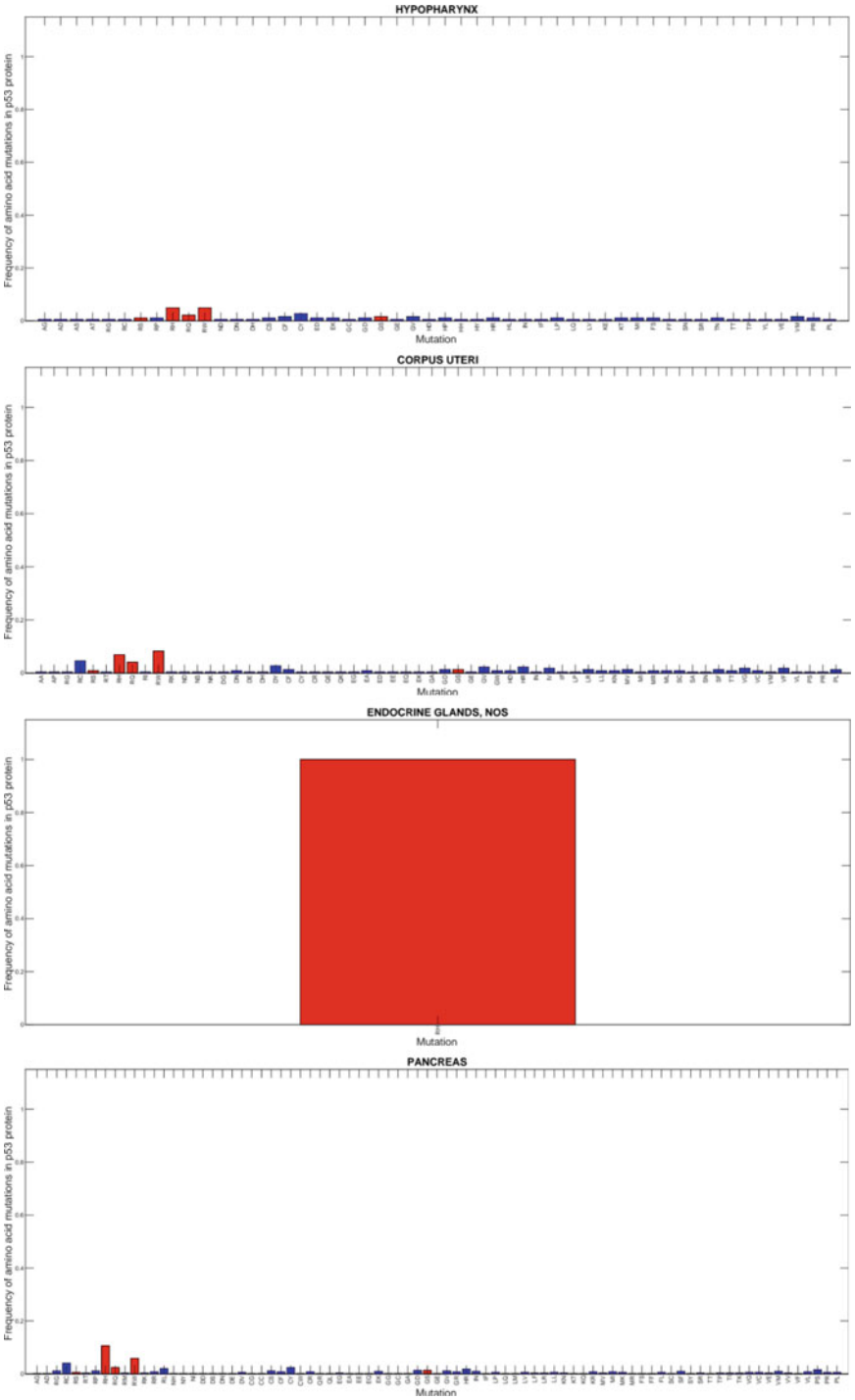

SKIN
ADRENAL GLAND
STOMACH
PROSTATE
Frequency of amino acid mutations in p53 protein
Mutation

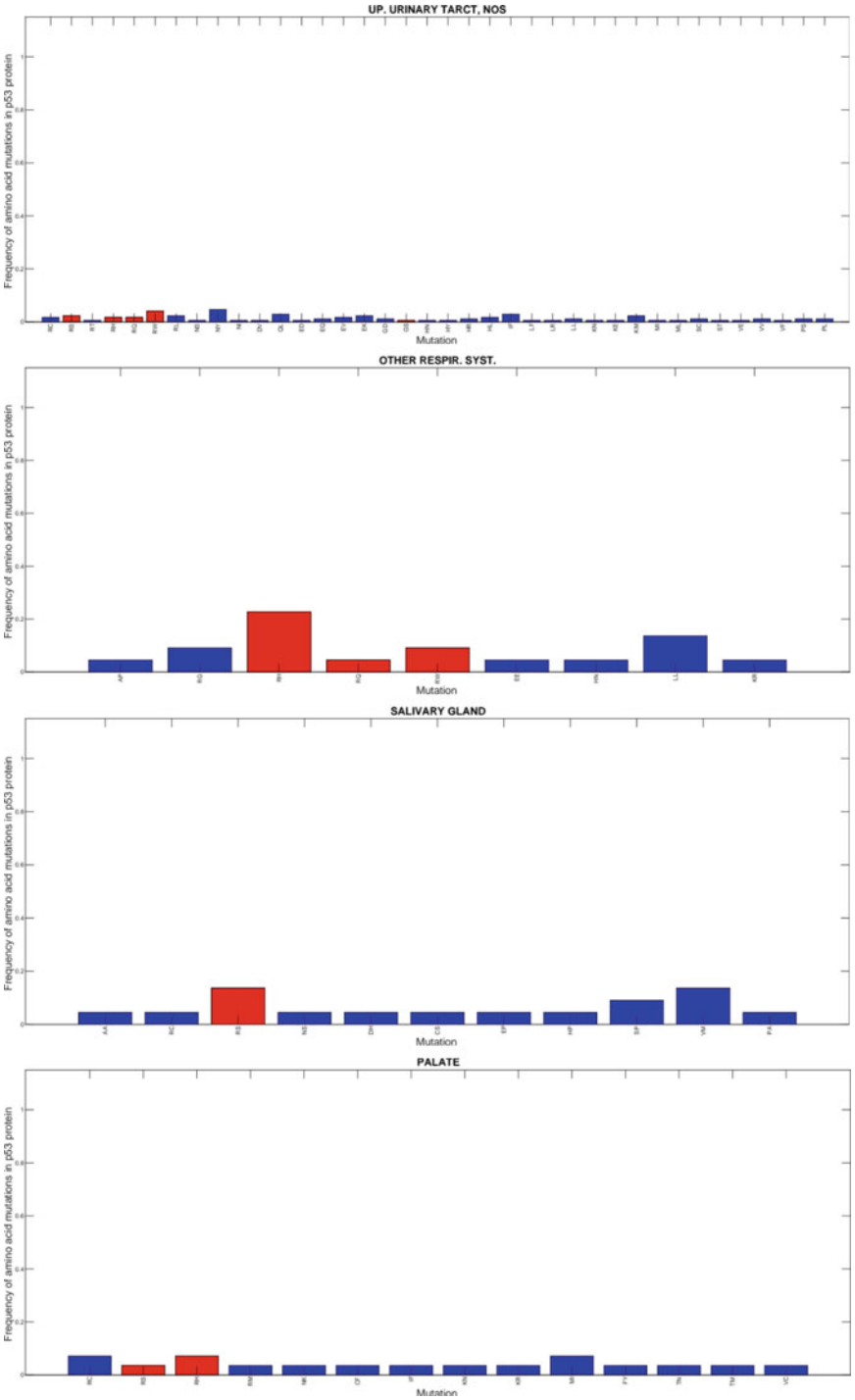

RENAL PELVIS
Frequency of amino acid mutations in p53 protein
Mutation
OVARY
Frequency of amino acid mutations in p53 protein
Mutation
NERVES
Frequency of amino acid mutations in p53 protein
Mutation
EYE AND ADNEXA
Frequency of amino acid mutations in p53 protein
Mutation

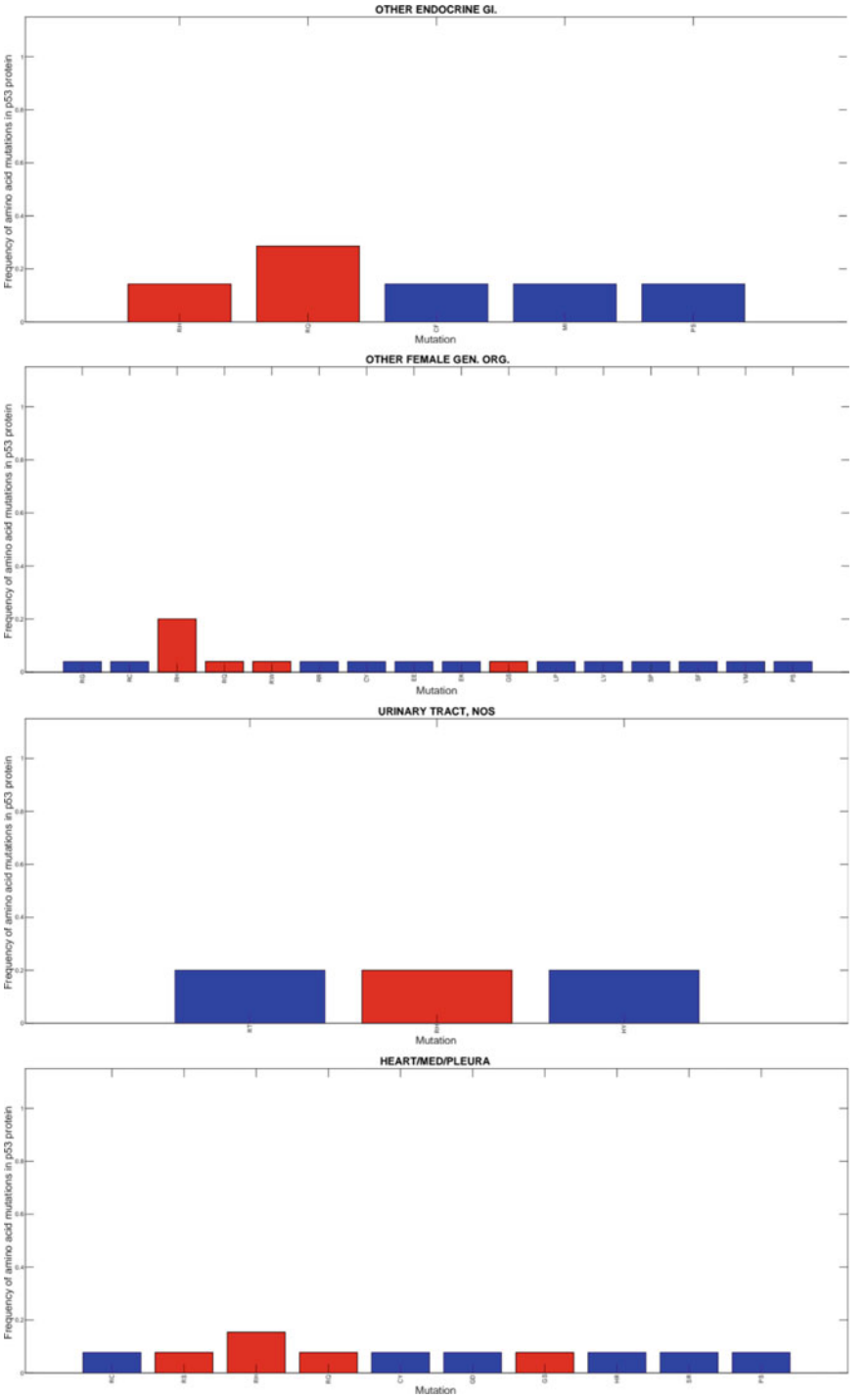

THYMUS
Frequency of amino acid mutations in p53 protein
Mutation
OTHER SITES
Frequency of amino acid mutations in p53 protein
Mutation
LYMPH NODES
Frequency of amino acid mutations in p53 protein
Mutation
OTHER MALE GEN. ORG.
Frequency of amino acid mutations in p53 protein
Mutation

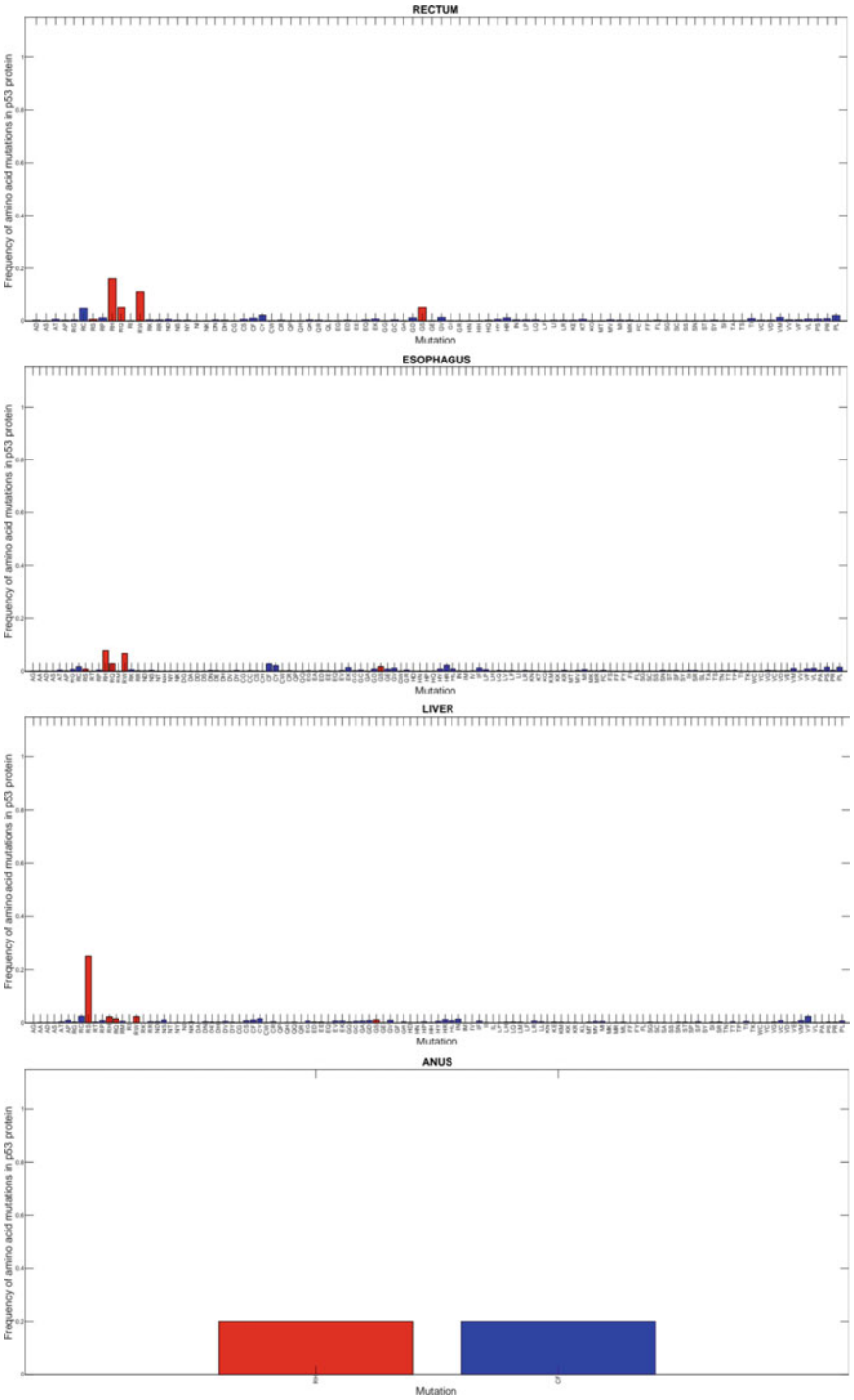

PYRIFORM SINUS
Frequency of amino acid mutations in p53 protein
Mutation
LARYNX
Frequency of amino acid mutations in p53 protein
Mutation
BLADDER
Frequency of amino acid mutations in p53 protein
Mutation
MENINGES
Frequency of amino acid mutations in p53 protein
Mutation

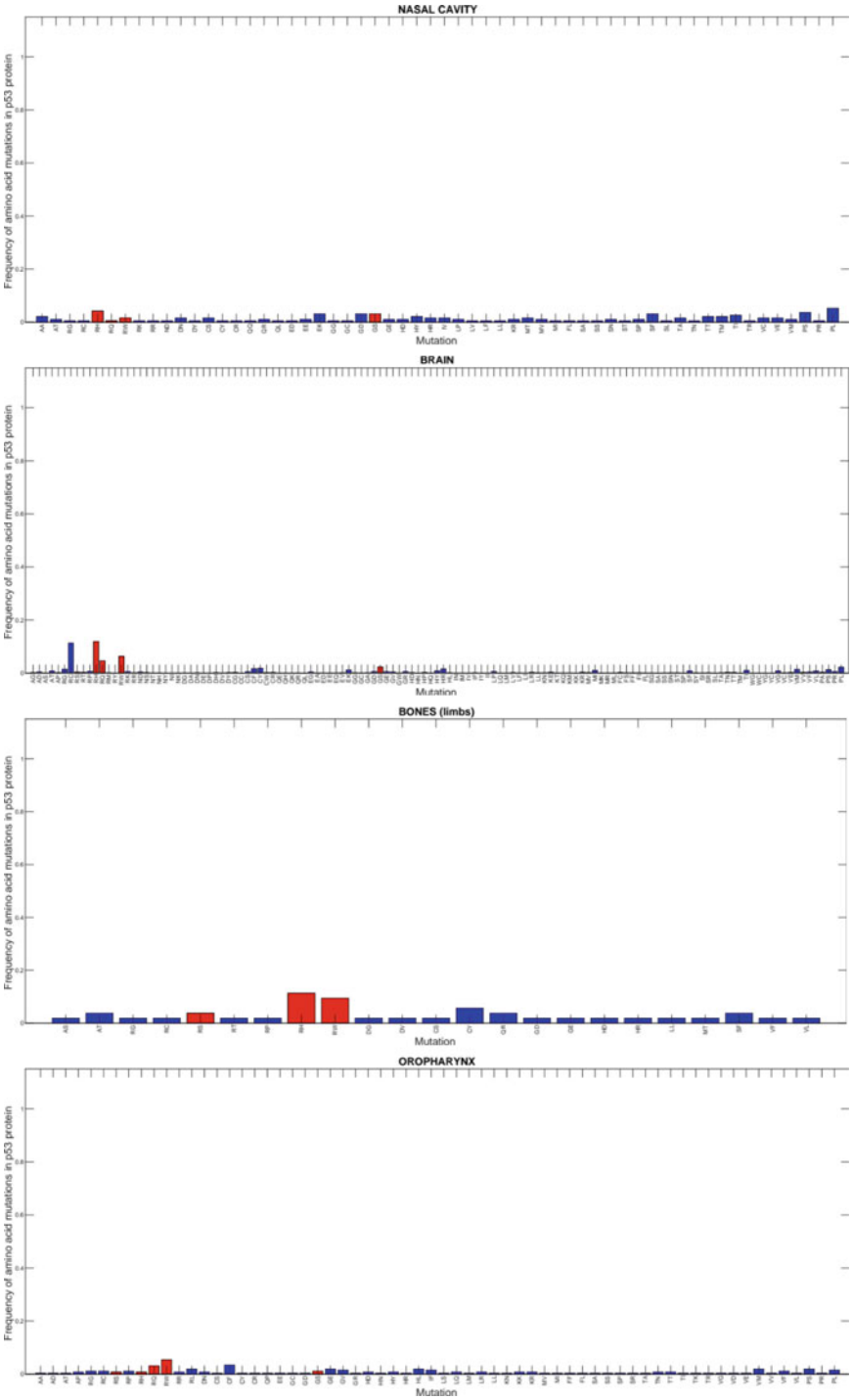

MOUTH (other)
KIDNEY
THYROID
VULVA
Frequency of amino acid mutations in p53 protein
Mutation

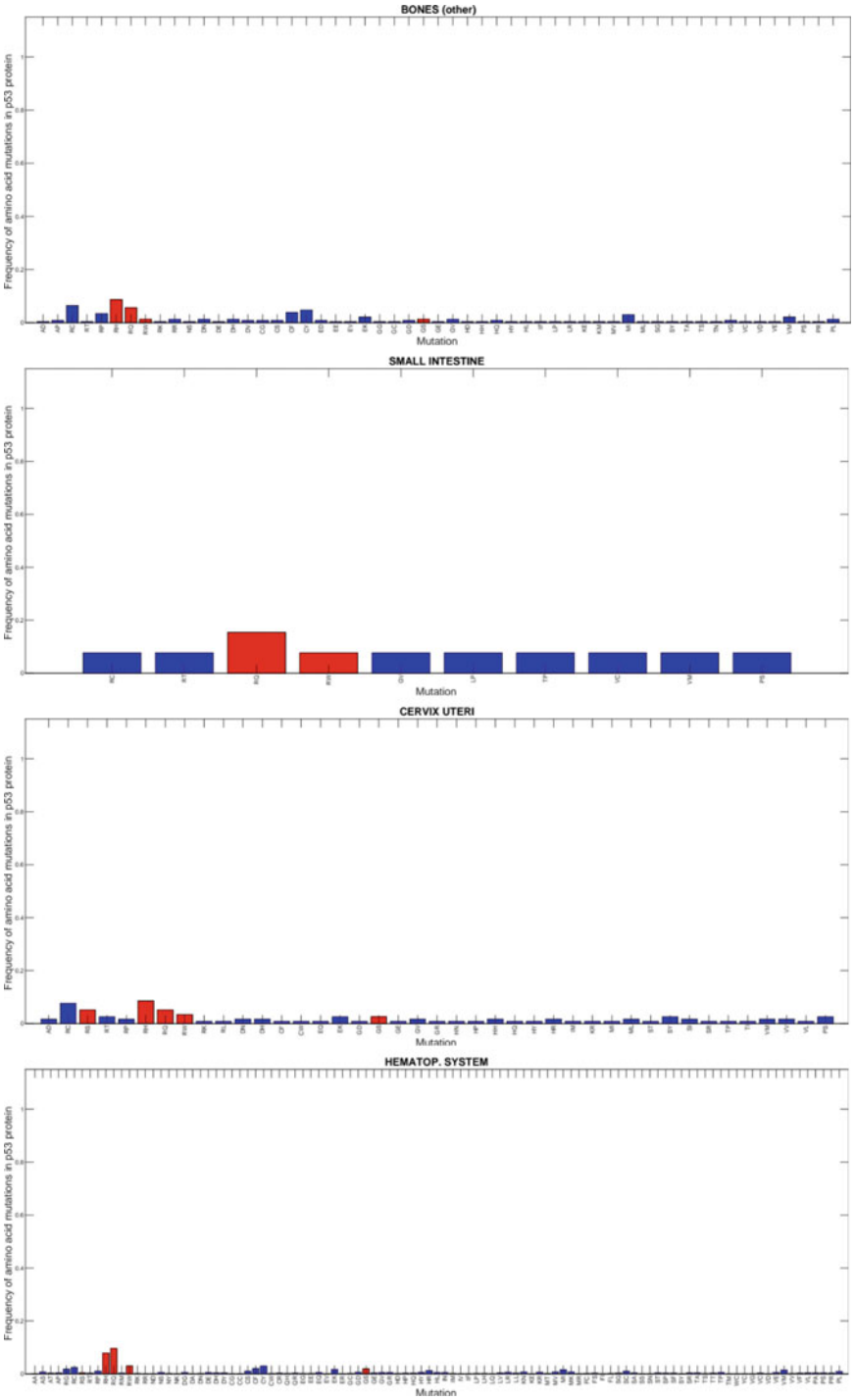

HYPOPHARYNX
Frequency of amino acid mutations in p53 protein
Mutation
CORPUS UTERI
Frequency of amino acid mutations in p53 protein
Mutation
ENDOCRINE GLANDS, NOS
Frequency of amino acid mutations in p53 protein
Mutation
PANCREAS
Frequency of amino acid mutations in p53 protein
Mutation

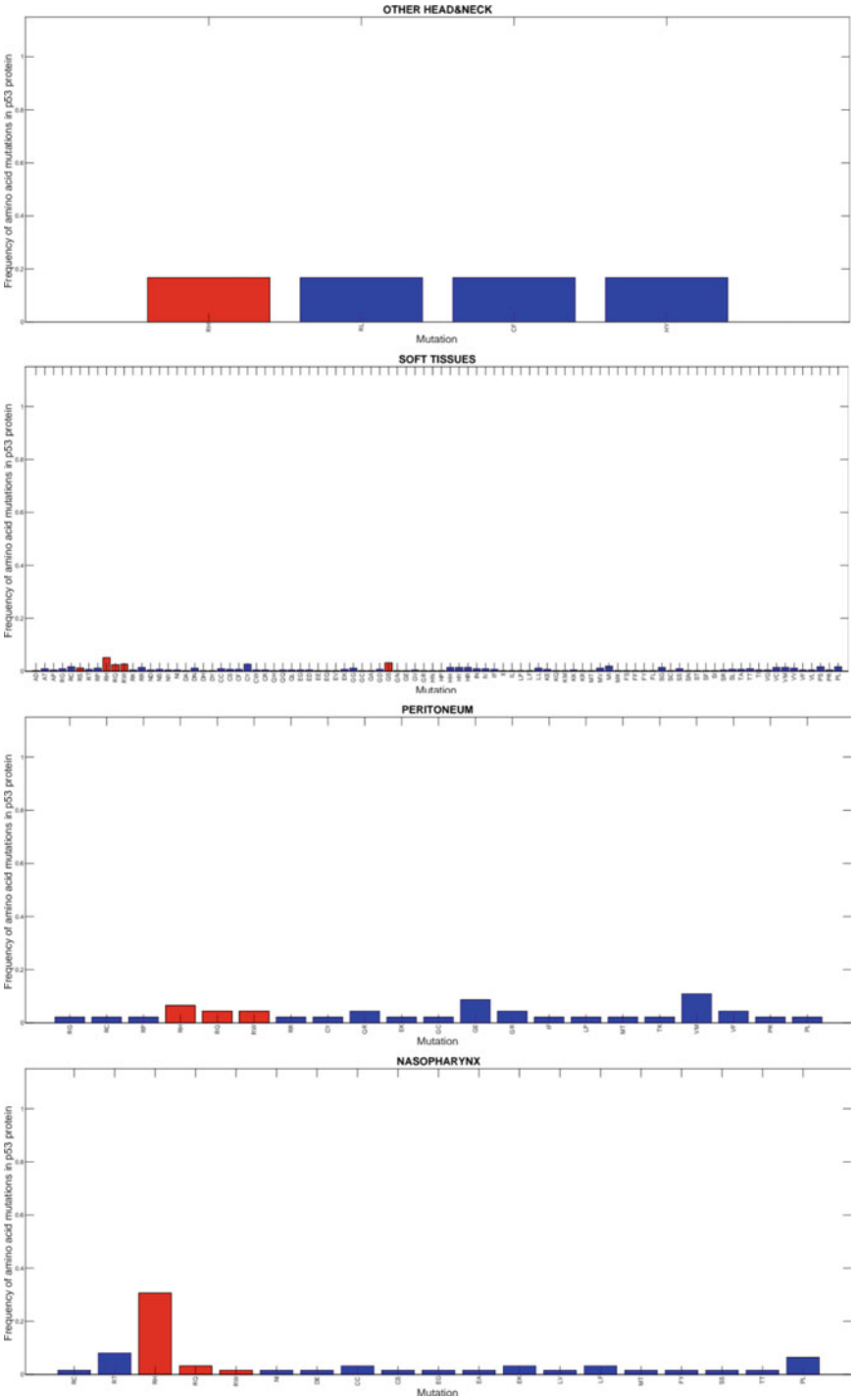

UP. URINARY TARCT, NOS
Frequency of amino acid mutations in p53 protein
Mutation
OTHER RESPIR. SYST.
Frequency of amino acid mutations in p53 protein
Mutation
SALIVARY GLAND
Frequency of amino acid mutations in p53 protein
Mutation
PALATE
Frequency of amino acid mutations in p53 protein
Mutation

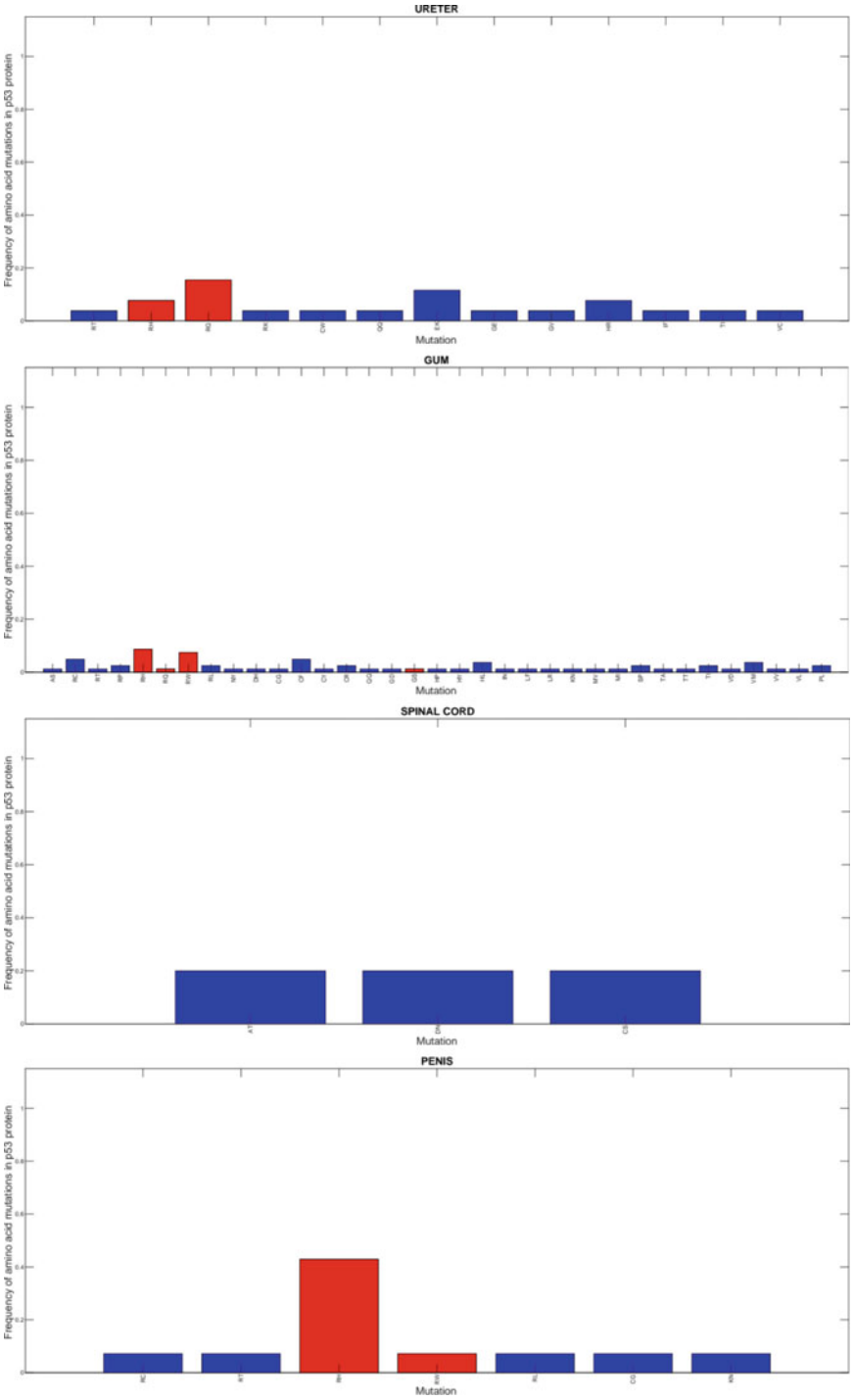

OTHER ENDOCRINE GI.
Frequency of amino acid mutations in p53 protein
Mutation
OTHER FEMALE GEN. ORG.
Frequency of amino acid mutations in p53 protein
Mutation
URINARY TRACT, NOS
Frequency of amino acid mutations in p53 protein
Mutation
HEART/MED/PLEURA
Frequency of amino acid mutations in p53 protein
Mutation

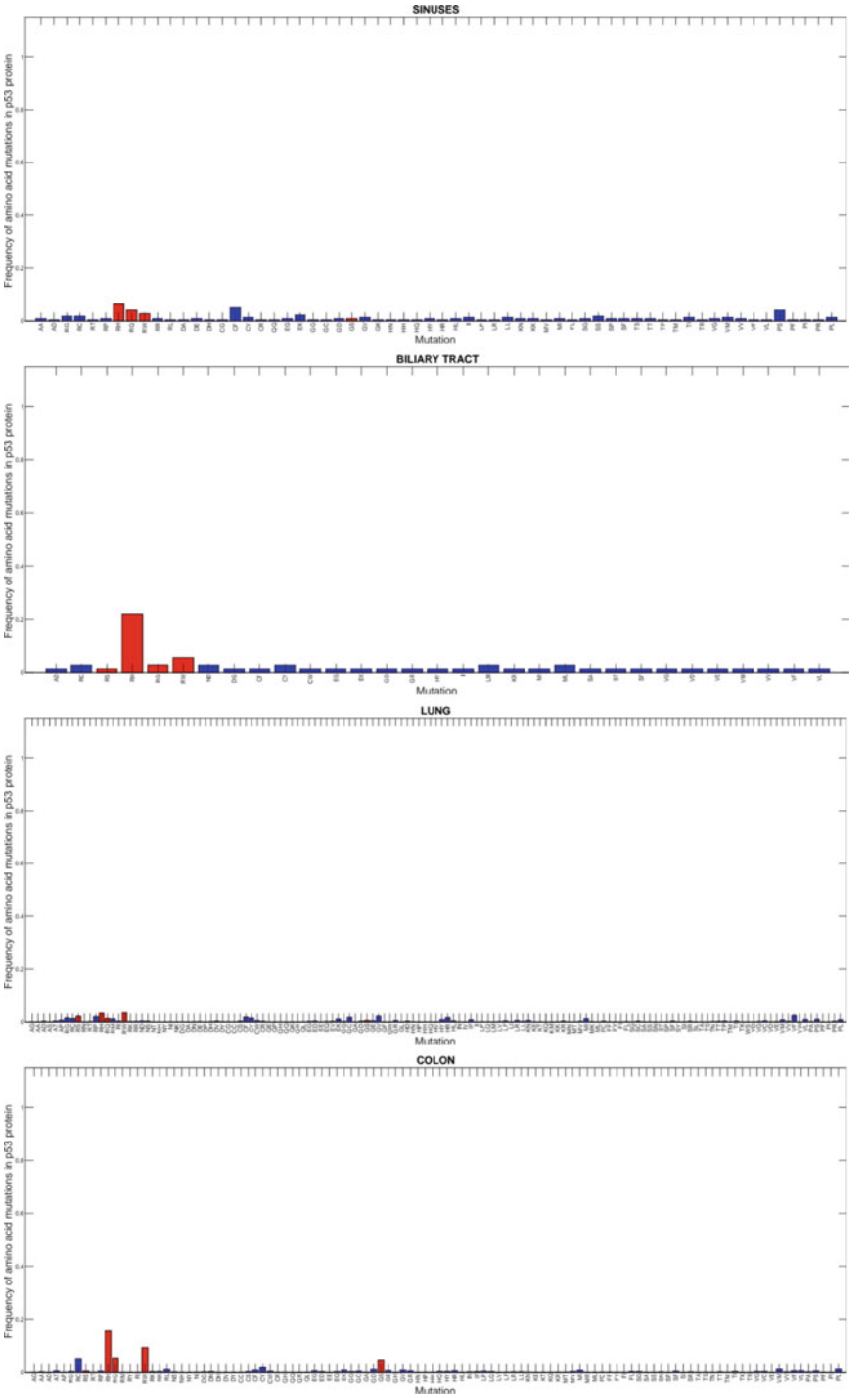

RECTUM
Frequency of amino acid mutations in p53 protein
Mutation
ESOPHAGUS
Frequency of amino acid mutations in p53 protein
Mutation
LIVER
Frequency of amino acid mutations in p53 protein
Mutation
ANUS
Frequency of amino acid mutations in p53 protein
Mutation

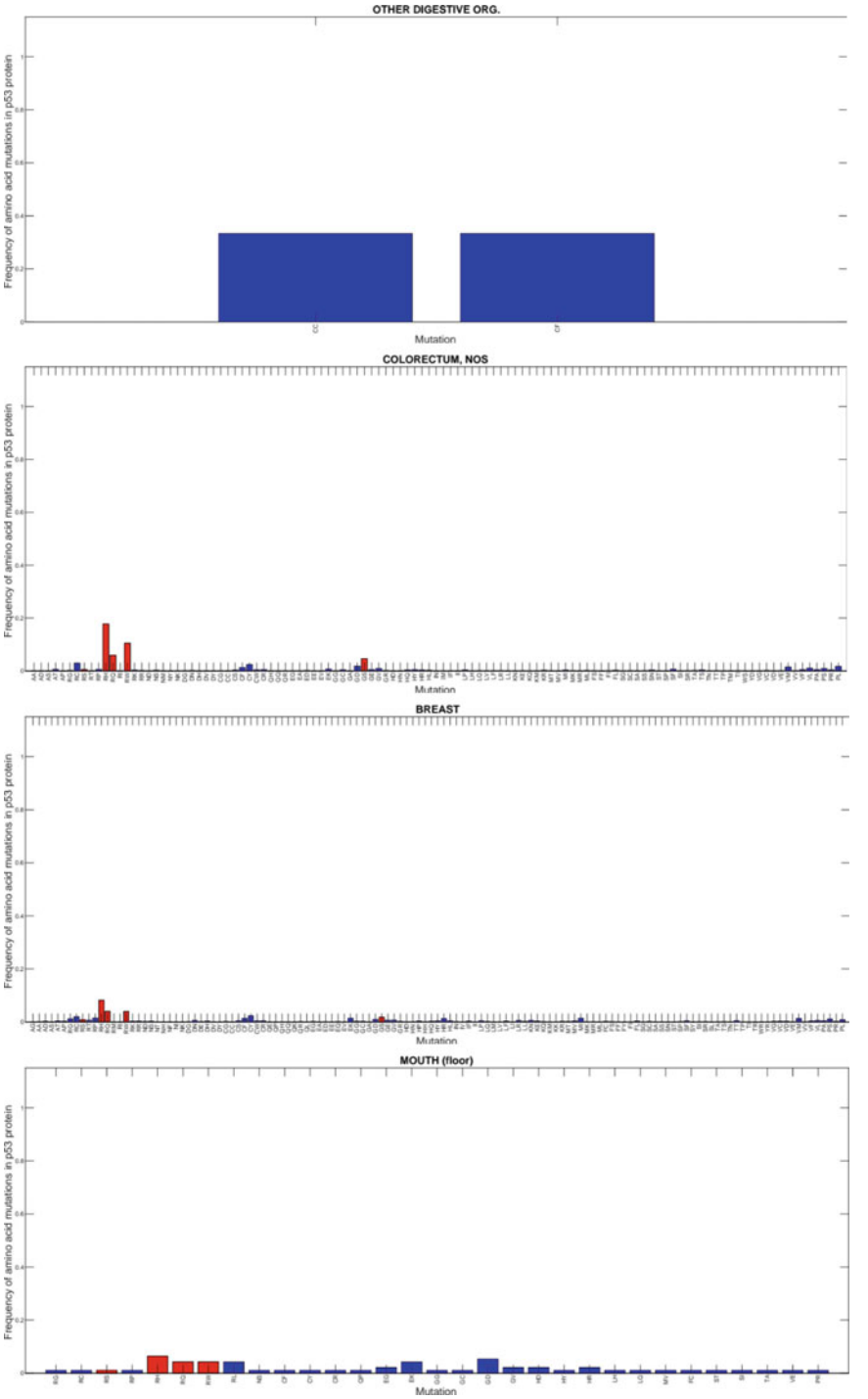

NASAL CAVITY
Frequency of amino acid mutations in p53 protein
Mutation
BRAIN
Frequency of amino acid mutations in p53 protein
Mutation
BONES (limbs)
Frequency of amino acid mutations in p53 protein
Mutation
OROPHARYNX
Frequency of amino acid mutations in p53 protein
Mutation

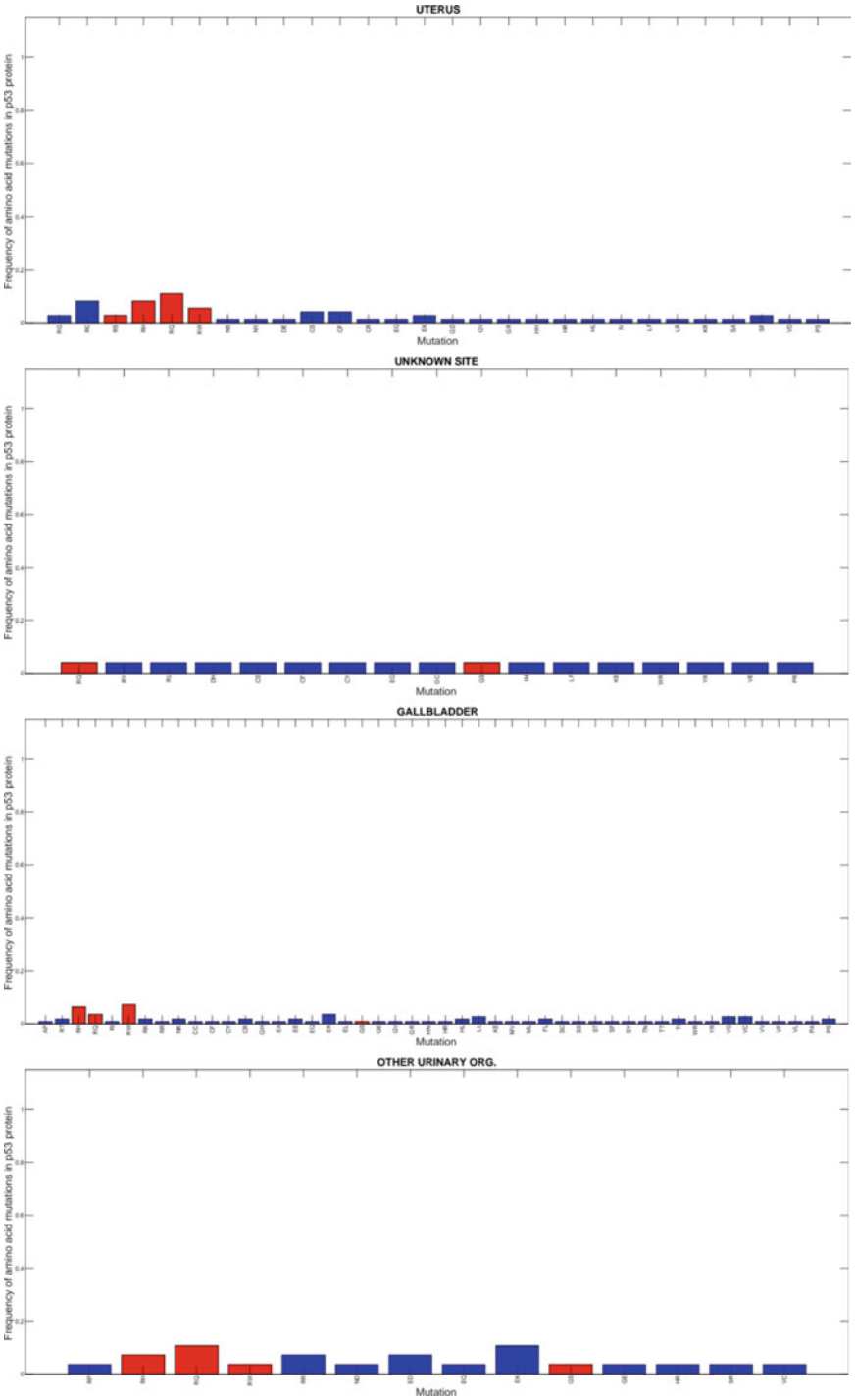

BONES (other)
Frequency of amino acid mutations in p53 protein
Mutation
SMALL INTESTINE
Frequency of amino acid mutations in p53 protein
Mutation
CERVIX UTERI
Frequency of amino acid mutations in p53 protein
Mutation
HEMATOP. SYSTEM
Frequency of amino acid mutations in p53 protein
Mutation

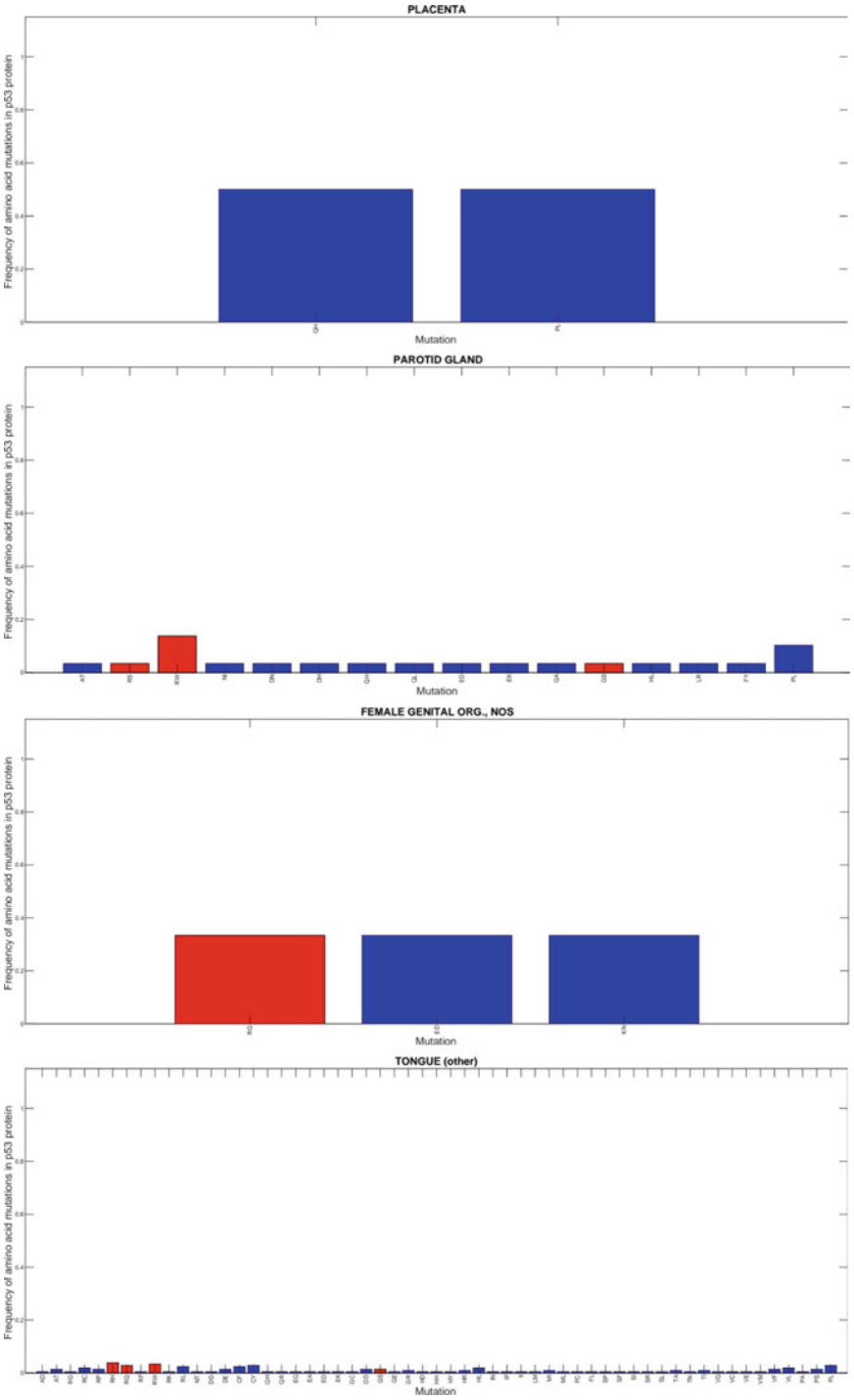

OTHER HEAD&NECK
Frequency of amino acid mutations in p53 protein
Mutation
SOFT TISSUES
Frequency of amino acid mutations in p53 protein
Mutation
PERITONEUM
Frequency of amino acid mutations in p53 protein
Mutation
NASOPHARYNX
Frequency of amino acid mutations in p53 protein
Mutation


URETER
Frequency of amino acid mutations in p53 protein
Mutation
GUM
Frequency of amino acid mutations in p53 protein
Mutation
SPINAL CORD
Frequency of amino acid mutations in p53 protein
Mutation
PENIS
Frequency of amino acid mutations in p53 protein
Mutation


SINUSES
Frequency of amino acid mutations in p53 protein
Mutation
BILIARY TRACT
Frequency of amino acid mutations in p53 protein
Mutation
LUNG
Frequency of amino acid mutations in p53 protein
Mutation
COLON
Frequency of amino acid mutations in p53 protein
Mutation


OTHER DIGESTIVE ORG.
Frequency of amino acid mutations in p53 protein
Mutation
COLORECTUM, NOS
Frequency of amino acid mutations in p53 protein
Mutation
BREAST
Frequency of amino acid mutations in p53 protein
Mutation
MOUTH (floor)
Frequency of amino acid mutations in p53 protein
Mutation


UTERUS
Frequency of amino acid mutations in p53 protein
Mutation
UNKNOWN SITE
Frequency of amino acid mutations in p53 protein
Mutation
GALLBLADDER
Frequency of amino acid mutations in p53 protein
Mutation
OTHER URINARY ORG.
Frequency of amino acid mutations in p53 protein
Mutation


PLACENTA
Frequency of amino acid mutations in p53 protein
Mutation
PAROTID GLAND
Frequency of amino acid mutations in p53 protein
Mutation
FEMALE GENITAL ORG., NOS
Frequency of amino acid mutations in p53 protein
Mutation
TONGUE (other)
Frequency of amino acid mutations in p53 protein
Mutation

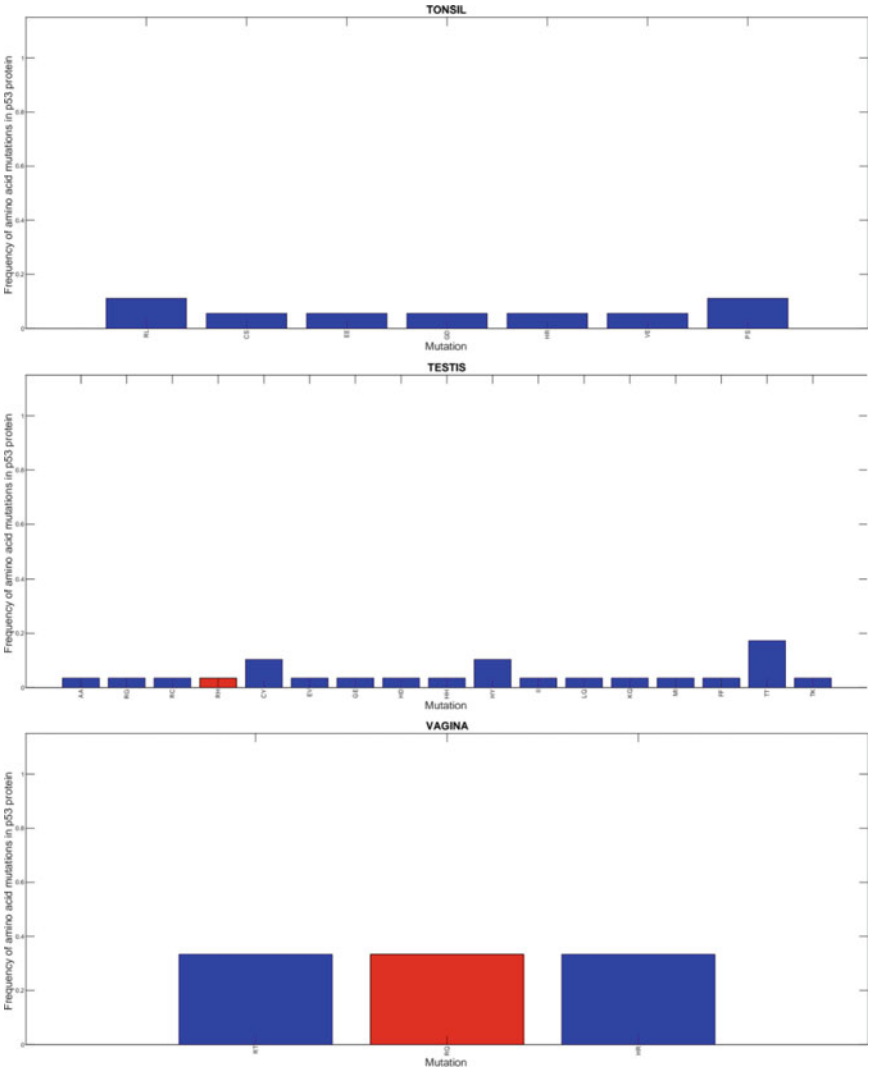
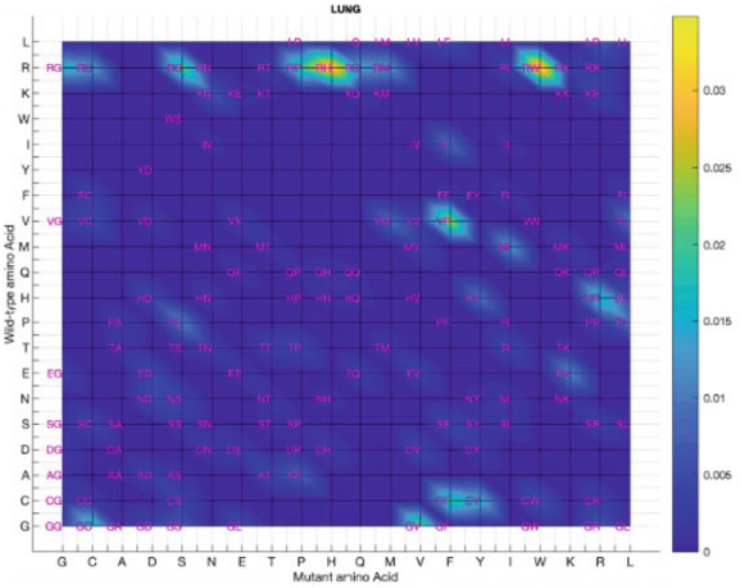


Fig. 9 Frequency of amino acid mutations in the p53 protein in different cancers. The data were extracted using IARC database. Each plot shows the mutations of the p53 protein in specific cancer types. The red bar shows the p53 hotspot mutations. In almost all of the cancer types at least one hotspot mutation exists and it is one the highest frequency mutations (in almost 84% of the studied cases)

(a)



(b)

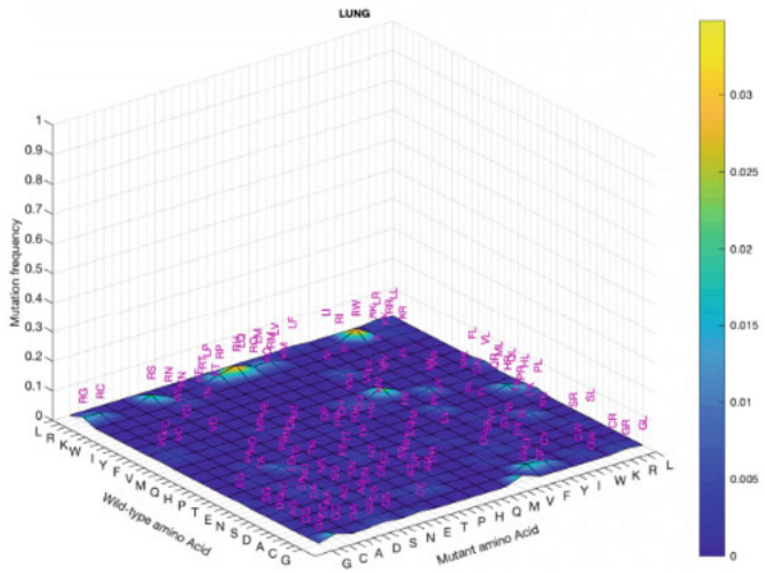


Fig. 10 The 2-D **(a)** and 3-D **(b)** plot of p53 mutation frequency in lung cancer obtained from the IARC database. The color bar changes from blue to yellow, which represents the mutation frequency from 0 to 0.04. Zero means there is no mutation from Wild-Type to that specific mutant reported in the database, and the higher the frequency of mutations, the more yellow it is represented as in both (**a**) and (**b**) plots. For each mutation, the Wild-Type to mutant is represented by the first letter representation of the amino acids shown in pink

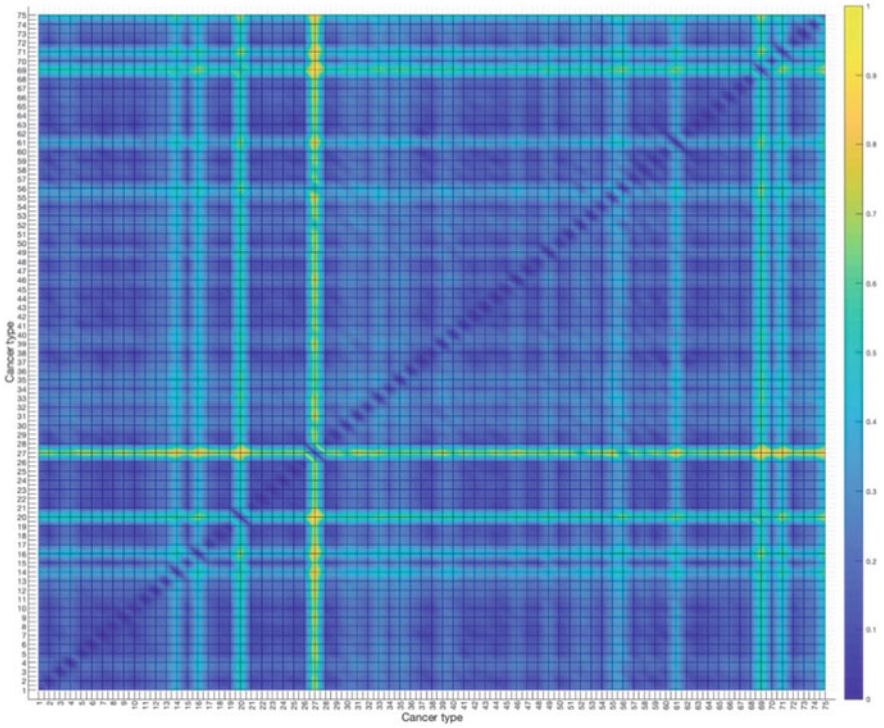


Fig. 11 The dissimilarity factors between different cancer types. Using Eq. (3), a 75 by 75 matrix of dissimilarities has been obtained with respect to each cancer type, $i = j$ columns are zero since they show the dissimilarity factors of each cancer to itself, and $i \neq j$ shows the dissimilarity of each cancer to the rest of the 74 cancer types. Color bar changes from blue to yellow to show the dissimilarity of each cancer to the reference cancer type

References

1. F. Austin, U. Oyarbide, G. Massey, M. Grimes, and S. J. Corey, *Synonymous Mutation in TP53 Results in a Cryptic Splice Site Affecting Its DNA Binding Site in an Adolescent with Two Primary Sarcomas*, Pediatr Blood Cancer **64**, (2017).
2. T. Hu and W. Banzhaf, *Nonsynonymous to Synonymous Substitution Ratio Ka/Ks: Measurement for Rate of Evolution in Evolutionary Computation, in Parallel Problem Solving from Nature – PPSN X*, edited by G. Rudolph, T. Jansen, N. Beume, S. Lucas, and C. Poloni (Springer, Berlin, Heidelberg, 2008), pp. 448–457.
3. K. Karakostis, S. Vadivel Gnanasundram, I. López, A. Thermou, L. Wang, K. Nylander, V. Olivares-Illana, and R. Fåhraeus, *A Single Synonymous Mutation Determines the Phosphorylation and Stability of the Nascent Protein*, J Mol Cell Biol **11**, 187 (2019).
4. E. H. Baugh, H. Ke, A. J. Levine, R. A. Bonneau, and C. S. Chan, *Why Are There Hotspot Mutations in the TP53 Gene in Human Cancers?*, Cell Death Differ **25**, 154 (2018).
5. L. Loewe, *Genetic Mutation*, Nature Education **1**, 113 (2008).
6. A. C. Joerger and A. R. Fersht, *The Tumor Suppressor P53: From Structures to Drug Discovery*, Cold Spring Harb Perspect Biol **2**, (2010).

7. S. Surget, M. P. Khoury, and J.-C. Bourdon, *Uncovering the Role of P53 Splice Variants in Human Malignancy: A Clinical Perspective*, https://doi.org/10.2147/OTT.S53876.
8. F. Perri, S. Pisconti, and G. Della Vittoria Scarpati, *P53 Mutations and Cancer: A Tight Linkage*, Ann Transl Med **4**, (2016).
9. C. W. Lee, M. Arai, M. A. Martinez-Yamout, H. J. Dyson, and P. E. Wright, *Mapping the Interactions of the P53 Transactivation Domain with the KIX Domain of CBP*, Biochemistry **48**, 2115 (2009).
10. C. W. Lee, M. A. Martinez-Yamout, H. J. Dyson, and P. E. Wright, *Structure of the P53 Transactivation Domain in Complex with the Nuclear Receptor Coactivator Binding Domain of CREB Binding Protein*, Biochemistry **49**, 9964 (2010).
11. N. Raj and L. D. Attardi, *The Transactivation Domains of the P53 Protein*, Cold Spring Harb Perspect Med **7**, (2017).
12. M. Ullah, *P53 Mutational Signature in Cancer*, International Journal of Vaccines & Vaccination **Volume 4**, 00070 (2017).
13. K. K. Walker and A. J. Levine, *Identification of a Novel P53 Functional Domain That Is Necessary for Efficient Growth Suppression*, Proc. Natl. Acad. Sci. U.S.A. **93**, 15335 (1996).
14. Y. Cho, S. Gorina, P. D. Jeffrey, and N. P. Pavletich, *Crystal Structure of a P53 Tumor Suppressor-DNA Complex: Understanding Tumorigenic Mutations*, Science **265**, 346 (1994).
15. R. Beckerman and C. Prives, *Transcriptional Regulation by P53*, Cold Spring Harb Perspect Biol **2**, a000935 (2010).
16. P. Chène, *The Role of Tetramerization in P53 Function*, Oncogene **20**, 2611 (2001).
17. G. Gaglia, Y. Guan, J. V. Shah, and G. Lahav, *Activation and Control of P53 Tetramerization in Individual Living Cells*, Proc. Natl. Acad. Sci. U.S.A. **110**, 15497 (2013).
18. O. Laptenko, D. R. Tong, J. Manfredi, and C. Prives, *The Tail That Wags the Dog: How the Disordered C-Terminal Domain Controls the Transcriptional Activities of the P53 Tumor-Suppressor Protein*, Trends Biochem. Sci. **41**, 1022 (2016).
19. A. R. Blanden, X. Yu, S. N. Loh, A. J. Levine, and D. R. Carpizo, *Reactivating Mutant P53 Using Small Molecules as Zinc Metallochaperones: Awakening a Sleeping Giant in Cancer*, Drug Discov Today **20**, 1391 (2015).
20. S. Kogan and D. R. Carpizo, *Zinc Metallochaperones as Mutant P53 Reactivators: A New Paradigm in Cancer Therapeutics*, Cancers (Basel) **10**, (2018).
21. S. N. Loh, *The Missing Zinc: P53 Misfolding and Cancer*, Metallomics **2**, 442 (2010).
22. C. Méplan, M.-J. Richard, and P. Hainaut, *Metalloregulation of the Tumor Suppressor Protein P53: Zinc Mediates the Renaturation of P53 after Exposure to Metal Chelators in Vitro and in Intact Cells*, Oncogene **19**, 46 (2000).
23. X. Yu, S. Kogan, Y. Chen, A. T. Tsang, T. Withers, H. Lin, J. Gilleran, B. Buckley, D. Moore, J. Bertino, C. Chan, S. D. Kimball, S. N. Loh, and D. R. Carpizo, *Zinc Metallochaperones Reactivate Mutant P53 Using an ON/OFF Switch Mechanism: A New Paradigm in Cancer Therapeutics*, Clin Cancer Res **24**, 4505 (2018).
24. L. Bouaoun, D. Sonkin, M. Ardin, M. Hollstein, G. Byrnes, J. Zavadil, and M. Olivier, *TP53 Variations in Human Cancers: New Lessons from the IARC TP53 Database and Genomics Data*, Hum. Mutat. **37**, 865 (2016).
25. K. H. Vousden and X. Lu, *Live or Let Die: The Cell's Response to P53*, Nat Rev Cancer **2**, 594 (2002).
26. S. Madhusudan and M. R. Middleton, *The Emerging Role of DNA Repair Proteins as Predictive, Prognostic and Therapeutic Targets in Cancer*, Cancer Treat. Rev. **31**, 603 (2005).
27. A. Comel, G. Sorrentino, V. Capaci, and G. D. Sal, *The Cytoplasmic Side of P53's Oncosuppressive Activities*, FEBS Letters **588**, 2600 (2014).
28. M. Weinfeld, R. S. Mani, I. Abdou, R. D. Aceytuno, and J. N. M. Glover, *Tidying up Loose Ends: The Role of Polynucleotide Kinase/Phosphatase in DNA Strand Break Repair*, Trends Biochem. Sci. **36**, 262 (2011).
29. D. W. Meek, *Regulation of the P53 Response and Its Relationship to Cancer1*, Biochemical Journal **469**, 325 (2015).

30. L. Gatti and F. Zunino, *Overview of Tumor Cell Chemoresistance Mechanisms*, Methods Mol. Med. **111**, 127 (2005).
31. D. W. Meek and C. W. Anderson, *Posttranslational Modification of P53: Cooperative Integrators of Function*, Cold Spring Harb Perspect Biol **1**, a000950 (2009).
32. S. Kachalaki, M. Ebrahimi, L. Mohamed Khosroshahi, S. Mohammadinejad, and B. Baradaran, *Cancer Chemoresistance; Biochemical and Molecular Aspects: A Brief Overview*, Eur J Pharm Sci **89**, 20 (2016).
33. C. L. Brooks and W. Gu, *The Impact of Acetylation and Deacetylation on the P53 Pathway*, Protein Cell **2**, 456 (2011).
34. A. J. Levine and M. Oren, *The First 30 Years of P53: Growing Ever More Complex*, Nat. Rev. Cancer **9**, 749 (2009).
35. K. A. Boehme and C. Blattner, *Regulation of P53 – Insights into a Complex Process*, Crit. Rev. Biochem. Mol. Biol. **44**, 367 (2009).
36. I. Goldstein, V. Marcel, M. Olivier, M. Oren, V. Rotter, and P. Hainaut, Understanding Wild-Type and Mutant P53 Activities in Human Cancer: New Landmarks on the Way to Targeted Therapies, Cancer Gene Ther. **18**, 2 (2011).
37. S.-J. Wang and W. Gu, *To Be, or Not to Be: Functional Dilemma of P53 Metabolic Regulation*, Curr Opin Oncol **26**, 78 (2014).
38. S. I. Omar and J. Tuszynski, *Ranking the Binding Energies of P53 Mutant Activators and Their ADMET Properties*, Chem Biol Drug Des **86**, 163 (2015).
39. S. I. Omar, M. G. Lepre, U. Morbiducci, M. A. Deriu, and J. A. Tuszynski, *Virtual Screening Using Covalent Docking to Find Activators for G245S Mutant P53*, PLoS One **13**, (2018).
40. G. Chillemi, S. Kehrloesser, F. Bernassola, A. Desideri, V. Dötsch, A. J. Levine, and G. Melino, *Structural Evolution and Dynamics of the P53 Proteins*, Cold Spring Harb Perspect Med **7**, (2017).
41. C. D. Wassman, R. Baronio, Ö. Demir, B. D. Wallentine, C.-K. Chen, L. V. Hall, F. Salehi, D.-W. Lin, B. P. Chung, G. W. Hatfield, A. Richard Chamberlin, H. Luecke, R. H. Lathrop, P. Kaiser, and R. E. Amaro, *Computational Identification of a Transiently Open L1/S3 Pocket for Reactivation of Mutant P53*, Nat Commun **4**, 1407 (2013).
42. M. G. Lepre, S. I. Omar, G. Grasso, U. Morbiducci, M. A. Deriu, and J. A. Tuszynski, *Insights into the Effect of the G245S Single Point Mutation on the Structure of P53 and the Binding of the Protein to DNA*, Molecules **22**, (2017).
43. *P14 ARF Links the Tumour Suppressors RB and P53* | Nature, https://www.nature.com/articles/25867.
44. S. Bell, C. Klein, L. Müller, S. Hansen, and J. Buchner, *P53 Contains Large Unstructured Regions in Its Native State*, Journal of Molecular Biology **322**, 917 (2002).
45. J. R. Bischoff, D. H. Kirn, A. Williams, C. Heise, S. Horn, M. Muna, L. Ng, J. A. Nye, A. Sampson-Johannes, A. Fattaey, and F. McCormick, *An Adenovirus Mutant That Replicates Selectively in P53-Deficient Human Tumor Cells*, Science **274**, 373 (1996).
46. M. V. Blagosklonny, *P53: An Ubiquitous Target of Anticancer Drugs*, Int. J. Cancer **98**, 161 (2002).
47. F. McCormick, *Cancer Gene Therapy: Fringe or Cutting Edge?*, Nat Rev Cancer **1**, 130 (2001).
48. T. Strachan and A. P. Read, *Human Molecular Genetics 2* (Wiley-Liss, New York, 1999).
49. B. Vogelstein, D. Lane, and A. J. Levine, *Surfing the P53 Network*, Nature **408**, 307 (2000).
50. M. Tollis, A. M. Boddy, and C. C. Maley, *Peto's Paradox: How Has Evolution Solved the Problem of Cancer Prevention?*, BMC Biology **15**, 60 (2017).
51. L. M. Abegglen, A. F. Caulin, A. Chan, K. Lee, R. Robinson, M. S. Campbell, W. K. Kiso, D. L. Schmitt, P. J. Waddell, S. Bhaskara, S. T. Jensen, C. C. Maley, and J. D. Schiffman, *Potential Mechanisms for Cancer Resistance in Elephants and Comparative Cellular Response to DNA Damage in Humans*, JAMA **314**, 1850 (2015).
52. S. S. Mello and L. D. Attardi, *Not All P53 Gain-of-Function Mutants Are Created Equal*, Cell Death Differ. **20**, 855 (2013).
53. M. Ullah, *P53 Mutational Signature in Cancer*, International Journal of Vaccines & Vaccination (2017).

54. M. Olivier, M. Hollstein, and P. Hainaut, *TP53 Mutations in Human Cancers: Origins, Consequences, and Clinical Use*, Cold Spring Harb Perspect Biol **2**, a001008 (2010).
55. N. Shirai, T. Tsukamoto, M. Yamamoto, T. Iidaka, H. Sakai, T. Yanai, T. Masegi, L. A. Donehower, and M. Tatematsu, *Elevated Susceptibility of the P53 Knockout Mouse Esophagus to Methyl-N-Amylnitrosamine Carcinogenesis*, Carcinogenesis **23**, 1541 (2002).
56. M. Yamamoto, T. Tsukamoto, H. Sakai, N. Shirai, H. Ohgaki, C. Furihata, L. A. Donehower, K. Yoshida, and M. Tatematsu, *P53 Knockout Mice (-/-) Are More Susceptible than (+/-) or (+/+) Mice to N-Methyl-N-Nitrosourea Stomach Carcinogenesis*, Carcinogenesis **21**, 1891 (2000).
57. A. C. Blackburn and D. J. Jerry, *Knockout and Transgenic Mice of Trp53: What Have We Learned about P53 in Breast Cancer?*, Breast Cancer Res **4**, 101 (2002).
58. S. North, F. El-Ghissassi, O. Pluquet, G. Verhaegh, and P. Hainaut, *The Cytoprotective Aminothiol WR1065 Activates P21 Waf-1 and down Regulates Cell Cycle Progression through a P53-Dependent Pathway*, Oncogene **19**, 9 (2000).
59. S. Emamzadah, L. Tropia, and T. D. Halazonetis, *Crystal Structure of a Multidomain Human P53 Tetramer Bound to the Natural CDKN1A (P21) P53-Response Element*, Mol. Cancer Res. **9**, 1493 (2011).
60. A. C. Minella, J. Swanger, E. Bryant, M. Welcker, H. Hwang, and B. E. Clurman, *P53 and P21 Form an Inducible Barrier That Protects Cells against Cyclin E-Cdk2 Deregulation*, Curr. Biol. **12**, 1817 (2002).
61. G. He, Z. H. Siddik, Z. Huang, R. Wang, J. Koomen, R. Kobayashi, A. R. Khokhar, and J. Kuang, *Induction of P21 by P53 Following DNA Damage Inhibits Both Cdk4 and Cdk2 Activities*, Oncogene **24**, 2929 (2005).
62. N. P. Pavletich, K. A. Chambers, and C. O. Pabo, *The DNA-Binding Domain of P53 Contains the Four Conserved Regions and the Major Mutation Hot Spots*, Genes Dev. **7**, 2556 (1993).
63. D. N. Cooper and Nature Publishing Group, *Nature Encyclopedia of the Human Genome* (Nature Pub. Group, London; New York, 2003).
64. A. E. Teschendorff and S. Severini, *Increased Entropy of Signal Transduction in the Cancer Metastasis Phenotype*, BMC Syst Biol **4**, 104 (2010).
65. K. Kayser, G. Kayser, S. Eichhorn, U. Biechele, M. Altiner, H. Kaltner, F. Y. Zeng, E. V. Vlasova, N. V. Bovin, and H. J. Gabius, *Association of Prognosis in Surgically Treated Lung Cancer Patients with Cytometric, Histometric and Ligand Histochemical Properties: With an Emphasis on Structural Entropy*, Anal. Quant. Cytol. Histol. **20**, 313 (1998).
66. U. Agrell, *Draft of a General Stochastic Theory of Cancer and Its Possible Experimental Verification with Monoclonal Multiplication of Repairing and Immunological Systems*, Med. Hypotheses **20**, 261 (1986).
67. J. West, G. Bianconi, S. Severini, and A. E. Teschendorff, *Differential Network Entropy Reveals Cancer System Hallmarks*, Sci Rep **2**, (2012).
68. R. Berretta and P. Moscato, *Cancer Biomarker Discovery: The Entropic Hallmark*, PLoS One **5**, (2010).
69. W. N. van Wieringen and A. W. van der Vaart, *Statistical Analysis of the Cancer Cell's Molecular Entropy Using High-Throughput Data*, Bioinformatics **27**, 556 (2011).
70. J. E. Dumont, S. Dremier, I. Pirson, and C. Maenhaut, *Cross Signaling, Cell Specificity, and Physiology*, Am. J. Physiol., Cell Physiol. **283**, C2 (2002).
71. J. E. Riggs, *Carcinogenesis, Genetic Instability and Genomic Entropy: Insight Derived from Malignant Brain Tumor Age Specific Mortality Rate Dynamics*, J. Theor. Biol. **170**, 331 (1994).
72. M. Tarabichi, A. Antoniou, M. Saiselet, J. M. Pita, G. Andry, J. E. Dumont, V. Detours, and C. Maenhaut, *Systems Biology of Cancer: Entropy, Disorder, and Selection-Driven Evolution to Independence, Invasion and "Swarm Intelligence"*, Cancer Metastasis Rev **32**, 403 (2013).
73. K. Metze, R. L. Adam, G. Kayser, and K. Kayser, *Pathophysiology of Cancer and the Entropy Concept, in Model-Based Reasoning in Science and Technology*, edited by L. Magnani, W. Carnielli, and C. Pizzi, Vol. 314 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2010), pp. 199–206.
74. K. K, K. G, and M. K, *The Concept of Structural Entropy in Tissue-Based Diagnosis.*, Anal Quant Cytol Histol **29**, 296 (2007).

75. R. A. Gatenby and B. R. Frieden, *Information Dynamics in Carcinogenesis and Tumor Growth*, Mutat. Res. **568**, 259 (2004).
76. M. C. Daly and I. M. Paquette, *Surveillance, Epidemiology, and End Results (SEER) and SEER-Medicare Databases: Use in Clinical Research for Improving Colorectal Cancer Outcomes*, Clin Colon Rectal Surg **32**, 61 (2019).
77. *SEER Datasets*, www.seer.cancer.gov/.
78. S. Lukman, D. P. Lane, and C. S. Verma, *Mapping the Structural and Dynamical Features of Multiple P53 DNA Binding Domains: Insights into Loop 1 Intrinsic Dynamics*, PLoS ONE **8**, e80221 (2013).

Oscillations in Biological Systems

Sándor Kovács

1 Introduction

As it is well known, many physical, chemical and biological phenomena are modelled by parabolic equations, among these one of the most frequently examined type is the reaction-diffusion equation. One of the fascinating features of these equations is the variety of special types of solutions they exhibit. Certain systems of this type have, for example, travelling wave solutions or rotating waves (cf. [14]) or via bifurcation analysis one can find a new class of solutions (cf. [13]).

In this chapter we consider the autonomous systems of reaction-diffusion equations

$$\mathbf{u}_t = D\Delta_{\mathbf{r}}\mathbf{u} + \mathbf{f} \circ (\mathbf{u}, \mu), \tag{1}$$

on $\Omega \times \mathbb{R}_0^+ \ni (\mathbf{r}, t)$, with the usual zero flux boundary and non-negative initial condition

$$(\mathbf{n} \cdot \nabla_{\mathbf{r}})\mathbf{u} = \mathbf{0} \qquad \text{on} \qquad \partial\Omega \times \mathbb{R}_0^+, \tag{2}$$

and

$$\mathbf{u}(\cdot, 0) = \mathbf{u}_0(\cdot) \qquad \text{on} \qquad \overline{\Omega} \times \{0\}, \tag{3}$$

where D is a positive diagonal matrix:

S. Kovács (✉)
Department of Numerical Analysis, Eötvös Loránd University, Budapest, Hungary
e-mail: alex@ludens.elte.hu

R. P. Mondaini (ed.), *Trends in Biomathematics: Stability and Oscillations in Environmental, Social, and Biological Models*,
https://doi.org/10.1007/978-3-031-12515-7_4

$$D := \operatorname{diag}(d_1, \ldots, d_n),$$

the kinetic function

$$\mathbf{f} = (f_1, \ldots, f_n) : \mathbb{R}^n \times I \to \mathbb{R}^n$$

belongs to $\mathfrak{C}^1$, μ is a parameter in an open interval $I \subset \mathbb{R}$, Ω is a bounded domain in $\mathbb{R}^n$ with piecewise smooth boundary, $\mathbf{n}$ is the outer unit normal to $\partial\Omega$ and $\mathbf{u}_0$ is a bounded non-negative, resp. not identically vanishing smooth function.

Insomuch as system (1) is biologically motivated it is necessary to show that (1) is biologically well-posed. Usually, this means positivity, resp. dissipativeness, i.e.

- the solution

$$\mathbf{\Phi} = (\Phi_1, \ldots, \Phi_n) \in \overline{\Omega} \times \mathbb{R}_0^+ \to \mathbb{R}^n$$

  of (1) with non-negative initial data

$$\mathbf{u}_0 = (u_0^1, \ldots, u_0^n) \qquad \text{with} \qquad u_0^i \not\equiv 0 \quad (i \in \{1, \ldots, n\})$$

  remains non-negative for all $t \geq 0$ in their domain of existence, resp.
- all solutions of system (1) are bounded and therefore defined for all $t \geq 0$.

The first requirement can be formulated as follows: the positive quadrant of the phase space

$$\left\{\mathbf{u} = (u_1, \ldots, u_n) \in \mathbb{R}^n : \ u_k \geq 0 \ (k \in \{1, \ldots, n\}\right\} \tag{4}$$

is (positively) invariant. This motivates the following

Definition 1.1 *A closed subset $\Sigma \subset \mathbb{R}^n$ (positively) invariant region for the local solution defined by (1), if for suitable $T > 0$ any solution $\mathbf{\Phi}$ having all of its boundary and initial values in Σ satisfies*

$$\mathbf{\Phi}(\boldsymbol{r}, t) \in \Sigma \qquad ((\boldsymbol{r}, t) \in \overline{\Omega} \times [0, T)).$$

It is obvious that the set Σ in (4) is a closed subset.

In [5] one can find the following fundamental result about the existence of (positively) invariant region.

Theorem 1.1 *Let $m \in \mathbb{N}$ and consider the region Σ of the form*

$$\Sigma := \bigcap_{k=1}^{m} \{\boldsymbol{r} \in U : \ G_i(\boldsymbol{r}) \leq 0\},$$

where $U \subset \mathbb{R}^n$ *is an open subset and* $G_i : \mathbb{R}^n \to \mathbb{R}$ *are smooth functions* $(i \in \{1, \ldots, m\})$ *whose gradient* ∇G_i *never vanishes. If at each point* $\boldsymbol{r} \in \partial\Sigma$ *we have for all* $i \in \{1, \ldots, m\}$*:*

(i) $\nabla G_i(\boldsymbol{r})$ *is a left eigenvector of the diffusion matrix* D*;*
(ii) the functions G_i *are quasi-convex, i.e. for all* $\boldsymbol{r} \in U$*, resp. for all* $\boldsymbol{s} \in \mathbb{R}^n$ *the equality* $\langle \nabla G_i(\boldsymbol{r}), \boldsymbol{s} \rangle = 0$ *implies* $\langle \boldsymbol{s}, \nabla^2 G_i(\boldsymbol{r})\boldsymbol{s} \rangle \geq 0$*;*
(iii) $\langle \nabla G_i(\boldsymbol{r}), \boldsymbol{f}(\boldsymbol{r}, \mu) \rangle < 0$ $(\mu \in I)$

then Σ *is positively invariant for system* (1)*.*

As an example we show that the region

$$\Sigma := \left\{ (n, T) \in \mathbb{R}^2 : \ 0 \leq n \leq a, \ \alpha \leq T \right\}$$

is an invariant region for the parabolic system

$$\partial_t n = k_1 \Delta_{\mathbf{r}} n - n \exp(-E/RT), \qquad \partial_t T = k_2 \Delta_{\mathbf{r}} T + Qn \exp(-E/RT)$$

arising in the theory of combustion (cf. [10]) where the quantities T and n denote the temperature and concentration, respectively, of a combustible substance and k_1, k_2, N, E and Q are positive constants, $0 < n(\mathbf{r}, 0) < a$, $0 < \alpha \leq T(\mathbf{r}, 0)$. Indeed, for

$$f_1(n, T, \mu) := -n \exp(-E/RT), \qquad f_2(n, T) := Qn \exp(-E/RT)$$

and

$$G_1(n, T, \mu) := n - a, \quad G_2(n, T) := -n - \varepsilon \ (\varepsilon > 0) \quad \text{resp.} \quad G_3(n, T) := \alpha - T$$

where $\mu \in \{k_1, k_2, N, E, \varepsilon\}$ we have

$$\langle \nabla G_1, (f_1, f_2) \rangle_{n=a} = -a \exp(-E/RT) < 0,$$

$$\langle \nabla G_2, (f_1, f_2) \rangle_{n=-\epsilon} = -\varepsilon \exp(-E/RT) < 0,$$

resp.

$$\langle \nabla G_3, (f_1, f_2) \rangle_{n>0,\, T=\alpha} = -Qn \exp(-E/R\alpha) < 0.$$

As a further example we deal with the reaction-diffusion system proposed by A. Lemarchand and B. Nowakowski (cf. [18]) which describes the macroscopic evolution of two variable concentrations A and B and is given by the two deterministic equation

$$\left.\begin{aligned} \partial_t A &= d_A \Delta_{\mathbf{r}} A + f_1(A, B, \mu), \\ \partial_t B &= d_B \Delta_{\mathbf{r}} B + f_2(A, B, \mu) \end{aligned}\right\} \tag{5}$$

on $\overline{\Omega} \times \mathbb{R}_0^+$ where $\Omega \subset \mathbb{R}^2$ is a bounded, connected spatial domain with piecewise smooth boundary $\partial\Omega$, $\mathbf{f} := (f_1, f_2)$ with

$$f_1(A, B, \mu) := -\alpha A + \beta A^2 B, \qquad f_2(A, B, \mu) := \gamma - \delta B - \beta A^2 B \tag{6}$$

belongs to $\mathfrak{C}^1$, where $\mu \in \{\alpha, \beta, \gamma, \delta\}$, $d_A > 0$, $d_B > 0$ represent the diffusion coefficients, $A(\mathbf{r}, t)$ and $B(\mathbf{r}, t)$ are the concentrations of the species at time $t \in [0, +\infty)$ and place $\mathbf{r} \in \overline{\Omega}$.

We show now that the interior of the first quadrant of the phase space of is an invariant region.

Lemma 1.1 *All solutions* $\mathbf{\Phi} = (\Phi_1, \Phi_2) : \overline{\Omega} \times \mathbb{R}_0^+ \to \mathbb{R}^2$ *of* (5) *with positive initial values* $\Phi_1(0) > 0$, $\Phi_2(0) > 0$ *remain positive for all* $t \geq 0$ *in their domain of existence.*

Proof We have to show that the region

$$\Sigma := \left\{(A, B) \in \mathbb{R}^2 : \ A \geq 0, B \geq 0\right\}.$$

is positively invariant for (5). Let assume that $\mathbf{\Phi} = (\Phi_1, \Phi_2) : \overline{\Omega} \times \mathbb{R}_0^+ \to \mathbb{R}^2$ is a solution of (5) satisfying positive initial conditions. Clearly, $\Phi_1 \equiv \mathbf{0}$ is a solution of the first equation. Thus, by uniqueness we can argue that no solution $\Phi_1(\cdot, t)$ at any times $t \geq 0$ can become zero in finite time. It is obvious furthermore that $(0, -1)$ is a left eigenvector of the diffusion matrix

$$D := \begin{bmatrix} d_A & 0 \\ 0 & d_B \end{bmatrix}.$$

Thus, if we set

$$G(A, B) := -B \qquad ((A, B) \in \Sigma),$$

then

$$\langle \nabla G, (f_1, f_2) \rangle_{B=0} = -\gamma < 0 \qquad \text{in} \qquad \Sigma.$$

This proves that Σ is invariant for system (5). □

In what follows we shall consider system (5) restricted to $(\mathbb{R}_0^+)^2$ and show that all solutions stay bounded in $0 \leq t \in \mathbb{R}$ which implies the existence of solutions for every $t > 0$.

Lemma 1.2 *System* (5) *is dissipative.*

Proof Let $\mathbf{\Phi} = (\Phi_1, \Phi_2) : \overline{\Omega} \times \mathbb{R}_0^+ \to \mathbb{R}^2$ be a solution of (5). Thus, for the second component of $\mathbf{\Phi}$ we have

$$\dot{\Phi}_2 - d_B \Delta_{\mathbf{r}} \Phi_2 \leq \gamma - \delta \Phi_2$$

in its domain of existence and from the comparison principle (cf. [19, Thm. 10.1., p. 94]) we obtain on this domain $\Phi_2 \leq \Psi$ where Ψ is a function of time t satisfying

$$\Psi' = \gamma - \delta \Psi, \qquad \Psi(0) := \max_{\mathbf{r} \in \overline{\Omega}} \Phi_2(\mathbf{r}, 0).$$

Clearly, $\lim_{+\infty} \Psi = \gamma/\delta$ which implies that the function $\Phi_2(\mathbf{r}, \cdot)$ ($\mathbf{r} \in \overline{\Omega}$) is defined on the whole positive half line and

$$\limsup_{t \to +\infty} \max_{\mathbf{r} \in \overline{\Omega}} \Phi_2(\mathbf{r}, t) \leq \gamma/\delta.$$

The boundedness of Φ_1 follows similarly. Thus, we have proved that all solutions of (5) stay bounded in $t \in \mathbb{R}_0^+$ which implies the existence of solutions of (5) for every $t > 0$. □

Clearly, a spatially constant solution $\mathbf{\Phi}(\cdot) = (\Phi_1(\cdot), \Phi_2(\cdot))$ of system (1) satisfies boundary conditions (2) and the kinetic system

$$\dot{\mathbf{u}} = \mathbf{f} \circ (\mathbf{u}, \mu) \tag{7}$$

The equilibria $\overline{\mathbf{u}}$ of system (7) for which

$$\mathbf{f} \circ (\overline{\mathbf{u}}, \mu) = \mathbf{0} \qquad (\mu \in I) \tag{8}$$

holds are constant solutions of (1), (2) at the same time. If e.g. the equality $\beta\gamma^2 = 2\alpha\delta$ in system (5) hold then we have a unique interior equilibrium

$$\left(\overline{A}, \overline{B}\right) := \left(\frac{\gamma}{2\alpha}, \frac{\gamma}{2\delta}\right).$$

In order to investigate the local dynamical behavior of system (1) near the equilibrium $\overline{\mathbf{u}}$ of (7) we linearize (1) at these equilibria. The realisation of the linearization depends strongly on which type of solution is investigated.

The chapter is organised as follows. In the next section we show how to investigate the occurrence of rotating waves on two types of planar domains: on disk and annulus. In the section that follows we examine the possibility the occurrence of time periodic solution of (1) when the kinetic system (7) exhibits periodic solution, as well.

2 Bifurcation of Rotating Waves

In this section we are interested in the problem of finding rotating wave solution of (1)–(2). The kinetic function **f** in (1) is required to have the following properties

$$\textbf{(F1)} \quad \mathbf{f} \in \mathfrak{C}^2(\mathbb{R}^n \times I) \qquad \text{and} \qquad \textbf{(F2)} \quad \mathbf{f}(0, \mu) = 0 \ \ (\mu \in I).$$

Assumption **(F1)** implies that the kinetic term in (1) depends only on the parameter μ and the variables $u_1, \ldots, u_n$, furthermore its second order derivative of its components are continuous. Assumption **(F2)** requires that $\mathbf{\Phi}(\mathbf{r}) \equiv \mathbf{0}$ is a solution of (1)–(2) for all $\mu \in I$.

Rotating waves are nonuniform periodic solutions to partial differential equations which rotate with a nonzero angular velocity. Thus, rotating waves can exist mathematically only in problems that have at least SO(2) symmetry (cf. [11]), i.e. there is a function $R_\vartheta \in Lin(\mathbb{R}^2)$ with

$$\left[R_\vartheta\right] = \begin{bmatrix} \cos(\vartheta) & \sin(\vartheta) \\ -\sin(\vartheta) & \cos(\vartheta) \end{bmatrix} \quad \text{and} \quad R_\vartheta\left(\overline{\Omega}\right) = \overline{\Omega} \quad (\vartheta \in [0, 2\pi)).$$

The domains disk, resp. annulus

$$\Omega_d := \left\{(x, y) \in \mathbb{R}^2 \middle| x^2 + y^2 < 1\right\},$$

resp.

$$\Omega_a := \left\{(r, \vartheta) \in \mathbb{R}^2 \middle| 1 < r < \sigma := R_2/R_1, 0 \le \vartheta < 2\pi\right\} \quad (0 < R_1 < R_2).$$

have this property (cf. Fig. 1).

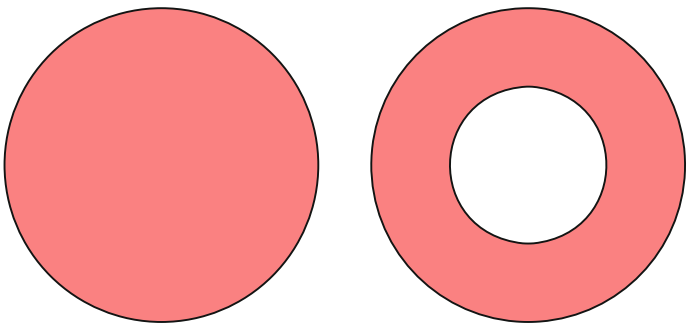


Fig. 1 $\Omega = \Omega_d$ and $\Omega = \Omega_a$

Definition 2.1 Let Ω be one of the radial symmetric domains Ω_d, Ω_a. A nontrivial non-negative solution $\mathbf{\Phi} : \overline{\Omega} \times \mathbb{R}_0^+ \to \mathbb{R}^n$ of (1) is called rotating wave if there is a function $\mathbf{T} : \overline{\Omega} \to \mathbb{R}^n$ and a number $0 \neq c \in \mathbb{R}$ (wave speed) such that

$$\mathbf{\Phi}(r, \vartheta; t) = \mathbf{T}(r, \vartheta - ct) \qquad ((r, \vartheta; t) \in \Omega \times (0, +\infty))$$

and

$$\mathbf{T}(r, \xi) = \mathbf{T}(r, \xi + 2\pi) \qquad (r \in (0, 1) \cup (1, \sigma),\ \xi \in [0, 2\pi))$$

hold.

Because we are looking for solutions $\mathbf{\Phi}$ of (1) for which

$$(\mathbf{n} \cdot \nabla_{\mathbf{r}})\mathbf{\Phi} = \mathbf{0} \qquad \text{on} \qquad \partial\Omega \times \mathbb{R}_0^+,$$

resp.

$$\mathbf{\Phi}(\cdot, 0) = \mathbf{\Phi}_0(\cdot) \geq \mathbf{0} \qquad \text{on} \qquad \overline{\Omega} \times \{0\}$$

hold, therefore using polar coordinates (r, ϑ) on Ω and denoting $\xi := \vartheta - ct$ one can easily see that chain rule implies

$$\partial_t \mathbf{\Phi} = -c\partial_\xi \mathbf{T}, \qquad (\mathbf{n} \cdot \nabla_{\mathbf{r}})\mathbf{\Phi} = \partial_r \mathbf{T} \qquad \text{and} \qquad \Delta_{\mathbf{r}}\mathbf{\Phi} = \Delta\mathbf{T},$$

where the Laplacian Δ is given by

$$\Delta := \frac{1}{r}\partial_r(r\partial_r) + \frac{1}{r^2}\partial_{\xi\xi}.$$

This means that $\mathbf{T}$ is a periodic function of period 2π in the second variable for which

$$D\Delta\mathbf{T} + c\partial_\xi \mathbf{T} + \mathbf{f}(\mathbf{T}, \mu) = \mathbf{0} \quad ((r, \xi) \in \Omega), \tag{9}$$

$$\partial_r \mathbf{T} = \mathbf{0} \quad ((r, \xi) \in \partial\Omega) \tag{10}$$

hold. Thus, we are interesting to seek those non-zero real numbers c for which system (9) and (10) has a non-trivial solution.

2.1 The Linearized Problem

Let $\overline{\mathbf{u}}$ denote one of the interior equilibria of the kinetic system (7). Moving the origin into $\overline{\mathbf{u}}$ by the coordinate transformation

$$z_1 := u_1 - \overline{u}_1, \qquad z_2 := u_2 - \overline{u}_2$$

and linearizing system (9) and (10) we get the linear boundary value problem

$$D\Delta\mathbf{z} + c\partial_\xi\mathbf{z} + Q(\mu)\mathbf{z} = \mathbf{0} \quad \text{in} \quad \Omega, \tag{11}$$

$$\partial_r\mathbf{z} = \mathbf{0} \quad \text{on} \quad \partial\Omega \tag{12}$$

where $Q(\mu) := \partial_1\mathbf{f}(\overline{\mathbf{u}}, \mu)$. The Eq. (12) has the form in case of the disc $\Omega = \Omega_d$:

$$\partial_r\mathbf{z}(1, \xi) = \mathbf{0} \qquad (\xi \in [0, 2\pi)),$$

and in case of the annulus $\Omega = \Omega_a$:

$$\partial_r\mathbf{z}(1, \xi) = \mathbf{0} = \partial_r\mathbf{z}(\sigma, \xi) \qquad (\xi \in [0, 2\pi)).$$

It is well know (cf. [4, 6, 9]) that if J_m, resp. Y_m denotes the Bessel function, resp. the Bessel function of second kind (c.f. Fig. 2) both of order m ($\in \mathbb{N}$) and

$$0 < \nu^d_{m,1} < \nu^d_{m,2} < \cdots < \nu^d_{m,n} < \ldots \qquad (n \in \mathbb{N}),$$

resp.

$$0 < \nu^a_{m,1} < \nu^a_{m,2} < \cdots < \nu^a_{m,n} < \ldots \qquad (n \in \mathbb{N})$$

are the roots of

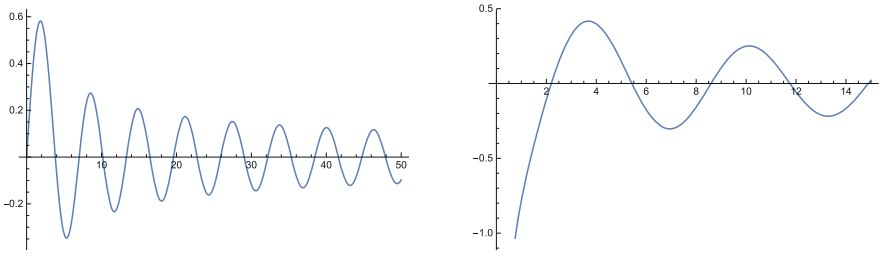


Fig. 2 The graphs of J_1 and of Y_1

$$J'_m(\cdot), \qquad \text{resp.} \qquad J'_m(\cdot\sigma)Y'_m(\cdot) - J'_m(\cdot)Y_m(\cdot\sigma)$$

then the eigenfunctions of the minus Laplacian on Ω_d, resp. Ω_a with homogeneous Neumann boundary conditions corresponding to the eigenvalues

$$\epsilon^k_{m,n} := (\nu^k_{m,n})^2 \qquad (k \in \{d, a\})$$

are the functions

$$\Omega \ni (r, \xi) \mapsto A(r)\exp(\imath m\xi)$$

where in case of the disc

$$A(r) :\equiv J_m\left(\nu^d_{m,n} r\right),$$

resp. in case of the annulus

$$A(r) :\equiv J_m(\nu^a_{m,n} r)Y'_m(\nu^a_{m,n}) - J'_m(\nu^a_{m,n})Y_m(\nu^a_{m,n} r).$$

Then the non-trivial solution of the (11) and (12) linear boundary value problem has the form (cf. [3])

$$\mathbf{T}(r, \xi) = A(r)\exp(\imath m\xi)\mathbf{e} \qquad ((r, \xi) \in \Omega) \tag{13}$$

where $\mathbf{e}$ is the eigenvector of the matrix

$$Q_{m,n}(\mu) := Q(\mu) - \epsilon^k_{m,n} D.$$

corresponding to the eigenvalue $\imath mc$. From symmetry considerations rotating wave solutions of (1) may rotate either clockwise or anticlockwise around the domain $\overline{\Omega}$ (cf. [1]). Given a solution with $c > 0$, there is another solution in the opposite direction with $c < 0$ so we will restrict our attention to the case c positive (or anticklockwise waves).

Thus, the linear boundary value problem (11)–(12) has non-trivial solution if and only if the matrix $Q_{m,n}(\mu)$ has purely imaginary eigenvalues. The eigenvalues z of $Q_{m,n}(\mu)$ are roots of the polynomial

$$z^2 - Tr\left(Q_{m,n}(\mu)\right) z + \det\left(Q_{m,n}(\mu)\right) \qquad (z \in \mathbb{C})$$

where

$$Tr\left(Q_{m,n}(\mu)\right) = Tr(Q(\mu)) - \epsilon^k_{m,n} Tr(D) \tag{14}$$

and

$$\det\left(Q_{m,n}(\mu)\right) = \det(D)\cdot\left(\epsilon_{m,n}^{k}\right)^2 - \Pi\cdot\epsilon_{m,n}^{k} + \det(Q(\mu)) \tag{15}$$

with

$$\Pi := -Tr\left(D\,(PQP)^T\right), \qquad \text{resp.} \qquad P := \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

In [2] and in [3] it was shown that for a parameter value $\mu_0 \in I$ the non-linear (9) and (10) has rotating wave solution only if the linear system (11) and (12) has non-trivial solution.

In case of system (5) the matrix $Q(\mu)$ has for $\mu = \alpha$ and $\overline{\mathbf{u}} = (\overline{A}, \overline{B})$ the form

$$Q(\alpha) = \begin{bmatrix} \alpha & \delta \\ -2\alpha & -2\delta \end{bmatrix},$$

provided $\beta\gamma^2 = 4\alpha^2\delta$ holds. Therefore we can prove the following

Theorem 2.1 *If the boundary value problem* (9) *and* (10) *with kinetic term* $\boldsymbol{f}$ *defined in* (5) *has a nontrivial solution, then*

$$d_A > d_B \tag{16}$$

must hold.

Proof The matrix $Q_{m,n}(\alpha)$ has purely imaginary eigenvalues when

$$\mathrm{Tr}\left(Q_{m,n}(\alpha)\right) = 0 \qquad \text{and} \qquad \det\left(Q_{m,n}(\alpha)\right) > 0. \tag{17}$$

The first condition in (17) holds if and only if

$$\alpha = \alpha_0 := \alpha_{m,n}^{k} = 2\delta + \epsilon_{m,n}^{k}(d_A + d_B). \tag{18}$$

When $\alpha = \alpha_{m,n}^{k}$, then

$$\det\left(Q_{m,n}(\alpha_{m,n}^{k})\right) = d_A d_B(\epsilon_{m,n}^{k})^2 + (2\delta d_A - \alpha_{m,n}^{k} d_B)\epsilon_{m,n}^{k}.$$

An easy calculation shows that in this case the polynomial

$$p(z) \equiv -d_B z^2 + 2\delta(d_A - d_B)z, \tag{19}$$

must have a positive root, which is valid if (16) holds. □

There are only finite number of eigenvalues $\epsilon_{m,n}^{k}$ of the minus Laplacian on Ω_k $(k \in \{d, a\})$ for which $\det\left(Q_{m,n}(\alpha_{m,n}^{k})\right) > 0$ holds. Because condition (16)

implies that there is a unique positive root of the polynomial p defined in (19), say $\widehat{\varepsilon}$, therefore rotating wave can bifurcate for system (5) with no-flux boundary conditions on Ω_k only from the eigenvalue $\epsilon^k_{m,n}$ for which $0 < \epsilon^k_{m,n} < \widehat{\varepsilon}$ holds.

2.2 The Nonlinear Problem

Note that the theorem in the last subsection gives necessary but not sufficient condition for bifurcation of rotating wave. To actually prove that there is a bifurcation at a critical value α_0 requires further analysis: certain transversality condition must be verified. In [2, 3, 13, 14] there was sketched a method, how the problem of finding rotating wave solution of (1) and (2) may be converted to one of finding non-trivial solution of operator equations in appropriate Banach spaces.

Clearly, introducing the new vector of variation $\mathbf{S} := \mathbf{T} - \overline{\mathbf{u}}$ where $\overline{\mathbf{u}}$ is the equilibrium of the kinetic system (cf. (8)), (9) and (10) assumes the form

$$D\Delta\mathbf{S} + c\partial_\xi\mathbf{S} + \mathbf{F}(\mathbf{S}, \mu) = \mathbf{0} \quad \text{in} \quad \Omega \tag{20}$$

$$\partial_r\mathbf{S} = \mathbf{0} \quad \text{on} \quad \partial\Omega \tag{21}$$

where $\mathbf{F}(\mathbf{0}, \mu) = \mathbf{0}$ $(\mu \in I)$ with $\mathbf{F} \in \mathfrak{C}^2((\mathbb{R}^+_0)^2 \times I, \mathbb{R}^2)$ holds for some open interval $I \subset \mathbb{R}$.

Using the implicit function theorem it can be shown (cf. e.g. [14] and [13]) that at the critical value $\alpha = \alpha_0$ in (18) the trivial solution $\mathbf{0}$ of the non-linear problem (20) and (21) undergoes a bifurcation causing rotating waves and (20) and (21) has the solution in case of the disc

$$\mathbf{\Xi}(r, \xi; s) = s \begin{bmatrix} \cos(n\xi) \\ -e_{m,n}\cos(n\xi + \varphi_{m,n}) \end{bmatrix} J_m(\nu^d_{m,n} r) + O(s)$$

and in the case of the annulus

$$\begin{aligned}\mathbf{\Xi}(r, \xi; s) = s &\begin{bmatrix} \cos(n\xi) \\ -e_{m,n}\cos(n\xi + \varphi_{m,n}) \end{bmatrix} \cdot \\ &\cdot \left(J_m(\nu^a_{m,n} r)Y'_m(\nu^a_{m,n}) - J'_m(\nu^a_{m,n})Y_m(\nu^a_{m,n} r)\right) + O(s),\end{aligned}$$

where

$$e_{m,n} := \frac{\sqrt{(\epsilon^k_{m,n} d_A - \alpha)^2 + \det\left(Q_{m,n}(\epsilon^k_{m,n})\right)}}{\delta}$$

and

$$\varphi_{m,n} := \tan^{-1}\left(\frac{\sqrt{\det\left(Q_{m,n}(\epsilon^k_{m,n})\right)}}{\alpha_0 - \epsilon^k_{m,n} d_A}\right) \qquad \text{with} \qquad \varphi_{m,n} \in (0, \pi/2).$$

Since s is considered to be small here, we this solution is called a *small amplitude rotating wave*.

3 Periodic Solutions of Reaction-Diffusion Systems

In this section we assume that $n = 2$, and the parameter dependence is not emphasized in the right hand side of (1), resp. (7), i.e. we deal with the kinetic system

$$\dot{\mathbf{u}} = \mathbf{f} \circ \mathbf{u} \tag{22}$$

and the parabolic system

$$\mathbf{u}_t = D\Delta_{\mathbf{r}}\mathbf{u} + \mathbf{f} \circ \mathbf{u}, \tag{23}$$

on a bounded spatial domain $\Omega \subset \mathbb{R}^2$ with piecewise smooth boundary with homogeneous Neumann boundary condition (2), resp. bounded non-negative initial condition (3), where D is a positive diagonal matrix: $D = \mathrm{diag}\{d_1, d_2\}$.

We assume that (22) has a non-constant orbitally asymptotically stable T-periodic solution

$$\mathbf{p} : [0, +\infty) \to \mathbb{R}^2, \qquad \mathbf{p}(t + T) \equiv \mathbf{p}(t),$$

and this solution is, at the beginning, a stable solution of the parabolic system (23), too. Varying one of the system parameters we consider the situation in which under certain conditions this spatially constant time periodic solution loses its stability and a spatially non-constant time periodic solution emerges.

Theorem 3.1 *[**Andronov-Witt**] Let be $\Phi : [0, +\infty) \to \mathbb{R}^2$ a fundamental matrix of the variational system*

$$\dot{\boldsymbol{v}} = (\boldsymbol{f}' \circ \boldsymbol{p})\boldsymbol{v} \tag{24}$$

with $\Phi(0) = I$ and M the monodromy matrix, i.e. $M = \Phi(T)$. The asymptotic orbital stability of p as a solution the kinetic system (22) depends on the modulus of the Floquet-multiplier of (24), i.e. on the modulus of the second eigenvalue μ_{20} of M /$\mu_{10} = 1$/. $\boldsymbol{p}$ is an orbitally asymptotically stable, resp. unstable solution of (23) if and only if $0 < \mu_{20} < 1$, resp. $\mu_{20} \geq 1$, i.e. $\delta < 0$, resp. $\delta > 0$ holds, where

$$\delta := \int_0^T \operatorname{div}(\boldsymbol{f}(\boldsymbol{p}(t))\,\mathrm{d}t.$$

Example 3.1 The system corresponding to the **Van der Pol's differential equation**

$$\ddot{u} + m(u^2 - 1)\dot{u} + u = 0 \tag{25}$$

has the form

$$\dot{u}_1 = u_2, \qquad \dot{u}_2 = m(1 - u_1^2)u_2 - u_1. \tag{26}$$

If $m > 0$ then system (26) has a non-constant periodic solution u_m with period T_m, but not in the strip $\|u\| < 1$. The variational system of (26) is

$$\dot{v}_1 = v_2, \quad \dot{v}_2 = -(1 + 2mu_m\dot{u}_m)v_1 + m(1 - u_m^2)v_2. \tag{27}$$

Thus, if

$$\delta = m\int_0^{T_m}(1 - u_m^2(t))\,\mathrm{d}t = mT_m - m\int_0^{T_m} u_m^2(t)\,\mathrm{d}t < 0$$

holds, the periodic solution u_m is orbitally asymptotically stable.

Example 3.2 If $\lambda, \omega > 0$, then

$$\mathbf{p}(t) := (\sqrt{\lambda}\cos(\omega t), \sqrt{\lambda}\sin(\omega t)) \quad (t \in [0, +\infty))$$

is a non-constant T-periodic solution of the autonomous system

$$\left.\begin{aligned} \dot{u}_1 &= \lambda u_1 - \omega u_2 - u_1(u_1^2 + u_2^2), \\ \dot{u}_2 &= \omega u_1 + \lambda u_2 - u_2(u_1^2 + u_2^2) \end{aligned}\right\} \tag{28}$$

where $T := 2\pi/\omega$. The variational system is

$$\dot{\mathbf{v}}(t) \equiv \begin{bmatrix} -2\lambda\cos^2(\omega t) & -\omega - \lambda\sin(2\omega t) \\ \omega - \lambda\sin(2\omega t) & -2\lambda\sin^2(\omega t) \end{bmatrix}\mathbf{v}(t).$$

Because

$$\delta = \int_0^T \left\{-2\lambda\cos^2(\omega t) - 2\lambda\sin^2(\omega t)\right\}\,\mathrm{d}t = -4\lambda\pi/\omega$$

the non-constant periodic solution $\mathbf{p}$ of (28) is orbitally asymptotically stable.

Example 3.3 (cf. [7]) If

$$\boldsymbol{\varphi}(t) := ((1/2)\sin(t) - t, t) \quad (t \in [0, +\infty))$$

is a derivo-periodic solution (cf. [8]) of the kinetic system (22) and the variational system (24) has the form

$$\dot{\mathbf{v}} = A\mathbf{v} \tag{29}$$

with

$$A(t) :\equiv \begin{bmatrix} \sin(t)/(2 - \cos(t)) - 2 & \cos(t) - 2 \\ 2/(2 - \cos(t)) & 1 \end{bmatrix},$$

then

$$\mathbf{p}(t) :\equiv \dot{\boldsymbol{\varphi}}(t) \equiv ((1/2)\cos(t) - 1, 1)$$

is a 2π-periodic solution of (29). It follows that

$$\int_0^{2\pi} \left\{ \frac{\sin(t)}{2 - \cos(t)} - 1 \right\} \mathrm{d}t = -2\pi < 0,$$

thus $\mathbf{p}$ is orbitally asymptotically stable.

Example 3.4 (Biochemical Oscillator) If $\nu, \mu, \eta > 0$ and the function g belongs to $\mathfrak{C}^1(\mathbb{R}^2, \mathbb{R})$ then certain biochemical systems can be modelled by

$$\left.\begin{aligned} \dot{u}_1 &= \nu - g(u_1, u_2) =: f_1(u_1, u_2), \\ \dot{u}_2 &= \eta\nu - \mu u_2 + g(u_1, u_2) =: f_2(u_1, u_2) \end{aligned}\right\} \tag{30}$$

where

$$g(0, u_2) \leq 0, \qquad g(u_1, 0) \geq 0 \qquad (u_1, u_2 \geq 0)$$

and

$$\partial_1 g(u_1, u_2) > 0, \qquad \partial_2 g(u_1, u_2) > 0 \qquad (u_1, u_2 > 0)$$

holds. If for all $u_2 > 0$

$$\lim_{u_1 \to +\infty} g(u_1, u_2) > u_2$$

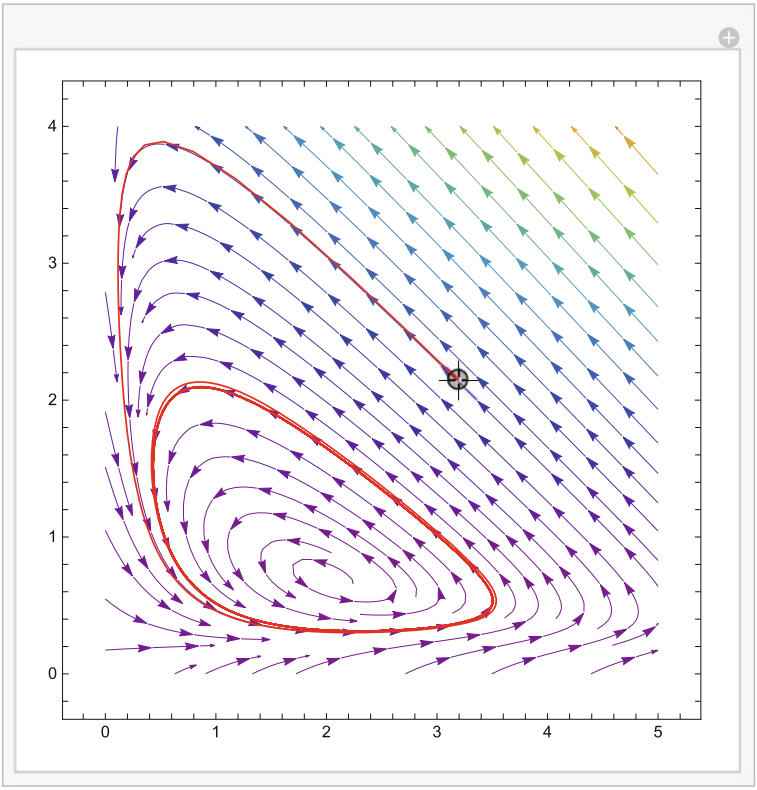


Fig. 3 Phase portrait of the system (30) in case $g(u_1, u_2) := u_1 u_2^2$

then (30) has a unique equilibrium (a, b) with $b = (1 + \eta)\nu/\mu$ in the positive quadrant of the phase space. If

$$\gamma := \partial_2 g(a, b) - \partial_1 g(a, b) - \mu > 0$$

then (a, b) is unstable and (30) has a T-periodic solution $\mathbf{p}$ which is orbitally asymptotically stable (Fig. 3).

Theorem 3.2 (cf. [12, 16]) *If*

- $\delta < 0$ *and* $d_1 = d_2$ *or the difference* $|d_1 - d_2|$ *is sufficiently small then* $\boldsymbol{p}$ *is also an orbitally asymptotically stable periodic solution of* (23)–(2).
- $\delta < 0$,

$$\int_0^T \partial_2 f_2(\boldsymbol{p}(t))\,\mathrm{d}t > 0,$$

for small $\epsilon > 0$ $d_2 = \epsilon$ and $d_1 = \epsilon^{-1}$, then $\mathbf{p}$ is an orbitally asymptotically stable solution of (22) *but unstable solution of* (23)–(2).

Clearly, the periodic solution in Example 3.2 remains orbitally asymptotically stable:

$$\int_0^{2\pi} -2\lambda \sin^2(\omega t)\,\mathrm{d}t = -2\lambda\pi < 0,$$

while the solution in Example 29 becomes unstable:

$$\int_0^{2\pi} \mathrm{d}t = 2\pi > 0.$$

The condition for change of stability in case of Example 3.4 is

$$\int_0^T \partial_2 g(\mathbf{p}(t))\,\mathrm{d}t > \mu T.$$

3.1 Bifurcation of Time-Periodic Patterns

The linearized system of (23) at $\mathbf{p}$ is

$$\mathbf{v}_t = D\Delta_{\mathbf{r}}\mathbf{v} + (\mathbf{f}' \circ \mathbf{p})\mathbf{v} \tag{31}$$

with boundary conditions

$$(\mathbf{n} \cdot \nabla_{\mathbf{r}})\mathbf{v} = 0 \qquad \text{on} \qquad \partial\Omega \times \mathbb{R}_0^+ \tag{32}$$

and smooth initial conditions

$$\mathbf{v}(\mathbf{r}, 0) = \mathbf{v}_0(\mathbf{r}) \geq 0 \qquad \text{on} \qquad \overline{\Omega} \times \{0\}, \tag{33}$$

Using the method of Fourier we obtain a sequence of solutions of (31) and (32):

$$\mathbf{\Lambda}_{kn}(\mathbf{r}, t) = \psi_n(\mathbf{r}) \cdot \boldsymbol{\varphi}_{nk}(t) \qquad \left((\mathbf{r}, t) \in \overline{\Omega} \times \mathbb{R}_0^+\right)$$

$$(n \in \mathbb{N}_0,\ k \in \{1, 2\}),$$

where ψ_n is the (eigenfunction-)solution of the problem

$$\Delta_{\mathbf{r}}\psi = -\lambda_n\psi, \qquad \partial_{\mathbf{n}}\psi\big|_{\partial\Omega} = 0$$

and

$$\boldsymbol{\varphi}_{nk} : [0, +\infty) \to \mathbb{R}^2 \qquad (k \in \{1, 2\})$$

are two linearly independent solutions satisfying

$$\dot{\boldsymbol{\varphi}} = \left(\mathbf{f}' \circ \mathbf{p} - \lambda_n D\right) \boldsymbol{\varphi} \tag{34}$$

for fixed n. In order to consider the initial condition (33) let us introduce the notation

$$\boldsymbol{\Lambda}_n := \int_\Omega \mathbf{v}_0 (\mathbf{r})\, \psi_n (\mathbf{r})\, \mathrm{d}\mathbf{r}.$$

Thus the solution of (31) and (32) has the form

$$\boldsymbol{\Lambda} (\mathbf{r}, t) = \sum_{n=0}^{\infty} \psi_n (\mathbf{r}) \exp (A_n t)\, \boldsymbol{\Lambda}_n \qquad \left((\mathbf{r}, t) \in \overline{\Omega} \times \mathbb{R}_0^+\right)$$

where

$$A_n := \mathbf{f}' \circ \mathbf{p} - \lambda_n D \qquad \text{and} \qquad \exp (A_n 0) = I.$$

Introducing the notation

$$\exp (A_n t)\, \boldsymbol{\Lambda}_n \equiv: \alpha_{n1} \boldsymbol{\omega}_{n1}(t) + \alpha_{n2} \boldsymbol{\omega}_{n2}(t)$$

and denoting the Floquet-multipliers of (34) by μ_{nk} $(n \in \mathbb{N}_0, k \in \{1, 2\})$ one can assume that in the stable case $\mu_{10} = 1$ holds and all other multipliers are in modulus less than one. If d_2 increases then at a certain critical value d_* the multiplier $\mu_{11} = 1$ while the rest of the multipliers stay in modulus less than 1. In this situation system (34) has one periodic solution $\boldsymbol{\omega}_{11}$, while another (linearly independent) solution tends exponentially to zero. In this case

$$\boldsymbol{\Lambda} (\mathbf{r}, t) - (\alpha_{n0} p(t) + \alpha_{11} \boldsymbol{\omega}_{11}(t) \psi_1 (\mathbf{r})) \longrightarrow \mathbf{0} \qquad (t \to \infty)$$

where

$$[0, +\infty) \ni t \mapsto \mathbf{P}(t) := \alpha_{n0} \mathbf{p}(t) + \alpha_{11} \psi_1 (\mathbf{r})\, \boldsymbol{\omega}_{11}(t)$$

is the time periodic spatially non-constant solution of (31) and (32), which is called **time-periodic pattern** (Fig. 4).

Finally, we note that this pattern $\mathbf{P}$ is only a solution of the linearized system (31) and (32). About the extension of this result to the nonlinear system (23)–(2) we refer the reader to [15].

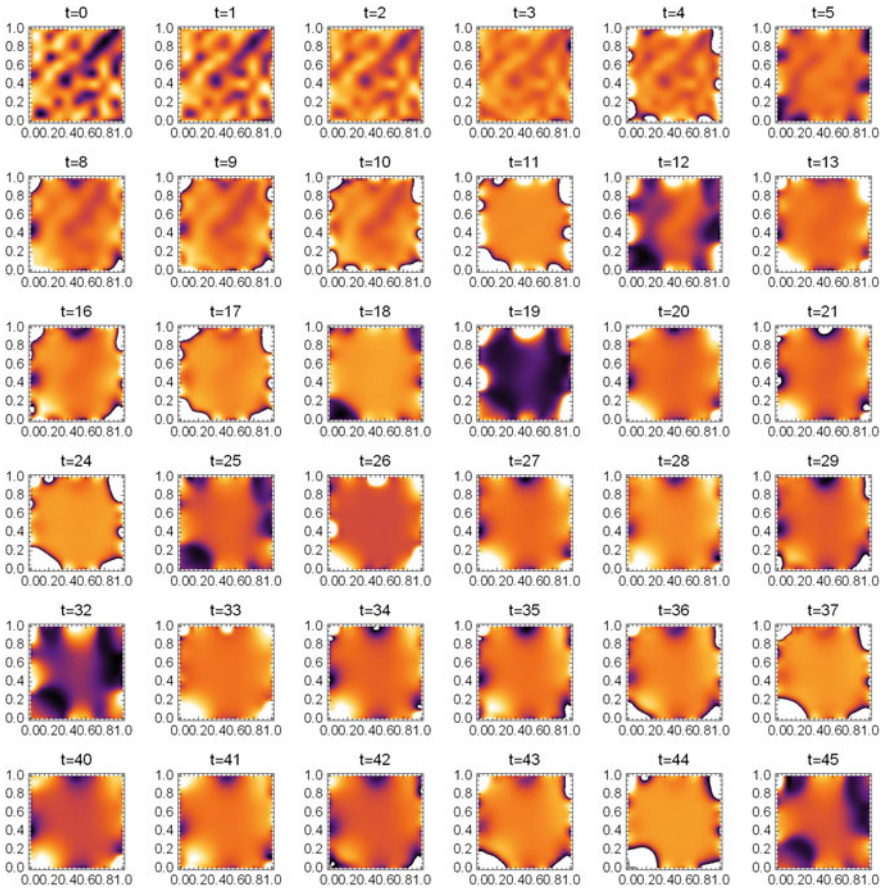


Fig. 4 Periodic pattern for system (30) in case $g(u_1, u_2) := u_1 u_2^2$

Acknowledgments The authors were supported in part by the European Union, co-financed by the European Social Fund (EFOP-3.6.3-VEKOP-16-2017-00001).

References

1. ALFORD, J. G.; AUCHMUTY, J. F. G.: *Rotating wave solutions of the FitzHugh-Nagumo equations*, J. Math. Biol. **53**(5) (2006), 797–819.
2. AUCHMUTY, J. F. G.: *Bifurcating waves*, Ann. New York Acad. Sci. **316** (1979), 263–278.
3. AUCHMUTY, J. F. G.: *Bifurcation analysis of reaction-diffusion equations v. rotating waves on a disc*, In Fitzgibbon (ed), Partial Differential Equations and Dynamical Systems III (W. E. Pitman), Boston, 1984, pp. 36–63.
4. BRITTON, N. F.: *Reaction-diffusion equations and their applications to biology*, Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], London, 1986.

5. CHUEH, K. N.; CONLEY, C. C.; SMOLLER, J. A.: *Positively invariant regions for systems of nonlinear diffusion equations*, Indiana Univ. Math. J. **26**(2) (1977), 373–392.
6. EVANS, L. C.: *Partial differential equations*, American Mathematical Society, Providence, RI, 1998.
7. FARKAS, M.: *On time-periodic patterns*, Nonlinear Analysis: Theory, Methods & Applications, **44**(5) (2001), 669–678.
8. FARKAS, M.: *Periodic motions*, Springer, New York, 1994.
9. FARKAS, M.: *Special functions* [in Hungarian], Műszaki Könyvkiadó, Budapest, 1964.
10. GAVALAS, G. R.: *Nonlinear diffusion equations of chemically reacting systems*, Springer, New York, 1968.
11. GOLUBITSKY, M.; LEBLANC, V. G.; MELBOURNE, I.: *Hopf Bifurcation from Rotating Waves and Patterns in Physical Space*, J. Nonlinear Sci. **10** (2000), 69–101.
12. HENRY, D.: *Geometric Theory of Semilinear Parabolic Equations*, Springer, New York, 1981.
13. KOVÁCS, S.: *Bifurcations in a human migration model of Scheurle-Seydel type. I: Turing bifurcation*, Internat J. Bifur. Chaos Appl. Sci. Engrg., **13**(5) (2003), 1303–1308.
14. KOVÁCS, S.: *Bifurcations in a human migration model of Scheurle-Seydel type. II: Rotating waves*, J. Appl. Math. Comput., **16**(1–2) (2004), 69–78.
15. KOVÁCS, S.: *Time-periodic patterns in reaction diffusion systems*, 2014 Workshop on Advances in Applied Nonlinear Mathematics, 18–19 September, Bristol, United Kingdom.
16. LEIVA, H.: *Stability of a periodic solution for a system of parabolic equations*, Appl. Anal., **60** (1996), 277–300.
17. LEMARCHAND, A.; BIANCA, C.: *Reaction-Diffusion Approach to Somite Formation*, IFAC-PapersOnLine, **48**(1) (2015), 346–351.
18. LEMARCHAND, A.; NOWAKOWSKI, B.: *Do the internal fluctuations blur or enhance axial segmentation?*, EPL, **94** (2011) 48004.
19. SMOLLER, J. A.: *Shock waves and reaction-diffusion equations*, Second edition. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 258. Springer-Verlag, New York, 1994.

Implementation of the Functional Response in Marine Ecosystem: A State-of-the-Art Plankton Model

Anal Chatterjee and Samares Pal

1 Introduction

The study of plankton-fish interaction is one of key area in marine ecology due to its global existence and importance. Most of plankton-fish interaction dynamics are based upon classical Lotka-Volterra mechanism, in which growth of zooplankton depend on phytoplankton and zooplankton predation rate of fish is the growth rate of fish with a conversion factor.

In complex dynamics, prey species make use of refuge to allow some degree of protection of the prey from predator. This prevents the predation of the prey population to some extent. In it's absence, prey density may increase due to a simultaneous increase of refuge that will trigger population outbreaks and forms multiple stable states. Recently, the scientists carried out many outcomes by considering refuge term in prey-predator system. The author in [1] examined the global stability of a stage-structured prey-predator model in presence of prey refuge. The researchers in [2] studied that prey refuge has a positive impact on Leslie-Gower-prey-predator model.

The phytoplankton (prey) community uses refuge to protect itself from the zooplankton (predator) community. This prevents the predation of the phytoplankton population to some extent [3–6]. Some researchers are opening new areas for study by incorporating refuge in the marine plankton system. The analysis of refuge in [7] showed that it has a strong effect in a phytoplankton-zooplankton spatiotemporal dynamics system. The authors in [8] establishes the vital role played

A. Chatterjee (✉)
Department of Mathematics, Barrackpore Rastraguru Surendranath College, Kolkata, India

S. Pal
Department of Mathematics, University of Kalyani, Kalyani, India

R. P. Mondaini (ed.), *Trends in Biomathematics: Stability and Oscillations in Environmental, Social, and Biological Models*,
https://doi.org/10.1007/978-3-031-12515-7_5

by phytoplankton refuge and toxin on the occurrence and termination of algal blooms in freshwater lake. The authors in [9] investigated the plankton dynamics in pond and the impact of zooplankton refuge on growth of tilapia (*Oreochromis niloticus*).

There are many environmental fluctuations which are not considered in deterministic models as they do not account for the variations in random parameters [10]. In such cases, a stochastic model is a better choice as they provide a realistic view of the natural system. Stochastic model of toxic producing phytoplankton-fish are studied recently in [11] for a clear understanding of the plankton dynamics. The authors in [11, 12] investigated that high and low value of environmental fluctuations may lead to the extinction of the population as well as population weakly persistent in the mean respectively. Recently, the scientists in [13] examined that white noise can affect directly the survival of plankton population in toxin producing phytoplankton-zooplankton interaction model.

Based on the above observations, we constructed a three-dimensional model focusing on the effects of phytoplankton refuge and zooplankton refuge on the marine ecosystem. Firstly, a mathematical model is designed by considering suitable functional responses. It is assumed that as per functional response type II the grazing rate of zooplankton is dependent on the phytoplankton. Zooplankton predation by fish follows a functional response type III, and the predation rate increases in a sigmoidal way with prey density. Also, many planktivorous fish feeds on chironomids, tubifex, or other bottom dwelling invertebrates. Thus, there is chance of switching to suitable prey which results in a type-III response considering there is absence of significant time-lag [14]. Further, the roles and the interaction of both plankton species and refuge phytoplankton and zooplankton are studied.

The paper is organized as follows. After the above introductory section, we move onto the model construction and the analysis of the deterministic model in Sect. 2 and Sect. 3 respectively. The stochastic counterpart is presented next, results on the numerical simulations are reported in Sect. 5 and discussed in the final section.

2 The Mathematical Model

Let P, Z and F be the concentration of the phytoplankton, zooplankton and fish population at time t with carrying capacity K and constant intrinsic growth rate r. Here α_1 and α_2 be the maximal zooplankton ingestion rate and conversion rate for the growth of zooplankton respectively ($\alpha_2 \leq \alpha_1$). Also, let γ_1 and γ_2 be the maximal planktivorous fish ingestion rate and conversion rate due to grazing of herbivorous zooplankton ($\gamma_2 \leq \gamma_1$). Further, d_1 and d_2 represent the mortality rates of the zooplankton and planktivorous fish biomass respectively. Let h be the harvesting rate of planktivorous fish population. We consider Holling type II and Holling III functional forms to illustrate the grazing phenomena with K_1 and K_2 as half saturation constants. We consider $m \in [0, 1)$ and $n \in [0, 1)$ are constant measures of the degree or strength of phytoplankton refuge and zooplankton refuge respectively.

With these above biological assumptions our model system is [15]

$$\left.\begin{aligned}
\frac{dP}{dt} &= rP(1-\frac{P}{K}) - \frac{\alpha_1(1-m)PZ}{K_1+(1-m)P} \equiv G_1(P,Z,F)\\
\frac{dZ}{dt} &= \frac{\alpha_2(1-m)PZ}{K_1+(1-m)P} - \frac{\gamma_1(1-n)^2Z^2F}{K_2+(1-n)^2Z^2} - d_1Z \equiv G_2(P,Z,F)\\
\frac{dF}{dt} &= \frac{\gamma_2(1-n)^2Z^2F}{K_2+(1-n)^2Z^2} - (d_2+h)F \equiv G_3(P,Z,F)
\end{aligned}\right\} \quad (1)$$

Firstly, we analyzed the system (1) with the following initial conditions,

$$P(0) > 0,\ Z(0) > 0,\ F(0) > 0. \quad (2)$$

Explicitly, the jacobian matrix at $\overline{E} = (\overline{P}, \overline{Z}, \overline{F})$ can be illustrate as

$$\overline{V} = \begin{bmatrix}
r - \frac{2r\overline{P}}{K} - \frac{K_1\alpha_1(1-m)\overline{Z}}{[k_1+(1-m)\overline{P}]^2} & -\frac{\alpha_1(1-m)\overline{P}}{K_1+(1-m)\overline{P}} & 0\\
\frac{K_1\alpha_2(1-m)\overline{Z}}{[K_1+(1-m)\overline{P}]^2} & \frac{\alpha_2(1-m)\overline{P}}{K_1+(1-m)\overline{P}} - \frac{2\gamma_1K_2(1-n)^2\overline{Z}\ \overline{F}}{[K_2+(1-n)^2\overline{Z}^2]^2} - d_1 & -\frac{\gamma_1(1-n)^2\overline{Z}^2}{K_2+(1-n)^2\overline{Z}^2}\\
0 & \frac{2\gamma_2K_2(1-n)^2\overline{Z}\ \overline{F}}{[K_2+(1-n)^2\overline{Z}^2]^2} & \frac{\gamma_2(1-n)^2\overline{Z}^2}{K_2+(1-n)^2\overline{Z}^2} - (d_2+h)
\end{bmatrix}. \quad (3)$$

3 Some Preliminary Results

3.1 Positive Invariance

By setting $X = (P,\ Z,\ F)^T \in \mathbf{R}^3$ and $G(X) = [G_1(X),\ G_2(X),\ G_3(X)]^T$, with $G : {\mathbf{R}_+}^3 \to \mathbf{R}^3$ and $G \in C^\infty(\mathbf{R}^3)$, Eq. (1) becomes

$$\dot{X} = J(X), \quad (4)$$

together with $X(0) = (P(0), Z(0), F(0)) = (X_1, X_2, X_3) = X_0 \in {\mathbf{R}_+}^3$. It is easy to check that whenever $X(0) \in {\mathbf{R}_+}^3$ with $X_i = 0$, for i=1, 2, 3, then $G_i(X)\ |_{X_i=0} \geq 0$. Then any solution of equation (4) with $X_0 \in {\mathbf{R}_+}^3$, say $X(t) = X(t; X_0)$, is such that $X(t) \in {\mathbf{R}_+}^3$ for all $t > 0$.

3.2 Equilibria

The system (1) possesses the following four equilibria: plankton free equilibrium $E_0 = (0, 0, 0)$ and zooplankton free equilibrium $E_{01} = (K, 0, 0)$, fish free equilibrium $E_1(P_1, Z_1, 0)$ and coexistence equilibrium $E^* = (P^*, Z^*, F^*)$.

3.2.1 Plankton Free Equilibrium

E_0 is always feasible. The eigenvalues evaluate from (3) at E_0 are $r > 0$, $-d_1 < 0$ and $-(d_2 + h)$. Thus, it is clearly indicates that E_0 is always unstable.

3.2.2 Zooplankton Free Equilibrium

E_{01} is always feasible. The eigenvalues evaluate from (3) at E_{01} are $-r$, $-(d_2 + h)$ and $\frac{\alpha_2(1-m)K}{K_1+(1-m)K} - d_1$. Therefore, E_{01} is asymptotically stable if

$$m > \frac{K\alpha_2 - d_1(K + K_1)}{K(\alpha_2 - d_1)} \tag{5}$$

hold.

3.2.3 Fish Free Equilibrium

The population level at $E_1(P_1, Z_1, 0)$ are $P_1 = \frac{K_1 d_1}{(\alpha_2 - d_1)(1-m)}$ and $Z_1 = \frac{rK_1\alpha_2[K(\alpha_2-d_1)(1-m)-K_1d_1]}{\alpha_1 K(1-m)^2(\alpha_2-d_1)^2}$. Feasibility at E_1 exists if $\alpha_2 > d_1$ and $m < \frac{K\alpha_2-d_1(K+K_1)}{K(\alpha_2-d_1)}$. By factorizing jacobian (3) at E_1 gives one explicit eigenvalue $\frac{\gamma_2(1-n)^2Z_1^2}{K_2+(1-n)^2Z_1^2} - (d_2+h)$ and the quadratic equation $\lambda^2 + P_1\left(\frac{r}{K} - \frac{\alpha_1(1-m)Z_1}{[K_1+(1-m)P_1]^2}\right)\lambda + \frac{\alpha_1\alpha_2K_1(1-m)^2P_1Z_1}{[K_1+(1-m)P_1]^3} = 0$. Clearly, two roots are negative real parts at E_1. Therefore, stability of E_1 is ensured by

$$R_0 = \frac{\gamma_2(1-n)^2Z_1^2}{(d_2+h)[K_2+(1-n)^2Z_1^2]} < 1. \tag{6}$$

3.2.4 Coexistence Equilibrium

The coexistence equilibrium at $E^* = (P^*, Z^*, F^*)$ are $Z^* = \frac{1}{(1-n)}\sqrt{\frac{(\mu_2+h)K_2}{[\gamma_2-(\mu_2+h)]}} = q$ while F^* is ensured by solving $\frac{[K_2\gamma_2(\alpha_2-d_1)(1-m)P^*-d_1K_1]}{[K_1+(1-m)P^*]\gamma_1(1-n)q[\gamma_2-(\mu_2+h)]}$ and $P^* = \frac{B+\sqrt{B^2+4AC}}{2A}$ which is ensured from equation $AP^* - BP^* - C = 0$ where $A = rK(1-m)$, $B = [rK(1-m) - rK_1]$ and $C = [rKK_1 - K\alpha_1(1-m)q]$.

At E^*, the jacobian matrix of system (1) can be written as

$$V^* = \begin{bmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & a_{23} \\ 0 & a_{32} & 0 \end{bmatrix},$$

where $a_{11} = \frac{\alpha_1(1-m)P^*Z^*}{[K_1+(1-m)P^*]^2} - \frac{rP^*}{K} < 0$, $a_{12} = -\frac{\alpha_1(1-m)P^*}{K_1+(1-m)P^*} < 0, a_{21} = \frac{K_1\alpha_2(1-m)Z^*}{[K_1+(1-m)P^*]^2} > 0$, $a_{22} = \frac{\gamma_1(1-n)^2Z^*F^*}{K_2+(1-n)^2Z^{*2}} - \frac{2\gamma_1K_2Z^*F^*}{K_2+(1-n)^2Z^{*2}} \in \mathbf{R}$, $a_{23} = -\frac{\gamma_1(1-n)^2Z^{*2}}{K_2+(1-n)^2Z^{*2}} < 0$, $a_{32} = \frac{2\gamma_2K_2(1-n)^2Z^*F^*}{[K_2+(1-n)^2Z^{*2}]^2} > 0$.
The characteristic equation is

$$y^3 + Q_1y^2 + Q_2y + Q_3 = 0 \tag{7}$$

where $Q_1 = -(a_{11} + a_{22})$, $Q_2 = a_{11}a_{22} - a_{12}a_{21} - a_{23}a_{32}$; $Q_3 = a_{11}a_{23}a_{32}$. By the Routh-Hurwitz criteria, all roots of above equation have negative real parts if and only if $Q_i > 0$, and $Q_1Q_2 - Q_3 > 0, i = 1, 2, 3$.
Here we consider two cases depending on the sign of a_{22}.

Case 1 When $a_{22>0}$, then $Q_1 > 0$ if $\frac{rP^*}{K} - \frac{\alpha_1(1-m)P^*Z^*}{[K_1+(1-m)P^*]^2} > \frac{\gamma_1(1-n)^2Z^*F^*}{K_2+(1-n)^2Z^{*2}} - \frac{2\gamma_1K_2Z^*F^*}{K_2+(1-n)^2Z^{*2}}$. Also, $Q_2 > 0$ if $-a_{12}a_{21} > -a_{11}a_{22} + a_{23}a_{32}$ since $a_{11}a_{22} < 0$, $a_{12}a_{21} < 0$ and $a_{23}a_{32} < 0$.
Here, clearly $Q_3 = a_{32}a_{11}a_{23} > 0$.

Case 2 When $a_{22<0}$ which indicates $Q_1 > 0$ and $Q_2 > 0$ if $a_{11}a_{22} > a_{12}a_{21} + a_{23}a_{32}$ and obviously, $Q_3 > 0$.

In addition, according to Routh-Hurwitz criterion, $Q_1Q_2 - Q_3 > 0$ must be satisfied if $Q_1Q_2 > Q_3$ which indicates that the system becomes locally asymptotically stable at E^* depending upon system parameters.

Remark 1 The system produces a Hopf-bifurcation at the coexistence equilibrium if the following two conditions are satisfied,

$$Q_1(n_c)Q_2(n_c) - Q_3(n_c) = 0, \quad Q_1'(n_c)Q_2(n_c) + Q_1(n_c)Q_2'(n_c) - Q_3'(n_c) \neq 0. \tag{8}$$

3.3 Hopf Bifurcation at Coexistence

Let us consider a value $n = n_c$ such that $Q_1(n_c)Q_2(n_c) - Q_3(n_c) = 0$. Then at $n = n_c$ the characteristic equation (7) becomes $(\rho + Q_1)(\rho^2 + Q_2) = 0$. Clearly the equation has three roots which are $\pm\sqrt{Q_2}i$ and $-Q_1$ i.e. two roots are pair of purely complex roots and third roots is negative. To examine the transversality condition, let us consider any point n of ϵ-neighborhood of n_c where $\rho_{1,2} = a(n) \pm ib(n)$. Putting this in (7) and separate the real and imaginary parts we get the following results:

$$a^3 - 3ab^2 + p_1(a^2 - b^2) + p_2a + p_3 = 0 \tag{9}$$

$$(3a^2b - b^3) + 2p_1ab + p_2 = 0 \tag{10}$$

Since $b(n) \neq 0$, then from (10) we have $b^2 = 3a^2 + 2Q_1a + Q_2$.
Putting the value of b^2 in (7) we have

$$8a^3 + 8Q_1a^2 + 2a(Q_1^2 + Q_2) + Q_1Q_2 - Q_3 = 0. \tag{11}$$

Now differentiating w.r.t n at $n = n_c$ we get the following results $\left[\frac{da}{dn}\right]_{n=n_c} = -\left[\frac{1}{2(Q_1^2+Q_2)}\frac{d}{dn}(Q_1Q_2 - Q_3)\right]_{n=n_c} \neq 0$ provided $\left[\frac{d}{dn}(Q_1Q_2 - Q_3)\right]_{n=n_c} \neq 0$ i.e. second condition of (8).

4 The Stochastic Model

Here, we study our system based on the environmental parameters and their fluctuations. All the parameters are assumed to be constant with time t. The stochastic stability of the coexistence equilibrium is tested.

A deterministic system can be extended to stochastic system by two process. Firstly, by replacing one of the environmental parameters by some random parameters and secondly, without changing any particular parameter in deterministic equations and including a randomly fluctuating driving force [16].

Here we choose the second method. The Gaussian white noise type stochastic perturbations of the state variables around their steady values E^* is very effective to model rapidly fluctuating phenomena which are proportional to the distances P, Z, F of each population from their equilibrium value P^*, Z^*, F^* [17]. Based on above assumption, the deterministic system (1) can be extended to the following stochastic model

$$\begin{aligned} dP &= G_1(P, Z, F)dt + \sigma_1(P - P^*)d\xi_t^1, \\ dZ &= G_2(P, Z, F)dt + \sigma_2(Z - Z^*)d\xi_t^2, \\ dF &= G_3(P, Z, F)dt + \sigma_3(F - F^*)d\xi_t^3 \end{aligned} \tag{12}$$

in which real constant parameters σ_1, σ_2 and σ_3 are considered to be the intensities of environmental fluctuations and $\xi_t^i = \xi_i(t)$, $i = 1, 2, 3$ to be the standard Wiener processes independent of each other [18].

The stochastic system (12) can be written in compact form like an Itō stochastic differential system

$$dX_t = G(t, X_t)dt + g(t, X_t)d\xi_t, \quad X_{t0} = X_0, \tag{13}$$

where the solution of above equation $X_t = (P, Z, F)^T$, for $t > 0$ is known as Itō process. Here G is the drift coefficient or it can be written as slowly varying continuous component. The diffusion coefficient is represented by $g = \text{diag}[\sigma_1(P - P^*), \sigma_2(Z - Z^*), \sigma_3(F - F^*)]$ of the diagonal matrix which can expressed by the rapidly varying continuous random component. $\xi_t = (\xi_t^1, \xi_t^2, \xi_t^3)^T$ can be considered as a three-dimensional stochastic process having scalar Wiener process components with increments $\Delta\xi_t^j = \xi_j(t + \Delta t) - \xi_j(t)$ which are independent Gaussian random variables $\mathbf{N}(0, \Delta t)$. The system (12) is known as multiplicative noise as the diffusion matrix g depends upon the solution of X_t.

4.1 Stochastic Stability of the Coexistence Equilibrium

The stochastic differential system (12) can be centered at its coexistence equilibrium E^* by introducing the perturbation vector $U(t) = (u_1(t), u_2(t), u_3(t))^T$, with $u_1 = P - P^*$, $u_2 = Z - Z^*$, $u_3 = F - F^*$. To derive the asymptotic stability in the mean square sense by the Lyapunov functions method, working on the complete nonlinear equations (12), could be attempted, following [19]. But for simplicity we deal with the stochastic differential equations obtained by linearizing (12) about the coexistence equilibrium E^*. The linearized version of (13) around E^* is given by

$$dU(t) = F_L(U(t))dt + g(U(t))d\xi(t), \tag{14}$$

where now $g(U(t)) = \text{diag}[\sigma_1 u_1, \sigma_2 u_2, \sigma_3 u_3]$ and

$$F_L(U(t)) = \begin{bmatrix} a_{11}u_1 + a_{12}u_2 + a_{13}u_3 \\ a_{21}u_1 + a_{22}u_2 + a_{23}u_3 \\ a_{31}u_1 + a_{32}u_2 + a_{33}u_3 \end{bmatrix} = MU,$$

and the coexistence equilibrium corresponds now to the origin $(u_1, u_2, u_3) = (0, 0, 0)$. Let $\Omega = \left[(t \geq t_0) \times R^3, t_0 \in R^+\right]$ and let $\Theta(t, X) \in C^{(1,2)}(\Omega)$ be a differentiable function of time t and twice differentiable function of X. Let further

$$L_\Theta(t, u) = \frac{\partial\Theta(t, u(t))}{\partial t} + f^T(u(t))\frac{\partial\Theta(t, u)}{\partial u} + \frac{1}{2}\text{tr}\left[g^T(u(t))\frac{\partial^2\Theta(t, u)}{\partial u^2}g(u(t))\right], \tag{15}$$

where

$$\frac{\partial\Theta}{\partial u} = \left(\frac{\partial\Theta}{\partial u_1}, \frac{\partial\Theta}{\partial u_2}, \frac{\partial\Theta}{\partial u_3}\right)^T, \quad \frac{\partial^2\Theta(t, u)}{\partial u^2} = \left(\frac{\partial^2\Theta}{\partial u_j \partial u_i}\right)_{i,j=1,2,3}.$$

With these positions, we now recall the following result, [20].

Theorem 1 *Assume that the functions* $\Theta(U,t) \in C_3(\Omega)$ *and* L_Θ *satisfy the inequalities*

$$r_1|U|^\beta \le \Theta(U,t) \le r_2|U|^\beta, \tag{16}$$

$$L_\Theta(U,t) \le -r_3|U|^\beta, \quad r_i > 0, \ \ i = 1,2,3, \ \ \beta > 0. \tag{17}$$

Then the trivial solution of (14) is exponentially β*-stable for all time* $t \ge 0$.

Remark 2 For $\beta = 2$ in (16) and (17), the trivial solution of (14) is exponentially mean square stable; furthermore, the trivial solution of (14) is globally asymptotically stable in probability [20].

Theorem 2 *Assume* $a_{ij} < 0$, $i, j = 1,2,3$, *and that for some positive real values of* ω_k, $k = 1,2$, *the following inequality holds*

$$\left[2(1+\omega_2)a_{22} + 2a_{32}\omega_2 + (1+\omega_2)\sigma_2^2\right]\left[2a_{23}\omega_2 + (\omega_1+\omega_2)\sigma_3^2\right] > \left[a_{12}\omega_1 + a_{22}\omega_2 + a_{23}(1+\omega_2) + a_{32}(\omega_1+\omega_2)\right]^2. \tag{18}$$

Then if $\sigma_1^2 < -2a_{11}$, *it follows that*

$$\sigma_2^2 < -\frac{2a_{22}(1+\omega_2) + 2a_{32}\omega_2}{1+\omega_2}, \quad \sigma_3^2 < -\frac{2a_{23}\omega_2}{\omega_1+\omega_2}, \tag{19}$$

where

$$\omega_1{}^* = \frac{a_{21}}{a_{11} - a_{12} - a_{32}}, \quad \omega_2{}^* = \frac{a_{11}}{a_{12} - a_{11} + a_{32}}, \quad a_{22} < 0. \tag{20}$$

and the zero solution of system (12) is asymptotically mean square stable.

Proof We consider the Lyapunov function

$$\Theta(u(t)) = \frac{1}{2}\left[\omega_1(u_1+u_3)^2 + u_2{}^2 + \omega_2(u_2+u_3)^2\right],$$

where real positive constants ω_1 and ω_2 to be define later.

It is straightforward to verify that inequalities (16) are valid for $\beta = 2$. Moreover,

$$\begin{aligned} L_\Theta(u(t)) &= a_{11}\omega_1 u_1^2 + \left[a_{22}(1+\omega_2) + a_{32}\omega_2\right]u_2^2 + \left[a_{23}\omega_2\right]u_3^2 \\ &+ u_1u_2\left[a_{12}\omega_1 + a_{21}(1+\omega_2) + a_{32}\omega_1\right] + u_2u_3\left[a_{12}\omega_1 + a_{22}\omega_2 + a_{23}(1+\omega_2)\right. \\ &\left.+ a_{32}(\omega_1+\omega_2)\right] + u_3u_1\left[a_{11}\omega_1 + a_{21}\omega_2\right] \\ &+ \frac{1}{2}\mathrm{tr}\left[g^T(u(t))\frac{\partial^2\Theta}{\partial u^2}g(u(t))\right]. \end{aligned}$$

Now we evaluate that

$$\frac{\partial^2\Theta}{\partial u^2} = \begin{vmatrix} \omega_1 & 0 & \omega_1 \\ 0 & 1+\omega_2 & \omega_2 \\ \omega_1 & \omega_2 & \omega_1+\omega_2 \end{vmatrix},$$

so that we can estimate the trace term as

$$\mathrm{tr}\left[g^T(u(t))\frac{\partial^2\Theta}{\partial u^2}g(u(t))\right] = \omega_1\sigma_1{}^2u_1{}^2 + (1+\omega_2)\sigma_2{}^2u_2{}^2 + (\omega_1+\omega_2)\sigma_3{}^2u_3{}^2.$$

Introducing (20), the Lyapunov function turn into $L_\Theta(u(t)) = -u^T Qu$, where Q be the real symmetric matrix

$$Q = \begin{vmatrix} -a_{11}\omega_1 - \frac{1}{2}\omega_1\sigma_1^2 & 0 & 0 \\ 0 & -(1+\omega_2)a_{22} - \omega_2 a_{32} - \frac{1}{2}(1+\omega_2)\sigma_2^2 & Q_{23} \\ 0 & Q_{23} & Q_{33} \end{vmatrix}$$

where

$$Q_{23} = -\frac{a_{12}\omega_1 + a_{22}\omega_2 + a_{23}(1+\omega_2) + a_{32}(\omega_1+\omega_2)}{2}$$

and $Q_{33} = -a_{23}\omega_2 - \frac{1}{2}(\omega_1+\omega_2)\sigma_3^2$. Easily, the inequality $L_\Theta(u(t)) \leq -u^T Qu$ holds. On the other hand, (18) and (19) imply that Q is positive definite and therefore all its eigenvalues $\lambda_i(Q)$, $i = 1, 2, 3$, are positive real numbers. Let $\lambda_m = \min\{\lambda_i(Q),\ i = 1, 2, 3\} > 0$. From the previous inequality for $L_\Theta(u(t))$ we thus get

$$L_\Theta(u(t)) \leq -\lambda_m |u(t)|^2,$$

thus completing the proof.

Remark 3 Theorem 2 provides the necessary conditions for the stochastic stability of the coexistence equilibrium E^* under environmental fluctuations, [21]. Thus the internal parameters of the model together with the intensities of the environmental fluctuations help in maintaining the stability of the stochastic system.

5 Numerical Simulations

In this section, we undergo the analysis of the dynamic characteristics of plankton-fish species with the help of numerical simulations. We begin with a reference set of parametric values (cf. Table 1, [22]) in which the criterion for existence at $E^* =$

Table 1 A set of parametric values

Parameter	Definition	Default value
r	Constant intrinsic growth rate of phytoplankton	8
K	Carrying capacity of phytoplankton	1.5
α_1	Maximal zooplankton ingestion rate	2
α_2	Maximal zooplankton conversion rate	1.5
β_1	Maximal fish ingestion rate	1
β_2	Maximal fish conversion rate	0.6
d_1	Mortality rate of zooplankton	0.3
d_2	Mortality rate of fish	0.36
h	Harvesting rate of fish	0.015
K_1	Half saturation constant for phytoplankton	0.5
K_2	Half saturation constant for zooplankton	1
m	Measures of the degree or strength of phytoplankton refuge	0.4
n	Measures of the degree or strength of zooplankton refuge	0.35

(1.20, 1.97, 1.78) is satisfied and coexistence equilibrium is locally asymptotically stable (cf. Fig. 1a). Now by varying the different parametric values we study the dynamic behavior of system (1).

5.1 *Effects of n*

If the value of strength of zooplankton refuge $n = 0.3$ is increased, the system exhibits oscillations around E^*. But high value of $n = 0.8$, the system switches to oscillatory behavior around zooplankton free equilibrium E_1 (cf. Fig. 1b). Figures 2a–c depicts the different steady state behaviors of phytoplankton, zooplankton and fish population in the system (1) for the parameter n. Here, we see two Hopf bifurcation points at $n_c = 0.3211$ and 0.6442 (denoted by a red star (H)) with first Lyapunov coefficient being $-6.351217e^{-02}$ and $2.715545e^{01}$ which indicates that a stable and unstable limit cycle bifurcates from the H and loses its stability respectively. Here $n = 0.6443$ (LP) and $n = 0.6441$ (BP) denotes the limit point and branch point of the system (1) respectively where fish population goes to extinction. Further, we have plotted a family of limit cycles bifurcates from H points (cf. Fig. 2d).

5.2 *Effects of m*

Taking $m = 0.8$, the system exhibits oscillations around E^* (cf. Fig. 1c). Figures 3a–c illustrate the different steady state behaviour of each species in the system (1)

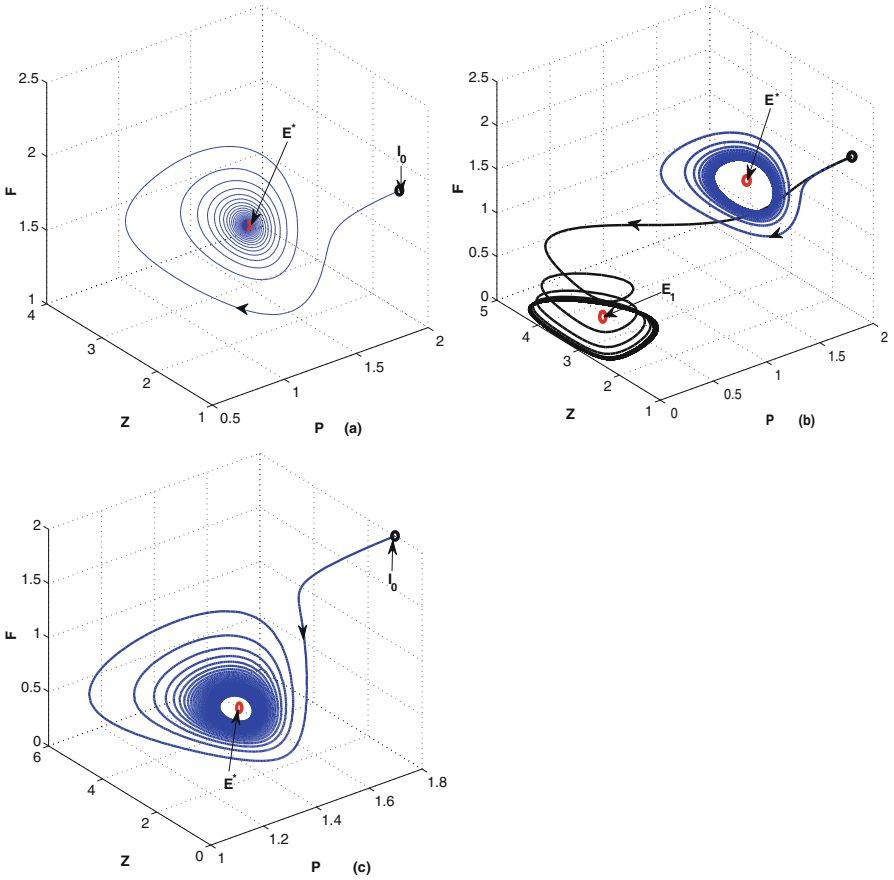


Fig. 1 **(a)** The equilibrium point E^* is stable for the parametric values as given in the Table 1. **(b)** The figure depicts oscillatory behavior around coexistence equilibrium point E^* of system (1) for $n = 0.3$ (blue line), zooplankton free equilibrium E_1 for $n = 0.8$ (black line) respectively. **(c)** The figure depicts oscillatory behavior around E^* of system (1) for $n = 0.8$ (blue line)

for the parameter m. Here, we see two Hopf bifurcation points at $m_c = 0.6069$ and 0.8021 (denoted by a red star (H)) with first Lyapunov coefficient being $-3.370275e^{-02}$ and $-1.339074e^{-02}$ which indicates that two stable limit cycle bifurcates from the H and loses its stability respectively. Here $m = 0.9102$ (BP) denotes branch point of the system (1). Further, we have displayed a family of limit cycles bifurcates from H points (cf. Fig. 3d).

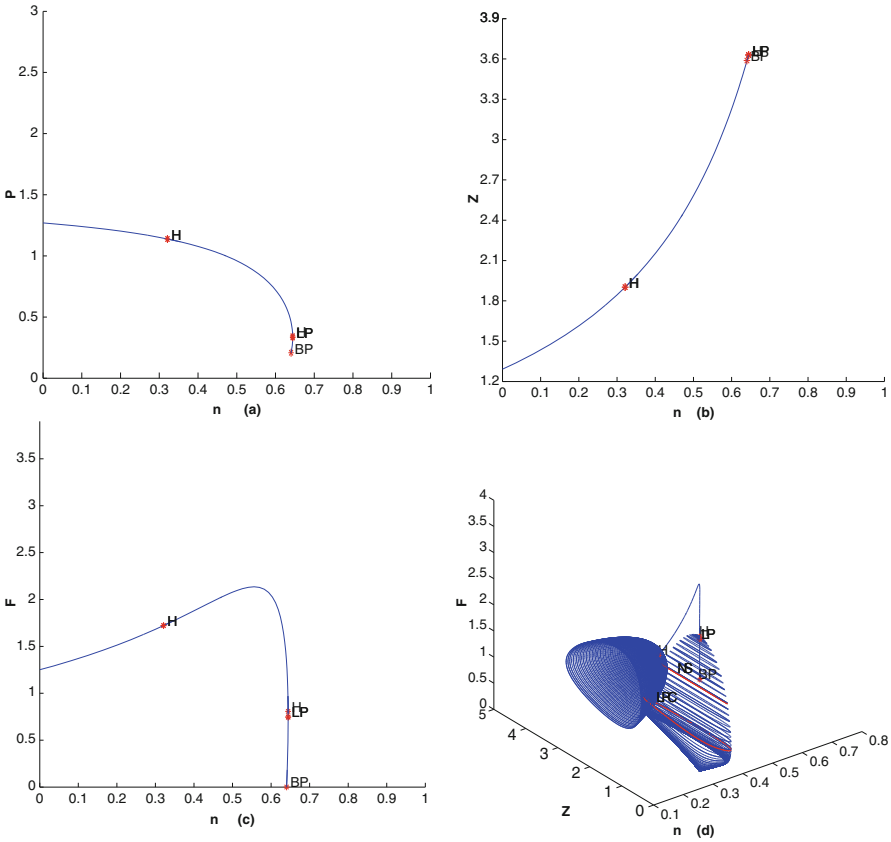


Fig. 2 **(a)** The figure depicts different steady-state behaviors of phytoplankton for the effect of n. **(b)** The figure depicts different steady-state behaviors of zooplankton for the effect of n. **(c)** The figure depicts different steady-state behaviors of fish for the effect of n. **(d)** The family of limit cycles bifurcate from the Hopf point H for n in (n,Z,F) space

5.3 *Effects of r*

From Figs. 4a–c it follows the system (1) has two Hopf bifurcation points at $r_c = 8.3570$ and 4.3791 with first Lyapunov coefficient being $-5.840029e^{-02}$ and $1.491169e^{+01}$, one limit point at 4.3776 and branch point at 4.4284 when we consider constant intrinsic growth rate of phytoplankton, i.e. r as a free parameter. To proceed further, a family of stable and unstable limit cycles bifurcating from Hopf points is plotted (cf. Fig. 4d).

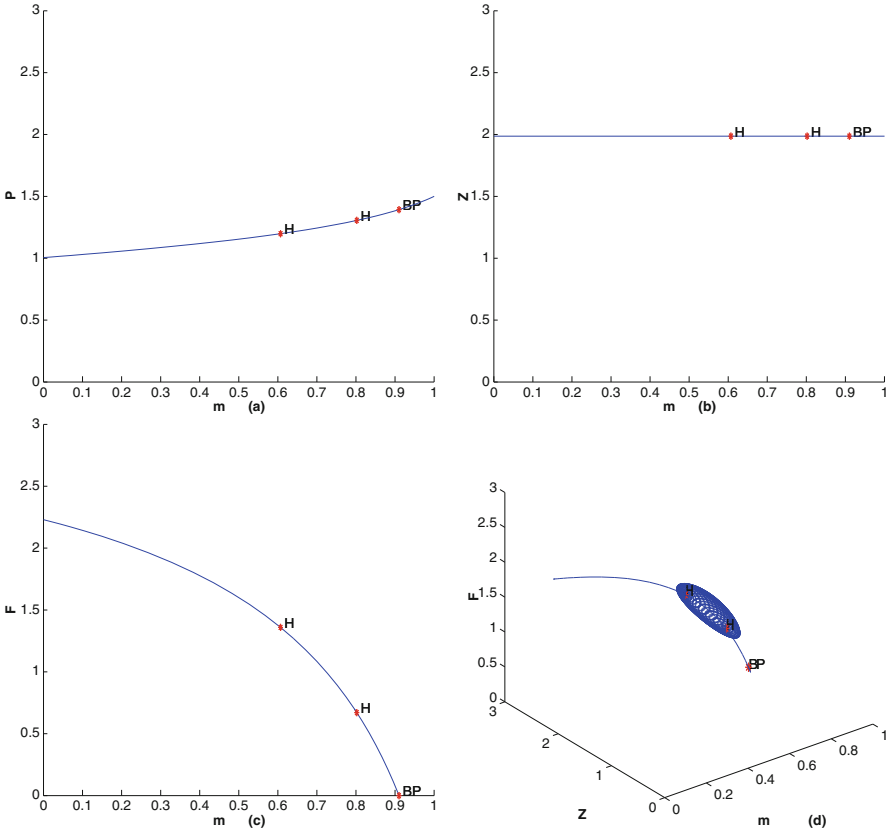


Fig. 3 **(a)** The figure depicts different steady-state behaviors of phytoplankton for the effect of m. **(b)** The figure depicts different steady-state behaviors of zooplankton for the effect of m. **(c)** The figure depicts different steady-state behaviors of fish for the effect of m. **(d)** The family of limit cycles bifurcate from the Hopf point H for m in (m,Z,F) space

5.4 Effects of h

To study the impact of harvesting on fish population we vary the parameter h. We note that the system (1) has one Hopf point at 0.1484 with first Lyapunov coefficient being $1.059854e^{+01}$ one limit point at 0.1486 and branch point at 0.1468 (cf. Fig. 5a). We have drawn a family of unstable limit cycles bifurcating from H points (cf. Fig. 5b).

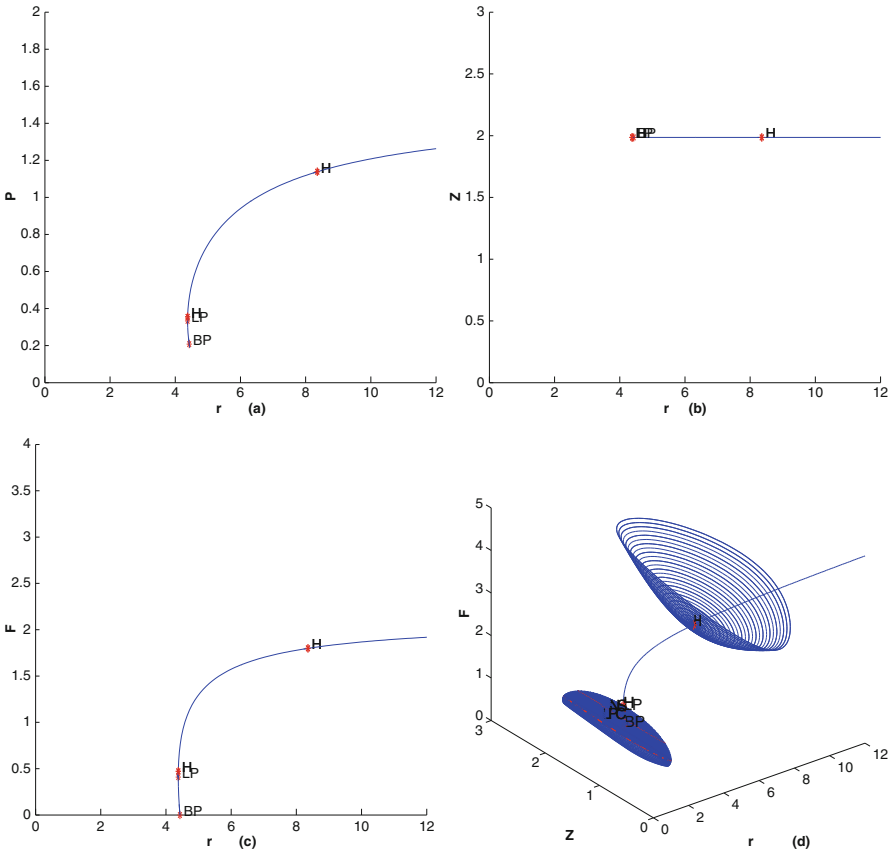


Fig. 4 **(a)** The figure depicts different steady-state behaviors of phytoplankton for the effect of r. **(b)** The figure depicts different steady-state behaviors of zooplankton for the effect of r. **(c)** The figure depicts different steady-state behaviors of fish for the effect of r. **(d)** The family of limit cycles bifurcate from the Hopf point H for r in (r,Z,F) space

5.5 *Hopf-Bifurcation*

For clear understanding of a dynamic change due to change in n, m and r, we have plotted three bifurcation diagrams separately (cf. Fig. 6a–c). Next, we have plotted two parameter bifurcation diagrams for $n-m$, $n-r$ and $m-r$ respectively (Figs. 7a–c) to show the stable zone at E^*. All the numerical results are summarized in Table 2.

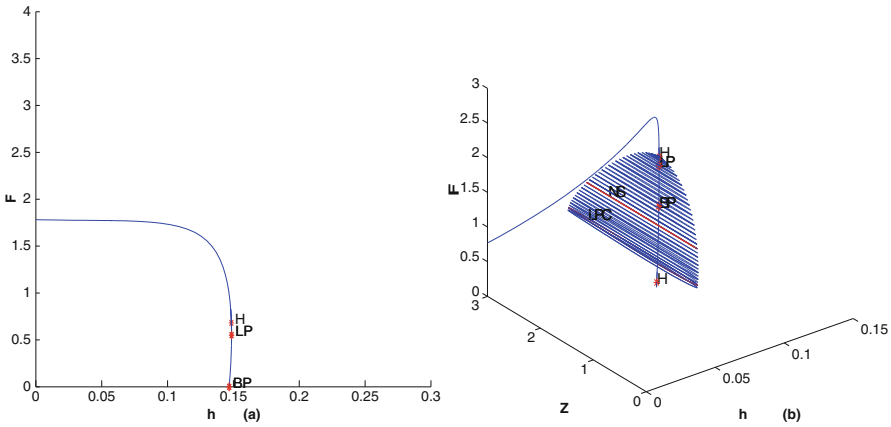


Fig. 5 **(a)** The figure depicts different steady-state behaviors of fish for the effect of h. **(b)** The family of limit cycles bifurcate from the Hopf point H for r in (r,Z,F) space

Table 2 Natures of equilibrium points

Parameters	Values	Eigenvalues	Equilibrium points
n	0.321158	$(-4.81561, \pm 0.404804i)$	Hopf (H)
	0.644284	$(-.0007, \pm 1.39674i)$	Hopf (H)
	0.644318	$((0, \pm 1.39308i)$	Limit Point (LP)
	0.640187	$(0, 0.13333 \pm 1.27889i)$	Branch Point (BP)
m	0.606983	$(5.50548, \pm 0.350157i)$	Hopf (H)
	0.802120	$(-6.55763, \pm 0.244601i)$	Hopf (H)
	0.910270	$(-7.29689, -0.0187574, 0)$	Branch Point (BP)
r	8.357091	$(-5.0378, \pm 0.404562i)$	Hopf (H)
	04.379199	$(-.00147, \pm 1.04446i)$	Branch Point (BP)
	4.377626	$(0, \pm 1.04446i)$	Limit Point (LP)
	4.428463	$(0, \pm 0.929514i)$	Branch Point (BP)
h	0.148487	$(-0.001428, \pm 1.39158i)$	Hopf (H)
	0.148622	$(0, 0.04471 \pm 1.38661i)$	Limit Point (LP)
	0.146818	$(0, 0.13453 \pm 1.26379i)$	Limit Point (LP)

5.6 *Environmental Fluctuations*

Next, we examine the dynamical behavior of the system in the presence of environmental disturbances. We apply the Euler-Maruyama method and investigate the stochastic differential equation numerically using MATLAB.

Firstly, we satisfy the conditions for asymptotic stability at E^* according to the mean square sense which depends on system parameters of (12) and σ_1, σ_2, σ_3. Taking $\sigma_1 = 0.1$, $\sigma_2 = 0.1$, $\sigma_3 = 0.1$, the values of intensities of the environmental perturbations with reference set of parametric values as in Table 1 for which all

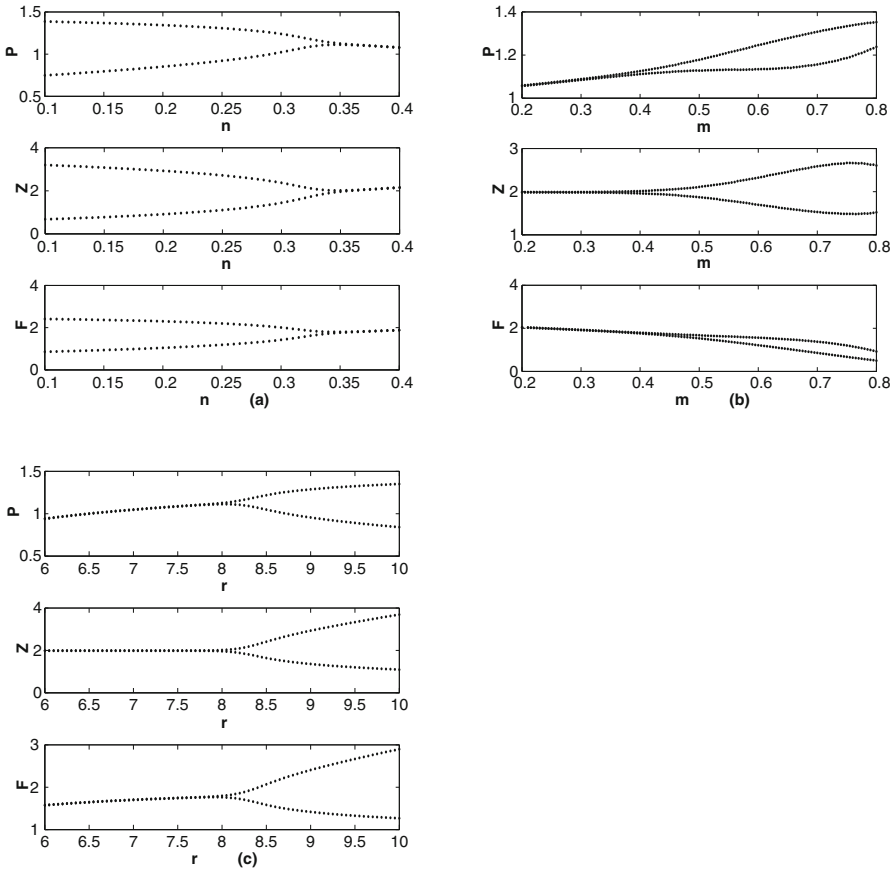


Fig. 6 **(a)** The bifurcation diagram for n. **(b)** The bifurcation diagram for m. **(c)** The bifurcation diagram for r

the three species coexist and the system is stochastically stable (cf. Fig. 8a). But it is clearly indicates that the coexistence equilibrium becomes unstable for higher values of intensities of the environmental perturbations, $\sigma_1 = 0.8$, $\sigma_2 = 0.8$, $\sigma_3 = 0.7$ (cf. Fig. 8b).

6 Discussion

We have formulated a mathematical model sketching the interaction of phytoplankton-zooplankton-fish species. The main focus is on the functional response in presence of refuge effects of phytoplankton and zooplankton on the marine ecosystem. The model parameters are also analysed by either varying one of them or combining some of them.

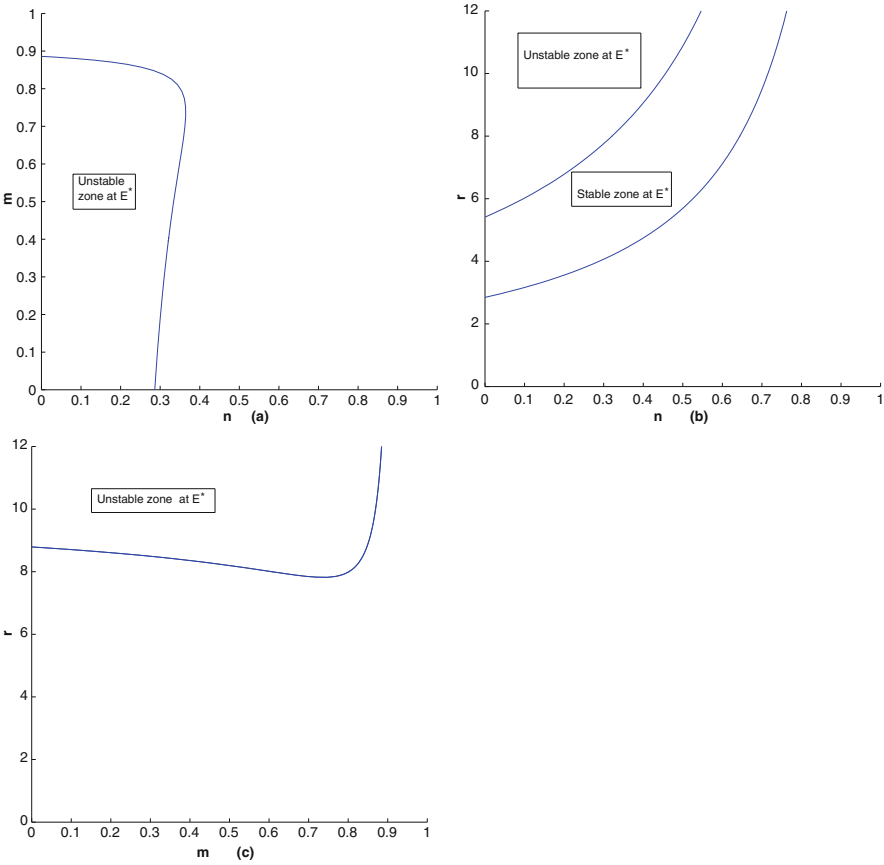


Fig. 7 **(a)** The two parameters bifurcation diagram for $n - m$. **(b)** The two parameters bifurcation diagram for $n - r$. **(c)** The two parameters bifurcation diagram for $m - r$

The stability of the three possible steady states namely plankton-free, the zooplankton-free and the coexistence equilibria are determined by studying the model analytically. The equilibrium states are observed to be related by transcritical bifurcations provided the parameter values matches suitable conditions. Hopf bifurcation at the coexistence equilibrium are obtained after analytical results and are backed by the numerical simulations. By changing the various parameters, persistent oscillations occur.

Numerically, oscillation of all population is observed when we reduce the strength of phytoplankton refuge and when the strength of zooplankton is increased. Further, same results are obtained by increasing the constant intrinsic growth rate of phytoplankton. The whole system is stabilized by harvesting rate of fish which plays a crucial role in marine ecosystem.

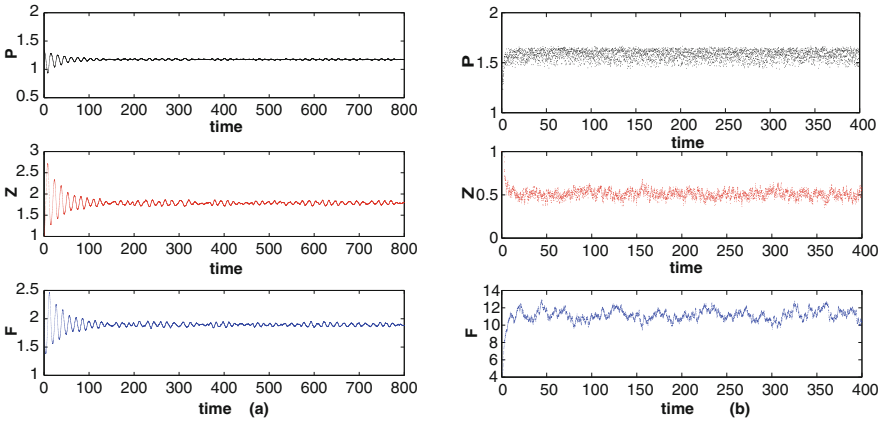


Fig. 8 **(a)** The figures depicts solution of system is stochastically stable for $\sigma_1 = 0.1, \sigma_2 = 0.1$ and $\sigma_3 = 0.1$. **(b)** The figures depicts solution of system is stochastically unstable for $\sigma_1 = 0.8, \sigma_2 = 0.8$ and $\sigma_3 = 0.7$

Based on the results,we can conclude that the strength of phytoplankton and zooplankton refuge, intrinsic growth rate of phytoplankton and harvesting rate of fish should be maintained within a range in order to avoid extinction of fish and recurrence bloom.

Environmental noise is further added to the model and it's low intensities makes the system stochastic asymptotic stable. High intensity values result in oscillations with high amplitudes. The model becomes stochastically stable if it fulfills certain conditions involving the maximum size of the environmental random fluctuations and the model parameters.

References

1. S. Khajanchi and S. Banerjee, Role of constant prey refuge on stage structure predator–prey model with ratio dependent functional response. *Applied Mathematics and Computation* **314**, 193–198 (2017).
2. X. Xie, Y. Xue, J. Chen and T. Li, Permanence and global attractivity of a nonautonomous modified Leslie-Gower predator-prey model with Holling-type II schemes and a prey refuge. *Advances in Difference Equations* **2016(1)**, 1–11 (2016).
3. S. Khajanchi and S. Banerjee, Role of constant prey refuge on stage structure predator–prey model with ratio dependent functional response. *Applied Mathematics and Computation* **314**, 193–198 (2017).
4. M. Haque, S. Rahman, E. Venturino and B. L. Li, Effect of a functional response-dependent prey refuge in a predator–prey model. *Ecological Complexity* **20**, 248–256 (2014).
5. D. Jana and S. Ray, Impact of physical and behavioral prey refuge on the stability and bifurcation of Gause type Filippov prey-predator system. *Modeling Earth Systems and Environment*, **2(1)**, 24 (2016) https://doi.org/10.1007/s40808-016-0077-y.

6. A. Das and G. P. Samanta, A prey-predator model with refuge for prey and additional food for predator in a fluctuating environment. *Physica A: Statistical Mechanics and its Applications* **538**, 122844 (2020).
7. W. Zhang and M. Zhao, Dynamical Complexity of a Spatial Phytoplankton-Zooplankton Model with an Alternative Prey and Refuge Effect. *Hindawi Publishing Corporation Journal of Applied Mathematics* Volume **2013**, Article ID 608073 (2013).
8. J. Li, Y. Song, H. Wan and H. Zhu, Dynamical analysis of a toxin-producing phytoplankton-zooplankton model with refuge. *Mathematical Biosciences and Engineering* **14(2)**, 529–557 (2017).
9. W. Sun, S. Dong, X. Zhao, Z. Jie, H. Zhang, and L. Zhang, Effects of zooplankton refuge on the growth of tilapia (Oreochromis niloticus) and plankton dynamics in pond. *Aquaculture international* **18(4)**, 647–655 (2010).
10. M. Bandyopadhyay and J. Chattopadhyay, Ratio-dependent predator-prey model: Effect of environmental fluctuation and stability. *Nonlinearity* **18** , 913–936 (2005).
11. T. Liao, C. Dai, H. Yu, Z. Ma, Q. Wang and M. Zhao, Dynamical analysis of a stochastic toxin-producing phytoplankton–fish system with harvesting. *Advances in Difference Equations* **2020(1)**, 1–22 (2020).
12. Z. Chen, S. Zhang and C. Wei, Dynamics of a stochastic phytoplankton-toxin phytoplankton–zooplankton model. *Advances in Difference Equations*, **2019(1)**, 1–18 (2019).
13. H. Liu, C. Dai, H. Yu, Q. Guo, J. Li, A. Hao, J. Kikuchi and M. Zhao, Dynamics induced by environmental stochasticity in a phytoplankton-zooplankton system with toxic phytoplankton. *Mathematical Biosciences and Engineering*, **18(4)**, 4101–4126 (2021).
14. W. W. Murdoch and J. Bence, General predators and unstable prey populations, in: Predation: direct and indirect impacts on aquatic communities. (W. C. Kerfoot and A. Sih, eds.), *University Press of New England*, Hanover, 17–30 (1987).
15. A. Chatterjee and S. Pal, Dynamical Analysis of Phytoplankton–Zooplankton Interaction Model by Using Deterministic and Stochastic Approach. In: Mondaini R.P. (eds) *Trends in Biomathematics: Chaos and Control in Epidemics, Ecosystems, and Cells. BIOMAT 2020.* Springer, Cham, 33–56 (2021).
16. P.K. Tapaswi and A. Mukhopadhyay, Effects of environmental fluctuation on plankton allelopathy, *J. Math. Biol.* 39, 39–58 (1999).
17. E. Beretta, V.B. Kolmanowskii and L. Shaikhet, Stability of epidemic model with time delays influenced by stochastic perturbations, *Math. Comp. Simul.* **45 (3-4)**, 269–277 (1998).
18. I.I. Gikhman and A.V. Skorokhod, The Theory of Stochastic Process-I, Springer, Berlin, (1979).
19. L. Shaikhet, Lyapunov Functionals and Stability of Stochastic Functional Differential Equations. Springer, Dordrecht, Heidelberg, New York, London, 2013.
20. V.N. Afanas'ev, V.B. Kolmanowskii and V.R. Nosov, Mathematical Theory of Control Systems Design, Kluwer Academic, Dordrecht, (1996).
21. M. Bandyopadhyay and J. Chattopadhyay, Ratio-dependent predator-prey model: Effect of environmental fluctuation and stability, *Nonlinearity* **18**, 913–936 (2005).
22. S. Pal and A. Chatterjee, Dynamics of the interaction of plankton and planktivorous fish with delay. *Cogent Mathematics* **2(1)**, 1074337 (2015).

Optimal Drug Therapy in a Multi-Pathways HIV-1 Infection Model with Immune Response Delay

Chittaranjan Mondal, Debadatta Adak, and Nandadulal Bairagi

1 Introduction

Human Immunodeficiency Virus (HIV) type 1 is a lentivirus that causes HIV-1 infection and is responsible for Acquired Immunodeficiency Syndrome (AIDS), a condition where a failure of the immune system causes life-threatening infections. When HIV infects the human body, it explicitly attacks $CD4^+T$ cells (helper T cells) that play an essential role in activating other players (e.g., cytotoxic T lymphocytes and B cells) of the immune system to fight against any invader. Transmission of HIV-1 among host cells occurs mainly following two mechanisms, viz., virus-to-cell and cell-to-cell [1–3]. The free plasma virus infects the healthy $CD4^+T$ cells. On the other hand, the infection spreads from one infected to another uninfected $CD4^+T$ cell in cell-to-cell transmission mode through synapse formation. Only recently, it has been acknowledged that the cell-to-cell disease transmission mode is not only a more efficient and faster mode of disease transmission compared to cell-free mode but also the predominant mode of viral spread in the in-host HIV infection [1–4].

A basic HIV in-host infection model considers the interaction between the host cells ($CD4^+T$) and virus particles without considering the role of cytotoxic T lymphocytes (CTL) or $CD8^+T$ cells, which on activated by $CD4^+T$ cells, kills the infected cells directly. Basic models also do not consider the infection spreading through cell-to-cell mode. Recently, some mathematical models have extended

C. Mondal · N. Bairagi (✉)
Centre for Mathematical Biology and Ecology, Department of Mathematics, Jadavpur University, Kolkata, India
e-mail: nbairagi.math@jadavpuruniversity.in

D. Adak
Department of Mathematics, Maharaja Bir Bikram University, Agartala, Tripura, India

R. P. Mondaini (ed.), *Trends in Biomathematics: Stability and Oscillations in Environmental, Social, and Biological Models*,
https://doi.org/10.1007/978-3-031-12515-7_6

the basic model by considering the multi-mode dissemination of disease [5–8]. Acknowledging the controlling role of CTL cells, Lai and Zou [9] considered a four-dimensional multi-pathways in-host HIV model. Since immune activation is not instantaneous but mediated by some time lag, Xu and Zhou [10] modified the model of Lai and Zou [9] taking into consideration the CTL activation delay and analyzed the model

$$\begin{aligned}
\dot{x}(t) &= s - dx(t) + rx(t)\left(1 - \frac{x(t) + \alpha y(t)}{K}\right) - \beta_1 x(t)v(t) \\
&\quad - \beta_2 x(t)y(t), \\
\dot{y}(t) &= \beta_1 x(t)v(t) + \beta_2 x(t)y(t) - \sigma y(t) - d_1 y(t)z(t), \\
\dot{v}(t) &= c\sigma y(t) - d_2 v(t), \\
\dot{z}(t) &= py(t-\tau) - d_3 z(t).
\end{aligned} \tag{1}$$

Here $x(t)$, $y(t)$, $v(t)$ and $z(t)$ represent, respectively, the concentrations of susceptible $CD4^+T$ cells (target cells), productively infected $CD4^+T$ cells, free plasma virus and CTL cells at time t. Target cells are infected by free virus particle as well as infectious $CD4^+T$ cells following mass action law with rate constants β_1 and β_2, respectively. The time required for the activation of CTL cells is represented by τ. Here s is the constant input rate, d is the death rate, r $(> d)$ is the proliferation rate, K is the maximum density of $CD4^+T$ cells, and α is a limitation coefficient of infected cells. The parameters d_1, d_2, and d_3 represent, respectively, the killing rate of infected cells by CTL, clearance rate of virus particles, and clearance rate of CTL. The virus replication factor is represented by c, and p is the production rate of CTL. Here all parameters are non-negative. For an elaborate explanation of this model, readers are referred to [10] and the references therein.

Antiretroviral therapies (ART) are medications that are used to treat HIV infection in the human body. It reduces viral load progression and increases the $CD4^+T$ cells count, thus opposing the onset of AIDS and increasing the life span of HIV-infected individuals. However, the continuous antiretroviral treatment causes toxicities, drug resistance and adherence. Therefore, antiretroviral drugs should be used optimally. Different mathematical models have been proposed and studied by the researchers to analyze the effect of ART drug therapies [12–23] on human body. These models, however, mainly consider the spread of infection through a single pathway, which is the cell-free mode. Recently, Mondal et al. [39] has analyzed a multi-pathways HIV control model in the absence of delay and CTL cells. In this work, we introduce different blockers in the delay-induced system (1) and then analyze it. It is to be mentioned that an infected cell may die naturally or through cell-lysis [29, 31]. In model (1), both types of death have been represented by a single term $\sigma y(t)$ and the production rate of new virus particles has been represented by $c\sigma y(t)$, where c is the number of new viruses produced per cell lysis. This may cause an overestimation of free virus in the blood plasma because the free virus can be created only through cell lysis, whereas no virus protein can be released

during normal cell death. Taking care of this fact, we split the total death rate (σ) of the infected cells into two parts: natural death, d, and death due to cell lysis, μ, (i.e., $\sigma = d + \mu$). We now introduce three blockers to reduce viremia. A control $u_1(t) \in [0, 1]$ is applied to reduce the transmission of infection through cell-free mode. This control is mainly reverse transcriptase inhibitor (RTI) drugs that block the synthesis of viral DNA from HIV-1 RNA, thereby reducing viral infectivity. A second control $u_2(t) \in [0, 1]$ is used to block the cellular mechanisms required for synapse formation, the primary mechanism of cell-to-cell transmission of HIV. We call it a synapse-forming inhibitor (SI). The third control $u_3(t) \in [0, 1]$ is applied to prevent HIV-1 protease from cleaving the HIV-1 polyprotein into functional units, popularly known as protease inhibitor (PI). Introducing these modifications, the model (1) reads

$$\begin{aligned}
\dot{x}(t) &= s - dx(t) + rx(t)\Big(1 - \frac{x(t) + \alpha y(t)}{K}\Big) - (1 - u_1(t))\beta_1 x(t)v(t) \\
&\quad - (1 - u_2(t))\beta_2 x(t)y(t), \\
\dot{y}(t) &= (1 - u_1(t))\beta_1 x(t)v(t) + (1 - u_2(t))\beta_2 x(t)y(t) - (d + \mu)y(t) \\
&\quad - d_1 y(t)z(t), \\
\dot{v}(t) &= (1 - u_3(t))c\mu y(t) - d_2 v(t), \\
\dot{z}(t) &= py(t - \tau) - d_3 z(t).
\end{aligned} \tag{2}$$

The initial conditions are taken as

$$\begin{aligned}
&x(\theta) = \phi_1(\theta) > 0,\ y(\theta) = \phi_2(\theta) > 0,\ v(\theta) = \phi_3(\theta) > 0, \\
&z(\theta) = \phi_4(\theta) > 0,\ \theta \in [-\tau, 0],\ \text{where} \\
&\phi = (\phi_1, \phi_2, \phi_3, \phi_4) \in C = C([-\tau, 0], \mathbb{R}^4_+) \\
&\text{with } \phi_i(\theta) \geq 0\ (\theta \in [-\tau, 0], i = 1, 2, 3, 4), \\
&(\phi_1(0), \phi_2(0), \phi_3(0), \phi_4(0)) \in C.
\end{aligned} \tag{3}$$

The objectives are to (1) observe how far these blockers can control the viremia, (2) know whether multi-drug therapy is beneficial over mono-drug therapy, considering toxicities of the antiviral drugs, (3) explore how immune response delay affect the plasma viral load in the presence and absence of the blockers and (4) have insights about the optimal dose of the blockers.

The paper is arranged in the following sequence. In Sect. 2, we present some preliminary results. Section 3 represents the system analysis and its simulations when controls are constant. Section 4 deals with the optimal management of the system when controls are time-dependent. A comparison of different control schemes is presented here. The paper ends with a discussion in Sect. 5.

2 Preliminary Results

2.1 Existence and Uniqueness of Solution

Let $N = C([-\tau, 0], \mathbb{R}^4_+)$ be the Banach space of continuous real-valued functions from $[-\tau, 0]$ *to* $\mathbb{R}^4_+$ with sup-norm $\|\phi\| = sup_{-\tau\le\theta\le 0}\{|\phi_1(\theta)|, |\phi_2(\theta)|, |\phi_3(\theta)|, |\phi_4(\theta)|\}$. Following fundamental theory of functional differential equations [24], for any $\phi \in N$ and initial conditions (3), the system (2) has a unique solution

$$\Omega(t, \phi) = (x(t, \phi), y(t, \phi), v(t, \phi), z(t, \phi)).$$

One can easily prove the following lemma from [10].

Lemma 2.1 *Let $\Omega(t, \phi) = (x(t), y(t), v(t), z(t))$ be a solution of the system (2) with initial conditions (3). Then the solution is positively invariant provided $u_1(t) = u_2(t) = 1$ does not hold simultaneously and $u_3(t) \in [0, 1)$. The solution is uniformly bounded on the region*

$$\Theta = \left\{ \Omega(t, \phi) \in \mathbb{R}^4_+ \mid 0 < x(t) \le x_0, 0 \le y(t) \le \frac{M}{\delta}, \right.$$
$$\left. 0 \le v(t) \le \frac{(1-u_3)cd_1M}{\delta d_2}, 0 \le z(t) \le \frac{pM}{\delta d_4} \right\}, \tag{4}$$

where $x_0 = \frac{K}{2r}[(r-d) + \sqrt{(r-d)^2 + \frac{4rs}{K}}]$, $\delta = min\{d, (d+\mu)\}$, $M = s + \frac{rK}{4}$. *Moreover, there exists* $\Psi_0 > 0$ *such that* $\liminf_{t\to\infty} x(t) > \Psi_0$, *i.e., $x(t)$ is uniformly bounded away from zero.*

3 Model Analysis with Fixed Controls

Control parameters $u_i, i = 1, 2, 3,$ may be either constant or time-dependent. In the immediate next section, we analyze the model treating control parameters as constant.

3.1 Basic Reproduction Number

A threshold quantity, R_0, called the basic reproduction number, is used to determine whether an infection will spread or be washed out. It is considered to be an essential threshold quantity for the elimination of infection. Typically, R_0 is defined by the expected number of secondary cases produced by an infected cell in a completely

susceptible host population [25]. Using the next-generation matrix [26], one can easily prove the following proposition.

Proposition 3.1 *If R_0 be the basic reproduction number of the system (2) and R_{01}, R_{02} are the respective basic reproduction numbers corresponding to the cell-free infection mode (i.e., $\beta_2 = 0$) and cell-to-cell infection mode (i.e., $\beta_1 = 0$) then $R_0 = R_{01} + R_{02}$, where $R_0 = \frac{[(1-u_1)(1-u_3)c\mu\beta_1+(1-u_2)\beta_2 d_2]x_0}{d_2(d+\mu)}$, $R_{01} = \frac{(1-u_1)(1-u_3)c\mu\beta_1 x_0}{d_2(d+\mu)}$, $R_{02} = \frac{(1-u_2)\beta_2 d_2 x_0}{d_2(d+\mu)}$ and $x_0 = \frac{K}{2r}[(r-d) + \sqrt{(r-d)^2 + \frac{4rs}{K}}]$ is the equilibrium density of $CD4^+T$ cells in infection-free state.*

Proof The basic reproduction number is determined using the next generation matrix [26]. The Jacobian matrix J_{11} of the system (2) at $(x_0, 0, 0, 0)$ is

$$J_{11} = \begin{pmatrix} -d + r(1 - \frac{2x_0}{K}) & \frac{-r\alpha x_0}{K} - (1-u_2)\beta_2 x_0 & -(1-u_1)\beta_1 x_0 & 0 \\ 0 & (1-u_2)\beta_2 x_0 - (d+\mu) & (1-u_1)\beta_1 x_0 & 0 \\ 0 & (1-u_3)c\mu & -d_2 & 0 \\ 0 & p & 0 & -d_3 \end{pmatrix}.$$

The sub-matrix of J_{11} associated with the infectious compartments can be written as

$$\begin{aligned} J_{12} &= \begin{pmatrix} (1-u_2)\beta_2 x_0 - (d+\mu) & (1-u_1)\beta_1 x_0 \\ (1-u_3)c\mu & -d_2 \end{pmatrix} \\ &= \begin{pmatrix} (1-u_2)\beta_2 x_0 & (1-u_1)\beta_1 x_0 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} (d+\mu) & 0 \\ -(1-u_3)c\mu & d_2 \end{pmatrix} = F - V. \end{aligned}$$

Then the next generation matrix [26] is defined as

$$FV^{-1} = \frac{1}{d_2(d+\mu)} \begin{pmatrix} (1-u_1)(1-u_3)c\mu\beta_1 x_0 + (1-u_2)d_2\beta_2 x_0 & 0 \\ (1-u_1)(d+\mu)\beta_1 x_0 & 0 \end{pmatrix}^T.$$

The basic reproduction number R_0 is then obtained from the spectral radius [26] of the matrix FV^{-1} as

$$\begin{aligned} R_0 &= \frac{[(1-u_1)(1-u_3)c\mu\beta_1 + (1-u_2)\beta_2 d_2]x_0}{d_2(d+\mu)} \\ &= (1-u_1)\beta_1 x_0 \frac{1}{(d+\mu)}(1-u_3)c\mu\frac{1}{d_2} + (1-u_2)\beta_2 x_0 \frac{1}{(d+\mu)}. \end{aligned} \tag{5}$$

In the expression of R_0, the terms $(1-u_1)\beta_1$ and $(1-u_2)\beta_2$ indicate, respectively, the effective infection rate corresponding to the cell-free and cell-to-cell infection modes; x_0 is the equilibrium value of $CD4^+T$ cells in the absence of infection; $\frac{1}{(d+\mu)}$ is the average life span of infected $CD4^+T$ cells; $(1-u_3)c\mu$ is the number

of virus produced from the infected $CD4^+T$ cells and $\frac{1}{d_2}$ is the average life span of virus.
For cell-free infection mode (i.e., $\beta_2 = 0$), J_{11} can be rewritten as

$$J_{21} = \begin{pmatrix} -d + r(1 - \frac{2x_0}{K}) & \frac{-r\alpha x_0}{K} & -(1-u_1)\beta_1 x_0 & 0 \\ 0 & -(d+\mu) & (1-u_1)\beta_1 x_0 & 0 \\ 0 & (1-u_3)c\mu & -d_2 & 0 \\ 0 & p & 0 & -d_3 \end{pmatrix}.$$

The sub-matrix of J_{21} associated with the infectious compartments can be written as

$$\begin{aligned} J_{22} &= \begin{pmatrix} -d_1 & (1-u_1)\beta_1 x_0 \\ (1-u_3)Nd_1 & -d_2 \end{pmatrix} \\ &= \begin{pmatrix} 0 & (1-u_1)\beta_1 x_0 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} (d+\mu) & 0 \\ -(1-u_3)c\mu & d_2 \end{pmatrix} = F_1 - V_1. \end{aligned}$$

One can then similarly compute the basic reproduction number in the case of cell-free infection mode as

$$R_{01} = \frac{(1-u_1)(1-u_3)c\mu\beta_1 x_0}{d_2(d+\mu)} = (1-u_1)\beta_1 x_0 \frac{1}{(d+\mu)}(1-u_3)c\mu\frac{1}{d_2}.$$

Similarly, the basic reproduction number R_{02} corresponding to the cell-to-cell infection mode (i.e., $\beta_1 = 0$) is

$$R_{02} = \frac{(1-u_2)d_2\beta_2 x_0}{d_2(d+\mu)}.$$

Thus, the basic reproduction number of the system is the sum of the basic reproduction numbers of two subsystems and is given by $R_0 = R_{01} + R_{02}$. □

3.2 Equilibria of the System

It is easy to see that the system (2) has two equilibrium points. A disease-free equilibrium point $E_0 = (x_0, 0, 0, 0)$, where $x_0 = \frac{K}{2r}[(r-d) + \sqrt{(r-d)^2 + \frac{4rs}{K}}]$ and an infected equilibrium point $E^* = (x^*, y^*, v^*, z^*)$, where $x^* = \frac{1}{2A}(-2B + \sqrt{B^2 - 4As})$, $y^* = \frac{d_3(d+\mu)}{d_1 p}(\frac{x^* R_0}{x_0} - 1)$, $v^* = \frac{(1-u_3)c\mu y^*}{d_2}$, $z^* = \frac{py^*}{d_3}$. Here x^* is the positive root of the quadratic equation

$$Ax^{*2} - Bx^* - s = 0, \tag{6}$$

where $A = \frac{r}{K}(1 + \frac{\alpha d_3(d+\mu)R_0}{d_1 p x_0}) + \frac{d_3(d+\mu)^2 R_0^2}{d_1 p {x_0}^2}$ (> 0) and $B = (r - d) + \frac{r\alpha d_3(d+\mu)}{d_1 K p} + \frac{d_3(d+\mu)^2 R_0}{d_1 p x_0}$ (> 0). As $A > 0$, $B > 0$, Eq. (6) always has a unique positive root. Note that y^* exists if and only if $R_0 > \frac{x_0}{x^*}$. Susceptible $CD4^+T$ cells attains its maximum value x_0 in the absence of infection, giving $x_0 > x^*$. Hence, a sufficient condition for the existence of y^* is $R_0 > 1$. If y^* exists, then v^* exists if $0 \leq u_3 < 1$. Thus, there exists a unique infected equilibrium point E^* if

(i) $R_0 > 1$, (ii) $u_1 = 1 = u_2$ does not hold simultaneously and

(iii) $u_3 \in [0, 1)$.

3.3 Stability of the Equilibrium Points

Consider the perturbations

$$X(t) = x(t) - \bar{x},\ Y(t) = y(t) - \bar{y},\ V(t) = v(t) - \bar{v},\ Z(t) = z(t) - \bar{z},$$

where $(\bar{x},\ \bar{y},\ \bar{v},\ \bar{z})$ is any arbitrary equilibrium point of the system (2). Then the linearized system can be written in the matrix form as

$$\frac{dQ}{dt} = MQ(t) + NQ(t - \tau),$$

where

$$M = \begin{pmatrix} a_{11} & a_{12} & a_{13} & 0 \\ a_{21} & a_{22} & a_{23} & a_{24} \\ 0 & a_{32} & a_{33} & 0 \\ 0 & 0 & 0 & a_{44} \end{pmatrix},\ N = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & b_{42} & 0 & 0 \end{pmatrix}, \tag{7}$$

$$Q(t) = \big(X(t)\ Y(t)\ V(t)\ Z(t)\big)^T,$$

and the entries of the matrices M and N are

$$\begin{cases} a_{11} = -d + r\left(1 - \frac{2\bar{x} + \alpha\bar{y}}{K}\right) - (1 - u_1)\beta_1\bar{v} - (1 - u_2)\beta_2\bar{y}, \\ a_{12} = -\frac{r\alpha\bar{x}}{K} - (1 - u_2)\beta_2\bar{x},\ a_{13} = -(1 - u_1)\beta_1\bar{x}, \\ a_{21} = (1 - u_1)\beta_1\bar{v} + (1 - u_2)\beta_2\bar{y},\ a_{22} = (1 - u_2)\beta_2\bar{x} - (d + \mu) - d_1\bar{z}, \\ a_{23} = (1 - u_1)\beta_1\bar{x},\ a_{24} = -d_1\bar{y},\ a_{32} = (1 - u_2)c\mu,\ a_{33} = -d_2, \\ a_{44} = -d_3,\ b_{42} = p. \end{cases}$$

The corresponding characteristic equation is given by

$$\Phi(\lambda, \tau) = \det\left(\lambda I - M - e^{-\lambda\tau} N\right) = 0. \tag{8}$$

We have the following theorem for the stability of the infection-free equilibrium point.

Theorem 3.1 *The infection-free equilibrium point E_0 is locally and globally asymptotically stable for all delay $\tau \geq 0$ if $R_0 < 1$.*

Proof The characteristic equation in this case becomes

$$\begin{aligned}(\lambda + \frac{s}{x_0} + \frac{r x_0}{K})(\lambda + d_3)\Big(\lambda^2 + (d_2 + (d+\mu) - (1-u_2)\beta_2 x_0)\lambda \\ + d_2(d+\mu)(1-R_0)\Big) = 0.\end{aligned} \tag{9}$$

This equation has two negative real roots, $\lambda_1 = -(\frac{s}{x_0} + \frac{r x_0}{K})$ and $\lambda_2 = -d_3$. The other two are the roots of the equation

$$\lambda^2 + (d_2 + (d+\mu) - (1-u_2)\beta_2 x_0)\lambda + d_2(d+\mu)(1-R_0) = 0. \tag{10}$$

Thus, if $R_0 < 1$ then both roots of the Eq. (10) have negative real parts, implying that E_0 is locally asymptotically stable.

To prove the global stability of the disease-free equilibrium E_0, we use Fluctuation Lemma [11]. The following result is true for a continuous and bounded function $f(t)$:

$$f^{\infty} = \limsup_{t\to\infty} f(t), \quad f_{\infty} = \liminf_{t\to\infty} f(t).$$

Since the solutions $x = x(t)$, $y = y(t)$, $v = v(t)$ and $z = z(t)$ of the system (2) are continuous and bounded,

$$\begin{aligned}0 < x_{\infty} \leq x^{\infty} < \infty,\ 0 \leq y_{\infty} \leq y^{\infty} < \infty, \\ 0 \leq v_{\infty} \leq v^{\infty} < \infty,\ 0 \leq z_{\infty} \leq z^{\infty} < \infty.\end{aligned}$$

From Lemma 2.1, $x = x(t)$ is bounded in $(0, x_0]$, for all $t \geq 0$. From Fluctuation Lemma [11], the last three equations of (2) can then be written as

$$\begin{aligned}(d+\mu)y^{\infty} &\leq (1-u_1)\beta_1 x_0 v^{\infty} + (1-u_2)\beta_2 x_0 y^{\infty}, \\ d_2 v^{\infty} &\leq (1-u_3)c\mu y^{\infty}, \\ d_3 z^{\infty} &\leq p y^{\infty}.\end{aligned} \tag{11}$$

First two inequalities of (12) lead to

$$y^{\infty}(1 - R_0) \leq 0, \quad \text{where } R_0 \text{ is defined in (5).} \tag{12}$$

Now, suppose $R_0 < 1$, which is the local stability condition of E_0. Then the inequality (12) gives $y^{\infty} \leq 0$. Since y^{∞} is the supremum of $y(t)$, then y^{∞} is nonnegative. Therefore, the possible value of y^{∞} is $y^{\infty} = 0$ provided $R_0 < 1$, and hence $\limsup_{t\to\infty} y(t) = 0$. If $y^{\infty} = 0$ and as $v(t), z(t)$ are nonnegative, from the last two inequalities of (12), we obtain $v^{\infty} = 0$ and $z^{\infty} = 0$. Therefore, $\limsup_{t\to\infty} v(t) = 0$ and $\limsup_{t\to\infty} z(t) = 0$. Following the Fluctuation Lemma [11], the first equation of system (2) yields

$$s - dx^{\infty} + rx^{\infty}\left(1 - \frac{x^{\infty} + \alpha y}{K}\right) - (1-u_1)\beta_1 x^{\infty} v - (1-u_2)\beta_2 x^{\infty} y = 0. \tag{13}$$

From Lemma 2.1, all solutions are nonnegative and if $R_0 < 1$, then $\limsup_{t\to\infty} y(t) = 0$ and henceforth $\limsup_{t\to\infty} v(t) = 0$, $\limsup_{t\to\infty} z(t) = 0$. Therefore, if $R_0 < 1$, the solution of $y(t)$ should be $y(t) = 0$ and hence $v(t) = 0$, $z(t) = 0$. The Eq. (13) then becomes

$$s - dx^{\infty} + rx^{\infty}\left(1 - \frac{x^{\infty}}{K}\right) = 0$$
$$(x^{\infty} - x_n)(x_0 - x^{\infty}) = 0,$$

This equation has two roots, $x_0 = \frac{K}{2r}[(r-d) + \sqrt{(r-d)^2 + \frac{4rs}{K}}] > 0$ is the equilibrium density of $CD4^+T$ cells in the infection-free state, and $x_n = \frac{K}{2r}[(r-d) - \sqrt{(r-d)^2 + \frac{4rs}{K}}]$ (< 0). Since x^{∞} is nonnegative, therefore $x^{\infty} = \limsup_{t\to\infty} x(t) = x_0$. Therefore, the disease-free steady state, E_0, is globally asymptotically stable if $R_0 < 1$. Note that the stability of the disease-free equilibrium does not depend on the value of τ. □

Below we prove the stability of the infected equilibrium point in two cases.

Theorem 3.2 *If E^* exists and $\tau = 0$, then the infected steady state E^* is locally asymptotically stable if and only if $a_1 > 0$, $(a_3+b_1) > 0$, $(a_4+b_2) > 0$, $a_1(a_2+b_0)(a_3+b_1)-(a_3+b_1)^2-a_1^2(a_4+b_2) > 0$, where $a_i > 0, b_j > 0, i = 1, 2, 3, 4, j = 0, 1, 2$, are given below.*

Proof The characteristic equation (8) at E^* is given by

$$\Phi(\lambda, \tau) = \lambda^4 + a_1\lambda^3 + a_2\lambda^2 + a_3\lambda + a_4 + (b_0\lambda^2 + b_1\lambda + b_2)e^{-\lambda\tau} = 0, \tag{14}$$

where

$$\begin{cases} a_1 = (\frac{s}{x^*} + \frac{rx^*}{K}) + d_2 + d_3 + \frac{(1-u_1)(1-u_3)c\mu\beta_1 x^*}{d_2}, \\ a_2 = (\frac{s}{x^*} + \frac{rx^*}{K})(d_2 + d_3 + \frac{(1-u_1)(1-u_3)c\mu\beta_1 x^*}{d_2}) \\ \quad + d_3(d_2 + \frac{(1-u_1)(1-u_3)c\mu\beta_1 x^*}{d_2}) + \frac{(d+\mu)y^* R_0}{x_0}(\frac{r\alpha x^*}{K} + (1-u_2)\beta_2 x^*), \\ a_3 = d_3(\frac{s}{x^*} + \frac{rx^*}{K})(d_2 + \frac{(1-u_1)(1-u_3)c\mu\beta_1 x^*}{d_2}) \\ \quad + \frac{(d+\mu)y^* R_0}{x_0}[(1-u_1)(1-u_3)c\mu\beta_1 x^* \\ \quad + (d_2+d_3)(\frac{r\alpha x^*}{K} + (1-u_2)\beta_2 x^*)], \\ a_4 = \frac{d_3(d+\mu)y^* R_0}{x_0}[\frac{d_2 r\alpha x^*}{K} + \frac{d_2(d+\mu)x^* R_0}{x_0}],\ b_0 = d_1 p y^*, \\ b_1 = d_1 p y^*(d_2 + \frac{s}{x^*} + \frac{rx^*}{K}),\ b_2 = d_1 d_2 p y^*(\frac{s}{x^*} + \frac{rx^*}{K}). \end{cases}$$

Note that $a_i, b_j, i = 1,2,3,4, j = 0,1,2$ are all positive. At $\tau = 0$, the characteristic equation (14) then becomes

$$\Phi(\lambda, 0) = \lambda^4 + a_1\lambda^3 + (a_2 + b_0)\lambda^2 + (a_3 + b_1)\lambda + (a_4 + b_2) = 0.$$

According to Routh-Hurwitz criteria, E^* will be locally asymptotically stable if and only if the conditions mentioned in the theorem hold. □

Theorem 3.3 *Assume that E^* is stable in the absence of delay and the conditions given in the Theorem (3.2) are satisfied. If (18) has at least one positive root following one of the conditions specified in (20), then there exists a critical value $\tau = \tau^*$, where τ^* is defined in (21), for which E^* is locally asymptotically stable if $\tau \in [0, \tau^*)$ and unstable for $\tau > \tau^*$. The switching of stability occurs through a Hopf bifurcation at $\tau = \tau^*$.*

For $\tau > 0$, we first investigate whether the Eq. (14) has a pair of purely imaginary roots of the form $\lambda = \pm i\omega^*$, $\omega^* \in \mathbb{R}^+ - \{0\}$ for some parametric conditions. In such a case, putting $\lambda = i\omega^*$ in (14), one gets

$$\begin{aligned} &(i\omega^*)^4 + a_1(i\omega^*)^3 + a_2(i\omega^*)^2 + a_3(i\omega^*) + a_4 \\ &+ (b_0(i\omega^*)^2 + b_1(\omega^*) + b_2)e^{-(i\omega^*)\tau} = 0. \end{aligned}$$

Separating real and imaginary parts, we have

$$\begin{cases} S_1 = S_3\cos(\omega^*\tau) - S_4\sin(\omega^*\tau), \\ S_2 = S_3\sin(\omega^*\tau) + S_4\cos(\omega^*\tau), \end{cases} \tag{15}$$

where

$$\begin{aligned} &S_1 = (\omega^*)^4 - a_2(\omega^*)^2 + a_4,\ S_2 = a_1(\omega^*)^3 - a_3\omega^*, \\ &S_3 = b_0(\omega^*)^2 - b_2,\ S_4 = b_1\omega^*. \end{aligned} \tag{16}$$

Summing up the squares of the equations in (15), we get

$$S_1^2 + S_2^2 = S_3^2 + S_4^2 \tag{17}$$

Let $z = (\omega^*)^2$. Then (17) becomes

$$H(z) = z^4 + A_1 z^3 + A_2 z^2 + A_3 z + A_4 = 0, \tag{18}$$

where $A_1 = a_1^2 - 2a_2$, $A_2 = a_2^2 + 2a_4 - 2a_1 a_3 - b_0^2$, $A_3 = a_3^2 - 2a_2 a_4 - b_1^2 + 2b_0 b_2$, $A_4 = a_4^2 - b_2^2$. Therefore, $\Phi(\lambda, \tau) = 0$ has a purely imaginary root $i\omega^*$ if $H(z) = 0$ has a positive real root. Differentiation of (18) yields

$$H'(z) = 4z^3 + 3A_1 z^2 + 2A_2 z + A_3.$$

Let $y = z + \frac{A_1}{4}$, so that $H'(z) = 0$ becomes

$$y^3 + n_1 y^2 + n_2 = 0,$$

where $n_1 = \frac{A_2}{2} - \frac{3A_1^2}{16}$, $n_2 = \frac{A_1^3}{32} - \frac{A_1 A_2}{8} + \frac{A_3}{4}$.
Define, $\Gamma = (\frac{n_2}{2})^2 + (\frac{n_1}{3})^3$, $\rho = \frac{-1+i\sqrt{3}}{2}$. We then get [10]

$$\begin{aligned}
y_1 &= (-\frac{n_2}{2} + \sqrt{\Gamma})^{1/3} + (-\frac{n_2}{2} - \sqrt{\Gamma})^{1/3}, \\
y_2 &= (-\frac{n_2}{2} + \sqrt{\Gamma})^{1/3}\rho + (-\frac{n_2}{2} - \sqrt{\Gamma})^{1/3}\rho^2, \\
y_3 &= (-\frac{n_2}{2} + \sqrt{\Gamma})^{1/3}\rho^2 + (-\frac{n_2}{2} - \sqrt{\Gamma})^{1/3}\rho, \\
z_l &= y_l - \frac{A_1}{4},\ l = 1, 2, 3.
\end{aligned} \tag{19}$$

Following [40], the existence of positive roots of the equation $H(z) = 0$ can then be asserted as

$$\begin{aligned}
&(i)\ \text{If } A_4 < 0, \text{ then } H(z) \text{ has at least one positive root.} \\
&(ii)\ \text{If } A_4 \geq 0 \text{ and } \Gamma \geq 0, \text{ then } H(z) = 0 \text{ has positive roots} \\
&\quad \text{iff } z_1 > 0 \text{ and } H(z_1) < 0. \\
&(iii)\ \text{If } A_4 > 0 \text{ and } \Gamma < 0, \text{ then } H(z) = 0 \text{ has positive roots} \\
&\quad \text{iff there exists at least one } z_* \in \{z_1, z_2, z_3\} \\
&\quad \text{such that } z_* > 0 \text{ and } H(z_*) \leq 0.
\end{aligned} \tag{20}$$

Without loss of generality, we assume that $H(z) = 0$ has four positive roots, say $z_k^*, k = 1, 2, 3, 4$. Let $\omega_k^* = \sqrt{z_k^*}$. From (15), one then finds

$$\cos(\omega_k^*\tau^*) = H_1 = \frac{S_1S_3 + S_2S_4}{S_3^2 + S_4^2},$$
$$\sin(\omega_k^*\tau^*) = H_2 = \frac{S_1S_3 - S_2S_4}{S_3^2 + S_4^2}.$$

Here S_1, S_2, S_3, S_4 are calculated form (16) with $\omega^* = \omega_k^*$. Define

$$\tau_j^{(k)} = \begin{cases} \frac{1}{\omega_k^*}[\arccos(H_1) + 2\pi j], & H_2 \geq 0, \\ \frac{1}{\omega_k^*}[2\pi - \arcsin(H_1) + 2\pi j], & H_2 < 0, \end{cases}$$

where, $k = 1, 2, 3, 4$, $j = 0, 1, 2, 3, \ldots$, and

$$\tau^* = \min_{1\leq k\leq 4, j\geq 0} \tau_j^{(k)}. \tag{21}$$

Let $\hat{\omega}_k^*$ be the value of ω_k^* ($k = 1, 2, 3, 4$) for which τ^* is obtained. Hence $\hat{\omega}_k^{*3} = \sqrt{\hat{z}_k^*}$.

To show that the transversality condition of Hopf bifurcation at $\tau = \tau^*$ is also hold, we differentiate the characteristic equation (14) with respect to τ to obtain

$$(4\lambda^3 + 3a_1\lambda^2 + 2a_2\lambda + a_3)\frac{d\lambda}{d\tau} + (2b_0\lambda + b_1)e^{-\lambda\tau}\frac{d\lambda}{d\tau}$$
$$- e^{-\lambda\tau}(b_0\lambda^2 + b_1\lambda + b_2)(\lambda + \tau\frac{d\lambda}{d\tau}) = 0.$$

One then finds

$$\left(\frac{d\lambda}{d\tau}\right)^{-1} = \frac{4\lambda^3 + 3a_1\lambda^2 + 2a_2\lambda + a_3}{-\lambda(b_0\lambda^2 + b_1\lambda + b_2)} + \frac{2b_0\lambda + b_1}{\lambda(b_0\lambda^2 + b_1\lambda + b_2)} - \frac{\tau}{\lambda}$$
$$= \frac{4\lambda^3 + 3a_1\lambda^2 + 2a_2\lambda + a_3}{-\lambda(\lambda^4 + a_1\lambda^3 + a_2\lambda^2 + a_3\lambda + a_4)} + \frac{2b_0\lambda + b_1}{\lambda(b_0\lambda^2 + b_1\lambda + b_2)} - \frac{\tau}{\lambda}.$$

Using (17), the value of $\left(\frac{d\lambda}{d\tau}\right)^{-1}$ at $\tau = \tau^*$ and $\lambda = i\hat{\omega}_k^*$ reads

$$\left[\left(\frac{d\lambda}{d\tau}\right)^{-1}\right]_{\lambda=i\hat{\omega}_k^*,\ \tau=\tau^*} = \frac{N_1N_2 + N_4N_5}{\hat{\omega}_k^*N_3} + \frac{i\tau^*}{\hat{\omega}_k^*},$$

where $N_1 = \left(4\hat{\omega}_k^{*3} - 2a_2\hat{\omega}_k^*\right) + i\left(a_3 - 3a_1\hat{\omega}_k^{*2}\right)$, $N_2 = \left(\hat{\omega}_k^{*4} - a_2\hat{\omega}_k^{*2} + a_4\right) + i\left(a_1\hat{\omega}_k^{*3} - a_3\hat{\omega}_k^*\right)$, $N_3 = \left(b_2 - b_0\hat{\omega}_k^{*2}\right)^2 + \left(b_1\hat{\omega}_k^*\right)^2$, $N_4 = 2b_0\hat{\omega}_k^* - ib_1$, $N_5 = \left(b_2 - b_0\hat{\omega}_k^{*2}\right) - ib_1\hat{\omega}_k^*$.
Substituting $\hat{\omega}_k^* = \sqrt{\hat{z}_k^*}$, we then have

$$\left[Re\left(\frac{d\lambda}{d\tau}\right)^{-1}\right]_{\lambda=i\hat{\omega}_k^*,\ \tau=\tau^*} = \frac{4\hat{z}_k^{*3} + 3A_1\hat{z}_k^{*2} + 2A_2\hat{z}_k^* + A_3}{N_3}$$

$$= \frac{H'(\hat{z}_k^*)}{N_3} \neq 0, \text{ if } H'(\hat{z}_k^*) \neq 0,$$

where $H(z)$ is defined in (18) .
Therefore, the sign of $\left[Re\left(\frac{d\lambda}{d\tau}\right)^{-1}\right]_{\lambda=i\hat{\omega}_k^*,\ \tau=\tau^*}$ is same as $H'(\hat{z}_k^*)$. The direction of Hopf bifurcation depends on the sign of the transversality condition. If the value of the transversality condition is positive, then the stability of E^* will change from a stable to an unstable state through a Hopf bifurcation in the forward direction. In contrast, its negative value implies the change from an unstable state to a stable equilibrium through a Hopf bifurcation in the backward direction.

3.4 Simulation Results

System (2) in the absence of delay and control has thirteen parameters. So the question is which parameters are essential and should be selected for further investigation. For this, we have performed a sensitivity analysis (see Fig. 1) of the system parameters (see Table 1) using the Latin Hypercube sampling method. It shows that c and β_2 are the most sensitive parameters. So, we fix other parameter values and consider c and β_2 as the variable parameters for further study.

In Fig. 2, we have drawn the stability region of the disease-free equlilibrium E_0 and the existence region of the infected equilibrium E^* in $c - \beta_2$ parameter plane when other parameters remain fixed with $u_1 = u_2 = u_3 = 0.5$.

Time evolutions of the system (2) with constant controls (Fig. 3, upper row) show that the disease-free equilibrium E_0 is stable when the parameters are selected from the lower region of Fig. 2 ($c = 4$, $\beta_2 = 0.0001$). In fact, the delay has no effect on the stability of E_0. The disease-free equilibrium is globally asymptotically stable for all $\tau \geq 0$ (figures not given). If $c = 75$, $\beta_2 = 0.001$ then the parameter values satisfy the stability conditions of the non-delayed system. Figure 3 (lower row) shows that the non-delayed system is locally asymptotically stable. Figure 4 is the bifurcation diagram of the infected population with respect to the delay parameter, τ. It shows that the system is stable for $\tau < \tau^*$, unstable for $\tau > \tau^*$, and a Hopf bifurcation

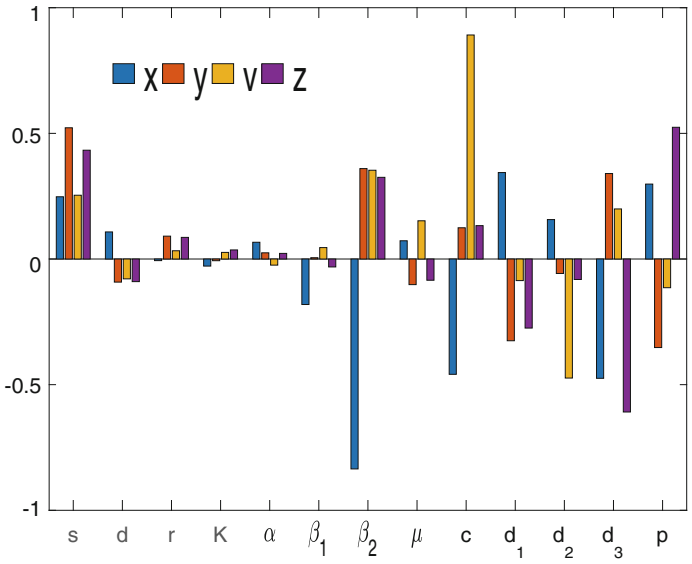


Fig. 1 Sensitivity analysis of the parameters (see Table 1) following Latin hypercube sampling-partial ranked correlation coefficients ($p < 0.00001$). Here $u_1 = u_2 = u_3 = 0$

occurs at $\tau = \tau^* = 5.1051$ days. Time evolutions of the system populations for two particular values of τ ($\tau = 5$ and $\tau = 5.2$) represent the stable and unstable behaviour of the system (Fig. 5).

4 The Optimal Control Problem

In the previous section, we discussed the effect of constant control. Here we assume that the control parameters are time-dependent. Our objective here is to maximize the number of healthy $CD4^+T$ cells, CTLs, and minimize the number of infected $CD4^+T$ cells and virus particles using three controls. At the same time, we want to reduce the deleterious side effects of the drugs. For this, we define the objective functional

$$J(u_1, u_2, u_3) = \int_0^{t_f} \left(A_1 x(t) + A_2 z(t) - \left(A_3 y(t) + A_4 v(t) + \frac{B_1 u^2{}_1(t)}{2} + \frac{B_2 u^2{}_2(t)}{2} + \frac{B_3 u^2{}_3(t)}{2} \right) \right) dt \tag{22}$$

subject to the states (2). In the integrand, the first two terms represent the benefit and the next two terms indicate the deleterious effects. Here $A_i > 0$, $i = 1, 2, 3, 4$, are

Table 1 Parameter descriptions and their values with references

Parameter	Description	Range of parameters	Default values	References
s	Constant input rate of $CD4^+T$ cells	0-10 cells mm^{-3} day^{-1}	10	[27, 29–31]
d	Death rate of susceptible $CD4^+T$ cells	0.07-0.1 day^{-1}	0.01	[10, 30, 31]
β_1	Cell-free disease transmission coefficient	0.000025-0.5 virons mm^{-3} day^{-1}	0.00025	[10, 27, 29, 31]
β_2	Cell-to-cell disease transmission coefficient	0.00001-0.7 mm^{-3} infected cells day^{-1}	varies	
r	Proliferation rate of $CD4^+T$ cells	0.03–3 day^{-1}	0.1	[10, 27, 29, 31]
K	Maximum density of $CD4^+T$ cells where proliferation stops	1500 mm^{-3}	1500	[10, 29, 31]
α	Limitation coefficient of infected cells imposed on the proliferation of $CD4^+T$ cells	≥ 1	1.2	[9, 10]
d_1	Killing rate of infected cells by CTL	0.812 day^{-1}	0.812	[10, 32]
d_2	Clearance rate of virus particles	2.4–3 day^{-1}	2.4	[10, 29, 31]
d_3	Clearance rate of CTL	1.618 day^{-1}	1.618	[10, 32]
c	Virus replication factor	10–2500 virions $cell^{-1}$	varies	[27, 28]
p	Production rate of CTL	0.05 day^{-1}	0.05	[10, 33]
μ	Lysis death rate of infected $CD4^+T$ cells	0.2–0.5 day^{-1}	0.4	[29, 31]

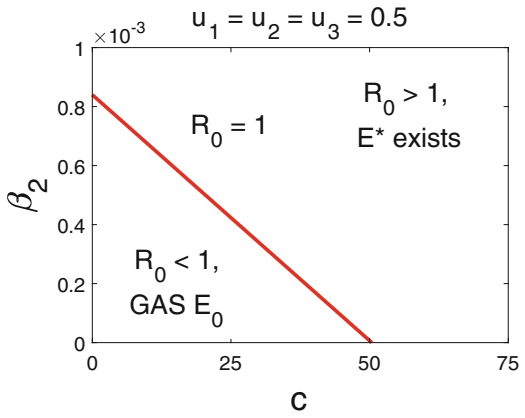


Fig. 2 Stability region of the disease-free equilibrium E_0 and the existence region of the endemic equilibrium E^* in $c - \beta_2$ plane. Parameters are $s = 10, d = 0.02, \beta_1 = 0.00025, r = 0.03, K = 1500, \alpha = 1.2, \mu = 0.4, d_1 = 0.812, d_2 = 3, d_3 = 1.618, p = 0.05$ and $u_1 = u_2 = u_3 = 0.5$

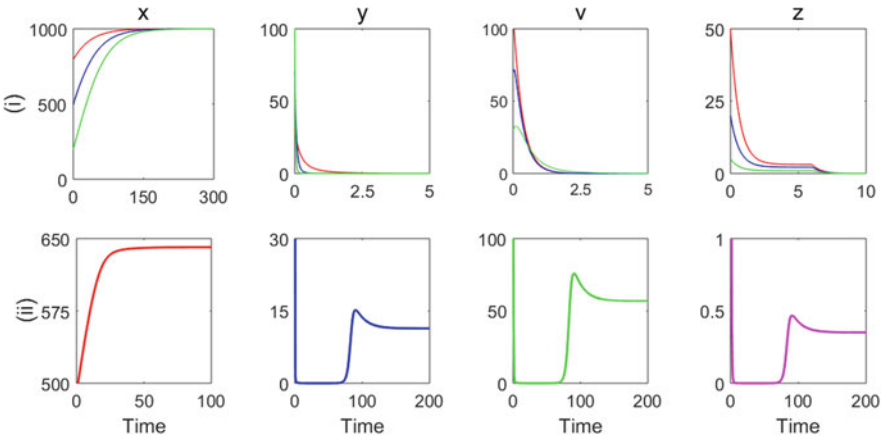


Fig. 3 Upper row: Time evolutions of the non-delayed system show that the disease-free equilibrium E_0 is stable when parameters are selected from the lower region of Fig. 2 with $c = 4, \beta_2 = 0.0001$. Lower row: Figures show that the non-delayed system is locally asymptotically stable for $c = 75, \beta_2 = 0.001$. Other parameters are as in Fig. 2

the weight constants, which balance the size of the terms $x(t)$,$z(t)$,$y(t)$ and $v(t)$. Three square terms are the respective cost of $u_1(t), u_2(t)$ and $u_3(t)$, and B_1, B_2 and B_3 are the weight parameters employed relative to the cost implication of the controls u_1, u_2 and u_3. Here t_f is the final time, where treatment stops. Therefore, our object is to find the optimal control triplet $(\hat{u_1}, \hat{u_2}, \hat{u_3})$ such that

$$J(\hat{u_1}, \hat{u_2}, \hat{u_3}) = \max_{(u_1,u_2,u_3)\in\Xi} J(u_1, u_2, u_3), \tag{23}$$

where, $\Xi = \{(u_1(t), u_2(t), u_3(t)) : u_i \; measurable, 0 \leq u_i(t) \leq 1, t \in [0, t_f], i = 1, 2, 3\}$.

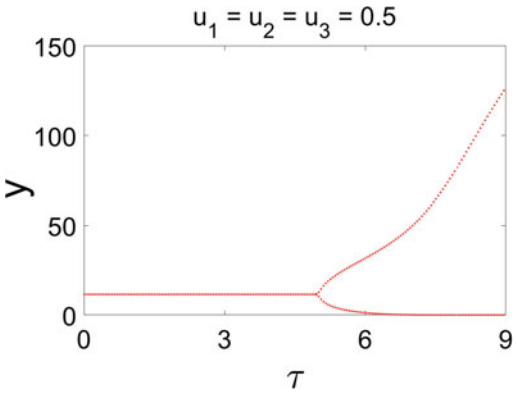


Fig. 4 Bifurcation diagram of the infected $CD4^+T$ cell population with respect to the delay, τ. The y population is stable for $\tau < \tau^*$ and unstable for $\tau > \tau^*$, where $\tau^* = 5.1051$ days. Here $u_1 = u_2 = u_3 = 0.5$, $c = 75$, $\beta_2 = 0.001$ and other parameters are as in Fig. 2

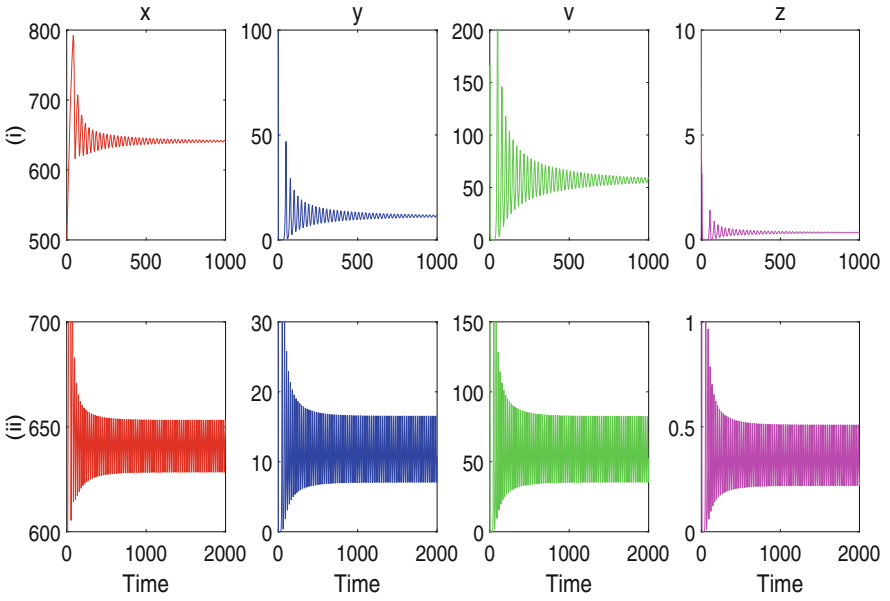


Fig. 5 Upper row: The infected equilibrium E^* is stable when $\tau = 5 (< \tau^*)$. Lower row: The infected equilibrium E^* is unstable when $\tau = 5.2 (> \tau^*)$, where $\tau^* = 5.1051$. Here $u_1 = u_2 = u_3 = 0.5$ and $c = 75$, $\beta_2 = 0.001$. Other parameters are as in Fig. 4

4.1 Existence of an Optimal Control Triplet

Proposition 4.1 *There exists an optimal control triplet* $(\hat{u_1}, \hat{u_2}, \hat{u_3}) \in \Xi$ *with time dependent control problem (2) that maximizes the objective functional* $J(u_1, u_2, u_3)$, *i.e.,* $J(\hat{u_1}, \hat{u_2}, \hat{u_3}) = \max_{(u_1,u_2,u_3)\in\Xi} J(u_1, u_2, u_3)$.

Proof We use an existence result of Flaming and Rishel [34] and check the following properties:

(a_1) The set of controls and corresponding state variables is nonempty.
(a_2) The control set Ξ is convex and closed.
(a_3) The right hand side of the state system is continuous, bounded above by a sum of the bounded control and state, and can be written as a linear function of u_i with coefficients depending on the state and time.
(a_4) The integrand of the objective functional is concave on Ξ.
(a_5) There exists constants $c_1,\ c_2 > 0$ and $b > 1$ such that the integrand of the objective functional is bounded above by $c_2 - c_1(|u_1(t)|^2 + |u_2(t)|^2 + |u_3(t)|^2)^{b/2}$.

In order to verify these properties, we use a result from Lukes [35] for the existence of solutions of (2) with bounded coefficients and (a_1) is satisfied. By the definition of Ξ, (a_2) is satisfied. As our control system is linear in u_1, u_2 and u_3, the right-hand side of (2) satisfies (a_3) as the solutions are bounded. The integrand of the objective functional is concave for the control set Ξ and hence a_4 is satisfied. For the last condition

$$\begin{aligned} &A_1x + A_2z - \left(A_3y + A_4v + \frac{B_1{u^2}_1}{2} + \frac{B_2{u^2}_2}{2} + \frac{B_3{u^2}_3}{2}\right) \\ &\leq A_1x + A_2z - (\frac{B_1{u_1}^2}{2} + \frac{B_2{u_2}^2}{2} + \frac{B_3{u_3}^2}{2}) \\ &\leq c_2 - c_1(|u_1|^2 + |u_2|^2 + |u_3|^2)^{\frac{b}{2}}, \end{aligned}$$

where c_2 depends on the upper bound of x and z, $b > 1$ and $c_1 > 0$ as $B_1, B_2, B_3 > 0$. Hence, we conclude that there exists an optimal control triplet. □

Pontryagin's minimum principle [36] and state delay provides a necessary condition for an optimal control triplet $(\hat{u_1}, \hat{u_2}, \hat{u_3})$. This principle converts (2), (22) and (23) into a problem which maximizes the Hamiltonian (H)

$$\begin{aligned} H(t; x, y, v, z; y_\tau; u_1, u_2, u_3; \lambda_1, \lambda_2, \lambda_3, \lambda_4) &= A_3y(t) + A_4v(t) + \frac{B_1{u_1}^2(t)}{2} \\ &+ \frac{B_2{u_2}^2(t)}{2} + \frac{B_3{u_3}^2(t)}{2} - A_1x(t) - A_2z(t) \\ &+ \lambda_1\dot{x}(t) + \lambda_2\dot{y}(t) + \lambda_3\dot{v}(t) + \lambda_4\dot{z}(t) \end{aligned} \tag{24}$$

where $\lambda_i, i = 1, 2, 3, 4$, are the co-state or adjoint variables. Applying Pontryagin's minimum principle with state delay [36], we obtain the following Proposition.

Proposition 4.2 *Suppose $(\hat{u}_1, \hat{u}_2, \hat{u}_3)$ is an optimal control triplet of (22) subject to the system (2) and $(\hat{x}, \hat{y}, \hat{v}, \hat{z})$ is the corresponding optimal solutions of (2), then there exists co-state or adjoint variables $\lambda_i, (i = 1, 2, 3, 4)$ such that the following conditions are satisfied with the delay-induced controlled system (2).*

i. Co-state equations:

$$\begin{aligned}
\dot{\lambda}_1(t) &= A_1 + \lambda_1(t)\Big(d + (1-\hat{u}_1(t))\beta_1\hat{v}(t) \\
&\quad + (1-\hat{u}_2(t))\beta_2\hat{y}(t) - r(1-\frac{2\hat{x}(t)+\alpha\hat{y}(t)}{K})\Big) \\
&\quad - \lambda_2(t)\Big((1-\hat{u}_1(t))\beta_1\hat{v}(t) + (1-\hat{u}_2(t))\beta_2\hat{y}(t)\Big), \\
\dot{\lambda}_2(t) &= -A_3 + \lambda_1(t)\Big(\frac{r\alpha\hat{x}(t)}{K} + (1-\hat{u}_2(t))\beta_2\hat{x}(t)\Big) \\
&\quad - \lambda_2(t)\Big((1-\hat{u}_2(t))\beta_2\hat{x}(t) - (d+\mu) - d_1\hat{z}(t)\Big) \\
&\quad - \lambda_3(t)(1-\hat{u}_3(t))c\mu - \chi_{[0,t_f-\tau]}p\lambda_4(t+\tau), \\
\dot{\lambda}_3(t) &= -A_4 + \lambda_1(t)(1-\hat{u}_1(t))\beta_1\hat{x}(t) \\
&\quad - \lambda_2(t)(1-\hat{x}_1(t))\beta_1\hat{x}(t) + \lambda_3(t)d_2, \\
\dot{\lambda}_4(t) &= A_2 + \lambda_2(t)d_1\hat{y}(t) + \lambda_4(t)d_3,
\end{aligned} \tag{25}$$

with transversality conditions $\lambda_i(t_f) = 0, i = 1, 2, 3, 4.$

ii. Optimality conditions:

$$\begin{aligned}
&H(\hat{x}, \hat{y}, \hat{v}, \hat{z}; \hat{u_1}, \hat{u_2}, \hat{u_3}; \lambda_1, \lambda_2, \lambda_3, \lambda_4) \\
&= H(\hat{x}, \hat{y}, \hat{v}, \hat{z}; u_1, u_2, u_3; \lambda_1, \lambda_2, \lambda_3, \lambda_4),
\end{aligned}$$

which implies

$$\begin{aligned}
\hat{u_1}(t) &= min\Big\{1, max\{0, \frac{(\lambda_2(t)-\lambda_1(t))\beta_1\hat{x}(t)\hat{v}(t)}{B_1}\}\Big\}, \\
\hat{u_2}(t) &= min\Big\{1, max\{0, \frac{(\lambda_2(t)-\lambda_1(t))\beta_2\hat{x}(t)\hat{y}(t)}{B_2}\}\Big\}, \\
\hat{u_3}(t) &= min\Big\{1, max\{0, \frac{\lambda_3(t)c\mu\hat{y}(t)}{B_3}\}\Big\}.
\end{aligned} \tag{26}$$

Proof By Pontryagin's minimum principle with state delay [36], the co-state equations and its transversality conditions can be solved by

$$\begin{aligned}
\dot{\lambda_1}(t) &= -\frac{\partial H(t)}{\partial x}, \qquad \lambda_1(t_f) = 0, \\
\dot{\lambda_2}(t) &= -\frac{\partial H(t)}{\partial y} - \chi_{[0,t_f-\tau]}\frac{\partial H(t+\tau)}{\partial y_\tau}, \qquad \lambda_2(t_f) = 0, \\
\dot{\lambda_3}(t) &= -\frac{\partial H(t)}{\partial v}, \qquad \lambda_3(t_f) = 0, \\
\dot{\lambda_4}(t) &= -\frac{\partial H(t)}{\partial z}, \qquad \lambda_4(t_f) = 0,
\end{aligned} \tag{27}$$

where $\chi_{[0,t_f-\tau]}$ is the characteristic function defined as

$$\chi_{[0,t_f-\tau]} = \begin{cases} 1 \text{ if } t \in [0, t_f - \tau], \\ 0, \text{ otherwise.} \end{cases}$$

The optimal control triplet $\hat{u_1}, \hat{u_2}$ and $\hat{u_3}$ are solved from the optimality conditions

$$\begin{aligned}
&\frac{\partial H(t)}{\partial u_1} = 0, \ at\ u_1(t) = \hat{u_1}(t); \ \frac{\partial H(t)}{\partial u_2} = 0, \ at\ u_2(t) = \hat{u_2}(t); \\
&\frac{\partial H(t)}{\partial u_3} = 0, \ at\ u_3(t) = \hat{u_3}(t)
\end{aligned} \tag{28}$$

and we get

$$\begin{aligned}
\frac{\partial H(t)}{\partial u_1} &= B_1 u_1(t) + (\lambda_1(t) - \lambda_2(t))\beta_1 x(t) v(t) = 0, \\
\frac{\partial H(t)}{\partial u_2} &= B_2 u_2(t) + (\lambda_1(t) - \lambda_2(t))\beta_2 x(t) y(t) = 0, \\
\frac{\partial H(t)}{\partial u_3} &= B_3 u_3(t) - \lambda_3(t) c\mu y(t) = 0.
\end{aligned} \tag{29}$$

We calculate $\hat{u_1}(t)$, $\hat{u_2}(t)$ and $\hat{u_3}(t)$ from the bounds of Ξ in the form of (26). □

Therefore, the optimal system associated with the system (2) is represented by

$$
\begin{aligned}
\frac{d\hat{x}}{dt} =& s - d\hat{x}(t) + r\hat{x}(t)(1 - \frac{\hat{x}(t) + \alpha\hat{y}(t)}{K}) \\
& - (1 - \hat{u}_1(t))\beta_1\hat{x}(t)\hat{v}(t) - (1 - \hat{u}_2(t))\beta_2\hat{x}(t)\hat{y}(t), \\
\frac{d\hat{y}}{dt} =& (1 - \hat{u}_1(t))\beta_1\hat{x}(t)\hat{v}(t) + (1 - \hat{u}_2(t))\beta_2\hat{x}(t)\hat{y}(t) \\
& - (d + \mu)\hat{y}(t) - d_1\hat{y}(t)\hat{z}(t), \\
\frac{d\hat{v}}{dt} =& (1 - \hat{u}_3(t))c\mu\hat{y}(t) - d_2\hat{v}(t), \\
\frac{d\hat{z}}{dt} =& p\hat{y}(t - \tau) - d_3\hat{z}(t), \\
\dot{\lambda}_1(t) =& A_1 + \lambda_1(t)\Big(d + (1 - \hat{u}_1(t))\beta_1\hat{v}(t) \\
& + (1 - \hat{u}_2(t))\beta_2\hat{y}(t) - r(1 - \frac{2\hat{x}(t) + \alpha\hat{y}(t)}{K})\Big) \\
& - \lambda_2(t)\Big((1 - \hat{u}_1(t))\beta_1\hat{v}(t) + (1 - \hat{u}_2(t))\beta_2\hat{y}(t)\Big), \\
\dot{\lambda}_2(t) =& - A_3 + \lambda_1(t)\Big(\frac{r\alpha\hat{x}(t)}{K} + (1 - \hat{u}_2(t))\beta_2\hat{x}(t)\Big) \\
& - \lambda_2(t)\Big((1 - \hat{u}_2(t))\beta_2\hat{x}(t) - (d + \mu) - d_1\hat{z}(t)\Big) \\
& - \lambda_3(t)(1 - \hat{u}_3(t))c\mu - \chi_{[0,t_f-\tau]}p\lambda_4(t + \tau), \\
\dot{\lambda}_3(t) =& - A_4 + \lambda_1(t)(1 - \hat{u}_1(t))\beta_1\hat{x}(t) \\
& - \lambda_2(t)(1 - \hat{x}_1(t))\beta_1\hat{x}(t) + \lambda_3(t)d_2, \\
\dot{\lambda}_4(t) =& A_2 + \lambda_2(t)d_1\hat{y}(t) + \lambda_4(t)d_3, \\
& x(0) = x_0 > 0,\ y(0) = y_0 > 0,\ v(0) = v_0 > 0,\ z(0) = z_0 > 0, \\
& \lambda_i(t_f) = 0,\ i = 1, 2, 3, 4,
\end{aligned}
\tag{30}
$$

where

$$
\begin{aligned}
\hat{u_1}(t) &= min\Big\{1, max\{0, \frac{(\lambda_2(t) - \lambda_1(t))\beta_1\hat{x}(t)\hat{v}(t)}{B_1}\}\Big\}, \\
\hat{u_2}(t) &= min\Big\{1, max\{0, \frac{(\lambda_2(t) - \lambda_1(t))\beta_2\hat{x}(t)\hat{y}(t)}{B_2}\}\Big\}, \\
\hat{u_3}(t) &= min\Big\{1, max\{0, -\frac{\lambda_3(t)c\mu\hat{y}(t)}{B_3}\}\Big\}.
\end{aligned}
$$

4.2 Simulation Results

Here we solve the optimal systems (30) numerically by combination of forward and backward difference approximation methods [37]. The treatment period is continued for 300 days and therefore, the time interval is considered as $[0, t_f]$, where $t_f = 300$. The initial values of the state variables are taken as $x(0) = 500$, $y(0) = 100$, $v(0) = 100$ and $z(0) = 10$ [38]. Since the weight parameters A_1 and A_3 are associated with the $CD4^+T$ cells, we assign the same values for them. Considering the same harmful effects of all the inhibitors, the same value for the weight parameters B_1, B_2, B_3 are considered. Other parameter values remain as in Table 1. We consider both the mono-drug and multi-drug therapies and compare their efficacies in controlling viremia under different delays.

4.2.1 Mono-Drug Therapy

Figure 6 represents various outcomes of mono-drug therapies for the system (30) with $\tau = 1$. The first column of this figure shows the counts of various immune cells and plasma viruses when the inhibitor u_1 is only administered (the case $(u_1 \neq 0, u_2 = 0, u_3 = 0)$). The last row of this figure gives the respective control profile. These figures show that no inhibitor can completely remove the infection, and infected cells persist in all three cases. However, in the case of blocker u_3, virus counts go below detection level, though it exists in the other two cases, but infected $CD4^+T$ cells (y) still persist. It happens due to the cell-to-cell dissemination of infection. Observe that healthy $CD4^+T$ cells count is relatively low in the mono-drug therapy $(u_1 = 0, u_2 \neq 0, u_3 = 0)$ compared to other two cases. Thus, u_2 is the worst mono-drug therapy. The respective control profiles indicate that all the controls should be applied with full efficacy for the entire treatment period, except for some occasional reduction in the u_1 control. This analysis shows that the mono-drug therapy with the blocker u_3 is relatively a better performer because its application can reduce the free virus particles, thereby reducing the chances of cell-free infection.

Figure 7 shows the time evolutions of the optimal system (30) for $\tau = 2$ days. These figures show that plasma concentrations of $CD4^+T$ cells, CTLs and virus particles oscillate in the case of $(u_1 \neq 0, u_2 = 0, u_3 = 0)$ and $(u_1 = 0, u_2 \neq 0, u_3 = 0)$, but they are stable in the case of mono-drug therapy $(u_1 = 0, u_2 = 0, u_3 \neq 0)$ with higher value of susceptible $CD4^+T$ cells and lower value of infected $CD4^+T$ cell counts. The free virus particles count (v) also remains below the detectable level in the latter case. The control profile (last row) also shows a significant difference for u_3 inhibitor compared to the other two controls, where oscillations are predominant.

Similar simulation results are presented when the delay is further increased to $\tau = 2.5 > \tau^*$ (Fig. 8), where $\tau^* = 2.37$ days. These figures show that plasma concentrations of $CD4^+T$ cells, CTLs, and virus particles oscillate in all three

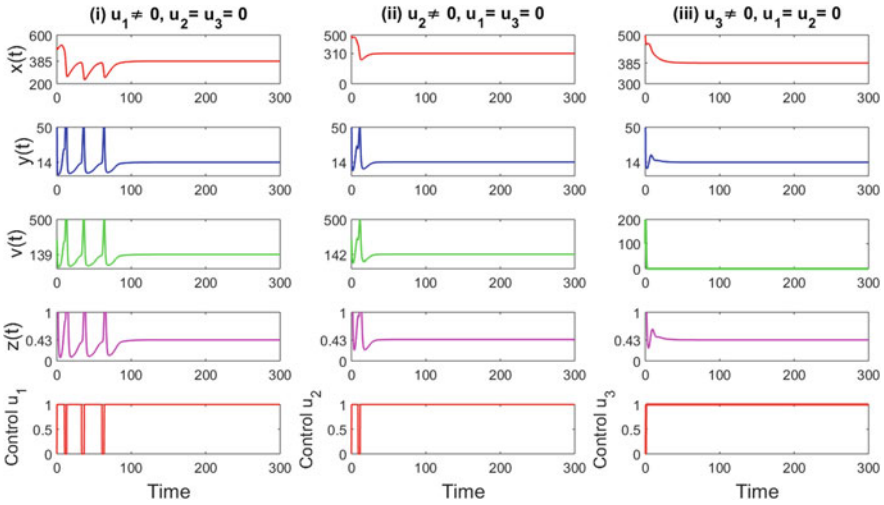


Fig. 6 The time variations of susceptible $CD4^+T$ cells, infected $CD4^+T$ cells, virus particles and CTLs due to different mono-drug therapy with $\tau = 1$ day. The last row represents the optimal controls corresponding to each mono-drug therapy. Here $\tau = 1$ and other parameters are $s = 10, d = 0.02, \beta_1 = 0.00025, r = 0.03, K = 1500, \alpha = 1.2, \mu = 0.4, d_11 = 0.812, d_22 = 3, d_3 = 1.618; p = 0.05; c = 75, \beta_2 = 0.002, A_1 = A_3 = 5, A_2 = A_4 = 2$ and $B_1 = B_3 = B_3 = 0.1$

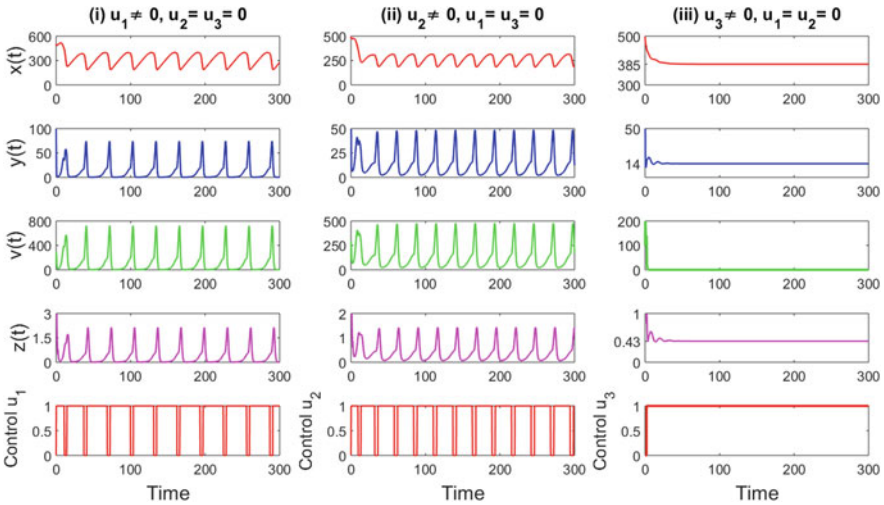


Fig. 7 The time variations of susceptible $CD4^+T$ cells, infected $CD4^+T$ cells, virus particles and CTLs due to different mono-drug therapies with $\tau = 2$ days. The last row represents the control profiles corresponding to each mono-drug therapy. Here $\tau = 2$ and other parameters are as in Fig. 6

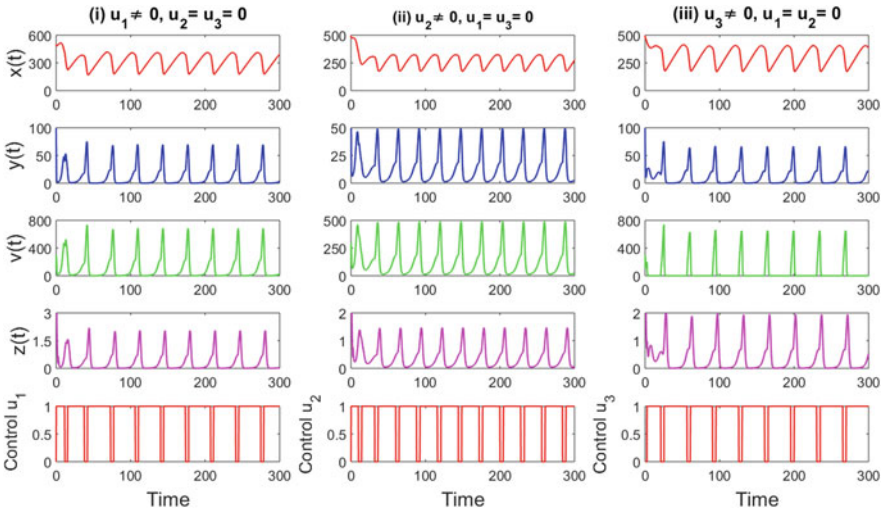


Fig. 8 Concentrations of susceptible $CD4^+T$ cells, infected $CD4^+T$ cells, virus particles and CTLs with respect to time under different mono-drug therapies with $\tau = 2.5$ days. The last row represents the control profiles corresponding to each mono-drug therapy. Here $\tau = 2.5$ and other parameters are as in Fig. 6

control strategies. The control profile also oscillates in each case. Thus, a large immune activation delay causes significant changes in the plasma counts and control profile. Therefore, no mono-drug therapy is capable of controlling the viremia in an HIV-1 infected individual if immune response delay is high.

4.2.2 Multi-Drug Therapy

In the case of a multi-drug therapy, we observed that infection can be removed (i.e., where $y = 0$) in three options. The multi-drug option ($u_1 \neq 0, u_2 = 0, u_3 \neq 0$) cannot eliminate infection, implying that it is the worst combination, but in the other three cases, both the infected cells and virus particles are eliminated (Fig. 9). It is to be noted that the control u_2 should be used while using multi-drug therapy to eliminate the infection. However, the control u_2 was not an efficient blocker when administered alone. Control profiles are all most same for three cases except the one ($u_1 \neq 0, u_2 = 0, u_3 \neq 0$), where the infection persists. Considering the cost of drugs and its side effects, any of the multi-drug therapies ($u_1 \neq 0, u_2 \neq 0, u_3 = 0$) or ($u_1 = 0, u_2 \neq 0, u_3 \neq 0$) may be used instead of ($u_1 \neq 0, u_2 \neq 0, u_3 \neq 0$).

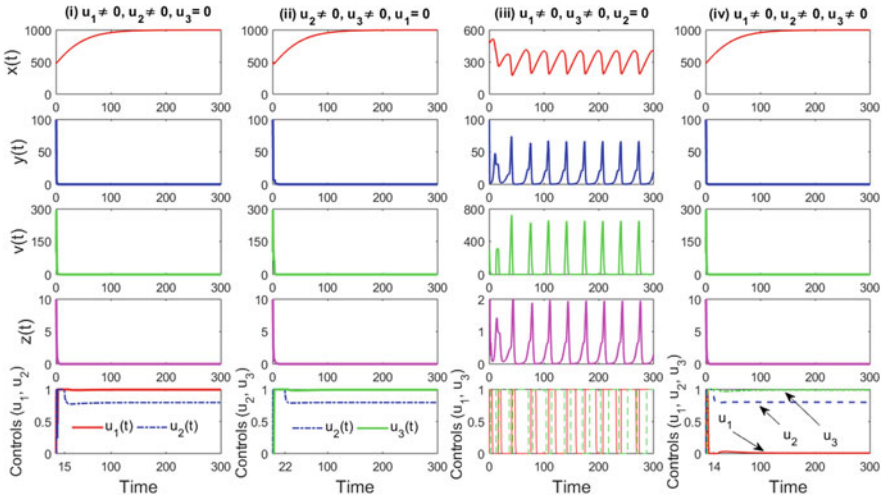


Fig. 9 The time variations of susceptible $CD4^+T$ cells, infected $CD4^+T$ cells, virus particles and CTLs due to different mono-drug therapies with $\tau = 2.5$ days. The last row represents the control profiles corresponding to each mono-drug therapy. Here $\tau = 2.5$ and other parameters are as in Fig. 6

5 Discussion

Recent experimental studies show that cell-to-cell disease transmission mode is a more efficient and faster mode of disease transmission than the cell-free mode in the case of in-host HIV-1 infection [1, 2, 4]. This paper has studied a multi-pathways HIV-1 infection model with immune activation delay. The model is further modified with three different blockers that may be used in controlling the viremia. An RTI control ($u_1(t)$) is used to inhibit the synthesis of viral DNA from HIV-1 RNA to inhibit the viral infectivity. A synapse-forming inhibitor ($u_2(t)$) is used to block the cellular mechanisms required for synapse formation. The third control, a protease inhibitor ($u_3(t)$) is applied to stop the process of free virus formation. We first prove that solutions of our system remain positive for all future time assuming positive initial values. It is also shown that the solutions are uniformly bounded. The analytical results are presented in two phases. In the first phase, we assume that the considered controls are time-independent constant controls. In the second phase, we relaxed this assumption and considered the controls as time-dependent. Applying the next-generation matrix, we calculated the basic reproduction number (R_0) of the system with constant controls and showed the disease-free equilibrium is locally and globally asymptotically stable if $R_0 < 1$. The infected steady state, if it exists, is locally asymptotically stable under some parametric restrictions. Delay may, however, cause instability in the system. There exists some critical value (τ^*) of the delay parameter below which the system is stable and above which it is unstable. The stability switching occurs through a Hopf bifurcation. If

the controls are time-dependent variables, then we defined an objective functional to maximize the healthy $CD4^+T$ cells & CTL cells and minimize the infected $CD4^+T$ cells & virus particles. We derived the necessary conditions for optimal control by applying Pontryagin's minimum principle. It is analytically shown that an optimal control triplet exists that maximizes the objective functional. We have demonstrated the effect of different control measures with mono-drug and multi-drug therapies with different delays using extensive simulation results. It is shown that removing infection is not possible, and the infected cells persist in all three mono-drug protocols. However, in the case of blocker u_3, virus counts (v) go below the detection level, but infected $CD4^+T$ cells (y) persist. This, however, does not happen in the other two controls, where both the infected cells and virus particles survive. Infected $CD4^+T$ cells persist, but the non-zero virus count may be possible due to the presence of cell-to-cell dissemination of infection and the use of blocker u_3. Such a result has not been shown in any previous study. It is observable that $CD4^+T$ cells count in this case is low compared to the other two cases. Thus, u_2 is the worst mono-drug therapy, and u_3 is better. However, when immune response delay increases, then plasma concentrations of $CD4^+T$ cells, CTLs, and virus particles oscillate in all three mono-control strategies, showing uncontrolled behavior. The control profile also oscillates in each case. In the case of a multi-drug therapy, we observed that infection could be removed in three options, where control u_2 is present. The multi-drug option ($u_1 \neq 0, u_3 \neq 0, u_2 = 0$), where u_2 is absent, is the worst one. Our study thus shows that immune response delay significantly affects the system dynamics. If CTL's response is quicker, then $CD4^+T$ cells count may remain stable but fails to do so if response time increases.

Acknowledgments Research of CM supported by CSIR; File No:09/096(0865)/2016-EMR-I. Research of NB is supported by CSIR; Ref. No.: 25(0294)/18/EMR- II.

References

1. Hübner, W. et al., Quantitative 3D video microscopy of HIV transfer across T cell virological synapses. Science 323 (2009) 1743–1747.
2. Zhong, P., Agosto, L. M., Munro, J. B. & Mothes, W. (2013). Cell-to-cell transmission of viruses. Current Opinion in Virology, 3(1), 44–50.
3. Zhong, P. et al., Cell-to-cell transmission can overcome multiple donor and target cell barriers imposed on cell-free HIV, PloS One 8(1) (2013) e53138, DOI: 10.1371/journal.pone.0053138 2013.
4. Chen, P., Hübner, W., Spinelli, M. A. & Chen, B. K. (2007). Predominant mode of human immunodeficiency virus transfer between T cells is mediated by sustained Env-dependent neutralization-resistant virological synapses. Journal of Virology, 81(22), 12582–12595.
5. Song, C., Xu, R., & Bai, N. (2021). Dynamics of a within-host virus infection model with multiple pathways: Stability switch and global stability. International Journal of Bifurcation and Chaos, 31(13), 2150195.
6. Adak, D., & Bairagi, N. (2018). Analysis and computation of multi-pathways and multi-delays HIV-1 infection model. Applied Mathematical Modelling, 54, 517–536.

7. Lai, X., & Zou, X. (2014). Modeling HIV-1 virus dynamics with both virus-to-cell infection and cell-to-cell transmission. SIAM Journal on Applied Mathematics, 74(3), 898–917.
8. Yang, Y., Zou, L., & Ruan, S. (2015). Global dynamics of a delayed within-host viral infection model with both virus-to-cell and cell-to-cell transmissions. Mathematical Biosciences, 270, 183–191.
9. Lai, X. & Zou, X. (2015). Modeling cell-to-cell spread of HIV-1 with logistic target cell growth. Journal of Mathematical Analysis and Applications, 426(1), 563–584.
10. Jinhu Xu, Yicang Zhou, Bifurcation analysis of HIV-1 infection model with cell-to-cell transmission and immune response delay,Mathematical Bioscience and Engineering,13(2016),343–367. %bibitemPB13 Perelson, A. S., Ribeiro, R. M., Modeling the within-host dynamics of HIV infection, BMC Biology 11 (2013) 96, https://doi.org/10.1186/1741-7007-11-96.
11. Hirsch, W.M., Hanisch, H. and Gabriel, J.P., 1985. Differential equation models of some parasitic infections: methods for the study of asymptotic behavior. Communications on Pure and Applied Mathematics, 38(6), pp.733–753.
12. Nowak, M. A., Bonhoeffer, S., Shaw, G. M. & May, R. M. (1997). Anti-viral drug treatment: dynamics of resistance in free virus and infected cell populations. Journal of Theoretical Biology, 184(2), 203–217.
13. Kirschner, D., Lenhart, S., Serbin, S., Optimal control of the chemotherapy of HIV, J. Math. Biol. 35 (1997) 775–792.
14. Bairagi, N., & Adak, D. (2014). Global analysis of HIV-1 dynamics with Hill type infection rate and intracellular delay. Applied Mathematical Modelling, 38(21–22), 5047–5066.
15. Orellana, J., 2009. Optimal control for HIV multitherapy enhancement.
16. Adak, D., & Bairagi, N. (2019). Bifurcation analysis of a multidelayed HIV model in presence of immune response and understanding of in-host viral dynamics. Mathematical Methods in the Applied Sciences, 42(12), 4256–4272.
17. Perelson, A.S. and Nelson, P.W., 1999. Mathematical analysis of HIV-1 dynamics in vivo. SIAM review, 41(1), 3–44.
18. Bairagi, N., & Adak, D. (2017). Dynamics of cytotoxic T-lymphocytes and helper cells in human immunodeficiency virus infection with Hill-type infection rate and sigmoidal CTL expansion. Chaos, Solitons & Fractals, 103, 52–67.
19. Rong, L. and Perelson, A.S., 2009. Modeling HIV persistence, the latent reservoir, and viral blips. Journal of theoretical biology, 260(2), 308–331.
20. Wodarz, D. and Nowak, M.A., 2002. Mathematical models of HIV pathogenesis and treatment. BioEssays, 24(12), 1178–1187.
21. Fister, K. R., Lenhart, S., Mcnally, J. S., Optimizing chemotherapy in an HIV model, Electron. J. Diff. Equations 32 (1998) 1–12.
22. Joshi, H. R., Optimal control of an HIV immunology model, Optim. Control Appl. Methods 23 (2002) 199–213.
23. Zhou, Y., Liang, Y., Wu, J., An optimal therapy for HIV multitherapy, J. Compu. Appl. Math. 263 (2014) 326–337.
24. J. K. Hale, S. M. Verduyn Lunel, Introduction to Functional Differential Equations, Springer-Verlag,1993.
25. Diekmann, O., Heesterbeek, J. A. P. & Metz, J. A. (1990). On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. Journal of Mathematical Biology, 28(4), 365–382.
26. Driessche, P Van Den, Watmough, J., Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission, Math. Biosci. 180 (2002) 29–48.
27. De Leenheer, P., Smith, H. L., Virus dynamics: a global analysis, SIAM J. Appl. Math. 63 (2003) 1313–1327.
28. Dixit, N. M., Perelson, A. S., Complex patterns of viral load decay under antiretroviral therapy: influence of pharmacokinetics and intracellular delay, J. Theor. Biol. 226 (2004) 95
29. Culshaw, R. V and Ruan, S., A delay-differential equation model of HIV infection of $CD4^+$T-cells. Math. Biosci. 165 (2000) 27–39.

30. Nelson, P. W., Murray, J. D. & Perelson, A. S. (2000). A model of HIV-1 pathogenesis that includes an intracellular delay. Mathematical Biosciences, 163(2), 201–215.
31. Wang, Y., Zhou, Y., Wu, J. & Heffernan, J. (2009). Oscillatory viral dynamics in a delayed HIV pathogenesis model. Mathematical Biosciences, 219(2), 104–112.
32. Ciupe, M. S., Bivort, B. L., Bortz, D. M. & Nelson, P. W. (2006). Estimating kinetic parameters from HIV primary infection data through the eyes of three different mathematical models. Mathematical Biosciences, 200(1), 1–27.
33. Pawelek, K. A., Liu, S., Pahlevani, F. & Rong, L. (2012). A model of HIV-1 infection with two time delays: mathematical analysis and comparison with patient data. Mathematical Biosciences, 235(1), 98–109.
34. Fleming, W., Rishel, R., Deterministic and stochastic optimal control, Springer verlag, New York, 1975.
35. Lukes, D., Differential equations: Classical to controlled , Math. Sci. and Engg.,Academic Press, New York, 1982.
36. Göllmann, L., Kern, D. & Maurer, H. (2009). Optimal control problems with delays in state and control variables subject to mixed control–state constraints. Optimal Control Applications and Methods, 30(4), 341–365.
37. Hattaf, K. and Yousfi, N., 2012. Optimal control of a delayed HIV infection model with immune response using an efficient numerical method. ISRN Biomathematics, 2012.
38. Ngina, P., Mbogo, R.W. & Luboobi, L.S., 2018. Modelling optimal control of in-host HIV dynamics using different control strategies. Computational and Mathematical Methods in Medicine, 2018.
39. Mondal, C., Adak, D. and Bairagi, N., 2019. Optimal control in a multi-pathways HIV-1 infection model: a comparison between mono-drug and multi-drug therapies. International Journal of Control, pp.1–18.
40. Li, X. and Wei, J., 2005. On the zeros of a fourth degree exponential polynomial with applications to a neural network model with delays. Chaos, Solitons & Fractals, 26(2), pp.519–526.

Mathematical Models of Acoustically Induced Vaporization of Encapsulated Nanodroplets

K. Jiang, M. Ghasemi, A. Yu, and S. Sivaloganathan

1 Introduction

Acoustic droplet vaporization (ADV) is a process in which droplets of liquid are vaporized through the application of ultrasound acoustic waves. Since its initial description in 1995 [8], the potential usefulness of this process within the medical field has been explored through proposed and demonstrated applications such as drug delivery, imaging, and embolic therapy.

The exploitation of this process for medical applications involves the introduction of droplets, typically several hundred nanometers in size, of an appropriately selected biocompatible organic compound into a patient's bloodstream. Ultrasound is then applied at the required site to induce the vaporization of these nanodroplets noninvasively and thus achieve the desired clinical purpose. Perfluorocarbons (PFCs) are the most explored class of compounds for this application due to their physicochemical properties, their lightest species having a boiling point below human body temperature, yet remaining metastable in the liquid state due to the additional Laplace pressure within the droplet arising from surface tension [9]. These droplets are typically encapsulated with either an albumin layer [10], a polymeric shell [12, 13], or fluorinated surfactants [14] to prevent their rapid dissolution within the bloodstream, thus improving their lifetime.

The distribution of administered drugs within the body poses a key challenge within the field of cancer chemotherapy. Severe side effects are caused by the

K. Jiang · M. Ghasemi · A. Yu
University of Waterloo, Waterloo, ON, Canada
e-mail: k27jiang@uwaterloo.ca; m23ghasemi@uwaterloo.ca; alfred.yu@uwaterloo.ca

S. Sivaloganathan (✉)
Center for Mathematical Medicine, The Fields Institute, Toronto, ON, Canada
e-mail: ssivaloganathan@uwaterloo.ca

R. P. Mondaini (ed.), *Trends in Biomathematics: Stability and Oscillations in Environmental, Social, and Biological Models*,
https://doi.org/10.1007/978-3-031-12515-7_7

cytotoxic effects of these drugs on healthy tissue. In addition, anomalous tumour vascularization and high interstitial pressure result in spatial drug gradients within a tumor, leading to the possible survival of cancer cells where drug concentration is low [15]. This challenge can be addressed through the development of targeted drug delivery techniques that increase the concentration of drugs within the tumour whilst minimizing the impact on healthy tissue. The small size of the aforementioned encapsulated nanodroplets and their stability within the bloodstream allow them to both circulate around the body and accumulate within tumour tissue through the enhanced permeability and retention (EPR) effect [16]. Targeted drug delivery can then potentially be achieved via the selective release of any encapsulated drugs by inducing ADV through the noninvasive application of a focused ultrasound field.

A limitation of conventional ultrasound imaging techniques is the difficulty encountered when imaging tissue microvasculature due to the small size of the microvessels and the poor resolution of signal between the blood and the surrounding tissue [17]. The ADV of encapsulated nanodroplets can turn them into micrometer-scale gas bubbles which oscillate under the influence of the applied acoustic field. This enhances the backscattered signal, boosting the contrast between the vasculature and the surrounding tissue, thus allowing us to overcome the above limitation [17]. The expansion in volume can also occlude blood vessels feeding a tumour, starving them of nutrients and thus damaging or eliminating them. The exploitation of this process for therapeutic benefits is known as embolotherapy [18].

The increase in stability caused by the Laplace pressure across the droplet surface, together with the encapsulating shell implies that a certain threshold of acoustic energy must be exceeded in order to achieve vaporization. The dependence of this threshold on various parameters such as acoustic frequency [11, 19, 20], pulse length [20–22], pulse repetition frequency [23], droplet size [19, 23] and concentration [22], as well as temperature [20, 22, 23] has been studied experimentally *in vitro*. A successful mathematical model of the ADV process and its dependence on some or all of the above factors can be used to optimize this process through the appropriate selection of an encapsulating shell with suitable physical characteristics, as well as the parameters of the applied acoustic field. It may also yield useful behavioural predictions in conditions that have yet to be tested experimentally.

Current physical models of ADV assume the presence of a spherical vapor bubble within the liquid droplet and an important aspect is the examination of the evolution of this bubble over time [4, 5]. Hence they build upon and incorporate early mathematical models of spherical bubble behavior. One such model is Lord Rayleigh's examination of the collapse of a spherical cavitation bubble within a liquid under constant pressure at a distance [28]. This was extended by Plesset to incorporate the possibility of time-varying external pressure [29]. The combined works form the basis of what has come to be known as the Rayleigh-Plesset framework and has since been extended further to incorporate the effects of surface tension [30], viscosity [31], heat diffusion [30, 32], and compressibility [33].

This paper presents two recent models of the ADV process. These models utilize the general Rayleigh-Plesset equation accounting for the effects of surface tension, viscosity, and heat diffusion (although still assuming incompressibility).

They also incorporate the effect of mass flux across the bubble surface through evaporation/condensation, as well as pressure contributions from the encapsulating shell. The first model, developed by Guédra et. al. [4] assumes that the shell exhibits linear elasticity. The second model, developed by Lacour et. al. [5] builds upon this by introducing nonlinear elasticity by treating the shell as a Mooney-Rivlin solid to account for large deformations.

2 Modelling

As shown in the above schematic (Fig. 1), the model represents the radius of the vapor bubble, the inner and outer radius of the shell as three concentric spheres with radii $R(t), a(t), b(t)$ respectively. It is assumed that the space within the inner vapor bubble, between the vapor bubble and the shell, and outside of the shell contains PFP vapor, liquid PFP, and water respectively. To model the evolution of these values over time, the following dynamics are taken into account:

1. Vapor behavior within the bubble
2. Mass flux and heat transfer across the bubble surface
3. Heat transfer and fluid flow within the liquid between the bubble surface and the shell
4. Heat transfer across the shell

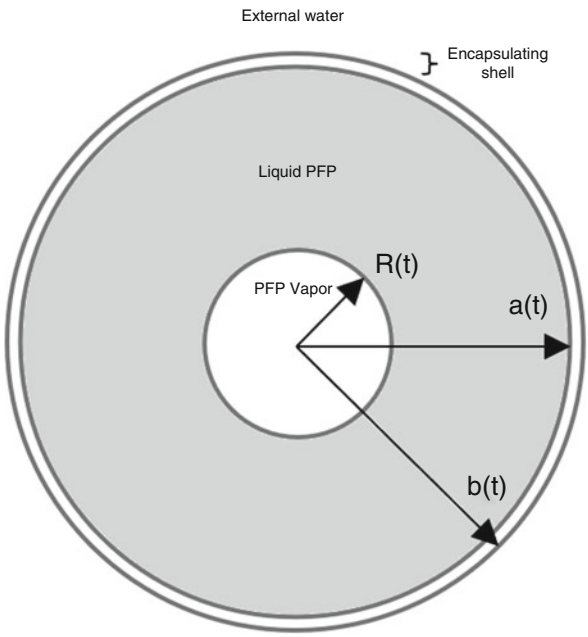


Fig. 1 Schematic of the model setup

5. Heat transfer and fluid flow within the external liquid

2.1 Evolution of Bubble Radius Over Time

The evolution of the bubble radius over time is described by the generalized Rayleigh-Plesset equation [5] which is obtained as follows. Its derivation begins with the Navier-Stokes equation for radially symmetric flows

$$-\frac{1}{\rho}\frac{\partial p}{\partial r} = \frac{\partial u}{\partial t} + u\frac{\partial u}{\partial r} - \nu\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial u}{\partial r}\right) - \frac{2u}{r^2}\right] \tag{1}$$

Where, ρ is the density, p is the pressure, $u := u(r,t)$ is the radial velocity at radial distance r from the bubble center at time t, and ν is the kinematic viscosity. By the conservation of mass, the radial velocity u must satisfy

$$u(r,t) = \left(\frac{R}{r}\right)^2 U(t) \tag{2}$$

Where $U(t) := \lim_{r\to R^+} u(r,t)$ is the velocity of the inner liquid right at the surface of the bubble. Plugging this into Eq. (1) results in

$$-\frac{1}{\rho}\frac{\partial p}{\partial r} = \frac{R^2\dot{U}}{r^2} + \frac{2R\dot{R}U}{r^2} - 2\frac{R^4U^2}{r^5} \tag{3}$$

Noting that the viscosity terms within the right-hand side of Eq. (1) cancel each other out as a result. (3) can then be integrated across the three domains of inner liquid $[R,a]$, shell $[a,b]$, and outer liquid $[b,\infty]$ to obtain the following [4]

$$R\dot{U} + \frac{4\dot{R}U - U^2}{2} - a\ddot{a} - \frac{3}{2}\dot{a}^2 = \frac{p_L(R) - p_L(a)}{\rho_L} \tag{4}$$

$$b\ddot{b} + \frac{3}{2}\dot{b}^2 = \frac{p_E(b) - p_\infty}{\rho_E} \tag{5}$$

$$b\ddot{b} - a\ddot{a} + \frac{3}{2}(\dot{b}^2 - \dot{a}^2) = \frac{p_S(a) - p_S(b)}{\rho_S} + \sigma_{rr}^S(a) - \sigma_{rr}^S(b) \tag{6}$$

Where subscripts L, S, E are used to denote properties corresponding to the inner liquid, shell, and outer liquid respectively. p_L and p_E represent the hydrostatic pressure within the inner and outer liquid respectively. The shell is assumed to be a viscoelastic material having both elastic stress and viscous stress terms p_S and σ_{rr}^S. In addition, taking into account surface tensions, viscosities, shell elasticity,

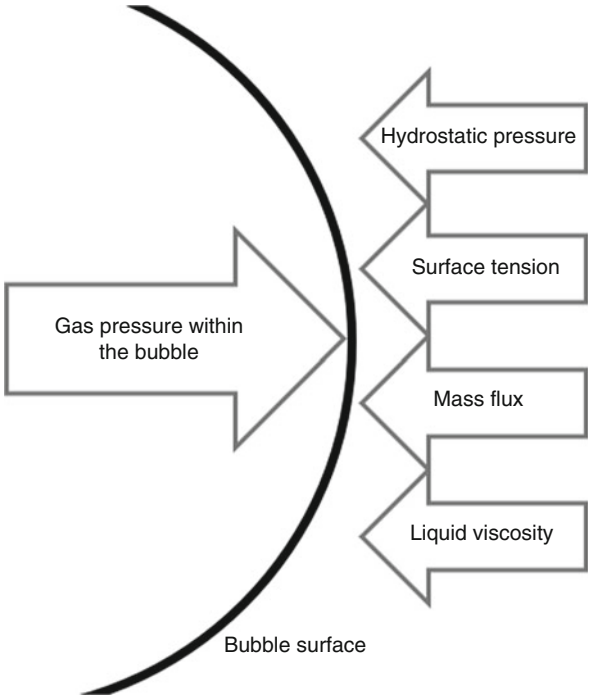


Fig. 2 Normal stress contributions from the outside and within the bubble

as well as the mass flux across the bubble surface, the continuity of normal stress across the bubble surface, shell inner surface, and shell outer surface (Fig. 2) gives the following [1, 6]

$$p_L(R) + JU = p_V(R) - 4\eta_L \frac{U}{R} - 2\frac{\sigma}{R} + JU_V \tag{7}$$

$$p_L(a) = p_S(a) - \sigma_{rr}^S(a) - 4\eta_L \frac{\dot{a}}{a} + \frac{2\sigma_1}{a} \tag{8}$$

$$p_E(b) = p_S(b) - \sigma_{rr}^S(b) - 4\eta_E \frac{\dot{b}}{b} - \frac{2\sigma_2}{b} \tag{9}$$

Where J is the mass flux across the bubble surface, $U_V := \lim_{r \to R^-} u(r,t)$ is the gas velocity at the inner surface of the bubble, η is the dynamic viscosity of the medium denoted by its subscript, and $\sigma, \sigma_1, \sigma_2$ are the surface tensions corresponding to the bubble surface, shell inner surface, and shell outer surface.

Combining Eqs. (4)–(9) results in the generalized Rayleigh-Plesset equation which describes the evolution of the bubble radius over time.

$$\begin{aligned}(R^2\dot{U} + 2R\dot{R}U)\left(\frac{\rho_E - \rho_S}{b} + \frac{\rho_S - \rho_L}{a} + \frac{\rho_L}{R}\right)\\ -\frac{R^4U^2}{2}\left(\frac{\rho_E - \rho_S}{b^4} + \frac{\rho_S - \rho_L}{a^4} + \frac{\rho_L}{R^4}\right) = p_R - p_\infty\end{aligned} \tag{10}$$

Where p_∞ is the pressure at infinity arising from an applied acoustic field as follows.

$$p_\infty(t) = p_0 - p_a \sin(2\pi f t) \tag{11}$$

Where p_0 is the ambient pressure, p_a is the amplitude of the acoustic wave, and f is the acoustic frequency. By the conservation of mass, the mass flux J is given by [3]:

$$J = \rho_L(U - \dot{R}) = \rho_V(U_V - \dot{R}) \tag{12}$$

Given this mass flux, together with Eq. (2), and noting that the shell is impermeable, the differential equations describing the evolution of R, a, b over time are given by.

$$\dot{R} = U - \frac{J}{\rho_L} \tag{13}$$

$$\dot{a} = \frac{R^2}{a^2}U \tag{14}$$

$$\dot{b} = \frac{R^2}{b^2}U \tag{15}$$

2.2 Pressure Contributions

In the process of obtaining Eq. (10), we see that the pressure at the bubble surface p_R can be broken down into contributions from viscosity, surface tension, mass flux, and shell response respectively.

$$p_R = p_V - \frac{2\bar{\sigma}}{R} - \frac{4\bar{\eta}U}{R} + \Phi + \mathcal{S} \tag{16}$$

Where

$$\bar{\eta} = \eta_L + (\eta_S - \eta_L)\left(\frac{R}{a}\right)^3 + (\eta_E - \eta_S)\left(\frac{R}{b}\right)^3 \tag{17}$$

$$\bar{\sigma} = \sigma + \frac{\sigma_1}{a} + \frac{\sigma_2}{b} \tag{18}$$

p_V is the pressure of the gas within the bubble. It can be determined using the Clausius-Clapeyron relation if the temperature is given (to be discussed in Sect. 2.3).

$$p_V(T) = p_{ref} \exp\left(\frac{L}{r_V}\left(\frac{1}{T_{ref}} - \frac{1}{T}\right)\right) \tag{19}$$

Where L is the latent heat of vaporization, r_V is the specific gas constant, and p_{ref} is the known pressure at a reference temperature T_{ref}

Φ is the contribution from the mass flux terms appearing in Eq. (7)

$$\Phi = J(U_V - U) \tag{20}$$

Using Eq. (12), this can be re-expressed as

$$\Phi = J^2\left(\frac{1}{\rho_V} - \frac{1}{\rho_L}\right) \tag{21}$$

The mass flux J itself is obtained from the conservation of energy across the bubble surface [4], assuming that the temperature distribution is uniform within the bubble (Sect. 2.3).

$$J = -\frac{K_L}{L}\frac{\partial T}{\partial r}\Big|_{r \to R^+} \tag{22}$$

Where K_L is the thermal conductivity of the inner liquid.

$\mathcal{S}$ is the contribution from the shell response which will be discussed in Sect. 2.4.

2.3 *Temperature Profile*

The determination of the pressure within the gas bubble within Eq. (19) as well as the mass flux across the bubble surface (22) requires the temperature within the bubble, as well as the temperature gradient immediately outside the bubble.

2.3.1 Bubble Surface Temperature

The modelling of the bubble surface temperature requires the determination of the gas velocity field within the bubble. This begins by assuming that it behaves as an ideal gas. The corresponding enthalpy equation is as follows [2]

$$\rho_V C_p \frac{dT}{dt} - \frac{dp}{dt} = \nabla \cdot (K_V \nabla T) \tag{23}$$

Where ρ_V is the gas density, C_p is its heat capacity under constant pressure, $T := T(r, t)$ is its temperature, and K_V is its thermal conductivity. The gas velocity field within the bubble $\vec{v}$ is introduced through the continuity equation.

$$\frac{\partial \rho_V}{\partial t} + \nabla \cdot (\rho_V \vec{v}) = 0 \tag{24}$$

The following ideal gas property is also used

$$C_p \rho_V T = \frac{\gamma p}{\gamma - 1} \tag{25}$$

Where $\gamma := \frac{C_p}{C_v}$ is the ratio of the specific heat at constant pressure to the specific heat at constant volume. Combining Eqs. (23)–(25), keeping in mind the assumption that the vapor density is constant in space, the following differential equation is obtained.

$$\frac{dp}{dt} + \gamma p \nabla \cdot \vec{v} = (\gamma - 1) \nabla \cdot (K_V \nabla T) \tag{26}$$

This is re-expressed as follows using the assumption of spherical symmetry.

$$\frac{\partial (r^2 v)}{\partial r} = \frac{r^2}{\gamma p} \left[\frac{\gamma - 1}{r^2} K_V \frac{\partial}{\partial r} \left(r^2 \frac{\partial T}{\partial r} \right) - \dot{p} \right] \tag{27}$$

Which is integrated to obtain the vapor velocity field within the bubble

$$v = \frac{1}{\gamma p} \left[(\gamma - 1) K_V \frac{\partial T}{\partial r} - \frac{r \dot{p}}{3} \right] \tag{28}$$

We know that $U_V = \lim_{r \to R^-} v(r, t)$. In addition, since the temperature field within the bubble is assumed to be spatially uniform, this reduces to

$$U_V = -\frac{R \dot{p}}{3 \gamma p} \tag{29}$$

Which can be plugged into the combination of Eqs. (12) and (22) to obtain

$$R \dot{p} + 3 \gamma p \left(\dot{R} - \frac{1}{\rho_V} \frac{K_L}{L} \frac{\partial T}{\partial r} \Big|_{r \to R^+} \right) = 0 \tag{30}$$

The Clausius-Clapeyron relation given in Eq. (19) provides the time rate of change of pressure, which is then used to obtain the final differential equation governing the

evolution of bubble surface temperature (hereby denoted T_S) over time.

$$\frac{L}{r_v T_S^2}\frac{dT_S}{dt} = \frac{3\gamma}{R}\left(\frac{K_L}{\rho_V L}\frac{\partial T}{\partial r}\Big|_{r\to R^+} - \dot{R}\right) \tag{31}$$

Where the gas density ρ_V can be obtained from the following ideal gas relation.

$$\rho_V = \frac{p_V}{r_V T} \tag{32}$$

2.3.2 Temperature Within the Inner and Outer Liquids

The temperature profiles within the inner and outer liquids are given by the energy equations [4, 7]

$$\frac{\partial T}{\partial t} + u(r,t)\frac{\partial T}{\partial r} = \frac{D_m}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial T}{\partial r}\right) + \frac{12\eta_m}{\rho_m c_m}\left(\frac{u(r,t)}{r}\right)^2 \tag{33}$$

Where $m \in \{L, E\}$ represents the medium for which the temperature profile is being evaluated and $D := \frac{K}{\rho c}$ is the thermal diffusivity. The heat flux is assumed to be continuous across the shell.

$$K_L\frac{\partial T}{\partial r}\Big|_{r\to a^-} = K_E\frac{\partial T}{\partial r}\Big|_{r\to b^+} \tag{34}$$

And the other boundary conditions are given by the temperature of the bubble surface and the temperature at infinity, which is held constant.

$$\lim_{r\to R^+} T(r,t) = T_S(t) \tag{35}$$

$$\lim_{r\to\infty} T(r,t) = T_\infty \tag{36}$$

2.4 Shell Contribution

In this review we cover two approaches to modelling the pressure term contributed by the shell's response—the Kelvin-Voigt model utilized in Guédra and Coulouvrat [4], and a later model by Lacour, Guédra and Coulouvrat [5] that models the shell as a hyperelastic material using the Mooney-Rivlin strain energy density function to incorporate nonlinear effects that arise as a result of large deformations.

The Kelvin-Voigt model represents the radial component of the viscous stress tensor that appears in Eqs. (6), (8), and (9) as follows

$$\sigma_{rr}^{S}(r) = -4\frac{a^2}{R^3}\left(G_S\frac{a}{3}\left[1-\frac{R_{10}^3}{R_1^3}\right]+\eta_S\frac{\dot{R}_1}{R_1}\right) \tag{37}$$

Where G_S is the shear modulus of the shell and η_S the shear viscosity. This gives the following expression as the shell response

$$\mathcal{S} = -\frac{4}{3}G_S\left(1-\frac{R_{10}^3}{R_1^3}\right)\left(1-\frac{R_1^3}{R_2^3}\right) \tag{38}$$

The Mooney-Rivlin model begins with Mooney's constitutive relation [34]

$$\Psi \approx \frac{G_S}{4}[(1+\beta)(I_1-3)+(1-\beta)(I_2-3)] \tag{39}$$

Where Ψ is the strain energy density function, G is the shear modulus, β is a fitting parameter, and I_1 and I_2 are the first two invariants of the left Cauchy-Green tensor $b = FF^T$. By expressing the two invariant terms in terms of the principal stretch $\lambda := \frac{r}{r_0}$, the strain energy density function can be reexpressed as

$$\Psi \approx \frac{G_S}{4}\sum_{i=-1}^{1}(1+i\beta)\left(\lambda^{-4i}+2\lambda^{2i}-3\right) \tag{40}$$

The shell response can be placed in the following form.

$$\mathcal{S} = \int_{\frac{a}{a_0}}^{\frac{b}{b_0}}\frac{\Psi'(\lambda)}{\lambda^3-1}d\lambda \tag{41}$$

Which can then be integrated to obtain

$$\mathcal{S} = G_S\left[\sum_{i=-1,i\neq 0}^{1}\sum_{k=0}^{1}\frac{1+i\beta}{-i-3k}\lambda^{-i-3k}\right]_{\frac{a}{a_0}}^{\frac{b}{b_0}} \tag{42}$$

3 Numerical Implementation

The set of equations in the previous section can be consolidated into a model of the four main state variables U, R, a, and b. The mass flux J is implicated in multiple sites within the model and its evaluation in turn requires the modelling of the temperature profile $T(r,t)$ of the system. To that end, the spatial terms within

the energy equation (33) are discretized using centralized finite differences. The following variable substitutions were also carried out to keep the boundary points fixed [3].

$$x = \frac{r - R}{a - R}; \forall r \in (R, a) \tag{43}$$

$$y = \frac{l}{r - b + l}; \forall r \in (b, \infty) \tag{44}$$

Where

$$l = B\sqrt{\frac{D_E}{\omega}} \tag{45}$$

B being a chosen parameter.

The internal and external liquids were discretized evenly into M and N spatial points respectively. I.e.

$$\forall i \in \{1, \ldots, M\}; x_i = i\Delta x; \text{ where } \Delta x = \frac{1}{M} \tag{46}$$

$$\forall j \in \{1, \ldots, N\}; y_j = j\Delta y; \text{ where } \Delta y = \frac{1}{N} \tag{47}$$

This results in a system of ODEs describing the evolution of temperature at $\{x_1, \ldots, x_M\}$, and $\{y_1, \ldots, y_N\}$. This system needs to be completed with boundary and initial conditions at $r = R \Leftrightarrow x_0 = 0$ and $r \to \infty \Leftrightarrow y_0 = 0$. The evolution of the bubble surface temperature is described by Eq. (31). By discretizing

$$\left.\frac{\partial T}{\partial r}\right|_{r \to R^+} \approx \frac{T(x_1, t) - T(x_0, t)}{\Delta x} \tag{48}$$

And rearranging Eq. (31), the evolution of bubble surface temperature can also be expressed as a function of the current temperature profile.

$$\frac{dT_S(t)}{dt} = \frac{\partial}{\partial t} T(x_0, t) \approx \frac{3T(x_0, t)^2 r_v \gamma}{LR} \left(\frac{K_L}{\rho_V L} \frac{T(x_1, t) - T(x_0, t)}{\Delta x} - \dot{R} \right) \tag{49}$$

On the other hand, the temperature at infinity is held constant, so

$$\frac{\partial}{\partial t} T(y_0, t) = 0 \tag{50}$$

The continuity of heat flux across the shell described in Eq. (34) was implemented by introducing an additional variable $T_{shell}(t)$ used to obtain the central finite

differences approximating the one-sided derivatives in Eq. (34).

$$\left.\frac{\partial T}{\partial r}\right|_{r\to a^-} = \frac{\partial}{\partial x}T(x_M,t) \approx \frac{T_{shell}(t) - T(x_{M-1},t)}{2\Delta x} \tag{51}$$

$$\left.\frac{\partial T}{\partial r}\right|_{r\to b^+} = \frac{\partial}{\partial y}T(y_N,t) \approx \frac{T(y_{N-1},t) - T_{shell}(t)}{2\Delta y} \tag{52}$$

$T_{shell}(t)$ can then be calculated at each time step by plugging the above into Eq. (34).

To avoid potential numerical difficulties [4], the initial temperature profile is interpolated between the initial bubble surface temperature and the temperature at infinity as follows [24]:

$$T(r,0) = T(\infty,0) - [T(\infty,0) - T_S(0)]\exp\left(-\frac{r - R(0)}{\delta - R(0)}\right) \tag{53}$$

Where δ was chosen to be within $[R(0), 2R(0)]$. The initial bubble surface temperature was obtained by applying the Clausius-Clapeyron relation to the initial pressure, which in turn was the combination of the ambient pressure and the Laplace pressure arising from the surface tension of the interfaces within the system.

$$T_S(0) = \left[\frac{1}{T_{ref}} - \frac{r_V}{L}\ln\left(\frac{p_0 + \frac{2\sigma}{R(0)} + \frac{2\sigma_1}{a(0)} + \frac{2\sigma_2}{b(0)}}{p_{ref}}\right)\right]^{-1} \tag{54}$$

The evolution of the entire system can then be expressed in the following form.

$$\frac{d}{dt}\begin{bmatrix} T(x_0,t) \\ T(x_1,t) \\ \vdots \\ T(x_M,t) \\ T(y_N,t) \\ \vdots \\ T(y_1,t) \\ T(y_0,t) \\ U \\ R \\ a \\ b \end{bmatrix} = f\left(\begin{bmatrix} T(x_0,t) \\ T(x_1,t) \\ \vdots \\ T(x_M,t) \\ T(y_N,t) \\ \vdots \\ T(y_1,t) \\ T(y_0,t) \\ U \\ R \\ a \\ b \end{bmatrix}, p_\infty(t)t\right) \tag{55}$$

This was then solved using the numerical method implemented by *ode15s* within Matlab® [35].

Table 1 Parameters used within the simulations

	Unit	PFP [25–27]	Shell [6]	H_2O
K	W/m K	0.056		0.6
ρ	kg/m^3	1590	1100	998
c	J/kg K	1046		4200
η	mN s/m^2	0.64	0.05	0.001
r_V	J/kg K	28.8		
L	kJ/kg	88		
γ		1.05		

4 Preliminary Results

4.1 Parameters

The physical properties of PFP, the shell material, and the external water are listed in the following Table 1.

The additional parameters involved include the vapor pressure of PFP at the reference temperature of 25 °C, $P_{ref} = 83.99$ kPa [27]. The ambient temperature and pressure were chosen as $T_\infty = 37$ °C, $P_0 = 1$ atm. B, the parameter used in the variable transformation of radial distance external to the encapsulated droplet shown in Eq. (45), was chosen to be 10. δ the parameter involved in setting up the initial temperature profile was chosen to be $1.1R(0)$. σ, the surface tension coefficient of th bubble was 0.0095 N/m. The surface tension coefficients corresponding to the shell inner and outer surfaces were assumed to be 0. β, the fitting parameter appearing in Eq. (39) was chosen to be 1.

The results (Figs. 3 and 4) exhibit two possibilities for the system: continued expansion through induced droplet vaporization, and bubble collapse. It appears that a larger amplitude of the driving acoustic wave imparts more energy into the system and bring it closer to vaporization. A stiffer shell with a higher shear modulus also provides more resistance acting against the bubble expansion, thus increasing the ADV threshold. The effect of nonlinearity introduced by the Mooney-Rivlin model also gives rise to a larger shell response and hence further increasing the energy required for ADV.

5 Conclusion

Acoustic droplet vaporization (ADV) is a mechanism that holds great promise for potential clinical applications ranging from contrast agents in ultrasound, to the administration of drug cargos at specific target sites. The development of appropriate mathematical models can lead to a deeper understanding of the mechanisms and factors involved in the ADV process. Although both models discussed in this paper are based on the Rayleigh-Plesset equation—the models used different constitutive

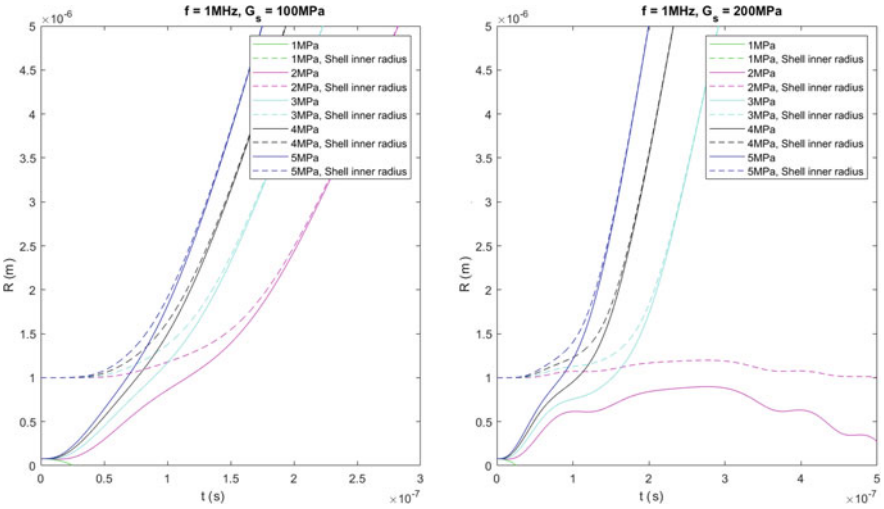


Fig. 3 Evolution of vapor bubble radius and shell inner radius as given by the linear elasticity model under varying amplitudes of the acoustic waves (1 to 5 MPa) and two different shell conditions: left) $G_S = 100$ MPa, and right) $G_S = 200$ MPa

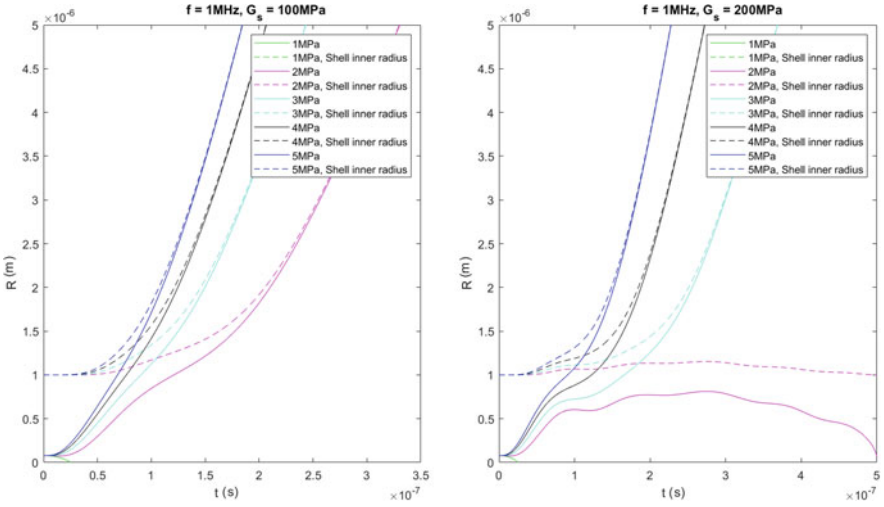


Fig. 4 Evolution of vapor bubble radius and shell inner radius as given by the nonlinear elasticity model under varying amplitudes of the acoustic waves (1 to 5 MPa) and two different shell conditions: left) $G_S = 100$ MPa, and right) $G_S = 200$ MPa

equations. The first model uses linear elasticity theory, whilst the second utilizes a Mooney-Rivlin model that is more suitable for large shell deformations. These models can be used to develop some understanding of the mechanisms that drive the ADV process, as well as to determine key factors in the vaporization process. The

main purpose of this paper is to present and compare these two models presented in the literature to study the ADV process.

Acknowledgments Work partially supported by an NSERC CREATE Grant and Discovery Grants (AY, SS) from the National Science and Engineering Research Council, Canada.

References

1. A. Prosperetti, *Meccanica* **14**, 34 (1979).
2. A. Prosperetti, *J. Fluid. Mech.* **222**, 587 (1991).
3. Y. Hao and A. Prosperetti, *Phys. Fluids* **11**, 2008 (1999).
4. M. Guédra and T. Valier-Brasier, F. Coulouvrat, *J. Acoust. Soc. Am.* **138**, 3656 (2015).
5. T. Lacour, M. Guédra, T. Valier-Brasier, F. Coulouvrat *J. Acoust. Soc. Am.* **143**, 23 (2018).
6. C. Church, *J. Acoust. Soc. Am.* **97**, 1510 (1995).
7. L. D. Landau and E. M. Lifshitz *Fluid Mechanics* Vol. 6, (1966).
8. F. Forsberg, et. al. *J. Ultrasound Med.* **14**, 949 (1995).
9. C.-Y. Lin and W. G. Pitt *BioMed. Res. Int.* **2013**, 404361 (2013).
10. O. D. Kripfgans, et. al. *Ultrasound Med. Biol.* **26**, 1177 (2000).
11. O. D. Kripfgans, M. L. Fabiilli, P. L. Carson, J. B. Fowlkes *J. Acoust. Soc. Am.* **116**, 272 (2004).
12. E. Pisani, et. al. *Adv. Funct. Mater.* **18**, 2963 (2008).
13. N. Rapoport *WIREs Nanomed. Nanobiotechnol.* **4**, 492 (2012).
14. K. Astafyeva, et. al. *J. Mater. Chem.* **3**, 2892 (2015).
15. N. Rapoport, Z. Gao, A. Kennedy *J. Natl. Cancer Inst.* **99**, 1095 (2007).
16. A. Iyer, G. Khaled, J. Fang, H. Maeda *Drug Discov. Today* **11**, 812 (2006).
17. D. Cosco, E. Fattal, M. Fresta, N. Tsapis *J. Fluorine Chem.* **171**, 18 (2015).
18. M. Zhang, et. al. *Ultrasound Med. Biol.* **36**, 1691 (2010).
19. A. L. Martin, et. al. *Ultrasound Med. Biol.* **38**, 1799 (2012).
20. R. Williams, et. al. *Ultrasound Med. Biol.* **39**, 475 (2013).
21. A. H. Lo, et. al. *IEEE Trans. Ultrason. Ferroelectr.* **54**, 933 (2007).
22. N. Reznik, R. Williams, P. N. Burns *Ultrasound Med. Biol.* **37**, 1271 (2011).
23. M. L. Fabiilli, et. al. *IEEE Trans. Ultrason. Ferroelectr.* **56**, 1006 (2009).
24. O. Shpak, L. Stricker, M. Versluis, D. Lohse *Phys. Med. Biol.* **58**, 2523 (2013).
25. P. S. Sheeran, P. A. Dayton *Curr. Pharm. Des.* **18**, 2152 (2012).
26. A. A. Doinikov, P. S. Sheeran, A. Bouakaz, P. A. Dayton *Med. Phys.* **41**, 102901 (2014).
27. FluoroMed® Perfluoropentane http://www.fluoromed.com/products/perfluoropentane/ (Last viewed Oct 7th, 2021)
28. L. Rayleigh *Lond. Edinb. Dubl. Philos. Mag.* **34**, 94 (1917).
29. M. S. Plesset *J. Appl. Mech.* **16**, 277 (1949).
30. M. S. Plesset, S. A. Zwick *J. Appl. Phys.* **25**, 493 (1954).
31. B. E. Noltingk, E. A. Neppiras *Proc. Phys. Soc. London, Sect. B* **63** 674 (1950).
32. Y. Yasui *J. Phys. Soc. Jpn.* **65** 2830 (1996).
33. J. B. Keller, I. I. Kolodner *J. Appl. Phys.* **27**, 1152 (1956).
34. M. Mooney *J. Appl. Phys.* **11**, 582 (1940).
35. Matlab® R2021a, The MathWorks® Inc., Natick, Massachusetts, USA.

Existence of Closed Trajectories in Lotka–Volterra Systems in $\mathbb{R}^+_n$

A. Bratus, V. Tikhomirov, and R. Isaev

1 Introduction

The existence of closed phase trajectories in the Lotka–Volterra system [1, 2], also known as predator-prey equations, made it possible to explain many phenomena of periodic changes in the number of species observed in nature. A. Kolmogorov and G. Gause proposed another version of this model [3, 4], which allowed recognising a variety of behaviours, including the occurrence of a limit cycle. These studies have been advanced in numerous publications, covered in [5, 6]. Of particular interest is the possibility of closed trajectories in ecological communities of the ⟪food chain⟫ type, which can ensure the sustainable development of these systems [6].

Consider the Lotka–Volterra system in $\mathbb{R}^n_+$:

$$\frac{du_i}{dt} = u_i(r_i - (Au)_i). \tag{1}$$

Here, $u_i(t)$ denotes the number of type i species at time $t > 0$ and r_i is the Malthusian coefficients of i-th species, $r = (r_1, r_2, \ldots, r_n) \in \mathbb{R}^n$. We imply

A. Bratus (✉)
Russian University of Transport, Moscow Center of fundamental and applied mathematics, Lomonosov Moscow State University, Moscow, Russia
e-mail: alexander.bratus@yandex.ru

V. Tikhomirov · R. Isaev
The faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University, Moscow, Russia
e-mail: zedum@cs.msu.ru

R. P. Mondaini (ed.), *Trends in Biomathematics: Stability and Oscillations in Environmental, Social, and Biological Models*,
https://doi.org/10.1007/978-3-031-12515-7_8

$(Au)_i = \sum_{j=1}^{n} a_{ij}u_j(t)$. At the initial time moment, we have $u_i(0) = u_i^0 \geq 0,\ i = \overline{1,n}$.

Matrix $A = \{a_{ij}\}_{n\times n}$ describes the nature of the interaction between species, where $\det(A) \neq 0$. The values a_{ij} describe the effect of exposure of j population on i population. If $i = j$, then the coefficient a_{ii} reflects the intraspecific competition of species. Let A^T mean a transposed matrix A. The matrix A can be represented as a sum of two matrices: $A = B + C$, where $B = B^T,\ C = -C^T$. Moreover, $B = \frac{A+A^T}{2},\ C = \frac{A-A^T}{2}$. The state-space of the system (1) is the positive octant $\mathbb{R}^n_+ = \{u \in \mathbb{R}^n : u_i \geq 0,\ \forall i\}$. Let us introduce the notation $\text{int}\mathbb{R}^n_+ = \{u \in \mathbb{R}^n : u_i > 0,\ \forall i\};\ \text{bd}\mathbb{R}^n_+ = \mathbb{R}^n_+ \setminus \text{int}\mathbb{R}^n_+$.

Boundary points $\mathbb{R}^n_+$ belong to the coordinate planes $u_i = 0$, which corresponds to the extinction (or absence) of the i species.

Definition 1 The system (1) is called nondegenerate if for any initial conditions $u_i^0 > 0,\ i = \overline{1,n}$ there is a number $\delta > 0$, that $\lim\limits_{t\to\infty}\big(\inf(u_i(t,u_0))\big) > \delta,\ i = \overline{1,n}$.

Definition 2 The system (1) is uniformly bounded if for any initial conditions $u_i^0 > 0\ i = \overline{1,n}$ there exists such a number K, independent of the initial conditions that $\lim\limits_{t\to\infty}\big(\sup(u_i(t,u_0))\big) \leq K,\ i = \overline{1,n}$.

The condition for the nondegeneracy of the system (1) in the absence of absorbing equilibrium positions, i.e., such equilibrium positions that are located on the boundary $\mathbb{R}^n_+(\text{bd}\mathbb{R}^n_+)$, in which trajectories can leave $\mathbb{R}^n_+$. This condition can be formulated in the following form [7]. There exists a vector $p = (p_1, p_2, \ldots, p_n) \in \text{int}\mathbb{R}^n_+$, such that for any fixed point of the system (1) $\overline{u} = (\overline{u}_1, \overline{u}_2, \ldots \overline{u}_n) \in \text{bd}\mathbb{R}^n_+$, the inequality holds:

$$\sum_{i=1}^{n} p_i(r_i - (Au)_i) > 0.$$

A necessary condition for the uniform boundedness of the system is the existence of a single fixed point $\overline{u} \in \text{int}\mathbb{R}^n_+$, such that for any trajectory $u(t) = (u_1(t), u_2(t), \ldots, u_n(t)) \in \mathbb{R}^n_+$ following equality is applied [7]:

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t u(t)dt = \overline{u},\quad A\overline{u} = r,\ r = (r_1, r_2, \ldots, r_n).$$

Theorem 1 *Let a nondegenerate uniformly bounded system* (1) *have a single equilibrium* $\overline{u} \in int\mathbb{R}^n_+$. *If the matrix B is definite or* $C = 0$, *then the system has no closed phase trajectories* $u(t) \in int\mathbb{R}^n_+$. *If* $B = 0$, *then all trajectories of the system* (1) *lie on the surface* $M \subseteq int\mathbb{R}^n_+$, *which is homeomorphic to the ball.*

Proof Consider the function:

$$V(u) = \sum_{i=1}^{n} \left(\overline{u}_i - u_i(t) + \overline{u}_i \ln \frac{u_i(t)}{\overline{u}_i} \right).$$

Here, $\overline{u} = (\overline{u}_1, \ldots, \overline{u}_n)$ stands for the equilibrium point of the system (1):

$$A\overline{u} = r.$$

Let us calculate the derivative of the function $V(u)$ along the trajectories of the system (1):

$$\dot{V}(u) = (A(\overline{u} - u), \overline{u} - u) = (B(\overline{u} - u), \overline{u} - u).$$

If there is a closed trajectory $\gamma_t = \left\{ u \in \text{int}\mathbb{R}^n_+,\ u_i = u_i(t),\ i = 1, 2, \ldots, n \right\}$, then

$$\int_0^T \dot{V}(u)dt = 0.$$

In the latter expression, $T > 0$ is a period of closed trajectory γ_t. This condition is certainly not satisfied if the matrix B is definite. If $B = 0$, then the function $V(u)$ is an integral of the system.

Since

$$V(\overline{u}) = 0,\ \frac{\partial V}{\partial u_i}(\overline{u}) = 0,\ a_i = \frac{\partial^2 V}{\partial u_i^2}(\overline{u}) > 0,\ \frac{\partial^2 V}{\partial u_i \partial u_j}(\overline{u}) = 0,\ i \neq j,$$

then $V(u)$ is a strictly convex function that reaches its minimum value at a single point $\overline{u} \in \text{int}\mathbb{R}^n_+$. All trajectories of the system (1) belong to the set $V(u) = const$. In the neighbourhood of the equilibrium, we get the approximation:

$$V(u) = \sum_{i=1}^{n} a_i (u_i - \overline{u}_i)^2 + o(||u - \overline{u}||^2).$$

According to the Morse–Palais lemma [8], there is a reversible differentiable change of variable $u = \psi(w)$, such that:

$$V(w) = \sum_{i=1}^{n} \lambda_i w_i^2 = const > 0,\ \lambda_i > 0, i = 1, 2, \ldots, n.$$

Let us prove that the system (1) has no closed trajectories for $C = 0$. Consider the function:

$$F(u) = \sum_{i=1}^{n} r_i u_i - (Bu, u).$$

The derivative of the function $F(u)$ along the trajectories of the system has the form:

$$\dot{F}(u) = \sum_{i=1}^{n} u_i \left(r_i - (Bu)_i\right)^2 > 0, \ u \in \text{int}\mathbb{R}^n_+$$

Therefore, there is no such value $T > 0$, at which the integral value would turn to zero:

$$\int_0^T \dot{F}(u)dt = 0.$$

The following example shows that if the matrix $B \neq 0$ is not definite, closed trajectories in the Lotka–Volterra system may exist [10, 11]. Consider now the case of a food chain with three types, without taking into account intraspecific competition:

$$\begin{cases} \dot{u}_1 = u_1(r_1 - a_{12}u_2), \\ \dot{u}_2 = u_2(-r_2 + a_{21}u_1 - a_{23}u_3), \\ \dot{u}_3 = u_3(-r_3 + a_{32}u_3). \end{cases} \tag{2}$$

Here, $a_{ij} > 0, \ i \neq j, \ r_i > 0, \ i = 1, 2, 3, \ u = (u_1, u_2, u_3) \in \mathbb{R}^3_+$. The system has a single equilibrium $\overline{u} \in \mathbb{R}^3_+$. Consider the function:

$$V_3(u) = \sum_{i=1}^{n} C_i \left(u_i - \overline{u}_i \ln u_i\right).$$

If we choose the set of constants $C_i, \ i = 1, 2, 3$ in such way that $C_i a_{i,i+1} = C_{i+1} a_{i+1,i}, \ i = 1, 2$, then the function $V_3(u)$ is the integral of the system(2), i.e., $\dot{V}_3(u) = 0$ [9].

All phase trajectories of the system (2) belong to the set:

$$\Xi = \left\{u \in \text{int}\mathbb{R}^n_+ : V_3(u) = const\right\}.$$

Consider points $(u_1, u_2, u_3), u_3 \neq \overline{u}_3$, belonging to Ξ. From the implicit function theorem it follows that in a neighbourhood of such points there exist functions $u_3 = u_3(u_1, u_2)$. Moreover, we can express its derivative as:

$$\frac{\partial u_3}{\partial u_i} = -\frac{V_{u_i}}{V_{u_3}} = -\frac{C_i \cdot \left((u_i - \overline{u}_i)\right) \cdot u_3}{C_3 \cdot \left((u_3 - \overline{u}_3)\right) \cdot u_i}.$$

Moreover, the equality holds true:

$$C_2\Big(1-\frac{\overline{u}_2}{u_2}\Big)\frac{\partial u_3}{\partial u_1}-C_1\Big(1-\frac{\overline{u}_1}{u_1}\Big)\frac{\partial u_3}{\partial u_2}=0. \tag{3}$$

The integral of this first order partial differential equation has the form:

$$\Phi_1(u_1,u_2)=C_1\Big(u_1-\overline{u}_1\ln u_1\Big)+C_2\Big(u_2-\overline{u}_2\ln u_2\Big).$$

The function $u_3 = \Phi(u_1,u_2)$ is a solution to the Eq. (3), which represents a strictly convex surface in a neighborhood of the point $(\overline{u}_1,\overline{u}_2)$ in space (u_1,u_2,u_3). Wherein, it has the derivatives:

$$\frac{\partial\Phi(\overline{u}_1,\overline{u}_2)}{\partial u_i}=0,\quad \frac{\partial^2\Phi(\overline{u}_1,\overline{u}_2)}{\partial u_i^2}>0,\quad \frac{\partial^2\Phi(\overline{u}_1,\overline{u}_2)}{\partial u_1\partial u_2}=0,\quad i=1,2.$$

Using the same arguments as in the proof of Theorem 1, we conclude that the function $\Phi(u_1,u_2)$, due to a reversible replacement, can be represented in new coordinates as follows:

$$\Phi\left(u_1(w_1,w_2),u_2(w_1,w_2)\right)=\lambda_1w_1^2+\lambda_2w_2^2,\quad \lambda_1,\lambda_2>0.$$

Therefore, the closed level lines of this surface represent phase trajectories. The result obtained can be generalized to for general equations of food chain $n = 2m+1$. In this case,

$$V_n(u)=\sum_{i=1}^{2m+1}C_i\cdot(u_i-\overline{u}_i\ln u_i)\,,$$

where nonnegative constants $C_i, i = 1,2,\ldots,n$ satisfy the conditions: $C_ia_{i,i+1} = C_{i+1}a_{i+1,i},\ i = 1,2,\ldots,2m,\ n = 2m-1$. These functions are integrals of the system and belong to the set

$$\Xi_n=\left\{u\in \mathrm{int}\mathbb{R}^n_+ : V_n(u)=const\right\}.$$

Using the implicit function theorem in a neighborhood of points:

$$\begin{aligned}
M_1&=(u_1,u_2,\overline{u}_3,\overline{u}_4,\ldots,\overline{u}_{2m-2},u_{2m-1}),\quad u_{2m-1}\neq\overline{u}_{2m-1},\\
M_2&=(\overline{u}_1,\overline{u}_2,u_3,u_4,\ldots,\overline{u}_{2m-2},u_{2m-1}),\quad u_{2m-1}\neq\overline{u}_{2m-1},\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
M_m&=(\overline{u}_1,\overline{u}_2,\overline{u}_3,\overline{u}_4,\ldots,u_{2m-3},u_{2m-2},u_{2m-1}),\quad u_{2m-1}\neq\overline{u}_{2m-1},
\end{aligned}$$

we can sequentially calculate the partial derivatives $\frac{\partial u_{2m-1}}{\partial u_k}$ for $k = 1, 2; k = 3, 4; \ldots;$
$k = 2m-2, k = 2m-1$ at $M_1, M_2, \ldots, M_m$ respectively.
Then, the functions $u^k_{2m-1} = \Phi_k(u_{2k-1}, u_{2k}),\ k = 1, 2, \ldots, m-1$ satisfy a first order partial differential equation of the form:

$$\frac{\partial u^k_{2m-1}}{\partial u_k} C_{k+1}\Big(1 - \frac{\overline{u}_{k+1}}{u_k}\Big) - \frac{\partial u^k_{2m-1}}{\partial u_{k+1}} C_k\Big(1 - \frac{\overline{u}_k}{u_k}\Big) = 0.$$

The integrals of these equations are given by the equalities:

$$V_k(u_k, u_{k+1}) = C_k\Big(u_k - \overline{u}_k \ln u_k\Big) + C_{k+1}\Big(u_{k+1} - \overline{u}_{k+1} \ln u_{k+1}\Big)$$

Here, $k = 1, 2; k = 3, 4; \ldots k = 2m-2, 2m-1$ respectively. As a result, repeating the previous reasoning, we obtain for the functions $u^k_{2m-1} = \Phi_k(u_{2k-1}, u_{2k}), k = 1, 2, \ldots, m-1$ the statement that these functions in a neighborhood of points $M_1, M_2, \ldots, M_m$, represent surfaces in spaces $(u_1, u_2, u_{2m-1}), (u_3, u_4, u_{2m-1}), \ldots, (u_{2m-3}, u_{2m-2}, u_{2m-1})$ respectively. The cross sections of such surfaces by the planes $u^k_{2m-1} = const,\ (u_{2m-1} \neq \overline{u}_{2m-1})\ k = 1, 2, \ldots, m$, are closed phase trajectories of the food chain systems. The result can be formalized as the theorem below.

Theorem 2 *Consider the Lotka-Volterra system of odd dimension $n = 2m+1,\ m \geq 2$, representig a food chain without intraspecific competition. If it has a single equilibrium $\overline{u} \in \mathbb{R}^{2m+1}_+$, then there is a closed phase trajectory that belongs to $(m-1)$-dimensional torus $K_{m-1} \in \mathbb{R}^{2m+1}_+$.*

Proof Let $\gamma_1, \gamma_2, \ldots, \gamma_{m-1}$ be closed phase trajectories of systems in spaces $(u_1, u_2), (u_3, u_4), \ldots, (u_{2m-3}, u_{2m-2})$ respectively. The direct multiplication of these phase trajectories represents the phase trajectory belonging to $(m-1)$-dimensional torus $K_{m-1} \in \mathbb{R}^{2m+1}_+$. If the sections of surfaces $u^k_{2m-1} = const,\ k = 1, 2, \ldots, m-1$ are chosen in such a way that the time of movement along each closed trajectory has a rational value, then there is a closed phase trajectory belonging to the torus $K_{m-1} \in \mathbb{R}^{2m+1}_+$.

Acknowledgments The work was supported by a grant from the Ministry of Science and Education.

References

1. Volterra V. *Mathematical theory of the struggle for existence. Moscow. The science. 1976.(Theorie mathematique de la lute pour la vie Vito Volterra Gauther – Villars, 1931)* (in Russian)

2. Lotka A.J. *Elements of mathematical biology. N.-Y. Dower, 1956*
3. Kolmogorov A. *Sulla teoria di Volterra della lotta per l" esistize.G. Inst. Ital. Attuari, 1936. 7,*#1*, 74–80* 1962
4. Gause G.F., Smorodova N.P., Witt A.A *Further studies of interaction between predator and prey. J. Anim. Ecol. 1936,5,* #1*, 1–18*
5. Svirezhev Yu., Logofet D. *Stability of biological communities, M. Science. 1978* (in Russian)
6. Bazykin A. *Nonlinear dynamics of interacting populations. 2003. Moscow. Izhevsk* (in Russian)
7. Hofbauer J., Sigmund K. *The Theory of Evolution and Dynamical System. 2003. Cambridge University Press*
8. Morse M. *The Calluses of Variation in the Large. N.-Y., 1934.*
9. Arnold V. *Geometric Methods in the Theory of Ordinary Differential Equations, 2002. MTsNMO.*
10. Bratus A., Novozhilov A., Platonov A.*Dynamical systems and models of biology // Moscow. Fizmatlit, 2010.* (in Russian)
11. Bratus A., Tikhomirov V. *// On closed trajectories of the Lotka–Volterra system of equations in multidimensional space // Lomonosov readings. 2020* . (in Russian)

Asymptotic Stability Analysis for Chaotic Predator-Prey Model via Active Controlled Hybrid Compound Difference Synchronization Strategy

Ayub Khan and Harindri Chaudhary

1 Introduction

The classical predator-prey model (PPM), mathematically and typically realized as Lotka-Volterra (LV) equations, goes back to 1920s when Alfred J. Lotka [1] and Vito Volterra [2] have proposed independently these LV equations on population dynamics to depict an interaction among various constituents elements in a system via competition or cooperation. Elucidations of such systems may include biological species, countries, businesses, and others. In the beginning, LV model has been described as a biological term, yet it is utilised to many diversified areas of research [3–9]. Interest in and applications are still rising [10–17]. While several additions as well as modifications to the classical PPM have been considered in the literature, an interesting twist, namely, that grown-up prey eat baby and juvenile predators, has been recently considered in the paper [18]. The author firstly sets up the relevant equations and then computes the ecological limit cycles. Most importantly, Arneodo et al.[19] in the year 1980 have explained that generalized Lotka-Volterra system (GLVS) may depict chaos phenomena for a specifically chosen system parameters. Furthermore, Samardzija and Greller [20] have illustrated in 1988 that GLVS may behave chaotically from the stable state through rise of a fractal torus. The most striking feature of chaotic system is its depiction of high sensitivity to initial conditions.

After the phenomenal work by Poincare [21], Lorenz [22] in 1963 described chaos phenomenon comprehensively in a deterministic system while examining a weather prediction model. Early concept of chaos synchronization among chaotic systems was emerged from pioneered studies conducted by Pecora and Carroll

A. Khan · H. Chaudhary (✉)
Department of Mathematics, Jamia Millia Islamia, New Delhi, India
e-mail: akhan12@jmi.ac.in

R. P. Mondaini (ed.), *Trends in Biomathematics: Stability and Oscillations in Environmental, Social, and Biological Models*,
https://doi.org/10.1007/978-3-031-12515-7_9

[23] in the 1990s using master-slave/leader-follower/drive-response configuration. Presently, we know of many types of chaos synchronization and chaos control techniques [13, 24–33] etc. to perform asymptotic stability analysis on chaotic systems. Researchers [34] discussed optimal synchronization problem among Lotka-Volterra systems via optimal control strategy. Further, researchers [27, 35] investigated adaptive control scheme to synchronize chaotic GLVSs. Also, researchers [13] proposed and studied combination difference anti-synchronization among chaotic GLVSs using adaptive control methodology. Specifically, Er-Wei Bai and Karl E Lonngren [36] in 1997 first reported active control strategy (ACS) of controlling chaos found in nonlinear dynamical systems. In addition, compound synchronization was initiated by Sun et al. [37] in 2013. Some further researches [30, 38] have been observed in this direction.

Considering the above discussed facts, our primal focus in this article is to investigate hybrid compound difference synchronization (HCDS) technique in four identical chaotic GLVSs via ACS. The remainder of the article is framed as follows: Sect. 2 presents some mathematical preliminaries to be used in the coming up sections. Section 3 designs the asymptotic stability theory for HCDS among chaotic systems using ACS. Section 4 deals with few structural properties of chaotic GLVS. Further, the active nonlinear controllers are properly designed to achieve HCDS strategy. Additionally, in view of Lyapunov stability theory, we have discussed the biological control laws in detail to achieve the asymptotical stability globally of error dynamics of the given system. Section 5 consists of numerical simulations, which are performed using MATLAB software, to illustrate the efficacy and superiority of discussed technique. Further, a comparative study has been done which exhibits the significance of considered strategy over earlier published literature. In the end, we have also stated some conclusions and the future scope of the presented work in Sect. 6.

2 Mathematical Preliminaries

We, in this section, describe a methodology to investigate hybrid compound difference synchronization (HCDS) technique based on master-slave configuration in four chaotic systems to be utilized in the coming up sections. Suppose that the scaling master system is given by

$$\dot{u}_{m1} = g_1(u_{m1}), \tag{1}$$

and the base second master systems is defined as

$$\dot{u}_{m2} = g_2(u_{m2}), \tag{2}$$

$$\dot{u}_{m3} = g_3(u_{m3}). \tag{3}$$

Corresponding to the aforesaid master systems, assume that the slave system is represented by

$$\dot{u}_{s4} = g_4(u_{s4}) + V(u_{m1}, u_{m2}, u_{m3}, u_{s4}), \tag{4}$$

where $u_{m1} = (u_{m11}, u_{m12}, \ldots, u_{m1n})^T \in R^n$, $u_{m2} = (u_{m21}, u_{m22}, \ldots, u_{m2n})^T \in R^n$, $u_{m3} = (u_{m31}, u_{m32}, \ldots, u_{m3n})^T \in R^n$, $u_{s4} = (u_{s41}, u_{s42}, \ldots, u_{s4n})^T \in R^n$ are the state vectors of the respective chaotic/ hyperchaotic systems (1), (2), (3) and (4), $g_1, g_2, g_3, g_4 : R^n \to R^n$ are continuous nonlinear vector functions, $V = (V_1, V_2, \ldots.., V_n)^T : R^n \times R^n \times R^n \times R^n \to R^n$ are the properly designed controllers.

Hybrid compound difference synchronization (HCDS) technique is defined as

$$E = Su_{s4} - Pu_{m1}(Ru_{m3} - Qu_{m2}),$$

where $P = diag(p_1, p_2, \ldots.., p_n)$, $Q = diag(q_1, q_2, \ldots.., q_n)$, $R = diag(r_1, r_2, \ldots.., r_n)$, $S = diag(s_1, s_2, \ldots.., s_n)$ and $S \neq 0$.

Definition 2.1 The chaotic master systems (1)–(3) are said to achieve HCDS with the slave system (4) if

$$lim_{t\to\infty}\|e(t)\| = lim_{t\to\infty}\|Su_{s4}(t) - Pu_{m1}(t)(Ru_{m3}(t) - Qu_{m2}(t))\| = 0,$$

where $\|.\|$ represents Euclidean norm.

Remark 2.2 Describing matrices P, Q, R and S as scaling matrices. In addition, P, Q, R, S could be augmented as matrices having functions in state vectors u_{m1}, u_{m2}, u_{m3} and u_{s4} as their entries.

Remark 2.3 If $S = I$ and $P = Q = R = \xi I$, for $\xi = 1$ it turns into complete compound difference synchronization and for $\xi = -1$ it converts to anti-compound difference synchronization. Coexistence of anti and complete synchronization composes HCDS technique. Thus, the HCDS error attains the following form :

$$e = u_{s1} - \xi u_{m1}(u_{m3} - u_{m2}), \tag{5}$$

where $\xi = diag(\xi_1, \xi_2, \ldots, \xi_n)$ with $\xi_i = (-1)^{i+1}$ for $i = 1, 2, 3, \ldots.n$.

3 Asymptotic Stability Analysis via ACS

In this section, we present the synchronization phenomena to achieve HCDS approach among three chaotic master systems (1)–(3) and one chaotic slave system

(4). In this regard, we next properly design the controllers which are described as:

$$V_i = \zeta_i - (g_4)_i - M_i e_i, \tag{6}$$

where
$\zeta_i = \xi_i (g_1)_i (u_{m3i} - u_{m2i}) + \xi_i u_{m1i}((g_3)_i - (g_2)_i)$, for $i = 1, 2, \ldots\ldots, n$.

Theorem 3.1 *The considering systems (1)–(4) would attain the proposed HCDS technique globally and asymptotically, if the nonlinear active controllers are designed as given in Eq.(6).*

Proof Let the error be defined as

$$e_i = u_{s4i} + \xi_i u_{m1i}(u_{m3i} - u_{m2i}), \; for \;\; i = 1, 2, \ldots.., n.$$

The error dynamical system transforms to

$$\begin{aligned} \dot{e_i} &= \dot{u}_{s4i} + \xi_i \dot{u}_{m1i}(u_{m3i} - u_{m2i}) + \xi_i u_{m1i}(\dot{u}_{m3i} - \dot{u}_{m2i}) \\ &= ((g_4)_i + V_i) + \xi_i (g_1)_i (u_{m3i} - u_{m2i}) + \xi_i u_{m1i}((g_3)_i - (g_2)_i) \\ &= ((g_4)_i + V_i) + \zeta_i, \end{aligned}$$

where
$\zeta_i = \xi_i (g_1)_i (u_{m3i} - u_{m2i}) + \xi_i u_{m1i}((g_3)_i - (g_2)_i)$, for $i = 1, 2, \ldots\ldots, n$.
This implies that

$$\begin{aligned} \dot{e_i} &= ((g_4)_i - \zeta_i - (g_4)_i - M_i e_i) + \zeta_i \\ &= -K_i e_i \end{aligned} \tag{7}$$

The typical Lyapunov function $V(e(t))$ is formulated as:

$$\begin{aligned} V(e(t)) &= \frac{1}{2} e^T e \\ &= \frac{1}{2} \Sigma e_i^2 \end{aligned} \tag{8}$$

Time-derivative of $V(e(t))$, as defined in Eq. (8), is given by

$$\dot{V}(e(t)) = \&\Sigma e_i \dot{e}_i$$

Using Eq (8), we obtain

$$\begin{aligned} \dot{V}(e(t)) &= \Sigma e_i (-M_i e_i). \\ &= -\Sigma K_i e_i^2). \end{aligned} \tag{9}$$

We now select $(M_1, M_2, \ldots\ldots, M_n)$ so that $\dot{V}(e(t))$ presented by Eq. (9) turns into a negative definite function. Therefore, by applying Lyapunov stability theory [39], we have

$$lim_{t\to\infty} e_i(t) = 0 \quad for \quad (i = 1, 2, 3).$$

Thus, the systems (1)–(4) have attained desired HCDS technique using ACS. □

4 A Numerical Example

We now consider GLVS as the master system:

$$\begin{cases} \dot{u}_{m11} = & u_{m11} - u_{m11}u_{m12} + \alpha_3 u_{m11}^2 - \alpha_1 u_{m11}^2 u_{m13}, \\ \dot{u}_{m12} = & -u_{m12} + u_{m11}u_{m12}, \\ \dot{u}_{m13} = & -\alpha_2 u_{m13} + \alpha_1 u_{m11}^2 u_{m13}, \end{cases} \tag{10}$$

where $(u_{m11}, u_{m12}, u_{m13})^T \in R^3$ is the state vector of system. Also, u_{m11} represents the prey population and u_{m12}, u_{m13} denote the predator populations. For parameters $\alpha_1 = 2.9851$,$\alpha_2 = 3$,$\alpha_3 = 2$ and initial values $(27.5, 23.1, 11.4)$, scaling master GLVS exhibits chaotic behaviour as displayed in Fig. 1a.
The base master systems are the similar chaotic GLVSs prescribed respectively as:

$$\begin{cases} \dot{u}_{m21} = & u_{m21} - u_{m21}u_{m22} + \alpha_3 u_{m21}^2 - \alpha_1 u_{m21}^2 u_{m23}, \\ \dot{u}_{m22} = & -u_{m22} + u_{m21}u_{m22}, \\ \dot{u}_{m23} = & -\alpha_2 u_{m23} + \alpha_1 u_{m21}^2 u_{m23}, \end{cases} \tag{11}$$

where $(u_{m21}, u_{m22}, u_{m23})^T \in R^3$ is the state vector of system.

$$\begin{cases} \dot{u}_{m31} = & u_{m31} - u_{m31}u_{m32} + \alpha_3 u_{m31}^2 - \alpha_1 u_{m31}^2 u_{m33}, \\ \dot{u}_{m32} = & -u_{m32} + u_{m31}u_{m32}, \\ \dot{u}_{m33} = & -\alpha_2 u_{m33} + \alpha_1 u_{m31}^2 w_{m33}, \end{cases} \tag{12}$$

where $(u_{m31}, u_{m32}, u_{m33})^T \in R^3$ is state vector of (11).
The slave system, defined by the similar GLVS, is prescribed as:

$$\begin{cases} \dot{u}_{s41} = & u_{s41} - u_{s41}u_{s42} + \alpha_3 u_{s41}^2 - \alpha_1 u_{s41}^2 u_{s43} + V_1, \\ \dot{u}_{s42} = & -u_{s42} + u_{s41}u_{s42} + V_2, \\ \dot{u}_{s43} = & -\alpha_2 u_{s43} + \alpha_1 u_{s41}^2 u_{s43} + V_3, \end{cases} \tag{13}$$

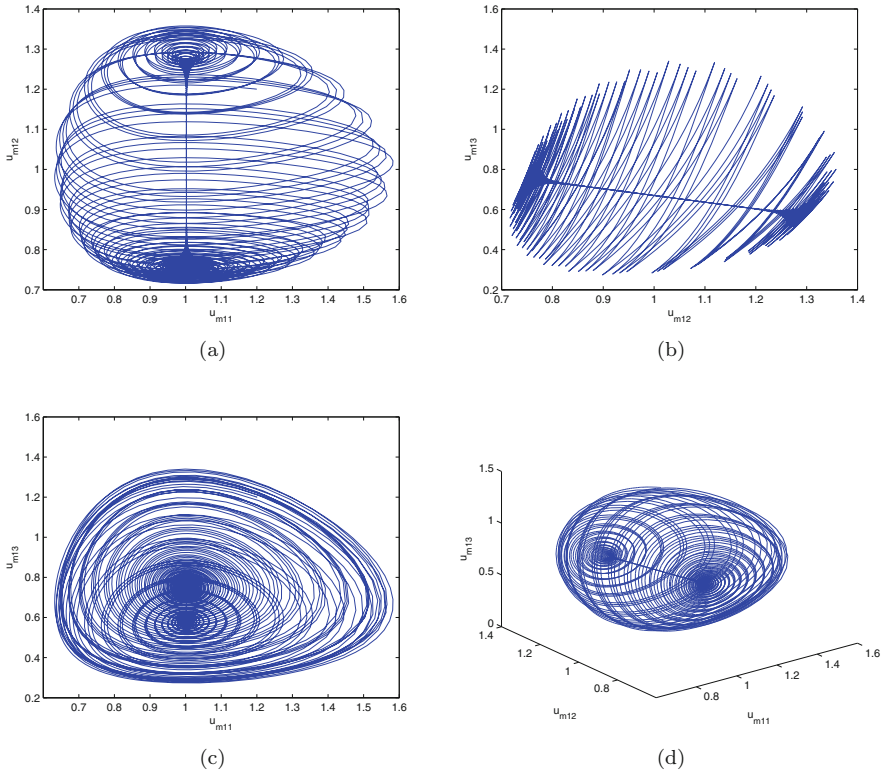


Fig. 1 Phase diagrams for chaotic GLVS in (**a**) $u_{m11} - u_{m12}$ plane, (**b**) $u_{m12} - u_{m13}$ plane, (**c**) $u_{m11} - u_{m13}$ plane, (**d**) $u_{m11} - u_{m12} - u_{m13}$ space

where $(u_{s41}, u_{s42}, u_{s43})^T \in R^3$ is the state vector of the system. Additionally, the researchers [20] carried out the detailed analytic study for (10)–(13). Furthermore, V_1, V_2 and V_3 are active controllers.

Next, the HCDS strategy is proposed for synchronizing the states of chaotic GLVSs. To achieve this, Lyapunov stability theory (LST) based ACS is employed and required asymptotic stability criterion is derived.

Here, we have $P = Q = R = diag(\xi_1, \xi_2, \xi_3)$, $S = diag(1, -1, 1)$.

Defining error functions (e_1, e_2, e_3) as

$$\begin{cases} e_1 = & u_{s41} - \xi_1 u_{m11}(u_{m31} - u_{m21}), \\ e_2 = & u_{s42} + \xi_2 u_{m12}(u_{m32} - u_{m22}), \\ e_3 = & u_{s43} - \xi_3 u_{m13}(u_{m33} - u_{m23}). \end{cases} \tag{14}$$

The major goal of this work is to design the active controllers V_i, $(i = 1, 2, 3)$ which ensure that error functions as mentioned in Eq.(13) satisfy

$$lim_{t\to\infty} e_i(t) = 0 \quad for \quad (i = 1, 2, 3).$$

The resulting error dynamics turns out to be

$$\begin{cases} \dot{e_1} = & \dot{u}_{s41} - \xi_1 \dot{u}_{m11}(u_{m31} - u_{m21}) - \xi_1 u_{m11}(\dot{u}_{m31} - \dot{u}_{m21}), \\ \dot{e_2} = & \dot{u}_{s42} + \xi_2 \dot{u}_{m12}(u_{m32} - u_{m22}) + \xi_2 u_{m12}(\dot{u}_{m32} - \dot{u}_{m22}), \\ \dot{e_3} = & \dot{u}_{s43} - \xi_3 \dot{u}_{m13}(u_{m33} - u_{m23}) - \xi_3 u_{m13}(\dot{u}_{m33} - \dot{u}_{m23}). \end{cases} \tag{15}$$

Using (10), (11), (12) and (13) in (15), the error dynamics becomes

$$\begin{cases} \dot{e_1} = & (u_{s41} - u_{s41}u_{s42} + \alpha_3 u_{s41}^2 - \alpha_1 u_{s41}^2 u_{s43} + V_1) \\ & +\xi_1(u_{m11} - u_{m11}u_{m12} + \alpha_3 u_{m11}^2 - \alpha_1 u_{m11}^2 u_{m13})(\xi_1 u_{m31} - \xi_1 u_{m21}) \\ & +\xi_1 u_{m11}(\xi_1(u_{m31} - u_{m31}u_{m32} + \alpha_3 u_{m31}^2 - \alpha_1 u_{m31}^2 u_{m33}) \\ & -\xi_1(u_{m21} - u_{m21}u_{m22} + \alpha_3 u_{m21}^2 - \alpha_1 u_{m21}^2 u_{m23}), \\ \dot{e_2} = & (-u_{s42} + u_{s41}u_{s42} + V_2) \\ & +\xi_2(-u_{m12} + u_{m11}u_{m12})(\xi_2 u_{m32} - \xi_2 u_{m22}) \\ & +\xi_2 u_{m12}(\xi_2(-u_{m32} + u_{m31}u_{m32}) - \xi_2(-u_{m22} + u_{m21}u_{m22})), \\ \dot{e_3} = & (\alpha_2 u_{s43} + \alpha_1 u_{s41}^2 u_{s43} + V_3) \\ & +\xi_3(\alpha_2 u_{m13} + \alpha_1 u_{m11}^2 u_{m13})(\xi_3 u_{m33} - \xi_3 u_{m23}) \\ & +\xi_3 u_{m13}(\xi_3(\alpha_2 u_{m33} + \alpha_1 u_{m31}^2 u_{m33}) - \xi_3(\alpha_2 u_{m23} + \alpha_1 u_{m21}^2 u_{m23})). \end{cases} \tag{16}$$

Let us now select the active controllers by the rule:

$$V_1 = \zeta_1 - (g_4)_1 - M_1 e_1, \tag{17}$$

where $\zeta_1 = \xi_1(g_1)_1(\xi_1 u_{m31} - \xi_1 u_{m21}) + \xi_1 u_{m11}(\xi_1(g_3)_1 - \xi_1(g_2)_1)$, as described in (5).

$$V_2 = \zeta_2 - (g_4)_2 - M_2 e_2, \tag{18}$$

where $\zeta_2 = \xi_2(g_1)_2(\xi_2 u_{m32} - \xi_2 u_{m22}) + \xi_2 u_{m12}(\xi_2(g_3)_2 - \xi_2(g_2)_2)$,

$$V_3 = \zeta_3 - (g_4)_3 - M_3 e_3, \tag{19}$$

where $\zeta_3 = \xi_3(g_1)_3(\xi_3 u_{m33} - \xi_3 u_{m23}) + \xi_3 u_{m13}(\xi_3(g_3)_3 - \xi_3(g_2)_3)$ and $M_1 > 0$, $M_2 > 0$, $M_3 > 0$ are gain constants.
By substituting the active controllers (17), (18) and (19) into (16), we obtain

$$\begin{cases} \dot{e}_1 = & -M_1 e_1, \\ \dot{e}_2 = & -M_2 e_2, \\ \dot{e}_3 = & -M_3 e_3. \end{cases} \tag{20}$$

Lyapunov function $V(e(t))$ is designed as:

$$V(e(t)) = \frac{1}{2}[e_1^2 + e_2^2 + e_3^2]. \tag{21}$$

It is clear that Lyapunov function $V(e(t))$ is positive definite in R^3.
Therefore, the time-derivative of Lyapunov function $V(E(t))$ as defined in (20) may be formulated as:

$$\dot{V}(e(t)) = e_1\dot{e}_1 + e_2\dot{e}_2 + e_3\dot{e}_3. \tag{22}$$

Combining Eq.(20) and Eq.(22), we obtain

$$\dot{V}(e(t)) = -M_1 e_1^2 - M_2 e_2^2 - M_3 e_3^2 < 0,$$

which confirms that $\dot{V}(e(t))$ is negative definite.
By applying LST [39], we therefore conclude that HCDS error dynamics is asymptotic stable globally, i.e., the considered synchronization error $e(t) \to 0$ asymptotically as $t \to \infty$ for each initial values $e(0) \in R^3$.

5 Numerical Simulations and Discussions

We in this section present some numerical simulation results for the demonstration of the effectivity of investigated HCDS technique among chaotic GLVSs via ACS. The 4th order Runge-Kutta algorithm is used to solve the given ordinary differential equations. Initial conditions of the considering master systems (10)–(12) and corresponding slave system (13) are (27.5, 23.1, 11.4), (1.2, 1.2, 1.2), (2.9, 12.8, 20.3) and (14.5, 3.4, 10.1) respectively. In addition, the control gains (M_1, M_2, M_3) are taken as 2. Also, Fig. 2a–c display the HCDS synchronized trajectories of states of three master (10)–(12) and one slave system (13) respectively. Furthermore, synchronization errors $(e_1, e_2, e_3) = (-32.25, 271.36, -207.64)$ converging to 0 as t tends to infinity as depicted in Fig. 2d. Therefore, the discussed HCDS scheme among three master and one slave systems is illustrated numerically.

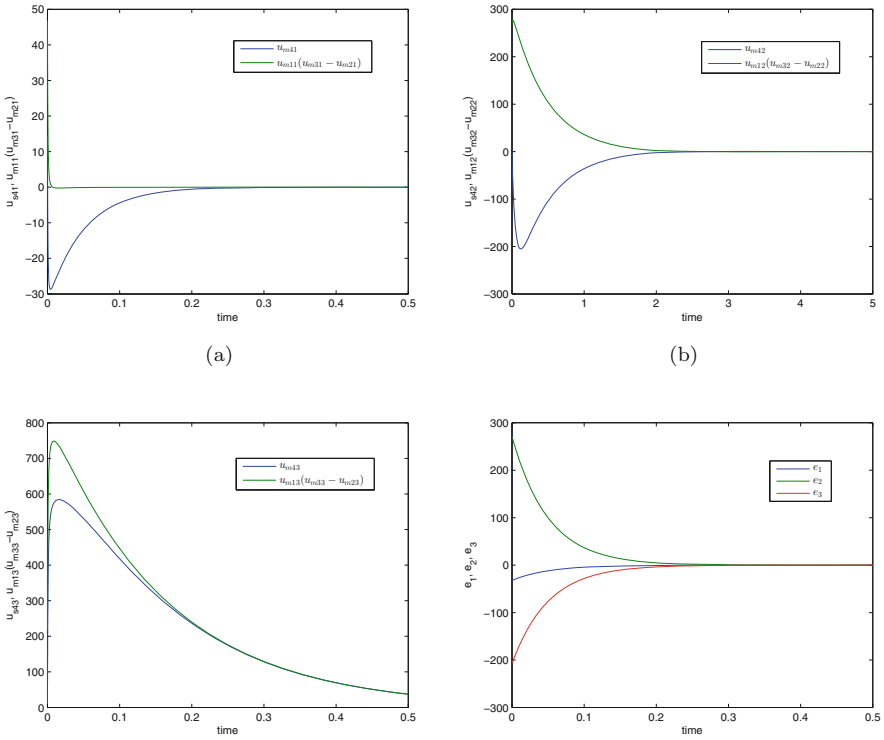


Fig. 2 HCDS synchronized trajectories for GLVS (**a**) between $w_{s41}(t)$ and $w_{m11}(t)(w_{m31}(t) - w_{m21}(t))$, (**b**) between $w_{s42}(t)$ and $w_{m12}(t)(w_{m32}(t) - w_{m22}(t))$, (**c**) between $w_{s43}(t)$ and $w_{m13}(t)(w_{m23}(t) - w_{m13}(t))$, (**d**) HCDS synchronized errors

5.1 A Comparative Study

Researchers [27] achieved hybrid synchronization in GLVSs via adaptive control method with same set of parameters. It is observed that synchronization error converges to zero at t =0.8(approx), while in this study, the HCDS technique is attained via ACS, in which it is noticed that the synchronization error converging to zero at t = 0.2(approx) as displayed in Fig. 2d. This justifies that our investigated HCDS scheme via ACS is much preferred over earlier published literature.

6 Conclusion

In this article, the HCDS scheme between four identical chaotic GLVSs via active control strategy has been investigated. Lyapunov stability theory and master-slave

configuration has been utilized to construct proper nonlinear active controllers. Additionally, numerical simulations using MATLAB software have been performed which suggest that the newly proposed control functions are highly effective in synchronizing and controlling the chaotic regime of GLVSs to desired set points that depicts the effectivity and supremacy of the considered HCDS scheme. Remarkably, the analytic theory is in excellent agreement with the numerical outcomes Also, in this direction we may expand the proposed HCDS approach in chaotic systems interfered with external disturbances as well as system uncertainties.

References

1. Alfred J Lotka, *Physical Review Letters* **64**(8), 821 (1990).
2. Francesco M Scudo, *Theoretical population biology* **2**(1), 1–23 (1971).
3. C Gavin, A Pokrovskii, M Prentice, and V Sobolev, *Journal of Physics:Conference Series, IOP Publishing* **55**, 80 (2006).
4. Pavlos Antoniou and Andreas Pitsillides, *Computer Communications* **33**(17), 2039–2047 (2010).
5. M Javidi and N Nyamoradi, *Applied mathematical modelling* **37**(20-210, 8946–8956 (2013).
6. Gurbir Perhar, Noreen E Kelly, Felicity J Ni, Myrna J Simpson, Andre J Simpson,and George B Arhonditsis, *Ecological informatics* **35**, 29–42 (2016).
7. Bi-Huei Tsai, Chih-Jen Chang, and Chun-Hsien Chang, *Energy* **100**, 416–424 (2016).
8. L Silva-Dias and A López-Castillo, *Mathematical biosciences* **300**, 36–46 (2018).
9. Alexandru Hening and Dang H Nguyen, *Journalof mathematical biology* **77**(1), 135–163 (2018).
10. AE Matouk and AA Elsadany, *Nonlinear Dynamics* **85**(3), 1597–1612 (2016).
11. P Gatabazi, JC Mba, E Pindza, and C Labuschagne, *Chaos, Solitons & Fractals* **122**, 47–57 (2019).
12. Junjing Xiong, Xiong Li, and Hao Wang, *Mathematical Biosciences* **1**(k1k2), 1–k1k2 (2019).
13. Taqseer Khan and Harindri Chaudhary, *Differential Equations and Dynamical Systems* **28**, 515–526 (2020).
14. Taqseer Khan and Harindri Chaudhary, *Applications & Applied Mathematics* **15**(2), 1135–1148 (2020).
15. Taqseer Khan and Harindri Chaudhary, *Poincare Journal of Analysis & Applications* **7**(2), 211–225 (2020).
16. Wael W Mohammed, ES Aly, AE Matouk, S Albosaily, and EM Elabbasy, *Results in Physics* **26**, 104432 (2021).
17. Taqseer Khan and Harindri Chaudhary, *Journal— MESA* **12**(2), 383–393 (2021).
18. Sami O Lehtinen, *Journal of Theoretical Biology* **510**, 110542 (2021).
19. Alain Arneodo, Pierre Coullet, and Charles Tresser, *Physics Letters A* **79**(4), 259–263 (1980).
20. Nikola Samardzija and Larry D Greller, *Bulletin of Mathematical Biology* **50**(5), 465–491 (1988).
21. Henri Poincaré, *Acta Mathematica* **13**(1), A3–A270 (1890).
22. Edward N Lorenz, *Journal of the atmospheric sciences* **20**(2), 130–141 (1963).
23. Louis M Pecora and Thomas L Carroll, *Physicalreview letters* **64**(8), 821 (1990).
24. Ajit K Singh, Vijay K Yadav, and S Das, *International Journal of Dynamics and Control* **5**(3), 756–770 (2017).
25. Chuandong Li and Xiaofeng Liao, *Chaos, Solitons & Fractals* **22**(4), 857–867 (2004).
26. Sanjay Kumar, Ahmed E Matouk, Harindri Chaudhary, and Shashi Kant, *International Journal of Adaptive Control and Signal Processing* **35**(4), 484–497 (2021).

27. Sundarapandian Vaidyanathan, *Internal Journal of PharmTech Research* **9**(1), 179–192 (2016).
28. Ping Zhou and Wei Zhu, *Nonlinear Analysis: Real World Applications* **12**(2), 811–816 (2011).
29. Ayub Khan and Harindri Chaudhary, *Arabian Journal of Mathematics* **9**(3), 597–611 (2020).
30. Lone Seth Jahanzaib, Pushali Trikha, Harindri Chaudhary, SM Haider, et al., *Journal of Mathematical and Computational Sciences* **10**(5), 1463–1480 (2020).
31. Vijay K Yadav, Vijay K Shukla, and Subir Das, *Chaos, Solitons &Fractals* **124**, 36–51 (2019).
32. Ayub Khan and Harindri Chaudhary, *Bloomsbury India*, 174–185 (2019).
33. Sundarapandian Vaidyanathan and Sivaperumal Sampath, *International Journal of Automation and Computing* **9**(3), 274–279 (2012).
34. Awad El-Gohary and MT Yassen, *Chaos, Solitons & Fractals* **12**(11), 2087–2093 (2001).
35. Sundarapandian Vaidyanathan, *International Journal of PharmTech Research* **8**(4), 622–631 (2015).
36. Er-Wei Bai and Karl E Lonngren, *Chaos, Solitons & Fractals* **8**(1), 51–58 (1997).
37. Junwei Sun, Yi Shen, Quan Yin, and Chengjie Xu, *Chaos: An Interdisciplinary Journal of Nonlinear Science* **23**(1), 013140 (2013).
38. Junwei Sun, Quan Yin, and Yi Shen, *Europhysics Letters* **106**(4), 40005 (2014).
39. Lawrence Perko, *SpringerScience & Business Media* (2013).

Modeling Competition in Motionless Populations

Hamlet Castillo Alvino, Marcos Marvá, and Ezio Venturino

1 Introduction

A key factor when modelling community dynamics consists in the way interactions take place. The vast majority of the research on community models, that follow in the wake of Lotka and Volterra [23], assumes that individuals are well mixed and that any one of them can interact with all the rest. Interestingly, epidemic and eco-epidemic models do take into account such a major feature distinguishing the type of transmission [4, 21], that is closely related to the way individuals interact. Recently, [1–3, 18] addressed a series of community models that implement a social structure determining how populations interact with each other. Some cases of predator-prey interactions of this type are instead discussed in [17].

In this paper we consider interference competition between two motionless populations, for which individual interaction dynamics is definitely different from mobile living beings [24]. Plants, of course, do belong to this category. Sessile species competition has been also reported in fungi [25], sponges [30], corals [5, 12, 14], giant clams, barnacles [10] or most of bivalves [27] are other examples, along with microorganism that grow up in colonies [20].

H. C. Alvino (✉)
Pontificia Universidad Católica "Madre y Maestra", STI Facultad Ciencias e Ingenierías, Escuela de Ciencias Naturales y Exactas, Santiago de los Caballeros, Republic of Dominican
e-mail: hh.castillo@ce.pucmm.edu.do

M. Marvá
Departamento de Física y Matemáticas, Universidad de Alcalá, Alcalá de Henares, Spain
e-mail: marcos.marva@uah.es

E. Venturino
Dipartimento di Matematica "Giuseppe Peano", Università di Torino, Torino, Italy
e-mail: ezio.venturino@unito.it

R. P. Mondaini (ed.), *Trends in Biomathematics: Stability and Oscillations in Environmental, Social, and Biological Models*,
https://doi.org/10.1007/978-3-031-12515-7_10

We thus consider two populations that homogeneously occupy two neighboring territories, but do not mix with each other, at least not significantly. Interactions among the two species populations can only occur through the common boundaries of the two cultures in consideration, that border each other. Note that these different domains occupied by the two populations may well consist also of islands or other more complicated geometrical shapes.

Thus, these considerations lead us to replace the classical 1–1 interactions among individuals of the same species with just those with the most immediate neighbors of a single individual [18]. In the model we thus must prevent that one individual may compete with all other individuals of the population that are in far away locations.

When the model is written in terms of *explicit carrying capacities*, the dynamics of interference competition is driven by the balance between intra- and interspecific competition [24, 31], or, if the model is formulated with *emerging carrying capacities*, by competitive strengths [6, 15, 26]. We prefer the latter approach, see [16]. In any case, the early theory of competition [11, 19, 29] understands coexistence as the result of the common interplay of inter- and intraspecific interactions. Specifically coexistence is obtained when intraspecific competition limits species density more strongly than interspecific competition. In the present model competition is by far different than in the classical model, since 1–1 interactions among all individuals are precluded. In particular, both intra- and interspecific competition are relaxed. Therefore, we expect competition outcomes to be different than in the classical model as in the recent extensions [8, 9]. Indeed, interactions seem to be milder in the model we present herein, so that we expect to find that species are more likely to coexist.

From a mathematical point of view, the model we present belongs to the general class of *competitive systems* in the positive cone, see [13] and the references therein. This fact precludes the existence periodic solutions and guarantees that solutions converge to an equilibrium point.

Our main results are the following. When a sessile population competes with a mobile population, all the competitive outcomes of the classical model are possible. Besides, conditional bi-stable coexistence in favor of the sessile population is possible, meaning that both a semi-trivial equilibrium and a coexistence equilibrium are (locally) asymptotically stable. Thus, whether the sessile population wins or both populations coexist depends on the system's initial conditions, i.e. the initial amount of individuals of each population.

When both competing populations are sessile only species exclusion, with outcome determined just by the system's initial conditions, or conditional tri-stable coexistence, i.e. coexistence, are possible, as it was already shown in [22].

The manuscript is organized as follows: in the next section we compare the classical logistic single population evolution with the corresponding one of a motionless population. In Sect. 3, we set the features of non-mobile populations for which the interactions occur only at the boundary of their respective domains into a mathematical formulation. Subsequently, we analyze the competition among mobile and motionless populations in Sect. 4. The interactions of two motionless

populations are instead analysed in Sect. 5. Finally, we discuss the results achieved for the various cases in Sect. 6.

2 The Single Population Case

The single population case has been briefly examined also in [18], as a motivation for further changes in the formulation of herd behavior models. Here however we focus on a population that does not move, which therefore has its own specific features, distinguishing it from the more commonly considered mobile populations.

The starting point for a single population is represented by the classical logistic (or Verhulst) model, namely

$$x' = \widehat{r}x - ax^2, \tag{1}$$

for which the the population settles at the equilibrium

$$x^c = \frac{\widehat{r}}{a}. \tag{2}$$

For one single plant living in a plantation, therefore surrounded just by other plants of the same species, because the interactions occur only with possibly a fraction $b \leq 1$ of the closest neighbors, the model becomes instead:

$$x' = \widehat{r}x - bax\sqrt{x}. \tag{3}$$

Equation (3) is a modification of the well-known logistic growth that takes into account that sessile living beings stand still and interact only with their most immediate neighbors. Thus, a given individual competes with the nearby ones that we assume are located at the boundary layer of its “vital space", that is, a circle area around the individual. Assuming that individuals are homogeneously distributed, the boundary of each individual’s *vital space* is proportional to the boundary of the territory occupied by this species, i.e., proportional to $\sqrt{x}$. Clearly, of the two possible equilibria of (3), the origin is unstable while the population thrives at level

$$x^* = \frac{\widehat{r}^2}{b^2a^2}. \tag{4}$$

Thus, qualitatively, the two models (1) and (3) behave in the same way. However, recalling that $b \leq 1$, whether a single plant living in a wood or plantation is better off than a corresponding animal individual living amidst his own consimilars, depends on the ratio between its net reproduction rate and the intraspecific competition rate. If the former exceeds the latter, the level at which the vegetable species settles is certainly higher than the corresponding animal population.

3 The Competing Sessile Populations Model

The departure model in this case is the classical Lotka-Volterra competition model with *emergent carrying capacities* [6, 15, 26] rather than explicit carrying capacities [23, 31]:

$$\begin{cases} x_1' = r_1 x_1 - a_{11} x_1^2 - a_{12} x_1 x_2 \\ x_2' = r_2 x_2 - a_{22} x_2^2 - a_{21} x_1 x_2 \end{cases} \tag{5}$$

where x_i and $r_i > 0$ stand for the amount of individuals and the intrinsic growth rate of species $i = 1, 2$, respectively. Coefficients $a_{ij} > 0$ account for intra- $(i = j)$ and interspecific $(i \neq j)$ competition, for $i, j = 1, 2$.

The modeling of sessile populations is rather different from the classical interacting populations of animals that can move around. As assumed in other investigations concerning herds of herbivores and their predators, [1, 2, 17] or interference competition [22], we assume here that interactions among different species, uniformly located in specific territories, occur through their common boundaries, as stated above. They are assumed to be smooth, motivating the use of the square root in the interaction terms.

Thus, we denote by x_1 and x_2 the densities of the populations, i.e., the number of individuals per surface unit, occupying an area S. Thus the species of the population i found in the periphery or neighborhood of their environment are proportional to the perimeter of the patch where the culture is located, whose length depends on $\sqrt{S}$. They are therefore proportional in number to the square root of the density, i.e. to $\sqrt{x_i}$, $i = 1, 2$. Different shapes could be accommodated by taking a different exponent, other than one-half, in the model formulation. However, in part based also on the results of [7] for which no fundamentally different results arise, and for the sake of simplicity, we confine ourselves to the assumption of a smooth boundary. This, as mentioned, entails the use of the exponent one half in the model formulation. We next examine in detail the intra- and inter- specific interaction terms:

- Inter-specific interactions take place on the boundary of each species' domain. In view of the previous discussion, as individuals are assumed to be homogeneously distributed, the interaction between species x_i and x_j takes the following form:

$$- d_i a_{ij} \sqrt{x_i} \sqrt{x_j}, \tag{6}$$

  where the minus sign denotes interactions harming the population i under consideration, a_{ij} stands for the competition interaction coefficient of species j on population i; $1 \geq d_i \geq 0$ is a constant that scales competition to the fraction of the common perimeter. If species do not interact, i.e. there is no common boundary, then $d_i = 0$. Therefore, $d_i a_{ij}$, $i \neq j$ stands for interspecific competition and includes information on the fraction of the boundary where

competition takes place, so that its interpretation is slightly different from the interaction coefficients of the classical model. In general $d_i \neq d_j$.

- For intraspecific dynamics, we must include the growth rate and possibly intraspecific competition and therefore use replicas of (3):

$$x_i' = \widehat{r}_i x_i - b_i a_{ii} x_i \sqrt{x_i}, \tag{7}$$

where $\widehat{r}_i$ stands for the net intrinsic growth rate, a_{ii} is the intraspecific competition rate. In this context, b_i in (7) stands for the proportion between the local boundary and the perimeter occupied by the entire population, so that $1 \geq b_i \geq 0$. It is assumed to be the same for all individuals of the same species.

Merging (7) and (6) yields the competing sessile populations model

$$\begin{cases} x_1' = \widehat{r}_1 x_1 - b_1 a_{11} x_1 \sqrt{x_1} - d_1 a_{12} \sqrt{x_1}\sqrt{x_2}, \\ x_2' = \widehat{r}_2 x_2 - b_2 a_{22} x_2 \sqrt{x_2} - d_2 a_{21} \sqrt{x_1}\sqrt{x_2}. \end{cases} \tag{8}$$

Theorem 3.1 *The positive solutions of system* (8) *are bounded from above.*

Proof Note that

$$x_i' = \widehat{r}_i x_i - b_i a_{ii} x_i \sqrt{x_i} - d_i a_{ij} \sqrt{x_j}\sqrt{x_i} < \widehat{r}_i x_i - a_{ii} x_i \sqrt{x_i} < 0$$

for $x_i > \left(r/a_{ij}\right)^2$. In particular, in the positive cone we define a suitable box B with one corner located at the origin and the opposite one at the point $V = \left((r_1/d_1 a_{12})^2, (r_2/d_2 a_{21})^2\right)$. The situation thus corresponds to the flow entering into B. ■

Remark 3.1 System (8) is a *competitive system* (sensu [13]) in the positive cone $\mathbb{R}^2_+ := (0, +\infty) \times (0, +\infty)$, in view of the fact that

$$\frac{\partial}{\partial x_j}\left(\widehat{r}_i x_i - b_i a_{ii} x_i \sqrt{x_i} - d_i a_{ij} \sqrt{x_j}\sqrt{x_i}\right) < 0, \qquad j \neq i,\ i, j = 1, 2,$$

and the flow of the system belongs to class $C^1\left(\mathbb{R}^2_+\right)$. Thus, Theorem 3.1 along with [13] imply that all the positive solutions of system (8) converge to an equilibrium point.

Remark 3.2 Note that as pointed out in [28], the right hand side of (8) does not satisfy the Lipschitz condition, with a consequent loss of uniqueness of the solution trajectories on the coordinate axes. Thus we need special care in treating vanishing populations when we change the variables of the system (8) to obtain the auxiliary system in order to eliminate the singularity. Therefore, to study the trivial and semi-

trivial equilibrium points, we have to turn to the original formulations (8), compare the approach of [3].

Hence, we will study the trivial and semi-trivial equilibria directly using model (8). The trivial and semi-trivial equilibria are the following points:

$$E_0 = (0,0), \quad E_1 = \left(\left(\frac{r_1}{b_1 a_{11}}\right)^2, 0\right), \quad E_2 = \left(0, \left(\frac{r_2}{b_2 a_{22}}\right)^2\right) \tag{9}$$

Note also that there are square root terms in system (8), so that the stability of the trivial and semi-trivial equilibrium points (9) cannot be assessed using the Jacobian matrix which, in turn, works when dealing with coexistence equilibria. However, square roots make the Jacobian matrix to be involved; we next introduce an equivalent singularity-free system to overcome such a problem.

4 Mobile and Sessile Populations Interactions

We consider here the interactions between a mobile population competing with a non-mobile one. Thus the system is a combination of both (5) and (7), giving:

$$\begin{cases} x_1' = r_1 x_1 - a_{11} x_1^2 - a_{12} x_1 x_2, \\ x_2' = r_2 x_2 - a_{22} b_2 x_2 \sqrt{x_2} - a_{21} x_1 x_2. \end{cases} \tag{10}$$

Proceeding as for system (8), it is easy to realize that the trajectories of (10) are bounded from above. To analyze the existence of equilibrium points and the long term behavior of the solutions of (10), we rewrite it in the following more convenient form:

$$\begin{cases} x_1' = r_1 x_1 \left(1 - \dfrac{a_{11}}{r_1} x_1 - \dfrac{a_{12}}{r_1} x_2\right), \\ x_2' = r_2 x_2 \left(1 - \dfrac{a_{22}}{r_2} b_2 \sqrt{x_2} - \dfrac{a_{21}}{r_2} x_1\right). \end{cases} \tag{11}$$

We rescale the above system with a special change of variables and parameters, namely $w_i = \frac{a_{ii}}{r_i} x_i, \quad c_{ij} = \frac{a_{ij} r_j}{a_{jj} r_i}$, to obtain

$$\begin{cases} w_1' = r_1 w_1 \left(1 - w_1 - c_{12} w_2\right), \\ w_2' = r_2 w_2 \left(1 - b_2 \sqrt{\dfrac{a_{22}}{r_2}} \sqrt{w_2} - c_{21} w_1\right). \end{cases} \tag{12}$$

A further rescaling in system (12) given by $\widehat{w}_2 = \frac{a_{22}}{r_2} w_2, \quad \widehat{c}_{12} = \frac{r_2}{a_{22}} c_{12}$, yields the so-called special auxiliary system:

$$\begin{cases} w_1' = r_1 w_1 \left(1 - w_1 - \widehat{c}_{12} \widehat{w}_2\right), \\ w_2' = r_2 \widehat{w}_2 \left(1 - b_2 \sqrt{\widehat{w}_2} - c_{21} w_1\right). \end{cases} \tag{13}$$

Let us recall that Theorem 3.1, as well as remarks 3.1 and 3.2 hold *mutatis mutandi*.

4.1 Equilibria

The trivial and semi-trivial equilibria of system (13) are:

$$E_0 = (0, 0) \quad E_1 = (1, 0) \quad \text{and} \quad E_2 = \left(0, \frac{1}{b_2^2}\right) \tag{14}$$

We consider the nullclines of system (13), that are given by

$$\widehat{w}_2 = f_1(w_1) = \frac{1 - w_1}{\widehat{c}_{12}}, \qquad \widehat{w}_2 = f_2(w_1) = \frac{(c_{21} w_1 - 1)^2}{b_2^2}.$$

The coexistence equilibria are denoted by $E_3^{\pm} = (w_1^{\pm}, f_1(w_1^{\pm}))$. They are given by the intersection of the nullclines, in this case a curve and a straight line. These solutions can be obtained from the roots of the following quadratic equation:

$$P_{sc}(w) = \frac{1}{\widehat{c}_{12} b_2^2} \left[-\widehat{c}_{12} c_{21}^2 w^2 + (2\widehat{c}_{12} c_{21} - b_2^2) w + (\widehat{c}_{12} - b_2^2) \right].$$

Thus

$$w_1 = \frac{2\widehat{c}_{12} c_{21} - b_2^2 \pm b_2 \sqrt{4\widehat{c}_{12} c_{21} (c_{21} - 1) + b_2^2}}{2\widehat{c}_{12} c_{21}^2}. \tag{15}$$

Imposing that the discriminant of expression (15)

$$D := 4\widehat{c}_{12} c_{21} (c_{21} - 1) + b_2^2 \tag{16}$$

is nonnegative, we find the real roots in the positive cone. The following Lemma 4.1 and Theorem 4.1 summarize these conditions.

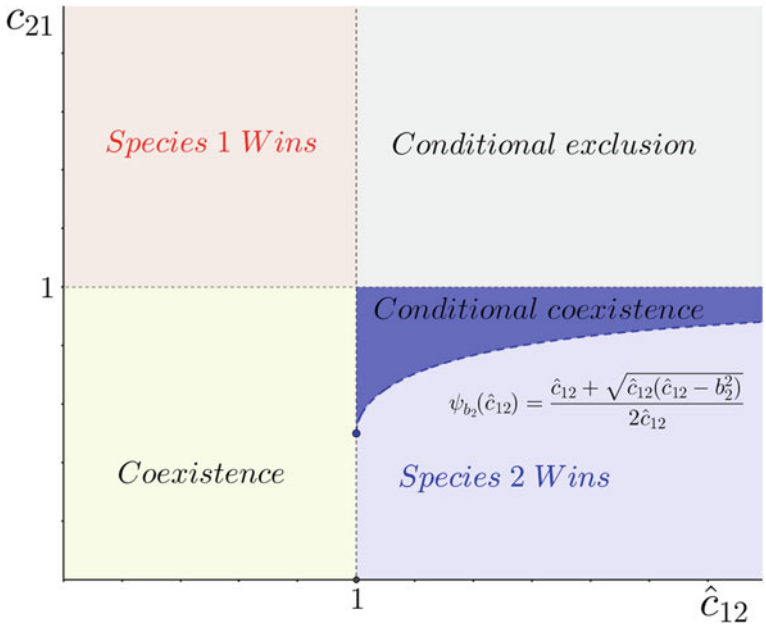


Fig. 1 Competition outcomes of system (10) in the $\widehat{c}_{12} - \widehat{c}_{21}$ parameter space as function of the competitive strengths as defined in Eq. (13). Species 1 (resp. 2) refers to the sessile (resp. mobile) population

Lemma 4.1 *Consider the function*

$$c_{21} = \psi_{b_2}(\widehat{c}_{12}) := \frac{\widehat{c}_{12} + \sqrt{\widehat{c}_{12}(\widehat{c}_{12} - b_2^2)}}{2\widehat{c}_{12}}, \tag{17}$$

then, ψ_{b_2} is an unimodal function such that:

(1) its domain is the set $\left\{x \in \mathbb{R}^+ \mid 1 \leq x < \infty\right\}$

(2) $c_{21}^* = \psi_{b_2}(1) = \frac{1+\sqrt{1-b_2^2}}{2}$ *and* $\lim_{\widehat{c}_{12} \to +\infty} \psi_{b_2}(\widehat{c}_{12}) = 1$

Proof It follows from direct calculations (Fig. 1). ■

Whether there is none, one or two equilibrium points is determined by the sign of the discriminant (16) of (15) and the quantities defined in the previous Lemma 4.1.

Theorem 4.1 *Consider the system* (13) *and the function* (17)*. Recalling (16) we find*

1. *System* (13) *has no equilibrium points in the positive cone if either*
   - *(a)* $D < 0$*, see the middle right panel in Fig. 2.*
   - *(b) Both* $1/c_{21} < 1$ *and* $1/\widehat{c}_{12} > 1/b_2^2$ *hold, see top left panel in Fig. 2. Alternatively we can require* $\widehat{c}_{12} < 1$ *and* $c_{21} > 1$.

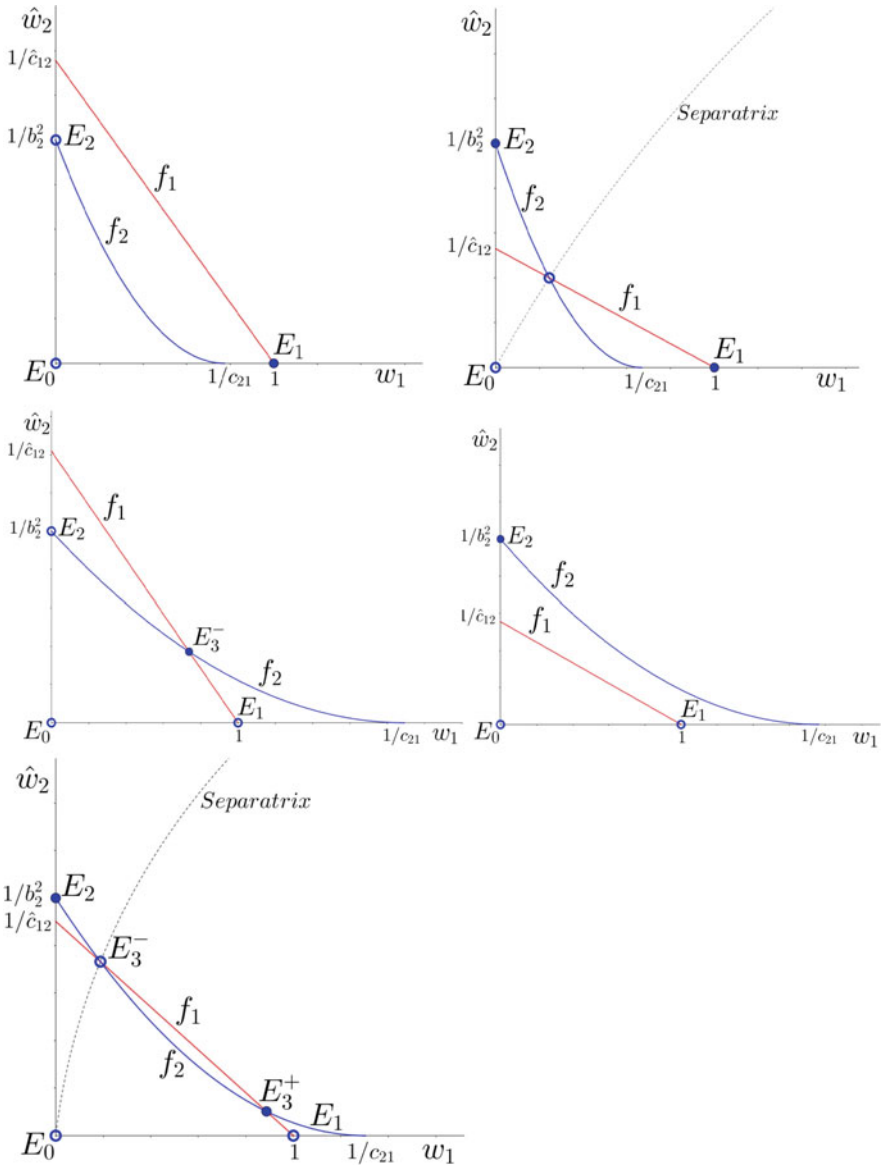


Fig. 2 Possible phase portraits of system (13). The horizontal (resp. vertical) axis refers to the sessile (resp. mobile) population. The curve f_1 (resp. f_2) stands for the nullcline of the sessile (resp. mobile) population. Solid points represent locally asymptotically stable equilibrium points while empty points represent unstable equilibria

2. *System* (13) *possesses a single equilibrium point in the positive cone (apart from the degenerated case* $D = 0$*) if either*

(a) *Both* $1/\widehat{c}_{12} < 1/b_2^2$ *and* $1/c_{21} < 1$ *hold, see top right panel in Fig. 2. An alternative formulation of the above inequalities is* $\widehat{c}_{12} > 1$ *and* $c_{21} > 1$.
(b) *Both* $1/\widehat{c}_{12} > 1/b_2^2$ *and* $1/c_{21} > 1$. *An alternative formulation of the above inequalities is* $\widehat{c}_{12} < 1$ *and* $c_{21} < 1$.

3. *System* (13) *has two equilibrium points in the positive cone if* $\widehat{c}_{12} > 1$ *and* $\psi_{b_2}(\widehat{c}_{12}) < c_{21} < 1$ *hold, see the bottom panel of Fig. 2.*

Proof It follows from direct calculations on (15) and (16) and geometrical considerations on the intersection of f_1 and f_2 with the axes. ■

4.2 Stability

We focus now on the stability of the existing equilibrium points.

Theorem 4.2 *Consider system* (13)*, assuming that* $r_1 > 0$ *and* $r_2 > 0$. *Then,*

1. *The origin is always unstable.*
2. *Assume now* $\widehat{c}_{12} > 1$ *and* $c_{21} > 1$. *Then by condition 2.(a) of Theorem 4.1, there exists a single coexistence equilibrium point* E_3^+*, that is unstable. There exists a separatrix line connecting the origin with* E_3^+ *that defines the basins of attraction of* E_1 *and* E_2.
3. E_1 *is globally asymptotically stable if and only if* $0 < \widehat{c}_{12} < 1$ *and* $c_{21} > 1$.
4. E_3^- *is globally asymptotically stable if and only if* $0 < \widehat{c}_{12} < 1$ *and* $0 < c_{21} < 1$.
5. *Finally, assume now* $0 < c_{21} < 1$ *and* $\widehat{c}_{12} > 1$. *Then, by the condition 2.(b) of Theorem 4.1,*

   (a) E_3^+ *and* E_2 *are locally asymptotically stable while* E_-^c *is unstable. There exist a separatrix connecting the origin with* E_-^c *that defines the basins of attraction of* E_2 *and* E_3^+*, if and only if* $\widehat{c}_{12} > 1$ *and* $\psi_{b_2}(\widehat{c}_{12}) < c_{21} < 1$.
   (b) E_2 *is globally asymptotically stable if and only if* $0 < c_{21} < \psi_{b_2}(\widehat{c}_{12})$ *and* $\widehat{c}_{12} > 1$.

Proof Let us consider the Jacobian matrix of (10)

$$J_F(w_1, \widehat{w}_2) := \begin{pmatrix} r_1(1 - w_2 - \widehat{c}_{12} w_2) & -r_1 \widehat{c}_{12} w_1 \\ -r_1 c_{21} \widehat{w}_2 & r_2(1 - b_2\sqrt{\widehat{w}_2} - c_{21} w_1) - \frac{1}{2} r_2 b_2 \sqrt{\widehat{w}_2} \end{pmatrix}$$

We analyze the characteristic equation and apply the Routh-Hurwitz criterion to the equilibrium points for the various cases of the statement of the theorem:

1. It is easy to see that in the system (13), $w_i' > 0$ for $w_1, \widehat{w}_2 \sim 0$.
2. The statement holds because the eigenvalues of $J_F(E_1)$ are $\lambda_1 = -r1$, $\lambda_2 = r_2(1 - c_{21})$. Therefore, E_3^- is a saddle and E_2 and E_1 are locally asymptotically

stable; in such case, there exists a separatrix line through both E_3^- and the origin that defines the basins of attraction of E_1 and E_2.

3. Direct calculations yield the eigenvalues of $J_F(E_2)$, $\lambda_1 = -r_2(1 - \frac{3}{2}b_2)$, $\lambda_2 = r_1(1 - \widehat{c}_{12})$. The statement implies that E_2 is stable while E_1 is unstable. Also theorem 4.1 shows that no equilibrium points exist in the positive cone, and the flow of the system makes E_2 globally asymptotically stable.
4. This statement follows by a standard analysis of the flow of the system.
5. We focus first in assessing the stability in case of two coexistence equilibrium points. This scenario in the system dynamics may arise, when coefficients vary, essentially in two different ways. On one hand, when D changes from being negative to positive. On the other hand, already there exits a single coexistence equilibrium point and the x_2-nullcline f_2 crosses one of the semi-trivial equilibrium points in such a way that a second one appears. The dynamical scenario is the same, no matter how it arises.

   Let us rewrite the Jacobian matrix in a more convenient form. System (13) is of the form $w_i' = w_i f_i(w_i, w_j)$, so that at any coexistence equilibrium point $E = (E_1, E_2)$ it follows that $f_1(E) = 0 = f_2(E)$. Thus:

$$J_F(E) := \begin{pmatrix} -r_1 E_1 & -r_1 \widehat{c}_{12} E_1 \\ -r_2 c_{21} E_2 & -\frac{1}{2} r_2 b_2 \sqrt{E_2} \end{pmatrix} \tag{18}$$

   From expression (15) for $D = 0$, the characteristic polynomial of (18) at the equilibrium point arising when f_1 and f_2 collide is

$$\lambda^2 + \left(\frac{b_2^2(c_{21} r_2 - 2r_1) + 4\widehat{c}_{12} c_{21} r_1}{4\widehat{c}_{12} c_{21}^2} \right) \lambda. \tag{19}$$

   Thus, one eigenvalue is $\lambda_1 = 0$ and the other one, λ_2, because of the hypotheses of this statement, is negative; in particular, both eigenvalues are simple. Thus, the eigenvalues are continuous under small perturbations of the parameters involved in expression (18). As E_3^- and E_3^+ appear, λ_2 keeps being negative in the corresponding Jacobian matrices while λ_1 becomes negative for $J_F(E_3^-)$ and positive for $J_F(E_3^+)$. This is shown by a standard analysis of the system flow after the bifurcation takes place.

   The second statement of 5. can be proved as 4. ■

5 Two Sessile Populations Interactions

System (8) can be rewritten in a more convenient form introducing new variables and rescaling coefficients by setting

$$x_i = z_i^2, \quad r_i = \frac{\widehat{r_i}}{2}, \quad c_{ii} = b_i a_{ii}, \quad c_{ij} = \frac{d_i a_{ij}}{2}, \quad i, j = 1, 2. \tag{20}$$

This yields the so-called *auxiliary system*:

$$\begin{cases} z_1' = r_1 z_1 - c_{11} z_1^2 - c_{12} z_2 \\ z_2' = r_2 z_2 - c_{22} z_2^2 - c_{21} z_1 \end{cases} \tag{21}$$

It is apparent that the non-negative semi-axes are not invariant for the flow of system (21), so that this system does not help in assessing the stability of the trivial and semi-trivial equilibrium points.

As before we address first the existence of equilibrium points and then analyze their stability.

5.1 Equilibria

As mentioned earlier the trivial and semi-trivial equilibria of system (8) are given by (9). As for the coexistence equilibria, we consider the nullclines of the auxiliary system (21) corresponding to system (8), given by

$$z_2 = f_1(z_1) = z_1 \frac{r_1 - c_{11} z_1}{c_{12}}, \quad z_1 = f_2(z_2) = z_2 \frac{r_2 - c_{22} z_2}{c_{21}}.$$

The nullclines cross each other at the origin and their curvature and location imply that they meet up to three times or none on the positive cone, see Fig. 3, the discussion below and the discussion and conclusion Sect. 6.

However, system (8) may exhibit two additional coexistence states in the positive cone. Substituting the second equation into the first one we obtain a fourth-degree equation:

$$Q_c(z_1) = \frac{1}{c_{12}^2} z_1 P_c(z_1) = 0, \tag{22}$$

where $Q_c(z_1) = f_2(f_1(z_1))$ and

$$P_c(z_1) = c_{22} c_{11}^2 z_1^3 - 2 c_{22} c_{11} r_1 z_1^2 + (c_{11} c_{12} r_2 + c_{22} r_1^2) z_1 + c_{12}(c_{12} c_{21} - r_1 r_2).$$

The geometry of the phase portrait implies that the coexistence equilibrium lies in the box with the origin and $(r_1/c_1, r_2/c_2)$ as opposite vertices. Therefore, we use Sturm's Theorem to account for the number of positive roots of equation $P_c(z_1) = 0$. Let us recall that the Sturm's sequence of equation $P_c(z_1) = 0$ is given by

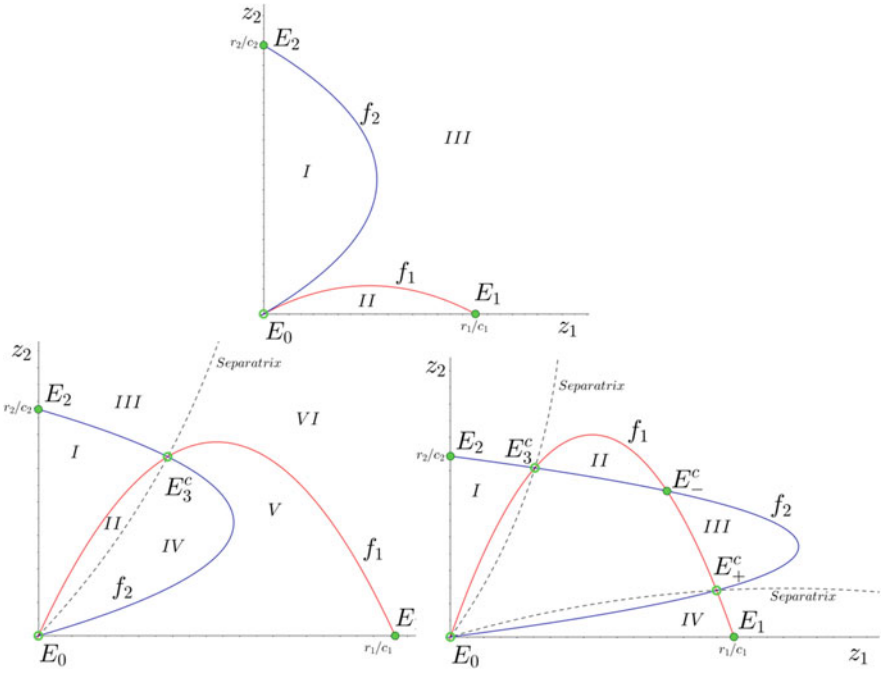


Fig. 3 Possible phase portrait of the corresponding auxiliary competition model (21). The horizontal (resp. vertical) axis refers to the sessile (resp. mobile) population. The curve f_1 (resp. f_2) stands for the nullcline of the sessile (resp. mobile) population. Solid points are locally asymptotically stable equilibria while non solid points are unstable equilibria. Left panel: for the parameter values $r_1 = 0.8$, $r_2 = 0.6$, $c_{12} = 1.2$, $c_{21} = 0.95$, $c_{11} = 0.47$, $c_{22} = 0.74$ condition (26) holds. Central panel: the parameter values $r_1 = 1$, $r_2 = 1$, $c_{12} = 1.51$, $c_{21} = 0.52$, $c_{11} = 0.66$, $c_{22} = 1.24$ fulfill conditions (27) and (28). Thus, the system exhibits the competitive exclusion principle, as for the classical competition model. Right panel: this scenario represents the tri-stable scenario. It is obtained for the parameter values $r_1 = 1$, $r_2 = 1$, $c_{12} = 0.28$, $c_{21} = 0.13$, $c_{11} = 0.47$, $c_{22} = 0.74$. In this case either one of the population could outcompete the other one, or both may coexist. In any case, the ultimate outcome of the system is determined just by the initial values

$$Seq_c(z_1) = \left\{P_c(z_1), P_c'(z_1), R_1(z_1), R_2(z_1)\right\},$$

where the second term is its derivative of $P_c(z_1)$, and the remaining terms $R_i(z_1)$, $i = 1, 2$, are the remainders of the Euclidean divisions:

$$R_1(z_1) = -\text{rem}(P_c(z_1), P'(z_1)), \quad R_2(z_1) = -\text{rem}(P'(z_1), R_1(z_1)).$$

Then evaluating each term of the Seq_c at the ends of interval $[0, r_1/c_{11}]$ we get the number of positive roots of $P_c(z_1) = 0$ as the number of changes of sign in $Seq_c(0)$ minus the number of change of signs in $Seq_c(\frac{r_1}{c_{11}})$. Thus, the actual Sturm's sequence is calculated on $P_c(z_1)$ instead of on $Q_c(z_1)$. Direct calculations lead to

$$P_c'(z_1) = 3c_{11}^2 c_{22} z_1^2 - 4c_{11}c_{22}r_1 z_1 + c_{11}c_{12}r_2 + c_{22}r_1^2 \tag{23}$$

$$R_1(z_1) := -\left(\frac{2r_2 c_{11}c_{12}}{3} - \frac{2c_{22}r_1^2}{9}\right) z_1 - \frac{9c_{11}c_{12}^2 c_{21} - 7c_{11}c_{12}r_1 r_2 + 2c_{22}r_1^3}{9c1} \tag{24}$$

$$R_2(z_1) := -\frac{9c_{11}c_{12}^2}{4(3c_{11}c_{12}r_2 - c_{22}r_1^2)^2}[4c_{11}^2 c_{12}r_2^3 + 27c_{11}c_{12}^2 c_{22}c_{21}^2 - 18c_{11}c_{12}c_{22}c_{21}r_1 r_2 - c_{11}c_{22}r_1^2 r_2^2 + 4c_{22}^2 c_{21}r_1^3] \tag{25}$$

Theorem 5.1 *System* (8) *has no equilibrium points in the positive cone if*

$$c_{12}c_{21} - r_1 r_2 > 0 \tag{26}$$

Proof Consider the Sturm's sequences $Seq(0)$

$$P_c(0) = c_{12}\left(c_{12}c_{21} - r_1 r_2\right), \qquad P_c'(0) = c_1 c_{12}r_2 + c_{22}r_1^2,$$

$$R_1(0) = \frac{-1}{9}\left(\frac{9c_{11}c_{12}^2 c_{21} - 7c_{11}c_{12}r_1 r_2 + 2c_{22}r_1^3}{c_{11}}\right)$$

and $R_2(0)$ is given by (25). On the other hand, $Seq(r_1/c_{11})$ consists of

$$P_c\left(\frac{r_1}{c_{11}}\right) = c_{12}^2 c_{21}, \quad P_c'\left(\frac{r_1}{c_{11}}\right) = r_2 c_{11}c_{12},$$

$$R_1\left(\frac{r_1}{c_{11}}\right) = \frac{-c_{12}(9c_{12}c_{21} - r_1 r_2)}{9}$$

and $R_2(r_1/c_{11})$ is, again, given by (25), since it does not depend on z_1. Rearranging terms

$$R_2(0) = \frac{-9c_{11}c_{12}^2}{4(3c_{11}c_{12}r_2 - c_{22}r_1^2)^2}[4\left(c_{11}^2 c_{12}r_2^3 + c_{22}^2 c_{21}r_1^3\right) + c_{11}c_{22}(27c_{12}^2 c_{21}^2 - 18c_{12}c_{21}r_1 r_2 - r_1^2 r_2^2)].$$

Note that the sign of $R_2(0) = R_2(r_1/c_{11})$ does not matter, since it is the same for $Seq(0)$ and $Seq(r_1/c_{11})$. Therefore, gathering signs yield $Seq(0) =$

$\{+,+,-,sign(R_2(0))\}$ and $Seq(r_1/c_{11}) = \{+,+,-,sign(R_2(0))\}$, which concludes the proof. ■

Assume now that condition (26) fails. Solving Eq. (22) under the assumption that $c_{12}c_{21} - r_1r_2 = 0$ yields two complex roots along with $z_1 = 0$ as unique real root with multiplicity 2. Recall that an equilibrium point in the third quadrant exists when (26) holds. This equilibrium collides with the origin when $c_{12}c_{21} - r_1r_2 = 0$ and appears in the positive cone as $c_{12}c_{21} - r_1r_2$ becomes negative. Besides, from the analysis of the Sturm's sequence for $c_{12}c_{21} - r_1r_2 < 0$ at least one, and up to three, positive coexistence equilibrium points exist.

Criteria leading to each outcome can be stated by assessing conditions that control the change of sign of the terms of the Sturm's sequence. Determining general conditions is a hard task, since the terms of $Seq(0)$ and $Seq(r_1/c_{11})$ depend on up to 6 parameters. In any case, we equate to zero each term with undetermined sign of the Sturm's sequence and solve each expression for one parameter to obtain conditions on the sign of each term. Solving the equations for c_i (resp. c_{ij}) yield conditions for the exclusion or conditional tri-stability that depend on the so-called handling time (resp., on the competition strength) of each species.

Theorem 5.2 *Consider system* (8) *and assume that*

$$c_{12}c_{21} - r_1r_2 < 0 \tag{27}$$

Then,

1. *Assume that*

$$9 < \frac{r_1r_2}{c_{12}c_{21}} \tag{28}$$

and consider the straight lines

$$c_{22}^{\pm}(c_{11}) = \frac{1}{8c_{21}r_1^3}c_{11}\Big[18c_{12}c_{21}r_1r_2 - 27c_{12}^2c_{21}^2 + r_1^2r_2^2 \pm\sqrt{(c_{12}c_{21} - r_1r_2)(9c_{12}c_{21} - r_1r_2)^3}\Big] \tag{29}$$

in the $c_{11} - c_{22}$ parameter space, arising by setting $R_2(0) = 0$, see (25)*. These lines define a sector region in the positive cone, see Fig. 4. Then,*

(a) The slope of $c_2 = c_2^{\pm}(c_{11})$ is positive.
(b) If (c_{11}, c_{22}) lies in between $c_2 = c_2^{\pm}(c_{11})$ three nontrivial equilibrium points $E_{\pm}^c$, E_3 in the positive cone exist. Ordered by its first component, $E_-^c < E_3^c < E_+^c$.
(c) If (c_{11}, c_{22}) does not lie in between $c_2 = c_2^{\pm}(c_{11})$ then there exists a single nontrivial equilibrium point E_3 in the positive cone.

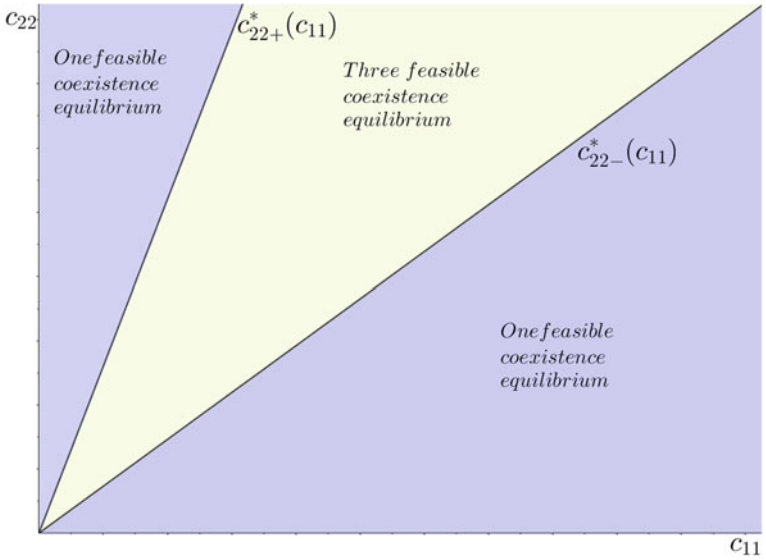


Fig. 4 Species competition outcomes in the $c_{11} - c_{22}$ parameter space, as function of the intraspecific competition $c_i, i, j = 1, 2$. The straight lines $c^*_{2\pm}(c_{11})$ show the thresholds values separating the regions where one or three coexistence equilibria exist in the positive cone. Namely, there exist three coexistence equilibria region consists of the (c_{11}, c_{22}) such that $c^*_{22-}(c_{11}) < c_{22} < c^*_{22+}(c_{11})$ (green colored area). Outside it, just one coexistence equilibrium is possible. The figure is generated with the parameter values $c_{12} = 0.5$, $c_{21} = 0.07$, $r_i = k_i = 1, i = 1, 2$

2. *Further, if condition*

$$9 > \frac{r_1 r_2}{c_{12} c_{21}} \tag{30}$$

holds, then there exists a single nontrivial equilibrium point E_3 in the positive cone. If condition (30) is an equality, $c_2^{\pm}$ collide in a single straight line.

Proof The equation $R_2(0) = 0$ is a second order equation in c_{22} of the form $-a(a_2 c_{22}^2 + a_1 c_{22} + a_0) = 0$. Solving it and rearranging terms we obtain expression (29). Note that condition (28), which, in particular, entails (27), ensures that the straight lines (29) are well defined, in the sense that the slopes are not complex but real numbers.

Note that a_0 and a_2 are positive; therefore, we need

$$a_1 = 27 c_{12}^2 c_{21}^2 - 18 c_{12} c_{21} r_1 r_2 - c_{12}^2 c_{21}^2$$

to be negative so that by solving $R_2(0) = 0$ in terms of c_{22} two positive roots are obtained. Rearranging terms, a_1 is equivalent to

$$a_1 = -(27 c_{12} c_{21} (c_{12} c_{21} - r_1 r_2) + r_1 r_2 (9 c_{12} c_{21} - r_1 r_2))$$

therefore, conditions (27) and (28) imply that $a_1 > 0$, which proves 1.(a).

Let us recall that conditions (27) and (28) imply that $Seq(0) = \{-, +, -, ?\}$ and $Seq(r_1/c_{11}) = \{+, +, ?, ?\}$, so that the only way of having three equilibrium points is that $R_2(0) > 0$ and $R_1(c_{11}/r_1) > 0$. Direct calculations show that the latter condition follows from (28) while the former holds for values of c_{11} and c_{22} that are between the straight lines $c_{22} = c_2^{\pm}(c_{11})$. This completes the proof of 1.(b).

The remaining statements follow in a similar way and the details are omitted. ■

In addition, solving the equations for c_{ij} (respectively c_i) yield conditions for exclusion or conditional tristability that depend on the so-called on the competition strength of each species (respectively the handling time). Recall that system (8) has no equilibrium points in the positive cone, independently of the parameter c_{ij} or c_i, in view of theorem 5.1.

Theorem 5.3 *Consider system* (8) *and assume that condition* (27) *holds. Then,*

1. *Assume further that condition* (28) *is fulfilled and*

$$\frac{2}{9}\frac{c_{22}r_1^2}{c_{11}r_2} < c_{12} < \frac{1}{4}\frac{c_{22}r_1^2}{c_{11}r_2} := \bar{c}_{12}, \tag{31}$$

Consider also the curves

$$c_{21}^{\pm}(c_{12}) = \frac{1}{27}\frac{c_{22}r_1(9c_{11}r_2c_{12} - 2c_{22}r_1^2) \pm 2\sqrt{c_{22}(c_{22}r_1^2 - 3c_{11}r_2c_{12})^3}}{c_{11}c_{22}c_{12}^2} \tag{32}$$

in the $c_{12}-c_{21}$ parameter space, obtained by setting $R_2(0)$ to zero, see expression (25)*. These curves define a region in the positive cone, shown in Fig. 5. Then,*

   (a) *If (c_{12}, c_{21}) lies in between $c_{21} = c_{21}^{\pm}(c_{12})$ then there exist three nontrivial equilibrium points $E_{\pm}^c, E_3$ in the positive cone, ordered by their first component, $E_-^c < E_3^c < E_+^c$.*
   (b) *If (c_{12}, c_{21}) does not lie in between $c_{21} = c_{21}^{\pm}(c_{12})$ then there exists a single nontrivial equilibrium point E_3 in the positive cone.*

2. *Besides, if condition* (30) *holds, then there exists a single nontrivial equilibrium point E_3 in the positive cone. If condition $c_{12} = \frac{1}{3}\frac{c_{22}r_1^2}{c_{11}r_2}$, $c_{21}^{\pm}$ coalesce into a single curve. The outcome is also a single positive coexistence equilibrium point.*

Proof Let us consider $R_2(0) = 0$, again it is a second order equation of the form $-b(b_2c_{21}^2 + b_1c_{21} + b_0) = 0$. Solving such an equation in c_{21} and arranging terms yields expression (32). Note that the rightmost condition (31) ensures that the curves (32) are well defined, i.e. the expression under the square root is positive. Coefficient b_2 is positive, so that equation $R_2(0) = 0$ possesses two positive roots if

$$b_1 < 0 \Leftrightarrow c_{12} > \frac{2c_{22}r_1^2}{9c_{11}r^2}$$

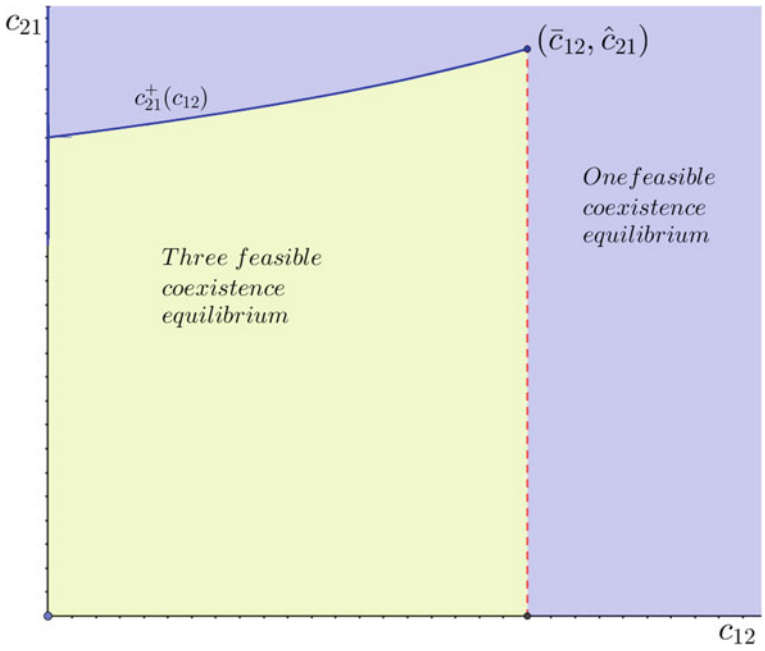


Fig. 5 Species competition outcomes as function of the intra-species competition parameter c_{ij}, $i, j = 1, 2$. The region in green delimited by $c_{21} = c_{21}^{+}(c_{12})$, the axes and $c_{12} = \bar{c}_{12}$ stands for the region in the parameter space where 3 positive (non trivial) equilibrium points arise. In the outer region (in purple) just one coexistence equilibrium is possible. The figure has been generated with the parameter values $c_{12} = 0.5$, $c_{21} = 0.07$, $r_i = k_i = 1$, $i = 1, 2$

and

$$b_0 > 0 \Leftrightarrow c_{12} < \frac{c_{22} r_1^2}{4 c_{11} r^2}.$$

Summing up, c_{12} fulfilling (31) ensures i) that $c_{21}^{\pm}$ are well defined and ii) $R_2(0)$ is positive if $c_{21}^{-}(c_{12}) < c_{21} < c_{21}^{+}(c_{12})$. Therefore, $Seq(0) = \{-, +, ?, +\}$ and $Seq(r_1/c_{11}) = \{+, +, +, +\}$, so that the only way of having three equilibrium points is that $R_1(0) < 0$, that is equivalent to:

$$c_{21} > \frac{r_1(7 c_{11} r_2 c_{12} - 2 c_{22} r_1^2)}{9 c_{11} c_{12}^2} := \tilde{c}_{21}(c_{12}).$$

Direct calculations show that $\tilde{c}_{21}(c_{12})$ crosses the c_{12} axis further away than $c_{21}^{-}(c_{12})$ does and that $\tilde{c}_{21}(c_{12}) < c_{21}^{-}(c_{12})$. Therefore, $c_{21}^{-}(c_{12}) < c_{21} < c_{21}^{+}(c_{12})$ implies $\tilde{c}_{21}(c_{12}) < c_{21}$, which completes the proof of 1.(b).

The remaining statement follows in the same way and is therefore omitted. ■

5.2 Stability

We gather in the following theorem the stability conditions of the equilibrium points found in the previous section.

Theorem 5.4 *Consider system* (8)*, assuming that* $r_1 > 0$ *and* $r_2 > 0$*. Then,*

1. *The origin is always unstable and the semi-trivial equilibrium points are always locally asymptotically stable.*
2. *Assume that condition* (26)*,* $c_{12}c_{21} - r_1 r_2 > 0$*, holds, so that there are no positive equilibrium points. Then there is a separarix through the origin (which is a saddle) delimiting the basins of attraction of the semi-trivial equilibrium points. Thus, one of the species goes extinct depending on the system initial conditions.*
3. *Assume now that conditions* (30) *hold. Then there exists a single coexistence equilibrium point* E_3^c*, that is unstable. There exists a separatrix line passing through* E_3^c *and the origin that defines the basins of attraction of* E_1 *and* E_2*.*
4. *Finally, if conditions* (28) *and* (31) *hold, there exist three nontrivial equilibrium points:* E_3^c *is locally asymptotically stable while* $E_\pm^c$ *are unstable. There exist separatrices passing through the origin and each of* $E_\pm^c$ *that define a region were* E_3^c *is located and define the basis of attraction of* E_1*,* E_2 *and* E_3^c*.*

Proof

1. Keeping in mind remark 3.2, consider the nullclines of system (8), given by

$$\begin{aligned} f_1 : x_2 &= \frac{1}{a_{12}^2}\left[r_1\sqrt{x_1} - b_1 a_{11} x_1\right]^2, \\ f_2 : x_1 &= \frac{1}{a_{21}^2}\left[r_2\sqrt{x_2} - b_2 a_{22} x_2\right]^2, \end{aligned} \tag{33}$$

   Condition (26) imply that the nullclines (33) divide the positive cone in three different regions, as displayed in Fig. 6. Region I is characterized for $x_1' < 0$ and $x_2' > 0$, and the flow points inward on the boundary of region I, so that it is a trapping region and solutions starting within the closure of region I are attracted by E_2, so that the origin is unstable. Proceeding as before, we find that region III is a trapping region and any solution starting within the closure of region III is attracted by E_1. It is immediate that E_1 and E_2 are locally asymptotically stable just by considering the flow of the system on the straight lines given by $\left\{(z_1, v_{2z_2}),\ z_1 \geq v_{2z_1}\right\}$ and $\left\{(v_{1z_1}, z_2),\ z_2 \geq v_{1z_2}\right\}$ respectively, where (v_{1z_k}, v_{2z_k}) are the coordinates of the vertex of the nullcline of z_k, $k = 1, 2$.
2. Consider a perfectly symmetric competition, i.e. $c_{12} = c_{21}$, $c_{11} = c_{22}$ and $r_1 = r_2$. Then, from the geometry of the phase portrait, the straight line $z_2 = z_1$ is the *stable manifold*, i.e. a forward invariant curve such that solutions starting on this line converge to the origin. Besides, it divides the positive cone on the basins of attraction of the semi-trivial equilibrium points. Furthermore, because of the continuity of the flow with respect to the system parameter and the uniqueness

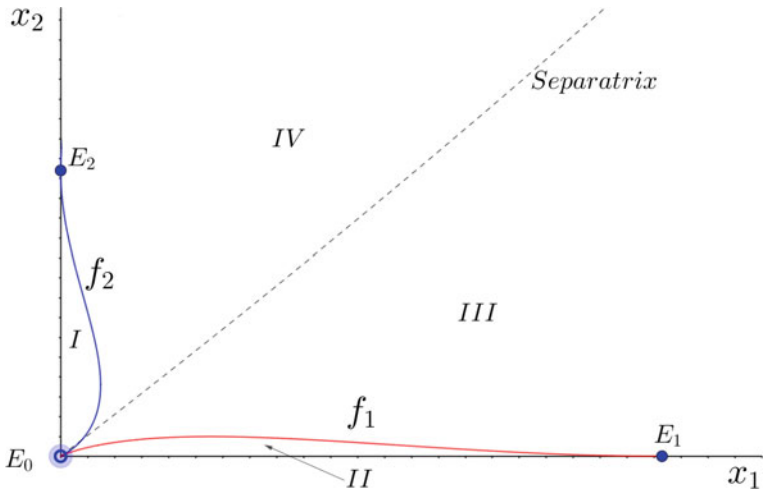


Fig. 6 The phase portrait for the competing populations. The figure is generated with the parameter values $a_{12} = 0.84$, $a_{21} = 1.10$, $a_{11} = 0.99$, $a_{22} = 0.75$, $r_i = 1$, $i = 1, 2$

of solutions on the positive cone, this invariant manifold changes continuously as the model parameters change as long as no positive equilibrium points appear in the positive cone.

3. This proof is an adaptation of the one of Theorem 4.2, item 2.a in [8]. Consider, as before, the perfectly symmetric case. Then the geometry of the phase plane ensures that E_3 is a saddle-node and that the straight line $z_2 = z_1$ is its stable manifold that defines the basins of attraction of the semi-trivial equilibrium points. As argued before, this setting remains qualitatively the same as the coefficients of the system vary continuously whenever no additional equilibrium points show up in the positive cone.
4. This statement follows *mutatis mutandi* from the previous one (see also the proof of Theorem 4.2, statement 2.a in [8]). ■

6 Discussion and Conclusions

We now interpret the mathematical results from a biological point of view, comparing the new model with the classical one. A first and important remark concerns both intra- and interspecific competition coefficients. They are expressed in the same units, $1/(time \times individuals)$ for intra- and interspecific competition, respectively, in the classical model. However the units in the sessile populations model are: $1/(time \times individuals)$ and $1/time$. If populations are counted as number of individuals, there is no difference, as all these coefficients become frequencies. However, if the populations represent animals or plants, for instance, a possibility

is to count them using biomass. In such case there is a difference in the units. But even if units are different, nullclines can be qualitatively compared, as well as the structure of each model possible outcome.

The main results are the following.

1. When a mobile and a sessile population compete, the dynamical outcomes are those of the classical model plus a bi-stable conditional coexistence.
2. Instead, when two sessile populations compete, the competition outcomes are reduced to competitive exclusion due to the system's initial conditions and tri-stable conditional coexistence.

Thus, sessility precludes global outcomes and puts the focus not only on the competitive abilities of each species, but also on the initial amount of individuals. Namely, there is neither global coexistence, for which species would coexist regardless of the initial amount of individuals of each one of them, nor global extinction where a "super-competitor" would rule out the other species independently of the initial composition of the community.

This is an interesting feature from the management point of view, since perturbations, whether human-driven or not, may drastically change the system outcome.

We next analyze in deeper detail each one of these features.

6.1 Sessile vs Mobile Species Competition

Theorem 4.2 concerns a mobile population 1 competing with a sessile population 2. It follows from this Theorem that all the dynamical outcomes of the classical model (5) are allowed and, in addition, bi-stable conditional coexistence in favor of the sessile population arises, except for a bi-stable conditional coexistence region that reduces the range in which the sessile species outcompetes the mobile one. The last dynamical outcome has already been observed in [8, 9] when modifying the classical competition model with Holling type II and IV competitive responses. Also in [1, 22] when considering social herd-induced behavior in one of the competing species.

The structure of the regions in the parameter space $\widehat{c}_{12} - c_{21}$ leading to each competitive outcome, see Fig. 2, is equivalent to the one of the classical model except for the bi-stable conditional coexistence in favor of the sessile species region. This region is set in the region where the classical model predicts that the mobile species would be excluded by the sessile species, and borders the species exclusion due to the system's initial conditions and (partially) the global coexistence regions, see Fig. 1. Therefore, in such a region the sessile species 2 cannot be eliminated by the mobile species 1. On the contrary, the mobile species has the chance of surviving via coexistence provided that the initial values are appropriate, see the bottom panel of Fig. 2. Thus, coexistence is more likely to occur than in the classical model.

Interestingly, consider interspecific competition coefficients $(\widehat{c}_{12}, c_{21})$ values belonging to the region where the sessile species 2 wins. Fixing $\widehat{c}_{12}$, as the effect of

the mobile species 1 on the sessile species 2, we find that $c_{21} < \phi_{b_2}(\widehat{c}_{12})$ increases and crosses the curve $c_{21} = \phi_{b_2}(\widehat{c}_{12})$, the mobile species 1 has the chance of surviving via coexistence with the sessile species. A further increment such that $c_{21} > 1$ makes the effect of the mobile species 1 on the sessile species 2 strong enough so that in the end only one species survives.

We summarize the results as follows:

- From a certain point of view, being sessile is not a serious handicap for a sessile species that faces competition with a mobile population. This means that the sessile population will survive for the same parameter value ranges as for the classical model, although for the parameters in the dark-blue region in Fig. 2 the sessile species will share the environment with the mobile population.
- In contrast, the mobile population is more likely to survive when facing a sessile population than a mobile one. This fact is reflected by the above-mentioned dark blue region in Fig. 1 that corresponds to the dynamical scenario depicted in the bottom panel of Fig. 2. Note that neither such a region nor nullclines configurations do exist in the classical model, when both populations are mobile.

6.2 *The Intra- and Interspecific Competition Effect for Nullclines*

We first analyze qualitatively the nullclines of system (8) versus the nullclines of the classical competition model (5). Let us note an important fact concerning both intra- and interspecific competition coefficients. These coefficients are somewhat non-comparable since they are expressed in different units, in the classical/sessile population models. Therefore, even if we plot both nullclines on the same axes, as in Fig. 7, we do not know how these parameters vary together. Keeping this in mind, we focus on the x_1 nullclines given by

$$x_2 = f_1(x_1) = \frac{1}{a_{12}}(\widehat{r}_1 - a_{11}x_1) \qquad x_2 = \widehat{f}_1(x_1) = \frac{x_1}{d_1^2\widehat{a}_{12}^2}(\widehat{r}_1 - b_1\widehat{a}_{11}\sqrt{x_1})^2 \tag{34}$$

As mentioned earlier the nullcline of species 1 defines the values of population size of species 2 that allow species 1 to thrive. For instance, the f_1 nullcline of the classical model is a straight line with negative slope, see Fig. 7, blue dotted straight line, which means that the larger x_1 is, the less tolerant to the presence of x_2 it is. In other words, it means that x_1 continues growing only if x_2 decreases.

Figure 7 represents the possible relative positions of the x_1 nullcline in the classical and sessile populations models.

A first claim is that at low x_1 population size the sessile model is highly tolerant to an increase of the number of individuals for both species 1 and species 2. This feature, which is at odds with the classical model, can be explained by

considering the peculiarities of the sessile model. Indeed, interactions take place at the boundary of either the vital surroundings of each individual (intraspecific case) or the region occupied by each population (interspecific case). On one hand, at low population size intraspecific competition does not play a major role as the population grows, since individuals only interact with the nearby ones and there are only a few of them. At the same time, interspecific interactions take place only at the boundary of the area occupied by species 1. Thus, population growth is the main driver of population dynamics at low densities. This trend is maintained at low densities, while $\widehat{f}_1$ is increasing. Direct calculations yield that the maximum of

$$\widehat{x}_2 := \widehat{f}_1(\widehat{x}_1) = \frac{1}{4^2}\frac{r_1^4}{b_1^2 a_{11}^2 a_{12}^2} \text{ is reached at } \widehat{x}_1 := \frac{1}{4}\left(\frac{r_1}{b_1 a_{11}}\right)^2.$$

A second claim is that from $\widehat{x}_1$ onwards the nullcline decreases, meaning that if x_1 increases slightly, so that intraspecific competition pressure increases, species 1 can keep growing only if species 2 reduces the interspecific competition pressure. That is, the trend is similar to that of the classical model,

A third feature is that nullclines $\widehat{f}_1$ and f_1 may not cross the horizontal axis at the same point. Indeed, the crossing points can be ordered in any way, as shown in Fig. 7.

Finally, note that both nullclines can meet essentially in four different ways, as shown in Fig. 7. Let us give an interpretation for instance to panel (A1). In the region

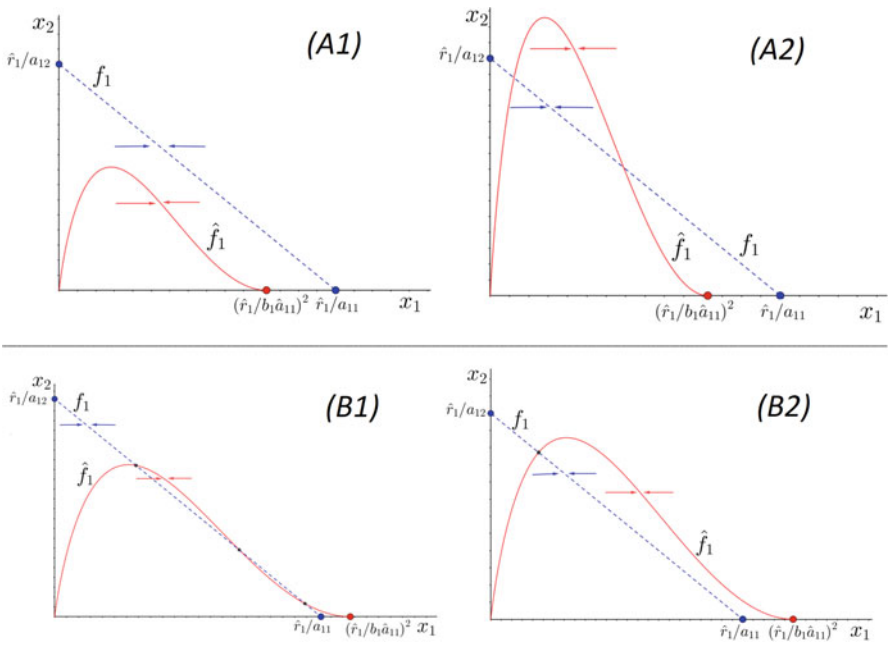


Fig. 7 The nullclines $x_1' = 0$ of systems (8) and classical model (5)

below $\widehat{f_1}$ (above f_1) x_1 increases (decreases) in both models. On the contrary, in the region between $\widehat{f_1}$ and f_1 species 1 would keep growing if represented by the classical model but would decrease if represented by the sessile populations model.

Note that the maximum of $\widehat{f_1}$ can be below (panel (A1)) or above (panel (A2)) f_1.

6.3 The Intra- and Interspecific Interaction and Coexistence Equilibria

Finally, we examine competition outcomes taking into account intra- and interspecific competition, that yield conditions for species exclusion or conditional tri-stability due to the system's initial conditions.

We may let either the intraspecific competition parameters c_{ii} vary for fixed values of the interspecific competition parameters c_{ij} or the other way around. Note that coefficient c_{ii} is a conglomerate of different factors that include intraspecific interaction a_{ii} and b_i, the relation between the perimeter of both the "vital space" around each individual and the perimeter of the area occupied by the whole population.

6.4 Varying Intra- Specific Competition Coefficients for Fixed Values of Interspecific Competition Coefficients

Theorem 5.2 indicates that under conditions (27) and (28) two possible outcomes are possible: either both species can coexist via tri-stable conditional coexistence or one of them goes extinct via the system's initial conditions.

In particular, with (c_{11}, c_{22}) lying between $c_{22} = c_{22}^{\pm}(c_{11})$, namely the straight lines defined in (29), there is tri-stable coexistence. Thus, consider fixed values of c_{ij} and r_i fulfilling the hypotheses of Theorem 5.2. Then, for each fixed value c_2^*:

- There is species coexistence via conditional tri-stability for $c_{11} \in (c_{11}^-, c_{11}^+)$ where $c_{22}^* = c_{22}^+(c_{11}^-)$ and $c_{22}^* = c_{22}^-(c_{11}^+)$.
- There is one species exclusion due to the system's initial conditions if $c_{11} \notin (c_{11}^-, c_{11}^+)$.

Everything works symmetrically if c_{11} is fixed and c_{22} varies. We may conclude that for each value of c_{ii} coexistence is possible for moderate values of $c_{jj}, i \neq j$. However, either low or large enough values of c_{ii} or c_{jj} would break coexistence.

Consistently with the classical model, numerical experiments show that the basin of attraction of E_1 is larger than the basin of attraction of E_2 for $c_{11} < c_{11}^-$. The result is reversed as $c_{11} > c_{11}^+$. This feature strongly suggests that the trade-

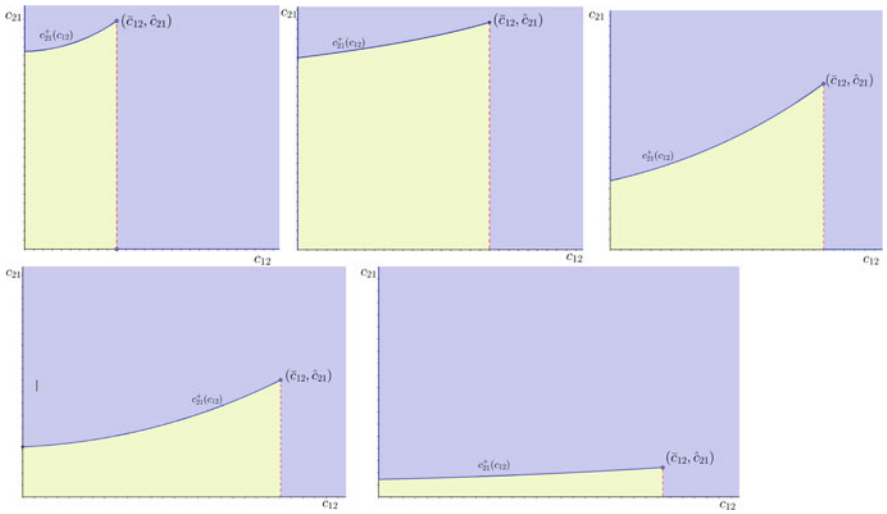


Fig. 8 Competition outcomes of system (8) as function of the interspecific competitive interaction c_{12}, c_{21} for a fixed value of c_{11} and increasing values of c_{22}. The figure is based analytic the expression or numerical calculations depending on each respective case and has been edited to improve it. Numerically fixed values of the parameters: $r_1 = 7.5$, $r_2 = 8$ $c_{11} = 0.9$ and of $c_{22} = 0.3, 1.1, 2.75, 3.3,$ and 16.7

off between intra- and interspecific competition forces works differently in sessile populations and in mobile populations.

6.5 *Inter-Specific Competition*

We fix now $c_{ii} > 0$ for $i = 1, 2$ and assume that condition (31) in Theorem 5.3 holds. Then, in a similar way as in the previous paragraph, there are conditions on the model coefficients that lead either to coexistence via conditional tri-stability or to species exclusion due to the system's initial conditions . In contrast, the curves defining the combination of interspecific competition coefficients leading to each outcome are nor straight lines and define a closed region on the positive cone, see Fig. 8. Inside such a region there is conditional coexistence while outside the competitive exclusion principle rules the system's outcome. Note that the shape of such a region changes as the other parameters change (in this case c_{22} changes as shown in the caption of Fig. 8). Note that the tri-stability region starts dropping towards the axis c_{12} as c_{22} increases.

Fixing c_{ii} and varying c_{ij}, this viewpoint is consistent with the classical model, meaning that a pair (c_{12}, c_{21}) close to the horizontal c_{12}-axis, i.e., $c_{12} > c_{21}$, makes the basin of attraction of E_2 larger than the one of E_1 and conversely.

Similarly, when we fix c_{ii}, moderate values of c_{ij} seem to promote species coexistence.

Acknowledgments The model was set up during a visit of M. Marvá in Torino; he kindly acknowledges the hospitality of E. Venturino. The paper was written during a visit of E. Venturino in Alcalá de Henares. He kindly acknowledges the support of the "Giner de los Rios" program of the University of Alcalá de Henares. H. Castillo-Alvino and M. Marvá are supported under the projects of the Ministerio de Economía y Competitividad (Spain) project MTM2014-56022-C2-1-P and the Ministerio de Ciencia Innovación y Universidades (Spain) project RTI2018-096884-B-C32. H. Castillo-Alvino is supported by a grant of the Pontificia Universidad Católica "Madre y Maestra" (Dominican Republic) under program "Programa de Formación Disciplinar de Docentes". Ezio Venturino is a member of the GNCS research group of INdAM. The E. Venturino's work has also been partially supported by the project "Modelli e metodi numerici in approssimazione, nelle scienze applicate e nelle scienze della vita" of the Dipartimento di Matematica "Giuseppe Peano".

References

1. Ajraldi V., Pittavino M., and Venturino E., Modeling herd behavior in population systems, *Nonlinear Real-Anal.* 12(4), 2319-2338 (2011).
2. Assis R.A., Pazim R., Malavazi M.C., Petry P.P. da C., Assis L.M.E. and Venturino E. A Mathematical Model to describe the herd behaviour considering group defense, *Applied Mathematics and Nonlinear Sciences*, 5(1), 11-24 (2020).
3. Banerjee M., Kooi B. W., Venturino E., An ecoepidemic model with prey herd behavior and predator-feeding saturation response on both healthy and diseased prey, *Mathematical Models in Natural Phenomena*, 12 (2), 133-161 (2017).
4. Begon M., Bennett M., Bowers RG., French NP., Hazel SM., Turner J. A clarification of transmission terms in host-microparasite models: numbers, densities and area, *Epidemiol Infect*, 129(1) 147-53 (2002).
5. Benayahu Y., Loya Y., Competition for Space among Coral-Reef Sessile Organisms at Eilat, Red Sea, *Bulletin of Marine Science* -Miami, 31(3), 514-522 (1981).
6. Bowers R.G., Boots A.W.M., Geritz S.A.H., Kisdi E., Evolutionary branching/speciation: contrasting results from systems with explicit or emergent carrying capacities, *Evol. Ecol. Res.*, 5, pp. 1-9 (2003).
7. Bulai I. M., Venturino E., Shape effects on herd behavior in ecological interacting population models, *Mathematics and Computers in Simulation*, 141 pp. 40-55 (2017).
8. Castillo-Alvino H., Marvá M., The competition model with Holling type II competitive response to interference time, *Journal of Biological Dynamics*, (2020)
9. Castillo-Alvino H., Marvá M., Group defense promotes coexistence in interference competition: The Holling type IV competitive response, *Mathematics and Computers in Simulation*, 198, 426–445 (2022). https://doi.org/10.1016/j.matcom.2022.02.031
10. Connell JH., The influence of interspecific competition and other factors on the distribution of the barnacle Chthamalus stellatus, *Ecology*, 42, 710-723 (1961).
11. Gause G.F., The Struggle For Existence, *Dover books on Biology* Baltimore, Hafner, New York 1934.
12. González-Rivero M., Yakob L., Mumby P. J., The role of sponge competition on coral reef alternative steady states, *Ecological Modelling*, 222(11) 1847-1853 (2011).
13. M. W. Hirsch, Systems of differential equations which are competitive or cooperative: I. Limit sets *SIAM Journal on Mathematical Analysis*, 13(2), 167-179 (1982)

14. Horwitz R., Hoogenboom M., Fine M., Spatial competition dynamics between reef corals under ocean acidification, *Sci Rep* 7, 40288 (2017)
15. Hoyle A., Bowers R.G., When is evolutionary branching in predator-prey systems possible with an explicit carrying capacity?, *Math. Biosci.*, 210, 1-16 (2007).
16. E. Kuno., Some strange propoerties of the logistic equation defined with r and K: inherent defect or artifacts?, *Res. Popul. Ecol.* 33 33-39 (1991).
17. Henri Laurie, Ezio Venturino, A two-predator one-prey model of population dynamics influenced by herd behaviour of the prey, *Theoretical Biology Forum* 111(1-2) 27-47 (2018).
18. Laurie H., Venturino E., Bulai I.M., Herding induced by encounter rate, with predator pressure influencing prey response, submitted, in "Current Trends in Dynamical Systems in Biology and Natural Sciences", Maira Aguiar, Carlos Braumann, Bob Kooi, Andrea Pugliese, Nico Stollenwerk, Ezio Venturino (Editors) Springer-SIMAI series, 2019, 63-92, 2020.
19. Lotka A.J., Elements of physical biology, *Williams and Wilkins*, Baltimore, 1925.
20. T. Lu, P.E. Saikaly, D.B. Oerther, Modelling the competition of planktonic and sessile aerobic heterotrophs for complementary nutrients in biofilm reactor, *Water Science and Technology*, 55(8-9) 227-235 (2007).
21. McCallum H., Barlow N., Hone J., How should pathogen transmission be modelled? *Trends Ecol Evol.* 16(6) 295-300 (2001).
22. Melchionda D., Pastacaldi E, Perri C., Venturino E., Social behavior-induced multistability in minimal competitive ecosystems, *J. Theor. Biol.*, 439, 24-38 (2018).
23. Murray J.G., Mathematical Biology I. An Introduction, Third Edition *Springer-Verlag* Berlin Heidelberg, 2001, pp. 94-98.
24. Ritchie M., Competition and coexistence of mobile animals, in *Competition and Coexistence*, U. Sommer and B. Worm, eds., Springer-Verlag, Berlin, 2002, pp. 109-131.
25. Shearer C.A., Fungal competition, *Canadian Journal of Botany*, 73(S1), 1259-1264 (1995).
26. Sieber M., Malchow H., Hilker F.M., Disease-induced modification of prey competition in eco-epidemiological models, *Ecological complexity*, 18, 74-82 (2014).
27. Stanley S.M., Effects of Competition on Rates of Evolution, With Special Reference to Bivalve Mollusks and Mammals, *Systematic Zoology*, 22(4), 486-506 (1973).
28. E. Venturino, S. Petrovskii, Spatiotemporal behavior of a prey-predator system with a group defense for prey, *Ecological Complexity*, 14, 37-47 (2013).
29. Volterra V., Variazioni e fluttuazioni del numero d'individui in specie animali conviventi, *Mem. R. Accad. Naz. dei Lincei*, 2, 31-113 (1926).
30. Wulff J. L., Ecological interactions of marine sponges, *Canadian Journal of Zoology*, 84(2), 146-166 (2006),
31. Zhang. Z., Mututalism or cooperation among competitors promotes coexistence and competitive ability, *Ecol. Model.*, 164, 271-282 (2003).

Mathematic Analysis of a $SIHV$ COVID-19 Pandemic Model Taking Into Account a Vaccination Strategy

Omar Khyar, Adil Meskaf, and Karam Allali

1 Introduction

Global health as well as the science of epidemiology are currently experiencing the greatest challenge in history. The pandemic caused by COVID-19, severe acute respiratory syndrome-related coronavirus SARS-COV-2, this disease which appeared in Wuhan, China, in December 2019, belongs to a large family of viruses that can cause various diseases in humans, ranging from the common cold to respiratory syndrome (MERS) and severe acute respiratory syndrome (SARS). All nations have entered a fierce race to find an effective cure or vaccine to curb the death rate among the populations which is growing day by day, as well as to bring a glimmer of optimism after the state of horror and panic that humanity has experienced and also save States from an unprecedented economic crisis after the total shutdown and confinement of the people that this mysterious COVID-19 has pushed the authorities to establish in the world whole.

Expected as the savior of humanity, many specialists consider the vaccine as our only means of regaining the "life before", but many questions arise about its effectiveness and its duration of protection whether in humans or population. All the epidemics that mankind has known have been eradicated by effective vaccines, it reveals the truth that a good vaccination strategy saves millions of lives around the world. The work of a vaccine is to excite and then prepare the immune system

O. Khyar (✉) · K. Allali
Laboratory of Mathematics and Applications, Faculty of Sciences and Techniques Mohammedia University Hassan II-Casablanca, Mohammedia, Morocco

A. Meskaf
Department of SEG, Faculty Polydisciplinary of El Jadida, University Chouaib Doukkali, El Jadida, Morocco

R. P. Mondaini (ed.), *Trends in Biomathematics: Stability and Oscillations in Environmental, Social, and Biological Models*,
https://doi.org/10.1007/978-3-031-12515-7_11

to recognize viruses, then, if the body is exposed to these same pathogens, it is immediately ready to destroy them, which helps prevent against the disease.

To fight COVID-19, it has been reported that more than 409.6 million doses of the vaccine have been administered worldwide. There is already a large gap between the vaccination programs of different countries, because of this and due to insufficient vaccine doses around the world, most countries have focused their early immunization efforts on priority groups such as people. Clinically vulnerable people in their sixties, sixties and over, with a few exceptions such as doctors, nurses, police and gendarmes as well as teachers and professors. There are ten different vaccines so far. The Pfizer-BioNTech vaccine, which has been shown to be 95 percent effective in reducing coronavirus infections, is currently in use in 77 countries. Most of the vaccines currently in use require two doses for a patient to be fully immunized, as recently the United States authorized a single dose vaccine called Johnson and Johnson—this data is compiled from government sources by the University of Oxford.

Mathematical modeling contributes enormously to the development of epidemiological research allowing a better understanding the evolution of pandemics. To illustrate this, Many mathematical models have been successfully developed in order to describe the evolution of infectious diseases [1–8]. The first mathematical models developed to study the evolution of the COVID-19 pandemic were inspired by SIR models [9–11]. Then, mathematics researchers proposed a SIQR models with quarantine strategy [12–16]. Kucharski et al. [17] have studied a mathematical model considering all the positive COVID-19 infected cases of Wuhan. An autonomous system of mathematical model to study the spread of COVID-19 in the Wuhan city have presented by Ndairou et al. [18] . The controlling status of COVID-19 of Wuhan city have analyzed by Prem et al. [19]. Also, the effective procedure of COVID-19 disease using isolation have described by Hellewell et al. [20].

In the same epic, we proposed a mathematical model describing all the scenarios of evolution of the COVID-19 pandemic with a vaccination strategy, we estimate that the infected individuals, after their recovery, can become susceptible after. We define our model consisting of four ordinary differential equations illustrating the interaction between the susceptible S, the real infected I, the hospitalized infected H and the vaccinated-treated individuals V (Fig. 1).

$$\begin{cases} \dfrac{dS}{dt} = \Lambda - \beta(1-\rho)SI + \eta I + \theta H + \varepsilon V - (d+\gamma)S, \\ \dfrac{dI}{dt} = \beta(1-\rho)SI - (d+\delta_1+\eta+\alpha)I\ , \\ \dfrac{dH}{dt} = \alpha I - (d+\delta_2+\theta)H\ , \\ \dfrac{dV}{dt} = \gamma S - (d+\varepsilon)V\ , \end{cases} \tag{1}$$

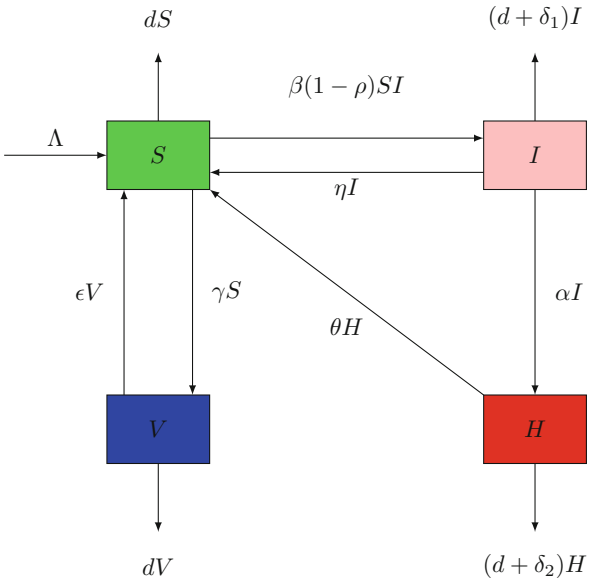


Fig. 1 The flowchart of $SIHV$ epidemic modeling of COVID-19

With

$$S(0) \geq 0,\ I(0) \geq 0,\ H(0) \geq 0,\ V(0) \geq 0.$$

Where Λ is the recruitment rate, β is the disease transmission rate, ρ is the portion of susceptible human would maintain proper precaution measure for disease transmission ($0 < \rho < 1$), ε is the rate of vaccinated individuals who became susceptible, η and θ are the recovery rates of real infected individuals and the confirmed infected, respectively, δ_1 is the COVID-19 induced death rate of real infected individuals, α is the rate of transmission from the class of real infected to the class of confirmed infected, δ_2 is the COVID-19 induced death rate of confirmed infected individuals, γ is the vaccinated susceptible individuals rate and d is the natural death rate of the population.

Our paper is organized as follows. In Sect. 2, we will study the local stability of our model. In the next section, we will prove the positivity and the boundedness results. After, we give the two equilibrium points and calculate the basic reproduction number of our COVID-19 epidemic model. Section 4 is devoted to illustrate our theoretical findings by numerical simulations, we will give also a comparison between the model results and COVID-19 clinical data. The last section concludes our work.

2 Positivity and Boundedness of Solutions

Since our problem is related to the population dynamics, we will prove that all model variables are positive and bounded. First, we will assume that all the parameters in our model are positive.

Proposition 2.1 *For any positive initial conditions $S(0)$, $I(0)$, $H(0)$ and $V(0)$, the variables of the model (1) $S(t)$, $I(t)$, $H(t)$ and $V(t)$ will remain positive for all $t > 0$.*

Proof We have the following results :

$$\dot{S}|_{S=0} = \lambda + \eta I + \theta H + \varepsilon V \geq 0,$$

$$\dot{I}|_{I=0} = 0 \geq 0,$$

$$\dot{H}|_{H=0} = \alpha I \geq 0,$$

et

$$\dot{V}|_{V=0} = \gamma S \geq 0,$$

this shows the positivity of solutions for all $t \geq 0$.
For the boundedness of the solutions, Let

$$N(t) = S(t) + I(t) + C(t) + V(t),$$

according to system (1), we have

$$\frac{dN(t)}{dt} = \Lambda - dN(t) - \delta_1 I(t) - \delta_2 C(t)) \leq d\left(\frac{\Lambda}{d} - N(t))\right),$$

then we have

$$\frac{d\left(\frac{\Lambda}{d} - N(t)\right)}{dt} + d\left(\frac{\Lambda}{d} - N(t)\right)dt \geq 0.$$

Therefore

$$\frac{\Lambda}{d} - N(t) \geq \left(\frac{\Lambda}{d} - N(0)\right)e^{-dt}$$

if $\frac{\Lambda}{d} - N(0) \geq 0$, i.e., $S(0) + I(0) + C(0) + V(0) \leq \frac{\Lambda}{d}$, then

$$S(t) + I(t) + C(t) + V(t) \leq \frac{\Lambda}{d}.$$

Thus the region

$$\Delta = \left\{ (S, I, C, V) \in \mathbb{R}^4_+ : S + I + C + V \leq \frac{\Lambda}{d} \right\}$$

is a positively invariant set of system (1).

3 Steady States and Local Stability

3.1 Basic Reproduction Number

The basic reproduction number denoted by R_0, is the average number of new infected cases generated by one infected individual when all the population are susceptible individuals [21]. In order to calculate the basic reproduction number, we will use the next generation matrix FV^{-1}, where F is the nonnegative matrix of new infection cases, and V is the matrix of the transition of infections associated to the model (1)

$$F = \begin{pmatrix} \beta(1-\rho)S_0 & 0 \\ 0 & 0 \end{pmatrix}, \qquad V = \begin{pmatrix} \delta_1 + d + \alpha + \eta & 0 \\ -\alpha & \delta_2 + d + \theta \end{pmatrix}.$$

So,

$$FV^{-1} = \begin{pmatrix} \frac{\beta(1-\rho)S_0}{\delta_1 + d + \alpha + \eta} & 0 \\ 0 & 0 \end{pmatrix},$$

with $S_0 = \frac{\Lambda(d+\epsilon)}{d(d+\epsilon+\gamma)}$. The basic reproduction number is the spectral radius of the matrix FV^{-1}. This fact implies that

$$R_0 = \frac{\Lambda\beta(1-\rho)(d+\epsilon)}{d(d+\epsilon+\gamma)(\delta_1 + d + \alpha + \eta)}.$$

3.2 Steady States

The steady states of our studied problem (1)are illustrated by the following theorem

Theorem 3.1 *The model (1) has a disease-free equilibrium E_f and an endemic equilibrium E_1.*

Proof To find the steady states of the system 1, we solve the following system

$$\begin{aligned}
\Lambda - \beta(1-\rho)\, SI + \varepsilon V + \theta H + \eta I - (d+\gamma)S &= 0,\\
\beta(1-\rho)\, SI - \delta_1 I - (d+\alpha+\eta)I &= 0,\\
\alpha I - \delta_2 H - (d+\theta)H &= 0,\\
\gamma S - (d+\varepsilon)V &= 0.
\end{aligned}$$

After a simple resolution, we obtain

- When $I = 0$ we find the disease-free equilibrium

$$E_f = \left(\frac{\Lambda(d+\epsilon)}{d(d+\epsilon+\gamma)}, 0, 0, \frac{\Lambda\gamma}{d(d+\epsilon+\gamma)}\right).$$

- When $I \neq 0$ we find the endemic equilibrium defined as follows $E_1 = \left(S^*, I^*, C^*, V^*\right)$,
  where

$$S^* = \frac{S_0}{R_0},$$

$$I^* = \frac{(R_0 - 1)}{\Gamma R_0},$$

$$H^* = \frac{\alpha(R_0 - 1)}{(\delta_2 + d + \theta)\Gamma R_0},$$

$$V^* = \frac{\gamma S_0}{(d+\epsilon)R_0}.$$

  Where $\Gamma = \dfrac{(d+\epsilon)(\delta_2+d+\theta)+\alpha(\delta_2+d)}{\Lambda(\delta_2+d+\theta)}$ and $S_0 = \dfrac{\Lambda(d+\epsilon)}{d(d+\epsilon+\gamma)}$.
  It's clear that E_1 is well defined when $R_0 > 1$.

3.3 Local Stability

3.3.1 Local Stability of the Disease-Free Equilibrium

The local stability of the disease-free equilibrium point $E_f = \left(\frac{\Lambda(d+\epsilon)}{d(d+\epsilon+\gamma)}, 0, 0, \frac{\Lambda\gamma}{d(d+\epsilon+\gamma)}\right)$, is given by the following result:

Prop 3.1 *When $R_0 < 1$, then the disease-free equilibrium point, E_0, is locally asymptotically stable.*

Proof The Jacobian matrix of the system (1) at E_0 is given by:

$$J_{E_f} = \begin{pmatrix} -(d+\gamma) & -\beta(1-\rho)S_0 + \eta & \theta & \epsilon \\ 0 & \beta(1-\rho)S_0 - (\delta_1 + d + \alpha + \eta) & 0 & 0 \\ 0 & \alpha & -(\delta_2 + d + \theta) & 0 \\ \gamma & 0 & 0 & -(d+\epsilon) \end{pmatrix}. \tag{2}$$

The characteristic polynomial of J_{E_0} is

$$P_{E_0}(\lambda) = [\lambda + d][\lambda + (d + \gamma + \epsilon)]\,[\lambda + (\delta_2 + d + \theta)]\,[\lambda + (\delta_1 + d + \alpha + \eta)(1 - R_0)].$$

Therefore, the eigenvalues of $J(E_0)$ are given as follow,

$$\begin{aligned} \lambda_1 &= -d < 0, \\ \lambda_2 &= -(d + \gamma + \epsilon) < 0, \\ \lambda_3 &= -(\delta_2 + d + \theta) < 0, \\ \lambda_4 &= -(\delta_1 + d + \alpha + \eta)(1 - R_0). \end{aligned}$$

clearly, λ_1, λ_2 and λ_3 are negative. However, λ_4 is negative when $R_0 < 1$.

Consequently E_f is locally asymptotically stable when $R_0 < 1$.

3.3.2 Local Stability of the Endemic Equilibrium

The local stability of the endemic equilibrium point $E_1 = \left(S^*, I^*, C^*, V^*\right)$, is given by the following result:

Prop 3.2 *When $R_0 > 1$ then the endemic equilibrium point E_1 is locally asymptotically stable.*

Proof The Jacobian matrix of the system (1) at E_1 is given by:

$$J_{E_1} = \begin{pmatrix} -\beta(1-\rho)I - (d+\gamma) & -\beta(1-\rho)S + \eta & \theta & \epsilon \\ -\beta(1-\rho)I & \beta(1-\rho)S - (\delta_1 + d + \alpha + \eta) & 0 & 0 \\ 0 & \alpha & -(\delta_2 + d + \theta) & 0 \\ \gamma & 0 & 0 & -(d+\epsilon) \end{pmatrix}. \tag{3}$$

The characteristic polynomial of J_{E_1} is

$$P_{E_1}(\lambda) = (d + \epsilon + \lambda)\left(\lambda^3 + A_1\lambda^2 + A_2\lambda + A_3\right)$$

such that:

$$A_1 = d + \delta_2 + \theta + \frac{\Lambda(d+\epsilon)}{S_0},$$

$$A_2 = \frac{\Lambda(d+\epsilon)}{S_0}(d + \delta_2 + \theta) + \frac{\beta(1-\rho)(d+\delta_1+\alpha)(d+\epsilon)I^*S_0}{R_0 S^*},$$

$$A_3 = \frac{\beta(1-\rho)I^*S_0}{R_0 S^*}\left(d + \delta_1 + \alpha + (d+\epsilon)(d+\delta_2+\theta) - \alpha\theta(d+\epsilon)\right).$$

The first eigenvalue of (3) is $\lambda_1 = -(d+\epsilon) < 0$, also it is easy to verify that $A_1 > 0$, $A_1A_2 - A_3 > 0$ and $A_3 > 0$ if $R_0 > 1$ then by using the Routh-Hurwitz Theorem, the other eigenvalues of (3) have negative real parts.

Consequently, E_1 is locally asymptotically stable when $R_0 > 1$.

4 Numerical Simulations

In this section, we will perform some numerical simulations in order to confirm our theoretical results and to check the impact of vaccination strategy in fighting against the spread of COVID-19 pandemic. Indeed, Fig. 2 shows the evolution of infection for $\Lambda = 1$, $\rho = 0.1$, $\eta = 0.1$, $\theta = 0.6$, $d = 0.1$,$\gamma = 1.2$, $\epsilon = 0.03$, $\beta = 1.3$, $\delta_1 = 0.7$, $\delta_2 = 0.6$ and $\alpha = 0.4$.

Figure 2 depicts the dynamics of all SIHV variables. In this figure, we observe that all curves drop to zero, except the curves representing the susceptible and vaccinated individuals. With the used parameters, the basic reproduction number is less than one ($R_0 = 0.8797 < 1$). Figure 3 shows the time evolution of our SICV four compartments model. With the used parameters, the basic reproduction number is more than one ($R_0 = 1.19721 > 1$). We observe that the trajectories representing real and confirmed infected individuals remain at a strictly positive level which means that the disease persists. Which is in good agreement with the

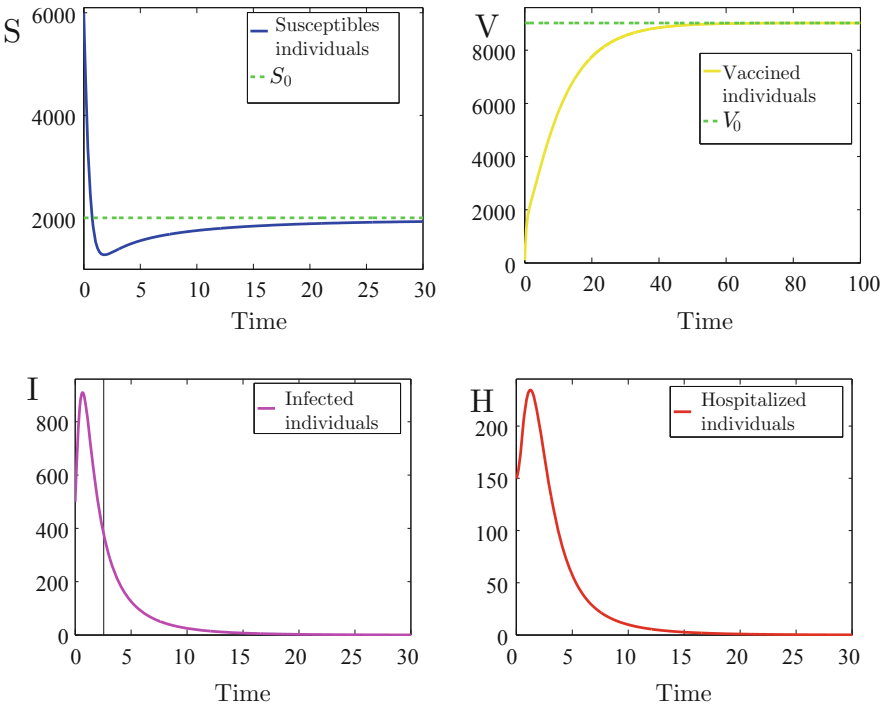


Fig. 2 The dynamical behavior of compartments S, I, H and V revealing the extinction of COVID-19 disease with $R_0 = 0.8797 < 1$

theoretical result concerning the stability of equilibria, the disease free and the endemic equilibrium points.

Figure 3 shows the evolution of infection for $\Lambda = 1$, $\rho = 0.1$, $\eta = 0.1$, $\theta = 0.6$, $d = 0.1$,$\gamma = 0.75$, $\epsilon = 0.01$, $\beta = 1.3$, $\delta_1 = 0.7$, $\delta_2 = 0.6$ and $\alpha = 0.35$.

4.1 Application to Morocco COVID-19 Clinical Data

We have chosen to make our comparison, the Moroccan clinical data during the period between September 12 and March 28 [22]. for the following parameter values: $\Lambda = 1$; $\rho = 0.1$; $a = 0.001$; $\eta = 0.1$; $\theta = 0.6$; $d = 0.1$; $\gamma = 0.6$; $\epsilon = 0.01$; $\beta = 1$; $\delta_1 = 0.7$; $\delta_2 = 0.6$; $\alpha = 0.2$. Figure 4 shows the time evolution of infected cases, we observe a significant good approach between the curves representing the model numerical results and the clinical data. Therefore, our model have shown its efficiency in approaching and predicting the second wave of COVID-19 pandemic.

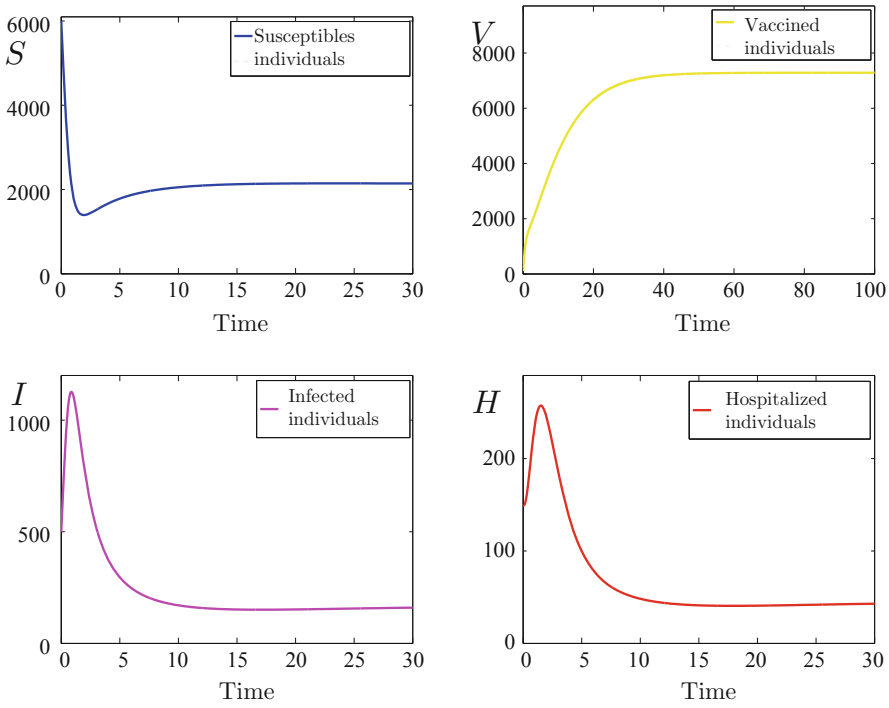


Fig. 3 The dynamical behavior of compartments S, I, H and V revealing the persistence of COVID-19 disease with $R_0 = 1.19721 > 1$

4.2 The Effect of the Vaccination Strategy on COVID-19 Pandemic Spread

In this subsection, we will study the role of the vaccination strategy in reducing the infection severity of COVID-19 pandemic. Figure 5 shows the time evolution of the real infected cases for the parameters $\Lambda = 0.95$, $\rho = 0$, $\eta = 0.1$, $\theta = 0.6$, $d = 0.1$, $\gamma = 0.1$, $\epsilon = 0.01$, $\beta = 2.6$, $\delta_1 = 0.2$, $d\delta_2 = 0.1$ and $\alpha = 0.3$. We observe the effect of vaccination strategy on reducing the spread of the COVID-19 infection. Indeed, by increasing the vaccination rate a significant reduce of the real infected individuals is observed which clearly reveals the impact of vaccination strategy in fighting against the spread of COVID-19 pandemic.

5 Discussion and Conclusion

Mathematical modeling contributes enormously in epidemiological research via both theoretical and numerical methods allowing a better understanding of the evo-

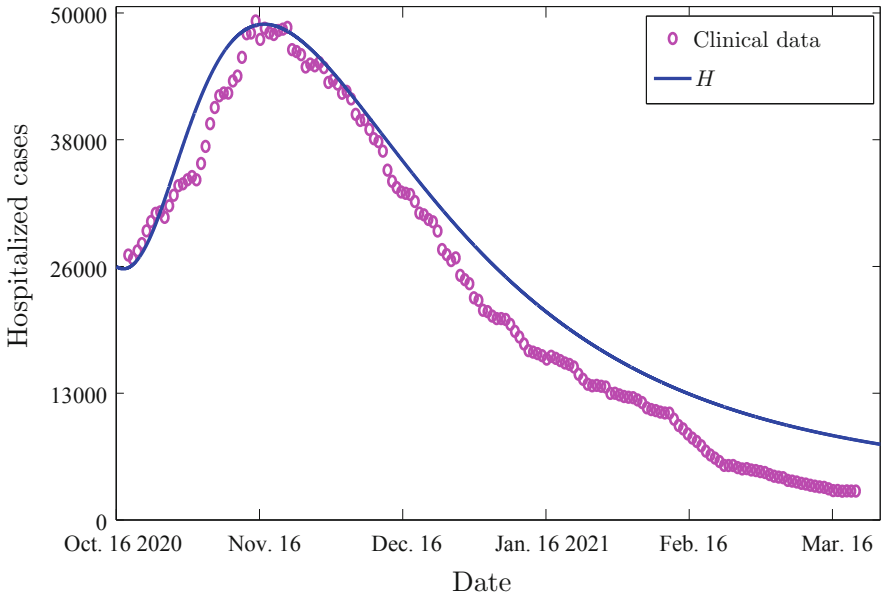


Fig. 4 The time evolution of infected cases trajectory (blue color). The clinical infected cases are illustrated by magenta circles

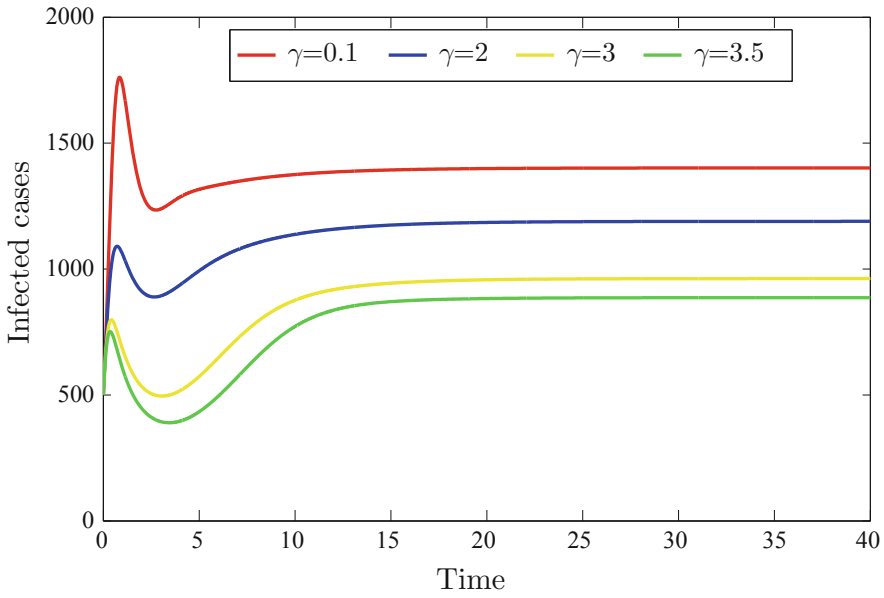


Fig. 5 The dynamical behavior of real infected cases for our model revealing the effect of vaccination rate on COVID-19 pandemic model

lution of the pandemic within populations and to unearth the interactions between the various factors responsible for the spread of infections between individuals, but also to provide conditions for the stability of the variables that cause the disease. Due to the rapid spread of COVID-19, scientific researchers are working day and night to find an ideal vaccine that eradicates this pandemic. Those efforts can bring the world back to a normal pre-COVID-19 normal life. The main objective is to find an adequate vaccination strategy to curb the rapid spread of the virus as well as to obtain collective immunity to prevent the appearance of new variants of COVID-19.

In this paper, we have studied a mathematical model describing the spread of the COVID-19 pandemic with a vaccination strategy. The model consisted of four compartments, namely, the susceptible S, the real infected I, the confirmed infected H and the vaccinated individuals V, this type of model takes the abbreviation $SIHV$. We have first studied the local stability of our model two state states by calculating the basic reproduction number of our COVID-19 epidemic model. Finally, we have confirmed our theoretical results by adequate numerical simulations. An interesting comparison was also made between the model theoretical results and the COVID-19 clinical data from Morocco between September 2020 and April 2021. It was also shown that vaccination strategy plays an essential role in controlling COVID-19 spread. We can conclude from our study that a good vaccination strategy leads to controlling COVID-19 in target populations.

References

1. Kermack, W.O., McKendrick, A.G.: A contribution to the mathematical theory of epidemics. Proc. R Soc. Lond. A. **115**, 700–721(1927)
2. Anderson, R., M. and R. M. May,(1999). Population biology of infectious disease I, Nature 180, 361–367.
3. Hethcote, H., W.,: The mathematics of infectious diseases, SIAM Rev. **42**, 599–653 (2000).
4. Xiao, D. M. and S. G. Ruan, (2007). Global analysis of an epidemic model with nonmonotone incidence rate, Math. Biosci. 208: 419–429.
5. Wang, J. J., Zhang, J. Z., Jin, Z.: Analysis of an SIR model with bilinear incidence rate. Nonlinear Anal. Real World Appl. **11**, 2390–2402 (2010)
6. Capasso, V., Serio, G.: A generalization of the kermack-mckendrick deterministic epidemic model. Math. Biosci. **42**, 43–61 (1978)
7. Nakata, Y., Kuniya, T.: Global dynamics of a class of SEIRS epidemic models in a periodic environment, J. Math. Anal. Appl. **363**, 230–237 (2010).
8. Liu, W. M., Levin, S. A., Iwasa, Y.: Influence of nonlinear incidence rates upon the behavior of sirs epidemiological models. J. Math. Biol. **23**, 187–204 (1986)
9. Ndaïrou, Faïçal, et al., Mathematical modeling of COVID-19 transmission dynamics with a case study of Wuhan. Chaos, Solitons and Fractals 135 (2020): https://doi.org/10.1016/j.chaos.2020.109846
10. Abdullah, Saeed Ahmad, et al., Mathematical analysis of COVID-19 via new mathematical model. Chaos, Solitons, and Fractals 143 (2021): 110585.
11. Khyar, Omar, and Karam Allali. Global dynamics of a multi-strain SEIR epidemic model with general incidence rates: application to COVID-19 pandemic. Nonlinear Dynamics 102.1 (2020): 489–509.

12. Upadhyay, Ranjit Kumar, et al. Age-group-targeted testing for COVID-19 as a new prevention strategy. Nonlinear Dynamics 101.3 : 1921–1932 (2020).
13. Malik, Arti, Khursheed Alam, and Nitendra Kumar. COEFFICIENT IDENTIFICATION IN SIQR MODEL OF INVERSE PROBLEM OF COVID-19. European Journal of Molecular and Clinical Medicine 7.09: (2020).
14. Odagaki, Takashi. Analysis of the outbreak of COVID-19 in Japan by SIQR model. Infectious Disease Modelling 5 691–698 (2020).
15. Crokidakis, Nuno. Modeling the early evolution of the COVID-19 in Brazil: Results from a SIQR model. International Journal of Modern Physics C (IJMPC) 31.10 (2020): 1–7.
16. Odagaki, Takashi. Exact Properties of SIQR model for COVID-19. Physica A, Statistical Mechanics and its Applications 564: 125564 (2021).
17. Kucharski AJ, Russell TW, Diamond C, Liu Y, Edmunds J, Funk S, et al. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. Lancet Infect Dis 2020
18. Ndariou F, Area I, Nieto JJ, Torres DF. Mathematical modeling of COVID-19 transmission dynamics with a case study of Wuhan. Chaos Solitons Fractals 2020. doi:10.1016/j.chaos.2020.109846.
19. Prem K, Liu Y, Russell TW, Kucharski AJ, Eggo RM, Davies N. The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan. China: a modelling study. The Lancet Public Health; 2020.
20. Hellewell J, Abbott S, Gimma A, Bosse NI, Jarvis CI, Russell TW, et al. Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. Lancet Global Health 2020;8:e488–96.
21. O. Diekmann, J. A. P. Heesterbeek, M. G. Roberts," The construction of next-generation matrices for compartmental epidemic models", Journal of the Royal Society Interface, 7(47), 873–885, 2010.
22. Statistics of Moroccan health ministry on COVID-19, https://www.sante.gov.ma/.

Derivation of Macroscopic Equations from Homogeneous Thermostatted Kinetic Equations in the Cancer-Immune System Competition

G. Morgado, L. Masurel, A. Lemarchand, and C. Bianca

1 Introduction

Biologists traditionally choose either the molecular scale or the macroscopic scale to describe the interactions between the immune system and cancer. They often introduce complex interaction networks between a large number of actors, including molecules, tissues, organs, the lymphatic and vascular systems, and involving feedback loops providing information on the inhibition or stimulation of an effect [1]. Harnessing the models developed by applied mathematicians and physicists, we choose to describe the competition between cancer cells and immune system cells at a mesoscopic scale [2–4]. This intermediate scale of description mobilizes tools of statistical physics, in particular kinetic theory [5]. In an analogous way to the Boltzmann equations associated with a diluted gas [5], the kinetic equations associated with a model of cellular interactions give the evolution of the distribution functions of the different cell types [6–8]. In order to reproduce the learning phenomena that concern both immune system cells and cancer cells [9, 10], we assign a quantity called activity [2] to each cell. Cell activity may grow during appropriate interactions. A cell with a high activity represents an educated cell. It should be noted that cell activity should not be confused with the notion of

G. Morgado (✉) · L. Masurel · A. Lemarchand
Sorbonne Université - CNRS, Laboratoire de Physique Théorique de la Matière Condensée, Paris Cedex 05, France
e-mail: gabriel.morgado@sorbonne-universite.fr

C. Bianca
École Supérieure d'Ingénieurs en Génie Électrique, Productique et Management Industriel Laboratoire Quartz, Cergy Pontoise Cedex, France
Laboratoire de Recherche en Eco-innovation Industrielle et Energétique, Cergy Pontoise Cedex, France

R. P. Mondaini (ed.), *Trends in Biomathematics: Stability and Oscillations in Environmental, Social, and Biological Models*,
https://doi.org/10.1007/978-3-031-12515-7_12

active matter, associated with the transduction of energy into oriented motion in far-from-equilibrium systems [11]. We have adapted the direct simulation Monte Carlo (DSMC) method [5, 8] to simulate the kinetic equations in an inhomogeneous system and shown that they include the description of clinically observed phenomena such as the 3Es of immunotherapy [12] for elimination, equilibrium, and escape of cancer, as well as pseudo-oscillations of the number and mean activity of immune cells [8].

In this work, we derive approximate macroscopic equations for the concentrations and mean activities of the immune and cancer cells from the kinetic equations for the distribution functions associated with a homogeneous system. Our aim is to determine if the homogeneous macroscopic equations are able to reproduce the 3Es and to address the question of possible oscillations. The paper is organized as follows. In Sect. 2, we present the model and the derivation of macroscopic equations from the kinetic equations. In Sect. 3, we discuss the results deduced from the macroscopic equations and compare them to DSMC simulations of the kinetic equations. Section 4 is devoted to the conclusion.

2 Model at Cell Scale

We consider a model involving three different types of cells, normal cells n, cancer cells c, and immune cells i [6–8]. Each cell has a given activity u which represents its level of adaptation towards the other cell types. The activity of an immune cell measures the efficiency with which it is able to kill a cancer cell. For a cancer cell, the activity corresponds to its level of aggressiveness towards normal cells and its ability to be invisible to immune cells. The activities of all cells are initially sampled from the Gaussian distribution

$$P_{\mathrm{R}}(u) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(u-\mu)^2}{2\sigma^2}\right) \tag{1}$$

of mean μ and variance σ^2. In vivo, cell future depends on complex biological phenomena that rule the birth and death processes of each cell population. In a purpose of modeling, we consider a minimal scheme of three processes involving autocatalytic interactions between the different cell types:

$$\begin{cases} c(u) + n(u') \xrightarrow{k_{cn}(u-u')H(u-u')} c(u+\epsilon) + c(u') \\ \qquad\qquad \mathrm{R} \qquad\qquad\qquad\quad n(u'') \end{cases} \tag{2}$$

$$i(u) + c(u') \xrightarrow{k_{ic}(u-u')H(u-u')} i(u+\epsilon) + i(u') \tag{3}$$

$$c(u) + i(u') \xrightarrow{k_{ci}(u-u')H(u-u')} c(u+\epsilon) + c(u') \tag{4}$$

where k_{cn}, k_{ic}, and k_{ci} are constant. According to the Heaviside step function, each process occurs only if the activity u of the cell which remains after the interaction is larger than the activity u' of the cell which changes of nature. The rate at which each process occurs is proportional to the difference of activities $(u - u')$ of the two interacting cells. The learning of the cancer cell or the immune cell whose nature does not change after the interaction is reproduced by the increase of its activity by ϵ, small compared to μ.

The first process includes the mutation of a normal cell into a cancer cell. When a normal cell n mutates into a cancer cell c, a new normal cell is injected by the reservoir R with an activity u'' sampled from the initial Gaussian distribution given in Eq. (1). The reservoir of normal cells ensures that the total number of normal cells remains constant. It also ensures that the mean activity of the normal cells remains almost constant, close to μ. The second process accounts for the elimination of a cancer cell by the immune system whereas the third process corresponds to the elimination of an immune cell by a cancer cell.

Similarly to a mechanical system maintained at constant temperature due to a thermostat, the system dissipates information according to a friction coefficient α and is submitted to a field E controlling the fluctuations of activity. Hence, the activity u of a cell is supposed to obey

$$\frac{\mathrm{d}u}{\mathrm{d}t} = E - \alpha u \tag{5}$$

with

$$\alpha = E\langle u \rangle / \langle u^2 \rangle \tag{6}$$

where $\langle . \rangle$ denotes an ensemble average. The expression of the friction α ensures that the second moment of the activity $\langle u^2 \rangle$ remains constant. The control of activity fluctuations reproduces the spreading of information through processes that are not explicitly considered in the model of interactions. In particular, the dissipation of information takes into account the death of all cell types [13, 14], which depletes the system of both highly and poorly educated cells.

2.1 Thermostatted Kinetic Framework

According to the thermostatted kinetic theory, the distribution function $f_j(t, u)$ of each cell type $j = c, i, n$ obeys the following kinetic equation [2, 8, 15]

$$\partial_t f_j(t, u) = -\partial_u((E - \alpha u) f_j(t, u)) + I_j \tag{7}$$

where I_j is the interaction operator. For the interaction scheme given in Eqs. (2–4), the three interaction operators are given by

$$
\begin{aligned}
I_c = & \int_{\mathbb{R}^+} k_{cn}(u-\epsilon-u')H(u-\epsilon-u')f_c(t,u-\epsilon)f_n(t,u')du' \\
& - \int_{\mathbb{R}^+} k_{cn}(u-u')H(u-u')f_c(t,u)f_n(t,u')du' \\
& + \int_{\mathbb{R}^+} k_{cn}(u'-u)H(u'-u)f_c(t,u')f_n(t,u)du' \\
& - \int_{\mathbb{R}^+} k_{ic}(u'-u)H(u'-u)f_i(t,u')f_n(t,u)du' \\
& + \int_{\mathbb{R}^+} k_{ci}(u-\epsilon-u')H(u-\epsilon-u')f_c(t,u-\epsilon)f_i(t,u')du' \\
& - \int_{\mathbb{R}^+} k_{ci}(u-u')H(u-u')f_c(t,u)f_i(t,u')du' \\
& + \int_{\mathbb{R}^+} k_{ci}(u'-u)H(u'-u)f_c(t,u')f_i(t,u)du'
\end{aligned} \tag{8}
$$

$$
\begin{aligned}
I_i = & \int_{\mathbb{R}^+} k_{ic}(u-\epsilon-u')H(u-\epsilon-u')f_i(t,u-\epsilon)f_c(t,u')du' \\
& - \int_{\mathbb{R}^+} k_{ic}(u-u')H(u-u')f_i(t,u)f_c(t,u')du' \\
& + \int_{\mathbb{R}^+} k_{ic}(u'-u)H(u'-u)f_c(t,u)f_i(t,u')du' \\
& - \int_{\mathbb{R}^+} k_{ci}(u'-u)H(u'-u)f_c(t,u')f_i(t,u)du'
\end{aligned} \tag{9}
$$

$$
\begin{aligned}
I_n = & - \int_{\mathbb{R}^+} k_{cn}(u'-u)H(u'-u)f_c(t,u')f_n(t,u)du' \\
& + P_{\mathrm{R}}(u) \int_{\mathbb{R}^+}\int_{\mathbb{R}^+} k_{cn}(u'-u'')H(u'-u'')f_c(t,u')f_n(t,u'')du'du''
\end{aligned} \tag{10}
$$

Each interaction term I_j accounts for all processes that change the nature or activity of cells of type $j = c, i, n$. As an example, we make precise how the interaction term I_i associated with the immune cells is built. The three first terms of the right-hand side of Eq. (9) are associated with the process given in Eq. (3). The first and third terms account for the positive contributions to the evolution of the distribution function of immune cells, i.e. to the processes leading to an immune cell of activity u after the interaction of an immune cell and a cancer cell. Specifically, the first term corresponds to the formation of an immune cell of activity u from an immune cell of activity $u - \epsilon$ before the interaction. The contribution

of all such interactions is obtained by integrating over all the activities u' of the interacting cancer cells. The third term is obtained by assigning the activity u of the disappearing cancer cell to a newly formed immune cell and integrating over all the activities u' of the interacting immune cells. The second term is negative and corresponds to the disappearance of immune cells of activity u due to their transformation into immune cells of activity $u + \epsilon$. The fourth term is related to the process given in Eq. (4).

The second term of the right-hand side of Eq. (10) accounts for the effect of the reservoir of normal cells that injects normal cells of activity u according to the distribution $P_R(u)$ given in Eq. (1) at the rate of consumption of normal cells imposed by the interaction between a cancer cell and a normal cell.

In order to simulate cell dynamics, we have adapted [6] the direct simulation Monte Carlo (DSMC) method initially introduced to simulate the Boltzmann equations associated with a dilute gas [5]. Figure 1a gives the evolution of the total number N_c of cancer cells and the total number N_i of immune cells for a critical value of the field E which controls the fluctuations of the activity. Figure 1b shows the evolution of the mean activities U_c and U_i in the same conditions. The initial decrease of the number of cancer cells followed by a pseudo-steady state and the final explosion of N_c illustrate a clinically observed phenomenon, the so-called 3Es of immunotherapy, i. e. the elimination, the equilibrium, and the escape of cancer from immunosurveillance [12]. The long induction period before the explosion is characteristic of the vicinity of a bifurcation between qualitatively different behaviors, the fast and monotonous explosion of cancer at weaker field and the control of cancer at stronger field [6–8].

2.2 Macroscopic Equations

The macroscopic quantity A_j for species $j = c, i, n$, which corresponds to the average value of the microscopic quantity a, is obtained by integration of the associated distribution function $f_j(t, u)$ over the activities

$$A_j = \langle a \rangle = \frac{1}{\rho_j} \int_{\mathbb{R}^+} a(t, u) f_j(t, u) \mathrm{d}u \tag{11}$$

where

$$\rho_j = \int_{\mathbb{R}^+} f_j(t, u) \mathrm{d}u \tag{12}$$

is the concentration of cells of type j in the system. The activity of the j cells is then

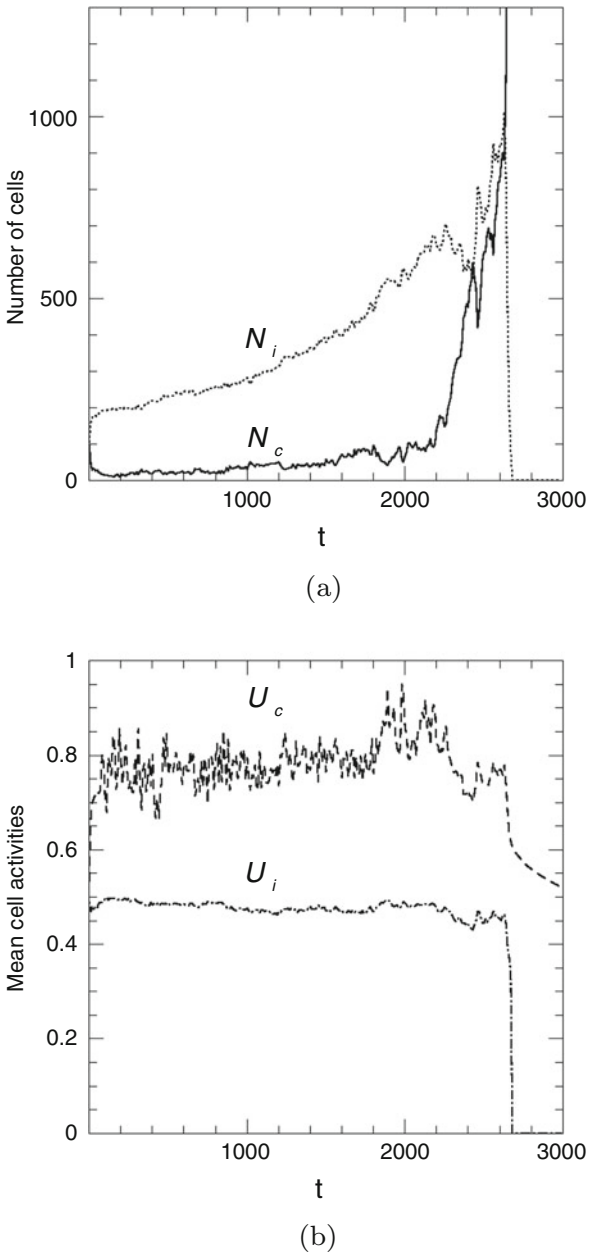


Fig. 1 Direct simulation of the kinetic equations: Illustration of the 3Es of immunotherapy. (**a**) Evolution of the total number N_c of cancer cells (solid line) and the total number N_i of immune cells (dotted line). (**b**) Evolution of the mean activity U_c of cancer cells (dashed line) and the mean activity U_i of immune cells (dash-dotted line). Parameter values: Total number of normal cells $N_n = 9800$, initial number of cancer cells $N_c(t = 0) = 100$, initial number of immune cells $N_i(t = 0) = 100$, initial activities of all cell types sampled from a Gaussian distribution of mean value $\mu = 0.5$ and standard deviation $\sigma = 0.2$, $\epsilon = 10^{-3}$, $k_{ic} = 10^{-2}$, $k_{ci} = 10^{-3}$, $k_{cn} = 10^{-6}$, $E = 7 \times 10^{-5}$

$$U_j = \frac{1}{\rho_j} \int_{\mathbb{R}^+} u f_j(t,u) \mathrm{d}u \tag{13}$$

Using Eq. (7), we derive the evolution equations for the concentrations ρ_j and the activities U_j

$$\frac{\mathrm{d}\rho_j}{\mathrm{d}t} = \int_{\mathbb{R}^+} I_j \mathrm{d}u \tag{14}$$

$$\frac{\mathrm{d}U_j}{\mathrm{d}t} = (E - \alpha U_j) - \frac{U_j}{\rho_j}\frac{\mathrm{d}\rho_j}{\mathrm{d}t} + \frac{1}{\rho_j}\int_{\mathbb{R}^+} u I_j \mathrm{d}u \tag{15}$$

We assume that the dependence of the distribution function f_j on the activity is a Dirac distribution

$$f_j(t,u) = \rho_j \delta(u - U_j) \tag{16}$$

We straightforwardly find that

$$\frac{\mathrm{d}\rho_n}{\mathrm{d}t} = 0 \tag{17}$$

$$\frac{\mathrm{d}U_n}{\mathrm{d}t} = 0 \tag{18}$$

The evolution equations of the concentrations of immune cells ρ_i and cancer cells ρ_c and the associated activities U_i and U_c are given by:

$$\frac{\mathrm{d}\rho_i}{\mathrm{d}t} = k_{ic}(U_i - U_c)H(U_i - U_c)\rho_i\rho_c - k_{ci}(U_c - U_i)H(U_c - U_i)\rho_i\rho_c \tag{19}$$

$$\begin{aligned}\frac{\mathrm{d}\rho_c}{\mathrm{d}t} =& k_{cn}\rho_n(U_c - U_n)H(U_c - U_n)\rho_c - k_{ic}(U_i - U_c)H(U_i - U_c)\rho_i\rho_c \\ &+ k_{ci}(U_c - U_i)H(U_c - U_i)\rho_i\rho_c\end{aligned} \tag{20}$$

$$\frac{\mathrm{d}U_i}{\mathrm{d}t} = E - \alpha U_i + k_{ic}(U_c - U_i + \epsilon)(U_i - U_c)H(U_i - U_c)\rho_c \tag{21}$$

$$\begin{aligned}\frac{\mathrm{d}U_c}{\mathrm{d}t} =& E - \alpha U_c + k_{cn}\rho_n(U_n - U_c + \epsilon)(U_c - U_n)H(U_c - U_n) \\ &+ k_{ci}(U_i - U_c + \epsilon)(U_c - U_i)H(U_c - U_i)\rho_i\end{aligned} \tag{22}$$

In the macroscopic description, Eq. (6) is meaningless and the values of the friction coefficient α are arbitrarily chosen.

3 Results

We first look for the stationary states of the macroscopic equations given in Eqs. (19)–(22). If one excepts the stationary states reached in the specific case where $U_n = E/\alpha$, the stationary states obey $U_i^S = E/\alpha$. The system admits two sets of stationary states $S_1 = (\rho_i^S, \rho_c^S, U_i^S = E/\alpha, U_c^S = E/\alpha)$ and $S_2 = (\rho_i^S, \rho_c^S = 0, U_i^S = E/\alpha, U_c^S(\rho_i^S))$. The values of ρ_i^S and ρ_c^S in the set S_1 depend on the initial condition $(\rho_i(t=0), \rho_c(t=0), U_i(t=0), U_c(t=0))$. In the set S_2, ρ_c^S always vanishes, the value of ρ_i^S is selected by the initial condition, and U_c^S is then determined by ρ_i^S. Different branches $U_c^S(\rho_i^S)$ are obtained depending on U_n. The stationary activity of the cancer cells obey

$$U_c^S(\rho_i^S) = U_n + \frac{k_{cn}\epsilon\rho_n - \alpha + \sqrt{(k_{cn}\epsilon\rho_n - \alpha)^2 + 4k_{cn}\rho_n(E - \alpha U_n)}}{2k_{cn}\rho_n} \tag{23}$$

$$\text{if} \quad U_n \leq E/\alpha - U_c^S(\rho_i^S)$$

$$U_c^S(\rho_i^S) = \frac{E}{\alpha} + \epsilon - \frac{\alpha}{k_{ci}\rho_i^S} \tag{24}$$

$$\text{if} \quad U_n \geq U_c^S(\rho_i^S) \quad \text{and} \quad \rho_i^S \geq \frac{\alpha}{k_{ci}\epsilon}$$

$$U_c^S(\rho_i^S) = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \tag{25}$$

$$\text{if} \quad U_n \leq U_c^S(\rho_i^S) \quad \text{and} \quad \frac{E}{\alpha} \leq U_c^S(\rho_i^S) \quad \text{where}$$

$$a = -(k_{cn}\rho_n + k_{ci}\rho_i^S)$$

$$b = -\alpha + k_{cn}\rho_n(2U_n + \epsilon) + k_{ci}\rho_i^S(2U_i + \epsilon)$$

$$c = E - k_{cn}\rho_n U_n(U_n + \epsilon) - k_{ci}\rho_i^S U_i(U_i + \epsilon)$$

Typical branches of stationary states as ρ_i^S varies are given in Fig. 2 for three values of U_n. For $U_n \leq E/\alpha$, the system exhibits a single stable stationary state. The specific case $U_n = E/\alpha$ shows a bifurcation for a critical value $\rho_i^{S,c}$. For $\rho_i^S \leq \rho_i^{S,c}$, the steady activity of the cancer cells U_c^S does not change with ρ_i^S. For $\rho_i^S > \rho_i^{S,c}$, the branch of constant U_c^S becomes unstable and the stable steady activity of the cancer cells increases as ρ_i^S increases. For $U_n > E/\alpha$, the behavior is more complex: The system admits simultaneously two stable stationary states and two unstable stationary states in an interval of ρ_i^S. For large values of ρ_i^S, a stable stationary state and an unstable one coexist.

Equations (19)–(22) are integrated numerically using Euler's method. Figure 3 shows the evolution of the concentrations and activities of cancer cells and immune cells for different values of the field E associated with the thermostat. For a

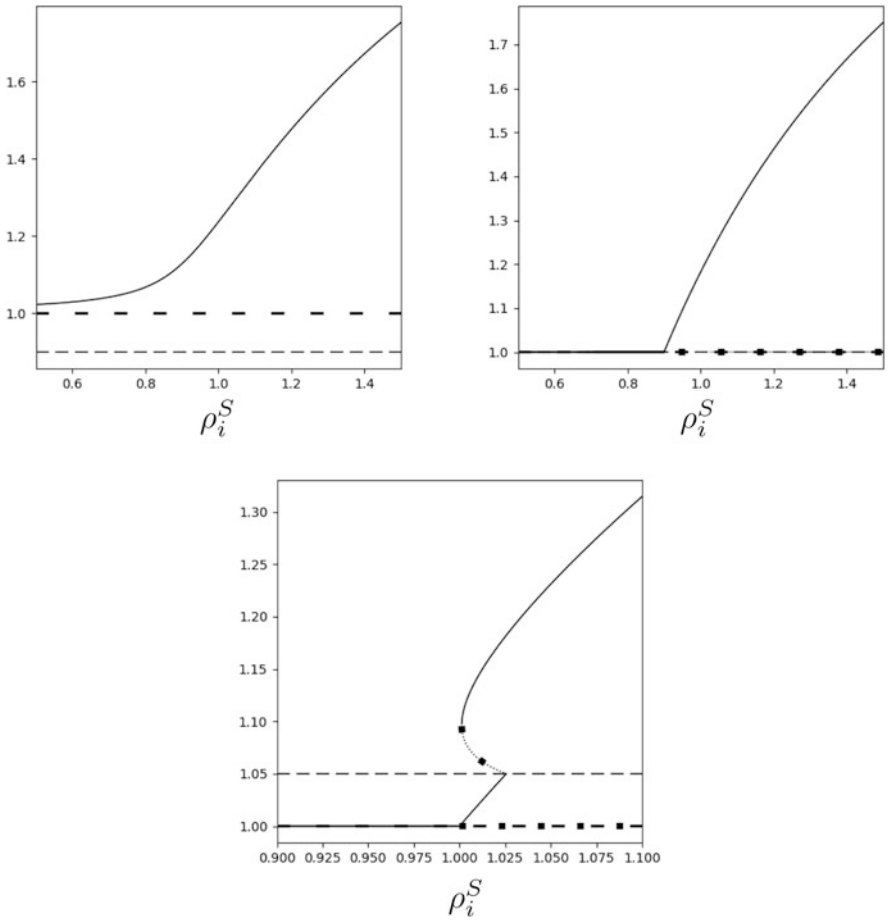


Fig. 2 Stationary states $(\rho_i^S, \rho_c^S = 0, U_i^S = E/\alpha, U_c^S)$ of Eqs. (19–22). Variations of U_c^S (solid lines for stable branches and squares for unstable branches) and U_i^S (thick long-dashed line) with respect to ρ_i^S for the following parameter values: $\epsilon = 2$, $k_{ic} = 0.5$, $k_{ci} = 0.1$, $k_{cn} = 0.01$, $E = \alpha = 0.2$, $\rho_n = 1$ for different values of U_n (dashed line): (top left) $U_n = 0.9$, (top right) $U_n = 1$, (bottom) $U_n = 1.05$

sufficiently small value $E = 0.1$, the system converges toward a stationary state associated with a nonvanishing value of ρ_c^S belonging to the set S_1 of steady states such that $U_i^S = U_c^S = E/\alpha$. For a sufficiently large value $E = 3$ of the field, the system converges toward a stationary state associated with a vanishing value $\rho_c^S = 0$ belonging to the set S_2 of steady states such that $U_i^S = E/\alpha$ and $U_c^S(\rho_i^S)$ obeys Eq. (25).

For an intermediate field value ($E = 0.8$), the system exhibits an interesting, non-stationary behavior. In this case, the system goes through three different phases. The first phase corresponds to a fast decrease of the concentration of cancer cells ρ_c and

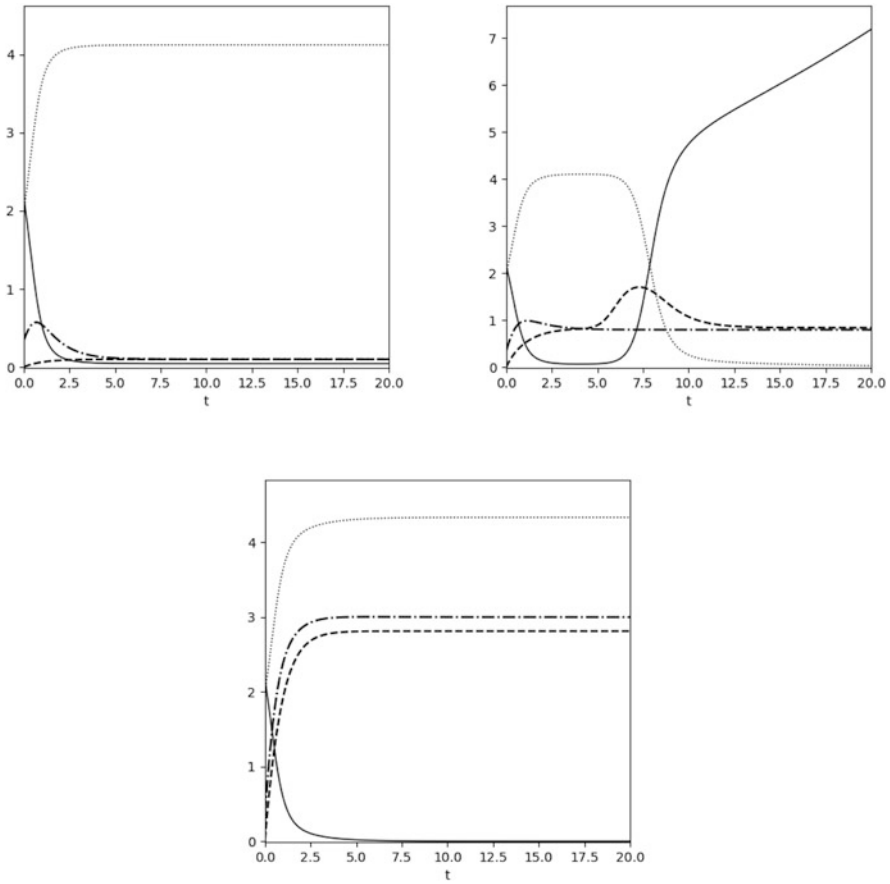


Fig. 3 Numerical integration of the macroscopic equations given in Eqs. (19)–(22). Evolutions of the concentration of cancer cells ρ_c (solid line), the concentration of immune cells ρ_i (dotted line), the activity of the cancer cells U_c (dashed line), and the activity of the immune cells U_i (dash-dotted line). The parameters are: $\rho_n(t=0)=1$, $\rho_i(t=0)=2$, $\rho_c(t=0)=6.33$, $U_n(t=0)=0.5$, $U_i(t=0)=0.36$, $U_c(t=0)=0$, $\epsilon=1.5$, $k_{ic}=1$, $k_{ci}=0.5$, $k_{cn}=0.1$, $\alpha=1$, $dt=0.01$. (top left) $E=0.1$, (top right) $E=0.8$, (bottom) $E=3$

a fast increase of the concentration of immune cells ρ_i, suggesting that the cancer will eventually fade out. The second phase shows a pseudo-steady state, where the concentration of immune cells ρ_i increases slowly and the concentration of cancer cells ρ_c looks stable and very small. When the activity of the cancer cells overcomes the activity of the immune cells, a third phase begins during which the rapid increase of the concentration of cancer cells is observed. The vanishing of the concentration of immune cells rapidly follows. This behavior is analogous to the one deduced from the simulations of the kinetic equations and given in Fig. 1. Hence, for well chosen values of the field E, the integration of the macroscopic equations typically leads to

the 3Es of immunotherapy [12], i.e. elimination, equilibrium, and escape of cancer from the control of the immune system. A more systematic sensitivity analysis of the macroscopic equations is necessary to explore all the complexity of the model, in particular with regards to possible pseudo-oscillations. The specific role of the fluctuations has also to be explored, using for example a Langevin approach or a master equation [16].

4 Conclusion

In this paper, we derive macroscopic equations from the kinetic equations associated with the interaction of cancer cells and immune system cells in the presence of a thermostat controlling the dissipation of cell activity. The first striking feature of the macroscopic equations is that they possess an infinite number of steady states for a given set of parameter values. The second remarkable property of the system is the strong dependence of the behavior on the initial conditions, which select a given steady state. For a constant value of the friction coefficient α and a small or large enough field E controlling the efficiency of thermalization, the system converges toward a steady state for which cancer is controlled. Interestingly, the 3Es of immunotherapy [12] are included in the macroscopic model and observed for an intermediate field value. The system certainly deserves further investigation, in particular to determine if the pseudo-oscillations of the number of cells observed in DSMC simulations require the presence of fluctuations and inhomogeneities [8].

References

1. D. S. Chen and I. Mellman, Oncology Meets Immunology: The Cancer-Immunity Cycle, *Immunity* **39**, 1–10 (2013).
2. C. Bianca, Onset of Nonlinearity in Thermostatted Active Particles Models for Complex Systems, *Nonlinear Analysis: Real World Applications* **13**, 2593–2608 (2012).
3. C. Bianca and C. Dogbe, Kinetic Models Coupled with Gaussian Thermostats: Macroscopic Frameworks, Nonlinearity **27**, 2771–2803 (2014).
4. C. Bianca, C. Dogbe and A. Lemarchand, The Role of Nonconservative Interactions in the Asymptotic Limit of Thermostatted Kinetic Models, *Acta Applicandae Mathematicae* **139**, 14–23 (2014).
5. G. A. Bird. *Molecular Gas Dynamics and the Direct Simulation of Gas Flows*, Oxford, New York, 1994.
6. C. Bianca and A. Lemarchand, Miming the Cancer-Immune System Competition by Kinetic Monte Carlo Simulations, *The Journal of Chemical Physics* **145**, 154108 (2016).
7. L. Masurel, C. Bianca and A. Lemarchand, On the Learning Control Effects in the Cancer-Immune System Competition, *Physica A* **506**, 462–475 (2018).
8. L. Masurel, C. Bianca and A. Lemarchand, Space-velocity thermostatted kinetic theory model of tumor growth, *Mathematical Biosciences and Engineering* **18**, 5525–5551 (2021).
9. S. J. Oiseth and M. S. Aziz, Cancer Immunotherapy: A Brief Review of the History, Possibilities, and Challenges Ahead, *Journal of Cancer Metastasis and Treatment* **3**, 250–261

(2017).
10. P. Sharma et al., Primary, Adaptive, and Acquired Resistance to Cancer Immunotherapy, *Cell* **168**, 707–723 (2017).
11. G. Gompper et al., The 2020 Motile Active Matter Roadmap, *Journal of Physics: Condensed Matter* **32**, 193001 (2020).
12. G. P. Dunn, L. J. Old and R. D. Schreiber, The Three Es of Cancer Immunoediting; *Annual Review of Immunology* **22**, 329–360 (2004).
13. D. C. Macallan, J. M. Borghans and B. Asquith, Human T-Cell Memory: A Dynamic View, *Vaccines* **5**, 5 (2017).
14. N Herranz and J. Gil, Mechanisms and functions of cellular senescence, *The Journal of Clinical Investigation* **128**, 1238 (2018).
15. B. Wennberg and Y. Wondmagegne, The Kac Equation with a Thermostatted Force Field, *Journal of Statistical Physics* **124**, 859–880 (2006).
16. C. W. Gardiner, *Handbook of stochastic methods*, Springer, Berlin, 1985.

Oscillations in a System Modelling Somite Formation

Sándor Kovács, Szilvia György, and Noémi Gyúró

1 Introduction

A minimal models of vertebrae formation were studied in [8] concerning periodic structures formation. The authors proposed two kinds of reaction-diffusion models, from which one is of clock-and-wavefront type and the other one is of Turing type. Our goal is to show that in case of the Turing type model the kinetic system as well the reaction-diffusion system exhibit oscillating solutions. The chapter is organised as follows. In the next section we introduce the model. In the section that follows we examine the existence and stability of some equilibria. In the third section we show the occurence of Hopf bifurcation in the kinetic system as well as in the parabolic system.

2 The Model

The model proposed by Annie Lemarchand and Bogdan Nowakowski (cf. [8]), which describes vertebrae formation is governed by

S. Kovács (✉)
Department of Numerical Analysis, Eötvös Loránd University, Budapest, Hungary
e-mail: alex@ludens.elte.hu

Sz. György · N. Gyúró
Eötvös Loránd University, Budapest, Hungary

R. P. Mondaini (ed.), *Trends in Biomathematics: Stability and Oscillations in Environmental, Social, and Biological Models*,
https://doi.org/10.1007/978-3-031-12515-7_13

$$\left.\begin{aligned} \partial_t A &= d_A \Delta_{\mathbf{r}} A + f_A(A, B), \\ \partial_t B &= d_B \Delta_{\mathbf{r}} B + f_B(A, B) \end{aligned}\right\} \tag{1}$$

on $\overline{\Omega} \times \mathbb{R}_0^+$ where Ω is a bounded, connected spatial domain with piecewise smooth boundary $\partial\Omega$, $d_A, d_B > 0$ represent the diffusion coefficients, $A(\mathbf{r}, t)$ and $B(\mathbf{r}, t)$ are the concentrations of the species at time $t \in [0, +\infty)$ and place $\mathbf{r} \in \overline{\Omega}$. The kinetic part of the model (1)

$$\left.\begin{aligned} \dot{A} &= f_A(A, B) := -\alpha A + \beta A^2 B, \\ \dot{B} &= f_B(A, B) := \gamma - \delta B - \beta A^2 B \end{aligned}\right\} \tag{2}$$

($\alpha, \beta, \gamma, \delta > 0$) was inspired from the Schnakenberg model (cf. [9])

$$\dot{A} = A^2 B - A, \qquad \dot{B} = -A^2 B + k_{Sch} \tag{3}$$

and the Gray-Schott model (cf. [4])

$$\dot{A} = -AB^2 - k_{GS}^1 A + k_{GS}^2, \qquad \dot{B} = AB^2 - k_{GS}^3 B - k_{GS}^4. \tag{4}$$

We are interested in solutions $\mathbf{\Phi} : \overline{\Omega} \times \mathbb{R}_0^+ \to \mathbb{R}^2$ of (2) that satisfy the no-flux boundary conditions

$$(\mathbf{n} \cdot \nabla_{\mathbf{r}}) \mathbf{S}(\mathbf{r}, t) = \mathbf{0} \quad \left((\mathbf{r}, t) \in \partial\Omega \times \mathbb{R}_0^+\right), \tag{5}$$

resp. non-negative initial conditions

$$\mathbf{S}(\mathbf{r}, 0) = \mathbf{S}_0(\mathbf{r}) \geq \mathbf{0} \quad \left((\mathbf{r}, t) \in \overline{\Omega} \times \{0\}\right) \tag{6}$$

where $\mathbf{S} := (A, B)$, and $\mathbf{n}$ denotes the outer unit normal to $\partial\Omega$.

3 The Kinetic System

It was mentioned in the original paper [8] that parameters $\alpha, \beta, \gamma, \delta$ are chosen such that the system possesses three steady states. It is easy to see that this is the case when

$$K := \beta\gamma^2 - 4\alpha^2\delta > 0 \tag{7}$$

holds. In this case the kinetic system (2) exhibits three equilibria in the first quadrant of the phase space, namely one on the boundary: $E_b := (0, \gamma/\delta)$ and two interior

equilibria (cf. Fig. 1.): $E_\pm := (A_\mp, B_\pm)$ where

$$A_\pm := \frac{\beta\gamma \pm \sqrt{\beta K}}{2\alpha\beta} \qquad \text{and} \qquad B_\pm := \frac{\alpha}{\delta} \cdot A_\pm.$$

In what follows we study the stability of possible equilibria $\overline{\mathbf{S}} := (\overline{S}_1, \overline{S}_2)$ of the kinetic system (2) and the possibility of Hopf bifurcations. The coefficient matrix of the system linearized at $\overline{\mathbf{S}}$ is

$$\mathfrak{A} := J_{(f_A, f_B)}(\overline{S}_1, \overline{S}_2) = \begin{bmatrix} -\alpha + 2\beta\overline{S}_1\overline{S}_2 & \beta\overline{S}_1^2 \\ -2\beta\overline{S}_1\overline{S}_2 & -\delta - \beta\overline{S}_1^2 \end{bmatrix}$$

with trace

$$\mathrm{Tr}(J_{(f_A, f_B)}(\overline{S}_1, \overline{S}_2)) = -\alpha + 2\beta\overline{S}_1\overline{S}_2 - \delta - \beta\overline{S}_1^2$$

and determinant

$$\det(J_{(f_A, f_B)}(\overline{S}_1, \overline{S}_2)) = \alpha\delta + \alpha\beta\overline{S}_1^2 - 2\beta\delta\overline{S}_1\overline{S}_2.$$

A simple linear stability analysis shows that E_b is always locally asymptotically stable, because the Jacobian of system (2) at these equilibrium point takes the form

$$J_b := J_{(f_A, f_B)}(0, \gamma/\delta) = \begin{bmatrix} -\alpha & 0 \\ 0 & -\delta \end{bmatrix}.$$

The Jacobians evaluated at $E_\pm$ have the form

$$J_+ := J_{(f_A, f_B)}(E_+) = \begin{bmatrix} \alpha & \frac{(\sqrt{\beta K} - \beta\gamma)^2}{4\alpha^2\beta} \\ -2\alpha & \frac{\gamma(\sqrt{\beta K} - \beta\gamma)}{2\alpha^2} \end{bmatrix}$$

and

$$J_- := J_{(f_A, f_B)}(E_-) = \begin{bmatrix} \alpha & \frac{(\sqrt{\beta K} + \beta\gamma)^2}{4\alpha^2\beta} \\ -2\alpha & -\frac{\gamma(\sqrt{\beta K} + \beta\gamma)}{2\alpha^2} \end{bmatrix}.$$

Based on the form of the characteristic polynomial

$$z^2 - \mathrm{Tr}(\mathfrak{A})z + \det(\mathfrak{A}) \qquad (z \in \mathbb{C})$$

it is easy to determine the stability of the equilibrium points. It is clear that

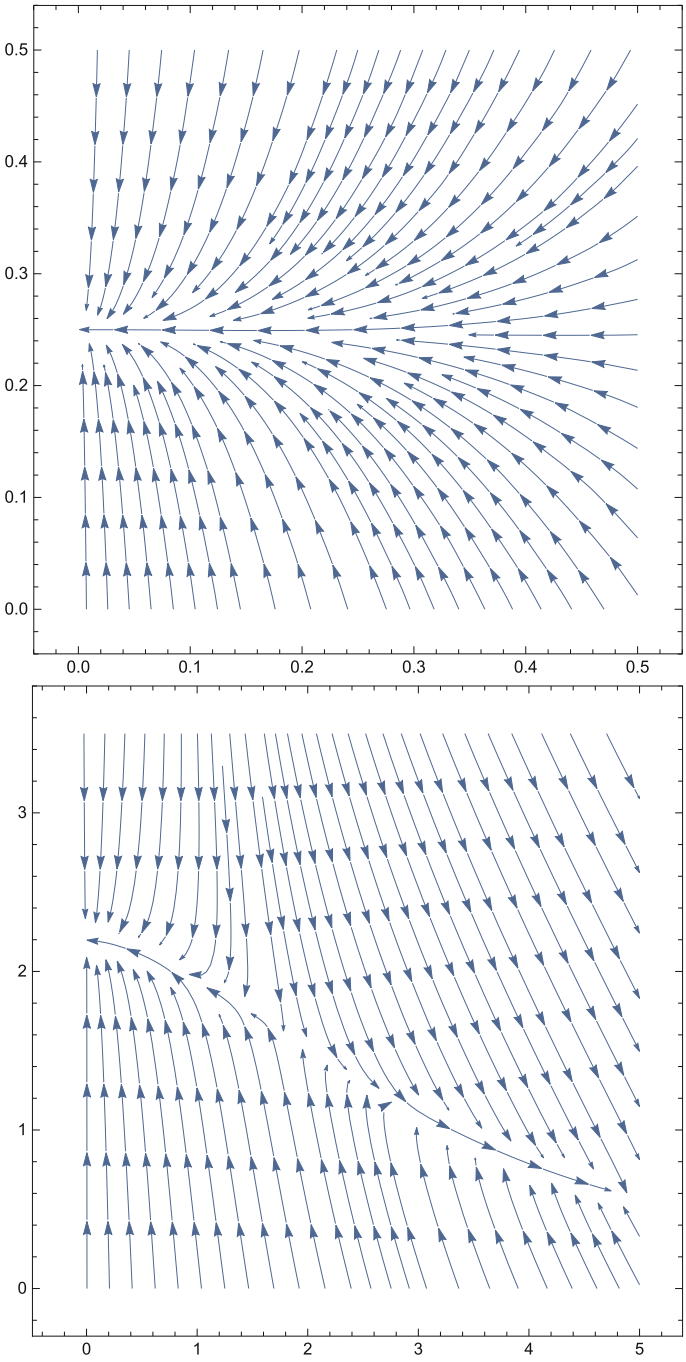


Fig. 1 A phase portrait of system (2) for $K < 0$, resp. $K > 0$

- the matrix J_+ is unstable, because

$$\mathrm{Tr}(J_+) = \alpha + \frac{\gamma\left(\sqrt{\beta K} - \beta\gamma\right)}{2\alpha^2}$$

and

$$\det(J_+) = \frac{K - \gamma\sqrt{\beta K}}{2\alpha},$$

furthermore $\det(J_+)$ is negative due to $K < \beta\gamma^2$;
- the matrix J_- is stable if and only if

$$\mathrm{Tr}(J_-) = \alpha - \frac{\gamma\left(\sqrt{\beta K} + \beta\gamma\right)}{2\alpha^2} < 0$$

hold, because

$$\det(J_-) = \frac{K + \gamma\sqrt{\beta K}}{2\alpha} > 0.$$

Since the determinant of J_- stays positive, then Hopf bifurcation can occur only if the trace is changing its sign. It is easy to calculate that if $\alpha > \delta$ then $\mathrm{Tr}(J_-) = 0$ if and only if

$$\beta = \frac{\alpha^4}{\gamma^2(\alpha - \delta)}$$

holds. Thus, by fixed α, γ, δ, the parameter β will play the role of the bifurcating parameter.

Theorem 3.1 *Suppose that*

$$\alpha > \delta \qquad \text{and} \qquad \alpha \neq 2\delta \tag{8}$$

hold, then at

$$\beta^* := \frac{\alpha^4}{\gamma^2(\alpha - \delta)} \tag{9}$$

the equilibrium $E_-(\beta)$ of (2) undergoes a Poincaré-Andronov-Hopf bifurcation: $E_-(\beta)$ loses its stability at β^* and system (2) has a branch of periodic solutions bifurcating from $E_-(\beta)$ near $\beta = \beta^*$ (cf. Fig. 2.).

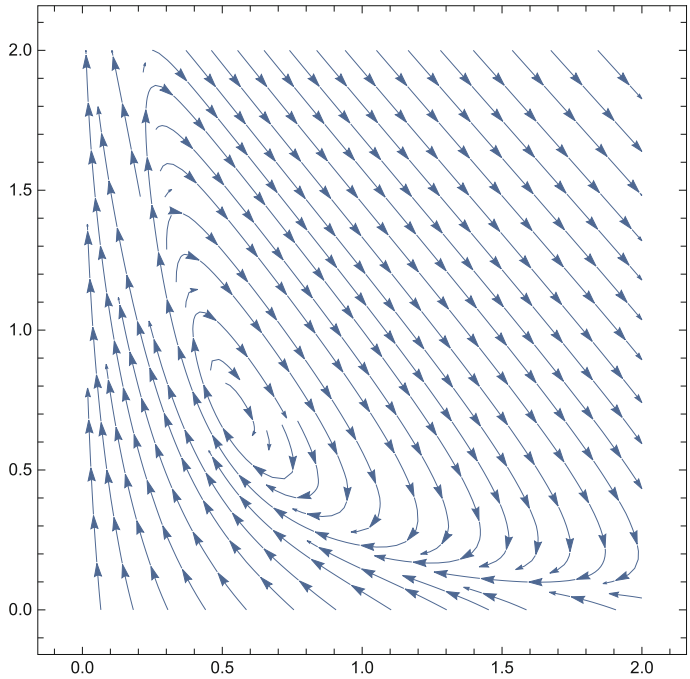


Fig. 2 A phase portrait of system (2) when (7), (8) and (9) hold

Proof The characteristic polynomial of the matrix $\mathfrak{A}$ at $E_-(\beta)$ has the form

$$\Delta_{\mathfrak{A}}(z,\beta) := z^2 - \mathrm{Tr}(\beta)z + \det(\beta) \qquad (z \in \mathbb{C})$$

where

$$\mathrm{Tr}(\beta) := \mathrm{Tr}\,(J_-(\beta)) \qquad \text{and} \qquad \det(\beta) := \det\,(J_-(\beta))\,.$$

Clearly,

$$\det(\beta^*) = \alpha(\alpha-\delta) > 0,$$

resp. from (8)

$$\mathrm{Tr}(\beta^*) = 0 \qquad \text{and} \qquad \frac{\partial}{\partial\beta}\,\mathrm{Tr}(\beta^*) = \frac{\gamma(\alpha-\delta)^2}{\alpha^3(\alpha-2\delta)} \neq 0$$

follows, which proves the statements of the theorem (cf. [5, 7]). □

4 The Parabolic System

In what follows we consider system (1) with homogeneous Neumann boundary conditions (5) and nonnegative initial conditions (6). Clearly, a spatially constant solution $\mathbf{\Phi}(\cdot) = (\Phi_1(\cdot), \Phi_2(\cdot))$ of system (1) satisfies boundary conditions (5) and system (2). The equilibria of system (2) are constant solutions of (1), (5) at the same time. In order to investigate the local dynamical behavior of system (1) near the equilibria E_b and $E_\pm$ of (2) we linearize (1) at these equilibria. The linearized system at the equilibrium point

$$\mathbf{S} = (S_1, S_2) \in \{E_b, E_\pm\}$$

with the same initial and boundary conditions has the form

$$\frac{\partial \mathbf{Z}}{\partial t} = D \cdot \Delta_{\mathbf{r}} \mathbf{Z} + \mathfrak{A}\mathbf{Z} \quad \text{in} \quad \Omega \times \mathbb{R}_0^+ \tag{10}$$

$$(\mathbf{n} \cdot \nabla_{\mathbf{r}})\, \mathbf{Z} = \mathbf{0} \quad \text{in} \quad \partial\Omega \times \mathbb{R}_0^+ \tag{11}$$

$$\mathbf{Z}(\mathbf{r}, 0) = \mathbf{Z}_0(\mathbf{r}) \quad \text{on} \quad \overline{\Omega} \times \{0\} \tag{12}$$

where

$$\mathfrak{A} := J_{(f_A, f_B)}(\mathbf{S}) =: \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

Using the method of eigenfunction expansions for the spatial domain Ω the solutions of problem (10) and (11) have the form

$$\mathbf{\Psi}(\mathbf{r}, t) = \sum_{n=0}^{\infty} \psi_n(\mathbf{r}) \exp(\mathfrak{A}_n t)\, \mathbf{\Psi}_{0_n} \qquad ((\mathbf{r}, t) \in \overline{\Omega} \times \mathbb{R}_0^+)$$

(cf. [6]), where for $n \in \mathbb{N}_0$

$$\mathfrak{A}_n := \mathfrak{A} - \lambda_n D, \qquad \mathbf{\Psi}_{0_n} := \int_\Omega \mathbf{Z}_0(\mathbf{r})\, \psi_n(\mathbf{r})\, \mathrm{d}\mathbf{r}$$

and λ_n is the nth eigenvalue of the minus Laplacian on Ω subject to homogeneous Neumann boundary conditions, resp. ψ_n is the corresponding normalized eigenfunction, i.e. λ_n and ψ_n are solutions of

$$\Delta\psi = -\lambda\psi, \qquad \left.\frac{\partial \psi}{\partial \mathbf{n}}\right|_{\partial\Omega} = 0.$$

It is well known (cf. [3]) that

$$0 = \lambda_0 \leq \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n \longrightarrow +\infty \qquad (n \to \infty)$$

and the eigenfunctions to different eigenvalues are orthogonal to each other.

According to [1, 2] the equilibrium $\mathbf{S}$ of (1), (5) is asymptotically stable if for all $n \in \mathbb{N}_0$ the matrix $\mathfrak{A}_n$ is stable, i.e. both eigenvalues of $\mathfrak{A}_n$ have negative real parts; furthermore $\mathbf{S}$ is unstable if for some index $n \in \mathbb{N}_0$ there exists an eigenvalue of $\mathfrak{A}_n$ with positive real part. The characteristic polynomial of the matrix $\mathfrak{A}_n$ have the form

$$\Delta_{\mathfrak{A}_n}(z) := z^2 - \mathfrak{T}_n z + \mathfrak{D}_n \qquad (z \in \mathbb{C}) \tag{13}$$

where

$$\mathfrak{T}_n := \operatorname{Tr}(\mathfrak{A}_n) = \operatorname{Tr}(\mathfrak{A}) - \lambda_n \operatorname{Tr}(D)$$

and

$$\mathfrak{D}_n := \det(\mathfrak{A}_n) = \lambda_n^2 \det(D) - \lambda_n \left(d_A a_{22} + d_B a_{11}\right) + \det(\mathfrak{A}).$$

Thus, if $\mathbf{S} = (S_1, S_2) = E_-$ then for all $\beta > 0$ the characteristic equation of $\mathfrak{A}_n$ has the form

$$\Delta_{\mathfrak{A}_n}(z, \beta) = z^2 - \mathfrak{T}_n(\beta) z + \mathfrak{D}_n(\beta) = 0 \qquad (z \in \mathbb{C},\ n \in \mathbb{N}_0)$$

where

$$\mathfrak{T}_n(\beta) := \alpha - \frac{\gamma\left(\beta\gamma + \sqrt{\beta K}\right)}{2\alpha^2} - \lambda_n(d_A + d_B)$$

and

$$\mathfrak{D}_n(\beta) := \lambda_n^2 d_A d_B + \left(\frac{d_A \gamma(\sqrt{\beta K} + \beta\gamma)}{2\alpha^2} - d_B\alpha\right)\lambda_n + \frac{K + \gamma\sqrt{\beta K}}{2\alpha}.$$

In order to have Hopf bifurcation one has to show that a pair of complex conjugate roots

$$\mu(\beta) \pm \imath\, \nu(\beta)$$

crosses the imaginary axis with non-zero velocity, that is for a $\beta_* > 0$

$$\mu(\beta_*) = 0, \qquad \nu(\beta_*) \neq 0 \qquad \text{and} \qquad \mu'(\beta_*) \neq 0$$

hold. This is fulfilled (cf. [5]) if exists $n \in \mathbb{N}_0$ and $\beta_* > 0$ such that

$$\mathfrak{T}_n(\beta_*) = 0, \qquad \frac{\partial}{\partial\beta}\mathfrak{T}_n(\beta_*) \neq 0, \qquad \mathfrak{D}_n(\beta_*) > 0 \tag{14}$$

and

$$\mathfrak{T}_m(\beta_*) \neq 0, \quad \mathfrak{D}_m(\beta_*) \neq 0 \qquad (n \neq m \in \mathbb{N}_0). \tag{15}$$

We have to remark that β^* in (9) is always a Hopf bifurcation value, since

$$\mathfrak{T}_0(\beta^*) = 0 \qquad \text{and} \qquad \mathfrak{T}_n(\beta^*) = -\lambda_n(d_A + d_B) < 0 \quad (n \in \mathbb{N}),$$

resp.

$$\mathfrak{D}_n(\beta^*) = \lambda_n^2 d_A d_B + \alpha\,(d_A - d_B)\,\lambda_n + \alpha(\alpha - \delta) > 0 \qquad (n \in \mathbb{N}_0)$$

if

$$d_A > d_B \qquad \text{and} \qquad \alpha > \delta \tag{16}$$

hold. This corresponds to the Hopf bifurcation of spatially homogeneous periodic orbits which have been known from Theorem 3.1. Apparently β^* is also the unique value for β for the Hopf bifurcation of spatially homogeneous periodic orbits (cf. Fig. 3.).

In what follows, we shall search for spatially non-homogeneous Hopf bifurcation value in case of $n \in \mathbb{N}$. For $0 \leq \beta \in \mathbb{R}$ let define

$$E(\beta) := \alpha - \frac{\gamma\left(\beta\gamma + \sqrt{\beta K}\right)}{2\alpha^2} = \alpha - \frac{\gamma\left(\beta\gamma + \sqrt{\beta(\beta\gamma^2 - 4\alpha^2\delta)}\right)}{2\alpha^2}$$

then

$$E(0) = \alpha > 0 \qquad \text{and} \qquad \lim_{\beta\to+\infty} E(\beta) = -\infty$$

and it follows from (7) that

$$E'(\beta) = -\frac{\gamma\left(\gamma + \frac{\beta\gamma^2 - 2\alpha^2\delta}{\sqrt{\beta(\beta\gamma^2 - 4\alpha^2\delta)}}\right)}{2\alpha^2} < 0 \qquad (\beta > 0).$$

This means that E is stricktly decreasing and in case of

$$\alpha > 2\delta + \lambda_n(d_A + d_B) \tag{17}$$

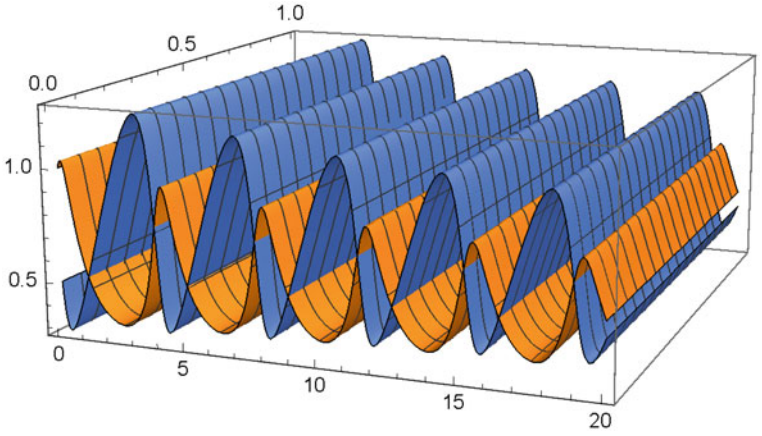


Fig. 3 Solution of system (1) when (9) and (16) hold

there is a unique solution $\beta = \beta_n > 0$ of the equation

$$E(\beta) = \lambda_n(d_A + d_B), \qquad \text{resp.} \qquad \mathfrak{T}_n(\beta) = 0.$$

Direct calculation shows that the unique positive solution has the form

$$\beta_n := \frac{\alpha^2 \, (\alpha - \lambda_n(d_A + d_B))^2}{\gamma^2 \, (\alpha - \delta - \lambda_n(d_A + d_B))}$$

holds.

Theorem 4.1 *The transversality condition, i.e.*

$$\mu'(\beta_n) < 0 \qquad (n \in \mathbb{N}_0) \tag{18}$$

is satisfied.

Proof It is easy to see that

$$\mu'(\beta_n) = \frac{1}{2} \cdot \frac{\partial}{\partial \beta} \mathfrak{T}_n(\beta_n) = \frac{1}{2} \cdot E'(\beta_n) < 0 \qquad (n \in \mathbb{N}_0)$$

which proves the statement of the theorem. □

It is also clear that for all $n \in \mathbb{N}$

$$\mathfrak{T}_n(\beta_n) = 0 \qquad \text{and} \qquad \mathfrak{T}_m(\beta_n) \neq 0 \quad (n \neq m \in \mathbb{N})$$

hold.

Next we will investigate whether

$$\mathfrak{D}_m(\beta_n) \neq 0 \qquad (m \in \mathbb{N}_0),$$

and in particular, $\mathfrak{D}_n(\beta_n) > 0$. It is easy to see that

$$\mathfrak{D}_m(\beta_n) = \lambda_m^2 d_A d_B + \left(\frac{d_A \gamma(\sqrt{\beta_n K} + \beta_n \gamma)}{2\alpha^2} - d_B \alpha \right) \lambda_m + \frac{K + \gamma\sqrt{\beta K}}{2\alpha}.$$

Because

$$\frac{d_A \gamma(\sqrt{\beta_n K} + \beta_n \gamma)}{2\alpha^2} - d_B \alpha = \frac{\gamma d_A}{2\alpha^2} \cdot (E + F) - d_B \alpha$$

where

$$E := \frac{\alpha^2 \left(\alpha - \lambda_n (d_A + d_B)\right)^2}{\gamma \left(\alpha - \delta - \lambda_n (d_A + d_B)\right)},$$

and

$$F := \sqrt{\frac{\alpha^4 (\alpha - \lambda_n (d_A + d_B))^2 (\alpha - 2\delta - \lambda_n (d_A + d_B))^2}{\gamma^2 (\alpha - \delta - \lambda_n (d_A + d_B))^2}}$$

we obtain the following result.

Theorem 4.2 *If an $n \in \mathbb{N}_0$ is chosen such that assumptions (17) and*

$$d_B < d_A \cdot \frac{\alpha - \lambda_n d_A}{\alpha + \lambda_n d_A} \tag{19}$$

hold, then in system (1) Poincaré-Andronov-Hopf bifurcation takes place: $E_-(\beta)$ loses its stability at β_n and system (1) has a branch of periodic solutions bifurcating from $E_-(\beta)$ near $\beta = \beta_n$.

Proof Obviously, if (17) holds then

$$F = \frac{\alpha^2 (\alpha - \lambda_n (d_A + d_B))(\alpha - 2\delta - \lambda_n (d_A + d_B))}{\gamma \left(\alpha - \delta - \lambda_n (d_A + d_B)\right)}.$$

Thus,

$$\frac{\gamma d_A}{2\alpha^2}\cdot(E+F) = d_A\cdot\left\{\frac{(\alpha-\lambda_n(d_A+d_B))^2}{\alpha-\delta-\lambda_n(d_A+d_B)}\right.$$

$$\left.+\frac{(\alpha-\lambda_n(d_A+d_B))(\alpha-2\delta-\lambda_n(d_A+d_B))}{\alpha-\delta-\lambda_n(d_A+d_B)}\right\}$$

$$= \frac{d_A(\alpha-\lambda_n(d_A+d_B))\,[2\alpha-2\delta-2\lambda_n(d_A+d_B)]}{2\,(\alpha-\delta-\lambda_n(d_A+d_B))}$$

$$= d_A(\alpha-\lambda_n(d_A+d_B)).$$

Hence, if condition (19) holds then

$$\frac{\gamma d_A}{2\alpha^2}\cdot(E+F) - d_B\alpha = -d_B(\alpha+\lambda_n d_A) + d_A(\alpha-\lambda_n d_A) > 0.$$

This means that $\mathfrak{D}_m(\beta_n) > 0$. This with the transversality condition (18) together proves Hopf bifurcation. □

Acknowledgments The authors were supported in part by the European Union, co-financed by the European Social Fund (EFOP-3.6.3-VEKOP-16-2017-00001).

References

1. Britton, N. F.: *Reaction-diffusion equations and their applications to biology*. Academic Press, Inc., London, 1986.
2. Casten, Richard G.; Holland, Charles J.: *Stability properties of solutions to systems of reaction-diffusion equations*, SIAM J. Appl. Math., **33**(2) (1977), 353–364.
3. Evans, L. C.: *Partial differential equations*, Second edition. Graduate Studies in Mathematics, 19. American Mathematical Society, Providence, RI, 2010.
4. Gray, P.; Scott, S. K.: *Autocatalytic reactions in the isothermal continuous stirred tank reactor: oscillations and instabilities in the system* $A + 2B \to 3B$; $B \to C$, Chem. Eng. Sci., **39** (1984), 1087–1097.
5. Hassard, B. D.; Kazarinoff, N. D.; Wan, Y. H.: *Theory and applications of Hopf bifurcation*, London Mathematical Society Lecture Note Series, 41. Cambridge University Press, Cambridge-New York, 1981.
6. Kovács, S.: *Turing bifurcation in a system with cross diffusion*, Nonlinear Anal., **59**(4) (2004), 567–581.
7. Kovács, S.; György, S.Z.; Gyúró, N.: *On an Invasive Species Model with Harvesting*, in: Trends in Biomathematics: Modeling Cells, Flows, Epidemics, and the Environment (ed. R. Mondaini), (Springer 2020), 299–334.
8. Lemarchand, A.; Nowakowski, B.: *Do the internal fluctuations blur or enhance axial segmentation?*, EPL, **94** (2011) 48004.
9. Schnakenberg, J.: *Simple chemical reaction systems with limit cycle behaviour*, J. Theoret. Biol. , **81**(3) (1979), 389–400.

Socio-Ecological Dynamics Generated by Hydrocarbon Exploration

J. M. Redondo, J. S. Garcia, and J. A. Amador

1 Introduction

Hydrocarbons meet the basic needs of today's societies such, as access to energy, heating, cooling, transportation, and production of inputs, as well as being important for promoting the industry and boosting economic growth. However, exploration and extraction activities have implications for the entire socio-ecological system, including environmental impacts, social problems, and governance asymmetries that lead to questioning the benefits of the use of hydrocarbons [1–3].

One of the environmental problems that hydrocarbon exploration brings with it is related to noise. Sound is an environmental feature used by some taxa for food, reproduction, navigation, and avoiding predators. Consequently, soundscape alterations have the potential to alter the behavior, physiology, and fitness of individuals [4]. In hydrocarbon exploration, noise pollution is caused by seismic exploration studies in which oil and gas reserves are identified, while during hydrocarbon exploitation, in addition to noise, the surface vegetation cover is removed for the construction of roads, camps, or platforms, and the increase in human activity that generates disturbance of wildlife and the migration of species necessary for the preservation of the ecosystem [4–6].

J. M. Redondo (✉)
Universidad Católica de Colombia, Bogotá, Colombia
e-mail: jmredondo@ucatolica.edu.co

J. S. Garcia
Universidad Nacional de Colombia, Medellin, Colombia
e-mail: jugarciaco@unal.edu.co

J. A. Amador
Instituto Humboldt, Bogotá, Colombia
e-mail: jamador@humboldt.org.co

R. P. Mondaini (ed.), *Trends in Biomathematics: Stability and Oscillations in Environmental, Social, and Biological Models*,
https://doi.org/10.1007/978-3-031-12515-7_14

It is highlighted that the implications of exploration and exploitation are different in intensity, duration, and frequency, being more impressive the exploitation activities, but opening the exploration door could open the exploitation door with it, so there is always the potential risk of losing the landscape ecological structure.

In the exploration of hydrocarbons, problems related to the impacts on biodiversity that can lead to the decrease of fauna in the exploration sites are recognized [7–9]. It is also known that hydrocarbon exploration is directly related to socioeconomic impacts that could determine the landscape dynamics. Some of the related positive impacts include business opportunities and local job creation [5, 10, 11], while negative impacts include price inflation, the generation of development expectations that may not come true, economic inequality, and the population transgression [1, 10].

This chapter focuses on the effects of hydrocarbon exploration in three different types of landscapes in the Amazon region: pristine landscapes, agricultural landscapes, and landscapes with licit and illicit crops. In this way, the coverage of the different land uses, biodiversity, soil productivity, the hiring of personnel for the economic activities of the landscape, and situations such as land retention and social coercion are articulated, with the absence or presence of hydrocarbon exploration.

For the analysis, a systemic approach based on the system dynamics[12, 13] methodology has been used, which allows obtaining a system of differential equations with which the simulation of the scenarios representing the three types of landscapes mentioned has been carried out. The simulation data are theoretical data that allow a phenomenon explanation, rather than a diagnosis or forecast in a specific situation. In this way, were obtained valuable arguments for decision-making via experimentation in controlled simulation environments.

In Sect. 2 the modeling process based on system dynamics is presented to obtain the mathematical model from which the simulations of the scenarios were carried out. In Sect. 3 the scenarios for the three proposed landscapes are evaluated, comparing the results of the presence and absence of hydrocarbon exploration on the variables of interest. Finally, in Sect. 4 the conclusions of this chapter are presented.

2 System Modeling

The modeling process involves the creation of a causal diagram and the discussion of its feedback structures (Sect. 2.1), the construction of the Forrester diagram, definition of the system equations, and the dimensional consistency validation (Sect. 2.2).

2.1 Causal Diagram

The proposed model is based on the causal diagram presented in Fig. 1, which is explained in the following lines. The causal diagram is explained firstly from

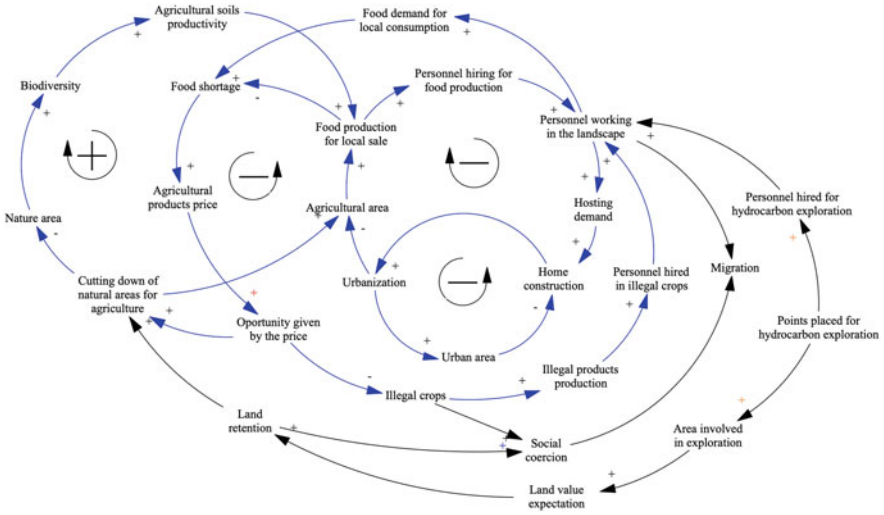


Fig. 1 Causal diagram

the landscape relationships without oil activities, while, towards the end, the hydrocarbon exploration activities are articulated.

The proposed model is based on the causal diagram presented in Fig. 1, which is explained in the following lines. The causal diagram is explained firstly from the landscape relationships without oil activities, while, towards the end, the hydrocarbon exploration activities are articulated.

Four types of general coverage were considered: natural, agricultural, illicit crops, and urban. It is assumed that the natural areas do not have the opportunity to increase in area, but they are diminished by agricultural activities and illicit crops, exclusively.

The agricultural crops vary due to the acquisition of natural areas, the adaptation of illicit crops to agricultural land, the adaptation of agricultural to illicit crops, and urbanization.

The illicit crops vary due to the transformation of natural areas, the adaptation of illicit crops to agriculture, and the adaptation of agriculture to illicit crops. Their urbanization is not considered because it is assumed that these crops are not usually very close to urban coverage.

Urban coverage, for its part, will always increase in area, although they do so very slowly.

The transformation drivers considered are the land retention due to the purchase speculation of the hydrocarbon industry and the price in the market of the agricultural products obtained, in comparison with the prices, generally very stable, of the products obtained in illicit crops (coca leaf, coca paste, and cocaine base).

Thus, the price of agricultural products increases the opportunity for profit in agricultural production, which leads to an increase in forest clearance for agricultural activities and a decrease in coverage of illicit crops. But forest clearing

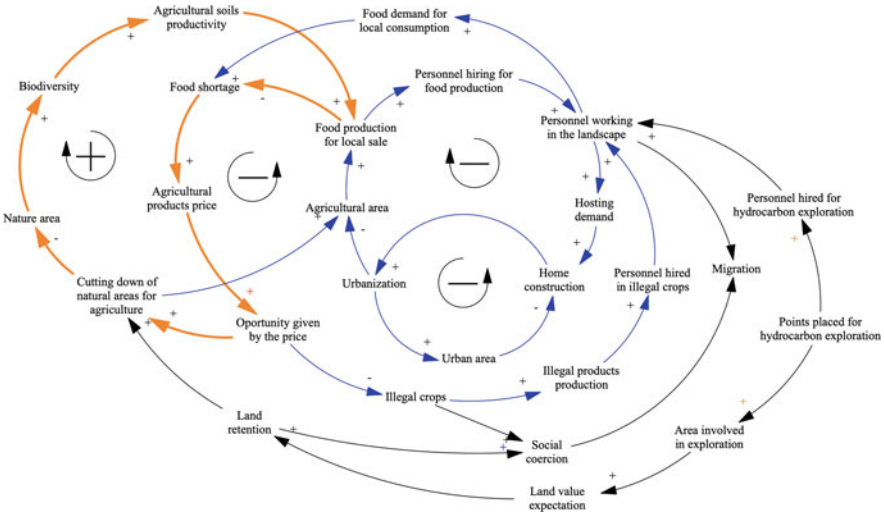


Fig. 2 First loop analyzed in the causal diagram

reduces the coverage of natural areas, which in turn decreases biodiversity, affecting the productivity of agriculture, thus reducing the food production for local sale. This decrease in production generates a food shortage for local consumption, making it scarce and, therefore, increasing the price of agricultural products. In this way, a positive feedback loop is formed (Fig. 2), which forms the vicious circle in which natural covers and biodiversity tend to disappear, thus affecting the productivity of landscapes to produce and supply food to the landscape population.

The price of agricultural products determines the opportunity for illicit crops, increasing the production of illegal products. This increase demands the hiring of people who will add within the total set of personnel working in the landscape, who demand food. If this demand for food is very high, either due to the number of people or the per capita amount of consumption, food will tend to become scarce, affecting the price of agricultural products for local consumption. In this way, a negative feedback loop is formed (Fig. 3) that limits the expansion of the coverage of illicit crops due to the need to supply food to the landscape population.

The total number of people who work in the landscape, which includes the personnel who work in agricultural crops, in illicit activities, and in goods and services, will demand a space to live, which will induce the construction of houses and, therefore, the increase in urbanization that will reduce agricultural coverage, affecting food production and, in turn, the person hired for agricultural activities, conditioning the total number of people who work in the landscape. The negative loop that emerges, (Fig. 4) explains that to maintain self-sufficiency, urbanization will have a limit on agricultural coverage or will have to import its supply.

In Fig. 5, is showed that the construction of houses for the accommodation of people who work in the landscape, including their families, leads to urbanization and

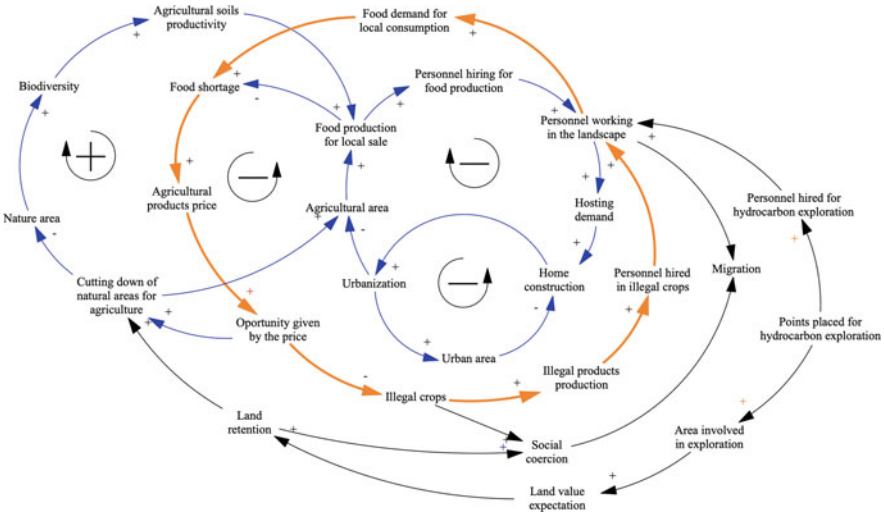


Fig. 3 Second loop analyzed in the causal diagram

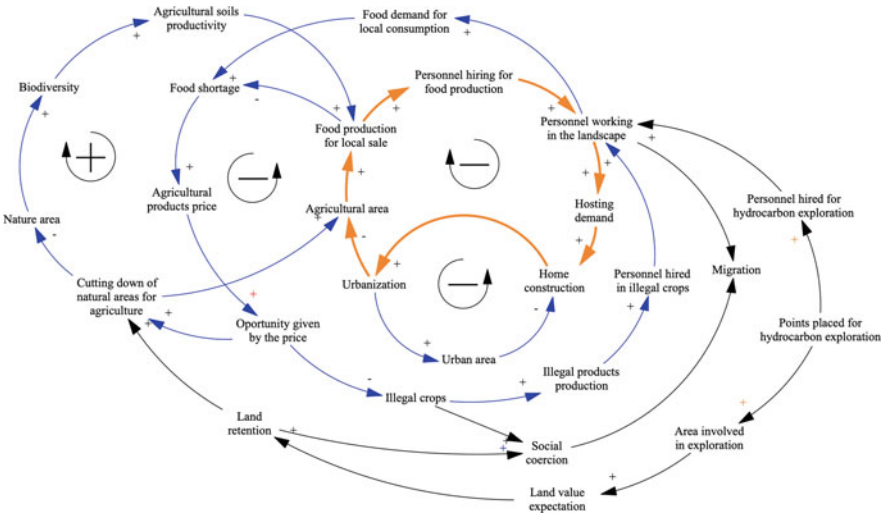


Fig. 4 Third loop analyzed in the causal diagram

an increase in urban coverage, and that enough urbanization reduces construction of housing. This negative loop shows that construction is limited by the need to urbanize, that is, the limit is given by the socioeconomic opportunity that the population finds in the landscape.

When the hydrocarbon exploration activity is carried out, the landscape is affected in two ways: by the hiring of personnel and by the area that is involved in the exploration.

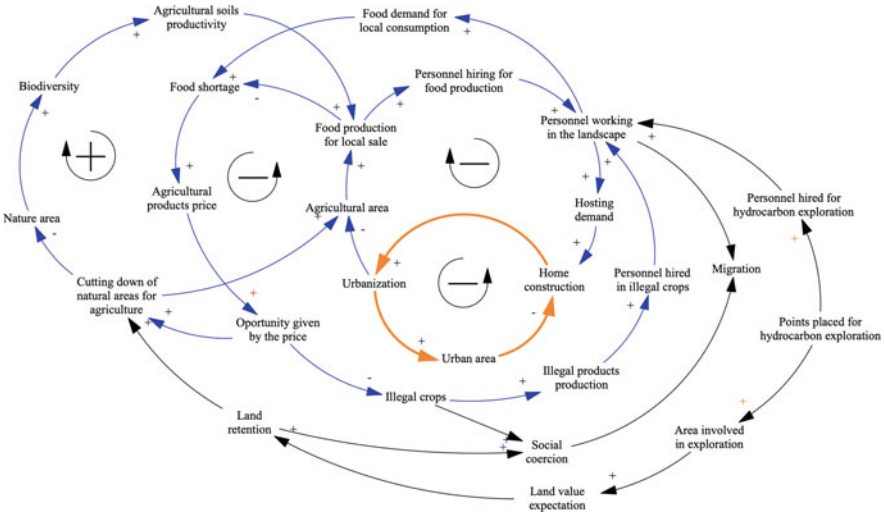


Fig. 5 Last loop analyzed in the causal diagram

In the first case and, according to the type of exploration carried out, the significant number of points used in the hydrocarbons exploration is considered, which increases the number of people working in this activity and the number of people who work in the landscape. This condition becomes changes in the local food supply and its prices, driving the urbanization in the landscape.

In the second case, the area involved in the hydrocarbon exploration, increases the expectation for the value of the land, increasing the land retention, which in turn leads to the forest clearing for agricultural activities.

This retention and the illegal activities lead to increased complaints of social coercion, which triggers the displacement or emigration of people from the region. In this way, while job opportunities in the landscape attract population to the landscape, social coercion forces them to seek opportunities in other places, from which it is concluded that exploration activity can weigh social coercion in those regions where the formality of land tenure and the presence of the state are scarce.

2.2 Forrester Diagram

From the causal diagram, were constructed the level and flow diagram presented in Fig. 6. The state variables were the four interested areas (natural areas, agricultural coverage, illegal crop coverage, and urban coverage), the four personnel hired to work in the landscape (for food production, goods and services, illicit crops, and exploration of hydrocarbons) and the local population, giving rise to the following differential system:

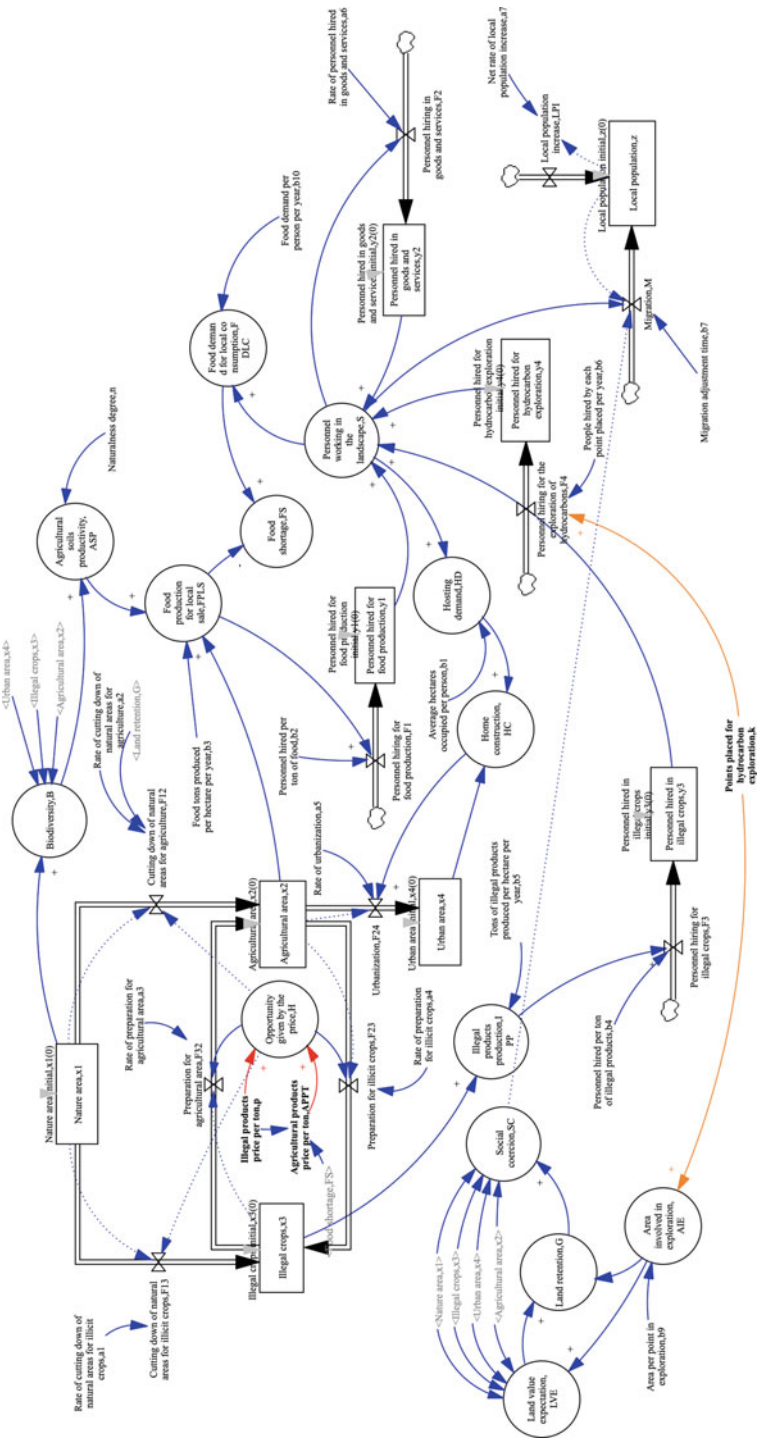


Fig. 6 Forrester diagram

$$\frac{dx_1}{dt} = \begin{cases} -a_1 x_1 & H < 0, G \leq 0 \\ -a_1 x_1 - a_2 x_1 & H < 0, G > 0 \\ -a_2 x_1 & H \geq 0 \end{cases} \tag{1}$$

$$\frac{dx_2}{dt} = \begin{cases} -a_4 x_2 - a_5 x_2 \frac{b_1(y_1+y_2+y_3+y_4)-x_4}{b_1(y_1+y_2+y_3+y_4)} & H < 0, G \leq 0 \\ a_2 x_1 - a_4 x_2 - a_5 x_2 \frac{b_1(y_1+y_2+y_3+y_4)-x_4}{b_1(y_1+y_2+y_3+y_4)} & H < 0, G > 0 \\ a_2 x_1 + a_3 x_3 - a_5 x_2 \frac{b_1(y_1+y_2+y_3+y_4)-x_4}{b_1(y_1+y_2+y_3+y_4)} & H \geq 0 \end{cases} \tag{2}$$

$$\frac{dx_3}{dt} = \begin{cases} a_1 x_1 + a_4 x_2 & H < 0 \\ a_3 x_3 & H \geq 0 \end{cases} \tag{3}$$

$$\frac{dx_4}{dt} = a_5 x_2 \frac{b_1(y_1 + y_2 + y_3 + y_4) - x_4}{b_1(y_1 + y_2 + y_3 + y_4)} \tag{4}$$

$$\frac{dy_1}{dt} = \frac{b_2 b_3 (1+n) x_1 x_2}{x_1 + x_2 + x_3 + x_4} \tag{5}$$

$$\frac{dy_2}{dt} = a_6(y_1 + y_2 + y_3 + y_4) \tag{6}$$

$$\frac{dy_3}{dt} = b_4 b_5 x_3 \tag{7}$$

$$\frac{dy_4}{dt} = b_6 k \tag{8}$$

$$\frac{dz}{dt} = a_7 z + \frac{y_1 + y_2 + y_3 + y_4 - z}{b_7} \left(1 - \frac{b_8 b_9 k + x_3}{x_1 + x_2 + x_3 + x_4}\right) \tag{9}$$

Where d/dt denotes the change in time, taking the year as the unit of time and the restrictions H and G are defined as follows:

$$H = \frac{-b_3 p(1+n) x_1 x_2}{b_{10}(y_1 + y_2 + y_3 + y_4)(x_1 + x_2 + x_3 + x_4)} \tag{10}$$

$$G = \frac{b_8 b_9 k}{x_1 + x_2 + x_3 + x_4} \tag{11}$$

All the variable and parameter abbreviations used in the model, and their measure units, are found in Table 1.

Table 1 Abbreviations and measure units used in the model

Name	Abbreviation	Unit	Type
Nature area	x1	Ha	State variable
Agricultural area	x2	Ha	State variable
Illegal crops	x3	Ha	State variable
Urban area	x4	Ha	State variable
Personnel hired for food production	y1	Person	State variable
Personnel hired in goods and services	y2	Person	State variable
Personnel hired in illegal crops	y3	Person	State variable
Personnel hired for hydrocarbon exploration	y4	Person	State variable
Local population	z	Person	State variable
Rate of cutting down of natural areas for illicit crops	a1	Percentage/year	Constant
Rate of cutting down of natural areas for agriculture	a2	Percentage/year	Constant
Rate of preparation for agricultural area	a3	Percentage/year	Constant
Rate of preparation for illicit crops	a4	Percentage/year	Constant
Rate of urbanization	a5	Percentage/year	Constant
Rate of personnel hired in goods and services	a6	Percentage/year	Constant
Net rate of local population increase	a7	Percentage/year	Constant
Average hectares occupied per person	b1	Ha/Person	Constant
Personnel hired per ton of food	b2	Person/Ton	Constant
Food tons produced per hectare per year	b3	Ton/(Ha*Year)	Constant
Personnel hired per ton of illegal products	b4	Person/Ton	Constant
Tons of illegal products produced per hectare per year	b5	Ton/(Ha*Year)	Constant
People hired by each point placed per year	b6	Person/(Point*Year)	Constant
Migration adjustment time	b7	Year	Constant
Area per point in exploration	b9	Ha/Point	Constant
Food demand per person per year	b10	Ton/(Person*Year)	Constant
Points placed for hydrocarbon exploration	k	Point	Constant
Naturalness degree	n	Dimensionless	Constant
Illegal products price per ton	p	US$/Ton	Constant
Land retention	G	Ha	Restriction
Opportunity given by the price	H	Dimensionless	Restriction

The mathematical model constructed makes the restriction H strictly negative, unless the willingness to pay for the products obtained in illicit crops disappears, in which case it will be null, while the restriction G is strictly positive, provided

that there are hydrocarbons exploration or an effective presence of the state in the exercise of formalizing land tenure.

3 Scenarios Assessment

From the mathematical model obtained, a simulation set of different scenarios was carried out, to review the prospective performance of the model and give some conclusions. The scenarios considered were: pristine landscape, agricultural landscape, and landscape of licit and illicit crops. Each scenario is evaluated with the presence and absence of hydrocarbon exploration, which is represented by placing exploration points on the landscape.

3.1 Scenario 1: Pristine Landscape

In this ideal scenario, the complete absence of socioeconomic dynamics is represented. This scenario was simulated considering that the initial coverage values are zero, except for the coverage of natural areas. The logging rates of natural areas have also been taken as zero and the local population is zero.

The simulations show a constant amount of natural areas, even with hydrocarbon exploration, see Fig. 7. Biodiversity simulates a value of 100 % and soil productivity of 125 %, while the production of licit and illicit crops appears at zero, see Fig. 7.

3.2 Scenario 2: Agricultural Landscape

In this scenario, a landscape with natural areas and socioeconomic dynamics revolving around agricultural production is represented. There is no presence of illicit crops. This scenario was simulated considering null the values of coverage of illicit crops, the clearing of natural areas for illicit crops, and the agricultural soils adaptation for illicit crops.

In the case in which there is no hydrocarbon exploration, there is less growth in urban areas from agricultural soils, while the coverage of natural areas remains intact (Fig. 8). The effect on biodiversity and soil productivity in this scenario leads to lower but constant values than those presented in the pristine scenario (Fig. 8).

On the other hand, with the hydrocarbons exploration in the landscape, the loss of natural areas and their conversion to agricultural and urban coverage is notorious, which in turn decreases both biodiversity and soil productivity (Fig. 9). Note also that the food production for local sale decreases drastically over time, thus showing a landscape that is highly dynamic due to its socioeconomic situation (Fig. 9).

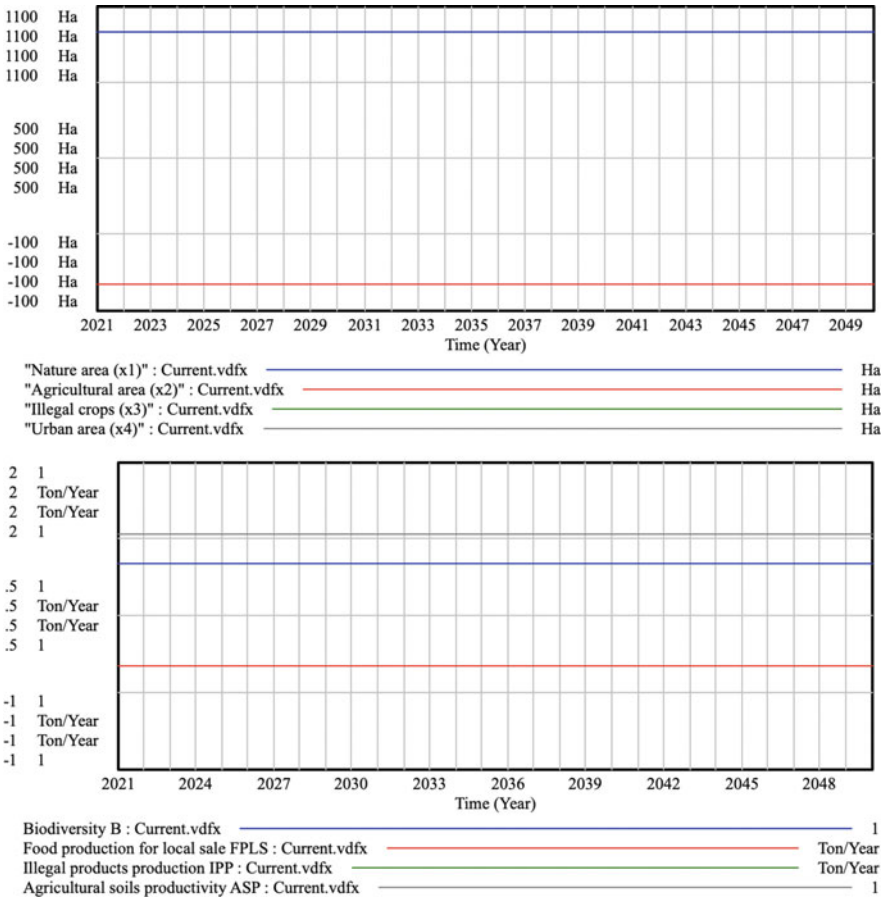


Fig. 7 Behavior of the covers, biodiversity, food production, production of illegal products and soil productivity in the pristine scenario

The social situation in this scenario shows that economic activity causes immigration to the landscape in conditions where coercion is non-existent (Fig. 8). However, with hydrocarbon exploration, there is a situation of constant coercion in the landscape becoming emigration.

3.3 Scenario 3: Landscape of Licit and Illicit Crops

In this scenario, a landscape with natural areas is represented, in which the socio-economic dynamics depend on agricultural production and illicit crops.

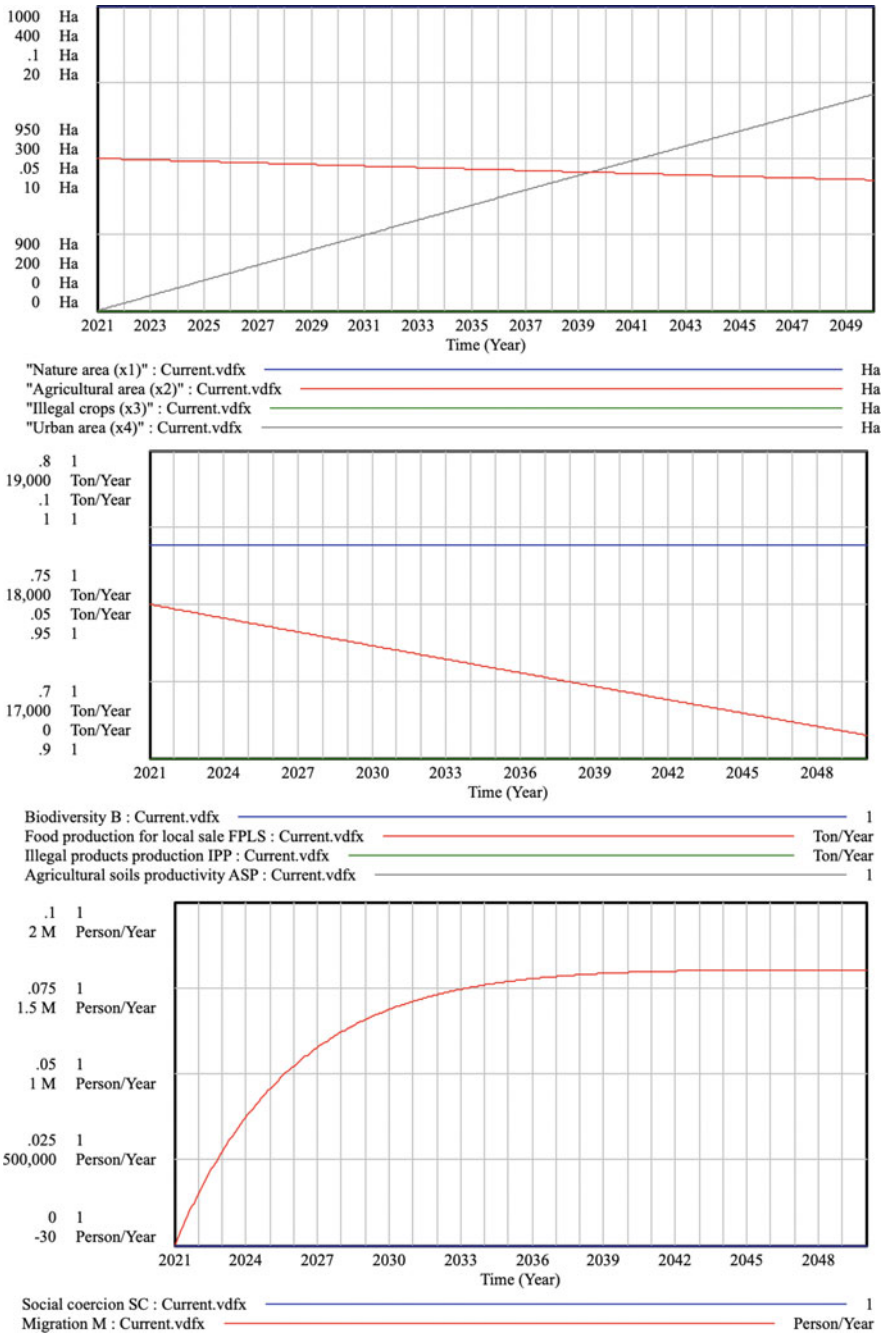


Fig. 8 Behavior of coverage, biodiversity, food production, production of illegal products and soil productivity in the agricultural landscape scenario

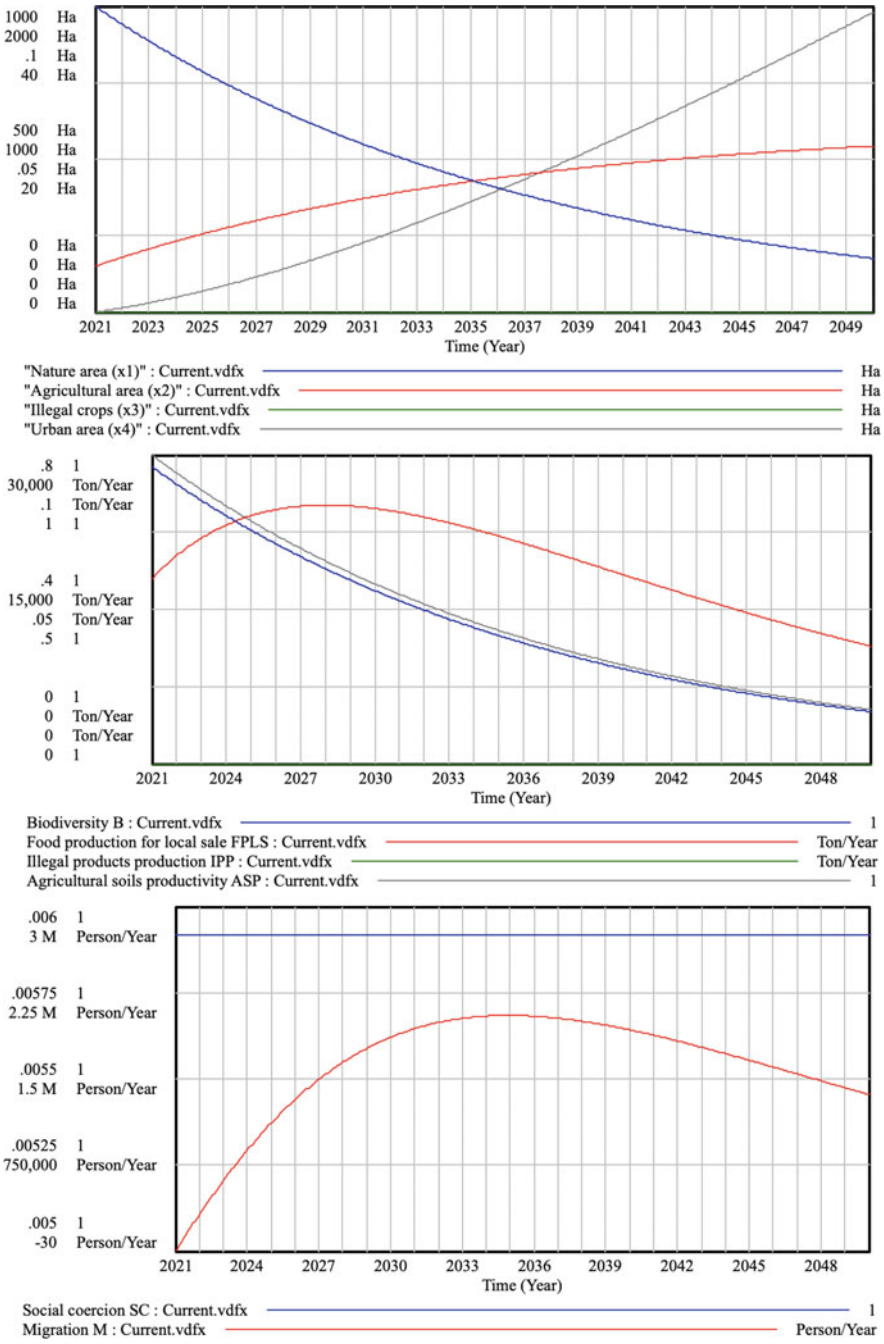


Fig. 9 Behavior of coverage, biodiversity, food production, production of illegal products and soil productivity in the agricultural landscape scenario with hydrocarbons exploration

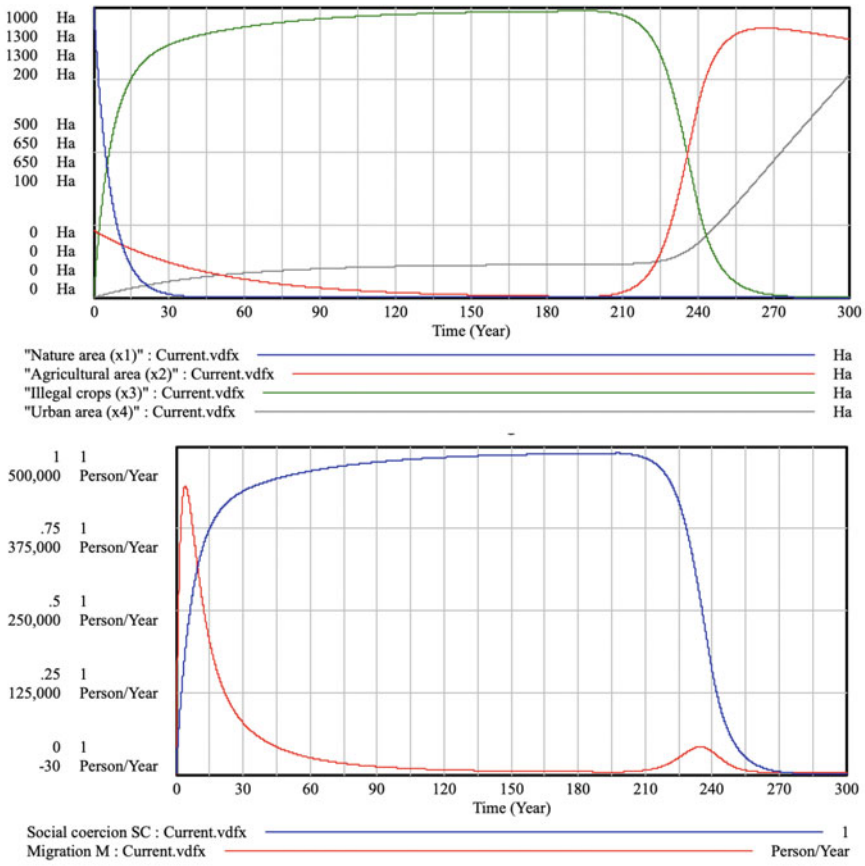


Fig. 10 Coverage and social behavior with licit and illicit crops

Without hydrocarbons exploration, illicit crops can grow above agricultural activity for a certain period (Fig. 10). However, with the increase in urban dynamics, the agricultural activity becomes dominant and the coverage of illicit crops tends to disappear. Natural areas are converted into lawful and illicit crop covers, which are also giving way to urban covers. In this way, a completely transformed landscape towards the urban-agricultural is obtained.

On the other hand, with hydrocarbons exploration, it is notable that the dynamics of behavior occur earlier in time, limiting the intensity and duration of illicit crops in the landscape, but accelerating its transformation towards urban agriculture (Fig. 11).

Regarding the social situation, it is notable that the presence of illicit crops in the landscape drastically increases coercion, which in turn increases the emigration of the landscape population. Once again, the hydrocarbon exploration activity acts as a social catalyst that does not prevent the existence of socioeconomic behavior but

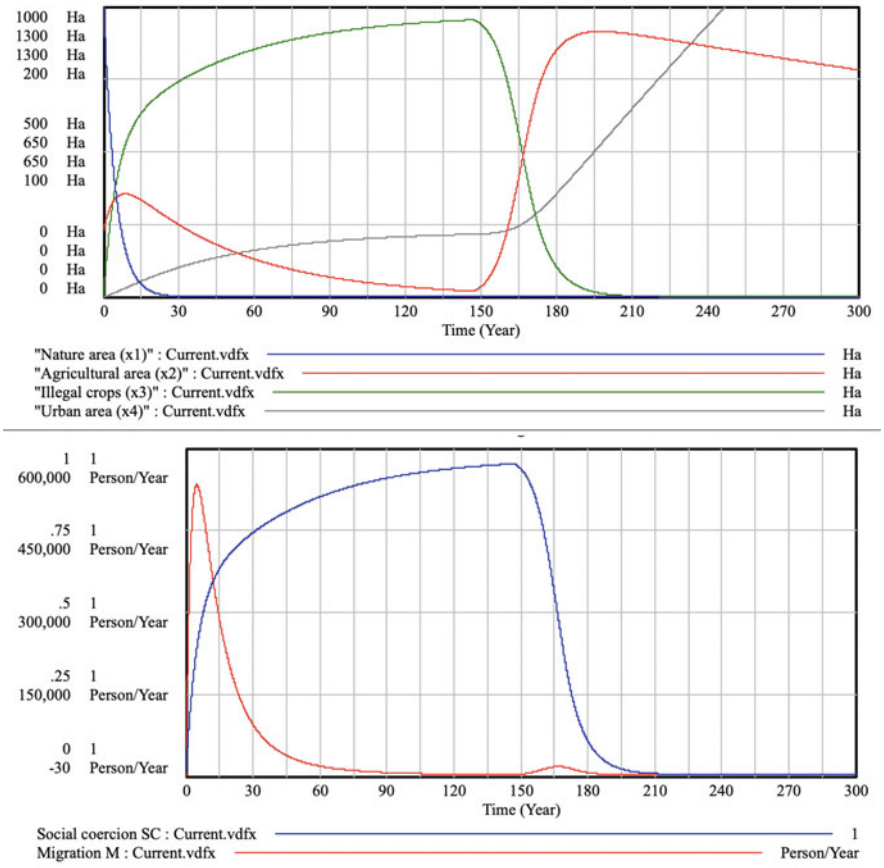


Fig. 11 Coverage and social behavior with licit and illicit crops, and hydrocarbons exploration

rather accelerates it. Check Fig. 10 for the absence of hydrocarbon exploration and Fig. 11 for its presence.

4 Conclusions

From the causal diagram, it is concluded that:

- There is a vicious circle in which natural covers and biodiversity tend to disappear, while the productivity of landscapes to produce and supply food to the landscape population is affected.
- There is a limit to the expansion of the coverage of illicit crops, given by the need to supply food to the population of the landscape.

- To maintain self-sufficiency, the urbanization of the landscape will have to impose a growth limit on agricultural coverage or it will have to import its supply.
- The landscape transformation is determined by the socioeconomic opportunity that the population finds in the landscape.

From the simulations carried out, it is concluded that:

- Exploration activity can intensify social coercion in the landscape, which will be weighted by informality in land tenure and the absence of state control.
- The cases in which the oil activity could bring undesirable consequences are the cases when the socio-economic dynamization is greater than the landscape is capable of assimilating.

References

1. D. A. Alemzero and N. Iqbal and S. Iqbal and M. Mohsin and N. J. Chukwuma and B. A. Shah, *Environ. Sci. Pollut. R.* **28** 5359-5371 (2021).
2. C. Rivard and G. Bordeleau and D. Lavoie and R. Lefebvre and P. Ladevèze and M. J. Duchesne and S. Séjourné and H. Crow and N. Pinet and V. Brake and A. Bouchedda and E. Gloaguen and J. M. E. Ahad X. and Malet and J. C. Aznar and M. Malo, *Mar. Petrol. Geol.* **100** 466–483 (2019).
3. W. U. Xishun and S. U. N.Zhangtao and Y. A. N. G. Tiantian and W. A. N. G. Shuling and Z. H. A. N. G. Wei, *Front. Mar. Geo.* **148** 1-17 (2020).
4. D. Cote and C. J. Morris and P. M. Regular and M. G. Piersiak, *Fish. Res.* **230**,105661 (2020).
5. P. Byakagaba and F. Mugagga and D. Nnakayima, *Ext. Ind. Soc.* **6** 358-366 (2019).
6. S. Kark and E. Brokovich and T. Mazor and N. Levin, *Conserv. Biol.* **29** 1573–1585 (2015).
7. T. Kumpula and B. C. Forbes and F. Stammler and N. Meschtyb, *Remote Sens-Basel* **4** 1046–1068 (2012)
8. M. L. Hornseth and K. E. Pigeon and D. MacNearney and T. A. Larsen and G. Stenhouse and J. Cranston and L. Finnegan, *Environ. Manage.* **62** 595–607 (2018).
9. M. P. J. A. Annevelink and J. A. J. Meesters and A. J. Hendriks, *Sci. Total. Environ.* **550** 431–438 (2016).
10. G. McDowell and J. D. Ford, *Appl. Geogr.* **46** 98-110 (2014).
11. E. E. Cordes and D. O. B. Jones and T. A. Schlacher and D. J. Amon and A. F. Bernardino and S. Brooke and R. Carney and D. M. DeLeo and K. M. Dunlop and E. G. Escobar-Briones and A. R. Gates and L. Génio and J. Gobin and L. A. Henry and S. Herrera and S. Hoyt and M. Joye and S. Kark and N. C. Mestre and A. Metaxas and S. Pfeifer and K. Sink and A. K. Sweetman and U. Witte, *Front. Env. Sc.* **4** (2016).
12. J. Sterman, *McGraw-Hill* (2000).
13. J. M. Redondo and C. Bustamante-Zamudio and J. Amador-Moncada and O. L. Hernandez-Manrique, *J. Phys. Conf. Ser.* **1414** 012010 (2019).

The Maximal Extension of the Strict Concavity Region on the Parameter Space for Sharma-Mittal Entropy Measures

R. P. Mondaini and S. C. de Albuquerque Neto

1 Introduction

The present work aims to complement a previous publication [1] where we have derived the Generalized Khinchin-Shannon (GKS) Inequalities [2] associated to Sharma-Mittal Entropy measures. We introduce here the result of maximal extension of the strict concavity region of this class of entropy measures. In Sect. 2, we present the state of art of the work with the Sharma-Mittal class as well as the Information measures associated to it. All the limit processes are then described carefully together with the consequences of these definitions and their properties with respect to the assumption of strict concavity. We think that the notation which has been adopted could appear as awkward, however it is very efficient for the derivation of all formulae to be presented here and specially for the proof of strict concavity of Sect. 3. A complete derivation is then done of the greatest lower bound of the successive epigraph regions, which leads to establish the maximal extension of the previously adopted strict concavity region of the scientific literature. In Sect. 4, we then derive some interesting additional matters to a subject already published in ref. [1], in terms of the difference between escort conditional probabilities and conditional escort probabilities. Some necessary development seems to be worthwhile here for a perfect understanding of the results based on the structure and properties of the probabilistic space and this is presented in appendices at the end of the paper.

R. P. Mondaini (✉) · S. C. de Albuquerque Neto
Federal University of Rio de Janeiro, Centre of Technology, COPPE, Rio de Janeiro, Brazil

R. P. Mondaini (ed.), *Trends in Biomathematics: Stability and Oscillations in Environmental, Social, and Biological Models*,
https://doi.org/10.1007/978-3-031-12515-7_15

2 The Sharma-Mittal Class of Entropy Measures

The usual expression of the Sharma-Mittal entropy measures [3], is given by:

$$(SM)^{(s,r)}_{j_1\ldots j_t} = \frac{(\alpha_{j_1\ldots j_t})^{\frac{1-r}{1-s}} - 1}{1-r} \tag{1}$$

These entropies are here defined on arrays of t columns of m rows each and the $\alpha_{j_1\ldots j_t}$ symbols are written as:

$$\alpha^{(s)}_{j_1\ldots j_t} = \sum_{a_1,\ldots,a_t} \Big(p_{j_1\ldots j_t}(a_1,\ldots,a_t)\Big)^s \tag{2}$$

where $p_{j_1\ldots j_t}(a_1,\ldots,a_t)$ stands for the probability of occurrence of t-sets of ordered values $(a_1,\ldots,a_t)$ in t columns $j_1,\ldots,j_t$ of a distribution obtained from an array of m rows and n columns (see Appendix 1).

All the values $a_1,\ldots,a_t$, are running over the values $1,\ldots,W$ and the j-values are ordered as [4]:

$$j_1 < j_2 < \ldots < j_n$$

and we have,

$$\begin{aligned} &j_1 = 1,\ldots,n-t+1\\ &j_2 = j_1+1,\ldots,n-t+2\\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ &j_{t-1} = j_{t-2}+1,\ldots,n-1\\ &j_t = j_{t-1}+1,\ldots,n \end{aligned} \tag{3}$$

There are then $\binom{n}{t} = \frac{n!}{t!(n-t)!}$ geometrical objects $p_{j_1\ldots j_t}$ and each of them has W^t correspondents $p_{j_1\ldots j_t}(a_1,\ldots,a_t)$.

For an equiprobable distribution [4],

$$p_{j_1\ldots j_t}(a_1,\ldots,a_t) = W^{-t}\,, \tag{4}$$

we then get

$$\alpha_{j_1\ldots j_t}(W^{-t}) = W^t(W^{-t})^s = W^{t(1-s)}\,,\ \forall t \tag{5}$$

and the corresponding value of the $(SM)_{j_1\ldots j_t}$ entropy will be written as:

$$(SM)^{(r)}_{j_1\ldots j_t}(W^{-t}) = \frac{\left(W^{t(1-s)}\right)^{\frac{1-r}{1-s}} - 1}{1-r} = \frac{W^{t(1-r)} - 1}{1-r}. \tag{6}$$

We also introduce an associated Information measure:

$$I^{(s,r)}_{j_1\ldots j_t} = -\frac{(SM)_{j_1\ldots j_t}}{(\alpha_{j_1\ldots j_t})^{\frac{1-r}{1-s}}} \tag{7}$$

For an equiprobable distribution, Eq. (4), this can be written as:

$$I^{(s,r)}_{j_1\ldots j_t}(W^{-t}) = -\frac{\frac{W^{t(1-r)} - 1}{1-r}}{W^{t(1-r)}} = \frac{1 - W^{t(1-r)}}{(1-r)W^{t(1-r)}} \tag{8}$$

We then list some special cases of entropies of the Sharma-Mittal class [3, 5–7] together with their equiprobable versions:

(a) The Havrda-Charvat entropy measure [5], with $r = s$

$$(HC)^{(s)}_{j_1\ldots j_t} = \frac{\alpha_{j_1\ldots j_t} - 1}{1-s} \tag{9}$$

$$(HC)^{(s)}_{j_1\ldots j_t}(W^{-t}) = \frac{W^{t(1-s)} - 1}{1-s} \tag{10}$$

(b) The Landsberg-Vedral entropy measure [6], with $r = 2 - s$

$$(LV)^{(s)}_{j_1\ldots j_t} = \frac{\alpha_{j_1\ldots j_t} - 1}{(1-s)\alpha_{j_1\ldots j_t}} = \frac{(HC)^{(s)}_{j_1\ldots j_t}}{\alpha_{j_1\ldots j_t}} \tag{11}$$

$$(LV)^{(s)}_{j_1\ldots j_t}(W^{-t}) = \frac{W^{t(1-s)} - 1}{(1-s)W^{t(1-s)}} = -I^{(r=s)}_{j_1\ldots j_t} \tag{12}$$

(c) The Renyi entropy measure [7]:

$$R^{(s)}_{j_1\ldots j_t} = \lim_{r\to 1}(SM)^{(s,r)}_{j_1\ldots j_t} = \frac{\log \alpha_{j_1\ldots j_t}}{1-s} \tag{13}$$

$$R^{(s)}_{j_1\ldots j_t}(W^{-t}) = \frac{\log W^{t(1-s)}}{1-s} = \frac{t(1-s)\log W}{1-s} = t\log W \tag{14}$$

(d) The "nonextensive Gaussian" entropy measure [8]:

$$G^{(r)}_{j_1\ldots j_t} = \lim_{s\to 1}(SM)^{(s,r)}_{j_1\ldots j_t} = \frac{e^{(1-r)(GS)_{j_1\ldots j_t}} - 1}{1-r} \tag{15}$$

$$G^{(r)}_{j_1\ldots j_t}(W^{-t}) = \frac{e^{t(1-r)\log W} - 1}{1-r} \tag{16}$$

where

$$(GS)_{j_1\ldots j_t} = -\sum_{a_1,\ldots,a_t} p_{j_1\ldots j_t}(a_1,\ldots,a_t)\log p_{j_1\ldots j_t}(a_1,\ldots,a_t) \tag{17}$$

$$(GS)_{j_1\ldots j_t}(W^{-t}) = t\log W \tag{18}$$

is the Gibbs-Shannon entropy measure, which is also obtained as the convenient limit in all previous entropy measures.

$$\lim_{s\to 1}(HC)^{(s)}_{j_1\ldots j_t} = \lim_{s\to 1}(LV)^{(s)}_{j_1\ldots j_t} = \lim_{s\to 1} R^{(s)}_{j_1\ldots j_t} = \lim_{r\to 1} G^{(r)}_{j_1\ldots j_t} = (GS)_{j_1\ldots j_t} \tag{19}$$

From Eqs. (1) and (2) the Generalized Khinchin-Shannon inequalities can be written according to ref. [1] as:

$$1 + (1-r)(SM)^{(s,r)}_{j_1\ldots j_t} \leq \prod_{l=1}^{t}\left[1 + (1-r)(SM)^{(s,r)}_{j_l}\right] \tag{20}$$

We can also then write for the information measure in [7]:

$$1 + (1-r)I^{(s,r)}_{j_1\ldots j_t} \geq \prod_{l=1}^{t}\left[1 + (1-r)I^{(s,r)}_{j_l}\right] \tag{21}$$

For the special cases of the Sharma-Mittal class, Eqs. (9), (11), (13), (15), and (17), we have,

$$1 + (1-s)(HC)^{(s)}_{j_1\ldots j_t} \leq \prod_{l=1}^{t}\left[1 + (1-s)(HC)^{(s)}_{j_l}\right] \tag{22}$$

$$1 + (1-s)(LV)^{(s)}_{j_1 \ldots j_t} \leq \prod_{l=1}^{t} \left[1 + (1-s)(LV)^{(s)}_{j_l} \right] \tag{23}$$

$$R^{(s)}_{j_1 \ldots j_t} \leq \sum_{l=1}^{t} R^{(s)}_{j_l} \tag{24}$$

$$1 + (1-r)G^{(r)}_{j_1 \ldots j_t} \leq \prod_{l=1}^{t} \left[1 + (1-r)G^{(r)}_{j_l} \right] \tag{25}$$

$$(GS)_{j_1 \ldots j_t} \leq \sum_{l=1}^{t} (GS)_{j_l} \tag{26}$$

The region of the parameter space corresponding to strict concavity of the Sharma-Mittal class of entropies is usually presented in the literature as the gray region in Fig. 1 or $C = \{(s, r) | r \geq s > 0\}$. This region is the epigraph region of the half straight line $r = s > 0$, corresponding to the Havrda-Charvat entropy, Eqs. (9) and (2).

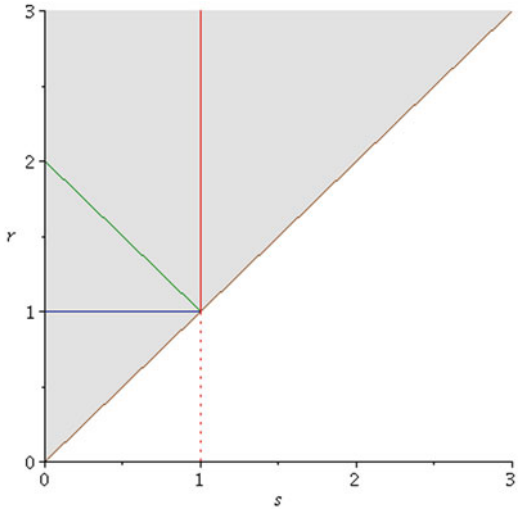


Fig. 1 The gray region is the epigraph of the brown half straight line ($r = s$) corresponding to the Havrda-Charvat entropy. This region is also assumed to be the strict concavity region in the literature. The The blue ($0 \leq s < 1, r = 1$), green ($0 \leq s < 1, 1 < r \leq 2$) segments of straight line do correspond to Renyi and Landsberg-Vedral entropies, respectively. The red half straight line ($r > 1, s = 1$) stands for the "Gaussian" entropy which has also been defined by Sharma-Mittal [3]. The point ($r = 1, s = 1$) does correspond to the Gibbs-Shannon entropy

3 The Maximal Extension of the Strict Concavity Region in the Parameter Space of Sharma-Mittal (SM) Entropy Measures

To undertake the analysis of maximal extension of the strict concavity region for SM entropies, we shall follow the techniques of construction of the probabilistic space in Appendix 1. In the next section, we also give some examples as numerical applications of the methods introduced here.

The requirement for strict concavity of the surface representing a multivariate function is the negative definiteness of the quadratic form associated to its Hessian matrix [1, 11]. This means that the principal minors of the Hessian matrix should be negative and positive alternately (negative those of odd order and positive those of even order).

We shall now introduce the Hessian matrix of the Sharma-Mittal class of entropies, Eqs. (1), (2), and (3). Its first derivative is given by:

$$\frac{\partial (SM)_{j_1 \ldots j_t}}{\partial p_{j_1 \ldots j_t}(a_1^{q_\nu}, \ldots, a_t^{q_\nu})} = \frac{s}{1-s} (\alpha_{j_1 \ldots j_t})^{\frac{s-r}{1-s}} \left(p_{j_1 \ldots j_t}(a_1^{q_\nu}, \ldots, a_t^{q_\nu}) \right)^{s-1} \tag{27}$$

A generic element of the Hessian matrix could then be written as

$$\begin{aligned} H_{q_\nu q_\xi} &= \frac{\partial^2 (SM)_{j_1 \ldots j_t}}{\partial p_{j_1 \ldots j_t}(a_1^{q_\nu}, \ldots, a_t^{q_\nu}) \partial p_{j_1 \ldots j_t}(a_1^{q_\xi}, \ldots, a_t^{q_\xi})} \\ &= s (\alpha_{j_1 \ldots j_t})^{\frac{s-r}{1-s}} \left(p_{j_1 \ldots j_t}(a_1^{q_\nu}, \ldots, a_t^{q_\nu}) \right)^{s-2} \\ &\quad \cdot \left[\frac{s(s-r)}{(1-s)^2} \frac{p_{j_1 \ldots j_t}(a_1^{q_\nu}, \ldots, a_t^{q_\nu})}{p_{j_1 \ldots j_t}(a_1^{q_\xi}, \ldots, a_t^{q_\xi})} \hat{p}_{j_1 \ldots j_t}(a_1^{q_\xi}, \ldots, a_t^{q_\xi}) - \delta_{\nu\xi} \right] \end{aligned} \tag{28}$$

where $\delta_{\nu\xi}$ is the Kronecker symbol. From Eq. (28) and Appendix 1, the principal minors are then given by

$$\begin{aligned} \det H_{q_\nu q_\xi}(\nu, \xi = 1, \ldots, k) &= (-1)^{k-1} s^k (\alpha_{j_1 \ldots j_t})^{k\left(\frac{s-r}{1-s}\right)} \left[\prod_{\nu=1}^{k} p_{j_1 \ldots j_t}(a_1^{q_\nu}, \ldots, a_t^{q_\nu}) \right]^{s-2} \\ &\quad \cdot \left[\frac{s(s-r)}{(1-s)^2} \sum_{\nu=1}^{k} \hat{p}_{j_1 \ldots j_t}(a_1^{q_\nu}, \ldots, a_t^{q_\nu}) - 1 \right], \quad k = 1, \ldots, M, 1 \leq M \leq m, \end{aligned} \tag{29}$$

where $\hat{p}_{j_1 \ldots j_t}(a_1^{q_\nu}, \ldots, a_t^{q_\nu})$ is the escort probability associated to $p_{j_1 \ldots j_t}(a_1^{q_\nu}, \ldots, a_t^{q_\nu})$ and we have,

$$\hat{p}_{j_1\ldots j_t}(a_1^{q_\nu},\ldots,a_t^{q_\nu}) = \frac{\left(p_{j_1\ldots j_t}(a_1^{q_\nu},\ldots,a_t^{q_\nu})\right)^s}{\sum\limits_{b_1^{q_\nu},\ldots,b_t^{q_\nu}}\left(p_{j_1\ldots j_t}(b_1^{q_\nu},\ldots,b_t^{q_\nu})\right)^s} \tag{30}$$

with $a_j^{q_\nu}; b_j^{q_\nu} = 1,\ldots,W; \nu = 1,\ldots,k; k = 1,\ldots,M.$

$$\sum_{a_1^{q_\nu},\ldots,a_t^{q_\nu}} \hat{p}_{j_1\ldots j_t}(a_1^{q_\nu},\ldots,a_t^{q_\nu}) = 1 \tag{31}$$

From Eqs. (29) and (30), and Eqs. (42)–(46) of Appendix 1, and Eqs. (64)–(68) of Appendix 2, the requirement of strict concavity could be given through the second square bracket of Eq. (29) and we can then write,

$$\left[\frac{s(s-r)}{(1-s)^2}\frac{\sum\limits_{\mu=1}^{k}\left(\frac{q_\mu}{m}\right)^s}{\sum\limits_{\mu=1}^{m}\left(\frac{q_\mu}{m}\right)^s} - 1\right] < 0, \quad k = 1,\ldots,M. \tag{32}$$

The curves $r(s)$ given by

$$\frac{s(s-r)}{(1-s)^2}\sigma_k(s) - 1 = 0, \quad k = 1,\ldots,M, \tag{33}$$

with

$$\sigma_k(s) = \frac{\sum\limits_{\mu=1}^{k}\left(\frac{q_\mu}{m}\right)^s}{\sum\limits_{\mu=1}^{m}\left(\frac{q_\mu}{m}\right)^s} \tag{34}$$

can be written as

$$r_k(s) = s - \frac{(1-s)^2}{s\sigma_k(s)}, \quad k = 1,\ldots,M. \tag{35}$$

The epigraph regions of these curves can be written as

$$C \cup C_k = \{(s,r)|r \geq s > 0\} \cup \left\{(s,r)\,\middle|\, s > r \geq s - \frac{(1-s)^2}{s\sigma_k(s)}\right\}, \quad k = 1,\ldots,M. \tag{36}$$

At the end of Sect. 2, we have emphasized that the epigraph region of the curve $r = s$ is $C = \{(s,r)|r \geq s > 0\}$, which is usually taken as the region of strict

concavity. Here we define the extended region of strict concavity as the epigraph region of the highest curve of the set (35)

$$r_m(s) = 2 - \frac{1}{s} \tag{37}$$

which is given by the union set:

$$C_{max} = C \cup C_m = \{(s,r)|r \geq s > 0\} \cup \left\{(s,r) \,\middle|\, s > r \geq 2 - \frac{1}{s}\right\}. \tag{38}$$

This extended region is depicted in Fig. 2 as the gray region.

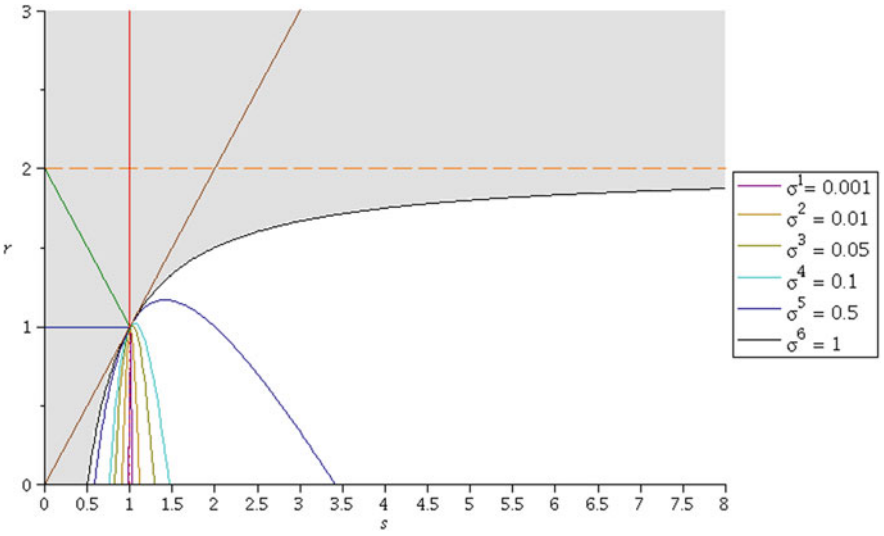


Fig. 2 The blue, green, red and brown lines do correspond to Renyi, Landsberg-Vedral, "Non-extensive" Gaussian and Havrda-Charvat entropies, respectively. The union of the regions $C = \{(s,r)|\, r \geq s > 0\}$ and $\mathbf{C}_m = \{(s,r)|\, s > r \geq 2 - 1/s\}$ is the epigraph region of the black line and corresponds to the extended region of strict concavity. Some $r_k(s)$ functions are also depicted and they correspond to $k = 1, 2, 3, 4, 5$, $m = 6$. We note that $r_6(0) = 0.5$ and the curve $r_6(s)$ is asymptotic to the straight line $r = 2$

Fig. 3 An example of a 8×3 array of 3-sets of amino acids

$$\begin{pmatrix} A\ C\ Y \\ A\ C\ Y \\ A\ C\ T \\ A\ T\ C \\ A\ Y\ C \\ A\ C\ Y \\ T\ C\ A \\ Y\ A\ T \end{pmatrix}$$

4 An Example of Systematic Derivation of $\sigma_k(s)$ Curves from Data Obtained from the Alignement of Protein Domains

In order to give an example of the construction of $\sigma_k(s)$ curves, we shall use arrays of t-amino acids which have been worked intensively on the presentation of results during the 21st BIOMAT International Symposium [10].

We shall provide an example of a 8×3 array of 3-sets of amino acids with 8 rows (Fig. 3). This example will be convenient for readers who intend to work with classification of amino acids distributions. According to equations of Appendices 1 and 2 there are 6 different groups on this array of equal t-sets of amino acids, $q_1, \ldots, q_6$, in the array (38). The symbol $\alpha_{j_1 j_2 j_3}$ could then be written as:

$$\begin{aligned} \alpha_{j_1 j_2 j_3} = &\left(p_{j_1 j_2 j_3}(A^{q_1}, C^{q_1}, Y^{q_1})\right)^s + \left(p_{j_1 j_2 j_3}(A^{q_2}, C^{q_2}, T^{q_2})\right)^s \\ &+ \left(p_{j_1 j_2 j_3}(A^{q_3}, T^{q_3}, C^{q_3})\right)^s + \left(p_{j_1 j_2 j_3}(A^{q_4}, Y^{q_4}, C^{q_4})\right)^s \\ &+ \left(p_{j_1 j_2 j_3}(T^{q_5}, C^{q_5}, A^{q_5})\right)^s + \left(p_{j_1 j_2 j_3}(Y^{q_6}, A^{q_6}, T^{q_6})\right)^s . \end{aligned} \tag{39}$$

We now refer to Eqs. (42)–(44) of Appendix 1 and we get

$$\alpha_{j_1 j_2 j_3} = \left(\frac{3}{8}\right)^s + \left(\frac{1}{8}\right)^s + \left(\frac{1}{8}\right)^s + \left(\frac{1}{8}\right)^s + \left(\frac{1}{8}\right)^s + \left(\frac{1}{8}\right)^s . \tag{40}$$

From Eqs. (34) and (42), we can also write for the $\sigma_k(s)$ functions:

$$\sigma_1(s) = \frac{\left(\frac{3}{8}\right)^s}{\alpha_{j_1 j_2 j_3}}\ ;\ \sigma_2(s) = \frac{\left(\frac{3}{8}\right)^s + \left(\frac{1}{8}\right)^s}{\alpha_{j_1 j_2 j_3}}\ ;\ \sigma_3(s) = \frac{\left(\frac{3}{8}\right)^s + \left(\frac{1}{8}\right)^s + \left(\frac{1}{8}\right)^s}{\alpha_{j_1 j_2 j_3}}\ ;$$

$$\sigma_4(s) = \frac{\left(\frac{3}{8}\right)^s + \left(\frac{1}{8}\right)^s + \left(\frac{1}{8}\right)^s + \left(\frac{1}{8}\right)^s}{\alpha_{j_1 j_2 j_3}}\ ;\ \sigma_5(s) = \frac{\left(\frac{3}{8}\right)^s + \left(\frac{1}{8}\right)^s + \left(\frac{1}{8}\right)^s + \left(\frac{1}{8}\right)^s + \left(\frac{1}{8}\right)^s}{\alpha_{j_1 j_2 j_3}}\ ;$$

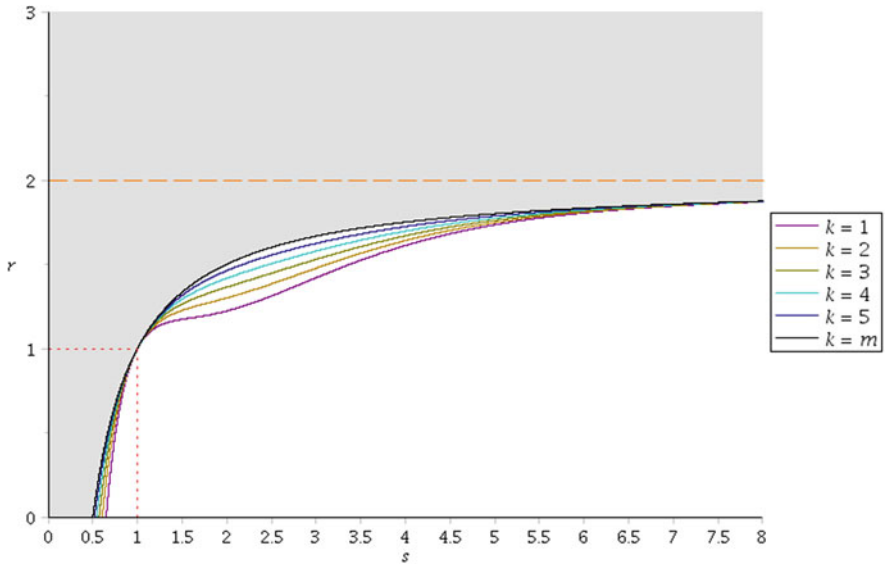


Fig. 4 The $r_k(s)$ curves and the structure of extended region of strict concavity as obtained from a 8×3 array of 3-sets of amino acids

$$\sigma_6(s) = \frac{\left(\frac{3}{8}\right)^s + \left(\frac{1}{8}\right)^s + \left(\frac{1}{8}\right)^s + \left(\frac{1}{8}\right)^s + \left(\frac{1}{8}\right)^s + \left(\frac{1}{8}\right)^s}{\alpha_{j_1 j_2 j_3}} = 1\,. \tag{41}$$

In Fig. 4, the related curves $r_k(s)$, Eq. (35), are depicted. As has been emphasized at the end of Sect. 3, Sect. 3, the epigraph region of the curve corresponding to $\sigma_6(s) = 1$ and asymptotic to $r = 2$, or $r_6(s) = 2 - 1/s$, does correspond to the extended region of strict concavity.

5 Concluding Remarks

In this work we have chosen to present in detail the extensions of the region of strict concavity on the space of parameters of the Sharma-Mittal class of entropy measures. We have emphasized the structure of the parameter space and we believe that this will be very useful for working with the several special cases of entropies in models of generalized Statistical Mechanics. Some special (s-r) regions of the parameter space and the curves $r(s)$ on them could specify models of interest on the study of diseases and their interconnection from the viewpoint of their evolution in terms of entropy values [12]. A detailed study has been undertaken on the Jaccard-like Symbol and its usefulness for analysing the distributions of amino acids in protein domain families [9]. A forthcoming comprehensive review will be

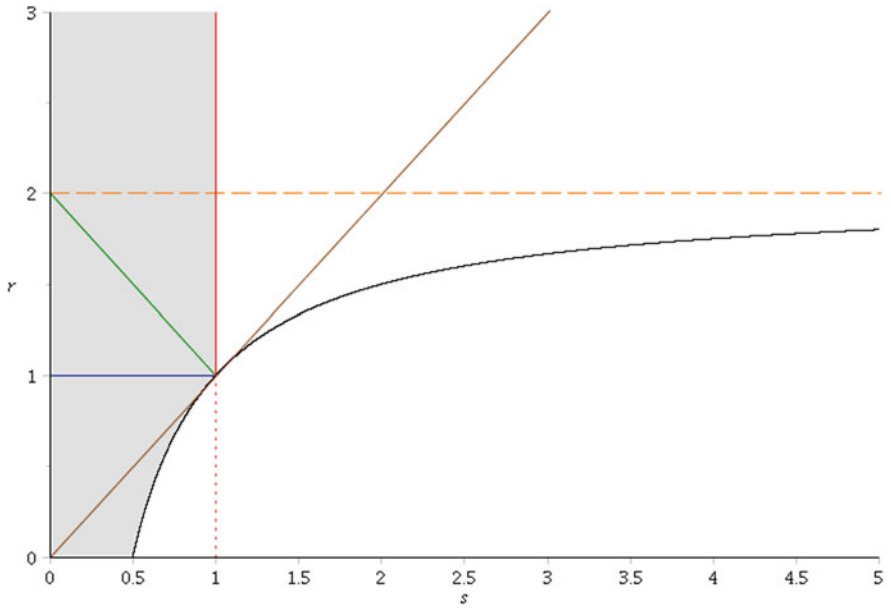


Fig. 5 The intersection of the extended region of strict concavity (Fig. 2) with the region of fully synergetic distributions of amino acids in arrays of m rows and n columns (inherited from the Hölder inequality). There are also synergetic distributions (non-Hölder specifically configurations of amino acids) on the complementary region $s > 1, r \geq 2 - 1/s$

also published with a study of the representative surfaces of entropy measures as obtained from a systematic parametrization method [9].

As has been emphasized in Appendix 2, the application of the additional criterium of fully synergetic distributions of amino acids in $m \times n$ arrays of m rows and n columns, will restrict still more the feasible region of the Sharma-Mittal entropy values associated to these arrays. This is done through the Generalized Khinchin-Shannon inequalities (Eq. (75)) and [13].

In Fig. 5, we then depict a gray region which is the intersection of the extended region of strict concavity (Fig. 2) with the region corresponding to full synergy of the probability distributions. Among all entropy measures belonging to the Sharma-Mittal entropy class and defined in Sect. 1, the "Gaussian non-extensive" entropy measure is the only one which remains for $r > 2$. This is taken as an insight to undertake the study of evolution of protein domain families and clans working with this entropy measure.

Appendix 1: The Construction and Properties of the Probabilistic Space

We fix the ideas for working with a probabilistic space by considering arrays of amino acids with m rows and n columns. We will then be able to characterize the protein domain families and the formation of several clans according the identification provided by expert biologists through alignment techniques. On some previous works [4], we have stressed that each protein domain family has as its representative at least one array of m rows and n columns. In this appendix, we summarize the usual properties of geometric objects associated to these arrays. These seem to be essential for the description of the structure of the probabilistic space and they also help to unveil some of its properties which have so far been unknown in the scientific literature.

First of all we should stress that the symbols $\alpha_{j_1\ldots j_t}$ of Eq. (2), could also be written as

$$\alpha_{j_1\ldots j_t} = \sum_{a_1^{q_\mu},\ldots,a_t^{q_\mu}} \left(p_{j_1\ldots j_t}(a_1^{q_\mu},\ldots,a_t^{q_\mu})\right)^s \leftrightarrow \sum_{\mu=1}^{m} \left(\frac{q_\mu}{m}\right)^s \; t{=}1,\ldots,n,\, k{=}1,\ldots,M,\, 1 \le M \le m. \tag{42}$$

where M is the number of μ-groups of amino acids and q_μ stands for the number of equal t-sets of the amino acids contained in the μth-group.

The significance of Eq. (42) comes from the definition of probability of occurrence:

$$p_{j_1\ldots j_t}(a_1^{q_\mu},\ldots,a_t^{q_\mu}) = \frac{n_{j_1\ldots j_t}(a_1^{q_\mu},\ldots,a_t^{q_\mu})}{m}\,, \tag{43}$$

and we have the correspondence:

$$q_\mu \;\leftrightarrow\; n_{j_1\ldots j_t}(a_1^{q_\mu},\ldots,a_t^{q_\mu}). \tag{44}$$

Since the number of groups is equal to M, this also means that there is a maximum of m different groups of t-sets or,

$$q_\mu \le m\,. \tag{45}$$

From eqs.(42) and (43), we write

$$\sum_{a_1^{q_\mu},\ldots,a_t^{q_\mu}} \frac{n_{j_1\ldots j_t}(a_1^{q_\mu},\ldots,a_t^{q_\mu})}{m} = 1 \;\leftrightarrow\; \sum_{\mu=1}^{m}\left(\frac{q_\mu}{m}\right) = 1\,. \tag{46}$$

We now pass to the probability unit vector of W components and to their generalization of geometric objects of $(W)^t$ components. From now on, in order to alleviate the notation, we do not take in consideration the super indices q_μ. For $t = 1$, $p_{j_1}(a_1)$ could be represented as a unit vector of W components:

$$p_{j_1}^{\mathsf{T}} = \left(p_{j_1}(1), \ldots, p_{j_1}(W)\right). \tag{47}$$

The escort vector associated to this vector is

$$\hat{p}_{j_1}^{\mathsf{T}} = \left(\frac{\left(p_{j_1}(1)\right)^s}{\alpha_{j_1}}, \ldots, \frac{\left(p_{j_1}(W)\right)^s}{\alpha_{j_1}}\right) \tag{48}$$

where

$$\alpha_{j_1} = \sum_{a_1} \left(p_{j_1}(a_1)\right)^s \tag{49}$$

and

$$1 = \sum_{a_1} \hat{p}_{j_1}(a_1) = \sum_{a_1} \frac{\left(p_{j_1}(a_1)\right)^s}{\alpha_{j_1}} \tag{50}$$

A geometric object of $(W)^2$ components could be also defined through a "★-product":

$$p_{j_1} \bigstar p_{j_2} = p_{j_1|j_2}\, p_{j_2}^{\mathsf{T}} \tag{51}$$

where $p_{j_1|j_2}$ will transform as a column vector.

The structure of this matrix product is given by

$$(W \times 1)(1 \times W) = (W \times W)$$

and their components could be written as:

$$p_{j_1} \bigstar p_{j_2} = \begin{pmatrix} p_{j_1 j_2}(1|1)p_{j_2}(1) & \ldots & p_{j_1 j_2}(1|W)p_{j_2}(W) \\ \vdots & \ddots & \vdots \\ p_{j_1 j_2}(W|1)p_{j_2}(1) & \ldots & p_{j_1 j_2}(W|W)p_{j_2}(W) \end{pmatrix} \tag{52}$$

where $p_{j_1 j_2}(a_1|a_2)$ are the components of the column vector of conditional probability $p_{j_1|j_2}$ of the distribution in column j_1 with a previous knowledge of the distribution in column j_2.

Since the components of the matrix 52 are joint probabilities, we then write,

$$p_{j_1} \bigstar p_{j_2} = \begin{pmatrix} p_{j_1 j_2}(1,1) & \dots & p_{j_1 j_2}(1,W) \\ \vdots & \ddots & \vdots \\ p_{j_1 j_2}(W,1) & \dots & p_{j_1 j_2}(W,W) \end{pmatrix} \tag{53}$$

where $p_{j_1 j_2}(a_1, a_2)$ are the components of the joint probability of occurrence of the 2-set (a_1, a_2) in columns j_1 and j_2.

Analogously a geometric object of $(W)^3$ components could be also defined by the $\bigstar$-product,

$$p_{j_1} \bigstar p_{j_2} \bigstar p_{j_3} = p_{j_1 j_2} \bigstar p_{j_3}^{\mathsf{T}} = p_{j_1 j_2 | j_3}\, p_{j_3}^{\mathsf{T}} = p_{j_1 j_2 j_3} \tag{54}$$

and the structure of this matrix product is

$$(W \times W \times 1)(1 \times W) = (W \times W \times W)\,.$$

The related $(W)^3$ components could be written as:

$$p_{j_1} \bigstar p_{j_2} \bigstar p_{j_3} = \begin{pmatrix} p_{j_1 j_2 j_3}(1,1,1) & \dots & p_{j_1 j_2 j_3}(1,1,W) \\ \vdots & \ddots & \vdots \\ p_{j_1 j_2 j_3}(W,1,1) & \dots & p_{j_1 j_2 j_3}(W,1,W) \\ \vdots & \ddots & \vdots \\ p_{j_1 j_2 j_3}(1,W,1) & \dots & p_{j_1 j_2 j_3}(1,W,W) \\ \vdots & \ddots & \vdots \\ p_{j_1 j_2 j_3}(W,W,1) & \dots & p_{j_1 j_2 j_3}(W,W,W) \end{pmatrix} \tag{55}$$

and so on and so forth for an object of $(W)^t$ components:

$$p_{j_1} \bigstar p_{j_2} \bigstar \dots \bigstar p_{j_t}^{\mathsf{T}} = p_{j_1 j_2 \dots j_{t-1}} \bigstar p_{j_t}^{\mathsf{T}} = p_{j_1 j_2 \dots j_{t-1} | j_t}\, p_{j_t}^{\mathsf{T}} = p_{j_1 j_2 \dots j_t} \tag{56}$$

with the structure of the matrix product given generally by

$$(W \times W \times \dots \times W \times 1)(1 \times W) = \underbrace{(W \times W \times \dots \times W)}_{t}\,.$$

The $(W)^2$, $(W)^3$ components of the associated escort geometric objects of the objects $p_{j_1} \bigstar p_{j_2}$ and $p_{j_1} \bigstar p_{j_2} \bigstar p_{j_3}$ are given by

$$\hat{p}_{j_1} \bigstar \hat{p}_{j_2} = \begin{pmatrix} \frac{\left(p_{j_1 j_2}(1,1)\right)^s}{\alpha_{j_1 j_2}} & \cdots & \frac{\left(p_{j_1 j_2}(1,W)\right)^s}{\alpha_{j_1 j_2}} \\ \vdots & \ddots & \vdots \\ \frac{\left(p_{j_1 j_2}(W,1)\right)^s}{\alpha_{j_1 j_2}} & \cdots & \frac{\left(p_{j_1 j_2}(W,W)\right)^s}{\alpha_{j_1 j_2}} \end{pmatrix} \tag{57}$$

and

$$\hat{p}_{j_1} \bigstar \hat{p}_{j_2} \bigstar \hat{p}_{j_3} = \begin{pmatrix} \frac{\left(p_{j_1 j_2 j_3}(1,1,1)\right)^s}{\alpha_{j_1 j_2 j_3}} & \cdots & \frac{\left(p_{j_1 j_2 j_3}(1,1,W)\right)^s}{\alpha_{j_1 j_2 j_3}} \\ \vdots & \ddots & \vdots \\ \frac{\left(p_{j_1 j_2 j_3}(W,1,1)\right)^s}{\alpha_{j_1 j_2 j_3}} & \cdots & \frac{\left(p_{j_1 j_2 j_3}(W,1,W)\right)^s}{\alpha_{j_1 j_2 j_3}} \\ \vdots & \ddots & \vdots \\ \frac{\left(p_{j_1 j_2 j_3}(1,W,1)\right)^s}{\alpha_{j_1 j_2 j_3}} & \cdots & \frac{\left(p_{j_1 j_2 j_3}(1,W,W)\right)^s}{\alpha_{j_1 j_2 j_3}} \\ \vdots & \ddots & \vdots \\ \frac{\left(p_{j_1 j_2 j_3}(W,W,1)\right)^s}{\alpha_{j_1 j_2 j_3}} & \cdots & \frac{\left(p_{j_1 j_2 j_3}(W,W,W)\right)^s}{\alpha_{j_1 j_2 j_3}} \end{pmatrix} \tag{58}$$

respectively.

To better understand Eq. (58), we introduce (Fig. 6).

The $(W)^t$ components of the geometric object of Eq. (56) are now written as:

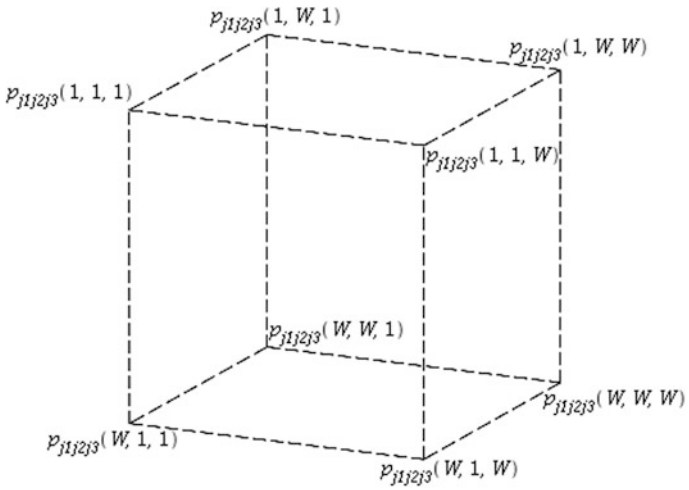


Fig. 6 A representative of $p_{j_1 j_2 j_3}$ object

$$\hat{p}_{j_1} \bigstar \hat{p}_{j_2} \bigstar \ldots \bigstar \hat{p}_{j_t} = \begin{pmatrix} \frac{\left(p_{j_1 \ldots j_t}(1,\ldots,1)\right)^s}{\alpha_{j_1 \ldots j_t}} & \cdots & \frac{\left(p_{j_1 \ldots j_t}(1,\ldots,1,W)\right)^s}{\alpha_{j_1 \ldots j_t}} \\ \vdots & \ddots & \vdots \\ \frac{\left(p_{j_1 \ldots j_t}(W,1,\ldots,1)\right)^s}{\alpha_{j_1 \ldots j_t}} & \cdots & \frac{\left(p_{j_1 \ldots j_t}(W,1,\ldots,1,W)\right)^s}{\alpha_{j_1 \ldots j_t}} \\ \vdots & \ddots & \vdots \\ \frac{\left(p_{j_1 \ldots j_t}(1,W,1,\ldots,1)\right)^s}{\alpha_{j_1 \ldots j_t}} & \cdots & \frac{\left(p_{j_1 \ldots j_t}(1,W,1,\ldots,1,W)\right)^s}{\alpha_{j_1 \ldots j_t}} \\ \vdots & \ddots & \vdots \\ \frac{\left(p_{j_1 \ldots j_t}(W,W,1,\ldots,1)\right)^s}{\alpha_{j_1 \ldots j_t}} & \cdots & \frac{\left(p_{j_1 \ldots j_t}(W,\ldots,W)\right)^s}{\alpha_{j_1 \ldots j_t}} \end{pmatrix} \tag{59}$$

In order to complete this appendix and to assure its usefulness on the derivation of the results presented in Sects. 3 and 4, we now treat the case of equiprobable probabilities of occurrence.

In the case of equiprobable distributions $p_{j_1 \ldots j_t}(a_1, \ldots, a_t) = W^{-t}$, we have from Eqs. (43) and (46),

$$1 = \sum_{\mu=1}^{m} \left(\frac{q_\mu}{m}\right) = m W^{-t} \tag{60}$$

This means that if $m \geq W^t$, we should understand this equation as

$$W^t W^{-t} + (m - W^t) \cdot 0 = 1 \tag{61}$$

We then see that for equiprobable distributions, the symbols σ_k from Sects. 3 and 4 do not depend on s and can be written as:

$$\sigma^k = \frac{\sum\limits_{\mu=1}^{k} \left(\frac{q_\mu}{m}\right)^s}{\sum\limits_{\mu=1}^{m} \left(\frac{q_\mu}{m}\right)^s} = \frac{k(W^{-t})^s}{m(W^{-t})^s} = \frac{k}{m} = \frac{k}{W^t}\,, \quad k = 1, \ldots, W^{t-1}, W^t\,, \tag{62}$$

where we have used Eq. (60).

The corresponding values of $r(s)$ could be written from Eq. (36) of Sect. 3, as:

$$r(s) = s - \frac{(1-s)^2}{s\sigma^k} = s - \frac{(1-s)^2}{sk} W^t\,. \tag{63}$$

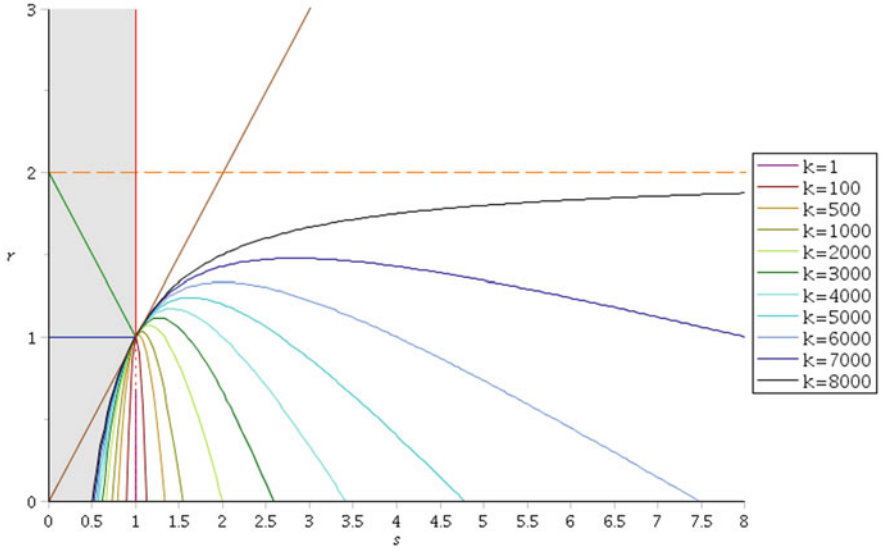


Fig. 7 The $r_k(s)$ curves and the structure of the $(s\text{-}r)$ space of parameters for some selected k-values of an equiprobable distribution of probabilities

In Sect. 3, we present a numerical application of Eq. (63), in order to continue the analysis of the structure of $(s\text{-}r)$ space of parameters.

In order to conclude this appendix, we work again with sets of 3-amino acids, with $W^3 = 8000$ for an equiprobable distribution and we choose to depict that the curves corresponding to the values: $k = 1, 100, 500, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000$, on Fig. 7. The last curve $r_{8000}(s) = 2 - \frac{1}{s}$ was already presented in Sect. 3 on Fig. 2.

Appendix 2: The Origin of the Generalized Khinchin-Shannon Inequalities

Let us consider the definition of the conditional of the escort probability:

$$\hat{p}_{j_1 \ldots j_t}(a_1, \ldots, a_{t-1}|a_t) = \frac{\hat{p}_{j_1 \ldots j_t}(a_1, \ldots, a_t)}{\hat{p}_{j_t}(a_t)} \tag{64}$$

From Eq. (30) we can write, analogously:

$$\hat{p}_{j_1\ldots j_t}(a_1,\ldots,a_t)=\frac{\Big(p_{j_1\ldots j_t}(a_1,\ldots,a_t)\Big)^s}{\sum\limits_{(b_1,\ldots,b_t)}\Big(p_{j_1\ldots j_t}(b_1,\ldots,b_t)\Big)^s} \tag{65}$$

$$\hat{p}_{j_t}(a_t)=\frac{\Big(p_{j_t}(a_t)\Big)^s}{\sum\limits_{a_t}\Big(p_{j_t}(a_t)\Big)^s} \tag{66}$$

In Eqs. (65) and (66), the symbols a_j, b_j, $1 \leq j \leq t$, are running over the one-letter code for the 20 amino acids

$$a_j; b_j = \text{A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y},\ 1 \leq j \leq t\,.$$

We then have from Eqs. (64)–(66):

$$\hat{p}_{j_1\ldots j_t}(a_1,\ldots,a_{t-1}|a_t)=\frac{\Big(p_{j_1\ldots j_t}(a_1,\ldots,a_{t-1}|a_t)\Big)^s}{\sum\limits_{(a_1,\ldots,a_t)}\hat{p}_{j_t}(a_t)\Big(p_{j_1\ldots j_t}(a_1,\ldots,a_{t-1}|a_t)\Big)^s} \tag{67}$$

We now remember the definition of the escort of the conditional probability

$$\overbrace{p_{j_1\ldots j_t}(a_1,\ldots,a_{t-1}|a_t)} \;=\frac{\Big(p_{j_1\ldots j_t}(a_1,\ldots,a_{t-1}|a_t)\Big)^s}{\sum\limits_{(a_1,\ldots,a_{t-1})}\Big(p_{j_1\ldots j_t}(a_1,\ldots,a_{t-1}|a_t)\Big)^s} \tag{68}$$

The left hand sides of Eqs. (67) and (68) are identical if all amino acids of the t-th column are equal. For instance,

$$j_t \longrightarrow \underbrace{(A, A, A, A, \ldots, A)}_{m} \tag{69}$$

This distribution will then lead to:

$$\hat{p}^{\mathsf{T}}_{j_t}=p_{j_t}=\underbrace{1,0,0,0,\ldots,0}_{20} \tag{70}$$

For any other distributions of amino acids in the jth column, the ordering of the two denominators on the right hand sides of Eqs. (67) and (68) has to be decided after choosing a protein domain family and its related distribution of probabilities of occurrence of amino acids. In order to undertake this study, we pay special attention to some functions of probabilities already defined in ref. [1], together an additional

definition, $X(a_t)$.

$$U \equiv \sum_{a_1,\ldots,a_{t-1}} \left[p_{j_1\ldots j_{t-1}}(a_1,\ldots,a_{t-1})\right]^s = \alpha_{j_1\ldots j_{t-1}} \tag{71}$$

$$J \equiv \sum_{a_1,\ldots,a_{t-1}} \left[\sum_{a_t} \hat{p}_{j_t}(a_t)\, p_{j_1\ldots j_t}(a_1,\ldots,a_{t-1}|a_t)\right]^s \tag{72}$$

$$Z \equiv \sum_{a_1,\ldots,a_t} \hat{p}_{j_t}(a_t)\left[p_{j_1\ldots j_t}(a_1,\ldots,a_{t-1}|a_t)\right]^s = \frac{\alpha_{j_1\ldots j_t}}{\alpha_{j_t}} \tag{73}$$

$$X(a_t) \equiv \sum_{a_1,\ldots,a_{t-1}} \left[p_{j_1\ldots j_t}(a_1,\ldots,a_{t-1}|a_t)\right]^s \tag{74}$$

The letters U, J, Z, X, have been chosen among those which do not codify the amino acids: B J O U X Z.

Actually, there is not any ordering between (74) and (71), (72), (73) for a generic occurrence of amino acids, as have been emphasized in ref. [1]. This assertion could be also proven "a fortiori" if the amino acid corresponding to a_t does not occur, or $X(a_t) = 0$.

It should be also stressed that the condition $J \geq Z$ i.e., the Hölder inequality [1, 12] which stands for $1 > s \geq 0$ is sufficient to guarantee the inequality $U \geq Z$, which leads to the Generalized Khinchin-Shannon inequalities to be obtained through iteration:

$$\alpha_{j_1\ldots j_{t-1}} \geq \frac{\alpha_{j_1\ldots j_t}}{\alpha_{j_t}}$$

$$t \to t-1 \Rightarrow \alpha_{j_1\ldots j_{t-2}} \geq \frac{\alpha_{j_1\ldots j_{t-1}}}{\alpha_{j_{t-1}}} \geq \frac{\alpha_{j_1\ldots j_t}}{\alpha_{j_t}\cdot\alpha_{j_{t-1}}}$$

$$t \to t-2 \Rightarrow \alpha_{j_1\ldots j_{t-3}} \geq \frac{\alpha_{j_1\ldots j_{t-2}}}{\alpha_{j_{t-2}}} \geq \frac{\alpha_{j_1\ldots j_t}}{\alpha_{j_t}\cdot\alpha_{j_{t-1}}\cdot\alpha_{j_{t-2}}}$$

$$\vdots$$

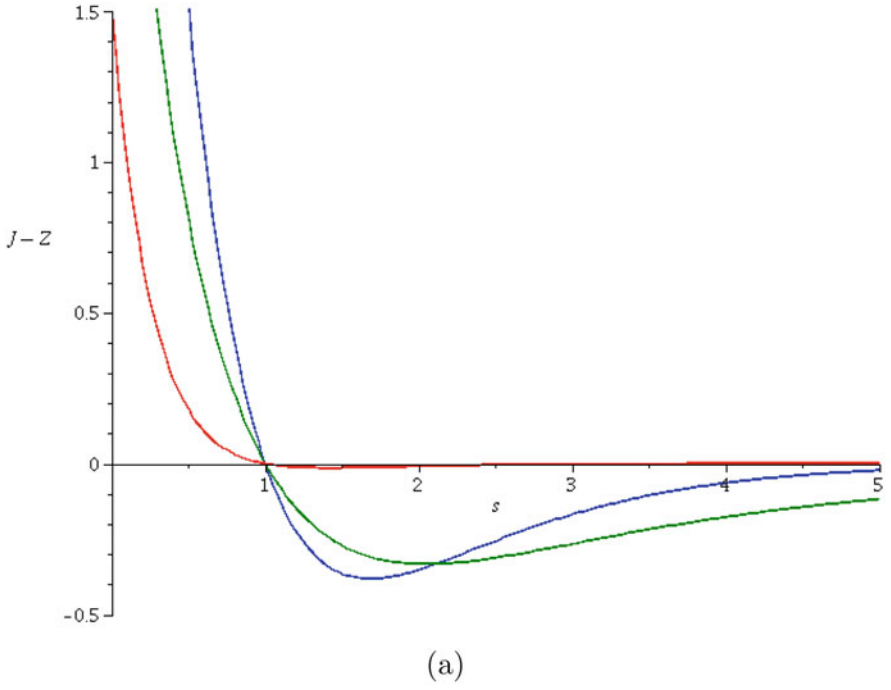
$$t \to 2 \Rightarrow \alpha_{j_1} \geq \frac{\alpha_{j_1 j_2}}{\alpha_{j_2}} \geq \frac{\alpha_{j_1\ldots j_t}}{\alpha_{j_t}\cdot\ldots\cdot\alpha_{j_3}\cdot\alpha_{j_2}}$$

$$\Rightarrow \alpha_{j_1\ldots j_t} \leq \prod_{l=1}^{t} \alpha_{j_l} \tag{75}$$

However, in Figs. 8 and 9, below, we have tried to characterize the upsurge of regions in which $J \geq Z$ for $s > 1$ by presenting the curves $U - Z$ and $J - Z$ for two 8×3 arrays obtained from the Pfam 27.0 database and a hypothetical 8×3 array which has been constructed aiming a better alignment. The results led us to conjecture that this would occur systematically on greater arrays to be obtained by working with recent versions of the database. This kind of work is now in progress and it will be published elsewhere.

Fig. 8 8×3 arrays of amino acids: (**a**) from Pfam PF01926, rows 25–32, columns 30-32; (**b**) from Pfam PF01926, rows 3–10, columns 3–5; (**c**) a hypothetical 8×3 array aiming a good alignment

$$\begin{pmatrix} K & A & Q \\ A & G & E \\ N & T & P \\ V & A & G \\ K & P & Q \\ V & P & G \\ P & K & C \\ R & P & G \end{pmatrix} \quad \text{(a)} \qquad \begin{pmatrix} V & I & T \\ A & I & I \\ V & I & A \\ V & I & A \\ M & V & V \\ V & L & I \\ A & I & V \\ A & L & I \end{pmatrix} \quad \text{(b)} \qquad \begin{pmatrix} A & A & A \\ A & C & C \\ A & C & A \\ A & C & A \\ A & C & A \\ C & A & A \\ C & C & A \\ C & C & A \end{pmatrix} \quad \text{(c)}$$



(a)

Fig. 9 The blue, green and red curves do correspond to the matrices (**a**), (**b**), and (**c**) of Fig. 8, respectively. (**a**) Study of the Hölder ($J \geq Z$) and Non-Hölder ($J < Z$) probability distributions from the 8×3 arrays of Fig. 8; (**b**) Study of the viability of Generalized Khinchin-Shannon Inequalities ($U \geq Z$) associated to Hölder probability distributions ($J \geq Z$) or $J \geq Z \Rightarrow U \geq Z$. However, there are distributions of amino acids such that $U \geq Z$ even where $J < Z$. The red curve is an example for $s > 1.4537$

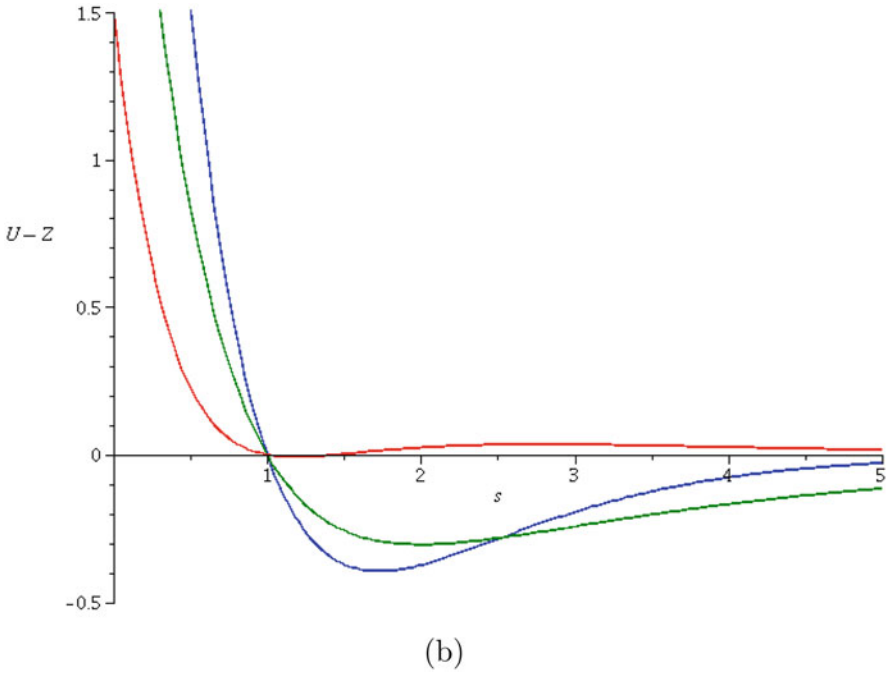


Fig. 9 (continued)

References

1. R. P. Mondaini, S. C. de Albuquerque Neto – Alternative Entropy Measures and Generalized Khinchin-Shannon Inequalities – *Entropy* **23** (2021) 1618, https://doi.org/10.3390/e23121618 and references therein.
2. A. I. Khinchin – Mathematical Fundations of Information Theory. Dover Publications: New York, NY, USA, 1957.
3. B. D. Sharma, D. P. Mittal – New Non-additive Measures of Entropy for Discrete Probability Distributions – *J. Math Sci* **10** (1975) 28–40.
4. R. P. Mondaini, S. C. de Albuquerque Neto – The Statistical Analysis of Protein Domain Family Distributions via Jaccard Entropy Measures. In *Trends in Biomathematics: Modeling Cells, Flows, Epidemics, and the Environment*; Mondaini, R. P., Ed.; Springer International Publishing, Cham, Switzerland, 2020, pp.169–207.
5. J. Havrda, F. Charvat – Quantification Method of Classification Processes. Concept of Structural α-entropy – *Kybernetica* **3** (1967) 30–35.
6. P. T. Landsberg, V. Vedral – Distributions and Channel Capacities in Generalized Statistical Mechanics – *Phys. Lett. A* **247** (1998) 211–217.
7. A. Renyi – On Measures of Entropy and Information. In *Contributions to the Theory of Statistics, Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, CA, USA, 20 June–30 July 1960*; Neyman, J., Ed.; University of California Press: Berkeley, CA, USA, 1961; Volume 1, pp. 547–561.
8. T. Oikonomou – Properties of the "non-extensive Gaussian" entropy – *Physica A* **381** (2007) 155-163.

9. R. P. Mondaini, S. C. de Albuquerque Neto – A Jaccard-like symbol and its usefulness in the derivation of amino acid distributions in protein domain families. In *Trends in Biomathematics: Chaos and Control in Epidemics, Ecosystem, and Cells*; Mondaini, R. P., Ed.; Springer International Publishing, Cham, Switzerland, 2021, pp.201-220.
10. R. P. Mondaini, S. C. de Albuquerque Neto – Generalized Khinchin-Shannon Inequalities and the Identification of Protein Domain Families via Jaccard-type Symbols. In *Trends in Biomathematics: Stability and Oscillations in Environmental, Social, and Ecology Models*; Mondaini, R. P., Ed.; Springer International Publishing, Cham, Switzerland, 2022, to be published.
11. J. E. Marsden, A. Tromba – Vector Calculus. W. H. Freeman and Company Publishers: New York, NY, USA, 2012.
12. G. H. Hardy, J. E. Littlewood, G. Pólya – Inequalities. Cambridge University Press: London, UK, 1934.
13. A. Moghadan, S. I. Omar, J. A. Tuszynski – Probability distributions of p53 mutations and their corresponding Shannon entropies in different cancer cell types. In *Trends in Biomathematics: Stability and Oscillations in Environmental, Social, and Ecology Models*; Mondaini, R. P., Ed.; Springer International Publishing, Cham, Switzerland, 2022, this issue.
14. R. P. Mondaini, S. C. de AlbuquerqueNeto – Essential Conditions for the Full Synergy of Probability Occurrence Distributions – *Entropy* **24** (2022) 993, https://doi.org/10.3390/e24070993

Computational Modeling of Cancer Response to Oncolytic Virotherapy: Improving the Effectiveness of Viral Spread and Anti Tumor Efficacy

H. Lefraich

1 Introduction

Oncolytic virotherapy represents a promising anti-cancer treatment approach, which involves viruses that have been selectively engineered to infect and destroy cancer cells, while sparing the surrounding healthy cells [1–4]. These oncolytic viruses can penetrate cancer cells either through receptor binding, or through fusion with plasma membrane, then they replicate by taking advantage of signaling pathways and common mutations inside those cancer cells. Despite the oncolytic viruses can enter also healthy cells, they do not usually replicate inside these cells and thus do not destroy them. In fact, the absence of such mutations in healthy cells tends to abort the replication cycle of the viruses. In 2015, the United States of America Food and Drug Administration (FDA), based on the recent advances in the understanding of tumor-virus interactions, approved the first genetically engineered OV (a Herpes Simplex Virus) as a therapy for the melanoma cancer [3]. However, despite the fact that multiple oncolytic viruses (OVs) are presently under clinical development [1, 5], this kind of therapy still has some challenges in terms of effectiveness (as confirmed by several clinical trials) [6]. This relatively low oncolytic effectiveness is not only due to premature virus clearance by circulating antibodies and various immune cells [7], but also the presence of physical barriers inside tumors (e.g., interstitial fluid pressure, extracellular matrix (ECM) deposits) that hinder virus movement [8]. In fact, the ECM has been recognized as a major barrier for anti-tumor efficacy as it plays a pivotal role in inhibiting virus spread [9–11]. To face the challenge of improving the intra-tumoral spread of oncolytic viruses, numerous experimental and clinical approaches are currently being considered,

H. Lefraich (✉)
Laboratory (MISI), Faculty of Sciences and techniques, Department of Mathematics and Computer Science, University Hassan First, Settat, Morocco

R. P. Mondaini (ed.), *Trends in Biomathematics: Stability and Oscillations in Environmental, Social, and Biological Models*,
https://doi.org/10.1007/978-3-031-12515-7_16

from genetically manipulating natural OVs to incorporate additional features for improving their efficiency [5], to modifications of the physical barriers (e.g., via ECM degradation) to improve virus spread [11]. Understanding more deeply the interactions between the extracellular matrix (ECM), uninfected and infected cancer cells, and oncolytic viruses is therefore of great importance to shed a light on the biological mechanisms that might improve OV spread by overcoming the physical barriers inside the tumor micro-environment.

During the last two decades, numerous mathematical models have been applied to gaining a broader understanding of the dynamics of virotherapy by analyzing more deeply the interaction between cancer cells and oncolytic viruses. Concerning ordinary differential equations models, one can refer to Wodarz in [12–14], Komarova and Wodarz [15], Novozhilov et al. [16], Bajzer et al. [17, 18], Tian in [19, 20] and others [21–24]. Wodarz [13] formulated a simple model with two differential equations where a virus is interacting with a population of uninfected and virus-infected tumour cells. Based on this work, Komarova and Wodarz [15] used a general approach by taking into account a class of models instead of a specific model and considered two populations: uninfected tumour cells and infected tumour cells. Wodarz in his paper [14] proposed a model with three populations where he modeled explicitly the viral population. Then, Tian [20] has proposed an improved model by incorporating the burst size. The burst size of a virus is the number of new virions released from a lysis of an infected cell. The model is given by

$$\begin{cases} \dfrac{dc}{dt} = \mu_1 c(1 - \dfrac{c+i}{C}) - \rho c v \\ \dfrac{di}{dt} = \rho c v - \delta_i i \\ \dfrac{dv}{dt} = b\delta_i i - \rho c v - \delta_v v \end{cases}$$

The populations of the model consist of the densities within three groups: uninfected cancer cells $c(t)$, infected cancer cells $i(t)$, and free virus $v(t)$. The constant C stands for the maximal tumor size. The term $\mu_1 c$ models the rate of growth of cancer cells, $\rho c v$ the rate of infection of cancer cells by the virus, $b\delta_i i$ the release of virions by infected cancer cells, $\delta_i i$ the rate of clearance of infected cancer cells, and $\delta_v v$ the rate of clearance of the virus. Besides, Novozhilov et al. [16] analyzed ratio based oncolytic virus infection models. Bajzer et al. [17] modeled the specific cancer virotherapy with measles virus by using three ordinary differential equations, and then they approached the optimization of viral doses, number of doses and timing in [18].

Because of the availability of temporal data the majority of these models approached the oncolytic viruses from temporal dynamics perspective. However, the main cause of mortality among cancer patients is the spread of primary

cancer cells to other parts of the body to establish secondary tumours (metastasis). Moreover, the diffusion of oncolytic viruses within the tumor can play a major role in the efficacy and success of the treatment [25, 26]. Consequently, the spatial dependency of the virotherapy needs also to be considered in mathematical models in order to realistically simulate the clinical observations. This shortcoming of ODE models necessitates the use of PDEs in the field of computational cancer biology. More recent advances in tumour imaging provided data on the spatial spread of tumours and viruses, which then encouraged the emergence of mathematical models investigating the spatial spread of these viruses. Thus, many partial differential equation models have been formulated to include both the spatial and temporal distributions of viruses and cells [27–32]. For instance, Wu et al. [27] and Wein et al. [28] proposed and analyzed a system of partial differential equations in order to describe some aspect of oncolytic viral therapy. Camara et al. [29] explored an important interaction between aggressively invasive glioma cancer cells and the free oncolytic virus "ONYX-015". Malinzi et al. [30] developed a reaction-diffusion system and considered the impact of virotherapy on the concentration of the tumor cells in the presence of CTL immune response. Then, Malinzi et al. [32] proposed a PDE based model that study the effect of the combination of oncolytic virotherapy with chemotherapy.

All suggested computational models have given valuable insights into oncolytic virotherapy dynamics. Nevertheless, there is a tremendous need to understand the dynamics of oncolytic virotherapy in the presence of extracellular matrix (ECM), particularly, to understand the complex interplay among cancer cells, oncolytic viruses and the extracellular matrix. In fact, tumour cells encounter a large amount of insoluble intact adhesive molecules of the extracellular matrix (ECM), which may promote their directed migration at different stages in the process of cancer invasion. A cell would migrate from a region of low concentration of relevant adhesive molecules towards a region with a higher concentration. This phenomenon is termed haptotaxis [33, 34]. Furthermore, the ECM is considered as a major barrier to virus motility by acting like a porous medium.

In this study, we are interested in a mathematical modeling and computational approach of oncolytic virotherapy which aims to help us improve our understanding of the physical barrier that inhibits the virus spread. Therefore, we propose a mathematical model of oncolytic virotherapy, that accounts for interaction between uninfected cancer cells, infected cancer cells, extracellular matrix (ECM) and oncolytic virus. Besides random motion, both uninfected and infected tumour cells migrate haptotactically towards higher ECM densities, moreover, in addition to degrading the static ECM upon contact the two cancer populations undergo an infection-induced transition mechanism conducted by virus particles which are released by infected cancer cells, and which attack the uninfected part of the tumor. One of the main contributions of this model consists in taking into account that the motility rate of virus particles is controlled by the population of ECM where lowering ECM leads to higher viral diffusion. Accordingly, we defined the diffusivity of tumour cells as a monotonically decreasing function of ECM density to model the obstruction of movement by the ECM.

This chapter is organized as follows. In Sect. 2, the mathematical model as well as its dimensionless form are presented. The variational formulation, the temporal discretization, and the finite element scheme of the nonlinear system are presented in Sect. 3. Finally, we present numerical experiments and summarize the observations of this study.

2 Mathematical Model

In this section, we present a two-dimensional mathematical model of oncolytic virotherapy. The model consists of four unknown variables namely the uninfected cancer cells density, the infected cancer cells density, the ECM density and the oncolytic virus particles density, which are denoted respectively by $c(x,t)$, $i(x,t)$, $u(x,t)$ and $v(x,t)$. We consider the system to hold on a bounded spatial domain $\Omega \subset \mathbb{R}^2$, representing a region of tissue. As the tumour cells depend on the closest blood vessel, we assume that the computational domain Ω is a disc with a centered hole inside representing the blood vessel (see Fig. 1). The boundary of the domain consists of two parts $\partial\Omega = \partial\Omega_1 \cup \partial\Omega_2$, where $\partial\Omega_1$ refers to the intern boundary delimited by the blood vessel wall, and $\partial\Omega_2$ refers to the extern boundary of Ω. Moreover, we estimate the radius of this region supported by the blood vessel to be $\frac{r_b}{\sqrt{BVF}}$, where BVF is the blood volume fraction [35, 36], and r_b is the radius of the blood vessel. Concerning the viruses, they attain the tumor through the blood vessel via $\partial\Omega_1$ and diffuse into it with no flux at $\partial\Omega_2$. Furthermore, viruses are provided continuously with a constant concentration v_b (for instance by using nanotechnology [37]). Guided by the in vitro experimental protocol in which invasion takes place within an isolated system, we consider zero-flux boundary conditions at the blood vessel wall $\partial\Omega_1$ and at the extern boundary of the disc for the uninfected cancer cells, the infected cancer cells, and also for the ECM.

Uninfected Cancer Cells $c(x,t)$ The underlying modeling assumptions are that in addition to random cell movement (with D_c as the random motility coefficient), uninfected cancer cells have a directed haptotactic movement towards higher ECM gradients (with η_c the haptotactic coefficient). Moreover, cancer cells apart from possibly proliferating logistically at rate μ_1 [38, 39], can decay due to virus infection at rate ρ. These hypotheses can be described by the following equation:

$$\frac{\partial c}{\partial t} = D_c \Delta c - \eta_c \nabla.(c \nabla u) + \mu_1 c(1 - \frac{c}{C} - \frac{i}{C}) - \rho c v,$$

where C is the carrying capacity of the cancer cells and the extracellular matrix.

Infected Cancer Cells $i(x,t)$ Similarly, we assume that the infected cancer cells can move randomly (with D_i as the random motility coefficient) and can migrate towards higher ECM gradients (with η_i the haptotactic coefficient). As mentioned

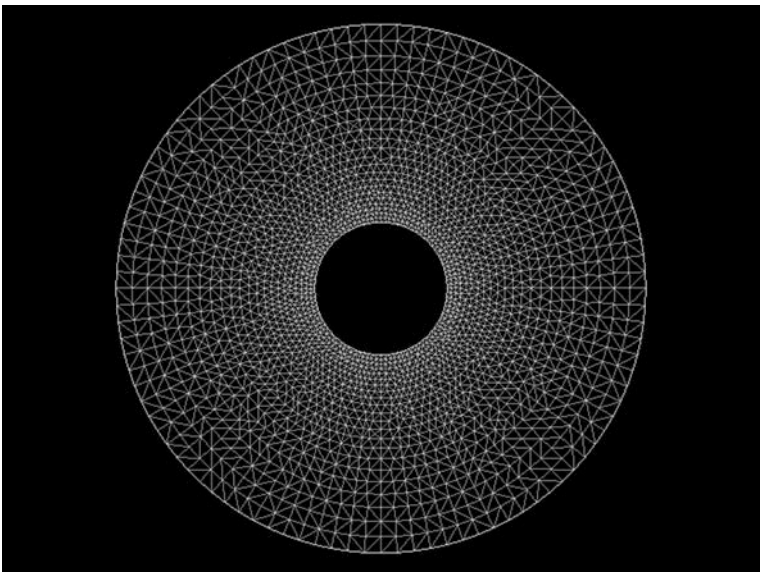


Fig. 1 Computational domain constructed. the computational domain Ω is a disc with a centered hole inside representing the blood vessel

above, these cells are infected at rate ρ by the oncolytic virus. Furthermore, these infected cells die owing to lysis at a rate δ_i. These assumptions are translated into the following equation:

$$\frac{\partial i}{\partial t} = D_i \Delta i - \eta_i \nabla.(i \nabla u) + \rho c v - \delta_i i.$$

Extracellular Matrix (ECM) $u(x,t)$ The ECM can be regarded as static in the sense that it does not move, and thus we may neglect any diffusion and migration terms, however, it undergoes a continuous remodelling by cells in the environment [40]. We represent this remodeling process by the difference between a logistic growth term with rate μ_2 and a degradation term (where α_c is the rate of ECM degradation by uninfected cancer cells, and α_i is the rate of ECM degradation by infected cancer cells). Thus, the evolution of the ECM is governed by the following equation:

$$\frac{\partial u}{\partial t} = -u(\alpha_c c + \alpha_i i) + \mu_2 u(1 - \frac{u}{C} - \frac{c}{C} - \frac{i}{C}),$$

Oncolytic Virus $v(x,t)$ We assume that the virus can spread randomly in ECM density-dependent manner through the environment with $D_v(u)$ the random motility coefficient which quantifies how virus particles outside infected cells can freely move around less dense ECM. We opt for the phenomenological form:

$$D_v(u) = D_v(1 - \frac{u}{C}),$$

where D_v is a reference diffusion in absence of ECM. Furthermore, the virus particles duplicate at rate b, the burst size of infected cancer cells, which release the new virions in the environment. However, the number of free virus particles reduction is mainly due to the natural virion's death at rate δ_v, and the trapping of these virus particles into the cancer cells at rate ρ. In summary, the governing equation for the density of virus particles is as follows:

$$\frac{\partial v}{\partial t} = \nabla(D_v(u)\nabla v) + bi - \rho cv - \delta_v v.$$

In summary, we obtain the following system of PDE-ODE equations, where all the parameters in the system are nonnegative:

$$\begin{cases} \dfrac{\partial c}{\partial t} = D_c \Delta c - \eta_c \nabla.(c\nabla u) + \mu_1 c(1 - \dfrac{c}{C} - \dfrac{i}{C}) - \rho cv & \text{in } \Omega \times]0,T[\,, \\ \dfrac{\partial i}{\partial t} = D_i \Delta i - \eta_i \nabla.(i\nabla u) + \rho cv - \delta_i i & \text{in } \Omega \times]0,T[\,, \\ \dfrac{\partial u}{\partial t} = -u(\alpha_c c + \alpha_i i) + \mu_2 u(1 - \dfrac{u}{C} - \dfrac{c}{C} - \dfrac{i}{C}) & \text{in } \Omega \times]0,T[\,, \\ \dfrac{\partial v}{\partial t} = \nabla(D_v(u)\nabla v) + bi - \rho cv - \delta_v v & \text{in } \Omega \times]0,T[\,. \end{cases} \tag{1}$$

Based on the aforementioned assumptions the system is closed with the following homogenous initial and boundary conditions:

$$\begin{cases} c(x,0) = c_0(x) \\ i(x,0) = i_0(x) \\ u(x,0) = u_0(x) \\ v(x,0) = v_0(x) \end{cases}$$

$$\begin{cases} D_c \dfrac{\partial c}{\partial n} + \eta_c c \dfrac{\partial u}{\partial n} = 0 & \text{in } \partial\Omega \times]0, T[\,, \\ D_i \dfrac{\partial i}{\partial n} + \eta_i i \dfrac{\partial u}{\partial n} = 0 & \text{in } \partial\Omega \times]0, T[\,, \\ D_v(u) \dfrac{\partial v}{\partial n} = 0 & \text{in } \partial\Omega_2 \times]0, T[\,, \\ v = v_b & \text{in } \partial\Omega_1 \times]0, T[\,, \end{cases}$$

where n is the outward normal vector to $\partial\Omega$ and v_b is the value of v at the blood vessel.

2.1 Dimensionless Form

Let $L = 0.1cm$ and $\tau = \frac{L^2}{D}$, (where $D \approx 10^{-6}\,\mathrm{cm}^2\,\mathrm{s}^{-1}$ according to [41]) be the characteristic length and time scale respectively. We define the dimensionless variables as follows:

$$\widetilde{c} = \frac{c}{C}, \quad \widetilde{i} = \frac{i}{C}, \quad \widetilde{u} = \frac{u}{C}, \quad \widetilde{v} = \frac{v}{C}, \quad \widetilde{x} = \frac{x}{L}, \quad \widetilde{t} = \frac{t}{\tau}$$

and new parameters via the following scaling:

$$\widetilde{D_c} = \frac{\tau D_c}{L^2}, \quad \widetilde{D_i} = \frac{\tau D_i}{L^2}, \quad \widetilde{D_v} = \frac{\tau D_v}{L^2}, \quad \widetilde{\eta_c} = \eta_c \frac{\tau C}{L^2}, \quad \widetilde{\eta_i} = \eta_i \frac{\tau C}{L^2}, \quad \widetilde{\mu_1} = \tau\mu_1,$$

$$\widetilde{\rho} = \tau\rho C, \quad \widetilde{\delta_i} = \tau\delta_i, \quad \widetilde{\alpha_c} = \tau\alpha_c C, \quad \widetilde{\alpha_i} = \tau\alpha_i C, \quad \widetilde{\mu_2} = \tau\mu_2, \quad \widetilde{b} = \tau b, \quad \widetilde{\delta_v} = \tau\delta_v$$

$$\widetilde{c(x,0)} = \frac{c(x,0)}{C}, \quad \widetilde{i(x,0)} = \frac{i(x,0)}{C}, \quad \widetilde{u(x,0)} = \frac{u(x,0)}{C}, \quad \widetilde{v(x,0)} = \frac{v(x,0)}{C}.$$

Henceforth, we drop the tilde over all variables for convenience. The dimensionless form of the model equations (1) can then be written in the following general form:

$$
\begin{cases}
\dfrac{\partial c}{\partial t} = D_c \Delta c - \eta_c \nabla.(c \nabla u) + \mu_1 c(1 - c - i) - \rho c v & \text{in } \Omega \times]0, T[\,, \\
\dfrac{\partial i}{\partial t} = D_i \Delta i - \eta_i \nabla.(i \nabla u) + \rho c v - \delta_i i & \text{in } \Omega \times]0, T[\,, \\
\dfrac{\partial u}{\partial t} = -u(\alpha_c c + \alpha_i i) + \mu_2 u(1 - u - c - i) & \text{in } \Omega \times]0, T[\,, \\
\dfrac{\partial v}{\partial t} = \nabla(D_v(u) \nabla v) + b i - \rho c v - \delta_v v & \text{in } \Omega \times]0, T[\,, \\
D_c \dfrac{\partial c}{\partial n} + \eta_c c \dfrac{\partial u}{\partial n} = 0 & \text{in } \partial\Omega \times]0, T[\,, \\
D_i \dfrac{\partial i}{\partial n} + \eta_i i \dfrac{\partial u}{\partial n} = 0 & \text{in } \partial\Omega \times]0, T[\,, \\
D_v(u) \dfrac{\partial v}{\partial n} = 0 & \text{in } \partial\Omega_2 \times]0, T[\,, \\
v = v_b & \text{in } \partial\Omega_1 \times]0, T[\,, \\
c(x, 0) = c_0(x) & \text{in } \Omega, \\
i(x, 0) = i_0(x) & \text{in } \Omega, \\
u(x, 0) = u_0(x) & \text{in } \Omega, \\
v(x, 0) = v_0(x) & \text{in } \Omega.
\end{cases}
\tag{2}
$$

We suppose that c_0, i_0, u_0, $v_0 \in L^2(\Omega)$ and are nonnegative.

3 Finite Element Scheme

The mathematical model (2) considered in the previous section is a highly nonlinear coupled system of partial and ordinary differential equations. Thus, we can't afford to solve it analytically. Aiming to solve it numerically, we present a finite element scheme in this section. First, we derive the variational formulation for the uninfected and infected cancer cells density equations, the ECM density equation and the virus density equation. Further, we present the temporal discretization of the model equations. Finally, we describe a fixed-point-iteration to handle the nonlinear terms in the system (2) and we provide an appropriate numerical scheme.

3.1 Variational Formulation

In order to show the numerical formulation of the problem, let $V = L^2(0, T; H^1(\Omega))$ be the space of approximate solutions and $W = H^1(\Omega)$ be the space of tests functions. Let W^h be a finite element space of Lagrange $P1$ included in W and $V^h = L^2(0, T; W^h)$ be the finite dimensional subspace of V. The Faedo-Galerkin formulation for the problem is given by, finding c_h, i_h, $u_h \in V$ and $v_h \in V$ such that $v_h = v_b^h$ in $\partial\Omega_1$:

$$\begin{cases} \left(\frac{\partial c_h}{\partial t}, \phi\right) + a_c(c_h, c_h, i_h, u_h, v_h, \phi) = 0 \\ \left(\frac{\partial i_h}{\partial t}, \phi\right) + a_i(i_h, c_h, u_h, v_h, \phi) = 0 \\ \left(\frac{\partial u_h}{\partial t}, \phi\right) + a_u(u_h, u_h, c_h, i_h, \phi) = 0 \\ \left(\frac{\partial v_h}{\partial t}, \psi\right) + a_v(v_h, c_h, i_h, u_h, \psi) = 0 \\ c_h(x, 0) = c_{h,0}(x) \quad \text{in } \Omega, \\ i_h(x, 0) = i_{h,0}(x) \quad \text{in } \Omega, \\ u_h(x, 0) = u_{h,0}(x) \quad \text{in } \Omega, \\ v_h(x, 0) = v_{h,0}(x) \quad \text{in } \Omega, \end{cases} \tag{3}$$

for all $\phi \in W^h$ and $\psi \in W^h$ such that $\psi = 0$ in $\partial\Omega_1$, and a. e. $t \in]0, T[$ where:

$$a_c(c_1, c_2, i, u, v, \phi) = D_c \int_\Omega \nabla c_1.\nabla\phi dx - \eta_c \int_\Omega c_1 \nabla u.\nabla\phi dx - \mu_1 \int_\Omega c_1(1-c_2-i)\phi dx + \rho \int_\Omega c_1 v\phi dx,$$

$$a_i(i, c, u, v, \phi) = D_i \int_\Omega \nabla i.\nabla\phi dx - \eta_i \int_\Omega i\nabla u.\nabla\phi dx - \rho \int_\Omega cv\phi dx + \delta_i \int_\Omega i\phi dx,$$

$$a_u(u_1, u_2, c, i, \phi) = \int_\Omega u_1(\alpha_c c + \alpha_i i)\phi dx - \mu_2 \int_\Omega u_1(1 - u_2 - c - i)\phi dx,$$

$$a_v(v, c, i, u, \psi) = \int_\Omega D_v(u)\nabla v.\nabla\psi dx - b \int_\Omega i\psi dx + \rho \int_\Omega cv\psi dx + \delta_v \int_\Omega v\psi dx,$$

$c_{h,0}(x)$, $i_{h,0}(x)$, $u_{h,0}(x)$ and $v_{h,0}(x)$ are respectively the projections of $c_0(x)$, $i_0(x)$, $u_0(x)$ and $v_0(x)$ on W^h.

3.2 Discrete Problem

In this section, we present the temporal discretization of the coupled variational system (3), where we discussed the application of Crank-Nicolson time discretization.

3.2.1 Temporal Discretization

Let $0 = t^0 < t^1 < ... < t^N = T$ be a decomposition of the considered time interval $]0, T[$, and $\delta_t = t^{n+1} - t^n$, $n = 0, 1, ..., N-1$ denotes the uniform time step. In addition, we use $c_h^n(x) := c_h(x, t^n)$, $i_h^n(x) := i_h(x, t^n)$, $u_h^n(x) := u_h(x, t^n)$ and $v_h^n(x) := v_h(x, t^n)$ to denote the approximation of the solutions at time t^n. We apply the implicit Crank-Nicolson discretization scheme, which is second order and A-stable, then the semi-discrete form of the system (3) reads:

For given c_h^{n-1}, i_h^{n-1}, u_h^{n-1} and v_h^{n-1} with $c_h^0 = c_{h,0}$, $i_h^0 = i_{h,0}$, $u_h^0 = u_{h,0}$ and $v_h^0 = v_{h,0}$, find c_h^n, i_h^n, $u_h^n \in W^h$ and $v_h^n \in W^h$ such that $v_h^n = v_b^h$ in $\partial\Omega_1$ such that:

$$\begin{cases} \left(\dfrac{c_h^n - c_h^{n-1}}{\delta_t}, \phi\right) + \dfrac{1}{2}a_c(c_h^n, c_h^n, i_h^n, u_h^n, v_h^n, \phi) = -\dfrac{1}{2}a_c(c_h^{n-1}, c_h^{n-1}, i_h^{n-1}, u_h^{n-1}, v_h^{n-1}, \phi) \\ \left(\dfrac{i_h^n - i_h^{n-1}}{\delta_t}, \phi\right) + \dfrac{1}{2}a_i(i_h^n, c_h^n, u_h^n, v_h^n, \phi) = -\dfrac{1}{2}a_i(i_h^{n-1}, c_h^{n-1}, u_h^{n-1}, v_h^{n-1}, \phi) \\ \left(\dfrac{u_h^n - u_h^{n-1}}{\delta_t}, \phi\right) + \dfrac{1}{2}a_u(u_h^n, u_h^n, c_h^n, i_h^n, \phi) = -\dfrac{1}{2}a_u(u_h^{n-1}, u_h^{n-1}, c_h^{n-1}, i_h^{n-1}, \phi) \\ \left(\dfrac{v_h^n - v_h^{n-1}}{\delta_t}, \psi\right) + \dfrac{1}{2}a_v(v_h^n, c_h^n, i_h^n, u_h^n, \psi) = -\dfrac{1}{2}a_v(v_h^{n-1}, c_h^{n-1}, i_h^{n-1}, u_h^{n-1}, \psi) \end{cases} \tag{4}$$

for all $\phi \in W^h$ and $\psi \in W^h$ such that $\psi = 0$ in $\partial\Omega_1$.

3.2.2 Numerical Scheme

The nonlinearity in the semi-discrete form of the system (4) besides the coupling between the equations makes the computations a very challenging task. In one hand, a fully implicit treatment of the nonlinear and coupled terms leads to a coupled nonlinear algebraic system and it will be a very tough task to solve it with a nonlinear solver, in the other hand an explicit treatment leads to a linearized system in which the equations are solved simultaneously. However, it may require a severe restriction on the time step. Accordingly, we suggest a fixed-point iteration method [42] to treat the nonlinear and coupled terms semi-implicitly. For instance, let us explain briefly the fixed point iteration steps for a nonlinear term in the uninfected cancer cells density equation in the time interval (t^{n-1}, t^n). Let $c_{h,0}^n = c_h^{n-1}$, $u_{h,0}^n = u_h^{n-1}$ and $i_{h,0}^n = i_h^{n-1}$ and we replace the nonlinear integral terms in the uninfected cancer cells density with:

$$\int_\Omega c_{h,k}^n \nabla u_{h,k}^n . \nabla \phi dx \simeq \int_\Omega c_{h,k}^n \nabla u_{h,k-1}^n . \nabla \phi dx,$$

$$\int_{\Omega} c_{h,k}^{n}(1-c_{h,k}^{n}-i_{h,k}^{n})\phi dx \simeq \int_{\Omega} c_{h,k}^{n}(1-c_{h,k-1}^{n}-i_{h,k-1}^{n})\phi dx$$

for $k = 0, 1, 2, \ldots$. We iterate until the residual of the system (4) is less than the prescribed threshold value $\left(10^{-8}\right)$ or until the maximal number of iterations is reached. Finally, we set $c_h^n = c_{h,k}^n$ and advance to the next time step. We handle the nonlinear and coupled terms in all other equations in a similar way by using the above prescribed iteration of fixed point type. Consequently, the linearized form of the system (4) in the interval (t^{n-1}, t^n) reads:

For given $c_{h,0}^n = c_h^{n-1}$, $i_{h,0}^n = i_h^{n-1}$, $u_{h,0}^n = u_h^{n-1}$ and $v_{h,0}^n = v_h^{n-1}$ with $c_h^0 = c_{h,0}$, $i_h^0 = i_{h,0}$, $u_h^0 = u_{h,0}$ and $v_h^0 = v_{h,0}$, find $c_{h,k}^n$, $i_{h,k}^n$, $u_{h,k}^n$ and $v_{h,k}^n$ satisfying $v_{h,k}^n = v_b^h$ in $\partial\Omega_1$ such that for all $\phi \in W^h$ and $\psi \in W^h$ with $\psi = 0$ in $\partial\Omega_1$

$$\begin{cases} (c_{h,k}^n, \phi) + \frac{\delta_t}{2} a_c(c_{h,k}^n, c_{h,k-1}^n, i_{h,k-1}^n, u_{h,k-1}^n, v_{h,k-1}^n, \phi) = \left(c_h^{n-1}, \phi\right) - \frac{\delta_t}{2} a_c(c_h^{n-1}, c_h^{n-1}, i_h^{n-1}, u_h^{n-1}, v_h^{n-1}, \phi) \\ (i_{h,k}^n, \phi) + \frac{\delta_t}{2} a_i(i_{h,k}^n, c_{h,k}^n, u_{h,k-1}^n, v_{h,k-1}^n, \phi) = \left(i_h^{n-1}, \phi\right) - \frac{\delta_t}{2} a_i(i_h^{n-1}, c_h^{n-1}, u_h^{n-1}, v_h^{n-1}, \phi) \\ (u_{h,k}^n, \phi) + \frac{\delta_t}{2} a_u(u_{h,k}^n, u_{h,k-1}^n, c_{h,k}^n, i_{h,k}^n, \phi) = \left(u_h^{n-1}, \phi\right) - \frac{\delta_t}{2} a_u(u_h^{n-1}, u_h^{n-1}, c_h^{n-1}, i_h^{n-1}, \phi) \\ (v_{h,k}^n, \psi) + \frac{\delta_t}{2} a_v(v_{h,k}^n, c_{h,k}^n, i_{h,k}^n, u_{h,k}^n, \psi) = \left(v_h^{n-1}, \psi\right) - \frac{\delta_t}{2} a_v(v_h^{n-1}, c_h^{n-1}, i_h^{n-1}, u_h^{n-1}, \psi) \end{cases} \tag{5}$$

4 Numerical Simulation and Results

In this section, we investigate numerically the model introduced previously. We aim to observe the cancer response to oncolytic virotherapy via tracking the model behaviour with respect to several aspects such as: cancer cells density, tumour load $(c + i)$, or cancer suppression.

For the numerical investigation, we first choose a baseline parameters (see Table 1) which are mainly based on the published studies in [29, 36, 41, 43] and present accordingly the spatial propagation of the oncolytic virus at several time stages. Second, we vary some of the parameters involved in virus dynamics and spread: the virus replication rate b, the infection rate of cancer cells ρ, and the death rate of infected cancer cells δ_i. Finally, based on the obtained results with different parameter values, we discuss the condition that can improve tumour suppression.

Visualizing the Oncolytic Viral Diffusion

In Figs. 2 and 3 we present the spatial propagation of the oncolytic virus for different relative time stages.

Table 1 Baseline parameters values for the model

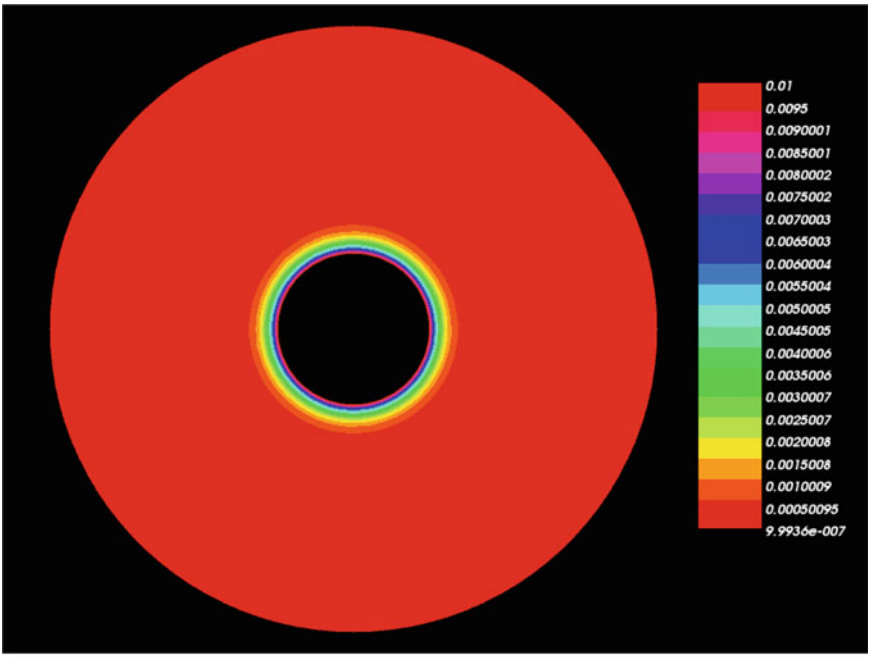
Parameter	Dimensional value	Non dimensional value	Unit	Reference
D_c (Diffusion coefficient of uninfected cancer cells)	0.00675	–	mm^2/h	[29]
D_i (Diffusion coefficient of infected cancer cells)	0.0054	–	mm^2/h	[29]
D_v (Diffusion coefficient of viruses in the reference case)	0.036	–	mm^2/h	[29]
η_c (Haptotaxis coefficient of uninfected cancer cells)	–	2.85×10^{-2}	$cm^2\ s^{-1}\ M^{-1}$	[43]
η_i (Haptotaxis coefficient of infected cancer cells)	–	2.85×10^{-2}	$cm^2\ s^{-1}\ M^{-1}$	[43]
μ_1 (Proliferation rate of uninfected cancer cells)	–	0.25	h^{-1}	[43]
ρ (Infection rate of cells by viruses)	–	79×10^{-3}	mm^2/h virus	[29]
δ_i (Death rate of infected cancer cells)	0.05	–	1/h	[29]
α_c (Degradation rate of ECM by uninfected cancer cells)	–	0.15	1/h	[41]
α_i (Degradation rate of ECM by infected cancer cells)	–	0.075	1/h	[41]
μ_2 (ECM remodelling rate)	–	0.015	1/h	[41]
b (Virus replication rate)	2	–	1/h	[31]
δ_v (Clearance rate of viruses)	0.025	–	1/h	[29]
C (Carrying capacity for cancer cells and ECM)	10^6	–	$Cells/mm^3$	[29]
r_b (Radius of the blood vessel)	0.01	–	mm	Estimated
BVF (Blood volume fraction)	0.05	–	–	[36]
v_b (Value of v at the blood vessel)	0.01×10^6	0.01	$Cells/mm^3$	Estimated
c_0 (Initial condition of uninfected cancer cells)	0.5×10^6	0.5	$Cells/mm^3$	Estimated
i_0 (Initial condition of infected cancer cells)	0	0	$Cells/mm^3$	Estimated
u_0 (Initial condition of ECM)	0.5×10^6	0.5	$Cells/mm^3$	Estimated
v_0 (Initial condition of viruses)	0	0	$Cells/mm^3$	Estimated

In the following, we will discuss the outcomes of varying some parameters in the model. The main focus here is on the key parameters that are relevant to the oncolytic viruses. All the other parameter remain the same as in Table 1.

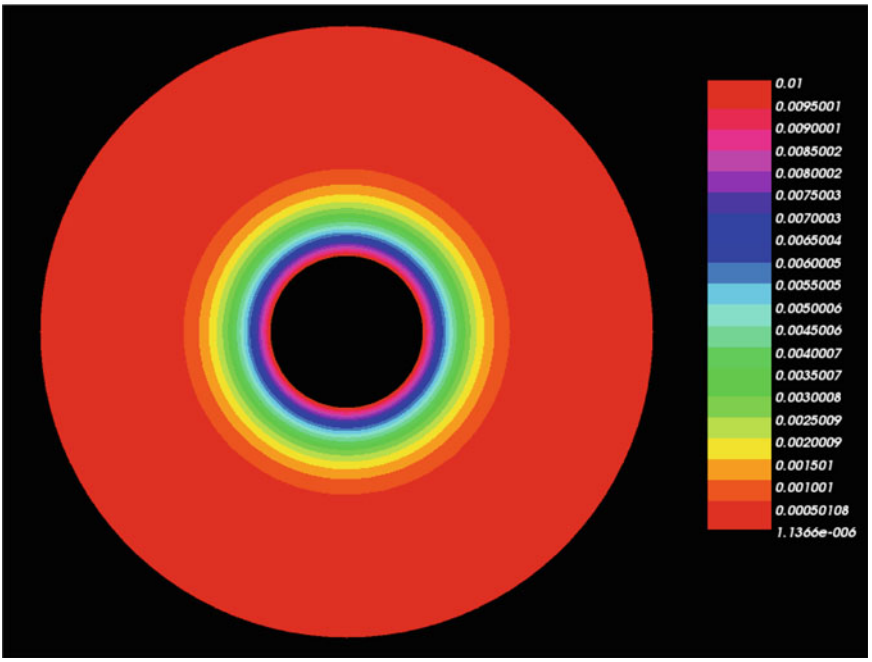
Virus Replication Rate (Burst Size) Known to be crucial within virotherapy treatment, the virus replication rate is the number of virus particles released by an infected cancer cell. We start our investigation by focusing first on the rate b at which the virus duplicate. We perform numerical experiments with $b = 0$, $b = 2$, $b = 4$, $b = 9$ and $b = 200$. Consequently, we found that there are two threshold value for the burst size. When the burst size is below the first threshold value the virotherapy always fails and the tumour grows to its maximum (carrying capacity) size (see Fig. 4). When the burst size is between the two threshold values, we obtain a partial success of the virotherapy as illustrated in Figs. 5 and 6. As an increase in the burst size will lead to a decrease in the tumour load, we can reach a minimum tumour load by genetically increasing the burst size of the virus up to the second threshold value, and still have a stable partial therapeutic success. However, once the burst size is above the second threshold, we obtain periodic oscillations with decay (Fig. 7), and if the burst size is large enough the tumour load can drop to an undetectable level then the cancer starts growing again (see Fig. 8). While the concept of oscillating tumour may seem abnormal, such behaviour have been seen in experimental observations and also in several ODE based viral dynamic models [15, 17, 23, 44, 45]. Furthermore, the oscillatory behaviour can be seen as analogous to the typical behaviour of predator-prey systems, where each population depends on the other for survival. The long period orbit can be viewed as a complete tumour eradication or tumour remission. In fact, a long period orbit can be biologically interpreted also as a complete tumour eradication: if the cancer cells population drops below certain level, this could signify extinction especially if we take into account increased clearance due to nutrient deficiency or a moderate reinforcement of virotherapy with another type of therapy. Long period orbits can also be considered as a remission or recurrence where reducing the amplitude of the orbit and increasing the period between the oscillations correspond to a more successful treatment. Furthermore, increasing exaggeratedly the burst size leads to an increase in the ECM density which will reduce the viral diffusion (see Fig. 8d).

Infection Rate of Cancer Cells In the following, we study the influence of viral infection rate (ρ) on the cancer dynamic. Experimental studies have shown that increasing the infection rate of cancer cells plays a pivotal role in the development of new anti-cancer therapies [46]. Aiming to investigate this aspect, we performed several simulation tests where we decreased and increased the baseline value $\rho = 79 \times 10^{-3}$ by a factor of three (to $\frac{\rho}{3}$ and 3ρ, respectively). As illustrated in Fig. 9 a three fold decrease in the infection rate leads to a poorer elimination of the cancer cells, compared to the case where the infection rate is increased three fold.

Death Rate of Infected Cancer Cells The death rate of infected cancer cells is the rate at which the infected cells are eliminated from the system by anti-viral

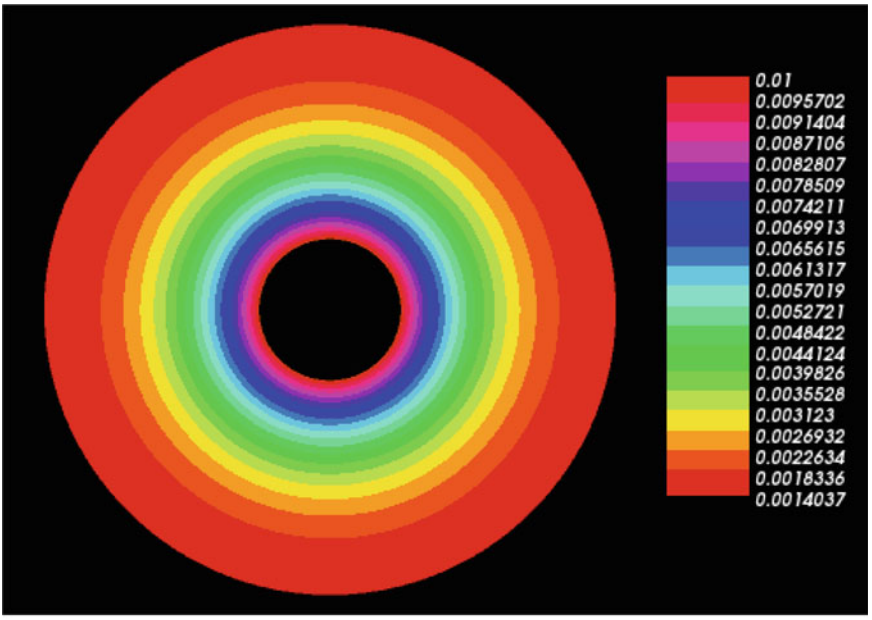


(a)

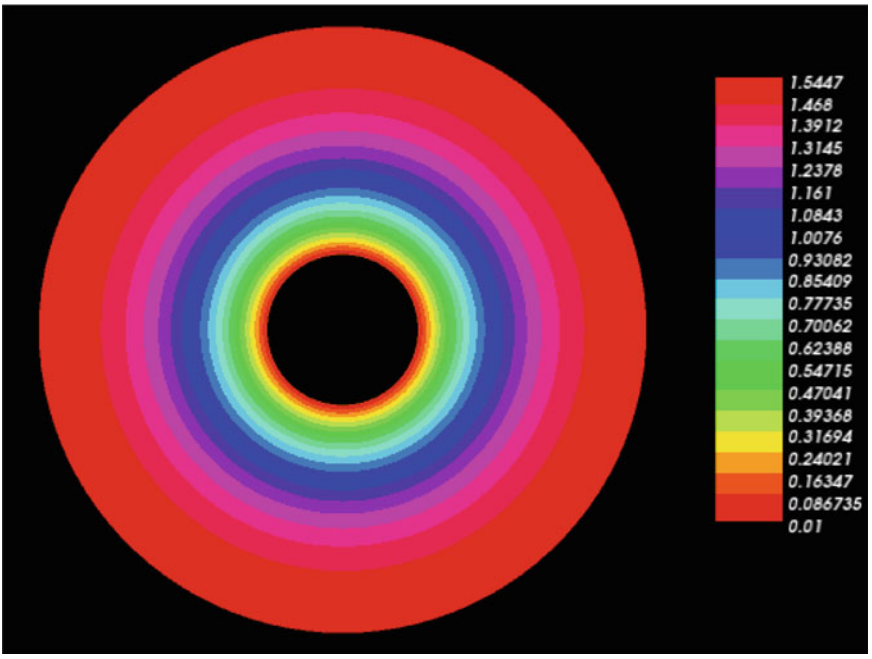


(b)

Fig. 2 Spatial propagation of the oncolytic virus for different relative time stages. (**a**) t=0. (**b**) t=10



(a)



(b)

Fig. 3 Spatial propagation of the oncolytic virus from relative time t=100 to T=720. (**a**) t=100. (**b**) T=720

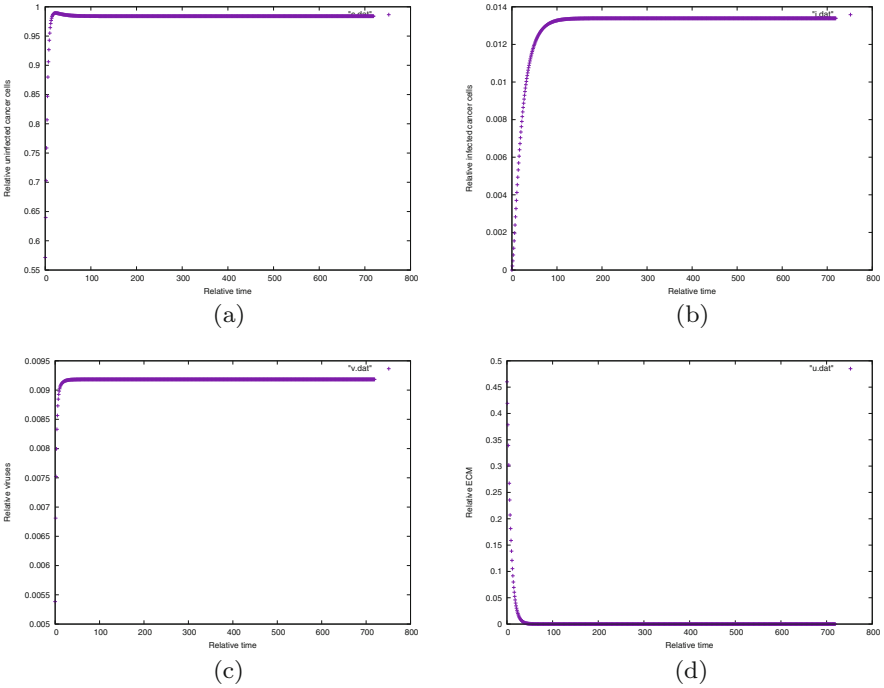


Fig. 4 Dynamics of the model at point $(0, 0.16)$ for $b = 0$. (**a**) Uninfected tumour cells. (**b**) Infected tumour cells. (**c**) Viruses. (**d**) ECM

immune response [47]. We reduce δ_i by a factor of four ($\delta_i/4$), then we notice that uninfected cancer cells are reduced dramatically (see Fig. 10d). This result is expected because the persistence of infected cancer cells promote the replication of more viruses inside these cells. Contrarily, increasing δ_i by a factor of four ($4\delta_i$) results in a weaker suppression of cancer cells as illustrated in Fig. 10a–c. However, the parameter δ_i alone does not show a clear influence towards the elimination of the tumour cells.

5 Conclusion

In this study, we introduced a system of partial differential equations coupled to an ordinary differential equation to simulate the treatment of cancer by using therapeutic viruses. The mathematical model illustrates the spatiotemporal dynamics between virotherapy, infected and uninfected cancer cells. The nonlinear terms in the coupled equations are handled semi-implicitly using an iteration of fixed-point type. The numerical simulations were carried out for different values of

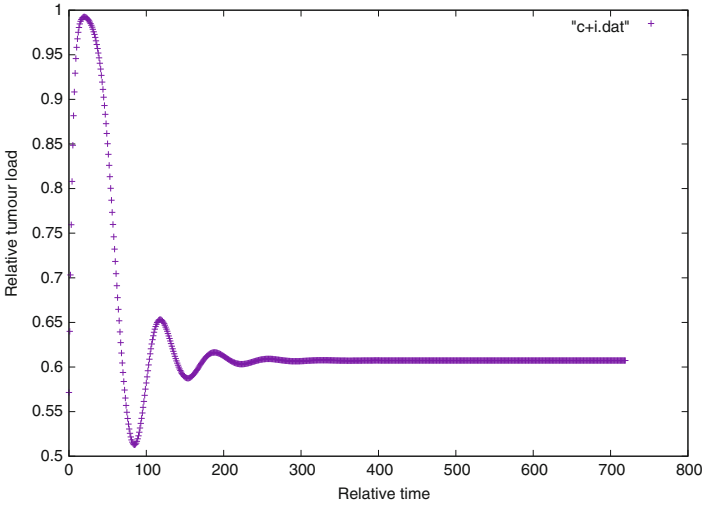


Fig. 5 Evolution of the tumour load at point $(0, 0.16)$ for $b = 2$

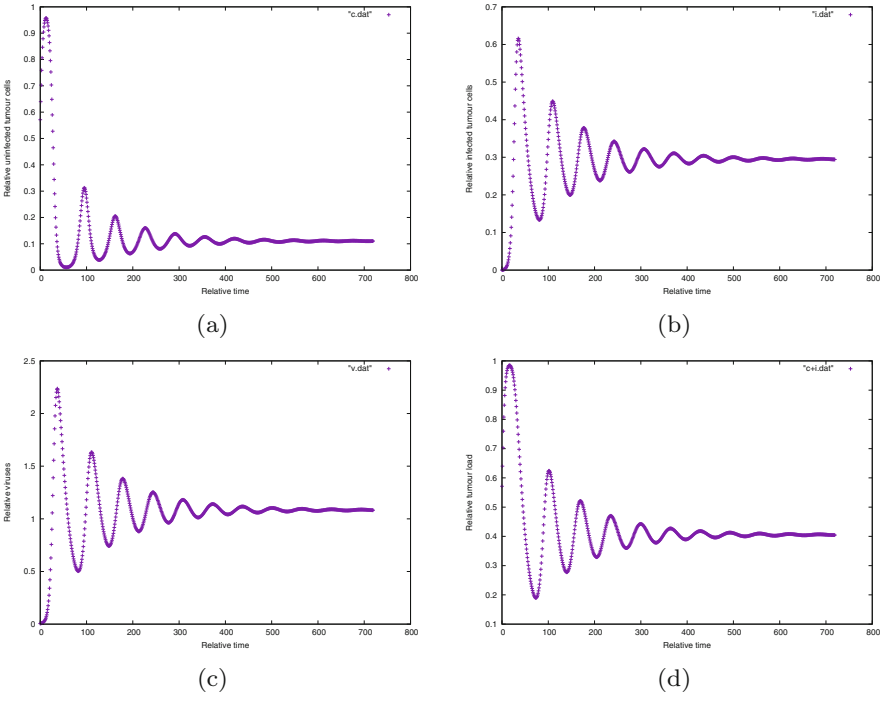


Fig. 6 Dynamics of the model at point $(0, 0.16)$ for $b = 4$. (**a**) Uninfected tumour cells. (**b**) Infected tumour cells. (**c**) Viruses. (**d**) Tumour load

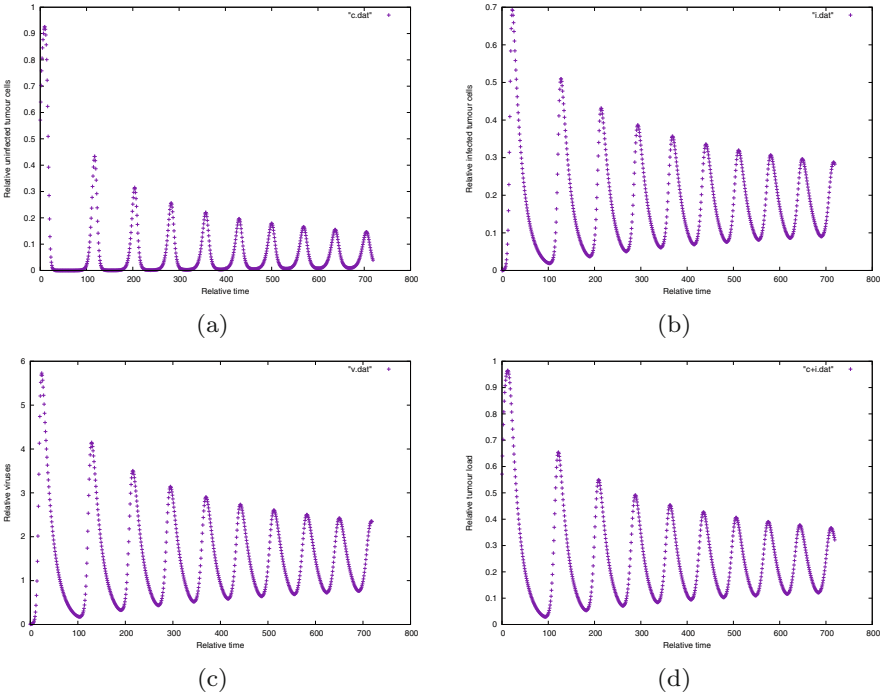


Fig. 7 Dynamics of the model at point $(0, 0.16)$ for $b = 9$. (**a**) Uninfected tumour cells. (**b**) Infected tumour cells. (**c**) Viruses. (**d**) Tumour load

the parameters related to virus dynamics and spread, namely, the burst size, the infection rate, and the clearance rate of infected cancer cells. The results showed that an improved therapy can be obtained by increasing the burst size, increasing the infection rate, and decreasing the death rate of infected cancer cells. However, since the tumour load can drop to an undetectable level then grows back, this mean that oncolytic virotherapy may not be able to eliminate all tumour cells from the body tissue. Thus, it's necessary to incorporate another treatment with virotherapy. In our study, the viruses are delivered continuously, a future research could include different delivery methods for viruses. Finally, another research point could be an investigation of the role of the immune system in the virotherapy.

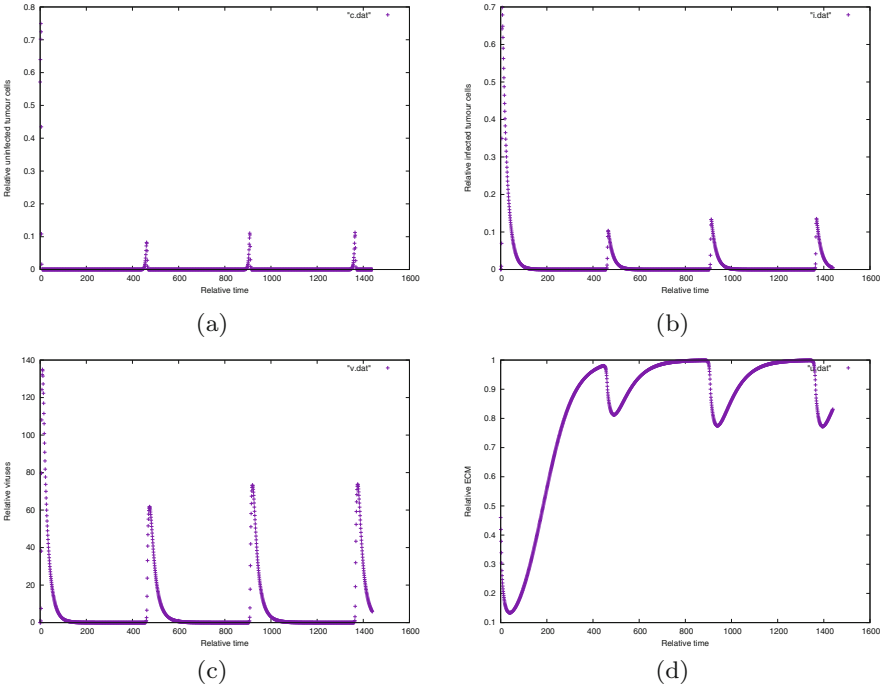


Fig. 8 Dynamics of the model at point $(0, 0.16)$ for $b = 200$. (**a**) Uninfected tumour cells. (**b**) Infected tumour cells. (**c**) Viruses. (**d**) ECM

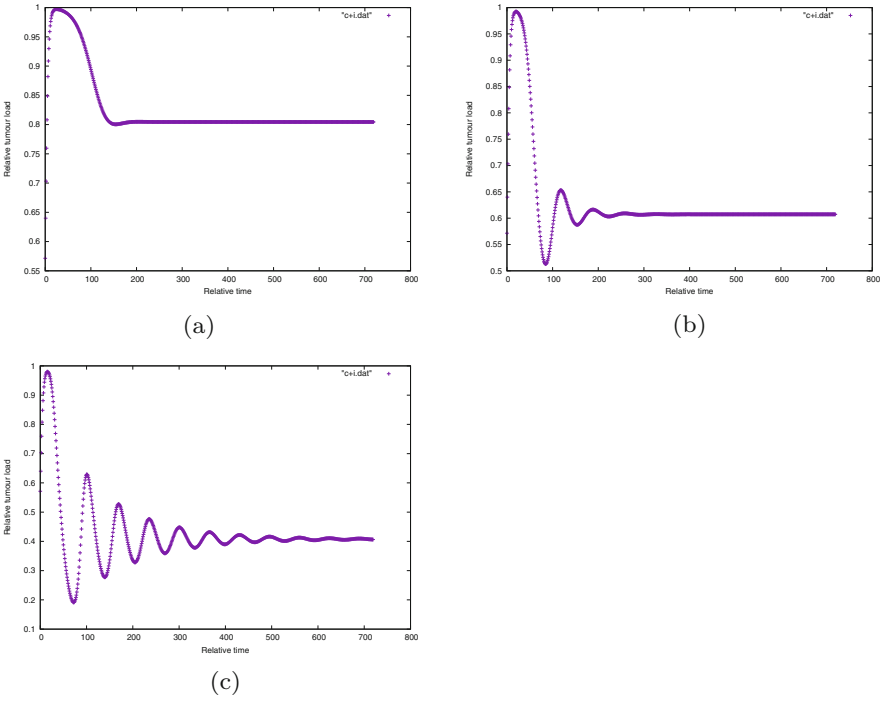


Fig. 9 Evolution in time of the tumour load at point $(0, 0.16)$ showing two variations of the baseline viral infection rate ρ (namely: (**a**). Tumour load for $\frac{\rho}{3}$; (**b**). Tumour load for ρ; and (**c**). Tumour load for 3ρ)

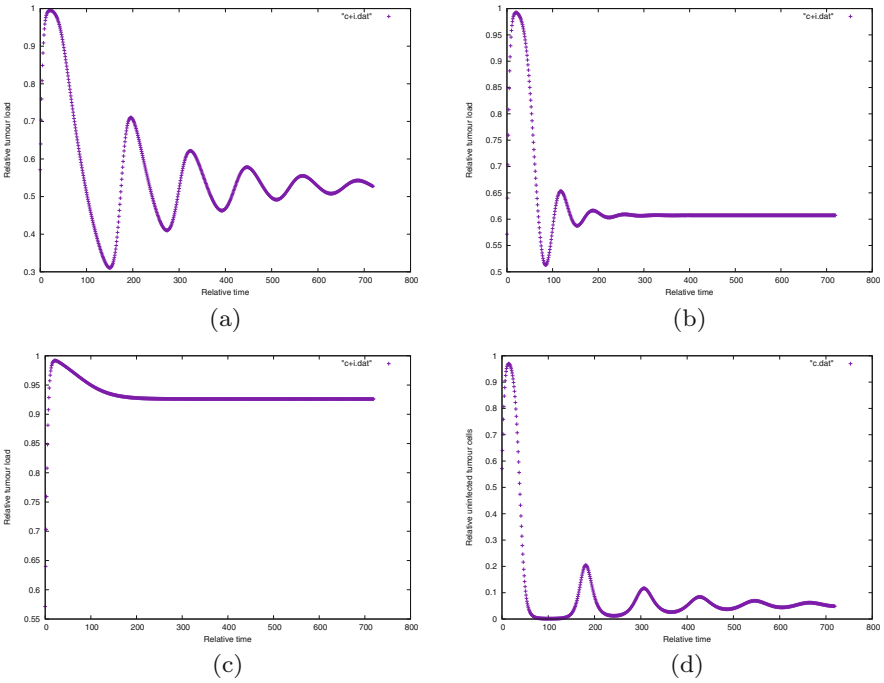


Fig. 10 Evolution in time of the tumour load at point $(0, 0.16)$ showing two variations of the baseline infected cancer cells death rate δ_i (namely: (**a**) Tumour load for $\delta_i/4$, (**b**) Tumour load for δ_i, and (**c**) Tumour load for $4\delta_i$); and (**d**) the evolution of Uninfected cancer cells for $\delta_i/4$

Reference

1. C. Fountzilas, S. Patel and D. Mhalingam, *Oncotarget* **8**, 1729 (2017).
2. H. L. Kaufman, F. J. Kohlhapp and A. Zloza, *Nat. Rev. Drug Discov* **14**, 642 (2015).
3. J. Pol, G. Kroemer and L. Galluzzi, *Oncoimmunology* **5**, e1115641 (2016).
4. S. J. Russel, K. W. Peng and J. C. Bell, *Nat. Biotechnol.* **30**, (2012).
5. A. Timalsina, J. P. Tian and J. Wang, *Bull. Math. Biol.* **79**, (2017).
6. H. Fukuhara, Y. Ino and T. Todo, *Cancer. Sci.* **107**, (2017).
7. R. Alemany, *Clinical and Translational Oncology* **15**, 182 (2013).
8. H. Wong, N. Lemoine and Y. Wang, *Viruses* **2**, 1 (2010).
9. B. Kaur, T. Cripe and E. Chiocca, *Curr Gen Ther* **9**, (2009).
10. N. Kuriyama, H. Kuriyama, C. Julin, K. Lamborn and M. Israel, *Hum Gene Ther* **11**, 1729 (2000).
11. N. Dmitrieva, L. Yu, M. Viapiano, T. Cripe, E. Chiocca, J. Glorioso and B. Kaur, *Clin. Cancer Res.* **17**, 1362 (2011).
12. D. Wodarz, *Gene Ther Mol Biol* **8**, 137 (2004).
13. D. Wodarz, *Cancer Research* **61**, 3501 (2001).
14. D. Wodarz, *Human Gene Therapy* **14**, 153 (2003).
15. N. Komarova and D. Wodarz, *Journal of Theoretical Biology* **263**, 530 (2010).
16. A. S. Novozhilov, F. S. Berezovskaya, E. V. Koonin and G. P. Karev, *Biology Direct* **1**, 1 (2006).
17. Z. Bajzer, T. Carr, K. Josic, S. G. Russel and D. Dingli, *Journal of Theoretical Biology* **252**, 109 (2008).
18. M. Biesecker, J.-H. Kimn, H. Lu, D. Dingli and Z. Bajzer, *Bulletin of Mathematical Biology* **vol. 72**, 469 (2010).
19. A. Friedman, P. Tian, G. Fulci, A. Chiocca and J. Wang, *Cancer Research* **Vol 66**, 2314 (2006).
20. P. Tian, *Mathematical Biosciences and Engineering* **Vol 8**, 841 (2011).
21. A. E.-A. Laaroussi, M. E. Hia, M. Rachik, E. Benlahmar and Z. Rachik, *Applied Mathematical Sciences* **8**, 929 (2014).
22. P. S. Kim, j. j. Crivelli, I.-K. Choi, C.-O. Yun and J. R. Wares, *Mathematical Biosciences and Engineering* **12**, 841 (2015).
23. D. Dingli, C. Offord, R. Myers, K.-W. Peng, T. Carr, K. Josic, S. Russell and Z. Bajzer, *Cancer Gene Therapy* **16**, 873 (2009).
24. M. A. Nowak and R. M. May, *Oxford University Press*, (2000).
25. C. Alvarez-Breckenridge, B. Choi, C. Suryadevara and E. Chiocca, *Curr. Opin. Virol.* **13**, 25 (2015).
26. M. A. Nowak, C. R. M. Bangham, *Science* **272**, 74 (1996). **2**, 131 (2011).
27. J. T. Wu, H. M. Byrne, D. H. Kirn and L. M. Wein, *Bulletin of Mathematical Biology* **vol. 63**, 731 (2001).
28. L. M. Wein, J. T. Wu and D. H. Kirn, *Cancer Research* **vol. 63**, 1371 (2003).
29. B. I. Camara, H. Mokrani and E. Afenya, *Mathematical Biosciences and Engineering* **10**, 565 (2013).
30. J. Malinzi, P. Sibanda and H. Mambili-Mamboundou, *Mathematical Biosciences* **263**, 102 (2015).
31. J. Malinzi, A. Eladdadi and P. Sibanda, *J. Biological Dynamics* **11**, 244 (2017).
32. J. Malinzi, R. Ouifki, A. Eladdadi, D. F. M. Torres and K. A. J. White, *Math. Biosci. Eng.* **15**, 1435 (2018).
33. S. B. Carter, *Nature***213**, 256 (1967).
34. A. S. G. Curtis, *Journal of Embryology and Experimental Morphology***22**, 305 (1969).
35. J. Pascal, E. L. Bearer, Z. Wang, E. J. Koay, S. A. Curley and V. Cristini, *Proceeding of the National Academy of Sciences* **vol. 110**, 14266 (2013).
36. Z. Wang, R. Kerketta, Y. L. Chuang, *PLoS Computational Biology* **vol. 12**, e1004969 (2016).
37. K. T. Nguyen, *Journal of Nanomedicine & Nanotechnology* **vol. 2**, (2011).
38. A. Laird, *J. Cancer***18**, 490 (1964).

39. C. Guiot, P. Degiorgis, P. Delsanto, P. Gabriele and T. Diesboeck, *Journal of Theoretical Biology* **225**, 147 (2003).
40. T. Cox and J. Erler, *Dis. Model Mech* **4**, 165 (2011).
41. M. A. J. Chaplain and G. Lolas, *Netw. Heterog. Media* **1**, 399 (2006).
42. S. Ganesan and L. Tobiska, *Internat. J. Numer. Methods Fluids* **57**, 119 (2008).
43. L. Peng, D. Trucu, P. Lin, A. Thompson and M. A. J. Chaplain, *Bulletin of Mathematical Biology* **79**, 389 (2017).
44. L. R. Paiva, C. Binny, S. C. Ferreira and M. L. Martins, *Cancer Research* **vol. 69**, 1205 (2009).
45. M. I. Titze, J. Frank, M. Ehrhardt, S. Smola, N. Graf and T. Lehr, *European Journal of Pharmaceutical Sciences* **vol. 97**, 38 (2017).
46. J. Maroun, M. M. Noz Alia, A. Ammayappan, A. Schulz, K. W. Peng and S. Russel, *Future Medicine* **12**, 193 (2017).
47. A. Filley and M. Dey, *Frontiers in Oncology* **106**, 8 (2017).

Propensity Matrix Method for Age Dependent Stochastic Infectious Disease Models

P. Boldog, N. Bogya, and Z. Vizi

1 Introduction

Quick and effective response to an emerging infection requires a large amount of information about the course of the epidemic (such as infection and recovery rate, latency period, etc.) and state of the art mathematical models. As it has been observed during the COVID-19 pandemic, the information at our disposal is highly limited, especially at the very beginning of the disease outbreak—when we have to act quickly. Thus, our models, and therefore their computer implementations have to be modified drastically on a daily basis.

Experts from the field of mathematical epidemiology, such as Rost et al. [1] and Barbarossa [2], usually apply deterministic models to predict the spread of the disease as a first approximation. Nevertheless, realistic models (with several classes of infection, age structure or spatial patches) tend to get difficult to analyze. Furthermore, deterministic models do not produce information about several important aspects of the epidemic, such as the variance of the state variables or the probability of certain events (like extinction) that are particularly important at the beginning of the outbreak.

Stochastic modeling approach offers feasible alternatives for tackling the above mentioned problems. Although the governing equation (the so called stochastic master equation) is often analitically intractable, numerical simulations of the corresponding Markov process and the average of the generated time series provides valuable information of several aspects of the epidemic.

Gillespie's stochastic simulation algorithm [4] (SSA) was originally designed to produce exact realisations of the stochastic master equation in case of coupled

P. Boldog (✉) · N. Bogya · Z. Vizi
University of Szeged, Bolyai Institute, Szeged, Hungary

R. P. Mondaini (ed.), *Trends in Biomathematics: Stability and Oscillations in Environmental, Social, and Biological Models*,
https://doi.org/10.1007/978-3-031-12515-7_17

chemical reaction systems. Since then, the scope of the SSA has been extended [7] in the last 40 years to stochastically simulate other chemical or biological phenomena. Gillespie's algorithm also yields a convenient method of simulation to gather information about an epidemic in the stochastic approach.

During the COVID-19 pandemic, one important lesson was that, quick implementation of—possibly complicated—stochastic epidemic models with several age groups is crucial for effective forecasting. The aim of the present paper is to provide a method that enables researchers to quickly prototype, build and modify robust stochastic epidemic models with age structure.

The structure of the paper is as follows: in the remaining subsections of the Introduction, we introduce Gillespie's first reaction method and provide a Python code for the case of the SIR model. In Sect. 2, we introduce our propensity matrix approach and extend our model implementation with age structure. Then, we demonstrate the flexibility of our approach by generalizing the model with other state variables to investigate an epidemiologically more realistic age-stratified model containing the E (exposed) and the D (dead) classes additionally. In Sect. 3, we introduce the update graph to make the algorithm much more efficient in terms of computation. Then in Sect. 4, we demonstrate the strength of stochastic epidemic models via some experiments with a hypothetical population. Finally, we summarize and discuss our results in Sect. 5.

1.1 Gillespie's Stochastic Simulation Algorithms

Considering a group of chemical species $(S_1, S_2, \ldots, S_n)$ reacting in a coupled system with reaction channels $(R_1, R_2, \ldots, R_m)$ in a well mixed environment, Gillespie's stochastic simulation algorithms are based on answering two questions [4]:

(i) What kind of reaction $(R_1, R_2, \ldots, R_m)$ will the next reaction be?
(ii) In what time (τ) will it occur?

Gillespie formulated two distinct, but mathematically identical versions of the SSA: the so called 'direct method' (DM) and the 'first-reaction method' (FRM). The main difference is in the way of deciding which reaction channel to fire. The *first-reaction method* generates m random time values $(\tau_1, \ldots, \tau_m)$ for the m possible reactions from the corresponding exponential distributions and selects the reaction channel with the least time to fire. Alternatively, the *direct method* only requires two random numbers (one from a uniform distribution and one from an exponential distribution). Considering the fact that if there are more than 2 possible reactions then the FRM requires more computation and memory than the DM, moreover drawing m random numbers from exponential distributions requires calculating m logarithms, we may conclude that the computational cost of the *first-reaction method* is even greater [3]. As we are about to work with robust systems of reactions, we only review and make use of the *direct method* in the following.

1.1.1 The Direct Method Algorithm

The direct method SSA creates stochastic realizations of the corresponding Markov chain that is continuous in time and discrete in state variables. Technically, starting from the vector of initial values $(X_1(0), \ldots, X_n(0)) \in \mathbb{N}_0^n$, the time series of the state variables $(X_1(t), \ldots, X_n(t)) \in \mathbb{N}_0^n$ is generated by constantly updating the state variables in properly generated subsequent times (τ) according to the answers to question (i) and (ii).

To this end, the reaction probability density function, $P(\tau, \mu)d\tau$, is defined [3]. That is, at time t, the next reaction in the reaction chamber will occur in the differential time interval $(t + \tau, t + \tau + d\tau)$ with the probability $P(\tau, \mu)d\tau$, and will be an R_μ reaction $(\mu \in 1, \ldots, m)$.

It can be shown that [3], with the procedure called conditioning, the two variable density function $P(\tau, \mu)$ can be written as the product of two one-variable probability density functions:

$$P(\tau, \mu) = P_1(\tau) \cdot P_2(\mu|\tau). \tag{1}$$

In particular, from the paper of Gillespie [3], it turns out to be

$$P(\tau, \mu) = a_\mu \cdot e^{-\tau \cdot a},$$

where the symbol a_μ stands for the so called propensity function that characterizes reaction R_μ, and may depend on the quality, and actual number of the reactants or the environment, etc. For convenience, we use the notation $a = \sum_{\mu=1}^{m} a_\mu$ for the sum of the propensity functions.

At time t let $P_1(\tau)d\tau$ be the probability that the next reaction will occur between times $t + \tau$ and $t + \tau + d\tau$, independent of which reaction it might be. Similarly, $P_2(\mu|\tau)$ is the probability that the next reaction will be the R_μ reaction, given that the next reaction occurs at $t + \tau$. The probability $P_1(\tau)$ is obtained by summing the reaction probability density function over all possible μ values:

$$P_1(\tau) = \sum_{\mu=1}^{m} P(\tau, \mu) = a \cdot e^{-\tau \cdot a}. \tag{2}$$

Substituting $P_1(\tau)$ into (1) and solving for $P_2(\mu|\tau)$ we obtain

$$P_2(\mu|\tau) = \frac{a_\mu}{a}. \tag{3}$$

It is clear that $\int_0^\infty P_1(\tau)d\tau = \int_0^\infty a \cdot e^{-\tau \cdot a} = 1$. Moreover, $\sum_{\mu=1}^{m} P_2(\mu|\tau) = \sum_{\mu=1}^{m} \frac{a_\mu}{a} = 1$.

With these notations and probability distributions, the direct method SSA can be described as follows.

(1) **Initialization:**

(i) set $t \leftarrow 0$,
(ii) set initial values $(X_1(0), \ldots, X_n(0)) \in \mathbb{N}_0^n$,
(iii) prescribe halting conditions C_H.

(2) **Calculate propensity functions** a_μ for all $\mu \in 1, 2, \ldots, m$.
(3) **Decide when the next reaction will occur:** choose τ according to Eq. (2),

(i) choose r_1 from $(0, 1)$ with a uniform distribution,
(ii) obtain $\tau = (1/a)\ln(1/r_1)$.

(4) **Decide which reaction occurs:** choose μ according to Eq. (3),

(i) choose r_2 from $(0, 1)$ with a uniform distribution,
(ii) take μ to be the integer for which

$$\sum_{j=1}^{\mu-1} a_j < r_2 a \leq \sum_{j=1}^{\mu} a_j.$$

(5) **Update time and state variables:**

(i) calculate change $(\Delta X_1, \ldots, \Delta X_n) \in \mathbb{Z}^n$ in number of reactants according to μ,
(ii) change the number of molecules according to

$$(X_1(t+\tau), \ldots, X_n(t+\tau)) = (X_1(t), \ldots, X_n(t)) + (\Delta X_1, \ldots, \Delta X_n),$$

(iii) set $t \leftarrow t + \tau$.

(6) **Halt if** $C_H = True$ **else continue the process with Step (2).**

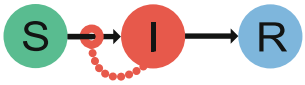
1.2 Gillespie's Algorithm for the SIR Model

The well known deterministic SIR model is a system of ordinary differential equations (4), and it is one of the simplest models to capture the spread of an infection in a population without demography. It separates the population into three classes: $S(t)$—the number of susceptible, $I(t)$—the number of infected, and $R(t)$—the number of recovered individuals in the population at time t, the symbol $()'$ stands for the usual time derivative. The model defines two reactions that we will call events from now on. Susceptible individuals *get infected* with rate β by making effective contact with infected individuals in the population. Infected individuals *recover* at a rate γ, and the per capita contact between S and I is $\frac{SI}{N}$. We consider a constant

Table 1 The table shows the possible events in the SSA in case of the SIR epidemic model

Event	Transition	Change ($\Delta S, \Delta I, \Delta R$)	Propensity	Probability
Infection	$S \to I$	$(-1, +1, 0)$	$a_1 = \beta \frac{SI}{N}$	$p_1 = \frac{a_1}{a}$
Recovery	$I \to R$	$(0, -1, +1)$	$a_2 = \gamma I$	$p_2 = \frac{a_2}{a}$

population size $N = S(t) + I(t) + R(t)$ and the following rate of change in the classes:



$$\begin{cases} S' &= -\beta \frac{SI}{N} \\ I' &= \beta \frac{SI}{N} - \gamma I \\ R' &= \gamma I. \end{cases} \tag{4}$$

In this section, by following the work of Gillespie [4] and Allen [5], we apply the *direct method* SSA to the SIR model (4). In the stochastic simulation algorithm "... the reaction constants are viewed not as reaction 'rates' but as reaction 'probabilities per unit time' ..." [3]. Thus, we follow the practice of Allen [5] and use the same notations for the corresponding parameters in both aproaches.

The code we implemented can be found in our GitHub repository [11]. It is important to note that, as we prepare to handle robust models in a flexible way, we will not follow the common practice of reducing the state variables with decoupling (4) by using the fact that N is constant in time.

In the stochastic SIR model, state variables $S(t), I(t), R(t) \in \{0, 1, \ldots, N\}$, and $t \in [0, \infty)$. In Table 1, we summarise the possible events with the corresponding transitions, change in state variables, and we also define the belonging propensity functions and probabilities. The intervent time τ is generated from the exponential distribution (2) with parameter $1/a$, where $a = \beta \frac{SI}{N} + \gamma I$ and the probability distribution of the possible events is $p_1 = a_1/a$ and $p_2 = a_2/a$ (with propensities $a_1 = \beta \frac{SI}{N}, a_2 = \gamma I$).

2 The Propensity Matrix Method for Age Dependent Stochastic Epidemic Models

In order to extend the SSA algorithm to handle the required age structure, we make the following definitions and assumptions:

- assume a population of N individuals,
- the population is divided into l *classes* ($l = 3$ in the SIR model) according to the characteristics of the epidemic: $X_1, \ldots, X_l$,

- the population is stratified into *n age groups*,
- this partition of the population leads us to the $n \cdot l$ *compartments* (that we will call *state variables* as well) at time t:

$$X_1^1(t), X_2^1(t), \ldots, X_{l-1}^n(t), X_l^n(t)$$

- an *event* (or reaction) during the simulation corresponds to a *transition* of an individual from one compartment to another in the same age group,
- we consider the epidemic in such a short time scale when aging is not significant, thus we assume there is no transition between age groups,
- we assume that there is no demography (birth or natural death) during this short time—if death is incorporated in the model, it only occurs due to the infection,
- the same transitions occur in every age group—possibly with zero rate.

Now, we define the propensity matrices as follows: for age group i, we define *propensity matrix* M^i with index i in the following way: element $M^i(j,k) = a^i_{j,k}$ of the propensity matrix represents the propensity function corresponding to the transition of an individual from compartment $X^i_j \to X^i_k$ (at age group i).

In case of particular models when it is clear from the context we would use the more convenient $a^i_{X_j,X_k}$ notation for the propensity of transition $X^i_j \to X^i_k$. For instance in case of the age structured SIR model we do not bother with the complicated indexing, simply note the propensity of transition $S^i \to I^i$ by $a^i_{S,I}$. It is clear that the propensity matrix is the adjacency matrix of the transition graph containing zero elements whenever a transition never happens and the propensities respectively (cf. Fig. 2).

During the simulation, we calculate the propensity matrix for every age group. Since, we observe the same transitions in every age group, we can outsource the repeating calculations into n cycles, obtaining a cleaner and more compact code.

Then we choose the intervent time τ from distribution (2), where a is the sum of all propensities over all propensity matrices:

$$a = \sum_{g=1}^{n} \sum_{f=1}^{l} \sum_{h=1}^{l} a^g_{f,h}, \tag{5}$$

and choose the next event from distribution

$$P_2\left((\text{transition } j \to k \text{ at age group } i)|\tau\right) = p^i_{j,k} = \frac{a^i_{j,k}}{a}. \tag{6}$$

Basically, one can obtain a transition from distribution (6) by choosing a random real $r \in (0,1)$ with uniform distribution and summing the propensities of all transitions over all age groups, then choosing transition $X_j \to X_k$ at age group i whenever

$$\sum_{g=1}^{n}\sum_{f=1}^{l}\sum_{h=1}^{l} a_{f,h}^{g} < r \cdot a \leq \sum_{g=1}^{i}\sum_{f=1}^{j}\sum_{h=1}^{k} a_{f,h}^{g}.$$

Eventually, we have to update time: $t \leftarrow t+\tau$ and the state variables with the change vector $(\Delta X_1^1, \ldots, \Delta X_l^n)$, where only $\Delta X_j^i = -1$ and $\Delta X_k^i = +1$ are nonzero elements:

$$(X_1^1(t+\tau), \ldots, X_l^n(t+\tau)) = (X_1^1(t), \ldots, X_l^n(t)) + (\Delta X_1^1, \ldots, \Delta X_l^n).$$

With these assumptions and notations we are prepared to present the propensity matrix algorithm that derives the time evolution of the state variables.

2.1 Algorithm Propensity Matrix Method

(0) Fix the order of the state variables: $(X_1^1, \ldots, X_l^n)$.
(1) **Initialize:** Set $t \leftarrow 0$, initial values $(X_1^1(0), \ldots, X_l^n(0))$, and halting condition C_H.
(2) **Calculate propensity matrices** M_i for all age groups $i \in \{1, \ldots, n\}$.
(3) **Choose intervent time** τ from $P(\tau) = a \cdot e^{-\tau \cdot a}$ where a comes from (5),
(4) **Choose the next reaction** from the distribution $p_{j,k}^i = \frac{a_{j,k}^i}{a}$ (6),
(5) **Update state variables and time.**
(6) **Halt** if $C_H = True$ else **continue** with step 2.

Propensity matrices of epidemic models are sparse, since, from one compartment, there are usually only one or very few transitions to other compartments. We also emphasise that this data structure provides a convenient way to easily modify our model by including new transitions (like waning immunity), or deleting transitions between age groups. Also, by extending the number of state variables, we can further introduce new classes (like exposed, latent, hospitalized, dead or recovered). In the following, we present the propensity matrices of two epidemic models.

2.2 The Age Structured SIR Model with Waning Immunity

When we extend model (4) with age structure we have to consider the effect of every infected age group on the susceptible population of age group i, furthermore to include waning immunity to (4) we consider the $R_i \rightarrow S_i$ transitions with age dependent rate ω_i, in the general case. Naturally, with $\omega_i = 0$ for $i \in \{1, \ldots, n\}$ we get back to the case where waning immunity is not applied in the process. Hence,

$$\begin{cases} a^i_{SI} = \frac{1}{N} S_i \cdot \sum_{f=1}^{n} \beta_f I_f \\ a^i_{SR} = \alpha_i S_i \\ a^i_{IR} = \gamma_i I_i \end{cases}$$

i	S	I	R
S	•	a^i_{SI}	a^i_{SR}
I	•	•	a^i_{IR}
R	•	•	•

Fig. 1 Propensity matrix and the corresponding propensities of the SIR model

with age dependent transmission rates β_i, recovery rates γ_i and waning rates ω_i we obtain the system of ODEs for age group i:

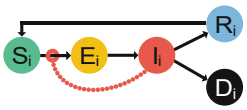
$$\begin{cases} S_i{}' = -\frac{1}{N} S_i \cdot \sum_{f=1}^{n} \beta_f I_f + \omega_i S_i \\ I_i{}' = \frac{1}{N} S_i \cdot \sum_{f=1}^{n} \beta_f I_f - \gamma I_i \\ R_i{}' = \gamma I_i - \omega_i S_i \end{cases} \tag{7}$$

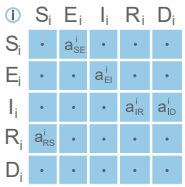
In the stochastic simulation formulation, let the order of state variables be ($X^i_1 = S_i, X^i_2 = I_i, X^i_3 = R_i$). Thus, the possible transitions and the belonging propensities (with the previously simplified noation) in age group i in this model are *infection* (a^i_{SI}), *waning immunity* (a^i_{SR}) and *recovery* (a^i_{IR}). The propensity matrix and the corresponding propensities can be seen in Fig. 1, where '•' stand for a zero element of the matrix. Python code for this model is included in the repository [11].

2.3 SEIRD Model with Age Structure and Waning Immunity

One strength of the propensity matrix technique is the flexibility in terms of computer implementation. For instance, several disease models require an exposed class E, as right after infection susceptibles usually do not show symptoms during the latency period of length $1/\epsilon_i$. Thus with rate ϵ people move from class E_i to class I_i. The corresponding propensity is $a^i_{EI} = \epsilon_i E_i$. We can further add class D to our model, that counts the death cases caused by the infection: with age dependent mortality rate $p_i\gamma$ people move from class I_i to D_i with the belonging propensity $a^i_{ID} = p_i\gamma I_i$. We also assume that recovered people loose immunity against the disease with rate ω, thus, the corresponding propensity is $a^i_{RS} = \omega_i R_i$.

Let the order of state variables at age group i be $X^i_1 = S_i, X^i_2 = E_i, X^i_3 = I_i, X^i_4 = R_i, X^i_5 = D_i$ and a^i_{jk} stands for the propensity of transition from





$$\begin{cases} S_i' = -\frac{1}{N_i} S_i \cdot \sum_{f=1}^{n} \beta_f I_f + \omega_i R_i \\ E_i' = \frac{1}{N_i} S_i \cdot \sum_{f=1}^{n} \beta_f I_f - \epsilon_i E_i \\ I_i' = \epsilon_i E_i - \gamma I_i \\ R_i' = (1-p_i)\gamma I_i - \omega_i R_i \\ D_i' = p_i \gamma I_i \end{cases}$$

Transition	Propensity
Infection	$a^i_{SE} = \frac{1}{N_i} S_i \cdot \sum_{f=1}^{n} \beta_f I_f$
Going infectious	$a^i_{EI} = \epsilon E_i$
Recovery	$a^i_{IR} = (1-p_i)\gamma I_i$
Death due to infection	$a^i_{ID} = p_i \gamma I_i$
Waning of immunity	$a^i_{RS} = \omega R_i$

Fig. 2 The SEIRD model with age structure and waning immunity. Possible transitions with the belonging propensities are: infection, going infectious from exposed, recovery, death due to infection and loss of immunity

compartment j to k at age group i. For flowchart of the process and the governing ODE model with the propensity matrix and the remaining propensities see Fig. 2.

3 Updating the Propensities

Gillespie [4] suggests that "...it is necessary to recalculate only those quantities a_v, corresponding to reactions R_v, whose reactant population levels were just altered..." in the reaction selection step.

Probably the most well known attempt to get around this concern belongs to Gibson and Bruck [8]. Their *Next Reaction Method* (NRM) may be regarded as an extension of Gillespie's original first reaction method. They define the so called reaction dependency graph that contains information about which propensity function (a_v) needs to be updated according to the chosen reaction in the selection step. However, in the detailed comparison of the DM, FRM and NRM by Cao et al. [6] it is found that "even with the best data structure, the NRM is less efficient than the DM except for a very specialized class of problems". As Scvehm [9] points out, this is mostly due to the fact that in case of the Gibson-Bruck method "the simulator engine spends most of its execution time for maintaining the priority queue of the tentative reaction times".

In this section we provide a dependency graph like method to make use of the fact that upon a transition event mostly only a small number of propensity functions have to be recalculated. For instance in case of the age structured SEIRD model in Sect. 2.3 the waning immunity event at age group i only changes state variable R_i and S_i hence it only influences propensity function a^i_{RS} and a^i_{SE}. All other propensities in age group i and every other propensities in the other age groups remain unchanged, that is $3 + (n-1) \cdot 5 = 5n - 2$ number of events in case of n age groups.

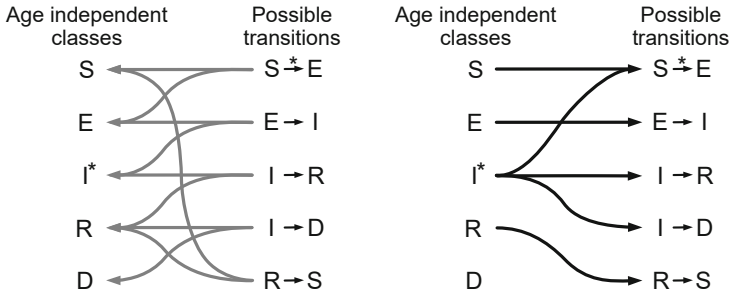


Fig. 3 Update graph of the SEIRD model. The figure shows the bipartite update graph that contains information about the classes and propensity functions to be updated upon event selection

In our methodology, in every iteration step after choosing the transition from distribution (6), we selectively update the propensity values based on this transition. We do this with the help of a suitable data structure that is a directed bipartite graph, called the *Update Graph* (UG) and is defined the following way: let C be the set of nodes associated with the *classes* (cf. Sect. 2) and let T be the other set of nodes representing the possible *transitions* between the classes. The edge $c \to t_{X_j,X_k}$ (where t_{X_j,X_k} is the transition from X_j to X_k) exists only if the class of a modified state variable $c \in C$ updates the propensity of transition $t_{X_j,X_k} \in T$. Also edge $t_{X_j,X_k} \to c$ exists only if $c = X_j$ or $c = X_k$. For convenience we note that from any node $t_{X_j,X_k} \in T$ there are always two edges pointing out, one to class X_j and the other one to class X_k.

However from class $c \in C$ there may be several edges pointing to different transitions, for instance in the $SEIRD$ model infection, recovery and death are all depending on the number of I individuals in a given age group. We also point out that (in case of the $SEIRD$ model) S_i individuals can be infected not only by individuals from compartment I_i but from all compartment I_g at all age groups, $g \in \{1, \ldots, n\}$. Therefore, change in any infectious compartment affects the propensity for transitions in all age groups. Thus upon an infection event all propensities a^g_{SE}, $g \in \{1, \ldots, n\}$ have to be updated. In more complicated models there may be several infectious classes (such as latent or in case of ebola models even dead individuals may cause infection). To this end, in general, we flag (*) the classes c^* that influence propensities across all age groups and also flag transitions t^* that have to be updated for all age group whenever a compartment from a flagged class changed. We handle this issue in the computer implementation.

For instance the two part of the UG of the SEIRD model with waning immunity is shown on Fig. 3. The left part of the figure shows edges $t_{X_jX_k} \to c$ and the right part shows edges $c \to t_{X_jX_k}$. Class I^* and transition $S \xrightarrow{*} E$ is flagged thus whenever an I^i compartment changes all propensities a^i_{SE} need to be updated.

During the simulation, after the event selection step we end up with the coordinate triple (i, j, k) that selects transition t_{X_j,X_k} at age group i. By using the

$t_{X_j,X_k} \to c$ edges of the graph (left figure) we update the state variables according to $X^i_j(t+\tau) = X^i_j(t) - 1$, and $X^i_k(t+\tau) = X^i_k(t) + 1$ and all other $X^i_f(t+\tau) = X^i_f(t)$ remain unchanged. Next we update the propensity a^i_{jk} according to the $c \to t_{X_j,X_k}$ edges. If we encounter and edge that points from a flagged class to a flagged transition then we updated all a^i_{jk} for $i \in \{1, \ldots, n\}$.

3.1 Propensity Matrix Method Extended with the Update Graph

(0) **Initialisation step:**

- Fix the order of the state variables: $(X^1_1, \ldots, X^n_l)$
- Construct the Update Graph according to Sect. 3
- Set $t \leftarrow 0$, initial values $(X^1_1(0), \ldots, X^i_l(0))$, and halting condition C_H.
- **Calculate propensity matrices** M_i for all age groups $i \in \{1, \ldots, n\}$

(1) **Selection step:**

- **Choose intervent time** from $P(\tau) = a \cdot e^{-\tau \cdot a}$ where a comes from (5),
- **Choose the next reaction** t_{X_j,X_k} at age group i from the distribution $p^i_{j,k} = \frac{a^i_{j,k}}{a}$ (6),

(2) **Update step:**

- **Time:** $t \leftarrow t + \tau$
- **Update state variables:** by using the $t_{X_j,X_k} \to c$ edges of the graph update $X^i_j(t+\tau) = X^i_j(t) - 1$, and $X^i_k(t+\tau) = X^i_k(t) + 1$ and all other $X^i_f(t+\tau) = X^i_f(t)$ remain unchanged,
- **Update propensities:** update the propensity a^i_{jk} according to the $c \to t_{X_j,X_k}$ edges. If we encounter and edge that points from a flagged class to a flagged transition then we updated all a^i_{jk} for $i \in \{1, \ldots, n\}$.

(3) **Halt** if $C_H = True$ else **continue** with step 2.

4 Experiments

In this section, we demonstrate the advantages of the stochastic approach and the flexibility of the propensity matrix—update graph method by experimenting with some real life problems.

During the following simulations, we consider a population of 200,000 individuals divided into three age groups (0–14, 15–59 and 60+), with the aggregated contact matrix from Prem et al. [10] and with the population distribution of Hungary from

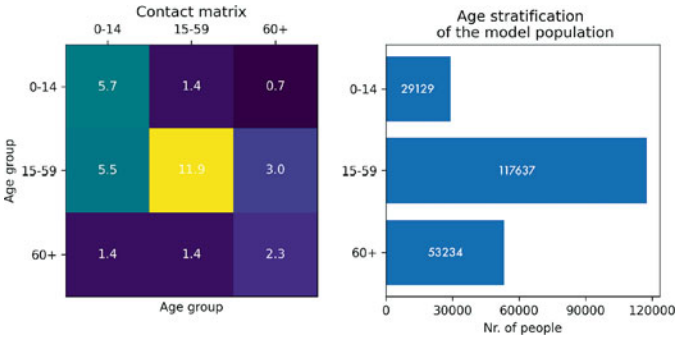


Fig. 4 Population related parameters of the demonstrated experiments. Agregated contact matrix (left) and age stratification (right) of the investigated model population. Both data collection corresponds to the Hungarian society

Table 2 Parameters of the demonstrated epidemic

Parameter	Value/age-dependent vector
Incubation period (ϵ^{-1})	5.2 (days)
Infectious period (γ^{-1})	5.0 (days)
Time spent immunized (ω^{-1})	180 (days)
Infection rate (β)	0.05
Probabilty of death (p_i)	(0.0000451, 0.00117, 0.0281)

the Hungarian Central Statistical Office (KSH) age-stratification [12] (cf. Fig. 4). We symmetrised the contact matrix according to Sec. 2.3.3. in Röst et al. [1].

Using the stochastic SEIRD model with immunity waning from Sec.2.3 we investigate the course of a COVID-19 epidemic with aggregated parameters from Rost et al. [1] (cf. Table 2). Every simulation starts with $I_1(0) = 10$ infected individuals in age group 1 at $t = 0$—all other individuals considered to be susceptible according to the age partition ($S_1(0) = 29119$, $S_2(0) = 117637$, $S_3(0) = 53234$). The governing propensity functions can be found in Sect. 2.3.

The left part of Fig. 5 shows the time series of the age aggregated state variables $\sum_{i=1}^{3} S_i$, $\sum_{i=1}^{3} E_i$, etc. from a single simulation outcome. We can observe that during the "first wave" the exposed class peaks before the infected class, and class S and R shows oscillation due to loss of immunity. The right part of Fig. 5 shows the time series of the three age groups of class S from the same simulation.

On Fig. 6 we focus on the peak size of the epidemic and show the effect of contact reduction in the early stage of the disease spread (on day 45 in this case). By 10, 20, and 30% of contact reduction we mean that we multiply every element of the contact matrix with 0.9, 0.8 and 0.7 respectively. The simulation was halted 10 days after the peak. We can notice that a uniform contact reduction of 30% in every age group may decrease the peak size by $1/3$, and in the meanwhile it also delays the time of the peak.

Montecarlo simulations, such as any version of Gillespie's SSA can only serve information about the process if we run several stochastic realisations and

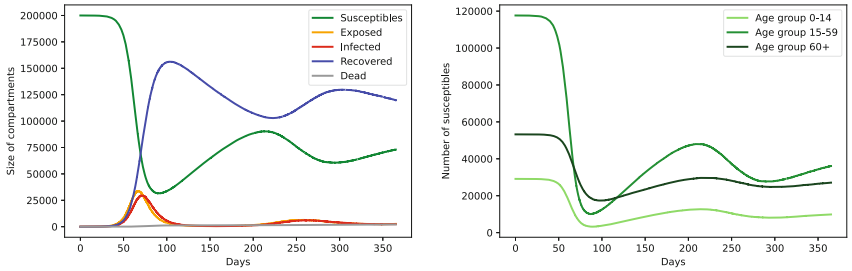


Fig. 5 Stochastic realisation of the SEIRD model with immunity waning. Time series of the age aggregated state variables (left), and the three susceptible age groups (right) from the same realisation of the stochastic SEIRD model. For the parameters cf. Fig. 4 and Table 2

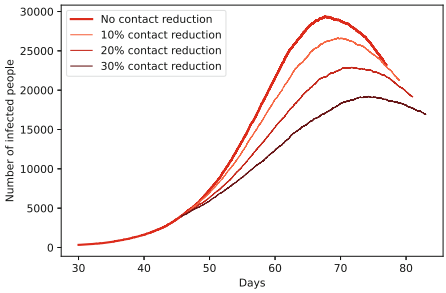


Fig. 6 Effect of contact reduction in the early stage of the epidemic. In this simulation we applied a 10, 20 and 30% uniform contact reduction in the total contact structure of the population. The simulation stops 10 days after the peak

investigate the statistics of them. Thus it is always important to know how many simulations we need for plausible conclusions.

In the following experiments we run a large number of simulations the following way: we obtain 100 simulations (for having enough data for statistical analysis) and for the daily sampled time series of every single compartment $X_j^i(t)$ we evaluate the mean $\mu_{X_j^i}(t)$ and variance $\sigma_{X_j^t}(t)$. We also calculate the mean value and variance for the time of the peak (μ_t, σ_t) and for the size of the peak (μ_s, σ_s). Then after each simulation we check whether 95% of the collected first peak sizes are in the confidence interval $[\mu_s - 2\sigma_s, \mu_s + 2\sigma_s]$ as well as in $[\mu_t - 2\sigma_t, \mu_t + 2\sigma_t]$ and stop the routine, if this condition is fulfilled.

On both part of Fig. 7, with solid red curve, we plotted the mean of the daily sampled time aggregated infected compartments $\sum_{i=1}^{3} \mu_{I_i}(t)$. The blue band on the left plot shows the minimal and maximal values we calculated for each days from the ensemble of the independent realisations of the process as well. Let us remark that the band at the top (around the peak) is almost flat, meaning that the different simulations produced similar peak sizes but different peak times. On the right part of Fig. 7 the highlighted rectangle shows the $2\sigma_t$ wide and $2\sigma_s$ high confidence interval around the maximum of the mean. Thus we may conclude that the peak of the epidemic occurs between day 63 and 79 and it is expected to be between approximately 27,900 and 28,600 with 95% probability.

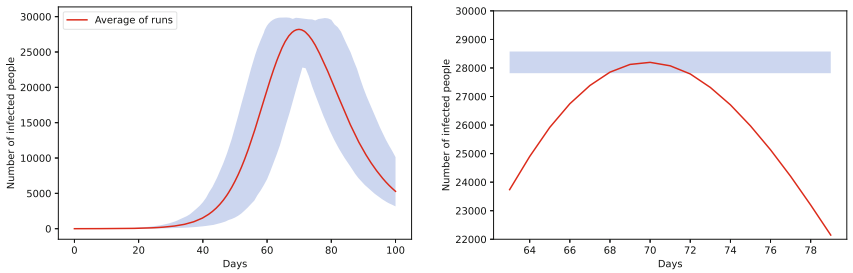


Fig. 7 Time series of the daily sampled aggregated infected classes. The red curve shows the daily sampled $\sum_{i=1}^{3} \mu_{I_i}(t)$ in case of 110 simulations around the peak of the first wave. The blue band on the left shows the daily sampled minimal and maximal values of the stochastic ensemble, the blue rectangle on the right represents the $2\sigma_t$ wide and $2\sigma_s$ high confidence interval around the maximum of the mean

5 Summary and Discussion

In this paper we introduced our approach that we use in our forecasting work to quickly implement age structured stochastic epidemic models. The core idea, the so called propensity matrix method, serves a data structure to handle the propensities and select the upcoming reaction/event to execute in a convenient way, in case of robust epidemic models. In Sect. 2 we introduced this method and the algorithm that obtains the time evolution of the state variables. In Sect. 3 we further improved the algorithm by introducing the so called update graph that helps us to speed up the algorithm by updating only the minimal number of propensities that are required to be updated due to the change in the state variables. Finally in Sect. 4 we showed some real life experiments to demonstrate the strength of the stochastic approach and the flexibility of our method.

We note here that after publicating the preprint on arXiv, we came across the article of Indurkhya and Beal [13] (2010), that also uses a similar directed bipartite graph (the so called *Update Dependency Graph*) to store information about the state variables and propensities to be updated. They also showed that their data structure is more memory efficient than the ones used in former aproaches as the memory they use scales with $O(n)$ instead of $O(n^2)$ in case of n reactions [13].

The speed of the simulation is usually a limiting factor of stochastic simulations. Generating a stochastic ensemble to analyse require several independent runs of the simulation and therefore is typically expensive in terms of computational time. From coding perspective interpreted languages like Python may have low performance. The simulation time may be reduced by several magnitudes using low-level languages like C or C++. The propensity matrix algorithm can be easily applied to such languages. We provide our sample code in Python to emphasize the algorithm and help readability as much as possible.

However, from the modeling perspective simulation of whole trajectories during years, from the outbreak untill the very end of the epidemic, is usually unnecessary

in practice. Information provided by stochastic simulations is particularly valuable (i) in the beginning of the epidemic—when only a small portion of the population is infected, (ii) in case of parameter estimation, (iii) when we want to make short-term forecasts, or (iv) we want to gain information of the variance of state variables or events—such like in Sect. 4 where we estimated the number of infected individuals near the peak of the epidemic. Thus, reducing the scope of the simulation to shorter time intervals is usually a convenient way of dealing with speed.

Incorporating demography and aging into our models may change the process drastically in a long term and may lead to much more realistic models. However, in case of short term simulations it usually leads to unnecessary complications and a number of (aging) events that occupies the simulation engine and leads to slower simulations. Moreover in case of a small number of age groups the change between two age groups during a short term simulation remains negligible. However, if unavoidable, we suggest that instead of Gillespie's [3] exact stochastic algorithm consider an approximating algorithm when aging events occur in discrete time steps, say on every 30th day of the simulation, and the appropriate portion of the population in compartment X_j^i (for $j \in \{0, \ldots, l\}$) moves to compartment X_j^{i+1} if $i \neq n$, dies whenever $i = n$. In the meanwhile appropriate portion of newborns need to be added into the susceptible, immunised or the infected compartments (according to the model assumptions) of age group 1.

Acknowledgments This work was done in the framework of the Hungarian National Development, Research, and Innovation (NKFIH) Fund 2020-2.1.1-ED-2020-00003. Work partially supported by grant NKFIH-1279-2/2020 of the Ministry for Innovation and Technology, Hungary. P.B. and Z.V. were partially supported by grant Functional Differential Equations in Mathematical Epidemiology National Research, Development and Innovation Office NKFI FK 124016 and the final phase was supported from the NKFIH Fund, project no. TKP2021-NVA-09.

References

1. G. Röst, F. Bartha, N. Bogya, P. Boldog, A. Dénes, T. Ferenci, K. Horváth, A. Juhász, C. Nagy, T. Tekeli, Z. Vizi. and B. Oroszi, *Viruses*. **12** (2020)
2. M. V. Barbarossa, *PLOS ONE*. **15**, 1–22 (2020,9)
3. D. Gillespie, *Journal Of Computational Physics*. **22**, 403–434 (1976)
4. D. Gillespie, *The Journal Of Physical Chemistry*. **81** (1977)
5. L. Allen, *Infectious Disease Modelling*. **2**, 128–142 (2017)
6. L. Yang Cao, *The Journal Of Chemical Physics*. **121**, 4059–4067 (2004)
7. D. Gillespie, A. Hellander and L. R. Petzold, *The Journal of Chemical Physics* **138**, 170901 (2013)
8. M Gibson and J. Bruck, *The Journal Of Physical Chemistry A*. **104** (2000,3)
9. M. Schwehm, Poster in German Conference on Bioinformatics. (2001)
10. K. Prem, *PLOS Computational Biology*. **13**, 1–21 (2017,9)
11. Z. Vizi, *GitHub Repository*. (2021), https://github.com/zsvizi/propensity-matrix-epidemics
12. Hungarian Central Statistical Office, *website*. Accessed: 2021-09-20, www.ksh.hu/stadat_files/nep/en/nep0003.html
13. S. Indurkhya, J. Beal, *PLOS ONE*. **5(1)** (2010)

Modeling the Insect-Vectors-Mediated Phytoplasm Transmission in Agroecosystems

Francesca Acotto, Mattia Berera, Giulia Malano, and Ezio Venturino

1 Introduction

The aim of this paper is to formulate and analyze particular insects-plants interactions, with the former being often harmful for crops, [1]. Specifically, we consider a leafhopper vector that introduces pathogenic phytoplasm into a susceptible population of seasonal plants. The agroecosystem is assumed to be possibly infected by different phytoplasm strains. The latter differ only by the presence or the absence of a specific effector protein, whose effect renders infected plants more palatable to the insect vector, that are assumed not to have alternative food sources. Indeed, it is known that the insect herbivores behavior towards plants can be mediated by molecular interactions, [3, 9]. The model is conceived to forecast the effect of this protein on the pathogen spread among the plant and insect populations. We envisage a closed epidemiological system with no insect migrations and consider only a fixed strain of phytoplasm.

Models for this situation have already appeared in the literature [6]. Also, the presence of effector proteins in agriculture has been demonstrated [2, 4, 5, 8], and can be used in various situations [7].

Two different scenarios are presented. In the first one we assume that the farmer regularly checks the crop, so that to control the spread of the pathogen, infected plants are possibly removed and replaced by healthy ones. Thus the total seasonal plant population is fixed in time. In the second one, this last restriction is removed and the infection may propagate freely among the plants, via the action of insect vectors.

F. Acotto · M. Berera · G. Malano · E. Venturino (✉)
Dipartimento di Matematica "Giuseppe Peano", Università di Torino, Torino, Italy
e-mail: ezio.venturino@unito.it

R. P. Mondaini (ed.), *Trends in Biomathematics: Stability and Oscillations in Environmental, Social, and Biological Models*,
https://doi.org/10.1007/978-3-031-12515-7_18

In particular, we would like to elucidate the role played by the effector protein in the agroecosystem dynamics.

In the next section we introduce the system controlled by the farmer, and analyze it assessing its equilibria and their stability. Section 3 contains the second situation, with a nonconstant plant population. A final discussion on the role of the effector proteins concludes the paper.

2 The Model Incorporating the Farmer's Behavior

2.1 Model Assumptions

As mentioned, here the number of plants in the agroecosystem is fixed: each diseased plant that dies or is discovered by the farmer is removed and is replaced with a new susceptible healthy plant. This assumption also allows us to disregard possible vertical phytoplasm transmission in the plant population.

Demographics is instead considered for the insect vector population. This is motivated by the fact that during the good season, while plants grow and produce fruits and seeds for the next year, insects instead experience generally several reproduction times. It is not uncommon that during the spring and summer, three or more generations are found, with average insects lifetimes spanning around a few weeks. No farmer's control action on the insects diffusion and their damages is assumed to occur. Also, here we allow the pathogen to be possibly vertically transmitted. The assumption however could be easily removed by setting the corresponding parameter to zero, in case it is found not to hold for specific species.

2.2 Model Setup

Let S represent the susceptible plants population, I represent the one of infected plants, U stand for the susceptible insects and V for those infected. Further, let C denote the whole plants population in the agroecosystem,

$$C = S + I. \tag{1}$$

The model under consideration consists of two simple equations for the susceptible and infected plants and another similar two for the insects. They are connected by the fact that each infected subpopulation is recruited via susceptible interactions with infected of the other subpopulation.

Susceptible plants become indeed infected when bitten by an insect-vector able to transmit the phytoplasm. Let us define a as the preferential attraction probability of insects towards infected plants. Let also λ be the phytoplasm transmission rate.

Now the term $(1-a)\lambda VS$ represents the rate at which vectors are attracted by, bite and transmit the phytoplasm to susceptible plants and therefore infect them. Note that in this term, no role is exerted by the healthy insects, because at least in this context, they do not harm the plants. In reality, they suck sap, therefore they damage the plants, but do not transmit the pathogenic phytoplasm, and this is what is of interest here. Furthermore, we model the farmers' behavior, who removes at rate μ the infected plants and replaces them by new susceptible ones. The plants dynamics is therefore captured by the following equations:

$$\frac{dS}{dt} = \mu I - (1-a)\lambda VS, \quad \frac{dI}{dt} = (1-a)\lambda VS - \mu I. \tag{2}$$

Insects are subject to demographics, with birth rate b, natural mortality n and subject to intraspecific competition. However, since they are partitioned among carriers and susceptibles, we must distinguish between their relative influences. We denote by c_{XY}, $X, Y \in \{U, V\}$ the negative action exerted by the subpopulation Y on the subpopulation X. As for the phytoplasm vertical transmission, we allow it, denoting by p the fraction of insect offsprings generated by carrier parents of the previous generation. Thus these new insects bear the phytoplasm since egg disclosure. Adult insects phytoplasm transmission occurs via biting of an infected plant. In this case we must take into account both plant subpopulations, because if a healthy insect bites a susceptible plant, it does not become infected. Let the plant-insect phytoplasm transmission be denoted by σ. This must be a function of susceptible insects and infected plants, thus $\sigma = \sigma(I, U)$. Now the insect has the choice of the plant to bite, and thus the rate at which it finds an infected plant depends on the latter abundance in the whole plant population. The fraction IC^{-1} gives the probability of picking one infected plant among all plants. We also denote by β the phytoplasm acquisition rate. Further, because phytoplasm cannot be transmitted in the absence of plants, we set the transmission to zero if $C = 0$. Overall, the pathogenic phytoplasm transmission from plant to insect is modeled via the function

$$\sigma(I, U) = \begin{cases} \beta \dfrac{IU}{C}, & C \neq 0 \\ 0 & C = 0 \end{cases}. \tag{3}$$

To complete the insects infection process, we need to take into account also their preferential attraction toward infected plants, a. We also assume that phytoplasm carrying does not alter insect natural mortality. Combining all these features and (3), we thus obtain for $C \neq 0$

$$\frac{dU}{dt} = b(1-p)V + bU - a\beta \frac{IU}{C} - c_{UU}U^2 - c_{UV}UV - nU, \tag{4}$$

$$\frac{dV}{dt} = bpV + a\beta \frac{IU}{C} - c_{VV}V^2 - c_{VU}VU - nV. \tag{5}$$

Remark 1 Note that for $C = 0$, the Eq. (2) are trivial, plants vanish, $S = I = 0$, and only the subsystem (4) and (5) remains, that will be analysed in detail below.

The complete model is made by Eqs. (2), (4), and (5), where all the parameters are assumed to be nonnegative. Further restrictions, in view of what we have discussed above, are $a \in (0, 1)$, $p \in [0, 1]$, these parameters being probabilities. Note that for the former the particular cases $a = 0$ and $a = 1$ are excluded, because the protein cannot render the infected or healthy plants entirely appetizing or unappetizing for the insect. The former conditions are supplemented by the following assumptions: healthy insects are more competitive than sick ones, thus some coefficients c_{XY} are comparable; namely,

$$c_{UV} \leq c_{VU}, \quad c_{UV} \leq c_{UU}, \quad c_{VV} \leq c_{VU}. \tag{6}$$

Also, the insect birth rate exceeds their mortality rate, i.e.

$$b \geq n; \tag{7}$$

indeed, if this were not the case, by adding the Eqs. (4) and (5) we find that the whole insect population would vanish. In fact its dynamics would become

$$\frac{d(U+V)}{dt} = b(V+U) - n(V+U) - c_{UU}U^2 - c_{UV}UV - c_{VV}V^2 - c_{VU}VU < 0.$$

A complete list of the parameters is given in Table 1.

2.3 Equilibria Analysis

2.3.1 The Insect-Only Subsystem

For $C = 0$, from (4) and (5) we obtain the equilibria $P_0 = (0, 0)$, $P_1 = (U_1, 0)$, $P_* = (U_*, V_*)$, with

$$U_1 = \frac{b-n}{c_{UU}}; \qquad V_* = \frac{1}{c_{VV}}(bp - c_{VU}U_* - n),$$

while U_* is a positive root of the quadratic $\sum_{k=0}^{2} a_k U^k = 0$ with

$$a_2 = \frac{c_{UV}c_{VU}}{c_{VV}} - c_{UU}, \quad a_0 = \frac{b(1-p)}{c_{VV}}(bp - n),$$
$$a_1 = b - n - \frac{c_{UV}}{c_{VV}}(bp - n) - \frac{c_{VU}}{c_{VV}}b(1-p),$$

Table 1 Description, dimensions and basic assumptions about the parameters, assumed all to be positive

	Description	Dimension	Assumptions
a	Insects preferential attraction probability toward infected plants	–	$a \in (0, 1)$
b	Insects natural birth rate	$\frac{1}{[t]}$	$b \geq n$
c_{XY}	Intraspecific competition rate of class X individuals over class Y individuals, $[X, Y \in \{U, V\}]$	$\frac{1}{[t]}$	$c_{UV} < c_{VU}, c_{UV} < c_{UU}, c_{VV} < c_{VU}$
C	Number of plants, fixed	–	$C = S + I > 0$
n	Insects natural mortality rate	$\frac{1}{[t]}$	
p	Probability of insects vertical phytoplasm transmission	–	$p \in [0, 1]$
β	Phytoplasm acquisition rate for insects from plants	$\frac{1}{[t]}$	
λ	Phytoplasm transmission rate from insects to plants	$\frac{1}{[t]}$	
μ	Infected plants mortality rate and replanting rate of susceptible plants	$\frac{1}{[t]}$	

which also must satisfy, to ensure nonnegativity of V_*,

$$U_* \leq \frac{bp - n}{c_{VU}}. \tag{8}$$

To be fulfilled, this implies also

$$bp \geq n \tag{9}$$

which in turn gives $a_0 \geq 0$. Thus by Descartes rule, a positive root U_* is obtained by requiring also $a_2 > 0$ or $a_1 > 0$. In summary, feasibility of P_* is ensured by (9) and at least one of the following conditions

$$bc_{VV} \leq nc_{VV} + c_{UV}(bp - n) + c_{VU}b(1 - p), \quad c_{UV}c_{VU} \leq c_{VV}c_{UU}. \tag{10}$$

Feasibility for P_1 holds unconditionally, being ensured by (7).

The Jacobian of (4) and (5) is

$$\widehat{\boldsymbol{J}} = \begin{bmatrix} b - n - 2c_{UU}U - c_{UV}V & b(1 - p) - c_{UV}U \\ -c_{VU}V & bp - 2c_{VV}V - c_{VU}U - n \end{bmatrix}.$$

At the origin its eigenvalues would give the stability condition $b < n$, which cannot hold in view of (7), so that it is always unstable.

For P_1 we find instead $n < b$, ensured by (7), and

$$(bp - n)c_{UU} < (b - n)c_{VU}. \tag{11}$$

At P_*, we apply the Routh-Hurwitz conditions. The trace condition is always satisfied,

$$c_{VV}U_*V_* + c_{UU}U_*^2 + b(1 - p)V_* > 0.$$

Stability is thus ruled only by the sign of the determinant; stable coexistence of healthy and infected insects occurs for

$$V_* > \frac{U_*}{b(1 - p)c_{VV}}[U_*(c_{VU}c_{UV} - c_{UU}c_{VV}) - c_{VU}b(1 - p)]. \tag{12}$$

2.3.2 The Complete Model

In view of the parameter assumptions, specifically here $\lambda, \mu > 0$ and $a \neq 1$, from (2) note that $I = 0$ implies $S = 0$ or $V = 0$. Thus the search for the system's equilibria is somewhat eased. We find the points

$$\widehat{E}_1 = (C, 0, 0, 0), \quad \widehat{E}_2 = \left(C, 0, \frac{b - n}{c_{UU}}, 0\right), \quad \widehat{E}_3 = \left(C - \widehat{I}_3, \widehat{I}_3, 0, \frac{b - n}{c_{VV}}\right),$$

where the latter exists only in the very special case of full vertical phytoplasm transmission:

$$p = 1, \tag{13}$$

and where

$$\widehat{I}_3 = \frac{(1 - a)\lambda(b - n)C}{c_{VV}\mu + (1 - a)\lambda(b - n)}.$$

The equilibrium $\widehat{E}_1$ corresponds to the insect-and-disease-free equilibrium, with the whole plantation being healthy and insects-free, while $\widehat{E}_2$ instead represents the disease-free point with both plants and insects thriving. Note also that equilibria with $S = 0$, while the other populations do not vanish, cannot occur in this situation, because the new susceptible plants are always introduced in the field by the farmer, to replace the diseased ones that he removes.

We now investigate the coexistence equilibrium. Combining the constraint equation (1) with the equation of the infected plants, we find

$$I = \frac{(1-a)\lambda C V}{\mu + (1-a)\lambda V}.$$

By substitution into the equilibrium equations obtained from (4) and (5), we find the following non-linear system:

$$c_{UU}U^2 + \frac{a(1-a)\beta\lambda}{\mu + (1-a)\lambda V}UV + c_{UV}UV + (n-b)U + b(p-1)V = 0,$$
$$c_{VV}V^2 - \frac{a(1-a)\beta\lambda}{\mu + (1-a)\lambda V}VU + c_{VU}VU + (n-bp)V = 0.$$

It turns out that this system is of order three, which makes an analytical study essentially impossible.

Alternatively, the coexistence point $\widehat{E}_4 = (C - \widehat{I}_4, \widehat{I}_4, \widehat{U}_4, \widehat{V}^*)$ has the population values

$$\widehat{I}_4 = \frac{(1-a)\lambda C\widehat{V}^*}{\mu + (1-a)\lambda\widehat{V}^*},$$
$$\widehat{U}_4 = \frac{(bp-n)\mu - (1-a)c_{VV}\lambda\widehat{V}^{*2} - [c_{VV}\mu + (1-a)(n-bp)\lambda]\widehat{V}^*}{(1-a)c_{VU}\lambda\widehat{V}^* + c_{VU}\mu + a(1-a)\beta\lambda},$$

with feasibility condition obtained by imposing the numerator of $\widehat{U}_4$ to be positive, namely

$$(bp-n)\mu \geq (1-a)c_{VV}\lambda\widehat{V}^{*2} + [c_{VV}\mu + (1-a)(n-bp)\lambda]\widehat{V}^* \qquad (14)$$

and where $\widehat{V}^*$ is a positive root of the fifth order algebraic equation

$$\widehat{a}_5V^5 + \widehat{a}_4V^4 + \widehat{a}_3V^3 + \widehat{a}_2V^2 + \widehat{a}_1V + \widehat{a}_0 = 0$$

whose coefficients depend on the model parameters. Note also that the condition $\widehat{I}_4 < C$ is easily seen to be always satisfied.

Both these analytic approaches essentially fail and we need to investigate this point numerically.

2.4 Local Stability Analysis

The Jacobian matrix associated with the model (2), (4), (5) is

$$\widehat{\boldsymbol{J}} = \begin{bmatrix} -(1-a)\lambda V & \mu & 0 & -(1-a)\lambda S \\ (1-a)\lambda V & -\mu & 0 & (1-a)\lambda S \\ 0 & -a\beta\dfrac{U}{C} & \widehat{\boldsymbol{J}}_{3,3} & \widehat{\boldsymbol{J}}_{3,4} \\ 0 & a\beta\dfrac{U}{C} & \widehat{\boldsymbol{J}}_{4,3} & \widehat{\boldsymbol{J}}_{4,4} \end{bmatrix},$$

with

$$\begin{aligned} \widehat{\boldsymbol{J}}_{3,3} &= b - a\beta\frac{I}{C} - 2c_{UU}U - c_{UV}V - n\,, \\ \widehat{\boldsymbol{J}}_{3,4} &= b(1-p) - c_{UV}U\,, \\ \widehat{\boldsymbol{J}}_{4,3} &= a\beta\frac{I}{C} - c_{VU}V\,, \\ \widehat{\boldsymbol{J}}_{4,4} &= bp - 2c_{VV}V - c_{VU}U - n\,. \end{aligned}$$

Because the first two rows of $\widehat{\boldsymbol{J}}$ are linearly dependent, at least one of the eigenvalues of $\widehat{\boldsymbol{J}}$ vanishes. Thus no equilibrium is asymptotically stable, but at most it can only be stable.

At $\widehat{E}_1$ the nonvanishing eigenvalues are $-\mu$, $b-n$, $bp-n$, but in view of (7) the second one is always positive, implying that this equilibrium is unconditionally unstable, unless $b=n$ in which case it is stable but not asymptotically.

At $\widehat{E}_2$ one eigenvalue is zero and another one is immediate, $n-b<0$. We then apply the Routh-Hurwitz criterion for the remaining minor of order two to respectively obtain, for the trace and the determinant,

$$\mu - \left[\left(p - \frac{c_{VU}}{c_{UU}}\right)b - \left(1 - \frac{c_{VU}}{c_{UU}}\right)n\right] > 0; \tag{15}$$

$$\mu\left[\left(p - \frac{c_{VU}}{c_{UU}}\right)b - \left(1 - \frac{c_{VU}}{c_{UU}}\right)n\right] + (1-a)a\beta\lambda\frac{b-n}{c_{UU}} < 0. \tag{16}$$

Now (15) follows if the bracket is negative, which is ensured by

$$(pc_{UU} - c_{VU})b < (c_{UU} - c_{VU})n. \tag{17}$$

Table 2 Feasibility and local stability conditions for the system (2), (4), (5) equilibria

Equilibria	Feasibility	Stability
$\widehat{E}_1 = (C, 0, 0, 0)$	—	For $b > n$ unstable; for $b = n$ stable but not asymptotically stable
$\widehat{E}_2 = (C, 0, \frac{b-n}{c_{UU}}, 0)$	$b > n$	(15), (16); sufficient: (18)
$\widehat{E}_3 = (C - \widehat{I}_3, \widehat{I}_3, 0, \frac{b-n}{c_{VV}})$	$p = 1, b > n$	(19) stable but not asymptotically stable
$\widehat{E}_4 = (C - \widehat{I}_4, \widehat{I}_4, \widehat{U}_4, \widehat{V}^*)$	$\widehat{V}^* > 0, U_4(\widehat{V}^*) > 0$	Numerical

Similarly, (16) can be ensured if in addition to (17) we require also

$$(pc_{UU} - c_{VU})b < (c_{UU} - c_{VU})n - (1-a)a\beta\lambda\frac{b-n}{\mu}. \tag{18}$$

At $\widehat{E}_3$ all the eigenvalues of the Jacobian are explicitly known, in view of (13); two of them are always negative, because of (7)

$$n - b < 0, \quad -\left[(1-a)\lambda\frac{b-n}{c_{VV}}\right] < 0.$$

Thus the remaining nonzero one alone determines the stability of this equilibrium:

$$\frac{1}{c_{VV}}\left(\frac{(1-a)a\beta\lambda}{\mu + (1-a)\lambda\frac{b-n}{c_{VV}}} + c_{UV}\right) < 1. \tag{19}$$

Coexistence is investigated numerically.

Table 2 summarizes the findings.

From Table 2 a transcritical bifurcation is observed for which $\widehat{E}_2$ emanates from $\widehat{E}_1$ as soon as b increases past n. In Figs. 1 and 2 a transcritical bifurcation relating equilibria $\widehat{E}_4$ and $\widehat{E}_2$ is shown to occur in terms of the bifurcation parameters β and λ respectively. The other parameter values for these figures are

$$\begin{aligned} a = 0.5, \quad b = 9, \quad c_{UU} = 0.15, \quad c_{UV} = 0.10, \\ c_{VU} = 0.25, \quad c_{VV} = 0.20, \quad n = 6, \quad \mu = 7. \end{aligned} \tag{20}$$

Also, the chosen initial conditions are

$$S_0 = 300, \quad I_0 = 0, \quad U_0 = 150, \quad V_0 = 1. \tag{21}$$

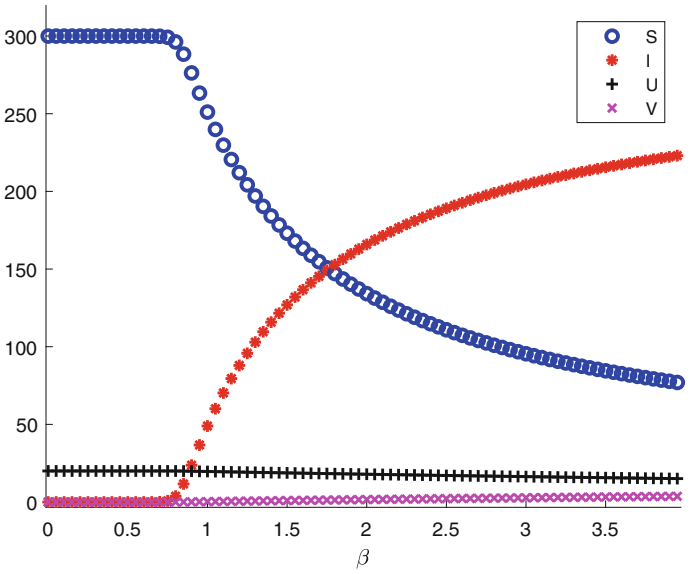


Fig. 1 Transcritical bifurcation from $\widehat{E}_4$ to $\widehat{E}_2$ as the bifurcation parameter β decreases below the threshold $\widehat{\beta} \approx 0.75$. The other parameter values are given in (20), with additionally $p = 0.5$ and $\lambda = 11$. Initial conditions given by Eq. (21)

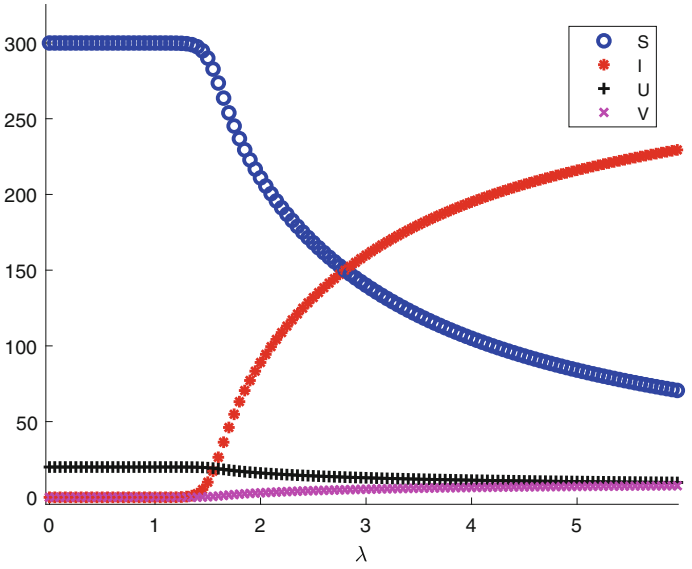


Fig. 2 Transcritical bifurcation from $\widehat{E}_4$ to $\widehat{E}_2$ as the bifurcation parameter λ decreases below the threshold $\widehat{\lambda} \approx 1.25$. The other parameter values are given in (20), with additionally $p = 1$ and $\beta = 2$. Initial conditions given by Eq. (21)

3 The Model with Variable Plant Population

In this second formulation, we account for plants intra-species competition, birth rate and natural mortality. Using the same variables already defined in Eqs. (2), (4), and (5) but introducing some new parameters in addition to the ones previously used, the current model becomes:

$$\frac{dS}{dt} = rS + rI - (1-a)\lambda VS - c_{SI}SI - c_{SS}S^2 - mS \tag{22}$$

$$\frac{dI}{dt} = (1-a)\lambda VS - c_{IS}IS - c_{II}I^2 - (\mu + m)I \tag{23}$$

$$\frac{dU}{dt} = b(1-p)V + bU - a\beta\frac{IU}{S+I} - c_{UU}U^2 - c_{UV}UV - nU \tag{24}$$

$$\frac{dV}{dt} = bpV + a\beta\frac{IU}{S+I} - c_{VV}V^2 - c_{VU}VU - nV. \tag{25}$$

In the Eq. (22), r represents the plants reproduction rate, and m their natural mortality. Note that the following assumption is necessary,

$$r > m, \tag{26}$$

because otherwise, summing Eqs. (22) and (23) we would obtain

$$\frac{d(S+I)}{dt} = (r-m)(S+I) - c_{SI}SI - c_{SS}S^2 - c_{IS}IS - c_{II}I^2 - \mu I < 0$$

and the whole plant population would vanish, making the model useless.

In addition, c_{SI} represents the intraspecific pressure exerted by infected over susceptible plants, while c_{SS} is the corresponding pressure due to healthy plants on other healthy ones. In Eq. (23), corresponding terms are c_{IS} denoting the pressure of susceptibles over infected plants, and c_{II} the intraspecific pressure of infected over other infected plants. The remaining Eqs. (24) and (25) for insect vectors are the same as in the previous model, namely Eqs. (4) and (5), with the only change due to the fact that now in the standard incidence term, the total plant population C is no longer constant, but must be replaced by $S + I$.

Assumptions similar to (6) on the new coefficients are made, namely

$$c_{SI} < c_{IS}, \quad c_{SI} < c_{SS}, \quad c_{II} < c_{IS}. \tag{27}$$

The newly introduced parameters, their meaning, units of measurement and assumptions are also reported in Table 3.

Table 3 Additional parameters in model (22)–(25) and related assumptions

	Description	Dimension	Assumptions
c_{XY}	Intraspecific competition coefficient of Y over X, with $X, Y \in \{S, I\}$	$\frac{1}{[t]}$	$c_{XY} > 0 \quad c_{SI} < c_{IS} \quad c_{SI} < c_{SS} \quad c_{II} < c_{IS}$
m	Natural plants death rate	$\frac{1}{[t]}$	$m > 0$
r	Natural plants birth rate	$\frac{1}{[t]}$	$r > m$

3.1 Equilibria Feasibility

Because Eqs. (24) and (25) are the same as for (4) and (5), apart from the replacement of C with $S + I$, the plant free case $S = I = 0$ has been already discussed in Sect. 2.3.1 and therefore is no longer considered in what follows.

The points that are possible equilibria for the system (22)–(25) are

$$E_1 = \left(\frac{r-m}{c_{SS}}, 0, 0, 0\right), \quad E_2 = \left(\frac{r-m}{c_{SS}}, 0, \frac{b-n}{c_{UU}}, 0\right)$$

the former being unconditionally feasible in view of (26), the latter also, using both (26) and (7). In addition, the particular healthy-insects-free case $E_3 = (S_3, I_3, 0, V_3)$ can be obtained, by assuming perfect vertical transmission

$$p = 1. \tag{28}$$

Easily from (24) and (25), $U_3 = 0$, $V_3 = V_2 = (b-n)c_{VV}^{-1}$. The plants populations can be obtained from (22)–(23). From the latter,

$$S = \Phi(I) = \frac{c_{II}I + \mu + m}{c_{IS}(I^\infty - I)}I, \quad I^\infty = \frac{(1-a)\lambda(b-n)}{c_{IS}c_{VV}}. \tag{29}$$

Thus $\Phi(I)$ has a negative zero and crosses the origin raising up to infinity at I^∞, this representing its only feasible branch, while for $I > I^\infty$ it is negative. Further, its derivative at the origin is positive,

$$\Phi'(0) = c_{VV}\frac{\mu + m}{(1-a)\lambda(b-n)} > 0. \tag{30}$$

Also, by adding the equilibrium equations corresponding to (22) and (23), we obtain a conic section

$$\Psi(I, S) = c_{II}I^2 + c_{SS}S^2 + (c_{IS} + c_{SI})SI - (r - m - \mu)I - (r - m)S = 0. \tag{31}$$

Discarding the degenerate cases, from its invariants, we observe that Ψ is an ellipse in case $4c_{II}c_{SS} > c_{IS} + c_{SI}$ and a hyperbola conversely. This conic goes through the origin and crosses the axes at the points $Q^I = (I^0, 0)$ and $Q^S = (0, S^0)$ with:

$$S^0 = \frac{r - m}{c_{SS}}, \quad I^0 = \frac{r - m - \mu}{c_{II}}.$$

Now, $S_0 > 0$ in view of (26), while the sign of I^0 is not determined. Note that if we try to assess the sign of the slope at the points at which the conic intersects the axes, by implicit differentiation, we find

$$\frac{dS}{dI} = \frac{2c_{II}I + (c_{IS} + c_{SI})S - (r - m - \mu)}{r - m - 2c_{SS}S - (c_{IS} + c_{SI})I},$$

so that

$$\frac{dS}{dI}|_{Q^I} = \frac{c_{II}I^0}{r - m - (c_{IS} + c_{SI})I^0}, \quad \frac{dS}{dI}|_{Q^S} = 1 - \frac{\mu + (c_{IS} + c_{SI})S^0}{r - m} < 1,$$

while at the origin we find

$$\Psi'(O) = \frac{dS}{dI}|_O = \frac{\mu}{r - m} - 1 > -1.$$

In spite of these calculations, the signs of the derivatives, although still useful to discard some possible situations arising below, are not enough to discriminate between the two possible types of conic sections. Therefore we must examine them both. Further implicit differentiation produces, after evaluation at the origin,

$$\frac{d^2S}{dI^2}|_O = -\frac{2}{2c_{SS} + r - m}\left[(c_{SS} + c_{IS} + c_{SI})\frac{dS}{dI}|_O + c_{II}\right].$$

Remark 2 From this expression, if $S'(0) > 0$ it follows $S''(0) < 0$. Thus at the origin, a positive slope must be coupled with a negative curvature.

Case A : Suppose Ψ is an ellipse: there are two cases:

(a) $r > m + \mu$ for which $I^0 > 0$; the feasible part of Ψ consists just of an arc in the first quadrant joining the points Q^I and Q^S;
(b) $r < m + \mu$ for which $I^0 < 0$; the feasible part of Ψ consists just of an arc joining the origin and the point Q^I.

Case B : Suppose Ψ is a hyperbola: there are six cases:

(1) $I^0 > 0$, $\Psi'(O) < 0$ and $\Psi''(O) < 0$: the feasible part of Ψ consists just of a convex arc in the first quadrant joining the points Q^I and Q^S;

(2) $I^0 > 0$, $\Psi'(O) > 0$ and $\Psi''(O) < 0$: the feasible part of Ψ consists just of an arc joining the origin and the point Q^I and a convex arc emanating from Q^S;
(3) $I^0 < 0$, $\Psi'(O) > 0$ and $\Psi''(O) < 0$: the feasible part of Ψ consists just of a concave arc emanating from the origin and a convex arc emanating from Q^S;
(4) $I^0 < 0$, $\Psi'(O) < 0$ and $\Psi''(O) > 0$: the feasible part of Ψ consists just of a concave arc emanating from Q^S;
(5) $I^0 < 0$, $\Psi'(O) < 0$ and $\Psi''(O) < 0$: the feasible part of Ψ consists just of a convex arc emanating from Q^S;
(6) finally, $I^0 > 0$, $\Psi'(O) > 0$ and $\Psi''(O) > 0$: the feasible part of Ψ consists of an arc joining the origin and the point Q^S and a concave arc emanating from Q^I; however, in view of Remark 2 this case cannot arise.

The susceptible S_3 and infected I_3 populations at equilibrium E_3 are obtained by the intersections of Ψ and Φ lying in the first quadrant. Examining the various situations, we are led to the following conclusions:

Case A : Ψ is an ellipse:

(a) the feasible intersection with Φ is always guaranteed;
(b) the feasible intersection with Φ is guaranteed if and only if $\Phi'(0) > \Psi'(O)$.

Case B : Ψ is a hyperbola:

(1) here $\Psi'(O) < 0 < \Phi'(0)$; exactly one intersection of $\Phi(I)$ and Ψ is guaranteed;
(2) (i) if $\Phi'(0) < \Psi'(O)$, given that Φ has the vertical asymptote, two intersections between Φ and the two feasible branches of Ψ are always guaranteed;
(ii) if $\Phi'(0) > \Psi'(O)$, given that Φ has the vertical asymptote, exactly one intersection of $\Phi(I)$ with the branch of Ψ emanating from Q^S exists;
(3) (i) if $\Phi'(0) < \Psi'(O)$, given that Φ has the vertical asymptote, two intersections between Φ and the two feasible branches of Ψ are always guaranteed;
(ii) if $\Phi'(0) > \Psi'(O)$, given that Φ has the vertical asymptote, exactly one intersection of $\Phi(I)$ with the branch of Ψ emanating from Q^S exists;
(4) here $\Psi'(O) < 0 < \Phi'(0)$; in this case there is always one feasible intersection between Φ and Ψ;
(5) here $\Psi'(O) < 0 < \Phi'(0)$; in this case there is always one feasible intersection between Φ and Ψ.

The coexistence equilibrium $E^* = (S^*, I^*, U^*, V^*)$ is investigated numerically.

3.2 Equilibria Stability

The Jacobian matrix $\boldsymbol{J}$ associated with the model (22)–(25) is the following:

$$\begin{bmatrix} \boldsymbol{J}_{1,1} & r - c_{SI}S & 0 & (a-1)\lambda S \\ -c_{IS}I - (a-1)\lambda V & \boldsymbol{J}_{2,2} & 0 & -(a-1)\lambda S \\ a\beta\frac{IU}{(S+I)^2} & a\beta\frac{IU}{(S+I)^2} - a\beta\frac{U}{S+I} & \boldsymbol{J}_{3,3} & \boldsymbol{J}_{3,4} \\ -a\beta\frac{IU}{(S+I)^2} & a\beta\frac{U}{S+I} - a\beta\frac{IU}{(S+I)^2} & \boldsymbol{J}_{4,3} & \boldsymbol{J}_{4,4} \end{bmatrix},$$

with

$$\begin{aligned} \boldsymbol{J}_{1,1} &= r - m - c_{SI}I - 2c_{SS}S + \lambda(a-1)V\,, \\ \boldsymbol{J}_{2,2} &= -m - \mu - 2c_{II}I - c_{IS}S\,, \\ \boldsymbol{J}_{3,3} &= b - n - 2c_{UU}U - c_{UV}V - a\beta\frac{I}{S+I}\,, \\ \boldsymbol{J}_{3,4} &= -c_{UV}U - b(p-1), \\ \boldsymbol{J}_{4,3} &= a\beta\frac{I}{S+I} - c_{VU}V, \\ \boldsymbol{J}_{4,4} &= bp - c_{VU}U - 2c_{VV}V - n. \end{aligned}$$

For the equilibrium E_1, we find the eigenvalues

$$m - r < 0, \quad \frac{c_{IS}(m-r)}{c_{SS}} - \mu - m < 0, \quad b - n \geq 0, \quad bp - n$$

so that it is unstable, unless $b = n$, in which case we obtain stability but not asymptotic stability.

Two eigenvalues of the Jacobian evaluated at E_2 are explicitly found, $m - r < 0$ and $n - b < 0$ both negative in view of (26) and (7). Using the Routh-Hurwitz criterion on the remaining minor of order two, we are led to the stability conditions

$$(bp - n) - (\mu + m) - \frac{c_{IS}}{c_{SS}}(r - m) - \frac{c_{VU}}{c_{UU}}(b - n) < 0, \tag{32}$$

$$\left(\frac{c_{IS}}{c_{SS}}(r - m) + \mu + m\right)\left(bp - n - \frac{c_{VU}}{c_{UU}}(b - n)\right) < \frac{\lambda\beta a(a-1)(b-n)}{c_{UU}}. \tag{33}$$

Further, estimating the second term on the left of condition (33) as follows

$$bp - n - \frac{c_{VU}}{c_{UU}}(b-n) = b\left[p - \frac{n}{b} - \frac{c_{VU}}{c_{UU}}(1-\frac{n}{b})\right]$$
$$= b\left(p - \frac{n}{b}\right)\left[1 - \frac{c_{VU}}{c_{UU}}\frac{1-\frac{n}{b}}{p-\frac{n}{b}}\right] > b\left(p - \frac{n}{b}\right)\left[1 - \frac{c_{VU}}{c_{UU}}\right]$$

the condition (33) can be rewritten as

$$bc_{UU}\left(1 - \frac{c_{VU}}{c_{UU}}\right)\left(p - \frac{n}{b}\right) < -\frac{\lambda\beta a(1-a)(b-n)c_{SS}}{c_{IS}(r-m) + c_{SS}(\mu+m)}.$$

Remark 3 Now (32) holds if

$$p < \frac{n}{b} < 1 \tag{34}$$

by (7) and (26). Conversely, using

$$bp - n - \frac{c_{VU}}{c_{UU}}(b-n) < (b-n)\left(1 - \frac{c_{VU}}{c_{UU}}\right) < 0,$$

it also holds if

$$\frac{n}{b} < p < 1, \quad c_{VU} > c_{UU}. \tag{35}$$

Remark 4 This last condition is certainly not met if either one of the following sets of inequalities hold:

$$\frac{n}{b} < p, \quad c_{VU} < c_{UU}; \qquad p < \frac{n}{b} < 1, \quad c_{VU} > c_{UU}. \tag{36}$$

Therefore, in such case, E_2 is certainly unstable.

At E_3 the characteristic equation factorizes into the product of two quadratic equations, stemming each from a suitable minor of order 2. We can apply once more the Routh-Hurwitz conditions to these submatrices, to find two sets of stability conditions, namely

$$\boldsymbol{J}(E_3)_{1,1} + \boldsymbol{J}(E_3)_{2,2} < 0, \tag{37}$$
$$\boldsymbol{J}(E_3)_{1,1}\boldsymbol{J}(E_3)_{2,2} + \left(c_{IS}I_3 + \frac{\lambda(b-n)(a-1)}{c_{VV}}\right)(r - c_{SI}S_3) > 0,$$

where

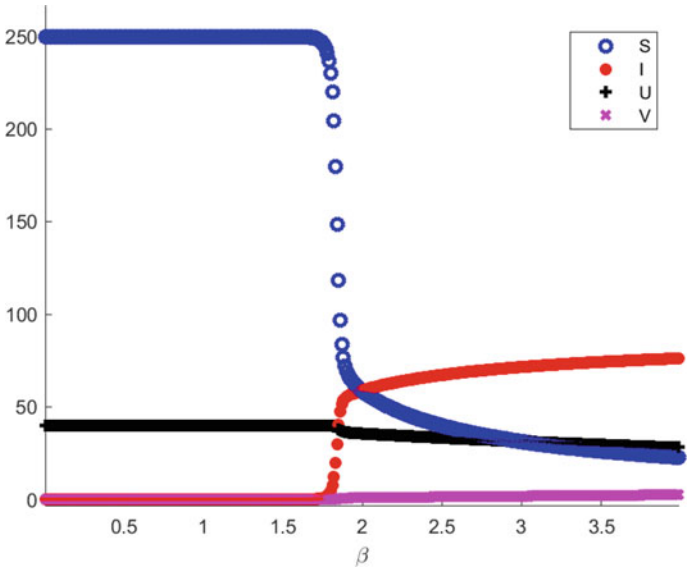


Fig. 3 Transcritical bifurcation from E_4 to E_2 as the bifurcation parameter β decreases below the threshold $\widehat{\beta} \approx 1.75$. The other parameter values are given by (39) and $p = 0.5$ and $\lambda = 12$. Initial conditions given below by Eq. (21)

$$\boldsymbol{J}(E_3)_{1,1} = r - m - c_{SI} I_3 - 2c_{SS} S_3 + \frac{\lambda(b-n)(a-1)}{c_{VV}},$$
$$\boldsymbol{J}(E_3)_{2,2} = -m - \mu - 2c_{II} I_3 - c_{IS} S_3$$

and

$$\boldsymbol{J}(E_3)_{3,3} + n - 2b + bp < 0, \quad \boldsymbol{J}(E_3)_{3,3}(n - 2b + bp) > \boldsymbol{J}(E_3)_{4,3} b(1-p) \tag{38}$$

with

$$\boldsymbol{J}(E_3)_{3,3} = b - n - \frac{c_{UV}(b-n)}{c_{VV}} - \frac{a\beta I^*}{S_3 + I^*}, \quad \boldsymbol{J}(E_3)_{4,3} = \frac{a\beta I^*}{S_3 + I^*} - \frac{c_{VU}(b-n)}{c_{VV}}.$$

Coexistence is shown numerically to be stable, for suitable parameter choices.

We also show transcritical bifurcations arising for the pair of equilibria E_4 to E_2 as the bifurcation parameters β (Fig. 3) and λ (Fig. 4) vary. The remaining parameter values used in these figures are

$$\begin{aligned} &a = 0.5, \quad b = 19, \quad c_{UU} = 0.1, \quad c_{UV} = 0.2, \quad c_{VU} = 0.4, \\ &c_{VV} = 0.3, \quad c_{SS} = 0.014, \quad c_{SI} = 0.023, \quad c_{IS} = 0.042, \\ &c_{II} = 0.035, \quad n = 15, \quad m = 0.5, \quad r = 4, \quad \mu = 0.3. \end{aligned} \tag{39}$$

Table 4 summarizes these findings.

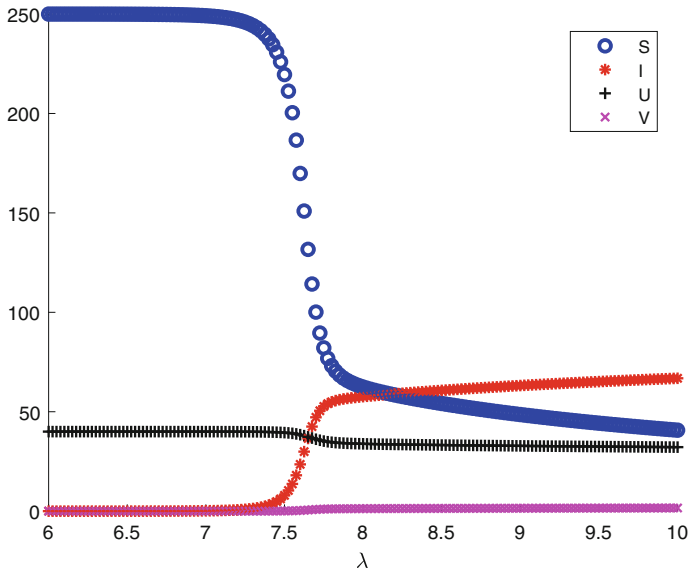


Fig. 4 Transcritical bifurcation from E_4 to E_2 as the bifurcation parameter λ decreases below the threshold $\widehat{\lambda} \approx 7.25$. The other parameter values are given by (39) and $p = 1$ and $\beta = 3$. Initial conditions given below by Eq. (21)

Table 4 Feasibility and local stability conditions for the system (22)–(25) equilibria, where the assumptions on the parameters (7) and (26) are implicitly used

Equilibria	Feasibility	Stability
$E_1 = (\frac{r-m}{c_{SS}}, 0, 0, 0)$	–	Unstable if $b > n$ simply stable if $b = n$
$E_2 = (\frac{r-m}{c_{SS}}, 0, \frac{b-n}{c_{UU}}, 0)$	–	(32), (33); sufficient (34), (35)
$E_3 = (S_3, I_3, 0, \frac{b-n}{c_{VV}})$	(28), see text: Cases A and B for $I_3 > 0$ and $S_3 > 0$	Numerical
$E_4 = (S^*, I^*, U^*, V^*)$	Numerical	Numerical

4 The Effector Protein Action on the Agroecosystem

For the model (2), (4), (5), we use the following reference parameter values:

$$b = 9, \quad c_{UU} = 0.15, \quad c_{UV} = 0.10, \quad c_{VU} = 0.25, \quad c_{VV} = 0.20, \quad (40)$$
$$n = 9, \quad p = 0.5, \quad \beta = 12, \quad C = 300, \quad \lambda = 11, \quad \mu = 7.$$

For model (22)–(25) the same initial conditions (21) are used, but the reference parameter values are

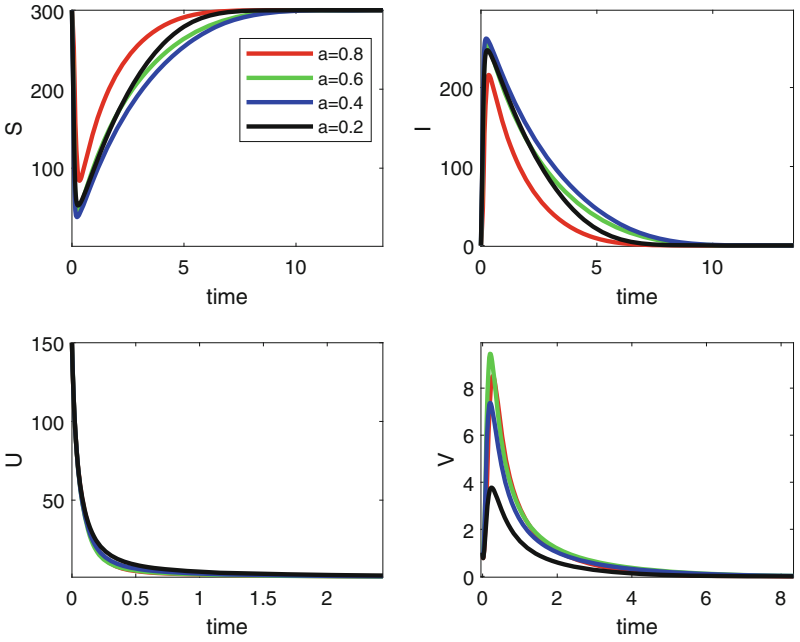


Fig. 5 Model (2), (4), (5): Equilibrium $\widehat{E}_1$ obtained with parameter values (40) and initial conditions (21)

$$b = 19, \quad c_{UU} = 0.1, \quad c_{UV} = 0.2, \quad c_{VU} = 0.4, \quad c_{VV} = 0.3, \tag{41}$$
$$c_{SS} = 0.014, \quad c_{SI} = 0.023, \quad c_{IS} = 0.042, \quad c_{II} = 0.035, \quad n = 19,$$
$$m = 0.5, \quad p = 0.5, \quad r = 4, \quad \beta = 11, \quad \lambda = 12, \quad \mu = 0.3.$$

Figures 5, 6, 7, 8, 9, 10 contain the simulations for model (2), (4), (5). In particular, Fig. 5 shows the solutions behavior in the particular case of $b = n$, while in all the other ones $b > n$. Note that all the possible system's equilibria are shown to arise, for suitable parameter choices. Clearly, the parameter a does not influence the final outcome, in the sense that changing it does not alter the point that is ultimately achieved. However, it does affect the speed at which the equilibrium is reached. Although in the transient phase there might be some slight differences, compare the frame for V in Fig. 6, in general a higher value of a helps in removing faster the infected and boosts the susceptible plants growth; these remarks hold in a lesser way for insects. An exception for the infected insects is however given by the coexistence equilibrium $\widehat{E}_4$ of Fig. 9, their number being increased by a larger value of the effector protein.

Completely similar remarks hold for the simulations involving model (22)–(25), Figs. 11, 12, 13, 14, 15, and 16. Again, Fig. 11 shows the particular case $b = n = 19$, while the remaining ones assume $b > n$.

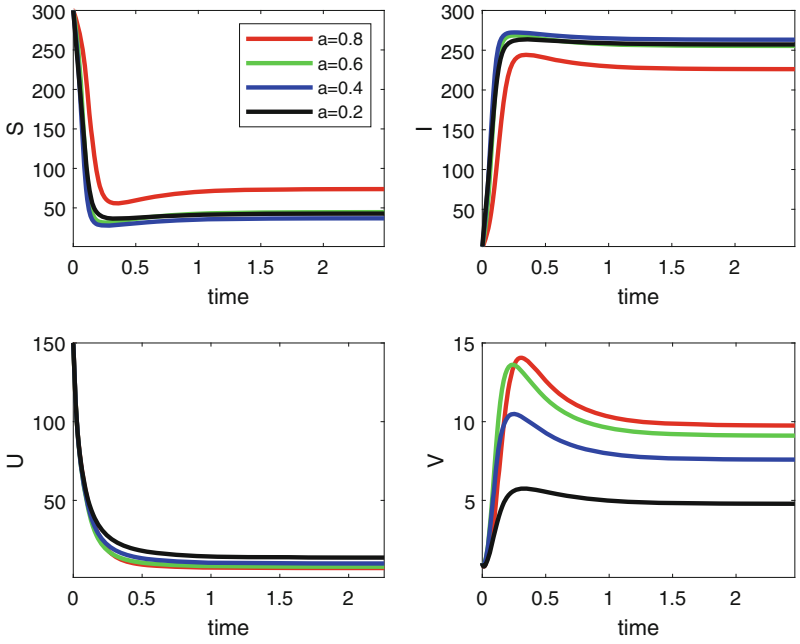


Fig. 6 Model (2), (4), (5): Equilibrium $\widehat{E}_4$ obtained with parameter values (40), but for $n = 6 < b$, and initial conditions (21)

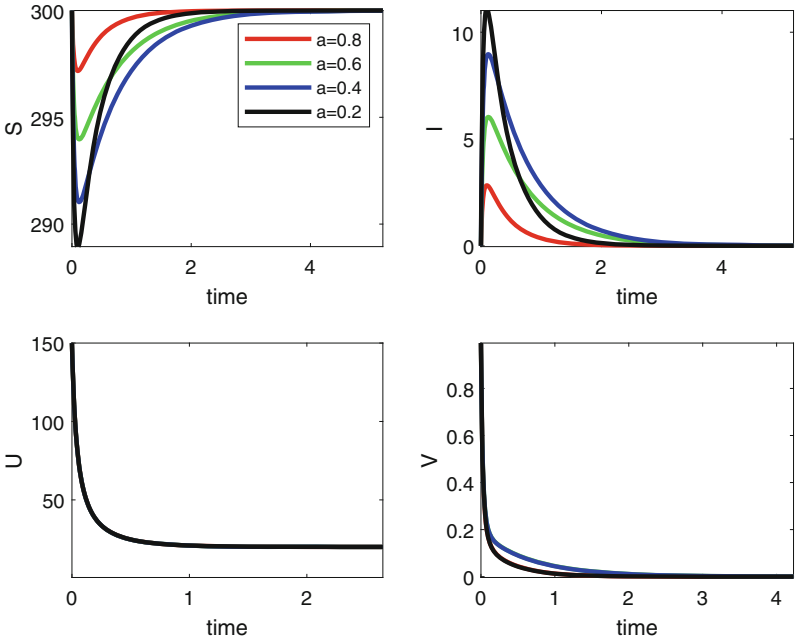


Fig. 7 Model (2), (4), (5): Equilibrium $\widehat{E}_2$ obtained with parameter values (40), but for $n = 6 < b$, $\beta = 3$, $\lambda = 2$ and initial conditions (21)

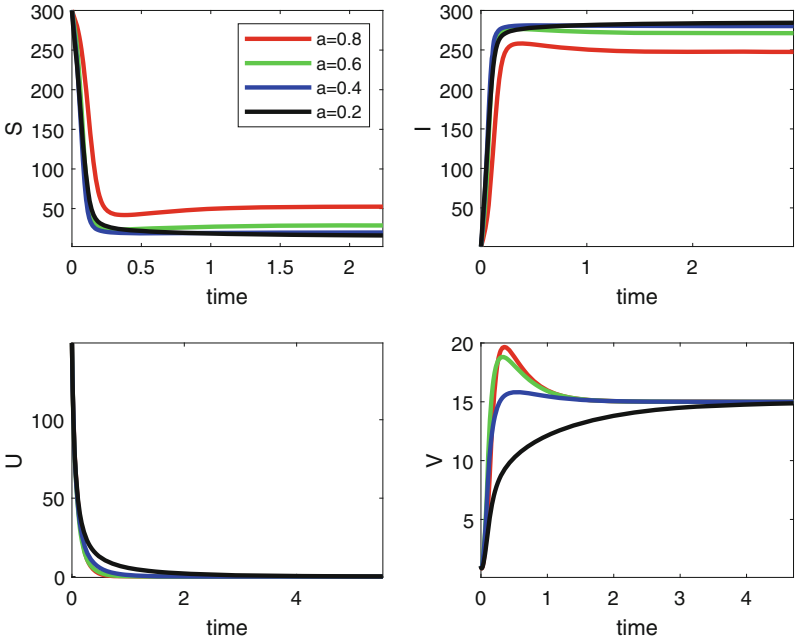


Fig. 8 Model (2), (4), (5): Equilibrium $\widehat{E}_3$ obtained with parameter values (40), but for $n = 6 < b$, $p = 1$ and initial conditions (21)

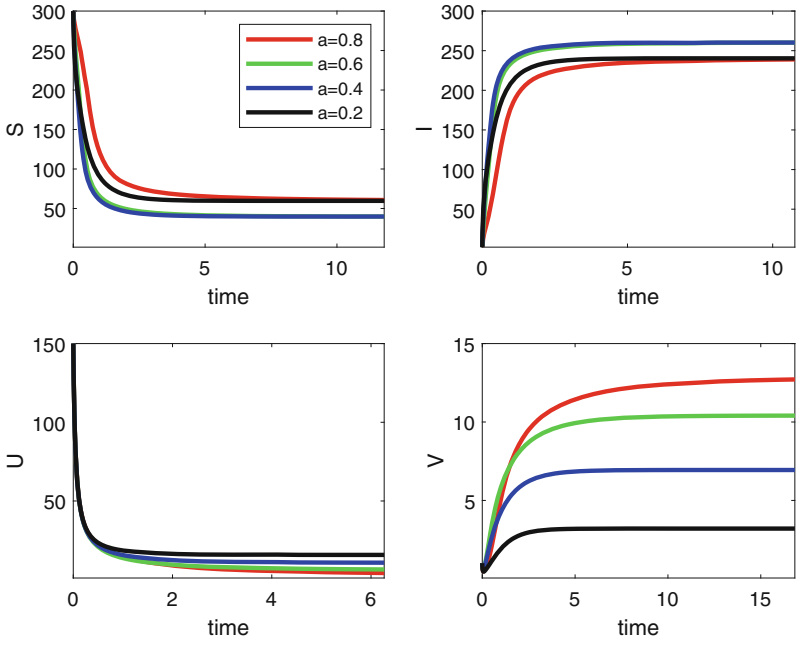


Fig. 9 Model (2), (4), (5): Equilibrium $\widehat{E}_4$ obtained with parameter values (40), but for $n = 6 < b$, $p = 1$, $\beta = 2$ and initial conditions (21)

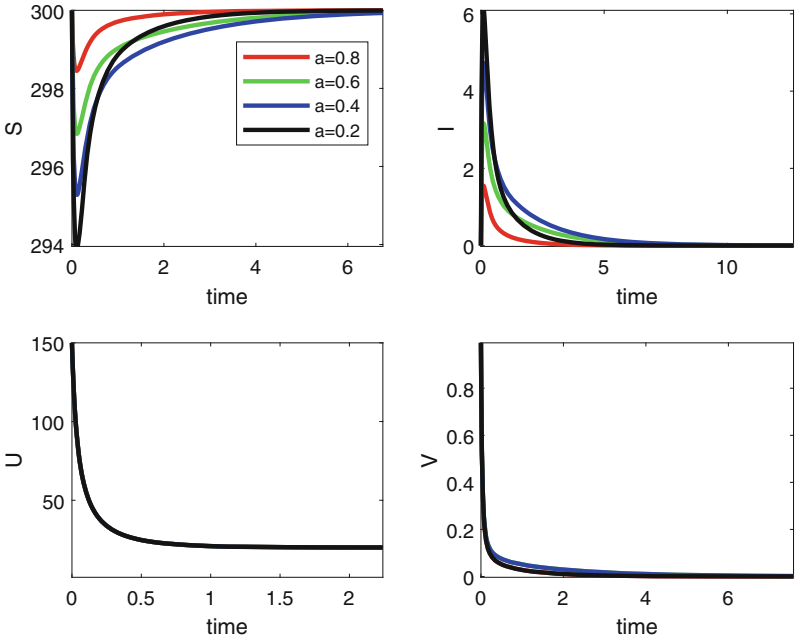


Fig. 10 Model (2), (4), (5): Equilibrium $\widehat{E}_2$ obtained with parameter values (40), but for $n = 6 < b$, $p = 1$, $\beta = 2$, $\lambda = 1$ and initial conditions (21)

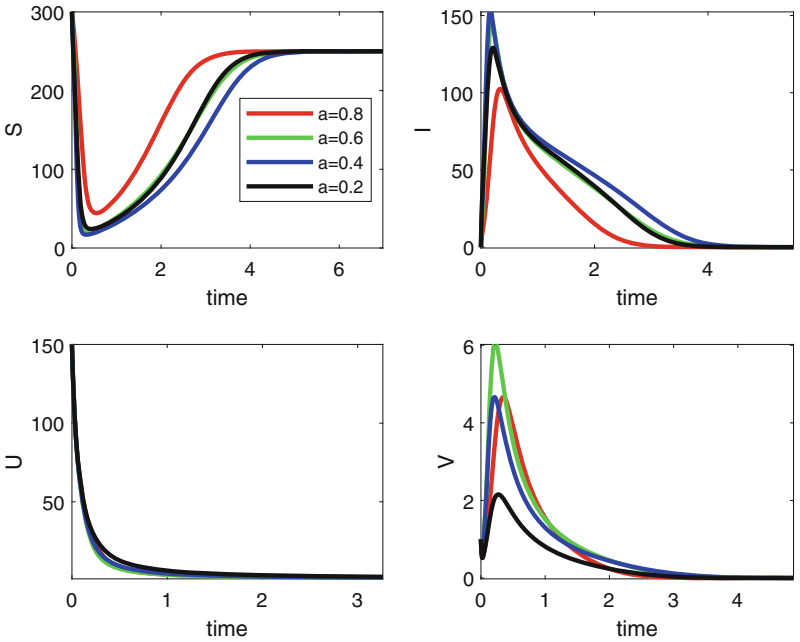


Fig. 11 Model (22)–(25): Equilibrium E_1 obtained with parameter values (41) and initial conditions (21)

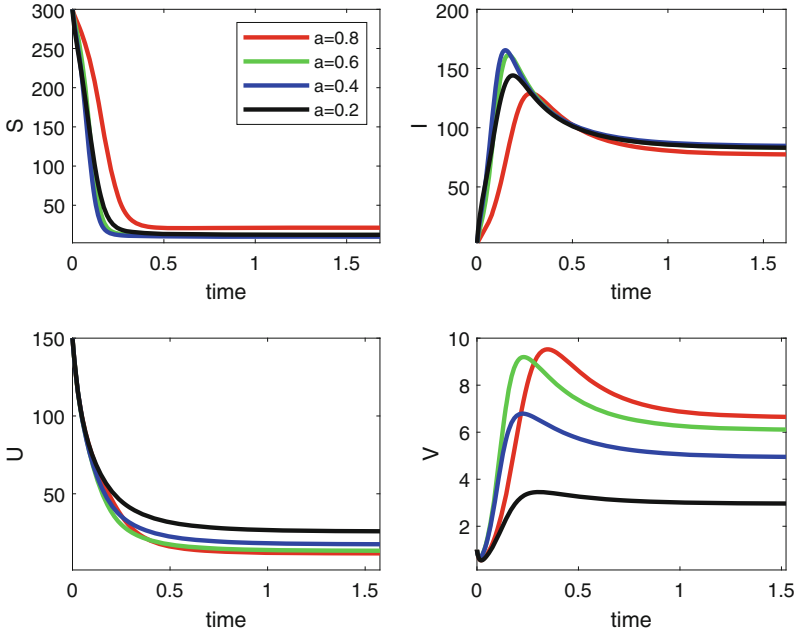


Fig. 12 Model (22)–(25): Equilibrium E_4 obtained with parameter values (41) but for $n = 15$ and initial conditions (21)

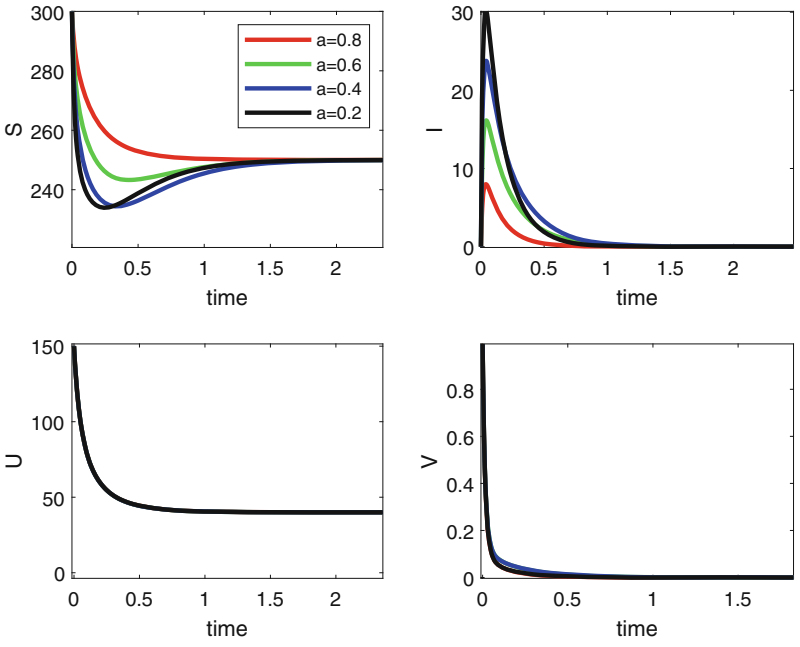


Fig. 13 Model (22)–(25): Equilibrium E_2 obtained with parameter values (41) but for $n = 15$, $\beta = 1$ and initial conditions (21)

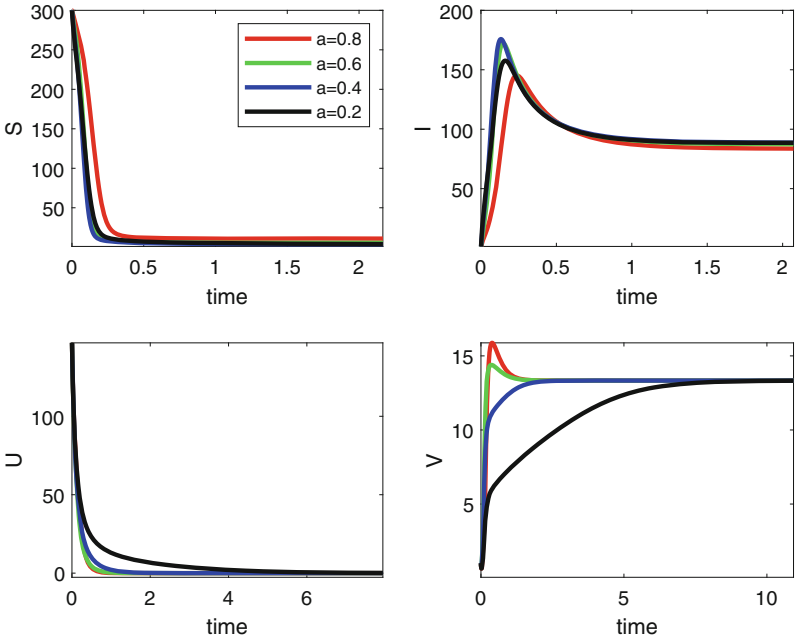


Fig. 14 Model (22)–(25): Equilibrium E_3 obtained with parameter values (41) but for $n = 15$, $p = 1$ and initial conditions (21)

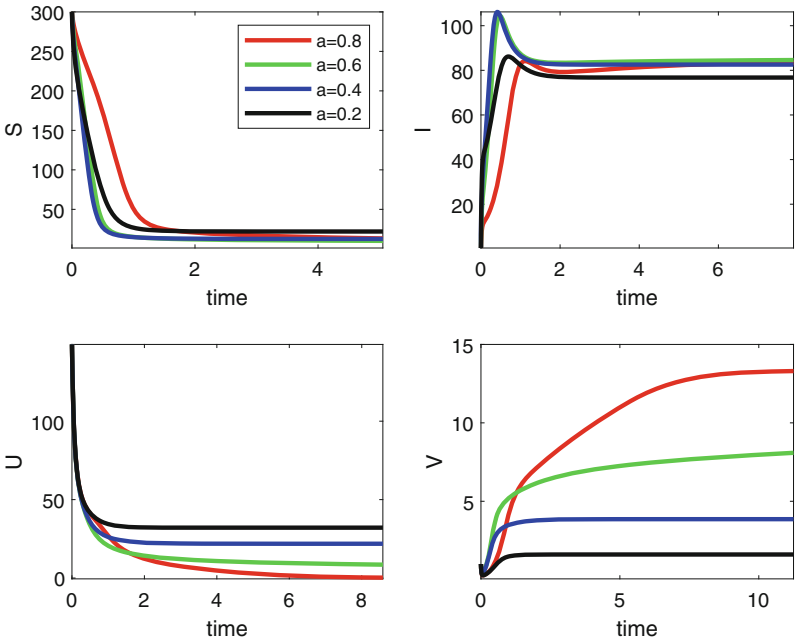


Fig. 15 Model (22)–(25): Equilibrium E_4 obtained with parameter values (41) but for $n = 15$, $p = 1$, $\beta = 3$ and initial conditions (21)

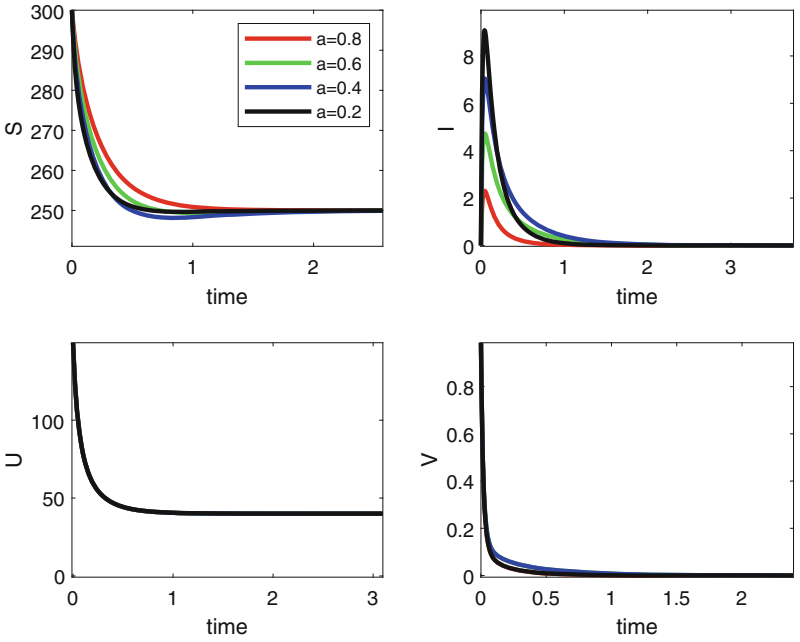


Fig. 16 Model (22)–(25): Equilibrium E_2 obtained with parameter values (41) but for $n = 15$, $p = 1$, $\beta = 3$, $\lambda = 3$ and initial conditions (21)

Acknowledgments Work partially supported by the project "Modelli e metodi numerici in approssimazione, nelle scienze applicate e nelle scienze della vita" of the Dipartimento di Matematica "Giuseppe Peano". Ezio Venturino is a member of the INdAM research group GNCS.

References

1. Bertaccini A., Plant Pathogens, Minor (Phytoplasms), *Encyclopedia of Microbiology* (Fourth Edition): 627-638 (2019).
2. Calari A., Paltrinieri S., Contaldo N., Sakalieva D., Mori N., Duduk B., Bertaccini A., Molecular evidence of phytoplasms in winter oilseed rape, tomato and corn seedlings, *Bulletin of Insectology* **64**, (Supplement): S157-S158 (2011).
3. Erb M., Reymond P., Molecular interactions between plants and insect herbivores, *Annual Review of Plant Biology* **70**, 527-557 (2019).
4. Hogenhout S.A., Bos J.I.B., Effector proteins that modulate plant-insect interactions, *Current Opinion in Plant Biology* **14**, 422-428 (2011).
5. Hogenhout S.A., Sugio A., The genome biology of phytoplasm: modulators of plants and insects, *Current Opinion in Microbiology* **15**, 247-254 (2012).
6. Hogenhout S.A., Kliot A., Marée A.F.M., Tomkins M., A multi-layered mechanistic modelling approach to understand how effector genes extend beyond phytoplasm to modulate plant hosts, insect vectors and the environment, *Current Opinion in Plant Biology* **44**, 39-48 (2018).

7. Makarova O., MacLean A.M., Nicolaisen M., Phytoplasm adapt to the diverse environments of their plant and insect hosts by altering gene expression, *Physiological and Molecular Plant Pathology* **91**, 81-87 (2015).
8. Sugio A., MacLean A.M., Kingdom H.N., Grieve V.M., Manimekalai R., Hogenhout S.A., Diverse targets of phytoplasm effectors: from plant development to defense against insects, *Annual Review of Phytopathology* **49**, 175-195 (2011).
9. Zwolińska A., Krawczyk K., Borodynko-Filas N., Pospieszny H., Non-crop sources of Rapeseed Phyllody phytoplasm ('*Candidatus* Phytoplasm asteris': 16SrI-B and 16SrI-(B/L)L), and closely related strains, *Crop Protection* **119**, 59-68 (2019).

An Interdisciplinary Model-Based Study on Emerging Infectious Disease: The Curse of Twenty-First Century

Anal Chatterjee and Suchandra Ganguly

1 Introduction

We have stepped in early days of 2021 and have carried with us the curse of 2020- COVID-19. It is wreaking havoc on the whole world at present, after its emergence in Wuhan in December 2019 and then global spread since February 2020 [1]. It has been declared as the Public Health Emergency of International concern in January 2020 and a pandemic in March 2020 by WHO. The first case was reported in Wuhan city of Hubei Province in south China on 31, December 2019 as unidentified pneumonia [2]. It is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). As of 12 January 2021, more than 90.9 million cases have been confirmed, with more than 1.94 million deaths attributed to COVID-19. The disease has been able to put the world into a halt. The disease affects individuals in different ways ranging from no, mild, moderate to even severe symptoms requiring hospitalization. The most common symptoms include fever, dry cough, tiredness and less common symptoms are aches and pains, sore throat, diarrhoea, conjunctivitis, headache, loss of taste or smell, a rash on skin, or discolouration of fingers or toes. The disease mainly spreads by airborne transmission. When an infected person coughs, sneezes or speaks the infectious droplets are emitted and can enter another individual by mouth, nose or eyes [3]. It can also spread via fomites when an healthy individual comes in contact of them and the virus reaches their mucous membranes. There is no particular drug available for treatment. Only symptomatic treatment is recommended as per countries policies. Thus, preventive

A. Chatterjee (✉)
Department of Mathematics, Barrackpore Rastraguru Surendranath College, Kolkata, India

S. Ganguly
University College of Nursing, JNMH, Kalyani, India

R. P. Mondaini (ed.), *Trends in Biomathematics: Stability and Oscillations in Environmental, Social, and Biological Models*,
https://doi.org/10.1007/978-3-031-12515-7_19

measures becomes more important than ever. Recommended preventive measures include social distancing, wearing face masks in public, ventilation and air-filtering, hand washing, covering one's mouth when sneezing or coughing, disinfecting surfaces, and monitoring and self-isolation for people exposed or symptomatic [4].

In this scenario, the awareness and adherence to the preventive measures becomes utmost important. Media plays a key role in this aspect. In today's era where media literally dictates one's decisions, it is essential that they play a fruitful role in such pandemic situations. During sudden outbreaks, the public needs access to timely and reliable information about the disease symptoms and its prevention [5]. Nowadays, social media are often seen as fast and effective platforms for searching, sharing, and distributing health information among the general population [6]. Media, thus becomes an essential weapon in our fight against COVID-19. The beneficial guidelines for preventing COVID-19 were reinforced among people through prominent advertisements on commonly used social media platforms. Facebook, Instagram, and television media posted the importance of 'social distancing' and 'stay at home' through free of cost and frequent, widespread ads. The printed media was utilized by supermarkets to promote their stores following the social distancing protocols. During road and air travel, there is continuous mention of ads like 'Stay home, stay safe,' 'Face covers mandatory in public,' 'COVID-19: less is more, avoid gatherings', 'give extra space with each other and on the road,' and 'wash your hands, stay healthy, avoid COVID-19'. This repetition is essential to consolidate the role of them in preventing the disease spread. This campaign was run extraordinarily by the media using all resources and its subtypes [7]. In our study, we have focused on this role of media and how it helps in reducing the spread of COVID-19.

In the present article, we formulated and analyzed a 4-compartment epidemiological model to study the impact of media on the spread of COVID-19, in a variable population with immigration. In the modeling process, we have assumed a population N which is the summation of susceptible unaware, susceptible aware and the infected classes respectively. The susceptible class (both aware and unaware) becomes diseased only by direct contact with the infected class. A part of the susceptible class will make conscious efforts to avoid being in contact with the infected under influence of media. The probability of contracting infection for individuals in aware class is less than those who are in unaware class. Further, we assume that a proportion of individuals recover and a fraction of these recovered individuals will join the aware susceptible class while the others will join the unaware susceptible class (may be due to ignorance, lacunae on their parts etc.). It is also assumed that the growth rate of the cumulative density of media coverage is proportional to the mortality caused by diseases in the infected population. Our study finds that when immigration is increased, the system becomes unstable. Also we found that the use of face masks and the efficiency of face mask, both are vital for maintaining a stable equilibrium. Further, we find that by increasing the implementation of media coverage above a threshold value, the system undergoes from stable to unstable through Hopf-bifurcation. Also, the proportion of infected individuals always decreases with an increase in the density of media coverage. In

the next section, we formulate a mathematical model and examine the equilibrium point and stability of the system. Numerical results are given in later section. Finally, the paper ends with a brief discussion.

2 Literature Review

The World is now facing one of the biggest health challenge in human history in the form of COVID-19. The authors in [7] analyzed the role of mass media and public health communications from December 31, 2019 to July 15, 2020. They reviewed that the media played a dual role in this pandemic situation. They proved advantageous for spreading essential health information, health guidelines, helped in adherence to hygienic practices through repeated advertisements. The media ran the COVID-19 data through live update dashboards which played a big role for providing current situation reports. A trend among people to use telehealth and telemedicine was also noted. But at the same time, various misinformation like unscientific cures, unverified medicines, etc. were also spread using various media platforms. Fear and panic among the general population was also promoted by various media platforms. The authors in [8] concluded that social media has both advantage and disadvantage. The proper use of this will lead to the spread of essential information while misuse will lead to the spread of false information, myths, etc. So, the author advised that to be responsible while disseminating information through social media. Study of the influence of social media on public health measures of COVID-19 via public health awareness and public health behavioral changes in Jordan [9] through quantitative approach was adopted. A web questionnaire was used and 2555 social media users were sampled. The findings revealed that there is a positive influence of media on public health protection against COVID-19 as a pandemic. The analysis of a mathematical model [10] to study the impact of awareness programs by media on the prevalence of infectious disease revealed that by increasing the rate of implementation of awareness programs by media, the number of infected individuals decline and the system remains stable upto a threshold value, after crossing which the system oscillates. The scientists in [11] developed a three-dimensional mathematical model to study the impact of media coverage on the spread and control of infectious diseases. Stability analysis of the model revealed that the disease-free equilibrium is globally-asymptotically stable when the basic reproduction number (R_0) is less than unity. When $R_0 > 1$, the media influence is found to be strong enough. A mathematical model was developed and used to assess the efficacy of face masks, hospitalization and quarantine on COVID-19 [12]. The results revealed the above-mentioned interventions efforts should be high to control the outbreak in a short period of time. It also revealed that the interventions strength should be increased to eliminate the disease but only the sole use of face mask may not be enough in doing so.

3 Basic Assumptions and Model Formulation

B_1: Let N(t) be the total population at time t the region under consideration. Here we consider that the total population is divided into three classes like susceptible unaware population (S_w), susceptible aware population (S_a) and infective population (I).

B_2: Let the rate of immigration of susceptible is A. Also, we consider that M be the cumulative density of media coverage driven by the media in that region at time t which is related to the infective. We assume that diseases spread due to the contact between the susceptible and the infective only.

B_3: It is assumed that susceptible avoid being in contact with the infective due to awareness through media coverage and forms a another class with a proportion λ called the aware susceptible. We assume that after treatment, a proportion of infected individuals recover and join susceptible class. After recovery, a fraction p of recovered people will join aware susceptible class whereas (1−p) will join unaware susceptible class.

B_4: It is notified that the growth rate of the cumulative density of media coverage is proportional to the disease induced mortality rate of the infected population. Here β represents the contact rate of unaware susceptible with infective class and λ be the dissemination rate of awareness through media among susceptible due to which they form a different class. Here β_1 is a fraction which denotes the reduced probability of contracting infection and its value lies between 0 and 1.

B_5: A proportion c_n of population wear face masks correctly and consistently in public places. Let ϵ_n be the efficacy of the face masks. Therefore, $Fn = 1 - \epsilon_n c_n$ represents the fraction which enters the infected class. A proportion h of aware population maintain social distance. The proper use of face masks reduces disease transmission effectively.

B_6: The parameters d, γ and α denote the natural death rate, recovery rate and disease induced death rate respectively. Here, λ_0 represents the transfer rate of aware individuals to unaware susceptible class. The implementation of the awareness through media is proportional to the number of disease induced deaths.

B_7: Let, the density of media coverage increase with increase in disease related death rate α. Here k be the proportionality constant which governs the implementation of awareness through media. We assume that μ_0 is the depletion rate of the media coverage due to ineffectiveness, social and psychological barriers in the population, etc. The parameter m represents the density level of media coverage on the disease from other region.

With these above assumptions our model system (Fig. 1) is

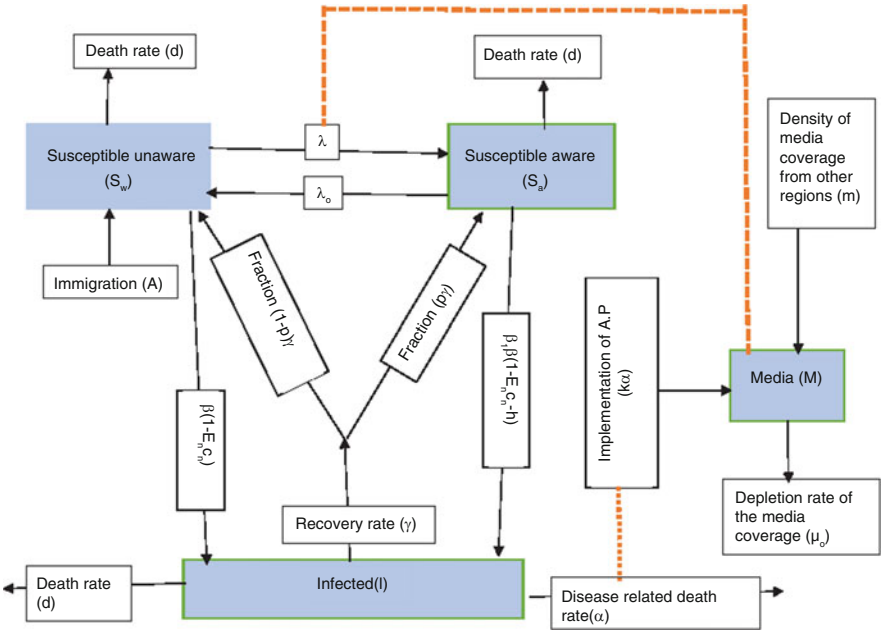


Fig. 1 Schematic diagram

$$\left.\begin{aligned}
\frac{dS_w}{dt} &= A - \beta F_n S_w I - \lambda S_w M - dS_w + \lambda_0 S_a + (1-p)\gamma I \equiv G_1(S_w, S_a, I, M)\\
\frac{dS_a}{dt} &= \lambda S_w M + p\gamma I - \beta_1\beta(F_n - h)S_a I - dS_a - \lambda_0 S_a \equiv G_2(S_w, S_a, I, M)\\
\frac{dI}{dt} &= \beta F_n S_w I + \beta_1\beta(F_n - h)S_a I - \gamma I - \alpha I - dI \equiv G_3(S_w, S_a, I, M)\\
\frac{dM}{dt} &= k\alpha I - \mu_0 M + m \equiv G_4(S_w, S_a, I, M).
\end{aligned}\right\} \tag{1}$$

The system (1) has to be analyzed with the following initial conditions,

$$S_w(0) > 0,\ S_a(0) \geq 0,\ I(0) \geq 0, M(0) \geq 0. \tag{2}$$

Using the fact that $N = S_w + S_a + I$, the system (1) transform to the following system:

$$\left.\begin{aligned}
\frac{dN}{dt} &= A - dN - \alpha I \equiv G_1(N, S_a, I, M)\\
\frac{dS_a}{dt} &= \lambda(N - S_a - I)M + p\gamma I - \beta_1\beta(F_n - h)S_a I - dS_a - \lambda_0 S_a \equiv G_2(N, S_a, I, M)\\
\frac{dI}{dt} &= \beta F_n(N - (1-\beta_1)S_a - I)I - \beta_1\beta h S_a I - (\gamma + \alpha + d)I \equiv G_3(N, S_a, I, M)\\
\frac{dM}{dt} &= k\alpha I - \mu_0 M + m \equiv G_4(N, S_a, I, M).
\end{aligned}\right\} \tag{3}$$

Now it is sufficient to discuss system (3) rather that system (1). Here the region of attraction which is given by the set $\Gamma = \{(N, S_a, I, M) \in R_+^4 : 0 \leq S_w, I \leq N \leq \frac{A}{d}, 0 \leq M \leq \frac{k\alpha(\frac{A}{d})+m}{\mu_0}\}$. According to existence and uniqueness theorem, the trajectories can not approach to unfeasible domain from positive octant which indicates that solution remain in positive octant. This ensure that the system is well defined.

Explicitly, the jacobian matrix at $\overline{E} = (\overline{N}, \overline{S_a}, \overline{I}, \overline{M})$ can be defined as

$$\overline{J} = \begin{bmatrix} -d & 0 & -\alpha & 0 \\ \lambda\overline{M} & m_{22} & m_{23} & \lambda(\overline{N} - \overline{S_a} - \overline{I}) \\ \beta F_n \overline{I} & -\beta F_n(1-\beta_1)\overline{I} - \beta_1\beta h\overline{I} & m_{33} & 0 \\ 0 & 0 & k\alpha & -\mu_0 \end{bmatrix}, \tag{4}$$

where $m_{22} = -(\lambda\overline{M} + \beta_1\beta(F_n - h)\overline{I} + d + \lambda_0)$,
$m_{23} = -\lambda\overline{M} + p\gamma - \beta_1\beta(F_n - h)\overline{S_a}$,
$m_{33} = \beta F_n\overline{N} - \beta F_n(1-\beta_1)\overline{S_a} - 2\beta F_n\overline{I} - \beta_1\beta h\overline{S_a} - (\gamma + \alpha + d)$.

4 Some Preliminary Results

4.1 Equilibria

The system (1) possesses the following equilibria: Disease free equilibrium (DFE) $E_0 = (\frac{A}{d}, \frac{m\lambda A}{d(m\lambda+(d+\lambda)\mu_0)}, 0, \frac{m}{\mu_0})$ and endemic equilibrium $E^* = (N^*, S_a^*, I^*, M^*)$.

4.1.1 Disease Free Equilibrium

E_0 is always feasible. The eigenvalues evaluate from (4) at E_0 are $-d < 0$, $-d - \lambda_0 < 0$, $-\mu_0$ and $(R_0 - 1)$. Thus, it is clearly indicates that E_0 is asymptotically stable if

$$R_0 = \frac{\beta A F_n}{d(\gamma + \alpha + d)} < 1 \tag{5}$$

hold. Here R_0 is the basic reproduction number of system (3). Clearly, E^* exists for $R_0 > 1$.

4.1.2 Endemic Equilibrium

The endemic equilibrium at $E^* = (N^*, S_a^*, I^*, M^*)$ are $N^* = \frac{\frac{\beta}{d}F_n[A-(\alpha+d)I^*]-(\gamma+\alpha+d)}{\beta(1-\beta_1)F_n+\beta\beta_1 h}$, $M^* = \frac{k\alpha I^*+m}{\mu_0}$ while $I*$ is ensured by solving

$$A_1 I^{*2} + A_2 I^* + A_3 = 0, \tag{6}$$

where $A_1 = \frac{\beta_1\beta^2(F_n-h)(\alpha+d)}{c} - \frac{\kappa\lambda\alpha^2}{d\mu_0} - \frac{k\alpha\lambda}{\mu_0} - \frac{k\alpha\lambda}{\mu_0}\frac{\beta}{dc}F_n(\alpha+d)$,
$A_2 = \frac{A\alpha k\lambda}{d\mu_0} - \frac{\lambda\alpha m}{d\mu_0} - \frac{m\lambda}{\mu_0} + p\gamma - \frac{\beta_1\beta^2}{dc}(F_n-h)F_nA + \frac{\beta_1\beta(F_n-h)(\gamma+\alpha+d)}{c} - (d + \lambda_0 + \frac{\lambda m}{\mu_0})\frac{\beta}{dc}F_n(\alpha+d) + \frac{k\alpha\lambda}{\mu_0 c}[\frac{\beta}{d}F_nA - (\gamma+\alpha+d)]$,
$A_3 = (d+\lambda_0+\frac{\lambda m}{\mu_0})F_n\frac{\beta A}{dc} + \frac{Am\lambda}{d\mu_0} - (d+\lambda_0+\frac{\lambda m}{\mu_0})(\frac{\gamma+\alpha+d}{c})$, where $c = \beta(1-\beta_1)(F_n) + \beta\beta_1 h$. Now for $I^* > 0$, solving (6) we get $I^* = \frac{-A_2 \pm \sqrt{A^2-4A_1A_3}}{2A_1}$.

At E^*, the jacobian matrix of system (3) can be written as

$$J^* = \begin{bmatrix} n_{11} & 0 & n_{13} & 0 \\ n_{21} & n_{22} & n_{23} & n_{24} \\ n_{31} & n_{32} & n_{33} & 0 \\ 0 & 0 & n_{43} & n_{44} \end{bmatrix},$$

where $n_{11} = -d < 0$, $n_{13} = -\alpha < 0, n_{21} = \lambda M^* > 0$, $n_{22} = -\lambda M^* - \beta\beta_1(F_n - h)I^* - d - \lambda_0 < 0$, $n_{23} = -\lambda M^* + p\gamma - \beta_1\beta(F_n-h)S_a < 0$, $n_{24} = \lambda(N^* - S_a^* - I^*) > 0, n_{31} = \beta F_n I^* > 0, n_{32} = -\beta(1-\beta_1)F_nI^* - \beta_1\beta h I^* < 0, n_{33} = -\beta F_n I^* < 0, n_{43} = k\alpha > 0, n_{44} = -\mu_0 < 0$.
Now the corresponding characteristic equation is

$$\omega^4 + Q_1\omega^3 + Q_2\omega^2 + Q_3\omega + Q_4 = 0,$$

where the coefficients Q_I, $I = 1, 2, 3, 4$ are $Q_1 = -(n_{11}+n_{22}+n_{33}+n_{44}) > 0$,
$Q_2 = n_{11}n_{22} + n_{22}n_{33} + n_{33}n_{11} + n_{11}n_{44} + n_{22}n_{44} + n_{33}n_{44} - n_{23}n_{32} - n_{13}n_{31}$,
$Q_3 = n_{13}n_{31}n_{44} + n_{23}n_{32}n_{44} + n_{11}n_{23}n_{32} + n_{13}n_{31}n_{22} - n_{11}n_{22}n_{44} - n_{11}n_{33}n_{44} - n_{22}n_{33}n_{44} - n_{11}n_{22}n_{33} - n_{13}n_{21}n_{32} - n_{24}n_{32}n_{43}$,
$Q_4 = n_{11}n_{22}n_{33}n_{44} + n_{13}n_{21}n_{32}n_{44} + n_{11}n_{24}n_{32}n_{43} - n_{11}n_{44}n_{23}n_{32} - n_{13}n_{22}n_{31}n_{44}$.

Now, $Q_2 > 0$ if $n_{23}n_{32} > (n_{11}n_{22} + n_{11}n_{33} + n_{22}n_{33} + n_{11}n_{44} + n_{22}n_{44} + n_{33}n_{44} - n_{13}n_{31})$.

Also, $Q_3 > 0$ if $(n_{13}n_{31}n_{44} + n_{13}n_{31}n_{22} - n_{11}n_{22}n_{44} - n_{11}n_{33}n_{44} - n_{22}n_{33}n_{44} - n_{11}n_{22}n_{33} - n_{24}n_{32}n_{43}) > n_{13}n_{21}n_{32} - n_{11}n_{23}n_{32} - n_{23}n_{32}n_{44}$. Then $Q_1Q_2 - Q_3 > 0$ if $Q_1Q_2 > Q_3$ as well as $Q_3(Q_1Q_2 - Q_3) - Q_1^2Q_4 > 0$ if $Q_3(Q_1Q_2 - Q_3) >$

$Q_1^2 Q_4$. Then, by the Routh-Hurwitz criterion, E^* is locally asymptotically stable which depending upon system parameters.

Remark The system could have a Hopf-bifurcation at the coexistence equilibrium if the following two conditions are satisfied,

$$Q_1(A_c)Q_2(A_c) - Q_3(A_c) = 0, \quad Q_1'(A_c)Q_2(A_c) + Q_1(A_c)Q_2'(A_c) - Q_3'(A_c) \neq 0. \tag{7}$$

4.2 Hopf Bifurcation at Coexistence

Theorem (Hopf-Bifurcation) *If $\psi_1(A) > 0$, then the equilibrium E^* of system (3) is locally asymptotically stable. If there exists $A_c \in R$ such that $\psi_1(A_c) = 0$ and $(\frac{d\psi_1}{dA}) \mid_{A_c} \neq 0$, then as A passes through A_c, a Hopf-bifurcation occurs at E^*.*

For positive equilibrium $E^* = (N^*, S_a^*, I^*, M^*)$, the characteristic equation is

$$\omega^4 + Q_1\omega^3 + Q_2\omega^2 + Q_3\omega + Q_4 = 0.$$

Define

$$\psi_1(A) = Q_1(A)Q_2(A)Q_3(A) - Q_3^2(A) - Q_1^2(A)Q_4(A). \tag{8}$$

Let $\omega_i (i = 1, 2, 3, 4)$ be the roots of above characteristic equation. Then we have

$$\begin{aligned} \omega_1 + \omega_2 + \omega_3 + \omega_4 &= -Q_1, \\ \omega_1\omega_2 + \omega_1\omega_3 + \omega_1\omega_4 + \omega_2\omega_3 + \omega_2\omega_4 + \omega_3\omega_4 &= Q_2, \\ \omega_1\omega_2\omega_3 + \omega_1\omega_3\omega_4 + \omega_2\omega_3\omega_4 + \omega_1\omega_2\omega_4 &= -Q_3, \\ \omega_1\omega_2\omega_3\omega_4 &= Q_4. \end{aligned} \tag{9}$$

If there exists $A_c \in R$ such that $\psi_2(A_c) = 0$, then by the Routh-Hurwitz criterion at least one root, say ω_1, has real part equal to zero. From the fourth equation of (8) it follows that $Im\ \omega_1 = \omega_0 \neq 0$, and hence there is another root, say ω_2, such that $\omega_2 = \overline{\omega}_1$. Since $\psi_2(A)$ is a continuous function of its roots, ω_1 and ω_2 are complex conjugate for A in an open interval including A_c. Therefore, the equation in (8) have the following form at A_c,

$$\begin{aligned} \omega_3 + \omega_4 &= -Q_1, \\ \omega_0^2 + \omega_3\omega_4 &= Q_2, \\ \omega_0^2(\omega_3 + \omega_4) &= -Q_3, \\ \omega_0^2\omega_3\omega_4 &= Q_4. \end{aligned} \tag{10}$$

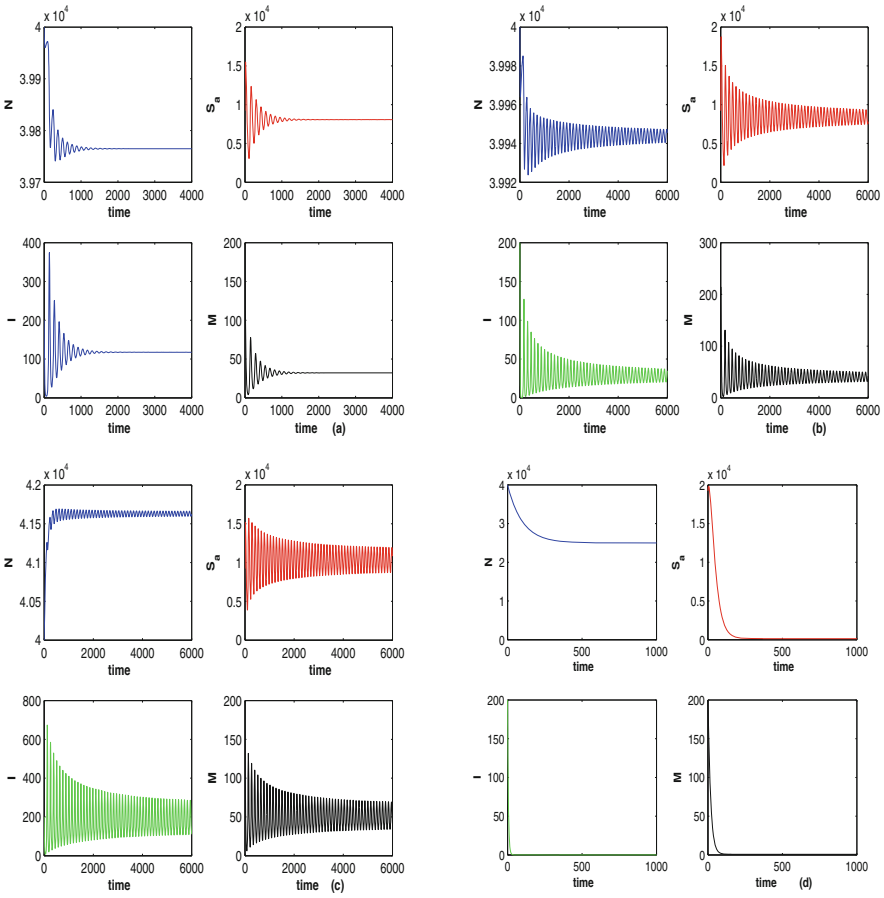


Fig. 2 **(a)** The equilibrium point E^* is stable for the parametric values as given in (11). **(b)** The figure depicts oscillatory behavior around the coexistence (endemic) equilibrium point E^* of system (1) for $k = 4.2$. **(c)** The figure depicts oscillatory behavior around coexistence (endemic) equilibrium point E^* of system (3) for $A = 450$. **(d)** The figure depicts disease free equilibrium E_0 for $A = 250$

If ω_3 and ω_4 are complex conjugate, from the first Eq. (9) it follows that $2Re\ \omega_3 = -Q_1 < 0$. If ω_3 and ω_4 are real, from the first and fourth equations of (9) it follows that $\omega_3 < 0$ and $\omega_4 < 0$. Also after some calculations it follows that $\frac{d}{d\gamma_2} Re(\omega_1)_{A=A_c} = -\frac{Q_1}{2[Q_1^2 Q_4 + (Q_1 Q_2 - 2Q_3)^2]} \frac{d\psi_1}{dA} \mid_{A_c} \neq 0$. Thus, we have the following result.

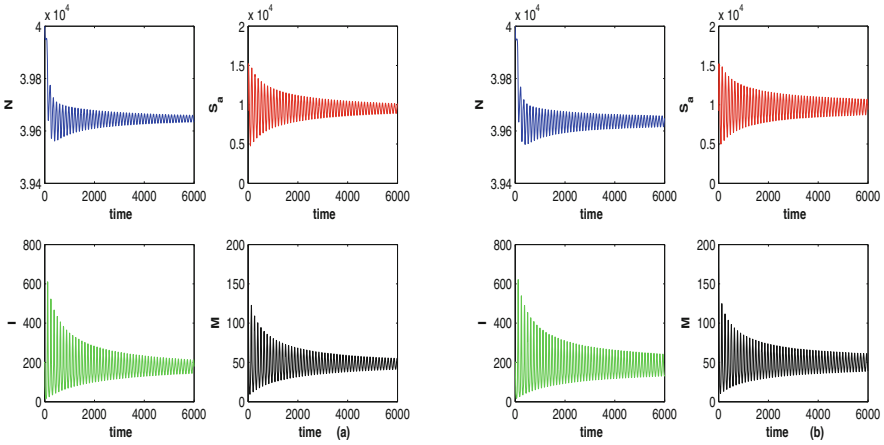


Fig. 3 **(a)** The figure depicts oscillatory behavior around coexistence (endemic) equilibrium point E^* of system (3) for $\epsilon_n = 0.1$. **(b)** The figure depicts oscillatory behavior around coexistence (endemic) equilibrium point E^* of system (3) for $c_n = 0.01$

5 Numerical Simulations

In this section, we study the impact of awareness programs with the help of numerical simulation. Here we investigate the effects of the various parameters on the qualitative behavior of the system, by using MATLAB. We begin with a parametric values[10, 13, 14]

$$A = 400, \beta = 0.00002, \beta_1 = .2, \lambda = 0.0002, \lambda_0 = .02, \gamma = .6, \alpha = 0.02,$$
$$d = 0.01, \mu_0 = 0.06, \epsilon_n = 0.5, c_n = .1, h = 0.02,\ k = 0.8,\ p = 0.05, m = .05. \tag{11}$$

Dealing with above set of parametric values, we note that the system is locally asymptotically stable at endemic equilibrium $E^* = (39637, 7834, 182, 37)$ in which $R_0 = 1.2698$(cf. Fig. 2a). Taking $k = 4.2$, the system exhibits oscillations around E^* (cf. Fig. 2b). Figure 2c illustrate the oscillatory behavior of each population for high value of A ($A = 450$). Analytical, we see that endemic equilibrium E^* exists if $A > \frac{d(\gamma+\alpha+d)}{\beta}$. We obtain the critical value of immigration rate $A = 333$, above which the endemic equilibrium exists. Taking $A = 250$, we observe that the system exhibits disease free equilibrium E_0 which satisfy our analytical finding (cf. Fig. 2d). It is interesting to see that low value of $\epsilon_n = 0.1$ and $c_n = 0.01$ play a big impact to destabilize the whole system respectively (cf. Fig. 3a, b). Now for clear understanding of dynamic change, we plot a bifurcation

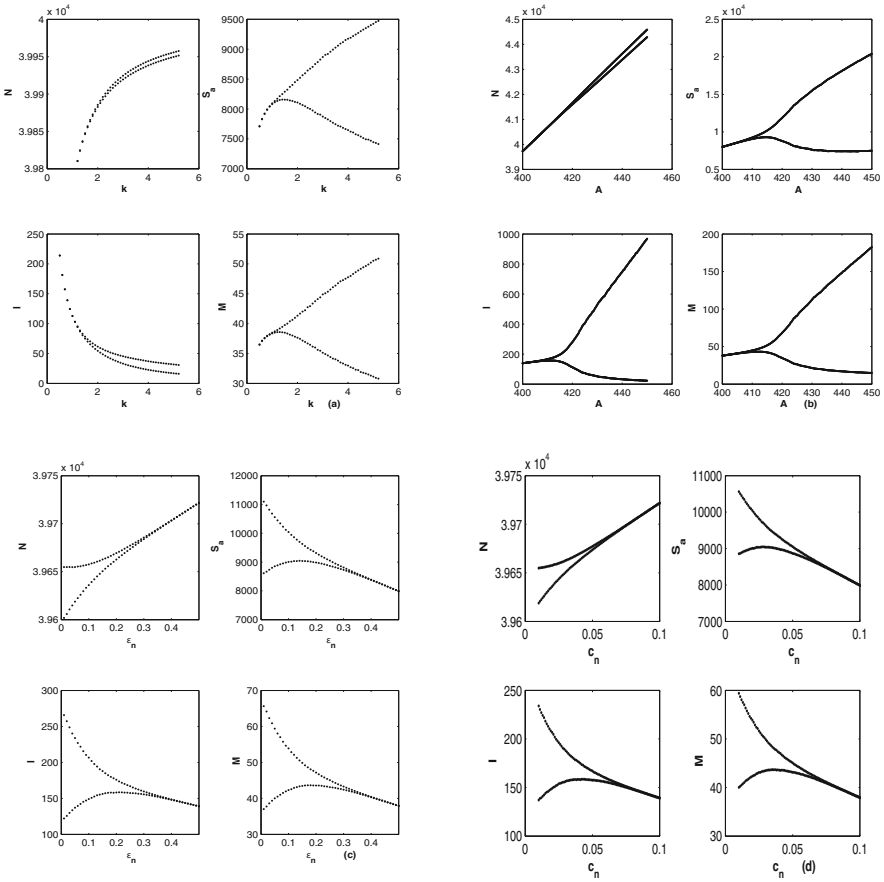


Fig. 4 **(a)** Bifurcation diagram for k. **(b)** Bifurcation diagram for A. **(c)** Bifurcation diagram for ϵ_n. **(d)** Bifurcation diagram for c_n

diagram with respect to k. Form Fig. 4a, it follows that lower values of k, the system is stable but above a threshold value of $k = k_c$, the system losses its stability and periodic solution arises through Hopf-bifurcation. Further, we also vary A as a free parameter, a bifurcation diagram (cf. Fig. 4b) indicates that the system looses its stability for high value of A after it crosses the critical value. Further, we plot another two bifurcation diagrams for efficacy of the face masks i.e. ϵ_n and masks compliance, c_n respectively. It is clear to see that the system looses stability for low value of these two parameters (cf. Fig. 4c, d). Figure 5a illustrates the different steady state behaviour of infected class in the system (3) for the parameter A. Here, we see a Hopf bifurcation points at $A = 419$ (denoted by a red star (H)) with eigenvalues -0.103689, -0.001029, $\pm 0.52642i$ and first Lyapunov coefficient being $-1.881669e^{-10}$ and generates a family of stable limit cycle bifurcates from the H and loses its stability. Here $A = 333(BP)$ denotes the branch point of the

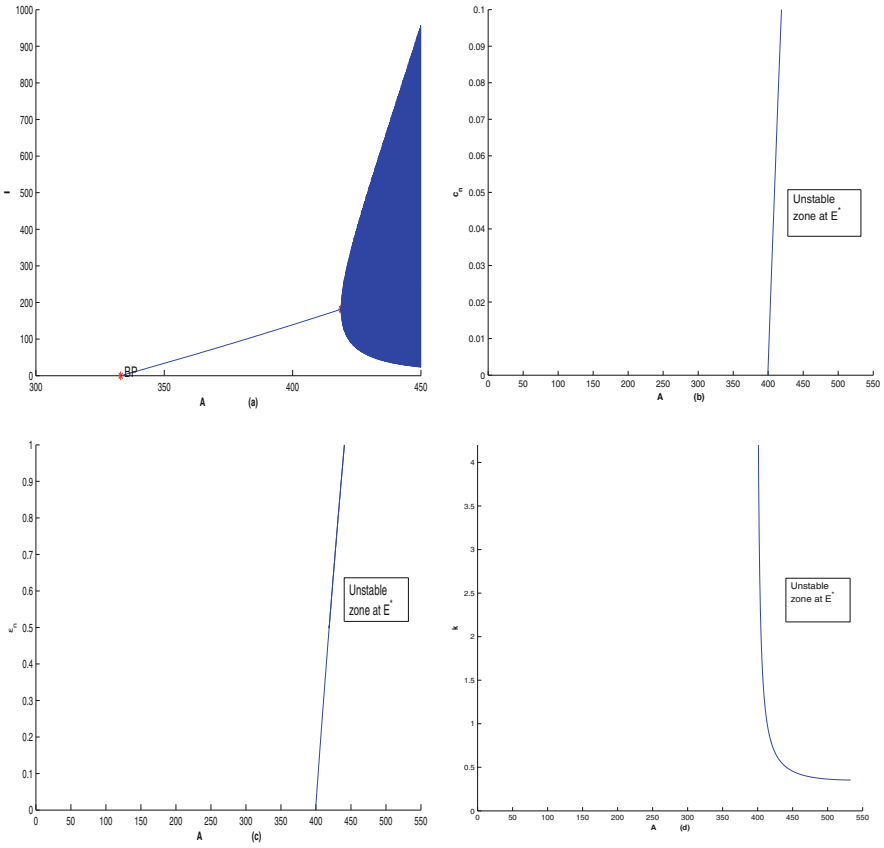


Fig. 5 **(a)** The figure depicts different steady-state behaviors of infected population for the effect of A. **(b)** Two parameter bifurcation diagram for $A - c_n$. **(c)** Two parameter bifurcation diagram for $A - \epsilon_n$. **(d)** Two parameter bifurcation diagram for $A - k$

system (3) with eigenvalues are o, -0.06, -0.03, -0.01. Figure 5b–d represent two parameters bifurcation diagrams for $A - c_n$, $A - \epsilon_n$ and $A - k$ respectively.

6 Discussion

The information and the awareness of the preventive strategy for COVID-19 is majorally emphasized through media coverage. So, in our paper we have analyzed a 4-compartment mathematical model. It is assumed that pathogens are transmitted via direct contact between the susceptible and the infective. The model exhibits two equilibria like the disease-free equilibrium and endemic equilibrium under certain conditions. Firstly, the model is studied analytically and shown that when the basic

reproduction number $R_0 < 1$, the system exhibits disease-free equilibrium. For $R_0 > 1$, it leads to the existence of an endemic equilibrium.

Our study indicates that if we increase the density of media coverage, the number of infected individuals decline. But after crossing the threshold value, system becomes unstable. The constant immigration may be one of the possible reasons of such outcomes. Further, we observe that lower value of immigration rate the system becomes disease-free equilibrium. Also, the efficacy of face mask and it's usage in public areas helps in keeping the system stable.

References

1. M. Mandal, S. Jana, S. K. Nandi, A. Khatua, S. Adak and T. K. Kar. A model based study on the dynamics of COVID-19: Prediction and control. *Chaos, solitons, and fractals*, 136, 109889 (2020) .
2. H. Lu, C. W. Stratton and Y. W. Tang. Outbreak of pneumonia of unknown etiology in Wuhan, China: the mystery and the miracle. *Journal of medical virology,* 92(4), 401–402 (2020).
3. N. V. Doremalen, T. Bushmaker, D. H. Morris, M. G. Holbrook, A. B. N. Williamson Gamble, A. B. N. et al. Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1. *New England Journal of Medicine*, 382, 1564–1567 (2020).
4. World Health Organization, Coronavirus disease 2019 (COVID-19). WHO situation report -73, 2020. Avaiable from: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200402-sitrep-73-covid-19.pdf.
5. P. Bastani, and M. A. Bahrami. COVID-19 Related Misinformation on Social Media: A Qualitative Study from Iran. *Journal of medical Internet research* (2020).
6. Y. Zhao and J. Zhang. Consumer health information seeking in social media: a literature review. *Health information and libraries journal*, 34(4), 268–283 (2017).
7. A. Anwar, M. Malik, V. Raees and A. Anwar. Role of mass media and public health communications in the COVID-19 pandemic. *Cureus*, 12(9) (2020).
8. D. A. González-Padilla and L. Tortolero-Blanco. Social media influence in the COVID-19 pandemic. *International braz j urol*, 46, 120–124 (2020).
9. H. Al-Dmour, A. Salman, M. Abuhashesh and R. Al-Dmour. Influence of social media platforms on public health protection against the COVID-19 pandemic via the mediating effects of public health awareness and behavioral changes: integrated model. *Journal of medical Internet research*, 22(8), e19996 (2020).
10. S. Samanta, S. Rana, A. Sharma, A.K. Misra and J. Chattopadhyay. Effect of awareness programs by media on the epidemic outbreaks: A mathematical model. *Applied Mathematics and Computation*, 219(12), 6965–6977 (2013).
11. J. Cui, Y. Sun, and H. Zhu. The impact of media on the control of infectious diseases. *Journal of dynamics and differential equations*, 20(1), 31–53 (2008).
12. A. K. Srivastav, P. K. Tiwari, P. K. Srivastava, M. Ghosh and Y. Kang. A mathematical model for the impacts of face mask, hospitalization and quarantine on the dynamics of COVID-19 in India: deterministic vs. stochastic[J]. *Mathematical Biosciences and Engineering*, 18(1): 182–213 (2021).
13. L. Zuo and M. Liu. Effect of Awareness Programs on the Epidemic Outbreaks with Time Delay. Abstract and Applied Analysis, *Hindawi Publishing Corporation*, Volume 2014, Article ID 940841 (2014).
14. L. Zuo, M. Liu and J. Wang. The Impact of Awareness Programs with Recruitment and Delay on the Spread of an Epidemic, Mathematical Problems in Engineering, *Hindawi Publishing Corporation*, Volume 2015, Article ID 235935 (2015).

Effect of Incubation Delay in an SIRS Model: Hopf Bifurcation and Stability Switch

Tanuja Das and Prashant K. Srivastava

1 Introduction

The infectious diseases spread in the population when the infectious agents are transmitted from infective to susceptible in a direct or indirect way. For instance, in diseases such as COVID-19, Measles, Gonorrhea, etc. the infectious agents are directly transmitted, whereas in diseases such as Malaria, Lyme disease, the infection is transmitted by the vectors. Once the healthy individual is exposed to the infectious agent, the infectious agent takes some time to develop in the host. This time period is considered as incubation period, after which a susceptible becomes infectious and may spread the infection further. This incubation delay length is not fixed and depending on the disease's biological mechanism, it may vary from several days (COVID-19) to several weeks (smallpox). The other factors which may impact the disease incubation period include the disease severity, people's immunity, and the amount of infectious agents people are exposed to.

In order to study the effect of the incubation period on the dynamics of a particular disease, one of the way to incorporate it in a model is via use of discrete time delay in a compartmental model. However, such introduction increases the complexity of the system from mathematical point of view as delay differential equation model systems possess solutions on infinite dimensional functional spaces. Still, it is very interesting to examine such models and study the dynamical properties as exhibited by the delay models. In works [7, 11, 15, 17], one can see the complex behavior of system for the positive time delays. Authors in [9, 12, 13] found the delay dependent conditions under which system become globally asymptotically stable. In works [4, 10], it is observed that system may have same stable nature for all

T. Das (✉) · P. K. Srivastava
Department of Mathematics, Indian Institute of Technology Patna, Patna, India

R. P. Mondaini (ed.), *Trends in Biomathematics: Stability and Oscillations in Environmental, Social, and Biological Models*,
https://doi.org/10.1007/978-3-031-12515-7_20

delay length conditionally. The incubation delay parameter leads the disease system through the local Hopf bifurcation in the works in [7, 11, 15–17], as a result local stability nature of the system will change. It is notable that the stability properties are affected in most cases by a critical value of the delay. For this delay threshold, the dynamical properties change significantly.

To impede the disease transmission, one of the most important ways is reducing the effective interaction between susceptible and infective people. The information about presence of disease in population induces inhibitory behaviour changes among healthy populations [1]. Looking at the influence of inhibitory effect in people, various form of incidence functions are used in disease modeling literature [3, 5, 6, 8, 18]. For example, the Monod—Haldane type incidence function $Sf(I) = S\frac{\beta I}{1+mI^2}$, $\beta, m > 0$ was introduced by Xiao and Ruan in 2007 [14]. This function reflects the non monotonic interaction behavior of people, which means upto a certain range $I = \frac{1}{\sqrt{m}}$ the susceptible interact with increasing number of infective (as $\frac{df}{dI} > 0$ for $I < \frac{1}{\sqrt{m}}$) and after that susceptible interact with decreasing number of infective (as $\frac{df}{dI} < 0$ for $I > \frac{1}{\sqrt{m}}$). Here the parameter m is introduced because of changing behavioral response of individual due to inhibitory effect. Authors in [19], studied this function for the delay model.

In this work an $SIRS$ model is proposed where the incidence rate is considered as $S\frac{\beta I}{1+mI^2}$. Also, the time delay is incorporated for the disease incubation period. In the next section we describe the delay differential equation system and establish the basic properties. In Sect. 3, local stability of all the equilibrium points along with Hopf bifurcation is provided. In Sect. 4, numerical simulation is performed to explore the analytical results of Sect. 3. At the end, the discussion along with the conclusion are provided.

2 Model Mathematical Form

The total population is divided into three classes—the susceptible class $S(t)$, the infective class $I(t)$ and the recovered class $R(t)$ at any time t. Healthy individuals are recruited in the susceptible class following logistic growth rate where b is the intrinsic growth rate and K is the maximum capacity of individual in susceptible class when there is no infection. The parameter β is representing the rate of disease transmission from susceptible class to infective class which follows a simplified Monod-haldane type incidence to represent the inhibitory behaviour of healthy individuals against the disease. Parameter m is related to the changing behavioral response of individual due to inhibitory effect [18]. The rate parameters α, and δ denote the recovery rate and mortality rate of infective individuals, respectively. The parameter γ is representing the rate at which individuals from the recovered class revert back to the susceptible class after losing their immunity. In case, $\gamma = 0$, then recovered individuals become permanent immune and remain in the recovered class forever. Parameter d represents the natural death rate of individuals in all the

three classes. Here, we assume that upon infection, individuals spend an incubation time before becoming infectious and this time is reflected via a discrete time delay τ in the incidence rate function in the infective class rate equation. Hence, the disease dynamics is represented by the below mentioned set of delay differential equations:

$$\begin{aligned}
\frac{dS(t)}{dt} &= bS(t)\left(1-\frac{S(t)}{K}\right)-\frac{\beta S(t)I(t)}{1+mI^2(t)}+\gamma R(t),\\
\frac{dI(t)}{dt} &= \frac{\beta S(t-\tau)I(t-\tau)}{1+mI^2(t-\tau)}-(d+\delta)I(t)-\alpha I(t),\\
\frac{dR(t)}{dt} &= \alpha I(t)-dR(t)-\gamma R(t),
\end{aligned} \tag{1}$$

with initial conditions $S(\theta) = S_0 \geq 0$, $I(\theta) = I_0 \geq 0$, $R(\theta) = R_0 \geq 0$ and $\theta \in [-\tau, 0]$, where $(S(\theta), I(\theta), R(\theta)) \in C([-\tau, 0], \mathbb{R}^3_+)$: a Banach space of continuous functions.

2.1 Positivity and Boundedness

In order ensure the positivity of all the population classes (S, I, R) we consider the initial population size so that all the populations remain positive for all time.

Beside, the sum of all the population classes, $N(t) = S(t) + I(t+\tau) + R(t+\tau)$ satisfies the equation,

$$\frac{dN}{dt} = bS\left(1-\frac{S}{K}\right)-d(N-S)-\delta I(t+\tau) \leq dN+(b+d)S_M,$$

where $S_M = \max\{S(0), K\}$. Then the biological feasible region of system (1) becomes,

$$\Gamma := \left\{(S, I, R) \in \mathbb{R}^3_+ : 0 \leq S, I, R \leq (b+d)S_M\right\}, \quad S_M = \max\{S(0), K\}.$$

2.2 Existence of Equilibrium Point

For the system (1), there always exists the disease free equilibrium points $E_0 = (0, 0, 0)$, and $E_1 = (K, 0, 0)$. Then following Ref.[2] the basic reproduction number is obtained as

$$\mathcal{R}_0 = \frac{\beta K}{d+\delta+\alpha}.$$

Table 1 Number of possible endemic equilibrium points for $\mathcal{R}_0 > 1$

Cases	Coefficients				Number of I_*	Endemic equilibrium point
	A_4	A_2	A_1	A_0		
$1 < \mathcal{R}_0 \leq 2$	+	+	+	−	1	E_*
$2 < \mathcal{R}_0$	+	−	+	−	3 or, 1	E_{3*}, E_{2*}, E_{1*} or, E_*

Also for system (1), there exists the endemic equilibrium point $E_* = (S_*, I_*, R_*)$ with $S_* = \frac{1+mI_*^2}{\beta}(d+\delta+\alpha)$, $R_* = \frac{\alpha I_*}{d+\gamma}$. I_* is the real positive solution of $h(I_*) = 0$, where,

$$h(I_*) := A_4 I_*^4 + A_3 I_*^3 + A_2 I_*^2 + A_1 I_* + A_0. \tag{2}$$

Here coefficients $A_4 = bm^2(d+\delta+\alpha)^2(d+\gamma)$, $A_3 = 0$, $A_2 = mb(d+\delta+\alpha)^2(d+\gamma)(2-\mathcal{R}_0)$, $A_1 = ((d+\gamma)(d+\delta)+d\alpha)\beta^2 K$ and $A_0 = b(d+\delta+\alpha)^2(1-\mathcal{R}_0)$.

Clearly, for $\mathcal{R}_0 < 1$ all coefficients of $h(I_*) = 0$ (A_4, A_3, A_2, A_1, A_0) are non-negative which means there is no sign change. Thus, no equilibrium point exists for $\mathcal{R}_0 < 1$. Further, for $\mathcal{R}_0 > 1$ we obtain $A_4 > 0$ and $A_0 < 0$ and so the equation $h(I_*) = 0$ has at least a real positive solution. This follows from the continuity of the $h(I^*)$ and $h(0) < 0$ and $h(I^*) \to +\infty$ as $I^* \to \infty$. Hence, an endemic equilibrium point exists for $\mathcal{R}_0 > 1$.

In fact there is a possibility that $h(I_*) = 0$ has multiple real positive roots. In Table 1 we show the number of possible real positive solution of $h(I_*) = 0$ and the corresponding endemic equilibrium points.

We summarise the above discussion in the following theorem.

Theorem 2.1 *For the system (1), no endemic equilibrium point exists when $\mathcal{R}_0 \leq 1$, a unique endemic equilibrium point exists when $1 < \mathcal{R}_0 \leq 2$ and multiple endemic equilibrium points may exist when $\mathcal{R}_0 > 2$.*

3 Stability Analysis

In this section we will analyze the local stability properties of equilibrium points. First we will determine the stability of endemic equilibrium point for all non negative time delay. The linearization system of model (1) at the endemic equilibrium point $E_* = (S_*, I_*, R_*)$ is given by

$$\frac{dY}{dt} = M_1 Y(t) + M_2 Y(t-\tau), \tag{3}$$

where,

$$M_1 = \begin{bmatrix} -\frac{bS_*}{K} & -\frac{\beta S_*(1-mI_*^2)}{(1+mI_*^2)^2} & \gamma \\ 0 & -(d+\delta+\alpha) & 0 \\ 0 & \alpha & -(d+\gamma) \end{bmatrix}, \quad M_2 = \begin{bmatrix} 0 & 0 & 0 \\ \frac{\beta I_*}{1+mI_*^2} & \frac{\beta S_*(1-mI_*^2)}{(1+mI_*^2)^2} & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{4}$$

and $Y(t) = [S(t), I(t), R(t)]^T$.
Then the corresponding characteristic equation of the linearized system (3) is given by

$$|M_1 + M_2 e^{-\lambda\tau} - \lambda I_{3\times 3}| = 0, \tag{5}$$

where $I_{3\times 3}$ is the third order identity matrix. Then Eq. (5) gives,

$$\begin{vmatrix} b(1-\frac{2S_*}{K}) - \frac{\beta I_*}{1+mI_*^2} - \lambda & -\frac{\beta S_*(1-mI_*^2)}{(1+mI_*^2)^2} & \gamma \\ \frac{\beta I_*}{1+mI_*^2}e^{-\lambda\tau} & \frac{\beta S_*(1-mI_*^2)}{(1+mI_*^2)^2}e^{-\lambda\tau} - (d+\delta+\alpha) - \lambda & 0 \\ 0 & \alpha & -(d+\gamma+\lambda) \end{vmatrix} = 0.$$

Or,

$$\phi(\lambda,\tau) := \lambda^3 + B_2\lambda^2 + B_1\lambda + B_0 + (C_2\lambda^2 + C_1\lambda + C_0)e^{-\lambda\tau} = 0, \tag{6}$$

where,

$$\begin{aligned} B_2 =& -b(1-\frac{2S_*}{K}) + \frac{\beta I_*}{1+mI_*^2} + 2d + \delta + \alpha + \gamma, \\ B_1 =& (d+\delta+\alpha)(d+\gamma) - (b(1-\frac{2S_*}{K}) - \frac{\beta I_*}{1+mI_*^2})(2d+\delta+\alpha+\gamma), \\ B_0 =& -(b(1-\frac{2S_*}{K}) - \frac{\beta I_*}{1+mI_*^2})(d+\delta+\alpha)(d+\gamma), \\ C_2 =& -\frac{\beta S_*(1-mI_*^2)}{(1+mI_*^2)^2}, \\ C_1 =& \frac{\beta S_*(1-mI_*^2)}{(1+mI_*^2)^2}(b(1-\frac{2S_*}{K}) - d - \gamma), \\ C_0 =& b(1-\frac{2S_*}{K})\frac{\beta S_*(1-mI_*^2)}{(1+mI_*^2)^2}(d+\gamma) - \alpha\gamma\frac{\beta I_*}{1+mI_*^2}. \end{aligned} \tag{7}$$

It is easily seen that Eq. (6) becomes an algebraic equation of λ when $\tau = 0$. Then following Routh-Hurwitz criterion we can say that all roots of Eq. (6) are with negative real part if following conditions are satisfied,

$$(B_2 + C_2)(B_1 + C_1) - (B_0 + C_0) > 0, \quad B_2 + C_2 > 0, \quad B_0 + C_0 > 0, \tag{8}$$

and at least one of the root of Eq. (6) is positive provided $B_0 + C_0 < 0$. Hence when $\tau = 0$, E_* is stable when conditions (8) are satisfied and is unstable when $B_0 + C_0 < 0$.

Equation (6) becomes a transcendental equation of λ when $\tau > 0$. Then it has infinitely many complex roots and finding all these roots is impossible. In order to determine the stability of endemic equilibrium E_*, we find only the pure imaginary roots of Eq. (6). Because then only the real part of roots cross the imaginary axis for changing the time delay and consequently stability nature of E_* may be changed. Let $\lambda = \iota\omega$ ($\omega \in \mathbb{R}/0$) is the purely imaginary root obtained at $\tau = \tau^*$, and so $\phi(\iota\omega, \tau^*) = 0$ (Eq. (6)). In the following we will evaluate the values of ω and τ^* and their existence conditions.

For this, we substitute $\lambda = \iota\omega$, ($\omega \in \mathbb{R}/0$) in (6) when $\tau = \tau^*$ and we have,

$$-\iota\omega^3 - B_2\omega^2 + B_1\iota\omega + B_0 + (-C_2\omega^2 + C_1\iota\omega + C_0)e^{-\iota\omega\tau^*} = 0. \tag{9}$$

Comparing the real and imaginary parts in (9), we get,

$$B_2\omega^2 - B_0 = (C_0 - C_2\omega^2)\cos\omega\tau^* + C_1\omega\sin\omega\tau^*, \tag{10}$$

$$\omega^3 - B_1\omega = C_1\omega\cos\omega\tau^* - (C_0 - C_2\omega^2)\sin\omega\tau^*. \tag{11}$$

Now squaring and adding both side of Eqs. (10) and (11), we obtain,

$$(B_2\omega^2 - B_0)^2 + (\omega^3 - B_1\omega)^2 = (C_0 - C_2\omega^2)^2 + C_1^2\omega^2. \tag{12}$$

Putting $\omega^2 = m_1$ in Eq. (12), we get,

$$\xi(m_1) = m_1^3 + D_2m_1^2 + D_1m_1 + D_0 = 0, \tag{13}$$

where $D_2 = B_2^2 - C_2^2 - 2B_1 = (b(1 - \frac{2S_*}{K}) - \frac{\beta I_*}{1+mI_*^2})^2 + 4m(d+\delta+\alpha)^2\frac{I_*^2}{(1+mI_*^2)^2} + (d+\gamma)^2 > 0$, $D_1 = B_1^2 - C_1^2 - 2B_0B_2 + 2C_0C_2$, and $D_0 = B_0^2 - C_0^2$.

Note that only the real positive value of m_1 in Eq. (13) can imply existence of $\omega \in \mathbb{R}/0$. Hence, we consider the cases when Eq. (13) has positive root(s)- unique or multiple (and distinct).

Also, since $D_2 > 0$, the Eq. (13) always has a simple positive real root $m_1 = \omega_1^2$ (say) when,

$$D_0 = B_0^2 - C_0^2 < 0. \tag{14}$$

So the characteristic equation $\phi(\iota\omega, \tau^*) = 0$ has a pair of purely imaginary roots $\pm\iota\omega_1$ when (14) is satisfied.

Define,

$$\tau^* = \tau_k^*(\omega) = \begin{cases} \frac{1}{\omega}[\cos^{-1}(g_1(\omega)) + 2\pi k], & \text{if } g_2(\omega) > 0, \\ \frac{1}{\omega}[2\pi(k+1) - \cos^{-1}(g_1(\omega))], & \text{if } g_2(\omega) < 0, \end{cases} \tag{15}$$

for all $k = 0, 1, 2, \ldots$, where $g_1(\omega) = \frac{C_1\omega(\omega^3 - B_1\omega) + (C_0 - C_2\omega^2)(B_2\omega^2 - B_0)}{(C_0 - C_2\omega^2)^2 + C_1^2\omega^2}$ and $g_2(\omega) = \frac{C_1\omega(B_2\omega^2 - B_0) - (\omega^3 - B_1\omega)(C_0 - C_2\omega^2)}{(C_0 - C_2\omega^2)^2 + C_1^2\omega^2}$.

The corresponding time delay values $\tau^* = \tau_{1_k}^*$ $(k = 0, 1, 2 \ldots)$ satisfying $\phi(\iota\omega_1, \tau^*) = 0$ are obtained using Eqs. (10) and (11) and given as follows,

$$\tau^* = \tau_{1_k}^*(\omega_1) = \begin{cases} \frac{1}{\omega_1}[\cos^{-1}(g_1(\omega_1)) + 2\pi k], & \text{if } g_2(\omega_1) > 0, \\ \frac{1}{\omega_1}[2\pi(k+1) - \cos^{-1}(g_1(\omega_1))], & \text{if } g_2(\omega_1) < 0, \end{cases} \tag{16}$$

for all $k = 0, 1, 2, \ldots$.

Lemma 3.1 *The characteristic equation (6), i.e., $\phi(\lambda, \tau) = 0$, has one conjugate pair of purely imaginary roots $\iota\omega_1$ when $\tau = \tau_{1_k}^*$ obtained from (16) for $k = 0, 1, \ldots$, provided condition (14) is true.*

Further, the Eq. (13) has two distinct positive real roots $m_{21} = \omega_{21}^2$, $m_{22} = \omega_{22}^2$ (say) if,

$$D_1 < 0,\ D_0 > 0,\ \ \Delta = 18D_2D_1D_0 - 4D_2^3D_0 + D_2^2D_1^2 - 4D_1^3 - 27D_0^2 > 0. \tag{17}$$

Thus the characteristic equation $\phi(\iota\omega, \tau) = 0$ (Eq. (6)) has two conjugate pairs of purely imaginary roots $\pm\iota\omega_{21}, \pm\iota\omega_{22}$ if conditions in (17) hold. The corresponding time delay values $\tau^* = \tau_{21_k}^*$ satisfy $\phi(\iota\omega_{21}, \tau_{21_k}^*) = 0$ and $\tau^* = \tau_{22_k}^*$ satisfy $\phi(\iota\omega_{22}, \tau_{22_k}^*) = 0$ for $k = 0, 1, 2, \ldots$, where $\tau_{21_k}^* = \tau_k^*(\omega_{21})$ and $\tau_{22_k}^* = \tau_k^*(\omega_{22})$ ($\tau_k^*(\omega)$ are defined in (15)).

Lemma 3.2 *The characteristic equation (6), i.e., $\phi(\lambda, \tau) = 0$, has two conjugate pairs of purely imaginary roots $\iota\omega_{21}, \iota\omega_{22}$ respectively when $\tau = \tau_{21_k}^*, \tau_{22_k}^*$ obtained from (15) for $k = 0, 1, \ldots$, if conditions in (17) are true.*

3.1 Existence of Hopf Bifurcation and Switch of Stability Property

It is known that the existence of purely imaginary characteristic roots may lead the system through Hopf bifurcation at the equilibrium point. So according to Lemmas 3.1 and 3.2, system (1) may exhibit Hopf bifurcation at the endemic equilibrium point for non zero time delay i.e., for $\tau = \tau_{1_k}^*, \tau_{21_k}^*, \tau_{22_k}^* > 0$, $k = 0, 1, 2, \ldots$. For system (1), there exists Hopf bifurcation at delay critical value $\tau_H = \tau_{1_k}^*, \tau_{21_k}^*, \tau_{22_k}^*$ provided,

(H_1) all roots of $\phi(\lambda, \tau_H) = 0$ are with negative real part except one pair of purely imaginary root $\iota\omega_H(\omega_H = \omega_1, \omega_{21}, \omega_{22})$, and

(H_2) $\left[Re\left(\frac{d\lambda}{d\tau}\right)^{-1}\right]\Big|_{\substack{\lambda=i\omega_H,\\ \tau=\tau_H}} \neq 0.$

Differentiating $\phi(\lambda, \tau) = 0$ with respect to τ we have,

$$\left(\frac{d\lambda}{d\tau}\right)^{-1} = \frac{3\lambda^2 + 2B_2\lambda + B_1}{\lambda e^{-\lambda\tau}(C_2\lambda^2 + C_1\lambda + C_0)} + \frac{2C_2\lambda + C_1}{\lambda(C_2\lambda^2 + C_1\lambda + C_0)} - \frac{\tau}{\lambda}.$$

So the transversality condition (H_2) becomes,

$$Re\left(\frac{d\lambda}{d\tau}\right)^{-1}\Big|_{\substack{\lambda=\iota\omega_H,\\ \tau=\tau_H}} = \frac{\xi'(\omega_H^2)}{(C_0 - C_2\omega_H^2)^2 + C_1^2\omega_H^2}. \tag{18}$$

Lemma 3.3 *Assume condition in (14) is satisfied and $\tau_{1_0}^*$ is the first positive critical value of τ. Then,*

(1) if all roots of characteristic equation (6), $\phi(\lambda, \tau) = 0$ are with negative real part for $\tau = 0$ (i.e., E_* is stable at $\tau = 0$), then all roots have negative real part for $0 < \tau < \tau_{1_0}^*$ and have positive real part for $\tau > \tau_{1_0}^*$. For $\tau = \tau_{1_0}^*$, all roots of (6) are with negative real part except one pair of purely imaginary root.

(2) If one root of characteristic equation (6), $\phi(\lambda, \tau) = 0$ is with positive real part for $\tau = 0$ (i.e., E_* is unstable at $\tau = 0$), then at least one root has positive real part for all $\tau > 0$.

Proof When (14) is satisfied, then we have $D_0 < 0$, $\xi(\infty) > 0$ and $m_1 = \omega_1^2$ is simple real positive root of $\xi(m_1) = 0$ (Eq. (13)), therefore $\xi'(\omega_1^2)$ is always positive. So from (18) we have,

$$Re\left(\frac{d\lambda}{d\tau}\right)^{-1}\Big|_{\substack{\lambda=\iota\omega_1,\\ \tau=\tau_{1_0}^*}} > 0, \quad \forall k = 0, 1, 2, \ldots. \tag{19}$$

i.e., the transversality direction (H_2) is positive in this case. So the purely imaginary roots for $\tau = \tau_{1_k}^*$ become complex with positive real part for $\tau > \tau_{1_k}^*, k = 0, 1, \ldots$. Hence the proof. □

Lemma 3.3 implies that $(H_1) - (H_2)$ are true for $\tau = \tau_{1_0}^*$ when E_* is stable at $\tau = 0$ and so $\tau_{1_0}^*$ is called the delay hopf critical value. Therefore, we can conclude the following theorem.

Theorem 3.1 *For the system (1), Hopf bifurcation occurrs with respect to time delay when $\tau = \tau_{1_0}^*$ and (8) holds. The solutions, converge to the endemic equilibrium point E_* for $\tau < \tau_{1_0}^*$ and bifurcate from the endemic equilibrium point E_* in periodic form for $\tau \geq \tau_{1_0}^*$.*

Lemma 3.4 *Assume conditions in (17) are satisfied and* $\tau^*_{21_0}, \tau^*_{22_0}$ *are the first two positive critical values of* τ *in* $\tau^*_{21_k}, \tau^*_{22_k}$ *for* $k = 0, 1, \ldots$.

(1) If all roots of characteristic equation (6), $\phi(\lambda, \tau) = 0$ *are with negative real part for* $\tau = 0$ *(i.e.,* E_* *is stable at* $\tau = 0$*), then all roots have negative real part for* $0 < \tau < \tau^*_{21_0}$, $\tau > \tau^*_{22_0}$ *and have positive real part for* $\tau \in (\tau^*_{21_0}, \tau^*_{22_0})$. *For* $\tau = \tau^*_{21_0}, \tau^*_{22_0}$, *all roots of (6) are with negative real part except one pair of pure imaginary root, provided* $\tau^*_{21_0} < \tau^*_{22_0}$.

(2) If one root of characteristic equation (6), $\phi(\lambda, \tau) = 0$ *is with positive real part for* $\tau = 0$ *(i.e.,* E_* *is unstable at* $\tau = 0$*), then all roots have positive real part for* $0 < \tau < \tau^*_{22_0}$, $\tau > \tau^*_{21_0}$ *and have negative real part for* $\tau \in (\tau^*_{22_0}, \tau^*_{21_0})$. *For* $\tau = \tau^*_{22}, \tau^*_{21}$, *all roots of (6) are with negative real part except one pair of pure imaginary root, provided* $\tau^*_{22_0} < \tau^*_{21_0}$.

Proof Again, since $D_0 > 0$, $\xi(\infty) > 0$ and $m_1 = m_{21} = \omega^2_{21}$, $m_1 = m_{22} = \omega^2_{22}$ are two simple real positive roots of $\xi(m_1) = 0$ (Eq. (13)) such that $\omega_{22} < \omega_{21}$, therefore $\xi'(\omega^2_{21})$ is positive and $\xi'(\omega^2_{22})$ is negative. So in this case, the transversality condition (H_2) becomes,

$$Re\left(\frac{d\lambda}{d\tau}\right)^{-1}\Big|_{\substack{\lambda=\iota\omega_{21},\\ \tau=\tau^*_{21_k}}} > 0, \quad Re\left(\frac{d\lambda}{d\tau}\right)^{-1}\Big|_{\substack{\lambda=\iota\omega_{22},\\ \tau=\tau^*_{22_k}}} < 0, \forall k = 0, 1, 2, \ldots. \tag{20}$$

Hence, the proof. □

Lemma 3.4 implies the conditions such that $(H_1) - (H_2)$ are hold for $\tau = \tau^*_{21_0}, \tau^*_{22_0}$ and so $\tau^*_{21_0}, \tau^*_{22_0}$ are called the delay hopf critical value. Therefore, we can conclude the following theorem.

Theorem 3.2 *For the system (1), Hopf bifurcation is occurred with respect to time delay when* $\tau = \tau^*_{21_0}, \tau^*_{22_0}$.

(1) The system solutions, converge to the endemic equilibrium point E_* *for* $0 < \tau < \tau^*_{21_0}$, $\tau > \tau^*_{22_0}$ *and bifurcate from the endemic equilibrium point* E_* *in periodic form for* $\in [\tau^*_{21_0}, \tau^*_{22_0}]$ *when (8) holds.*

(2) The system solutions, converge to the endemic equilibrium point E_* *for* $\in (\tau^*_{22_0}, \tau^*_{21_0})$ *and bifurcate from the endemic equilibrium point* E_* *in periodic form for* $0 < \tau < \tau^*_{22_0}$, $\tau > \tau^*_{21_0}$ *when (8) does not hold.*

Now we will illustrate the local stability criterion of disease free equilibrium points $E_0(0, 0, 0)$ and $E_1(K, 0, 0)$.

Theorem 3.3 *The disease free equilibrium point* E_0 *is unstable for all* $\tau \geq 0$.

Proof It can be verified easily. □

Theorem 3.4 *The disease free equilibrium point* E_1 *is locally asymptotically stable for* $\mathcal{R}_0 < 1$ *and is unstable for* $\mathcal{R}_0 > 1$ *for all* $\tau \geq 0$.

Proof The characteristic equation (6) at $E_1(K, 0, 0)$ becomes

$$\phi_0(\lambda, \tau) := \lambda^3 + B_{20}\lambda^2 + B_{10}\lambda + B_{00} + (C_{20}\lambda^2 + C_{10}\lambda + C_{00})e^{-\lambda\tau} = 0, \quad (21)$$

where $B_{20} = 2d + \delta + \alpha + \gamma + b$, $B_{10} = (d + \delta + \alpha)(b + d + \gamma) + b(d + \gamma)$, $B_{00} = b(d + \delta + \alpha)(d + \gamma)$, $C_{20} = -\beta K$, $C_{10} = -(b + d + \gamma)\beta K$, $C_{00} = -b(d + \gamma)\beta K$.

Equation (21) is a cubic polynomial in λ for $\tau = 0$. Thus, following Routh-Hurwitz criterion we can say that all roots of Eq. (21) are with negative real part for $\mathcal{R}_0 < 1$ as

$$(B_{20} + C_{20})(B_{10} + C_{10}) - (B_{00} + C_{00}) = (b + d + \gamma)\{(b(d + \gamma) + (b + d + \gamma)$$
$$(d + \delta + \alpha)(1 - \mathcal{R}_0)) + (d + \delta + \alpha)^2(1 - \mathcal{R}_0)^2\} > 0,$$
$$B_{20} + C_{20} = b + d + \gamma + (d + \delta + \alpha)(1 - \mathcal{R}_0) > 0,$$
$$B_{00} + C_{00} = b(d + \gamma)(d + \delta + \alpha)(1 - \mathcal{R}_0) > 0,$$

and one root of Eq. (21) is positive for $\mathcal{R}_0 > 1$ as $(B_{00} + C_{00})|_{\mathcal{R}_0 > 1} < 0$. Therefore, E_1 is locally asymptotically stable for $\mathcal{R}_0 < 1$ and is unstable for $\mathcal{R}_0 > 1$ when $\tau = 0$.

Equation (21) becomes a transcendental equation of λ when $\tau > 0$. Then proceeding similar calculations as was done for Eqs. (9)–(12), at E_1 Eq. (13) takes the following form,

$$\xi_0(m_1) := m_1^3 + D_{20}m_1^2 + D_{10}m_1 + D_{00} = 0, \quad (22)$$

where $D_{20} = B_{20}^2 - C_{20}^2 - 2B_{10} = b^2 + (d + \gamma)^2 + (d + \delta + \alpha)^2(1 - \mathcal{R}_0^2)$,
$D_{10} = B_{10}^2 - C_{10}^2 - 2B_{00}B_2 + 2C_{00}C_{20} = b^2(d + \gamma)^2 + 2(b + d + \gamma)\beta^2K^2 + (b + d + \gamma)^2(d + \delta + \alpha)^2(1 - \mathcal{R}_0^2)$,
and $D_{00} = B_{00}^2 - C_{00}^2 = b^2(d + \gamma)^2(d + \delta + \alpha)^2(1 - \mathcal{R}_0^2)$.
When $\mathcal{R}_0 < 1$, we have $D_{20}, D_{10}, D_{00} > 0$. So there does not exist any real positive root of Eq. (14) and so any purely imaginary root of Eq. (21) does not exist for any positive time delay. Therefore if E_1 is locally stable for $\tau = 0$, remains stable for all $\tau > 0$ when $\mathcal{R}_0 < 1$.
When $\mathcal{R}_0 > 1$ then $D_{00} < 0$ and so purely imaginary roots $\iota\omega_0$ exist at $\tau = \tau_0$. Since $Re\left(\frac{d\lambda}{d\tau}\right)^{-1}\Big|_{\substack{\lambda=\iota\omega_0,\\ \tau=\tau_0}} = \frac{\xi_0'(\omega_0^2)}{(C_{00} - C_{20}\omega_0^2)^2 + C_{10}^2\omega_0^2} > 0$, so similar to Lemma 3.3, E_1 remains unstable for all $\tau > 0$ as it is unstable for $\tau = 0$ when $\mathcal{R}_0 > 1$. Hence the proof. □

4 Numerical Observation

In this section, we exhibit the effect of incubation time delay on the stability properties of model system (1). We have discussed in Theorem 3.1 that endemic equilibrium point switches its stability property with varying time delay and consequently the dynamics of system will change and consequently a Hopf bifurcation appears. Therefore, for the purpose of exhibiting stability switch of equilibrium, we vary the time delay $\tau \in (0, 20]$ and fix the others parameters at $K = 1000$, $b = 0.0044$, $d = 0.003$, $\delta = 0.003$, $\beta = 0.001$, $\alpha = 0.26$, $m = 0.0001$, $\gamma = 0.0001$.

Continuing with these parametric value, we obtain basic reproduction number $\mathcal{R}_0 = 3.76 > 1$. The coefficients of Eq. (2) become $A_2 = -5.5 \times 10^{-8} < 0$, $A_1 = 0.00026 > 0$, $A_0 = -0.00085 < 0$ and we find unique endemic equilibrium point $E_* = (266.3, 3.34, 279.89)$ (following Table 1) for this selected parametric set of values. Endemic equilibrium E_* satisfies condition (8) when $\tau = 0$ as we have,
$(B_2 + C_2)(B_1 + C_1) - (B_0 + C_0) = 1.77 \times 10^{-6} > 0$, $B_2 + C_2 = 0.005 > 0$, $B_0 + C_0 = 2.7 \times 10^{-6} > 0$. Therefore E_* is locally asymptotically stable for $\tau = 0$. Again E_* satisfies condition in (14) when $\tau > 0$ as we get $D_0 = -1.5 \times 10^{-12} < 0$. Hence one pair of purely imaginary root $\pm\iota\omega_1 = \iota 0.016$ is obtained when $\tau = 9.8 = \tau^*_{1_0}$ (evaluated using (15)) according to the Lemma 3.1. The transversality condition (19) also holds true at E_* as $Re\left(\frac{d\lambda}{d\tau}\right)^{-1}\Big|_{\substack{\lambda=\iota 0.016,\\ \tau=9.8}} = 46.2$ is positive. So from Lemma 3.3, $\tau = \tau^*_{1_0} = 9.8$ is the critical value of delay for appearance of Hopf bifurcation in this case. Therefore Theorem 3.1 is valid which assures the existence of Hopf bifurcation at $\tau = \tau^*_{1_0} = 9.8$ and indicates the appearance of two different dynamical behavior for different values of τ, i.e., for $\tau < \tau^*_{1_0}$ and for $\tau \geq \tau^*_{1_0}$. When time delay is smaller than the critical value of delay (i.e., $\tau = 4 < 9.8$), then solution trajectories initiated from any population level always converge to the endemic equilibrium point E_*, as shown in Fig. 1(left). When time delay is larger than the critical value of delay (i.e., $\tau = 17 \geq 9.8$), then solution trajectories initiated from any population level converge to a closed periodic orbit around E_*, as shown in Fig. 1(right), thus showing oscillatory solutions. Figure 2 shows the time series plots for the $S(t)$, $I(t)$ and $R(t)$ populations with $\tau = 4$ (left) and $\tau = 17$ (right).

The closed periodic orbit generated at this delay length is seen in Fig. 3(left). Figure 3(right) depicts the Hopf bifurcation diagram with respect to the incubation delay bifurcation parameter. In this figure, the plane $\tau = \tau^*_{1_0}$ indicates the bifurcation value and before this critical value E_* is locally stable and after this delay critical value E_* is unstable and starts showing oscillatory solutions. This instability property of equilibrium will be continued for all $\tau > \tau^*_{1_0} = 9.8$. From Fig. 3(right) it is noticed that the period of the closed periodic orbits generated for $\tau \geq \tau^*_{1_0} = 9.8$ is increasing if the incubation delay length is increased after

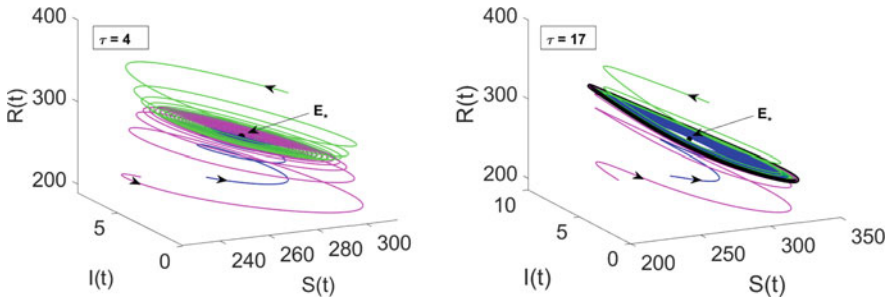


Fig. 1 Phase diagram of system (1) at $\tau = 4 < \tau^*_{1_0} = 9.8$ depicted stable E_* (left). Phase diagram of system (1) at $\tau = 17 > \tau^*_{1_0} = 9.8$ depicted unstable E_* (right)

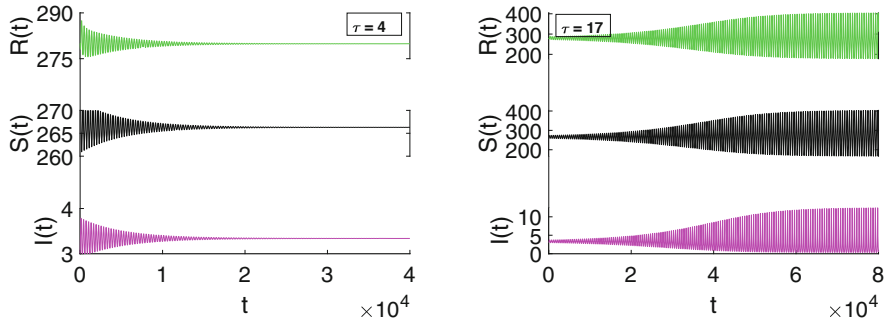


Fig. 2 Time series plots for the susceptible, infective, recovered populations converging to S_*, I_*, R_* respectively for $\tau = 4 < \tau^*_{1_0} = 9.8$ (left). Time series plot of all population showing oscillatory solutions for $\tau = 17 > \tau^*_{1_0} = 9.8$ (right)

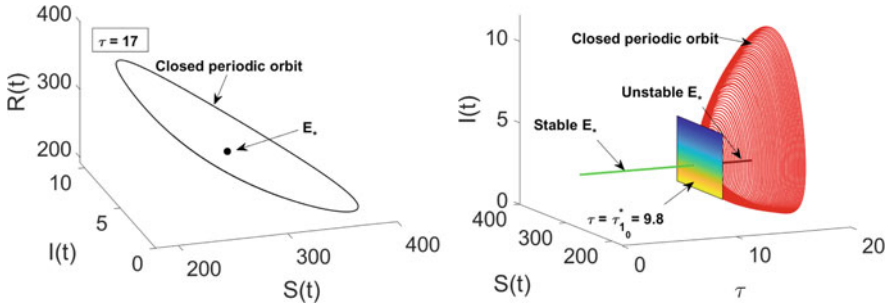


Fig. 3 Phase diagram of system (1) at $\tau = 17 > \tau^*_{1_0} = 9.8$ showing the closed periodic orbit (left). Hopf bifurcation diagram with respect to time delay τ depicting stability property switch of E_* from stable (green color) to unstable (black color) at $\tau = \tau^*_{1_0} = 9.8$ and the appearance of closed periodic orbits (red curves) bifurcated from unstable E_* after $\tau = \tau^*_{1_0} = 9.8$ (right)

$\tau = \tau^*_{1_0} = 9.8$. Therefore the fluctuation range of infective level will increase when incubation delay length is further increased after $\tau = \tau^*_{1_0} = 9.8$, as depicted in Fig. 4.

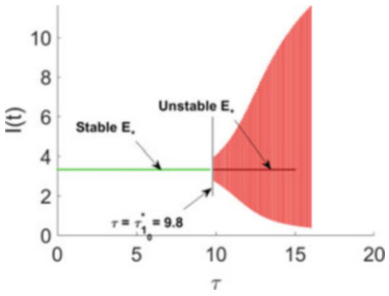


Fig. 4 Bifurcation plot depicting the fluctuation range of infective cases after $\tau > \tau^*_{1_0} = 9.8$

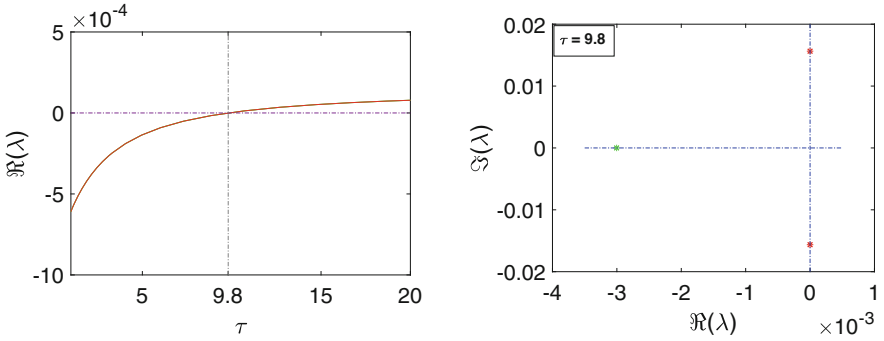


Fig. 5 Plot of maximum real part of the characteristic roots of Eq. (6) for $\tau \in [0, 20]$ (left). Plot of pair of purely imaginary characteristic roots of Eq. (6) for $\tau = \tau^*_{1_0} = 9.8$ (right)

We have discussed earlier in Sect. 3 that the stability property of endemic equilibrium point will switch due to Hopf bifurcation in system as the time delay parameter changes. This happens when the real part of characteristic roots (Eq. (6)) crosses the imaginary axis for varying time delay. Using the same parametric values as above, in Fig. 5(left) we plot the maximum real part of the characteristic roots of Eq. (6) and note that it crosses the imaginary axis at $\tau = 9.8$ and so the maximum real part of root for $\tau < 9.8$ becomes negative to positive for $\tau > 9.8$ and in Fig. 5(right) we see the purely imaginary characteristic roots ($\iota 0.016$) of Eq. (6) obtained for $\tau = 9.8$. Hence Fig. 5 assures the stability switch phenomenon of E_* depicted in the bifurcation diagram (Fig. 3)(right).

Now we discuss the effect of behavioral response m and disease transmission rate β on stability behaviour of system (1). For this purpose we take the time delay $\tau \in (0, 30)$ and fix others parameters at $K = 1000$, $b = 0.0044$, $d = 0.003$, $\delta = 0.003$, $\alpha = 0.26$, $\gamma = 0.0001$.

If m is varied within $(0.0001, 0.0005)$ and $\beta = 0.001$, then we obtain $\mathcal{R}_0 = 3.76 > 1$ and $A_2 \in (-2.74 \times 10^{-7}, -5.5 \times 10^{-8})$, $A_1 \in (0.00025, 0.00026)$, and $A_0 \in (-0.00084, -0.00086)$. A unique endemic equilibrium point E_* exists for all $m \in (0.0001, 0.0005)$. This E_* satisfies condition (8) for $\tau = 0$ because we get $(B_2 + C_2)(B_1 + C_1) - (B_0 + C_0) \in (1.77 \times 10^{-6}, 3.9 \times 10^{-6})$, $B_2 + C_2 \in (0.005, 0.0073)$, $B_0 + C_0 \in (2.6 \times 10^{-6}, 2.7 \times 10^{-6})$ and satisfies (14) for $\tau > 0$

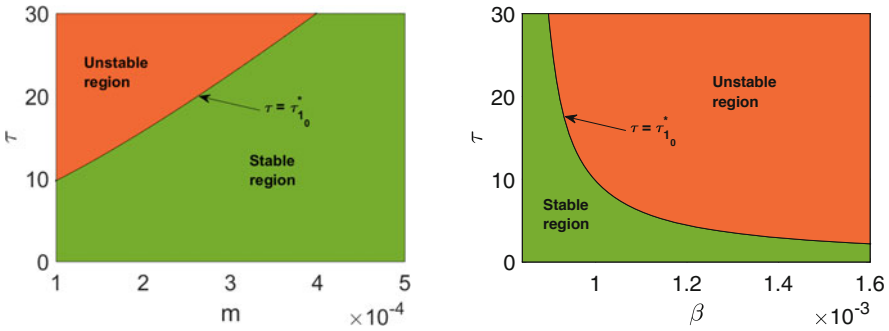


Fig. 6 Bifurcation plot in $\tau - m$ plane for $\tau \in (0, 30)$, $m \in (0.0001, 0.0005)$ depicted stability switch of E_* (left). Bifurcation plot in $\tau - \beta$ plane for $\tau \in (0, 30)$, $\beta \in (0.00085, 0.0016)$ depicted stability switch of E_* (right). The stability property of E_* is stable in the green color region and is unstable in the red color region

because we get $D_0 \in (-1.5 \times 10^{-12}, -1.3 \times 10^{-12})$ for $m \in (0.0001, 0.0005)$. Therefore critical value of delay parameter $\tau^*_{1_0} \in (9.8, 307)$ are obtained following Theorem 3.1 for $m \in (0.0001, 0.0005)$, and plotted in Fig. 6(left) by the black intersection curve of the red and green color region for the range of $\tau \in (0, 30)$.

Further, for $\beta \in (0.00084, 0.0016)$ and $m = 0.0001$, we obtain $\mathcal{R}_0 \in (3.1, 6.1)$ and $A_2 \in (-1.25 \times 10^{-7}, -3 \times 10^{-8})$, $A_1 \in (0.00018, 0.00066)$, and $A_0 \in (-0.0016, -0.0006)$. The unique endemic equilibrium point E_* is obtained for this range of $\beta \in (0.00084, 0.0016)$. Again, E_* satisfies condition (8) for $\tau = 0$ because we get $(B_2 + C_2)(B_1 + C_1) - (B_0 + C_0) \in (1.2 \times 10^{-6}, 2 \times 10^{-6})$, $B_2 + C_2 \in (0.0008, 0.002)$, $B_0 + C_0 \in (2.4 \times 10^{-6}, 3.1 \times 10^{-6})$ and satisfies condition in (14) for $\tau > 0$ because we get $D_0 \in (-4.8 \times 10^{-12}, -2.9 \times 10^{-14})$ for $\beta \in (0.00084, 0.0016)$. Therefore critical value of delay parameter $\tau^*_{1_0} \in (2, 1636)$ are obtained following Theorem 3.1 for $\beta \in (0.00084, 0.0016)$, and plotted in Fig. 6(right) by the black intersection curve of the red and green color region for the range of $\tau \in (0, 30)$.

Furthermore, Theorem 3.1 implies that E_* is locally asymptotically stable for $\tau < \tau^*_{1_0}$ and is unstable for $\tau \geq \tau^*_{1_0}$. Therefore in Fig. 6, E_* is stable in the green color region and is unstable in the red color region. It is noticed from Fig. 6(left) that higher value of behavioral response parameter m increases the chance so that system remains in stable mode for larger incubation time delay. Also, from Fig. 6(right), we note that the higher disease transmission rate β reduces this chance.

5 Conclusion

To understand the impact of incubation delay length on the model dynamics we considered a delayed $SIRS$ model by introducing the incubation period as delay parameter. We chose a simplified Monod Haldane type information induced

incidence function. We analysed the model and observed that stability of disease free equilibrium point is independent of incubation delay length. It is locally stable if basic reproduction number ($\mathcal{R}_0$) is less than unity and is unstable if ($\mathcal{R}_0$) is greater than unity. The unique endemic equilibrium (when it exists) for $\mathcal{R}_0 > 1$ is observed to be stable below the threshold value of the delay parameter and once the delay crosses it, there exists Hopf-bifurcation. Thus oscillations and stability loss are obtained in presence of incubation delay. We also note that system may remain stable for higher delay values if behaviour response parameter is large or disease transmission rate is small. Therefore, the stability of the system may be impacted by the incubation delay and will depend on its length.

Acknowledgments PKS acknowledges the financial support from SERB (DST) project no MTR/2017/000803.

References

1. V. Capasso and G. Serio *Mathematical biosciences*, A generalization of the Kermack-McKendrick deterministic epidemic model, **42**, 43–61, (1978).
2. P. V. Driessche and J. Watmough *Mathematical Biosciences*, Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission, **180**,29–48 (2002).
3. T. Das and P. K. Srivastava, *Chaos: An Interdisciplinary Journal of Nonlinear Science*, Nonlinear dynamical behavior of an SEIR mathematical model: Effect of information and saturated treatment, **31**, 043104 (2021).
4. R. Xu and Z. Ma, *Chaos, Solitons & Fractals*, Stability of a delayed SIRS epidemic model with a nonlinear incidence rate, **41**, 2319–2325 (2009).
5. Y. Liu and J. Cui, *International Journal of Biomathematics*, The impact of media coverage on the dynamics of infectious disease **01**, 65–74 (2008).
6. J. Cui, Y. S and H. Zhu, *Journal of Dynamics and Differential Equations* , The impact of media on the control of infectious diseases, **20**, 31–53 (2008).
7. Y. Yang and D. Xiao, *Discrete & Continuous Dynamical Systems-B*, Influence of latent period and nonlinear incidence rate on the dynamics of SIRS epidemiological models, **13**, 195 (2010).
8. A. Yadav and P. K. Srivastava, *Journal of Applied Mathematics and Computing*, The impact of information and saturated treatment with time delay in an infectious disease model, **66**, 277–305 (2021).
9. M. Li and X. Liu, *Discrete Dynamics in Nature and Society*, A delayed epidemic model with pulse vaccination, **2008**, (2008).
10. G. Huang, Y. Takeuchi, W. Ma and D Wei, *Bulletin of mathematical biology*, Global stability for delay SIR and SEIR epidemic models with nonlinear incidence rate, **72**, 1192–1207 (2010).
11. A. Kumar, K. Goel and others, *Theory in Biosciences*, A deterministic time-delayed SIR epidemic model: mathematical modeling and analysis, **139**, 67–76 (2020).
12. M. Li and X. Liu, *Abstract and Applied Analysis*, An SIR epidemic model with time delay and general nonlinear incidence rate, **2014**, (2014).
13. R Xu, *Journal of Mathematical Analysis and Applications*, Global stability of an HIV-1 infection model with saturation infection and intracellular delay, **375**, 75–81 (2011).
14. D. Xiao and S. Ruan, *Mathematical biosciences*, Global analysis of an epidemic model with nonmonotone incidence rate, **208**, 419–429 (2007).

15. Y. Enatsu, E. Messina, Y. Muroya, Y. Nakata, E. Russo, and Vecchio, Antonia, *Applied Mathematics and Computation*, Stability analysis of delayed SIR epidemic models with a class of nonlinear incidence rates, **218**, 5327–5336 (2012).
16. C. Yan and J. Jia, *Abstract and Applied Analysis*, Hopf bifurcation of a delayed epidemic model with information variable and limited medical resources, **2014**, (2014).
17. L. Liu, *Advances in Difference Equations*, A delayed SIR model with general nonlinear incidence rate, **2015**,1–11 (2015).
18. X. Liu, S. Wang, S. Liu and J. Li, *Mathematical Biosciences and Engineering*, An SEI infection model incorporating media impact, **14**, 1317–1335 (2017).
19. A. Kumar and Nilam, *Journal of Engineering Mathematics*, Mathematical analysis of a delayed epidemic model with nonlinear incidence and treatment rates, **115**, 1–20 (2019).

Modelling the Role of TV and Internet Coverage on Mitigating the Spread of Infectious Diseases

Rakesh Medda, Samares Pal, and Joydeb Bhattacharyya

1 Introduction

Infectious diseases are spreading throughout the world at an alarming rate which becomes a global concern. People have health issues all around the world, especially in underdeveloped countries [1]. Rural residents are in danger of death owing to poor health care [2, 30]. In India, public hospitals struggle to treat serious illnesses due to poor medical facilities [5]. Mathematical models are commonly recognized as one of the most effective techniques for forecasting the modes of transmission of many infectious diseases. In the event of a severe pandemic, when health care systems are unable to protect people from newly developing viruses, spreading first-hand knowledge on preventive measures via TV, social media, and the internet is an easy, quick, and cost-effective strategy to minimize disease burden [7, 21–23].

Media coverage through TV and internet has a major role and it is the most authoritative source of information. Because of this, it influences governments' health care engagement and thus has a significant influence on the control of the epidemic. By educating the public through the TV and internet coverage [4], people can reduce the spread of diseases by taking simple precautions such as social distancing [28], wearing face masks [13, 16], quarantine [10], vaccination [6], etc. The media is a major player in infectious disease outbreaks because of this.

During the last few decades, various studies of the efficacy of media efforts to suppress infectious diseases have been conducted [8, 11, 17, 18, 20, 25, 27]. The majority of these research assume that media will help minimize disease

R. Medda · S. Pal (✉)
Department of Mathematics, University of Kalyani, Nadia, West Bengal, India

J. Bhattacharyya
Department of Mathematics, Karimpur Pannadevi College, Nadia, West Bengal, India

R. P. Mondaini (ed.), *Trends in Biomathematics: Stability and Oscillations in Environmental, Social, and Biological Models*,
https://doi.org/10.1007/978-3-031-12515-7_21

transmission. Liu and Cui [18], for instance, provided a model in which the contact rate was modelled as a decreasing function of infectious individuals. Kiss et al. [12] proposed a model in which all conscious persons are classified as receptive or indifferent to media campaigning. They demonstrated that disseminating awareness information affected infection prevention. Funk et al. [8] established a SIRS model for spreading awareness during an epidemic outbreak and discovered that awareness campaigns can slow disease spread. Misra et al. [19] investigated the effect of awareness efforts on emerging diseases by assuming that media campaigns create an isolated aware class that is immune to infection. They determined that, while the disease cannot be completely removed from the population, it may be managed. Additionally, Misra et al. [20] investigated a paradigm in which susceptible classes become aware classes as a result of a Holling type-II functional reaction between the susceptible and the awareness programmes. In this model, delaying the execution of awareness programmes results in periodic solutions via Hopf-bifurcation. Samanta et al. [27] discussed a scenario in which the aware class is less likely to contract the illness than the non-conscious class. They discovered that increasing the growth rate of a media campaign results in a decrease in infected individuals, but only up to a certain point. Notably, the immigration rate is critical for regulating the system's dynamics. Some mathematical models are provided for certain infectious diseases to examine the influence of time delay in implementing awareness initiatives [9, 24] and It is noticed that by including time delay in the modelling process, the system becomes destabilized. Kumar et al. [14] recently developed an optimal control problem to examine the effect of information on illness prevalence. The authors assumed that information grows at a pace proportional to the saturation function of infected people. Agaba et al. [3] used a SIRS model to examine the effect on the dynamics of infectious diseases of the propagation of private awareness as a result of direct interaction between unaware and aware persons and public awareness as a result of information campaigns. Their findings indicate that both public and individual awareness have the potential to prevent illness spread.

The purpose of this study is to examine the role of TV and internet coverage on the occurrence of epidemic breakouts. In this current study, we consider media as a combination of TV and internet coverage. To begin, we determine how much media attention is necessary to mitigate the epidemic's impact. The model is extensively analyzed using mathematical and numerical techniques to handle these epidemiological difficulties.

In the rest of the paper: In the next section, a mathematical model is created to assess the role of TV and internet broadcasting on infectious disease. We investigate the model to show some basic properties such as positivity and boundedness of the system. We found disease-free and endemic equilibria and explored their local stability features. The fundamental reproduction number is stated. The existence of Transcritical and Hopf-bifurcation has been shown in subsequent sections. Numerical simulations are used to depict the effects of various epidemiologically significant parameters on the epidemic curve and gain insight into the proposed models' dynamics. In the last section, we conclude the paper with noteworthy findings.

2 Model Formulation

Let $N(t)$ be the total number of human populations in a particular region at time $t > 0$. The overall human population $N(t)$ is separated into three mutually exclusive classes: those who are prone to infection i.e., the susceptible population $X(t)$, those who are already infected i.e., the infected population $Y(t)$, and those who are aware of the infection i.e., the aware population $X_a(t)$. For our current study, we're going to assume that no new people currently infected with the disease will enter the investigated region at any time in the future; thus, we will use A to represent that constant influx of new individuals without prior knowledge of the ongoing epidemic.

We assume that the human population is evenly distributed and the disease is transmitted exclusively through direct contact between individuals who are infected and those who are susceptible to it at a rate of β. Aware individuals don't just mean people who are aware of the ongoing epidemics; rather, these individuals are those who are knowledgeable about the disease prevention and management process. The aware population may lose their awareness after a while by returning to the unaware class i.e. susceptible class at a rate $\lambda_0 X_a$. It is possible for those infected to recover from the disease by receiving appropriate treatment, or they can recover by the body's immune system going into action. Following recovery, a fraction σ of the recovered class γY will join the aware class as a result of previous exposure to disease while the remaining fraction $(1 - \sigma)$ will join the unaware class due to ignorance about the ongoing epidemic.

In addition, let $M(t)$ be the cumulative density of media-initiated awareness campaigns driven by TV and internet in that area at time t. Most of the media available to the public are delivered via TV and the internet, from which people can learn information about disease prevention. Thus, to make awareness more easily accessible, we have focused on two media: TV and the internet. While awareness propagation can increase behavior alterations such as self-quarantine and social distance, people will use these strategies to reduce their overall susceptibility. People's attention is predominantly focused on TV coverage, so as the number of aware individuals grows, it is reasonable to assume that the growth rate of internet coverage will slow down. The growth of the combined media coverage due to both TV and internet is assumed to be proportional to the infected populations, while the growth rate of internet coverage decreases by a factor $f(X_a) = \theta \dfrac{X_a}{\omega + X_a}$. Thus, the net growth rate of media coverage due to internet will be $\mu_i \left(1 - \theta \dfrac{X_a}{\omega + X_a}\right)$. In our model, we have considered media coverage as a convex combination of TV and the internet to propagate the process of mitigating the transmission of infectious diseases. Additionally, let M_0 be the constant baseline amount of advertisements in the investigated region. The effectiveness of TV and internet coverage is substantially reduced over time because of the inefficiency and psychological barriers; thus, advertisements through TV and internet become less influential and resulting in media decline with a disappearing rate δ. Further, advertising on TV and the

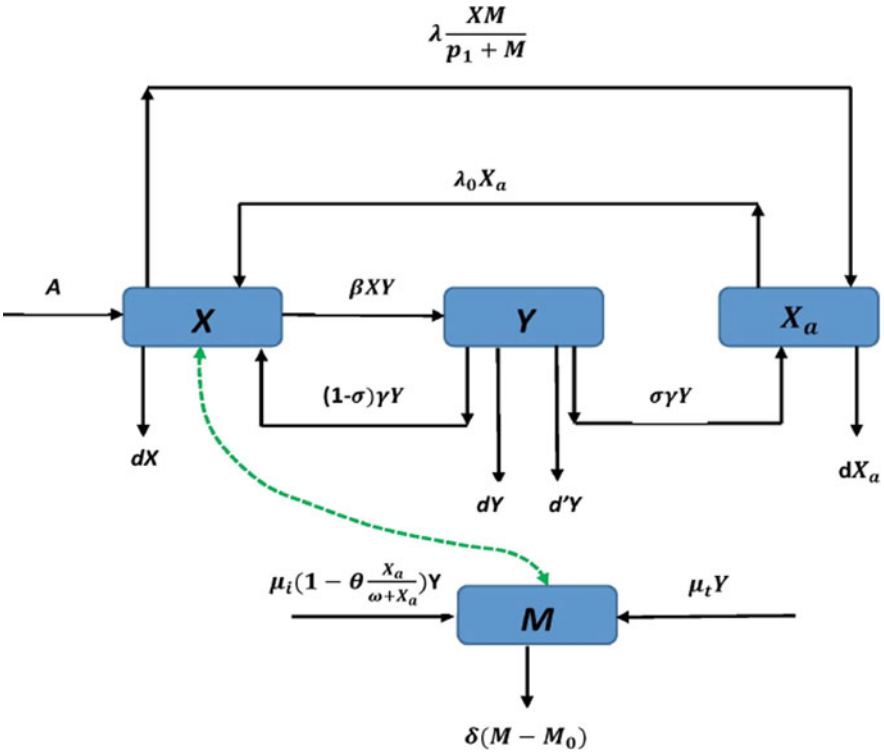


Fig. 1 Schematic diagram for the system (1)

internet is believed to have an impact on susceptible individuals and therefore it is believed that they form an isolated class known as the aware class to protect themselves from infection. Additionally, it is assumed that the advertisements on TV and internet have a limited effect on the susceptible population, and as a result, the susceptible population becomes aware at a rate of $\lambda\frac{XM}{p_1+M}$. The constant p_1 denotes the half-saturation point for the effect of TV and internet advertisements on unaware susceptible individuals, and it reaches half of its maximum value λX when the cumulative number of advertisements to aware the population reaches p_1. Infections are eradicated as a result of recovery, disease-related mortality, or natural death, at rates of γ, d' and d, respectively.

The dynamic variables within this study are diagrammatically represented in Fig. 1. Based on these model assumptions, the dynamics of the model is governed by the following system of non-linear differential equation:

$$\frac{dX}{dt} = A - \beta XY - \lambda\frac{XM}{p_1+M} + (1-\sigma)\gamma Y + \lambda_0 X_a - dX,$$

$$\begin{aligned}
\frac{dY}{dt} &= \beta XY - (\gamma + d^{'} + d)Y, \\
\frac{dX_a}{dt} &= \lambda \frac{XM}{p_1 + M} + \sigma\gamma Y - \lambda_0 X_a - dX_a, \\
\frac{dM}{dt} &= \alpha_1 \mu_i \left(1 - \theta \frac{X_a}{\omega + X_a}\right) Y + \alpha_2 \mu_t Y - \delta(M - M_0),
\end{aligned} \tag{1}$$

where $X(0) > 0$, $Y(0) \geq 0$, $X_a(0) > 0$, and $M(0) \geq M_0$.
Using the fact that $X + Y + X_a = N$, the above system reduces to the following system:

$$\begin{aligned}
\frac{dY}{dt} &= \beta(N - Y - X_a)Y - (\gamma + d^{'} + d)Y, \\
\frac{dX_a}{dt} &= \lambda \frac{(N - Y - X_a)M}{p_1 + M} + \sigma\gamma Y - \lambda_0 X_a - dX_a, \\
\frac{dN}{dt} &= A - dN - d^{'}Y, \\
\frac{dM}{dt} &= \alpha_1 \mu_i \left(1 - \theta \frac{X_a}{\omega + X_a}\right) Y + \alpha_2 \mu_t Y - \delta(M - M_0),
\end{aligned} \tag{2}$$

where $Y(0) \geq 0$, $X_a(0) > 0$, $N(0) > 0$, and $M(0) \geq M_0$. Epidemiological descriptions of model parameters involved in system (2) and their values used for numerical simulations are listed in Table 1.

2.1 Basic Properties

We will prove that the values of all the model (1) variables remain non-negative throughout the entire time period. For this, we shall need to prove the following claim.

2.2 Positivity of Solutions

Lemma 2.1 *The solution $X(t)$, $Y(t)$, $X_a(t)$ and $M(t)$ of the model system (1) with initial conditions $X(0) > 0$, $Y(0) \geq 0$, $X_a(0) > 0$, $M(0) \geq M_0$ are positive for all $t > 0$.*

Proof First, we will prove the positivity of $Y(t)$ and $X_a(t)$ as positivity of $X(t)$ depends on positivity of $Y(t)$ and $X_a(t)$. From the second equation of the model system (1), we have

Table 1 Model Parameters of system (2)

Parameters	Descriptions	Values	References
A	Immigration in the susceptible population class	3	Assumed
β	Contact rate of susceptible with infected individuals	0.00002	[19]
λ	Dissemination rate of awareness among the susceptible individuals	0.02	Assumed
λ_0	Transfer rate of aware individuals to susceptible class	0.0024	Assumed
p_1	Half saturation constant	160	Assumed
σ	Fraction of recovered individuals joining the aware class	0.1	Assumed
ω	Half saturation constant	6000	[23]
γ	Recovery rate of infected individuals	0.32	Assumed
d'	Disease induced death rate	0.0001	Assumed
d	Natural death rate of human population	0.00004	[23]
μ_i	The growth rate of broadcasting the information through internet	0.08	Assumed
μ_t	The growth rate of broadcasting the information through TV	0.003	Assumed
δ	Diminution rate of TV and internet coverage due to inefficiency and physiological barrier	0.06	[20, 27]
M_0	Baseline number of media coverage through TV and internet	10	Assumed
θ	The decay of internet coverage with the increase in the number of individuals	0.0005	[23]
α_1 and α_2	Convex combination constant ($\alpha_1 + \alpha_2 = 1$)	$\alpha_1 = 0.5, \alpha_2 = 0.5$	Assumed

$$Y(t) = Y(0)exp\left(\int_0^t [\beta X(s) - (\gamma + d^{'} + d)]ds\right), Y(0) > 0.$$

Which implies that $Y(t) \geq 0$ for all time $t \geq 0$
Again, from the fourth equation of the model system (1) can be written as;

$$\frac{dM(t)}{dt}exp(\delta t) + \delta M exp(\delta t) = \left[\left\{\alpha_1\mu_i\left(1-\theta\frac{X_a}{\omega+X_a}\right)+\alpha_2\mu_t)Y(t)+\delta M_0\right\}\right]exp(\delta t)$$

And, thus

$$M(t) = M(0)exp(-\delta t) + \int_0^t exp(-\delta(t-s))\left[\left\{\alpha_1\mu_i\left(1-\theta\frac{X_a(s)}{\omega+X_a(s)}\right)+\alpha_2\mu_t Y(s)+\delta M_0\right\}\right]ds.$$

where , $M(0) > 0$. This implies that $M(t) > 0$ for all $t > 0$. Similarly, we can show that $X_a(t) > 0$ for all $t > 0$.

Now from the first equation of the model system (1), we have

$$\left(\frac{dX(t)}{dt} + f_1(t)X(t)\right)exp\left(\int_0^t f_1(s)ds\right) = f_2(t)exp\left(\int_0^t f_1(s)ds\right)$$

Where $f_1(s) = \beta Y(s) + \lambda\frac{M(s)}{p_1+M(s)} + d$ and $f_2(t) = A + (1-\sigma)\gamma Y(t) + \lambda_0 X_a(t)$

$$X(t) = X(0)exp\left(-\int_0^t f_2(s)ds\right) + \exp\left(-\int_0^t f_2(s)ds\right)\times\int_0^t f_3(s)exp\left(\int_0^s f_2(u)du\right)ds, X(0) > 0.$$

This implies that $X(t) > 0$ for all $t > 0$. Thus, the solutions $X(t), Y(t), X_a(t)$ and $M(t)$ of the model system (1) with initial conditions $X(0) > 0$, $Y(0) \geq 0$, $X_a(0) > 0$, and $M(0) \geq M_0$ are positive for all $t > 0$. This completes the proof. □

Now, it is sufficient to study the model system (2) in detail rather than system (1).

Lemma 2.2 *For the solutions of model system (2), The region of attraction is given by the set,*

$$\Omega = \{(Y, X_a, N, M) \in \mathbb{R}^4_+ : 0 \leq Y, X_a \leq N \leq \frac{A}{d}, 0 \leq M \leq \left(\frac{(\alpha_1\mu_t + \alpha_2\mu_r)A}{\delta d} + M_0\right) = M_r\}$$

and attracts all solutions initiating in the interior of the positive orthant.

Proof From the third equation of system (2), we have

$$\frac{dN}{dt} = A - dN - d^{'}Y \leq A - dN.$$

Using comparison theorem , we get

$$0 \leq N(t) \leq \frac{A}{d} + (N(0) - \frac{A}{d}e^{-dt} \leq A - dN.$$

Thus, as $t \to \infty$, $0 \leq N \leq \frac{A}{d}$. Hence $0 \leq N \leq \frac{A}{d}$ for any $t > 0$. Since $X = N - Y - X_a \geq 0$, therefore $0 \leq Y, X_a \leq N \leq \frac{A}{d}$.
Now, from the fourth equation of the system (2), we obtain

$$\frac{dM}{dt} = \alpha_1\mu_i(1 - \theta\frac{X_a}{\omega + X_a})Y + \alpha_2\mu_t Y - \delta(M - M_0) \leq (\alpha_1\mu_i + \alpha_2\mu_t)\frac{A}{d} + \delta M_0 - \delta M_0.$$

Applying the differential inequality theory [15], we get

$$\limsup_{t\to\infty} M(t) \leq \frac{(\alpha_1\mu_i + \alpha_2\mu_t)A}{d\delta} + M_0 = M_r(say).$$

Hence $0 \leq M(t) \leq M_r$ for all $t > 0$. □

3 Equilibrium Analysis

In this section, all possible steady state solutions of the system (2) are analyzed. The system (2) has the following two equilibria:

(i) Disease-free equilibrium (DFE) $E^0 = \left(0, \frac{A\lambda M_0}{d\{\lambda M_0 + (\lambda_0 + d)(p_1 + M_0)\}}, \frac{A}{d}, M_0\right)$,

(ii) Endemic equilibrium $E^* = (Y^*, X_a^*, N^*, M^*)$.

The disease-free equilibrium E^0 of the system (2) exists always and the conditions for feasibility of endemic equilibrium of the system (2) is discussed in later subsection.

3.1 Basic Reproduction Number

The basic reproduction number R_0 is a global index that public health organisations use to determine the severity of an epidemic outbreak. Epidemiologically, the basic reproduction number indicates the number of secondary cases produced by an infectious individual in a population composed entirely of susceptible individuals. Here, we will determine reproduction number R_0 of the system (2) using the next-generation operator method [29].
Let, $x = (Y, X_a, N, M)^T$, $\mathscr{F}(x)$ be the matrix of new infection and $\mathscr{V}(x)$ be the matrix of transition terms. Then, the model system (2) can be written as:

$$\frac{dx}{dt} = \mathscr{F}(x) - \mathscr{V}(x),$$

where $\mathscr{F}(x) = \begin{bmatrix} \beta Y(N - Y - X_a) \\ 0 \\ 0 \\ 0 \end{bmatrix}$, $\mathscr{V}(x) = \begin{bmatrix} (\gamma + d' + d)Y \\ V_2 \\ -A + dN + dY \\ V_4 \end{bmatrix}$,

$V_2 = -\lambda(N - Y - X_a)\dfrac{M}{p_1 + M} - \sigma\gamma Y + (\lambda_0 + d)X_a$ and $V_4 = -\alpha_1\mu_i(1 - \theta\dfrac{X_a}{\omega + X_a})Y - \alpha_2\mu_t + \delta(M - M_0)$

Now, The Jacobian of $\mathscr{F}(x)$ and $\mathscr{V}(x)$ at DFE $E^0 = \left(0, \frac{A\lambda M_0}{d\{\lambda M_0 + (\lambda_0 + d)(p_1 + M_0)\}}, \frac{A}{d}, M_0\right)$ are represented by

$$F = \begin{bmatrix} \dfrac{\beta A}{d} & \dfrac{(\lambda_0 + d)(p_1 + M_0)}{\lambda M_0 + (\lambda_0 + d)(p + M_0)} & 0\,0\,0 \\ 0 & 0\,0\,0 \\ 0 & 0\,0\,0 \\ 0 & 0\,0\,0 \end{bmatrix} \text{ and}$$

$$V = \begin{bmatrix} \gamma + d' + d & 0 & 0 & 0 \\ \dfrac{\lambda M_0}{p_1 + M_0} - \sigma\gamma & \dfrac{\lambda M_0}{p_1 + M_0} + (\lambda_0 + d) & -\dfrac{\lambda M_0}{p_1 + M_0} & -\dfrac{\lambda p_1(\frac{A}{d} - X_a^0)}{(p_1 + M_0)^2} \\ d' & 0 & d & 0 \\ -\alpha_1\mu_i(1 - \theta\dfrac{X_a^0}{\omega + X_a^0}) - \alpha_2\mu_t & 0 & 0 & \delta \end{bmatrix}$$

The basic reproduction number is given by $\mathcal{R}_0 = \rho(FV^{-1})$, where ρ is the spectral radius of the next-generation matrix (FV^{-1}). Thus, from the model (2), we obtain the expression for $\mathcal{R}_0$ as

$$\mathcal{R}_0 = \frac{\beta A}{d(\gamma + d' + d)}\left(\frac{(\lambda_0 + d)(p_1 + M_0)}{\lambda M_0 + (\lambda_0 + d)(p_1 + M_0)}\right). \tag{3}$$

3.2 *Feasibility of the Endemic Equilibrium*

For system (2), an endemic equilibrium is denoted by $E^* = (Y^*, X_a^*, N^*, M^*)$, the components of which are positive solutions to the system's equilibrium equations. From the third equation of equilibrium equations, we get

$$N - Y = \frac{A - (d + d')Y}{d} \tag{4}$$

Using (4), from first equilibrium equation, we get

$$X_a = \frac{\beta(A - (d + d')Y) - d(\gamma + d' + d)}{\beta d} \tag{5}$$

It is worth noting here that in order for X_a to be positive, we must have $Y < \frac{\beta A - d(\gamma + d' + d)}{\beta(d' + d)} = Y_b$ (say) also noted $Y_b > 0$ if $R_0 > 1$.
From Eq. (5),we have

$$H(Y) = \left(1 - \theta \frac{X_a}{\omega + X_a}\right) = \frac{\beta d\omega + (1-\theta)[\beta(A - (d' + d)Y) - d(\gamma + d' + d)]}{\beta d\omega + \beta(A - (d' + d)Y) - d(\gamma + d' + d)} \quad (6)$$

Differentiating with respect to Y, we obtain

$$H'(Y) = \frac{\theta\beta^2 d\omega(d' + d)}{[\beta d\omega + \beta(A - (d' + d)Y) - d(\gamma + d' + d)]^2} \quad (7)$$

Clearly, $H'(Y) > 0$. Further form fourth equilibrium equation of system (2), we have

$$\frac{M}{p_1 + M} = \frac{\delta M_0 + (\alpha_1 \mu_i H(Y) + \alpha_2 \mu_t)Y}{\delta(p_1 + M_0) + (\alpha_1 \mu_i H(Y) + \alpha_2 \mu_t)Y} \quad (8)$$

substituting (4), (5) and (8) in second equilibrium equation, we obtain a equation in Y as:

$$\begin{aligned} P(Y) =& \lambda \left(\frac{\delta M_0 + (\alpha_1 \mu_i H(Y) + \alpha_2 \mu_t)Y}{\delta(p_1 + M_0) + (\alpha_1 \mu_i H(Y) + \alpha_2 \mu_t)Y}\right)\left(\frac{\gamma + d' + d}{\beta}\right) + \sigma\gamma Y \\ & - (\lambda_0 + d)\left[\frac{\beta(A - (d' + d)Y) - d(\gamma + d' + d)}{\beta d}\right] \end{aligned} \quad (9)$$

From Eq. (9), we have the following:

(i) $\beta P(0) = \left[\lambda \frac{M_0}{p_1 + M_0}(\gamma + d' + d) - (\lambda_0 + d)\left(\frac{\beta A}{d} - (\gamma + d' + d)\right)\right] <$ 0 if $R_0 > 1$.

(ii) $\beta P(Y_b) = \left[\lambda \left(\frac{\delta M_0 + (\alpha_1 \mu_i + \alpha_2 \mu_t)Y_b}{\delta(p_1 + M_0) + (\alpha_1 \mu_i + \alpha_2 \mu_t)Y_b}\right)(\gamma + d' + d) + \beta\sigma\gamma Y_b\right] > 0.$

(iii) $P'(Y) > 0$ in $(0, Y_b)$.

Thus, there exists an unique positive root of $P(Y)$ in the interval $(0, Y_b)$, say $Y = Y^*$, if $R_0 > 1$. Now, using the value of this $Y = Y^*$ in Eqs. (4), (5), and (8), we get positive values of N^*, X_a^* and M^*. Therefore, the endemic equilibrium $E^* = (Y^*, X_a^*, N^*, M^*)$ is feasible in $(0, Y_b)$, provided that $R_0 > 1$.

4 Local Stability Analysis

Theorem 4.1

(*i*) *The Disease free equilibrium* E^0 *is always feasible and its stable when* $R_0 < 1$ *and unstable whenever* $R_0 > 1$.

(*ii*) *The endemic equilibrium* E^* *is feasible whenever* $R_0 > 1$ *and is locally asymptotically stable if* $F_1 > 0$, $F_2 > 0$, $F_3 > 0$, $F_4 > 0$, *and* $F_1F_2F_3 > F_3^2 + F_1^2F_4$, *where* F_i*'s* $(i = 1, \ldots, 4)$ *are defined in the proof.*

Proof The Jacobian matrix of system (2) is given by

$$J = \begin{pmatrix} j_{11} & -\beta Y & \beta Y & 0 \\ j_{21} & j_{22} & j_{23} & j_{24} \\ -d' & 0 & -d & 0 \\ j_{41} & j_{42} & 0 & -\delta \end{pmatrix},$$

where,

$$j_{11} = \beta(N - 2Y - X_a) - (\gamma + d' + d),\ j_{21} = -\frac{\lambda M}{p_1 + M} + \sigma\gamma$$

$$j_{22} = -\frac{\lambda M}{p_1 + M} - (\lambda_0 + d),\ j_{23} = \frac{\lambda M}{p_1 + M},\ j_{24} = \frac{\lambda p_1(N - Y - X_a)}{(p_1 + M)^2},$$

$$j_{41} = \alpha_1\mu_i\left(1 - \theta\frac{X_a}{\omega + X_a}\right) + \alpha_2\mu_t,\ j_{42} = -\frac{\alpha_1\mu_i\theta\omega Y}{(\omega + X_a)^2}$$

Now, at the disease-free equilibrium point E^0, we compute the Jacobian, which is given by

$$J_{E^0} = \begin{pmatrix} j_{11}^0 - (\gamma + d' + d) & 0 & 0 & 0 \\ j_{21}^0 & j_{22}^0 & j_{23}^0 & j_{24}^0 \\ -d' & 0 & -d & 0 \\ j_{41}^0 & 0 & 0 & -\delta \end{pmatrix}$$

where

$$j_{11}^0 = \frac{\beta A}{d}\left(\frac{(\lambda_0 + d)(p_1 + M_0)}{\lambda M_0 + (\lambda_0 + d)(p_1 + M_0)}\right),\ j_{21}^0 = -\frac{\lambda M_0}{p_1 + M_0} + \sigma\gamma$$

$$j_{22}^0 = -\frac{\lambda M_0}{p_1 + M_0} - (\lambda_0 + d),\ j_{23}^0 = \frac{\lambda M_0}{p_1 + M_0},\ j_{24}^0 = \frac{\lambda p_1 A}{d(p_1 + M_0)^2}\left(\frac{(\lambda_0 + d)(p_1 + M_0)}{\lambda M_0 + (\lambda_0 + d)(p_1 + M_0)}\right),$$

$$j_{41}^0 = \alpha_1\mu_i\left(1 - \theta\frac{X_a^0}{\omega + X_a^0}\right) + \alpha_2\mu_t$$

and $X_a^0 = \dfrac{A\lambda M_0}{d\{\lambda M_0 + (\lambda_0 + d)(p_1 + M_0)\}}$.

The Jacobian matrix evaluated at E^0 has four eigenvalues: $-\delta$, $-d$, $-\left(\frac{\lambda M_0}{p_1+M_0}+\lambda_0+d\right)$ and $(\gamma+d'+d)[R_0-1]$. It is clear that first three eigenvalues are negative and last one is dependent on the value of R_0. If, $R_0>1$ then fourth eigenvalue is positive, and so the DFE(E^0) is unstable (cf. Fig. 2). Furthermore, when $R_0<1$, all the eigenvalues are negative and so DFE(E^0) is stable (cf. Fig. 2).

Now, at the endemic equilibrium E^*, the Jacobian matrix is denoted by J_{E^*} and given by

$$J_{E^*}=\begin{pmatrix} -\beta Y^* & -\beta Y^* & \beta Y^* & 0 \\ -j_{21}^* & -j_{22}^* & j_{23}^* & j_{24}^* \\ -d' & 0 & -d & 0 \\ j_{41}^* & -j_{42}^* & 0 & -\delta \end{pmatrix}$$

where

$$j_{21}^*=\frac{\lambda M^*}{p_1+M^*}-\sigma\gamma,\ j_{22}^*=\frac{\lambda M^*}{p_1+M^*}+(\lambda_0+d),\ j_{23}^*=\frac{\lambda M^*}{p_1+M^*},\ j_{24}^*=\frac{\lambda p_1(N^*-Y^*-X_a^*)}{(p_1+M^*)^2},$$

$$j_{41}^*=\alpha_1\mu_i(1-\theta\frac{X_a^*}{\omega+X_a^*})+\alpha_2\mu_t,\ j_{42}^*=\frac{\alpha_1\mu_i\theta\omega Y^*}{(\omega+X_a^*)^2}.$$

Now, we use Routh-Hurwitz criterion to demonstrate the local stability of the endemic equilibrium E^*. The characteristic equation of the matrix J_{E^*} is obtained as:

$$D(\zeta)\equiv\zeta^4+F_1\zeta^3+F_2\zeta^2+F_3\zeta+F_4=0, \tag{10}$$

where

$$F_1=\frac{\lambda M^*}{p_1+M^*}+\lambda_0+2d+\delta+\beta Y^*,$$

$$F_2=d\delta+\left(\frac{\lambda M^*}{p_1+M^*}+\lambda_0+d\right)(d+\delta)+j_{24}^*j_{42}^*+\beta(\lambda_0+2d+d'+\sigma\gamma+\delta)Y^*,$$

$$F_3=d\delta\left(\frac{\lambda M^*}{p_1+M^*}+\lambda_0+d\right)+dj_{24}^*j_{42}^*+\beta\delta(d'+d)Y^*+\beta(d+\delta)(\lambda_0+d+\sigma\gamma)Y^*,$$
$$+\beta d'(\lambda_0+d)Y^*+\beta j_{24}^*(j_{41}^*+j_{42}^*)Y^*$$

$$F_4=\beta d\delta(\lambda_0+d+\sigma\gamma)Y^*+\beta d'\delta(\lambda_0+d)Y^*+\beta dj_{24}^*(j_{41}^*+j_{42}^*)Y^*+\beta d'j_{24}^*j_{42}^*Y^*.$$

Thus, by applying the Routh-Hurwitz criterion to the preceding characteristic equation (10), we can assert that all eigenvalues of the Jacobian matrix J_{E^*} will be lie in the complex plane's left-half if the following condition is satisfied:

$$F_i>0\ (i=1,\ldots,4)\text{ and }\psi=F_1F_2F_3-F_3^2-F_1^2F_4>0.$$

Hence, so long as the above conditions are satisfied, the endemic equilibrium E^* is locally asymptotically stable. □

5 Existence of Transcritical Bifurcation

As a result of the preceding theorem (4.1), we can see that as R_0 increases through 1, the stability of disease free equilibrium point E^0 changes from stable to unstable. As a result, we employ R_0 as the bifurcation parameter.

Theorem 5.1 *The system (2) undergoes transcritical bifurcation at the disease free equilibrium E_0 when the bifurcation parameter $R_0 = 1$*

Proof Let, $Z = (Y, X_a, N, M)$ and $f_W(Z, \beta)$ vector represented by

$$f_W(Z,\beta) = \begin{pmatrix} f_1(Z,\beta) \\ f_2(Z,\beta) \\ f_3(Z,\beta) \\ f_4(Z,\beta) \end{pmatrix} = \begin{pmatrix} \beta(N-Y-X_a)Y - (\gamma + d^{'} + d)Y \\ \lambda\frac{(N-Y-X_a)M}{p_1+M} + \sigma\gamma Y - \lambda_0 X_a - dX_a \\ A - dN - d^{'}Y \\ \alpha_1\mu_i\left(1-\theta\frac{X_a}{\omega+X_a}\right)Y + \alpha_2\mu_t Y - \delta(M-M_0) \end{pmatrix}$$

By differentiating the above function with respect to β, we obtain

$$D_\beta f_W = \begin{pmatrix} (N-Y-X_a)Y \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Now, at $R_0 = 1$, the Jacobian matrix of the system (2) around the disease free equilibrium is given by

$$J(E^0) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ j^0_{21} & j^0_{22} & j^0_{23} & j^0_{24} \\ -d' & 0 & -d & 0 \\ j^0_{41} & 0 & 0 & -\delta \end{pmatrix}$$

Here, $J(E^0)$ has a simple zero eigenvalue with corresponding left eigenvector $u = \begin{pmatrix} 1 & 0 & 0 & 0 \end{pmatrix}^T$ and right eigenvector $v = \begin{pmatrix} 1 & v_2 & v_3 & v_4 \end{pmatrix}^T$, where $v_2 = \dfrac{\frac{d'}{d}j^0_{23} - j^0_{21} - \frac{j^0_{41}j^0_{24}}{\delta}}{j^0_{22}}$, $v_3 = \dfrac{d'}{d}$ and $v_4 = \dfrac{j^0_{41}}{\delta}$.

Now, at $R_0 = 1$ we have the following,

$$\left(u^T D_\beta f_W\right)_{E^0} = 0,$$

and

$$\left(u^T (D_Z D_\beta f_W) v)\right)_{E^0} = \frac{A}{d}\left(\frac{(\lambda_0 + d)(p_1 + M_0)}{\lambda M_0 + (\lambda_0 + d)(p_1 + M_0)}\right) \neq 0,$$

and also

$$\begin{aligned}\left(u^T ((D_{ZZ} f_W)(v, v))\right)_{E^0} &= \left(u^T \sum (e_i v^T D_Z (D_Z f_i)^T v)\right)_{E^0} \\ &= -\frac{2d(\gamma + d' + d)}{A(\lambda_0 + d)}\left[\frac{d'}{d}(\lambda_0 + d) + \frac{j^0_{41} j^0_{24}}{\delta} + \sigma\gamma\right] \neq 0.\end{aligned}$$

Therefore, from Sotomayor's theorem [26], we can conclude that the system (2) undergoes a transcritical bifurcation at E^0 when R_0 crosses $R_0 = 1$ (cf. Fig. 2). □

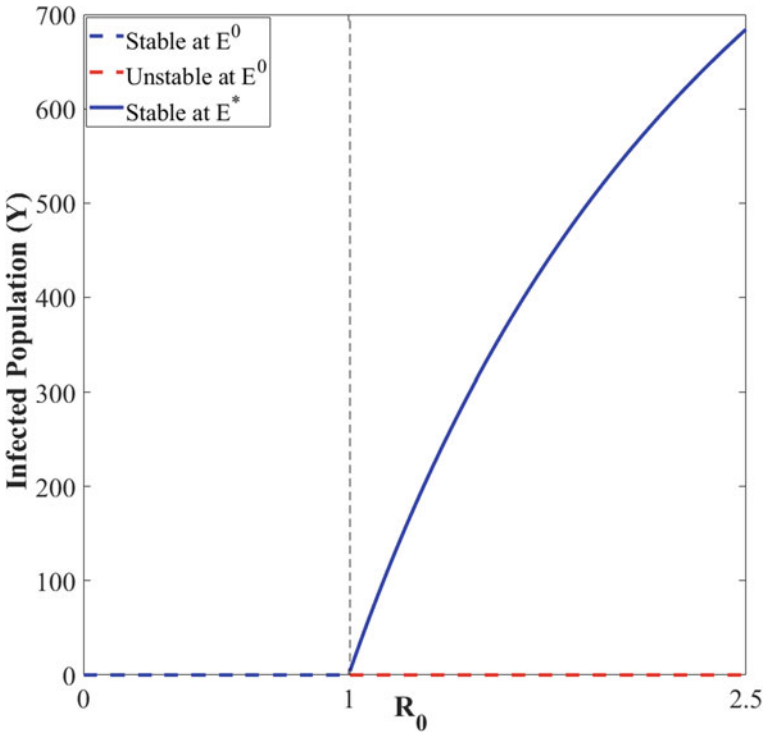


Fig. 2 One parameter bifurcation diagram with R_0 as a bifurcation parameter illustrating a transcritical bifurcation at $R_0 = 1$ when the transmission rate (β) crosses the critical value $\beta = 0.0000113$. For $R_0 < 1$, the system is stable at the DFE (E^0), whereas for $R_0 > 1$, the system is stable at the endemic equilibrium (E^*)

From Fig. 2, the condition $R_0 \leq 1$ ensures that the disease can be eradicated from the population for sufficiently low contact rate.

6 Existence of Hopf-Bifurcation

From Fig. 3 we observe that the system exhibits oscillations about the endemic equilibrium when α_1 crosses some critical value $\alpha_{1_c} \in (0, 0.4)$. To determine nature of bifurcation, we choose α_1 as a bifurcation parameter. For this, we express the coefficients in characteristic equation (10) as functions of α_1:

$$\zeta^4 + F_1(\alpha_1)\zeta^3 + F_2(\alpha_1)\zeta^2 + F_3(\alpha_1)\zeta + F_4(\alpha_1) = 0, \tag{11}$$

It is noted that all F_i's $(i = 1 - 4)$ are positive and let at $\alpha_1 = \alpha_{1_c}$ the following condition holds

$$(\psi(\alpha_1))_{\alpha_1=\alpha_{1_c}} = F_1(\alpha_{1_c})F_2(\alpha_{1_c})F_3(\alpha_{1_c}) - F_3^2(\alpha_{1_c}) - F_1^2(\alpha_{1_c})F_4(\alpha_{1_c}) = 0. \tag{12}$$

then at $\alpha_1 = \alpha_{1_c}$ the above Eq. (11) is written as:

$$\left(\zeta^2 + \frac{F_3}{F_1}\right)\left(\zeta^2 + F_1\zeta + \frac{F_1F_4}{F_3}\right) = 0 \tag{13}$$

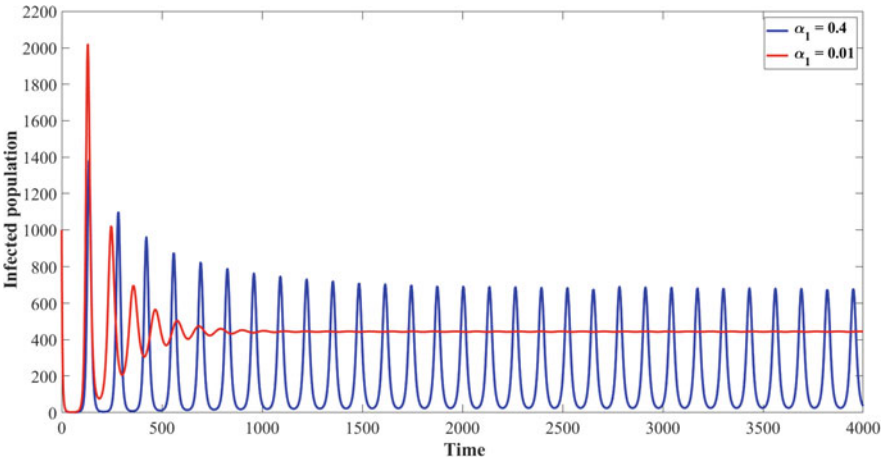


Fig. 3 Time evolution for the infected population density, exhibiting the existence of the stable endemic state at $\alpha_1 = 0.01$ (in red) and an oscillatory endemic state at $\alpha_1 = 0.4$ (in blue)

Let, ζ_i (i= 1, 2, 3, 4), are the roots of the above Eq. (13). It is clear that the Eq. (13) has two purely imaginary roots i.e. $\zeta_{1,2} = \pm i\omega_0$, where $\omega_0 = \sqrt{\frac{F_3}{F_1}}$. For Hopf-bifurcation to exist, all roots except $\pm i\omega_0$ must lie in the left half of the complex plane. To ascertain the nature of the remaining two roots (i.e., ζ_3, ζ_4), we have the following :

$$\zeta_3 + \zeta_4 = -F_1 \tag{14}$$

$$\omega_0^2 + \zeta_3\zeta_4 = F_2 \tag{15}$$

$$\omega_0^2(\zeta_3 + \zeta_4) = -F_3 \tag{16}$$

$$\omega_0^2\zeta_3\zeta_4 = F_4 \tag{17}$$

Now, first we check if ζ_3, ζ_4 are real, then from Eq. (17) we get ζ_3 and ζ_4 are of the same sign as $F_4 > 0$ and from Eq. (14) implies that they should be negative i.e., $\zeta_3 < 0$, $\zeta_4 < 0$. Further, if ζ_3 and ζ_4 are complex conjugate, then $2Re(\zeta_3) = -F_1$, i.e., ζ_3 and ζ_4 have negative real parts. Thus, in both case, the roots ζ_3, ζ_4 lie in the left half of the complex plane. Now, we verify the transversality condition to find out the interval under which Hopf-bifurcation occurs. For this let at any point $\alpha_1 \in (\alpha_{1_c} - \epsilon, \alpha_{1_c} + \epsilon)$ and $\zeta_{1,2} = \kappa \pm i\eta$, putting this values in Eq. (11), we get

$$\kappa^4 + F_1\kappa^3 + F_2\kappa^2 + F_3\kappa + F_4 + \eta^4 - 6\kappa^2\eta^2 - 3F_1\kappa\eta^2 - F_2\eta^2 = 0, \tag{18}$$

$$4\kappa\eta(\kappa^2 - \eta^2) - F_1\eta^3 + 3F_1\kappa^2\eta + 2F_2\kappa\eta + F_3\eta = 0 \tag{19}$$

Since, $\eta(\alpha_1) \neq 0$, from Eq. (19), we get

$$-(4\kappa + F_1)\eta^2 + 4\kappa^3 + 3F_1\kappa^2 + 2F_2\kappa + F_3 = 0 \tag{20}$$

putting this value of η^2 in Eq. (18), we have

$$\begin{aligned} &-64\kappa^6 - 96F_1\kappa^5 - 16(3F_1^2 + 2F_2)\kappa^4 - 8(F_1^3 + 4F_1F_2)\kappa^3 \\ &-4(F_2^2 + 2F_1^2F_2 + F_1F_3 - 4F_4)\kappa^2 - 2F_1(F_2^2 + F_1F_3 - 4F_4)\kappa \\ &-(F_1F_2F_3 - F_3^2 - F_1^2F_4) = 0 \end{aligned} \tag{21}$$

Differentiating above equation with respect to α_1 and by using $\kappa(\alpha_{1_c}) = 0$, we have

$$\left[\frac{d\kappa}{d\alpha_1}\right]_{\alpha_1=\alpha_{1_c}} = \left[\frac{\frac{d}{d\alpha_1}\psi(\alpha_1)}{-2F_1(F_2^2 + F_1F_3 - 4F_4)}\right]_{\alpha_1=\alpha_{1_c}} \tag{22}$$

using the value of $F_4(\alpha_{1_c}) = \left(\frac{(F_1(\alpha_{1_c})F_2(\alpha_{1_c})F_3(\alpha_{1_c})-F_3^2(\alpha_{1_c})}{F_1^2(\alpha_{1_c})}\right)$ from Eq. (12), we obtain

$$\left[\frac{d\kappa}{d\alpha_1}\right]_{\alpha_1=\alpha_{1_c}} = \left[\frac{\frac{d}{d\alpha_1}\psi(\alpha_1)}{-2F_1(F_1F_3+(2\frac{F_3}{F_1}-F_2)^2}\right]_{\alpha_1=\alpha_{1_c}} \neq 0$$

if $\left[\frac{d}{d\alpha_1}\psi(\alpha_1)\right]_{\alpha_1=\alpha_{1_c}} \neq 0$. Therefore, the transverality condition holds if $\left[\frac{d}{d\alpha_1}\psi(\alpha_1)\right]_{\alpha_1=\alpha_{1_c}} \neq 0$. Hence, we have the following theorem for Hopf-bifurcation existence.

Theorem 6.1 *The Hopf-bifurcation occurs in the reduced model system (2) around the endemic equilibrium E^* if the there exists $\alpha_1 = \alpha_{1_c}$ such that:*

(i) $F_i(\alpha_{1_c}) > 0$ $(i = 1, \ldots, 4)$,
(ii) $\psi(\alpha_{1_c}) = 0$,
(iii) $\left[\mathscr{R}\frac{d\zeta_i(\alpha_1)}{d\alpha_1}\right]_{\alpha_1=\alpha_{1_c}} \neq 0$ *for i= 1, 2*
i.e., $\left[\frac{d}{d\alpha_1}\psi(\alpha_1)\right]_{\alpha_1=\alpha_{1_c}} \neq 0$.

Due to the complexities in the algebraic expressions of $\psi(\alpha_1)$ and $\frac{d\psi}{d\alpha_1}$, we use numerical simulations to verify the transversality conditions for a Hopf bifurcation. From Fig. 4 we see that the system undergoes a Hopf bifurcation when α_1 crosses $\alpha_{1_c} = 0.28$.

To identify the changes in the infected population density with the changes in infection rate and baseline media cover, in Fig. 5a we plot a two-parameter bifurcation diagram with β and M_0 as bifurcation parameters. We observe that any increase in the infection rate increases the number of infected individuals in the system. However, the increase in the baseline media coverage helps in decreasing the infected population density. We also observe that if the baseline media coverage is high, the disease can be eradicated with a moderately low infection rate.

Fig. 5b represents the changes in the infected population with the changes in infection rate and recovery rate. We observe that any increase in the recovery rate helps in reducing the number of infected individuals. When the recovery rate is low, there is a surge in infected individuals in the system.

From Fig. 6a we observe that the increase in the dissemination rate of awareness among susceptible individuals plays a major role in eradicating the disease. However, if the infection rate is high, the disease can thrive in the population even with a high awareness level. Also, a sufficiently high infection rate can lead to oscillatory dynamics of the system about the endemic state.

From Fig. 6b we observe that restricting the immigration of the susceptible individuals can help eradicate the disease. However, even a moderate increase in immigration can lead to a persistence of the disease in the population. In this case,

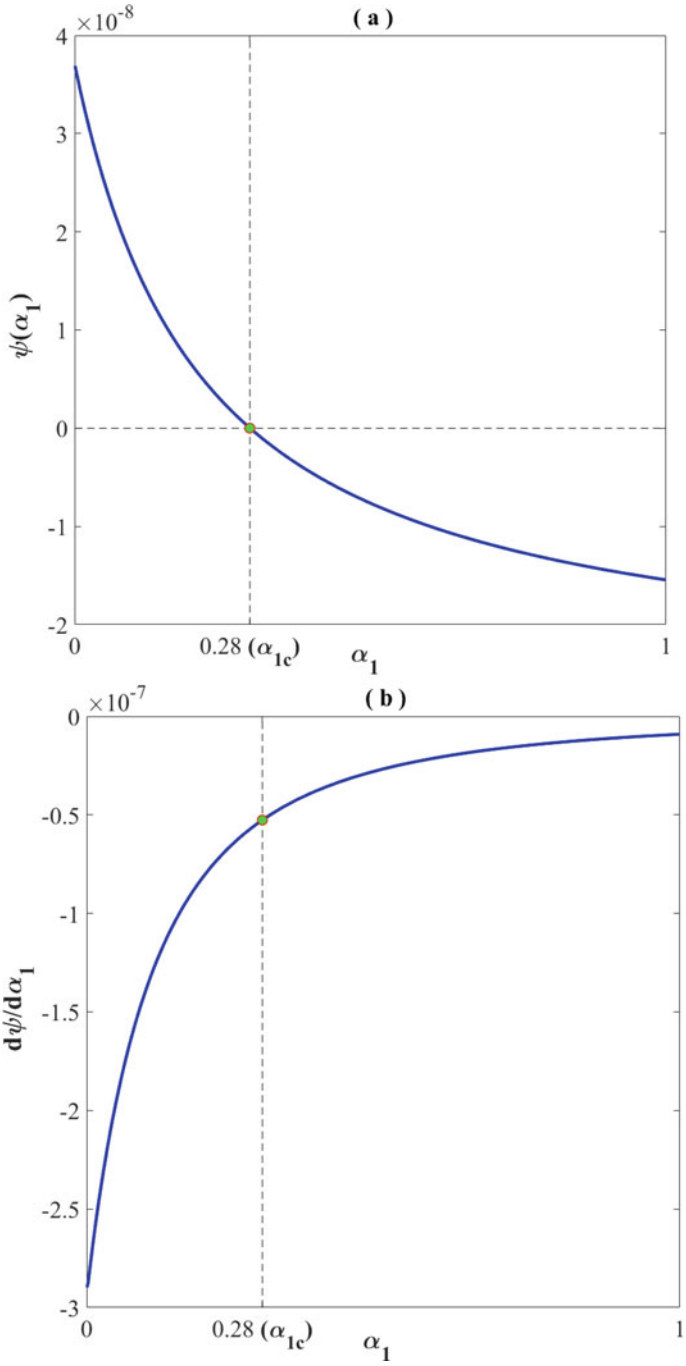


Fig. 4 Verification of the transversality conditions for Hopf bifurcation when α_1 crosses α_{1_c}

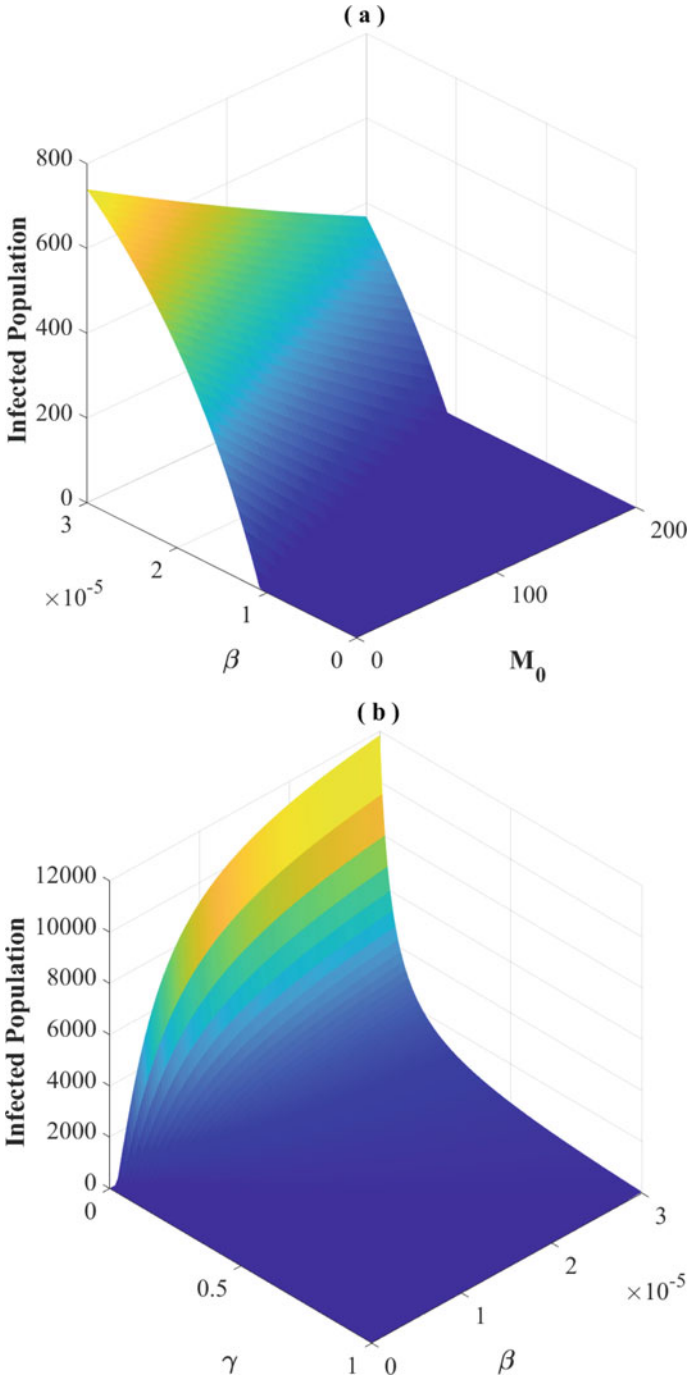


Fig. 5 The changes in the infected population density with the changes in (**a**) infection rate and baseline media cover, and (**b**) infection rate and recovery rate

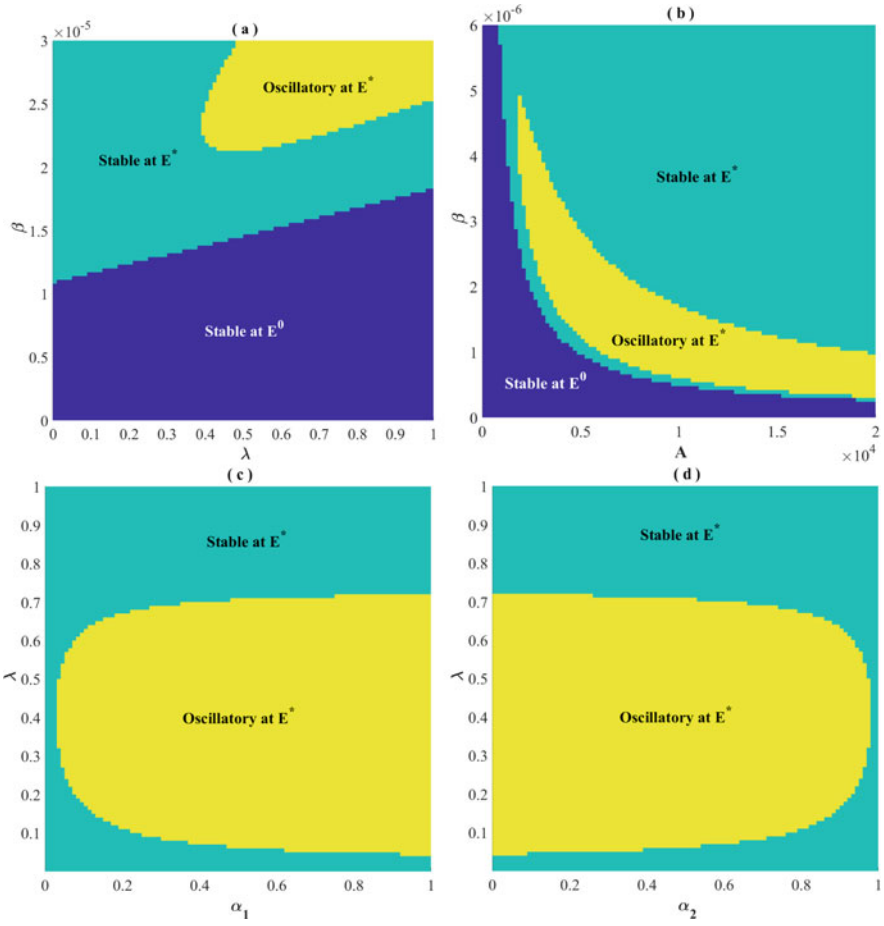


Fig. 6 Two parameter bifurcation of the system with respect to the parameters (**a**) λ and β; (**b**) A and β; (**c**) α_1 and λ; (**d**) α_2 and λ

the system exhibits dynamic instability followed by stability about the endemic equilibrium with the increase in the infection rate.

Further, we observe that the viewer's preference for the TV over the internet plays no significant difference in the system dynamics (c.f. Fig. 6c, d).

7 Discussion

Awareness campaigns through TV and internet advertisements promote awareness and bring behavioral changes in viewers on mitigating the spread of infectious disease. In this current work, we have proposed and scrutinized a nonlinear

mathematical model for infectious disease control by taking into account a convex combination of the roles of TV and internet coverage in imparting awareness among the human population. By investigating the proposed model, we obtain two possible feasible equilibria: the disease-free equilibrium (E^0) and the endemic equilibrium (E^*). We employed the next-generation technique to calculate an expression for the basic reproduction number R_0 and observed that whenever $R_0 < 1$ is satisfied, the disease-free equilibrium exists and is stable. However, when the value of R_0 is greater than unity, the disease-free equilibrium becomes unstable and the endemic equilibrium E^* appears. The endemic equilibrium is locally asymptotically stable under certain conditions. Taking R_0 as a bifurcation parameter, we observe that the system (2) switches its stability through a transcritical bifurcation. The existence of Hopf-bifurcation has been noticed when α_1 crosses $\alpha_{1_c} = 0.28$ (see Fig. 4). We have investigated the role of some key parameters of our model in managing the disease. It is observed that with the increase in the infection rate, more individuals in the system become infected, whereas an increase in the baseline media coverage contributes to a reduction in the infected population density. It is also observed that there is a surge in the number of infected individuals as long as the recovery rate is low. The results suggest that the spread of the disease can be minimized or even can be completely eradicated from the system if the distribution of awareness information through TV and the internet among the susceptible population is sufficiently high, ensuring a strong baseline awareness in the population.

The study found that at the early stage TV and internet coverage are quite effective. The effect of isolation and the rate of awareness among susceptible persons owing to TV and internet commercials are also found to help reduce disease spread. Aside from this, some other factors such as immunization, cleanliness, social distancing and voluntary quarantine are also essential aspects in the community for disease eradication.

Acknowledgments The research work of Rakesh Medda is financially supported by Council of Scientific and Industrial Research, Government of India in the form of Senior Research Fellowship (File No.: 09/106(0175)/2018-EMR-I). Joydeb Bhattacharyya is partially supported by the grants from West Bengal State Council of Science and Technology (WBSCST), Govt. of West Bengal, India (Grant No. ST/P/S&T/16G-06/2018). The research of Samares Pal is partially supported by Science and Engineering Research Board, Government of India (Grant No. MTR/2020/000542).

References

1. Owen O'Donnell. Access to health care in developing countries: breaking down demand side barriers. *Cadernos de saude publica*, 23:2820–2834, 2007.
2. Gram Vaani. Rural healthcare: towards a healthy rural india. *Rural Health Care: Towards a Healthy Rural India-Gramvaani-Community-Powered-Technology*, 3, 2013.
3. GO Agaba, YN Kyrychko, and KB Blyuss. Mathematical model for the impact of awareness on the dynamics of infectious diseases. *Mathematical biosciences*, 286:22–30, 2017.
4. Khaled Al-Surimi, Mohammed Khalifa, Salwa Bahkali, EL-Metwally Ashraf, and Mowafa Househ. The potential of social media and internet-based data in preventing and fighting

infectious diseases: from internet to twitter. In *Emerging and Re-emerging Viral Infections*, pages 131–139. Springer, 2016.
5. Vikas Bajpai. The challenges confronting public hospitals in india, their origins, and possible solutions. *Advances in Public Health*, 2014, 2014.
6. Noel T Brewer, Gretchen B Chapman, Frederick X Gibbons, Meg Gerrard, Kevin D McCaul, and Neil D Weinstein. Meta-analysis of the relationship between risk perception and health behavior: the example of vaccination. *Health psychology*, 26(2):136, 2007.
7. Xinghua Chang, Maoxing Liu, Zhen Jin, and Jianrong Wang. Studying on the impact of media coverage on the spread of covid-19 in hubei province, china. *Math. Biosci. Eng*, 17(4):3147–3159, 2020.
8. Sebastian Funk, Erez Gilad, Chris Watkins, and Vincent AA Jansen. The spread of awareness and its impact on epidemic outbreaks. *Proceedings of the National Academy of Sciences*, 106(16):6872–6877, 2009.
9. David Greenhalgh, Sourav Rana, Sudip Samanta, Tridip Sardar, Sabyasachi Bhattacharya, and Joydev Chattopadhyay. Awareness programs control infectious disease–multiple delay induced mathematical model. *Applied Mathematics and Computation*, 251:539–563, 2015.
10. Herbert Hethcote, Ma Zhien, and Liao Shengbing. Effects of quarantine in six endemic models for infectious diseases. *Mathematical biosciences*, 180(1–2):141–160, 2002.
11. Hem Joshi, Suzanne Lenhart, Kendra Albright, and Kevin Gipson. Modeling the effect of information campaigns on the hiv epidemic in uganda. *Mathematical Biosciences & Engineering*, 5(4):757, 2008.
12. Istvan Z Kiss, Jackie Cassell, Mario Recker, and Péter L Simon. The impact of information transmission on epidemic outbreaks. *Mathematical biosciences*, 225(1):1–10, 2010.
13. Ivar S Kristiansen, Peder A Halvorsen, and Dorte Gyrd-Hansen. Influenza pandemic: perception of risk and individual precautions in a general population. cross sectional study. *BMC public health*, 7(1):1–7, 2007.
14. Anuj Kumar, Prashant K Srivastava, and Yasuhiro Takeuchi. Modeling the role of information and limited optimal treatment on disease prevalence. *Journal of theoretical biology*, 414:103–119, 2017.
15. Vangipuram Lakshmikantham and Srinivasa Leela. *Differential and Integral Inequalities: Theory and Applications: Volume I: Ordinary Differential Equations*. Academic press, 1969.
16. Joseph TF Lau, Xilin Yang, HY Tsui, and Jean H Kim. Impacts of sars on health-seeking behaviors in general population in hong kong. *Preventive medicine*, 41(2):454–462, 2005.
17. Rongsong Liu, Jianhong Wu, and Huaiping Zhu. Media/psychological impact on multiple outbreaks of emerging infectious diseases. *Computational and Mathematical Methods in Medicine*, 8(3):153–164, 2007.
18. Yiping Liu and Jing-An Cui. The impact of media coverage on the dynamics of infectious disease. *International Journal of Biomathematics*, 1(01):65–74, 2008.
19. AK Misra, Anupama Sharma, and JB Shukla. Modeling and analysis of effects of awareness programs by media on the spread of infectious diseases. *Mathematical and Computer Modelling*, 53(5–6):1221–1228, 2011.
20. AK Misra, Anupama Sharma, and Vishal Singh. Effect of awareness programs in controlling the prevalence of an epidemic with time delay. *Journal of Biological Systems*, 19(02):389–402, 2011.
21. Arvind Kumar Misra and Rajanish Kumar Rai. A mathematical model for the control of infectious diseases: Effects of tv and radio advertisements. *International Journal of Bifurcation and Chaos*, 28(03):1850037, 2018.
22. Arvind Kumar Misra and Rajanish Kumar Rai. Impacts of tv and radio advertisements on the dynamics of an infectious disease: A modeling study. *Mathematical Methods in the Applied Sciences*, 42(4):1262–1282, 2019.
23. Arvind Kumar Misra, Rajanish Kumar Rai, and Yasuhiro Takeuchi. Modeling the control of infectious diseases: Effects of tv and social media advertisements. *Mathematical Biosciences & Engineering*, 15(6):1315, 2018.

24. Arvind Kumar Misra, Rajanish Kumar Rai, and Yasuhiro Takeuchi. Modeling the effect of time delay in budget allocation to control an epidemic through awareness. *International Journal of Biomathematics*, 11(02):1850027, 2018.
25. Farai Nyabadza, Christinah Chiyaka, Zindoga Mukandavire, and Senelani D Hove-Musekwa. Analysis of an hiv/aids model with public-health information campaigns and individual withdrawal. *Journal of Biological Systems*, 18(02):357–375, 2010.
26. Lawrence Perko. Nonlinear systems: Bifurcation theory. In *Differential Equations and Dynamical Systems*, pages 315–540. Springer, 2001.
27. Sudip Samanta, Sourav Rana, Anupama Sharma, Arvind Kumar Misra, and Joydev Chattopadhyay. Effect of awareness programs by media on the epidemic outbreaks: A mathematical model. *Applied Mathematics and Computation*, 219(12):6965–6977, 2013.
28. Chengjun Sun, Wei Yang, Julien Arino, and Kamran Khan. Effect of media-induced social distancing on disease transmission in a two patch setting. *Mathematical biosciences*, 230(2):87–95, 2011.
29. Pauline Van den Driessche and James Watmough. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical biosciences*, 180(1–2):29–48, 2002.
30. Abdallah IA Yagub. Collaboration between government and non-governmental organizations (ngos) in delivering curative health services in north darfur state, sudan-a national report. *Iranian journal of public health*, 43(5):561, 2014.

Potential Severity, Mitigation, and Control of Omicron Waves Depending on Pre-existing Immunity and Immune Evasion

F. A. Bartha, P. Boldog, T. Tekeli, Z. Vizi, A. Dénes, and G. Röst

1 Introduction

At the end of November 2021, according to the World Health Organization, the B.1.617.2 Delta variant accounted for 99% of COVID-19 cases around the world [1]. Yet in recent weeks, we have observed the rise of the B.1.1.529 lineage, designated as the Omicron variant. This variant was first reported to the WHO on 24th November, and as of 9th December, Omicron sequences have been found already in 63 countries [2], in spite of travel restrictions. Omicron transmission dynamics can be estimated from sharp turning points in epidemiological trends, widespread genome sequencing or by S–gene target failure (SGTF), which is a proxy for Omicron [3]. Omicron has out-competed Delta in South Africa in a short time [4], and rapid growth of cases has been observed in the United Kingdom [5] and Denmark [6].

There is accumulating evidence of high transmissibility and immune-evasion capability of the Omicron variant [3, 5, 7, 8]. Ongoing neutralization studies indicate a significant drop in vaccine efficacy [9], and increased frequency of reinfection has been reported in South Africa [8].

There is an urgent need to estimate the potential impact of this variant. Modelling for the United Kingdom has been posted on 11th December [5]. However, since countries have different levels of pre-existing immunity either from vaccination or from previous infections, and different non-pharmaceutical measures are in place, we can expect that countries will be affected differently. Our goal here is to provide a quick assessment of this threat, considering these country-specific factors.

F. A. Bartha · P. Boldog · T. Tekeli · Z. Vizi · A. Dénes (✉) · G. Röst
Bolyai Institute, University of Szeged, Szeged, Hungary
e-mail: denesa@math.u-szeged.hu

R. P. Mondaini (ed.), *Trends in Biomathematics: Stability and Oscillations in Environmental, Social, and Biological Models*,
https://doi.org/10.1007/978-3-031-12515-7_22

2 Methods

2.1 Potential Control of Delta and Omicron

The effective reproduction number corresponding to the Delta variant (denoted by $\mathcal{R}_t^{\Delta}$) at a given time characterizes the current transmission of the infection in a population. It can be obtained by the correction of the basic reproduction number ($\mathcal{R}_0^{\Delta}$) with the mitigating effect of the actual non-pharmaceutical interventions (NPI) and the population immunity level (p) in reducing transmission. That is, $\mathcal{R}_t^{\Delta} = \mathcal{R}_0^{\Delta} \cdot (1 - \text{NPI}) \cdot (1 - p)$. Similarly, if the Omicron is already present in the same population, its effective reproduction number is $\mathcal{R}_t^{o} = \mathcal{R}_0^{o} \cdot (1 - \text{NPI}) \cdot (1 - p_o)$, where $\mathcal{R}_0^{o}$ is the basic reproduction number of the Omicron variant, and p_o is the population immunity against this variant.

We introduce parameter e for the extent of immune evasion, expressing that immunity p to prior variants is reduced by $e \cdot p$ with respect to Omicron, that is $p_o = p \cdot (1 - e)$. Let p^{SA} be the population immunity in South Africa, and let $q = \mathcal{R}_t^{o}/\mathcal{R}_t^{\Delta}$ at the time of the emergence of Omicron. Then, we have the relation

$$\mathcal{R}_0^{o} \cdot (1 - \text{NPI}) \cdot (1 - p^{SA} \cdot (1 - e)) = q \cdot \mathcal{R}_0^{\Delta} \cdot (1 - \text{NPI}) \cdot (1 - p^{SA}). \qquad (1)$$

In a country with a combination of pre-existing immunity p and NPIs, to achieve $\mathcal{R}_0^{\Delta} \leq 1$ and $\mathcal{R}_0^{o} \leq 1$, one needs

$$\text{NPI} > 1 - 1/(\mathcal{R}_0^{\Delta} \cdot (1 - p)), \quad \text{and} \quad \text{NPI} > 1 - 1/(\mathcal{R}_0^{o} \cdot (1 - p_o), \qquad (2)$$

respectively.

2.2 Transmission Dynamics of Omicron

To estimate the potential of an Omicron wave, we employ a compartmental model of disease dynamics, which is monitoring the temporal change of the number of infected individuals in the population, separately with and without pre-existing immunity. The technical details of the model and the numerical code are described in the Supplement. Most parameters are estimated from the literature, while p and e are varied in a feasible interval. The differential equations are solved for many combinations of p and e, thus we can assess the total epidemic size, the peak size and the peak time of the Omicron wave for a range of scenarios.

2.3 Parametrization

Current estimates from South-Africa and United Kingdom indicate that in those countries $\mathcal{R}_t^o/\mathcal{R}_t^\Delta \approx 4$ [7, 10–12], which means that an Omicron-case generates four times more cases than a Delta-case. As baseline parameters, we choose $\mathcal{R}_0^\Delta = 6$ [13, 14], and for South Africa at the time of the emergence of Omicron we assume $q = \mathcal{R}_t^o/\mathcal{R}_t^\Delta = 4$, and $p^{SA} = 0.85$ [15], but these key parameters can easily be varied to explore the sensitivity of the outcome. The choice of p^{SA} is consistent with the observed $\mathcal{R}_t^\Delta$ in South Africa in a period before Omicron [10]. Previous experience in European countries shows that strict lockdowns can achieve an 82% drop in the effective reproduction number [16], hence we consider NPI < 0.82. For the compartmental model, following [5], the average incubation period is assumed to be 2.5 days, and the average infectious period is assumed to be 5 days, each following a gamma distribution.

3 Results

Since the NPIs affect the transmission of both variants in the same way, from $q = \mathcal{R}_t^o/\mathcal{R}_t^\Delta = 4$, we have $\mathcal{R}_0^o \cdot (1 - p_o) = 4 \cdot \mathcal{R}_0^\Delta \cdot (1 - p)$. Such a four-fold advantage of Omicron can emerge either from inherently higher transmissibility or a larger susceptible pool, when immunity obtained by vaccination or prior infection by other variants does not protect against the new variant as effectively. There is an obvious trade-off relationship between these two factors determining the transmission fitness: the higher the transmissibility, the lower the immune evasion must be to maintain the four-fold ratio of the effective reproduction numbers.

With our baseline parameters, from (1), we obtain the relation

$$\mathcal{R}_0^o(1 - 0.85 \cdot (1 - e)) = 3.6,$$

showing that if immune evasion is significant, then $\mathcal{R}_0^o$ must be more moderate. In particular, $\mathcal{R}_0^o < \mathcal{R}_0^\Delta$ whenever $e > 0.53$.

Consider a country with population immunity p, where Delta is controlled, then by (2), NPI $> 1 - 1/(6 \cdot (1 - p))$. The necessary NPIs to control Omicron can also be found from (1) and (2) as

$$\text{NPI} > 1 - (1 - 0.85 \cdot (1 - e))/(3.6 \cdot (1 - p \cdot (1 - e))).$$

The necessary NPIs are plotted in Fig. 1 as a function of p for various values of e. As we can see, the NPIs that are sufficient against Delta are not enough to stop the spreading of Omicron, for any considered combination of p and e. Since Omicron containment requires very stringent NPIs, the invasion of this variant is likely to result in widespread infection.

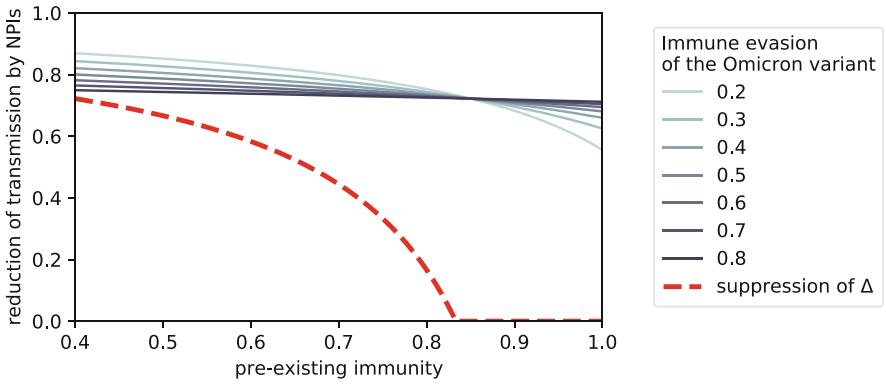


Fig. 1 The necessary level of non-pharmaceutical interventions (NPI) to control the Omicron and Delta (dashed) variants as a function of pre-existing immunity

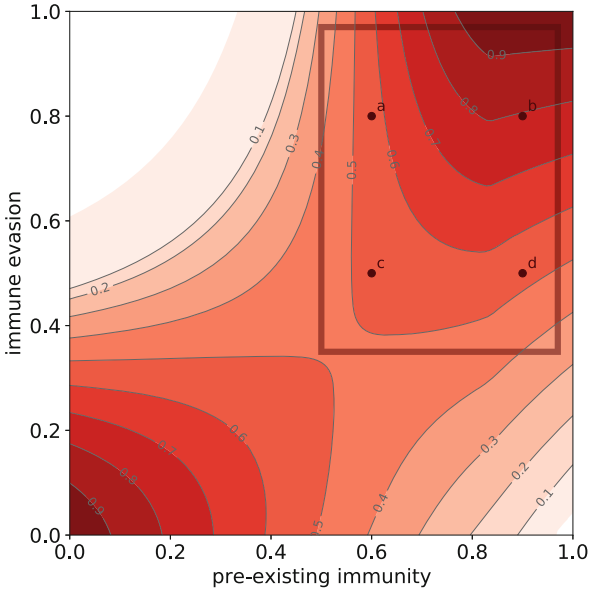


Fig. 2 Heatmap of the total number of infections (as fraction of the population) during Omicron wave, depending on pre-existing immunity and Omicron's immune evasion, assuming that no further mitigation measures will be implemented. We select examples from a feasible parameter region highlighted by a rectangle

To estimate the severity of future Omicron waves in countries where Delta is under control, we solve our transmission model for a range of parameters (p, e), and when population immunity is below the Delta herd immunity threshold, we employ the necessary NPIs (the red dashed curve in Fig. 1) to achieve $\mathcal{R}_t^{\Delta} \leq 1$. The heatmap in Fig. 2 shows the fraction of the population infected during the Omicron wave with respect to the parameters (p, e) space. One can see that the only scenarios not to have a significant outbreak are the following: the population has extremely high immunity and Omicron is not very immune evasive (bottom right corner); or the competitive advantage of Omicron emanates from high immune evasion rather

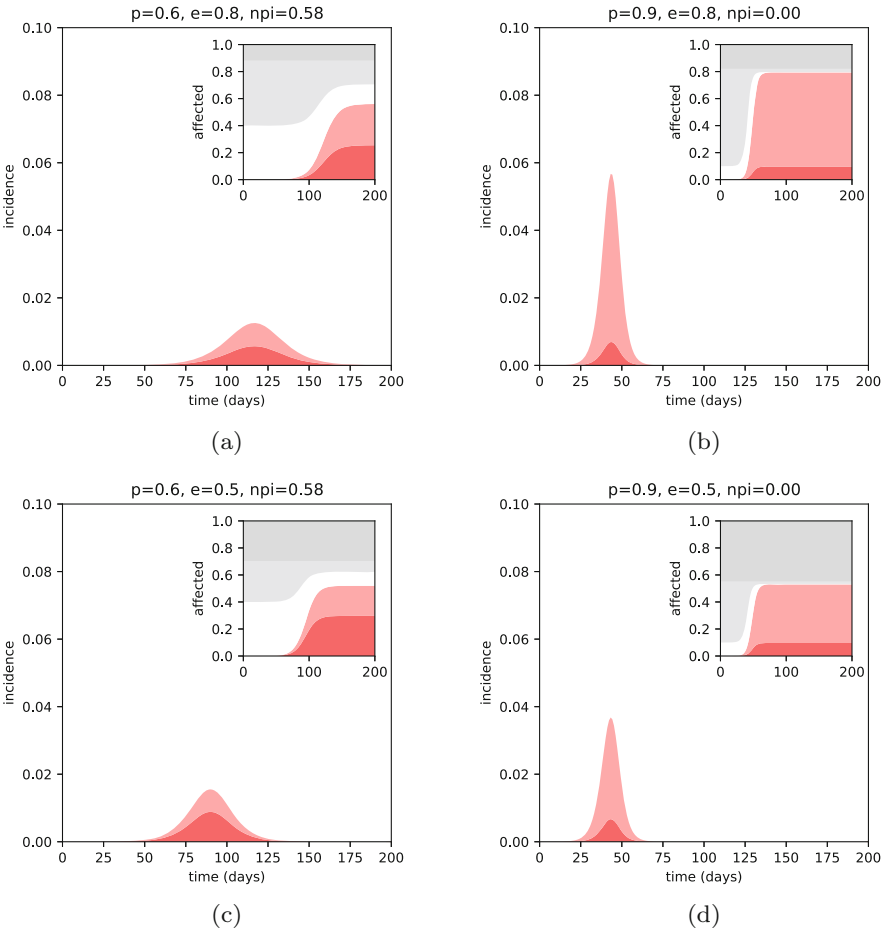


Fig. 3 Epidemic curves of the Omicron wave under various assumptions on the pre-existing immunity and the variant's immune evasion, in the absence any additional measures. (**a**) $p = 0.6$, $e = 0.8$, NPI $= 0.58$. (**b**) $p = 0.9$, $e = 0.8$, NPI $= 0.00$. (**c**) $p = 0.6$, $e = 0.5$, NPI $= 0.58$. (**d**) $p = 0.9$, $e = 0.5$, NPI $= 0.00$

than inherent transmissibility and at the same lockdown-like NPIs are in place (top left corner): none of which are plausible. We chose examples a, b, c, d from the highlighted rectangle of a feasible parameter region for more detailed investigation, but in our published code [17] we provide an interactive tool that makes exploring other scenarios very easy for the reader.

The time course of the Omicron outbreak in the four selected scenarios are depicted in Fig. 3. The dark red colour represents infected population without prior immunity, while the pink colour represents infected population having been vaccinated or previously infected. In the insets, the cumulative infected fraction is shown, and the grey colour represents the population with prior immunity, while

the lighter shade is the portion becoming available for Omicron-infections due to immune evasion.

Potential Omicron waves may have very different characteristics. In countries with very high population immunity, where Delta is contained with very mild or without NPIs, one can expect an extremely rapid increase of cases and a high peak in a matter of weeks, especially if Omicron is highly immune evasive (scenarios *b* and *d* in Fig. 3). If population immunity is moderate and strong NPIs are being employed to contain Delta, then the Omicron wave is more flattened, as long as the NPIs remain in place (scenarios *a* and *c* in Fig. 3). In this case, as opposed to the high immunity scenarios, the peak is lower if immune evasion is higher.

Besides the sheer number of infections, to assess the severity of the outbreaks, it is very important to take into consideration the prior immune status of the population getting infected. For example, despite in *b* and *d* the peak is much higher than in *a* and *c*, if Omicron infections of those with prior immunity turn out to be overwhelmingly mild, than the severity is much better reflected by the number of infected without prior immunity, that is the red curves without the pink part in Fig. 3. In this measure, the *a* and *c* are more severe. For scenarios *b* and *d*, reintroducing measures with NPI $= 0.4$ can reduce the peak size roughly to half, and delaying the peak by a month, as shown in Fig. 4.

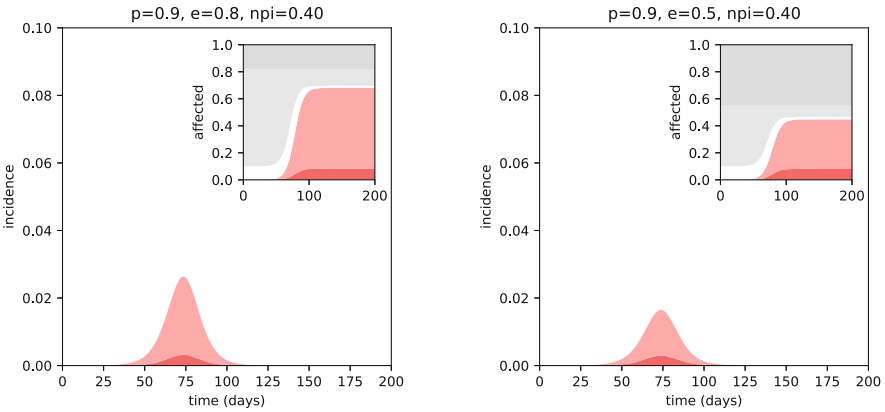


Fig. 4 Effect on the Omicron wave: introducing moderate NPIs in countries with high immunity against the Delta variant. Scenarios (3(b)) and (3(d)) from Fig. 3. (**a**) $p = 0.9$, $e = 0.8$, NPI $= 0.40$. (**b**) $p = 0.9$, $e = 0.5$, NPI $= 0.40$

4 Discussion

The Omicron variant is spreading across the globe with unprecedented speed. Hence there is great urgency to assess its potential consequences, despite all uncertainties about its epidemiological parameters. Our calculations suggest that any combination of current measures and population immunity and, that is just sufficient to suppress the Delta variant, will not be able to contain Omicron. Hence widespread Omicron transmission is expected. However, the impact of the Omicron wave on various countries can be very different depending on the level of pre-existing population immunity, and the immune evasion property of Omicron. The severity of the disease caused by Omicron is yet unclear, however previous immunity appears to reduce the risk of severe outcome [5]. To further improve the assessment, we constructed a transmission dynamics model that tracks the emerging infections in the groups with and without pre-existing immunity separately. This allowed us to investigate a range of scenarios, and we found that in countries with high immunity but mild NPIs a very sharp epidemic curve is expected with a high peak, but a smaller fraction of infections are from the high risk population (those without pre-existing immunity). On the other hand, countries with moderate immunity that employ stricter NPIs at the moment to fight Delta, can have a flatter Omicron curve, but with more infections from the high risk group. Once disease severity will be clarified with and without pre-existence immunity, the overall severity of the outbreak could be determined as a simple linear combination of the number of infected in those sub-populations. In case of high severity, reintroducing further NPIs might be needed to mitigate the Omicron waves, and that would also buy some time to expedite vaccination programs and booster roll-outs.

Acknowledgments The authors were supported by NKFIH and ITM (FK 138924, FK 124016, PD 128363, KKP 129877, 2020-2.1.1-ED-2020-00003, TUDFO/47138-1/2019-ITM). F.B. was also supported by ÚNKP-21-5 and the Bolyai Scholarship of the Hungarian Academy of Sciences.

Appendix: Supplementary Material

Trade-Off Between Inherent Transmissibility and Immune Evasion Capability

The contour formula (1) results in a trade-off between immune evasion and the spreading capability of the Omicron variant when all other parameters are kept fixed. In Fig. 5, we demonstrate this phenomenon fixing $\mathcal{R}_0^{\Delta}$ and q to their baseline values and allowing uncertainty of the pre-existing immunity (p^{SA}) in South Africa.

Note that the scenarios with $\mathcal{R}_0^{o} < \mathcal{R}_0^{\Delta}$ on Fig. 5 are to the right of the contour with special focus. Clearly, even when considering only $p^{SA} \approx 0.85$, the uncertainty of $\mathcal{R}_0^{o}$ is substantial as potential values lie within a range of [1.5, 24],

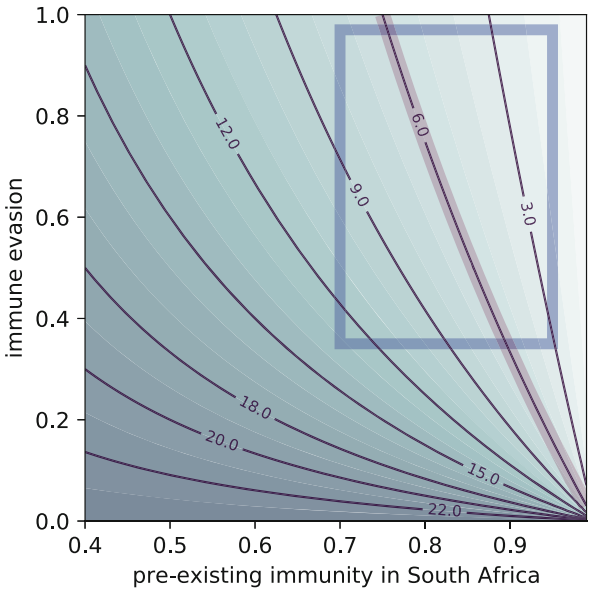


Fig. 5 $\mathcal{R}_0^o$-contours visualizing the trade-off between transmissibility ($\sim \mathcal{R}_0^o$) and immune evasion (e) of the Omicron variant. The rectangle marks a feasible parameter region for South Africa, the highlighted contour agrees with the transmissibility of Delta

where $24 = q \cdot \mathcal{R}_0^\Delta$ is the theoretical extreme when there is no immune evasion whatsoever. If we restrict our attention to a more feasible immune evasion region, similar to what was highlighted on the y-axis of Fig. 2, we may bound our estimate to $\mathcal{R}_0^o \in [1.5, 13]$.

Transmission Dynamics Model

The transmission dynamics is based on an SE_2I_4R model, generalized from $SEIR$ by assuming gamma-distributed incubation and infectious periods, using the Erlang parameters $n = 2$ and $m = 4$. The sixteen compartments $s(t)$, $s_p(t)$, $\ell_s^j(t)$, $\ell_p^j(t)$, $i_s^k(t)$, $i_p^k(t)$, $r_s(t)$ and $r_p(t)$—with $j = 1, 2$; $k = 1, 2, 3, 4$—represent fraction of the population being in different disease and immunity status. The compartment s designates population level susceptibility, without any pre-existing immunity to any of the variants, while s_p stands for susceptibles to Omicron despite of being immune to earlier variants due to a past infection or vaccination. Susceptibles (either from s or s_p) can get infected upon adequate contact with an infectious person (from one of i_s^k or i_p^k) with transmission rate β, and moved to the corresponding first latent compartment. For both pathways $\mathbf{i} = s, p$, the latent class comprises the two stages $\ell_{\mathbf{i}}^1$, $\ell_{\mathbf{i}}^2$. The length of the latent period is α^{-1}, thus, one moves with the rate 2α from the first latent stage to the second and thereon to the first infectious compartment, alike. The model features four infectious stages stretching over the infectious period γ^{-1}, again, for both pathways. Individuals transition from one infected stage to

the next one, as well as removed from the fourth with the rate 4γ. This linear chain ensures that the infectious period is gamma distributed. We do not model the clinical outcome of disease progression and all non-infectious individuals eventually enter an r class. Note that this does not pose any restrictions on assessing the severity of the Omicron wave as said severity is in direct correlation with the transmission dynamics of the epidemics captured by model time series and the corresponding peak and final size.

The above considerations are summarized in the compartmental ODE system:

$$
\begin{aligned}
s'(t) &= -\beta s(t) \sum_{k=1}^{4} (i_s^k(t) + i_p^k(t)), & s_p'(t) &= -\beta s_p(t) \sum_{k=1}^{4} (i_s(t)^k + i_p^k(t)),\\
\ell_s^{1\prime}(t) &= \beta s(t) \sum_{k=1}^{4} (i_s^k(t) + i_p^k(t)) - 2\alpha \ell_s^1(t), & \ell_p^{1\prime}(t) &= \beta s_p(t) \sum_{k=1}^{4} (i_s^k(t) + i_p^k(t)) - 2\alpha \ell_p^1(t),\\
\ell_s^{2\prime}(t) &= 2\alpha \ell_s^1(t) - 2\alpha \ell_s^2(t), & \ell_p^{2\prime}(t) &= 2\alpha \ell_p^1(t) - 2\alpha \ell_p^2(t),\\
i_s^{1\prime}(t) &= 2\alpha \ell_s^2(t) - 4\gamma i_s^1(t), & i_p^{1\prime}(t) &= 2\alpha \ell_p^2(t) - 4\gamma i_p^1(t),\\
i_s^{2\prime}(t) &= 4\gamma i_s^1(t) - 4\gamma i_s^2(t), & i_p^{2\prime}(t) &= 4\gamma i_p^1(t) - 4\gamma i_p^2(t),\\
i_s^{3\prime}(t) &= 4\gamma i_s^2(t) - 4\gamma i_s^3(t), & i_p^{3\prime}(t) &= 4\gamma i_p^2(t) - 4\gamma i_p^3(t),\\
i_s^{4\prime}(t) &= 4\gamma i_s^2(t) - 4\gamma i_s^4(t), & i_p^{4\prime}(t) &= 4\gamma i_p^2(t) - 4\gamma i_p^3(t),\\
r_s'(t) &= 4\gamma i_s(t), & r_p'(t) &= 4\gamma i_p(t),
\end{aligned}
$$

with the transmission rate $\beta = R_0\gamma(1-\text{NPI})$. Here, the level of non-pharmaceutical interventions NPI(t) is assumed to be constant in the timeframe of each numerical simulation. Finally, the computations were carried out using the initial values $s(0) = 1 - p$, $s_p(0) = e \cdot p$, and a small amount of initial Omicron-infected.

Supplementary Source Codes

The software used in our analysis and for generating all figures is available on Github [17] with an option for direct experiments in Google Colab. We implemented our solution as a Jupyter notebook with Python kernel using the standard libraries for computation (`scipy`, `numpy`) and visualization (`matplotlib`).

The code is structured as follows. First, global parameters are defined such as $\mathcal{R}_0^{\Delta}$, pSA, and q enabling effortless sensitivity analysis of the results. Then, the methodological core is set up to compute $\mathcal{R}_0^o$, suppressing NPIs in various scenarios, and to carry out numerical simulations of the ODE model. Finally, the code snippets for visualization are enabled and a multitude of figures are produced, many of which are equipped with an interactive interface to ease the exploration of the parameter space.

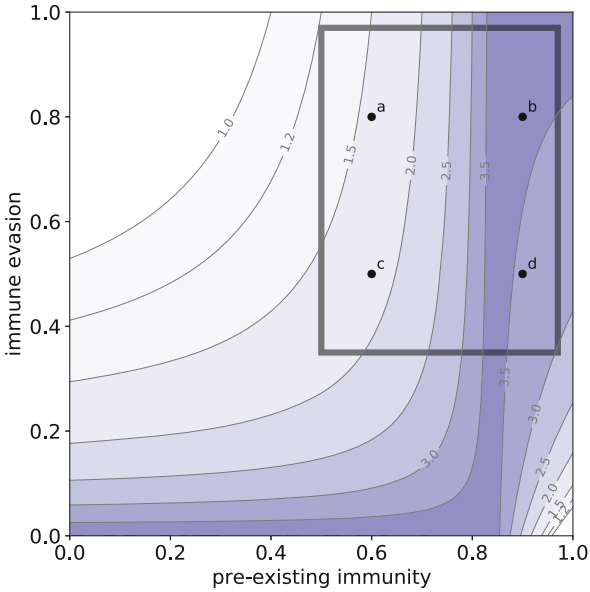


Fig. 6 Reproduction number of Omicron when introduced into a population where Delta is controlled

Invasion Reproduction Number of Omicron

The effective reproduction number at the time of introduction into a population is shown in Fig. 6. The values in the figure are in line with the empirically estimated value of $\mathcal{R}_t^o = 3.7$ for the UK [5].

Severity Assessment

The overall severity of an Omicron wave might be assessed by the number of infected individuals who had no prior immunity, as they are at higher risk of severe disease outcome. Figure 7 shows the comparison of heatmaps.

A striking difference is that while the peak size of all infected is the highest when pre-existing immunity and immune evasion are both very high, meanwhile this is a relatively favorable situation in terms of peak size of the infected without prior immunity. A similar comparison can be made between the total number of infections (Fig. 2) and counting only those without prior immunity Fig. 8.

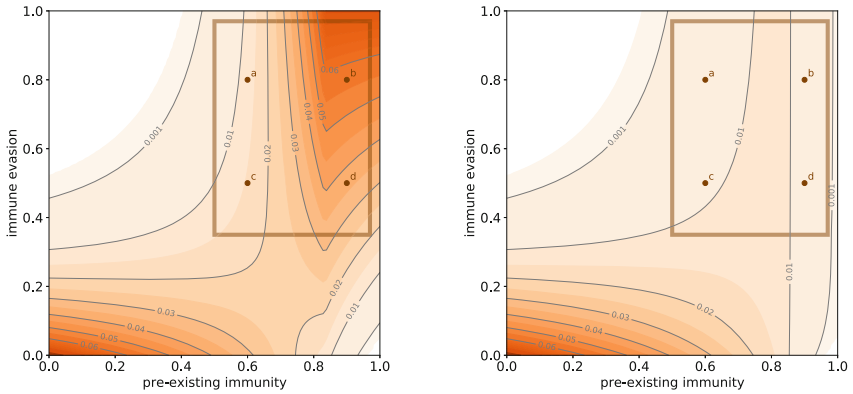


Fig. 7 Peak size of the Omicron wave in the whole population, and in the sub-population without pre-existing immunity. (**a**) Peak size of all infected. (**b**) Peak size of infected without prior immunity

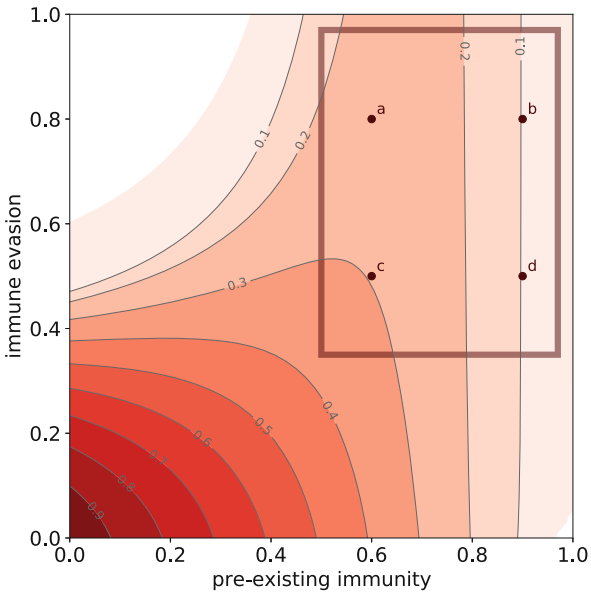


Fig. 8 Total number of infections without pre-existing immunity during the Omicron wave, as the fraction of the total population

Sensitivity to Parameters

Our publicly available code [17] make it easy for anyone to explore the sensitivity of the outcomes to the key parameters. A higher value p^{SA} would mean that the Omicron is less transmissible hence more controllable by NPIs, see Fig. 9.

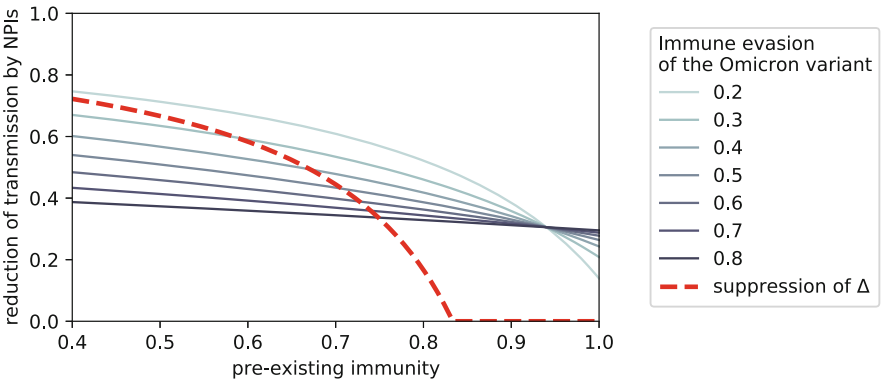


Fig. 9 The necessary level of non-pharmaceutical interventions to control the Omicron and the Delta variants as a function of pre-existing immunity, when $p^{SA} = 0.94$

Varying $\mathcal{R}_0^{\Delta}$ around 6 does not change the main features of the figures. Decreasing q makes the epidemic curves flatter, however for $q = 3$ the invasion reproduction number of Omicron is becoming already too small compared to the observations from UK [5].

References

1. WHO, Weekly epidemiological update on COVID-19 - 30 November 2021, Edition 68. *World Health Organization* **2021.11.30.** Available online: https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19---30-november-2021 *(accessed on 14 December 2021)*.
2. WHO, Enhancing Readiness for Omicron (B.1.1.529): Technical Brief and Priority Actions for Member States. *World Health Organization* **2021.12.10.** Available online: https://www.who.int/publications/m/item/enhancing-readiness-for-omicron-(b.1.1.529)-technical-brief-and-priority-actions-for-member-states *(accessed on 14 December 2021)*.
3. UKHSA. SARS-CoV-2 variants of concern and variants under investigation in England. *UK Health Security Agency* **2021.12.10.** Available online: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1040076/Technical_Briefing_31.pdf *(accessed on 14 December 2021)*.
4. The GISAID Initiative. Online database for variant tracking. *GISAID* **2021.** Available online: https://www.gisaid.org/hcov19-variants/ *(accessed on 14 December 2021)*.
5. R.C. Barnard, N.G. Davies, *et al.*, Modelling the potential consequences of the Omicron SARS-CoV-2 variant in England. *preprint* **2021.12.11.** Available online: https://cmmid.github.io/topics/covid19/omicron-england *(accessed on 14 December 2021)*.
6. SSI. Covid-19 Rapport om omikronvarianten B.1.1.529. *Statens Serum Institut* **2021.12.14.** Available online: https://files.ssi.dk/covid19/omikron/statusrapport/rapport-omikronvarianten-14122021-rg41 *(accessed on 14 December 2021)*.
7. C.A.B. Pearson *et al.*, Bounding the levels of transmissibility and immune evasion of the Omicron variant in South Africa, South African COVID-19 Modelling Consortium (SACMC) , https://www.sacmcepidemicexplorer.co.za/downloads/Pearson_etal_Omicron.pdf *(accessed on 22 December 2021)*.

8. J.R.C. Pulliam, C. van Schalkwyk, *et al.*, Increased risk of SARS-CoV-2 reinfection associated with emergence of Omicron in South Africa, *Science* **2022**, *376*, 6593. https://doi.org/10.1126/science.abn4947.
9. B.J. Gardner, A.M. Kilpatrick, Estimates of reduced vaccine effectiveness against hospitalization, infection, transmission and symptomatic disease of a new SARS-CoV-2 variant, Omicron (B.1.1.529). using neutralizing antibody titers, *medRχiv* **2021.12.12.** https://doi.org/10.1101/2021.12.10.21267594.
10. M. Figgins, T. Bedford, Estimates for Omicron across countries. *Github* **2021.** Available online: https://github.com/blab/rt-from-frequency-dynamics/blob/master/estimates/omicron-countries *(accessed on 14 December 2021)*.
11. L. Rossouw, Estimating the effective reproduction number of COVID-19 in South Africa. *Unsupervized* **2021.** Available online: https://unsupervised.online/static/covid-19/estimating_r_za.html *(accessed on 14 December 2021)*.
12. H. Nishiura, Covid-19 Report – 7 December 2021. *Japanese Ministry of Health Labour, and Welfare* **2021.12.07.** Available online *(in Japanese)*: https://www.mhlw.go.jp/content/10900000/000864204.pdf *(accessed on 14 December 2021)*.
13. M. Kang, H. Xin, *et al.*, Transmission dynamics and epidemiological characteristics of Delta variant infections in China. *medRχiv* **2021.08.13.** https://doi.org/10.1101/2021.08.12.21261991.
14. R. Sonabend, L.K. Whittles, *et al.*, Evaluating the Roadmap out of Lockdown: modelling step 4 of the roadmap in the context of B.1.617.2. *Imper. Coll. London* **2021.06.09.** Available online: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/993427/S1289_Imperial_Roadmap_Step_4.pdf *(accessed on 14 December 2021)*.
15. R. Chowdhury, K. Heng, *et al.*, Dynamic interventions to control COVID-19 pandemic: a multivariate prediction modelling study comparing 16 worldwide countries. *Eur. J. Epidemiol.* **2021**, *35*, 389–399. https://doi.org/10.1007/s10654-020-00649-w.
16. S. Flaxman, S. Mishra *et al.*, Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature* **2020**, *584*, 257–261. https://doi.org/10.1038/s41586-020-2405-7.
17. F.A. Bartha, P. Boldog, T. Tekeli, Zs. Vizi, A. Dénes, G. Röst, Supplementary codes for Potential severity and control of Omicron waves depending on pre-existing immunity and immune evasion. *Github* **2021.** Available online: https://github.com/epidelay/covid-19-omicron-waves.

Index

R. P. Mondaini (ed.), *Trends in Biomathematics: Stability and Oscillations in Environmental, Social, and Biological Models*,
https://doi.org/10.1007/978-3-031-12515-7