STEAM-H: Science, Technology, Engineering, Agriculture, Mathematics & Health

Bourama Toni *Editor*

The Mathematics of Patterns, Symmetries, and Beauties in Nature

In Honor of John Adam

Springer

STEAM-H: Science, Technology, Engineering, Agriculture, Mathematics & Health

STEAM-H: Science, Technology, Engineering, Agriculture, Mathematics & Health

Series Editor
Bourama Toni
Department of Mathematics
Howard University
Washington, DC, USA

This interdisciplinary series highlights the wealth of recent advances in the pure and applied sciences made by researchers collaborating between fields where mathematics is a core focus. As we continue to make fundamental advances in various scientific disciplines, the most powerful applications will increasingly be revealed by an interdisciplinary approach. This series serves as a catalyst for these researchers to develop novel applications of, and approaches to, the mathematical sciences. As such, we expect this series to become a national and international reference in STEAM-H education and research.

Interdisciplinary by design, the series focuses largely on scientists and mathematicians developing novel methodologies and research techniques that have benefits beyond a single community. This approach seeks to connect researchers from across the globe, united in the common language of the mathematical sciences. Thus, volumes in this series are suitable for both students and researchers in a variety of interdisciplinary fields, such as: mathematics as it applies to engineering; physical chemistry and material sciences; environmental, health, behavioral and life sciences; nanotechnology and robotics; computational and data sciences; signal/image processing and machine learning; finance, economics, operations research, and game theory.

The series originated from the weekly yearlong STEAM-H Lecture series at Virginia State University featuring world-class experts in a dynamic forum. Contributions reflected the most recent advances in scientific knowledge and were delivered in a standardized, self-contained and pedagogically-oriented manner to a multidisciplinary audience of faculty and students with the objective of fostering student interest and participation in the STEAM-H disciplines as well as fostering interdisciplinary collaborative research. The series strongly advocates multidisciplinary collaboration with the goal to generate new interdisciplinary holistic approaches, instruments and models, including new knowledge, and to transcend scientific boundaries.

More information about this series at http://www.springer.com/series/15560

Bourama Toni
Editor

The Mathematics of Patterns, Symmetries, and Beauties in Nature

In Honor of John Adam

Springer

Editor
Bourama Toni
Department of Mathematics
Howard University
Washington, DC, USA

ISSN 2520-193X ISSN 2520-1948 (electronic)
STEAM-H: Science, Technology, Engineering, Agriculture, Mathematics & Health
ISBN 978-3-030-84595-7 ISBN 978-3-030-84596-4 (eBook)
https://doi.org/10.1007/978-3-030-84596-4

Mathematics Subject Classification: 93-10, 37N25, 86A15, 35Q91, 97M60

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

Preface

What is mathematics and what do mathematicians do? Let us recall G.H Hardy's excellent description:

> "*A mathematician, like a painter or poet, is a maker of patterns. If his patterns are more permanent than theirs, it is because they are made of ideas.*"

With the unifying aesthetic principle "simple is beautiful," and the simplest possible things are of course imaginary.

This book *The Mathematics of Patterns, Symmetries, and Beauties in Nature* is in honor of Professor John Adam and his outstanding achievements; it features recent developments and techniques by world-renowned experts in the field of mathematics in nature; it will contribute to re-emphasize the relevance and depth of mathematical concepts used to decipher and understand the natural and physical world, including mathematics' ever expanding reach into the physical, biological, social/behavioral, and computational sciences. The volume provides an accessible summary of a wide range of active research topics, along with exciting new results. Topics include: mathematical modeling of the evolution dynamics of Covid-19, simulations of social distancing scenarios, mathematical modeling of tungiasis inflammation of the skin, modeling of a tick-killing robot, spatial-temporal of Covid-19 progression, Archimedean and non-Archimedean mathematical modeling.

The volume's unique feature is to gather in a single expert book the most recent theoretical developments as well as state-of-the-art applications.

It will certainly serve as a useful resource for both graduate students entering this research area and more established researchers, including a wide angle snapshot of this exciting and far-reaching research domain. It also facilitates an in-depth exchange of ideas on recent advances in the various aspects of the Mathematics of Patterns, Symmetries, and Beauties in Nature.

As such the volume is an important part of the multidisciplinary STEAM-H series (Science, Technology, Engineering, Agriculture, Mathematics and Health); the series brings together leading researchers to present their work in the perspective to advance their specific fields, and in a way to generate a genuine interdisciplinary interaction transcending disciplinary boundaries. All chapters therein were carefully

edited and peer reviewed; they are reasonably self-contained and pedagogically exposed for a multidisciplinary readership.

Contributions are invited only and reflect the most recent advances delivered in a high standard and self-contained in line with the goals of the series, that is:

1. To enhance multidisciplinary understanding between the disciplines by showing how some new advances in a particular discipline can be of interest to the other discipline, or how different disciplines contribute to a better understanding of a relevant issue at the interface of mathematics and the sciences.
2. To promote the spirit of inquiry so characteristic of mathematics for the advances of the natural, physical, and behavioral sciences by featuring leading experts.
3. To encourage diversity in the readers' background and expertise, while at the same time structurally fostering genuine interdisciplinary interactions and networking.

Current disciplinary boundaries do not encourage effective interactions between scientists; researchers from different fields usually occupy different academic buildings, publish in journals specific to their field, and attend different scientific meetings. Existing scientific meetings usually fall into either small gatherings specializing on specific questions, targeting specific and small group of scientists already aware of each other's work and potentially collaborating, or large meetings covering a wide field and targeting a diverse group of scientists but usually not allowing specific interactions to develop due to their large size and a crowded program. Here contributors focus on how to make their work intelligible, accessible to a diverse audience, which in the process enforces mastery of their own field of expertise.

This volume strongly advocates multidisciplinarity with the goal to generate new interdisciplinary approaches, instruments, and models including new knowledge, transcending scientific boundaries to adopt a more holistic approach. For instance, it should be acknowledged, following Nobel laureate and president of the UK's Royal Society of Chemistry, Professor Sir Harry Kroto, "that the traditional chemistry, physics, biology departmentalised university infrastructures—which are now clearly out-of-date and a serious hindrance to progress—must be replaced by new ones which actively foster the synergy inherent in multidisciplinarity." The National Institutes of Health and the Howard Hughes Medical Institute have strongly recommended that undergraduate biology education should incorporate mathematics, physics, chemistry, computer science, and engineering until "interdisciplinary thinking and work become second nature." Young physicists and chemists are encouraged to think about the opportunities waiting for them at the interface with the life sciences. Mathematics is playing an ever more important role in the physical and life sciences, engineering, and technology, blurring the boundaries between scientific disciplines.

The series, through contributed volumes such as the current one, is to be a reference of choice for established interdisciplinary scientists and mathematicians and a source of inspiration for a broad spectrum of researchers and research students, graduate and postdoctoral fellows; the sheer emphasis of these carefully selected

and refereed contributed chapters is on important methods, research directions, and applications of analysis including within and beyond mathematics. As such the volume implicitly promotes mathematical sciences, physical and life sciences, engineering, and technology education, as well as interdisciplinary, industrial, and academic genuine cooperation.

The current book, entitled *The Mathematics of Patterns, symmetries, and Beauties in Nature,*" as a whole certainly enhances the overall objective of the series, that is, to foster the readership interest and enthusiasm in the STEAM-H disciplines (Science, Technology, Engineering, Agriculture, Mathematics and Health), stimulate graduate and undergraduate research, and generate collaboration among researchers on a genuine interdisciplinary basis.

The STEAM-H series is hosted at Howard University, Washington DC, USA, an area that is socially, economically, intellectually very dynamic and home to some of the most important research centers in the USA. This series, by now well established and published by Springer, a world-renowned publisher, is expected to become a national and international reference in interdisciplinary education and research.

Washington, DC, USA Bourama Toni
June 2021

Acknowledgments

We are grateful to the many reviewers for their professionalism: they raised many questions and challenged the materials featured in this book with insightful comments and suggestions from concepts to exposition.

We would like to express our sincere appreciation to all the contributors for the excellent research work. They all made this volume a reality for the greater benefit of the community of Science, Technology, Engineering, Agriculture, Mathematics, and Health.

Contents

Contributors

Ephraim Agyingi School of Mathematical Sciences, College of Science, Rochester Institute of Technology, Rochester, NY, USA

Regina C. Almeida National Laboratory for Scientific Computing, Petrópolis, Brazil

Eric P. Choate Radford University, Radford, VA, USA

Norou Diawara Mathematics and Statistics Department, Old Dominion University, Norfolk, VA, USA

Holly D. Gaff Old Dominion University, Norfolk, VA, USA

Dilini Gamage Mathematics and Statistics Department, Old Dominion University, Norfolk, VA, USA

Aboudramane Guiro Département de Mathématiques, UFR Sciences et Techniques, Université Nazi Boni, Bobo-Dioulasso, Burkina Faso

Hueiwang Anna Jeng Community and Environmental Health, Old Dominion University, Norfolk, VA, USA

Blaise Koné LAboratoire de Mathématiques et Informatique (LAMI), Université Joseph Ki-Zerbo, Ouagadougou, Burkina Faso

Gustavo B. Libotte National Laboratory for Scientific Computing, Petrópolis, Brazil

Fran Sérgio Lobato Chemical Engineering Faculty, Federal University of Uberlândia, Uberlândia, Brazil

Sophia Maggelakis School of Mathematical Sciences, College of Science, Rochester Institute of Technology, Rochester, NY, USA

Sandra M. C. Malta National Laboratory for Scientific Computing, Petrópolis, Brazil

Agida G. Manizade Radford University, Radford, VA, USA

Jennifer L. Matthews United States Naval Academy, Annapolis, MD, USA

Stanislas Ouaro LAboratoire de Mathématiques et Informatique (LAMI), Universite Joseph Ki-Zerbo, Ouagadougou, Burkina Faso

Yuliya Petrova National Institute of Health, Bethesda, MD, USA

Gustavo M. Platt Graduate Program in Agroindustrial Systems and Processes, School of Chemistry and Food, Federal University of Rio Grande, Santo Antônio da Patrulha, Brazil

Renato S. Silva National Laboratory for Scientific Computing, Petrópolis, Brazil

Bourama Toni Department of Mathematics, Howard University, Washington, DC, USA

Alexis L. White University of Florida, Gainesville, FL, USA

Tamas Wiandt School of Mathematical Sciences, College of Science, Rochester Institute of Technology, Rochester, NY, USA

Introduction: Nature and Its Mathematics

Bourama Toni

What is mathematics? What do mathematicians do? Do they *discover* or do they *invent* mathematics? Artificial Intelligence (AI) is permeating all realms of human endeavors, prompting the question as to whether mathematical creativity is uniquely human. Let recall G.H. Hardy's excellent description:

> "*A mathematician, like a painter or poet, is maker of patterns. If his patterns are more permanent than theirs, it is because they are made of ideas*"

With the unifying aesthetic principle "simple is beautiful", and the simplest possible things are of course imaginary, some have said that mathematics and poetry carry the same principle and quest of beauty; indeed in both beauty seems to be generated by translation/displacement and unforeseen twists. (e.g. change of perspective casting problems in a new light often leads to a mathematical solution).

Augustus de Morgan (1806–1871) is quoted as saying that *the moving power of mathematical invention is not reason but imagination.* Abstraction is mathematics' primary strength.

Mathematics has been and continues to be used to decipher natural phenomena and nature itself, and to some extend the universe. This volume, titled "the Mathematics of Patterns, Symmetries and Beauties in Nature" in Honor of Professor John Adam aims at contributing to such endeavor: reveal and appreciate nature and its beauty through mathematical lenses. This is truly reflected and invited by this 1857 description of Nature by the French poet Charles Baudelaire, and we quote:

B. Toni (✉)
Department of Mathematics, Howard University, Washington, DC, USA
e-mail: bourama.toni@howard.edu

B. Toni (ed.), *The Mathematics of Patterns, Symmetries, and Beauties in Nature*,
STEAM-H: Science, Technology, Engineering, Agriculture, Mathematics & Health,
https://doi.org/10.1007/978-3-030-84596-4_1

La Nature est un temple où de vivants piliers
Laissent parfois sortir de confuses paroles;
L'homme y passe à travers des forêts de symboles
Qui l'observent avec des regards familiers.

Comme de longs échos qui de loin se confondent
Dans une ténébreuse et profonde unité,
Vaste comme la nuit et comme la clarté,
Les parfums, les couleurs et les sons se répondent.

Il est des parfums frais comme des chairs d'enfants,
Doux comme les hautbois, verts comme les prairies,
—Et d'autres, corrompus, riches et triomphants,

Ayant l'expansion des choses infinies,
Comme l'ambre, le musc, le benjoin et l'encens,
Qui chantent les transports de l'esprit et des sens.

Charles Baudelaire, Sonnet de 1857

Professor John Adam has contributed greatly through his work to exploring patterns in nature, showing us how to bypass the unnecessary complexities of natural phenomena and focus on the essentials. See for instance his book "Guesstimation: solving the world' problems on the back of a cocktail napkin". John has also published "Mathematical Modeling in Unusual Contexts: Rainbows, Halos and Glories." His mathematical interests in deciphering nature's patterns, symmetries and beauties include: snow crystals and their regularity and symmetry; 'Why the Sky is Blue"; The mathematical basis of a universal "law of quarter powers" for critters, large and small; mathematical theory of waves; wound healing in bones; mathematical theory of meteorological optics. Rays, Waves and Scattering was the topic of his most recent book. John's published books on the subject include, by the Princeton University Press: "Mathematics in Nature: Modeling Patterns in the Natural World"; "A Mathematical Nature Walk"; "X and the City: Modeling Aspects of Urban Life" which reveals mathematics in the metropolitan landscape.

Thus everything that is not forbidden by laws of nature is decipherable/explainable, given the right mathematical knowledge, which encompasses approaches grounded in the Archimedean and Non-Archimedean Principle.

In chapter "A mathematical model of thermography with application to tungiasis inflammation of the skin" by Ephraim Agyingi, Tamas Wiandt and Sophia Maggelakis, the authors use the Pennes equation to model the heat transfer in tungiasis-associated inflammation of the skin, and provide numerical simulations for the presence of tungiasis and other bacterial co-infection.

In chapter "Application and Modeling of a Tick-Killing Robot, TickBot" by Alexis White and Holly Gaff, the authors model mathematical tick population dynamics, quantify risk of tick-borne disease, and identify strategies to reduce that risk. They also explore from field studies data the use of a TickBot as a tick management tool, leading to supporting the use of scenarios with carbon-dioxide for greater reduction in questing tick populations.

Chapter "Mathematics of Covid-19" by Eric Choate, Agida Manizade and Yuliya Petrova introduces the general framework of "compartment models" used to model

epidemics as a first-order system of nonlinear ordinary differential equations in the context of the COVID-19 pandemic. The resulting model is then adapted to account for the rollout of a vaccination plan and its effect on the expected deaths.

In chapter "Spatio-temporal Modelling of Progression of the COVID-19 Pandemic" the authors Dilini Gamage, Jennifer Matthews, Norou Diawara and Hueiwang Anna Jeng seek and present suitable statistical models (Bayesian Conditional autoregressive (CAR), Moran measures) that can be used to predict progression of the COVID-19, for instance, allowing to monitor the spread of the virus in the targeted locations and time. Data on new COVID-19 cases data on four selected countries (and neighboring countries) over time were collected via Situation Reports published by the World Health Organization.

Chapter "Simulations of Social Distancing Scenarios and Analysis of Strategies to Predict the Spread of Covid-19" by Fran Lobato, Gustavo Libotte, Gustavo Platt, Regina Almeida, Renato Silva and Sandra Malta, presents a comprehensive view of covid-19 induced social distancing measures along two different strategies for predicting the spread of Covid-19 in the scope of computational modeling and to analyze scenarios to describe the impact of social distancing measures.

In chapter "Mathematical modeling of the evolution dynamics of the coronavirus disease 2019 (COVID-19) in Burkina Faso" by Aboudramane Guiro, Blaise Koné and Stanislas Ouaro, the authors develop a compartmental model adapted from real data of the country Burkina Faso along the categories of hospitalized, severely hospitalized patients, and dead persons, determine condition for a locally asymptotically stable disease-free equilibrium. As a result they could predict the disease dynamic to assist with national decision-making policies.

Chapter "Archimedean and Non-Archimedean Approaches to Mathematical Modeling" by Bourama Toni, first discusses cultural approaches to mathematical fundamental concepts, in particular the French and Russian schools of mathematics, as well as the artificial mathematical creativity to lead to the so-called post-human mathematics; the author then describes and illustrates the two main approaches to mathematical modeling; the classic Archimedean methodology featuring the Jacobian Feedback Loops Analysis and the most recent non-Archimedean/ultrametric/p-adic approach.

A Mathematical Model of Thermography with Application to Tungiasis Inflammation of the Skin

Ephraim Agyingi, Tamas Wiandt, and Sophia Maggelakis

1 Introduction

Many body infections and inflammations usually result in higher body temperatures. Among these is Tungiasis, an ectoparasitic infection caused by *Tunga penetrans* (female sand fleas). Tungiasis begins when a female sand flea in an attempt to reproduce, burrows entirely into the skin with its head, leaving only the posterior end which consists of the anus for wasted disposal, the copulatory organs for eggs dispersal, and the stigmata for respiration outside the epidermis. On entering the host, the sand flea starting with a maximum length of 1 mm develops and increases its volume by a factor of about 2,000–3,000 through several stages (ranging from I-V) into a pea-size white halo (of diameter up to 1 cm) spotting a black posterior end for the outlet [1]. The parasitic lesion often dies after a period of up to 5 weeks when all eggs have been dispensed, leaving a wound crater at the affected skin site which may become colonized leading to other forms of skin infections.

Tunga penetrans infestation is common to humans, domesticated animals, and many other general hosts [2]. However, pigs have been identified as the most important reservoir for Tunga penetrans [3, 4]. Although Tungiasis is not featured on the prominent list of neglected tropical diseases, it is prevalent in tropical and subtropical regions of the world, including Mexico to South America, the West Indies and Africa [5, 6]. Tungiasis is endemic where living conditions are unhygienic and also mostly affects the poor, especially children and the elderly. Clinical symptoms include itching, burning irritation, inflammation and ulceration

E. Agyingi · T. Wiandt · S. Maggelakis (✉)
School of Mathematical Sciences, College of Science, Rochester Institute of Technology, Rochester, NY, USA
e-mail: sxmsma@rit.edu; eoasma@rit.edu; tiwsma@rit.edu

B. Toni (ed.), *The Mathematics of Patterns, Symmetries, and Beauties in Nature*, STEAM-H: Science, Technology, Engineering, Agriculture, Mathematics & Health,
https://doi.org/10.1007/978-3-030-84596-4_2

leading to severe pain at the infection site. Recurring multiple lesions in the feet can lead to morbidity and consequently limited mobility. In poor communities, diagnosis is restricted to visual inspection and treatment is often crudely effected using sharp pointed objects (e.g. sharpened stick, needle etc.) to pry out the parasite from the skin.

In some cases, visual inspections even when detailed, often miss diagnosing Tungiasis infections especially those in early stages and those associated with acute morbidity. A very heavy skin pigmentation can also hinder visual diagnosis. Thermography as an additional tool will/may lead to effective diagnosis of Tungiasis infection. The procedure is noninvasive and uses an infrared camera to capture the thermal image (made up of the temperature at the surface of the skin) of the concerned region. Thermography is an established procedure and has been considered in screening several diseases including breast cancer [7], skin cancer [8], veterinary medicine [9, 10], and allergy [11]. Most recently, Schuster et al. [12] clinically considered thermography as a tool in diagnosing tungiasis-associated inflammation of the skin. They concluded that it was a useful tool and that it may help to diagnose hidden and atypical manifestations of tungiasis.

The general neglect given to tungiasis infections is also reflected in the mathematical literature where it has been given only little attention focusing on its transmission dynamics [13–16]. Mathematical models of thermography for breast tumors [17–21] and skin lesions [22–26] have been considered more often in the published literature. In this paper, we consider a mathematical model of thermography for tungiasis infection diagnosis. Following the approach in [26], we model tungiasis as a lesion and study the temperature profile generated at the skin surface using a variant of the Pennes equation for heat transfer. Tungiasis can lead to other bacterial infections during its tenure or after it dies. In view of this, we study coinfection, that is, tungiasis along with another bacteria and also the case where there is bacterial infection of the skin region following the death of the parasite.

The paper is organized as follows: we present the mathematical model for heat transfer in Sect. 2. The model provided here differs from that in [26] in that the different layers of the skin have not been taken into account. For simplicity, we have assumed a homogenous skin region. In section 3, we provide numerical simulations of the model for different degrees of inflamatory response to tungiasis infection and coinfection with bacterial. Closing remarks are stated in the Conclusion.

2 Mathematical Model

We present a model of heat transfer in tungiasis; we are considering a burrowed flea located in the dermis. Following the formulation presented in [25], we construct a mathematical model of heat transfer in a two-dimensional cross-section of the skin. The goal is to investigate whether skin temperature is a good indicator of the presence of tungiasis and whether the mathematical model fits the observed

temperature differences in case of the disease. We study the effects of extra heat production of inflammations on the surface temperature.

We assume that the burrowed fleas create spherical holes in the skin, and in the interest of simplicity, we shall present a spatially two-dimensional model. We consider a two-dimensional cross section of the skin tissue, containing a circularly shaped hole containing the burrowed flea. The temperature of the cross-sectional skin tissue is given by $T(x, y)$, where the x coordinate is the horizontal direction and y is the depth. The two dimensional cut is placed in the (x, y)-plane so that the $y = 0$ level corresponds to the bottom layer of the skin, and $y = d$ corresponds to the surface. The origin is placed so that $-a \leq x \leq a$, with the center of the hole created by the burrowed flea on the y-axis $x = 0$. We assume that there is a layer of inflamed tissue around the hole created by the flea.

The heat flow in this region is described by the Pennes equation [27]

$$\rho\bar{c}\frac{\partial T}{\partial t} = \nabla \cdot (K\nabla T) + m_b c_b(T_A - T) + S,$$

where ρ is the tissue's density, $\bar{c}$ is the tissue's specific heat, K is the tissue's thermal conductivity, m_b is the mass flow rate of blood, c_b is the blood's specific heat, T_A is the arterial blood temperature, and S is the metabolic heat generation.

We consider the steady state temperature, where the time derivative is zero. The equation describing the temperature $T_h(x, y)$ in the healthy region is given by

$$\Delta T_h = -\frac{S_h}{K_h} - \frac{m_b c_b(T_A - T_h)}{K_h},$$

where S_h is the metabolic heat generation rate of the healthy tissue and K_h is its thermal conductivity.

The equation is analogous for the inflamed region, with the heat generation rate of sand flea S considerably higher.

The boundary conditions of the various parts of the domain considered are listed below. The outer boundary conditions of the healthy region are provided by (i)−(iii). Items (iv)−(v) provide inner boundary conditions for the healthy region and outer boundary conditions for the region containing the flea.

(i) $T_h(x, 0) = T_b$ is the temperature of the body at the bottom layer.
(ii) At the skin surface, $y = d$, $-K\frac{\partial T_h}{\partial y} = \lambda(T_h - T_a)$, where T_a is the ambient temperature and λ is the surface heat transfer coefficient.
(iii) At the side boundaries $x = \pm a$, $T_h(-a, y) = T_h(a, y)$ is the temperature distribution of healthy tissue. By assuming homogeneity of the tissue, we reduce the problem into a one-dimensional equation of the form

$$\Delta T_h(x, y) = T_h''(y) - \frac{m_b c_b T_h}{K_h} = -\frac{S_h}{K_h} - \frac{m_b c_b T_A}{K_h},$$

where y is the tissue depth, $T_h(0) = T_b$, and $T_h'(d) = -\lambda(T_h - T_a)/K_h$. The solution of this equation is given by

$$T_h(y) = T_b - \eta + \eta\cosh(\omega\, y) + \left(\frac{\lambda(\eta + T_a - T_b) - \eta K_h\omega\sinh(\omega d) - \eta\lambda\cosh(\omega d)}{\lambda\sinh(\omega d) + K_h\omega\cosh(\omega d)}\right)\sinh(\omega\, y),$$

where $\eta = T_b - \frac{S}{m_b c_b} - T_A$ and $\omega = \sqrt{\frac{m_b c_b}{K_h}}$.

(iv) We assume that the temperature is continuous across the interface of the healthy tissue, the inflamed region, and the burrowed flea.

(v) The heat flux is continuous across the interface of the healthy tissue, the inflamed region, and the burrowed flea.

3 Results and Discussion

In this section, we present numerical simulations of the model presented above using pdetool, a MATLAB finite element solver. The results are for a single sand flea embbeded in a skin cross section. The values of the thermophysical parameter for healthy skin tissue were chosen within the range of published data [23, 25, 26] and are set $S_h = 0.009\,\text{W/cm}^3$, $K_h = 0.0042\,\text{W/((cm)°C)}$, $\lambda = 0.0005\text{W/((cm}^2\text{)°C)}$, $m_b = 0.0005\,\text{g/(ml s)}$ and $c_b = 4.2\,\text{J/g°C}$. The value of the arterial blood temperature was set at $T_A = 36°\text{C}$ and the body temperature, at the level $y = 0$, was set at $T_b = 36°\text{C}$. Other parameters used were chosen to investigate the behavior of the model.

By burrowing into the infection site, the parasite injure tissue cells and as a consequence triggers an inflammatory response. We consider cases with varying degrees of inflammatory response to the parasite. We assume that the generation of inflammatory mediators in the vicinity of the parasite leads to an increased rate of energy production. We therefore associate a high inflammatory response with higher heat generation. We also assume that the inflammatory response is uniformly distributed in the region surrounding the parasite. In all simulations, we assign the heat generation rate of the inflamed region a value that is a multiple of the heat generation rate of healthy tissue. Because of the rapid rate of development and the accompanying metabolic activity, we also assign the heat generation rate of the parasite a value that is a multiple of the heat generation rate of healthy tissue.

The results given in Fig. 1 are for the steady state temperature distribution within the skin cross section where the heat generation rate of healthy tissue is $S_h = 0.009\,\text{W/cm}^3$ and where the parasite heat generation rate is $10S_h$. The heat generation rate of the inflammation region is set at $5S_h$, $10S_h$, and $20S_h$ as depicted in Fig. 1a, b, c, respectively. The temperature distribution as shown in all cases is significantly higher in the proximity of the parasite and the intensity increases with an increase in the heat generation rate of the inflamed region. Figure 1d, simulates

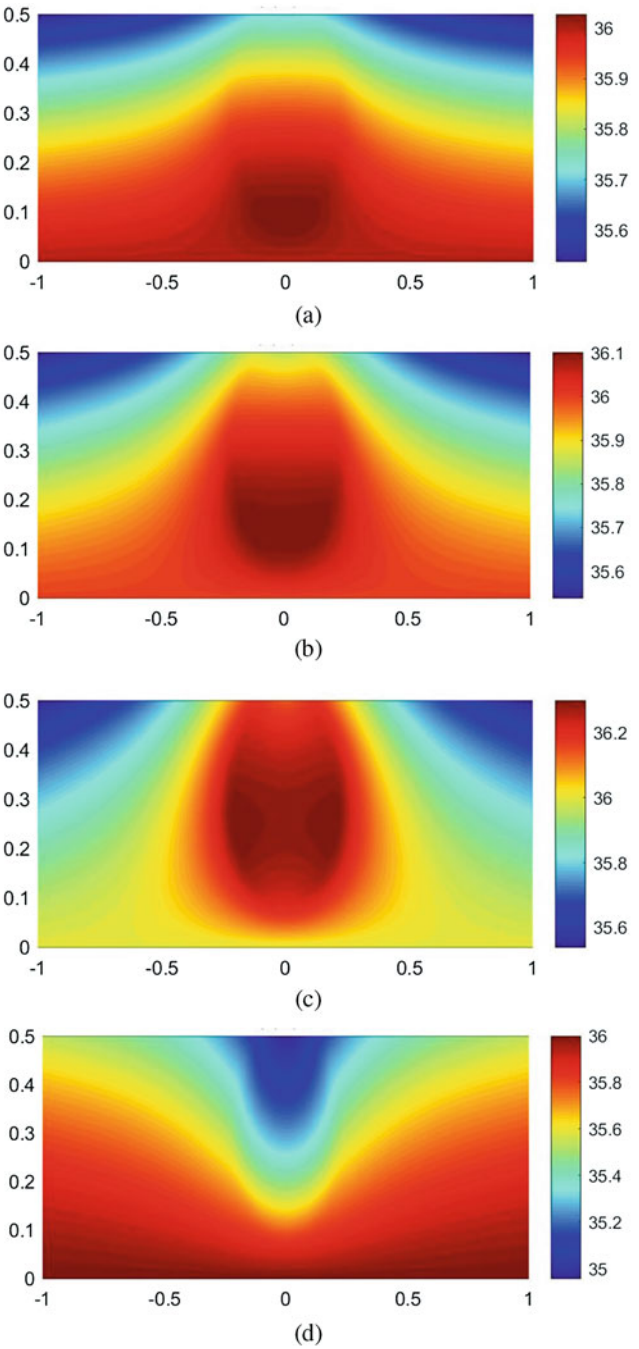


Fig. 1 Steady state temperature distribution within the skin section where the heat generating rate of inflamatory response is: (**a**) 5 times that of healthy tissue; (**b**) 10 times that of healthy tissue; and (**c**) describes 20 times that of healthy tissue. In (**d**) a non-viable flea is simulated

the steady state temperature of the skin section when the parasite is in stages IV-V, where it is no longer viable and eventually dies. Here, we maintain a heat generation rate of $10S_h$ for the inflammatory response and set that of the parasite to zero since it consist of dead cells.

A change in the steady state temperature within the skin section also results in a change in the steady state temperature at the skin surface as illustrated in Fig. 2. An increased temperature at the skin surface is observed for different inflammatory responses in the presence of the parasite. The biggest increase of about 0.6°C occurs directly above the parasite and is associated with the highest inflammatory response. Figure 2 also shows a decrease in the temperature at the skin surface over the dead parasite. We remark here that every sand flea upon death leaves behind a carcass of dead tissue at the infection site. A round depression is often formed in the outer layer of the skin after the carcass of the sand flea is expelled and the epidermis is reconstructed in a timely manner assuming no underlying issue such as bacterial infection is present.

Because cases do arise in which an abscess is formed around a tungia lesion [28], we next simulate sand flea and bacterial coinfection. In this case we also assumed that the metabolic activity by the bacterial infection leads to a higher heat generation rate compared to that of healthy skin tissue. Further, we supposed that inflammatory response is directed at both the sand flea and bacteria, and is contained within the

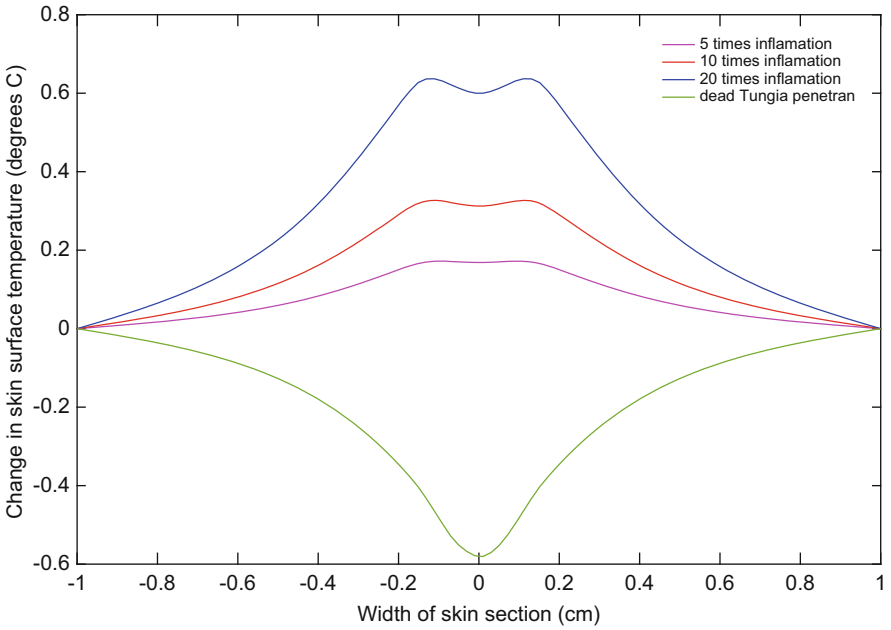


Fig. 2 Numerical simulation of the change in steady state temperature at the skin surface for various degrees inflamatory response. The degree of the inflamatory response corresponds to a multiple of the heat generating rate of healthy skin tissue

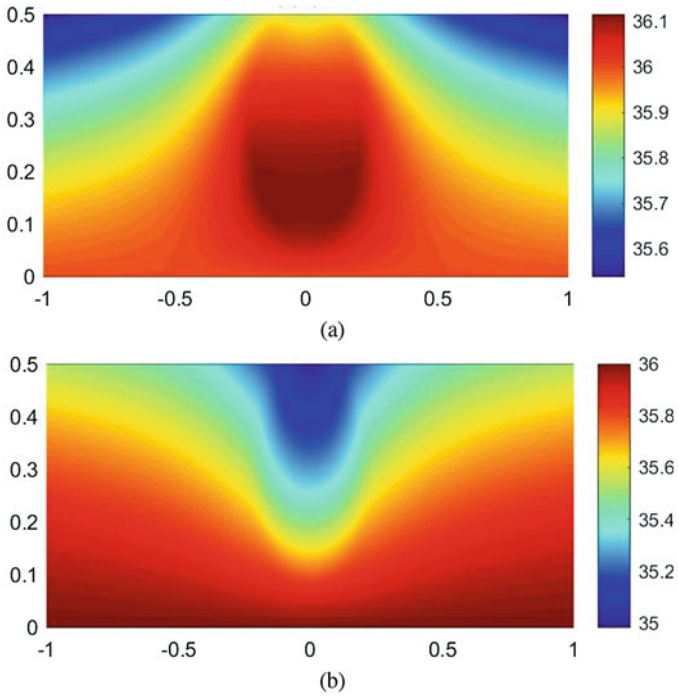
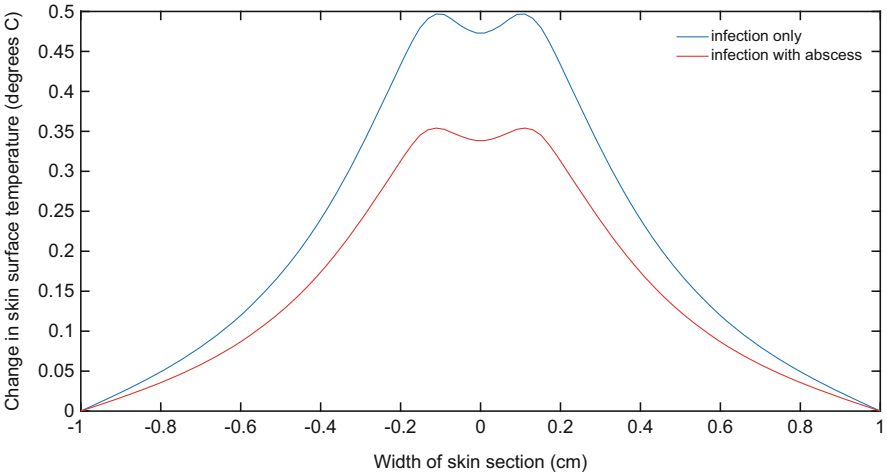


Fig. 3 Coinfection temperature distrubution within skin cross section: (**a**) temperature distribution for bacterial infection with live tungia penetran; and, (**b**) temperature distribution for bacterial infection with non-viable tungia penetran

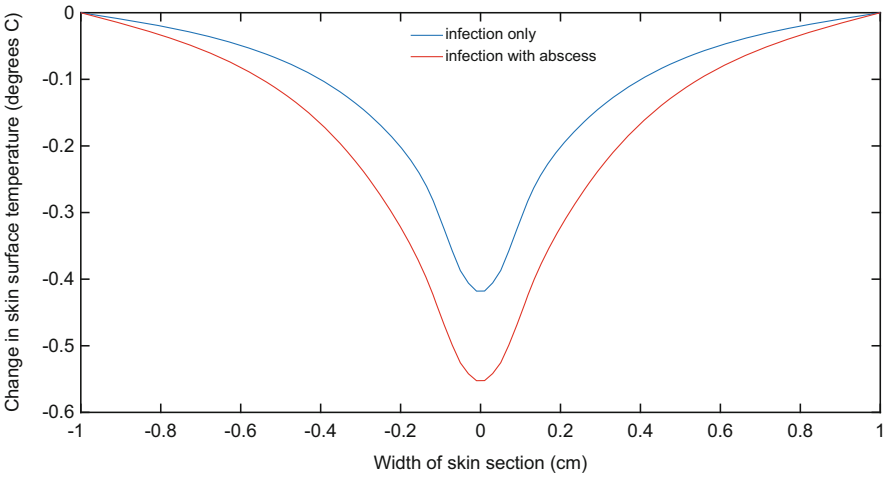
vicinity of the burrowed flea. Coinfection results are presented in Fig. 3, where the steady state temperature of the skin section is given in Fig. 3a for a viable sand flea, while Fig. 3b illustrates the steady state temperature of the skin section for a non-viable (that is dead) sand flea.

Some bacterial infections are very severe leading to the formation of an abscess. The results in Fig. 4 provide the change in the steady state temperature observed at the skin surface for two types of coinfections: one without and the other with abscess formation. The simulations in Fig. 4a are for a viable sand flea while that in Fig. 4b is for a non-viable sand flea given the presence of bacterial infection. We observe from Fig. 4a that the skin surface temperature is higher where an abscess is absent for a viable flea. The lower temperature in the case of an abscess can be attributed to the fact that its content is mostly made up of dead cells and as a consequence has a lower heat generating rate. In the case of a non-viable flea, Fig. 4b, the temperature at the skin surface decreases the most where coinfection with an abscess is present.

The results presented above clearly show that a significant increase or decrease in temperature at a region on the skin surface is a diagnostic indication of an internal underlying abnormality. An icreased temperature is associated with a viable burrowed sand flea, while a decreased temperature is associated with bacterial coinfection leading to the formation of an abscess after the death of the sand flea.



(a)



(b)

Fig. 4 Change in skin surface steady state temperature for coinfection: (**a**) bacterial infection with life tungia penetran; and, (**b**) bacterial infection with non-viable tungia penetran

4 Conclusion

We have presented a model of heat transfer in tungiasis-associated inflammation of the skin based on the Pennes bio-heat equation. We used the model to simulate the steady state temperature of a 2 dimensional cross section of the skin containing a burrowed sand flea; with and without bacterial coinfection. We also considered cases in which the sand flea is non-viable, representing the final stages IV-V of

the life cycle of a borrowed flea. We associated various degrees of inflammatory responses whether directed to the sand flea or bacteria with a multiple of the heat generating rate of healthy skin tissue. Numerical simulations show that the steady state temperature at the skin surface increases most with a very strong inflammatory response for a viable sand flea and is lower for coinfection with abscess formation. In the case of a non-viable flea, the temperature at the skin surface at the infection site is generally lower compared to that of the surrounding region, with the largest decrease observe for coinfection resulting in abscess formation. The latter result provides an insight to diagnosing not only tungiasis related coinfection but also other underlying abnormalities using thermography.

References

1. P. M. Linardi and D. Moreira de Avelar, Neosomes of tungid fleas on wild and domestic animals, *Parasitol Res* **113** (2014) 3517–3533
2. D. M. De Avelar, Sistematica e analise cladostica dasespecies neotropicais do genero Tunga Jarocki, 1838 (Si-phonaptera: Tungidae), Tese de Doutoramento, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil (2010).
3. S. Pampiglione, M. Trentini, F. Mattei Gentili, J.L.X.Mendes, C. Pampiglione, and F. Rivasi, Tunga penetrans(Insecta: Siphonaptera) in pigs in Sao Tome (equatorial Africa): epidemiological, clinical, morphologicaland histopathological aspects. *Revue Elev. Med. Pays Trop* **51** (1998) 201–205.
4. S. Ugbomoiko, L. Ariza, and J. Heukelbach, Pigs arethe most important animal reservoir for Tunga penetrans (jigger flea) in rural Nigeria. *Trop. Doct.* **38** (2008) 226–227.
5. J. Heukelbach, F. A. S. De Oliveira, G. Hesse, H. Feldmeier, Tungiasis: a neglected health problem of poorcommunities, *Trop. Med. Int. Health* **6(4)** (2001) 267–272.
6. J. Heukelbach, T. Wilcke, G. Harms, H. Feldmeier, Seasonal variation of tungiasis in an endemic community, *Am. J. Trop. Med. Hyg.* **72(2)** (2005) 145–149.
7. N. Arora, D. Martins, D. Ruggerio, E. Tousimis, A. J. Swistel, M. P. Osborne, and R. M. Simmons, Effectiveness of a noninvasive digital infrared thermal imaging system in the detection of breast cancer, *The American Journal of Surgery* **196** (2008) 523.
8. A. L. Shada, L. T. Dengel, G. R. Petroni, M. E. Smolkin, S. Acton, and C. L. Slingluff Jr., Infrared thermography of cutaneous melanoma metastases, *J. Surg. Res.* **182** (2013) e9.
9. M. Sathiyabarathi, S. Jeyakumar, A. Manimaran, G. Jayaprakash, H. A. Pushpadass, M. Sivaram, K. P. Ramesha, D. N. Das, M. A. Kataktalware, M. A. Prakash & R. D. Kumar, Infrared thermography: A potential noninvasive tool to monitor udder health status in dairy cows, *Veterinary world* **9(10)** (20161) 1075–1081.
10. K. Rainwater-Lovett, J. Pacheco, C. Packer &L. Rodriguez, Detection of foot-and-mouth disease virus infected cattle using infrared thermography *Veterinary journal (London, England : 1997)* **180** (2008) 317–24.
11. T. Rok, E. Rokita, G. Tato, T. Guzik & T. Sliwa, Thermographic imaging as alternative method in allergy diagnosis, *Journal of Thermal Analysis and Calorimetry* **127** (2017) 1163–1170.
12. A. Schuster, M. Thielecke, V. Raharimanga, C. E. Ramarokoto, C. Rogier, I. Krantz and H. Feldmeier, High-resolution infrared thermography: a new tool to assess tungiasis-associated inflammation of the skin, *Tropical Medicine and Health* **45** (2017) 23
13. R.l A. Nyang'inja, D. N. Angwenyi, C. M. Musyoka and T. O. Orwa, Mathematical modeling of the effects of public health education on tungiasis a neglected disease with many challenges in endemic communities, *Advances in Difference Equations* **2018** (2018) 426.

14. N. I. Matendechere, W. N. Mutuku, F. Nyabadza, Modelling the dynamics of jiggers infestation: insights from a theoretical model, *J. Math. Comput. Sci.* **9(4)** (2019) 473–500.
15. F. K. Mbuthia and I. Chepkwony, Mathematical Modelling of Tungiasis Disease Dynamics Incorporating Hygiene as a Control Strategy, *Journal of Advances in Mathematics and Computer Science* **33(5)** (2019) 1–8.
16. N. I. Matendechere, W. Mutuku, F. Nyabadza, A mathematical model for jiggers infestation incorporating effects of media campaigns, *Commun. Math. Biol. Neurosci.* **2019** (2019) 25.
17. Z. Deng and J. Liu, Mathematical modeling of temperature mapping over skin surface and its implementation in thermal disease diagnostics, *Computers in Biology and Medicine* **34** (2004) 495–521.
18. E. Agyingi, T. Wiandt, S, Maggelakis, Thermal detection of a prevascular tumor embedded in breast tissue, *Mathematical Biosciences and Engineering* **12** (2015) 907–915.
19. Q. Y. Lin, H. Q. Yang, S. S. Xie, Y. H. Wang, Z. Ye, and S. Q. Chen, Detecting early breast tumour by finite element thermal analysis, *Journal of Medical Engineering & Technology* **33** (2009) 4, 274–280.
20. M. Mital and E. P. Scott, Thermal Detection of Embedded Tumors Using Infrared Imaging, *Journal of Biomechanical Engineering*, **129** (2007) 33–39.
21. M. Pirtini Cetingl, C. Herman, A heat transfer model of skin tissue for the detection of lesions: sensitivity analysis, *Phys. Med. Biol.* **55** (2010) 5933–5951.
22. M. Pirtini Cetingl, C. Herman, Quantification of the thermal signature of a melanoma lesion, *International Journal of Thermal Sciences* **50** (2011) 421–431.
23. M. Bonmarin e F.-A. Le Gal, Lock-in thermal imaging for the early-stage detection of cutaneous melanoma: A feasibility study, *Comput. Biol. Med.* **47** (2014) 36–3.
24. A. Bhowmik, R. Repaka, S.C. Mishra, Thermographic evaluation of early melanoma within the vascularized skin using combined non-Newtonian blood flow and bioheat models, *Computers in Biology and Medicine* **53** (2014) 206–219.
25. E. Agyingi, T. Wiandt, S, Maggelakis, A quantitative model of cutaneous melanoma diagnosis using thermography, in Mathematical and Computational Approaches in Advancing Modern Science and Engineering, Blair, J., Frigaard, I.A., Kunze, H., Makarov, R., Melnik, R., Spiteri, R.J. (Eds.), Springer International Publishing Switzerland (2016).
26. E. Agyingi, T. Wiandt, S, Maggelakis, Thermal profiling of the evolution of cutaneous melanoma with multilayered computational models"; with Tamas Wiandt & Sophia Maggelakis *Journal of Coupled Systems and Multiscale Dynamics* Volume **4** (2016) 153–161.
27. H. H. Pennes, Analysis of tissue and arterial blood temperatures in the resting forearm, *J. Appl. Physiol.* **1** (1948) 93–122.
28. M. Eisele, J. Heukelbach, E. Van Marck, H. Mehlhorn, O. Meckes, S. Franck, H. Feldmeier, Investigations on the biology, epidemiology, pathology and control of Tunga penetrans in Brazil: I. Natural history of tungiasis in man, *Parasitol Res* **90** (2003) 87–99

Mathematics of COVID-19

Eric P. Choate, Agida G. Manizade, and Yuliya Petrova

1 Introduction

Severe Acute Respiratory Syndrome Corona Virus 2 is the official name of the virus responsible for the COVID-19 pandemic. This name was announced on February 11, 2020, and was chosen because the virus is genetically related to the coronavirus responsible for the SARS outbreak in 2003 [17]. On December 31, 2019, an outbreak of unknown etiology in Wuhan, Hubei Province, China, was identified by the authorities as a novel coronavirus, which has started a worldwide pandemic with significant casualties and economic implications. In response, the world's scientists united their efforts to battle this newfound threat. Several research groups and teams have made significant progress in developing and implementing vaccinations and treatments against COVID-19.

2 Development of the Vaccine

Messenger Ribonucleic Acid (mRNA) is a molecule similar to DNA. Unlike DNA, it is single-stranded and RNA strands have a backbone made of sugar and phosphate groups. mRNA vaccines operate by providing human cells with a piece of Spike protein—a protein found on the surface of the virus that causes COVID-19—that has

E. P. Choate · A. G. Manizade (✉)
Radford University, Radford, VA, USA
e-mail: echoate2@radford.edu; amanizade@radford.edu

Y. Petrova
National Institute of Health, Bethesda, MD, USA
e-mail: petrova.yuliya@nih.gov

B. Toni (ed.), *The Mathematics of Patterns, Symmetries, and Beauties in Nature*,
STEAM-H: Science, Technology, Engineering, Agriculture, Mathematics & Health,
https://doi.org/10.1007/978-3-030-84596-4_3

been rendered harmless and delivered into the body through the upper arm muscles. mRNA acts as instructions for the immune cells that allows them to identify and protect against viruses that display this protein on their surface [18].

By December 2020, there were two new mRNA vaccines for COVID-19 that had completed the required human trials. These were mRNA-1273, a vaccine developed at the National Institute of Health (NIH) and produced by Moderna Pharmaceuticals, and Tozinameran, the vaccine developed by Pfizer-BioNTech. Other labs have developed COVID-19 vaccines using different technologies, included but not limited to the University of Oxford's DNA-based vaccine, which was made in collaboration with AstraZeneca Pharmaceuticals, the recombinant protein Sanofi vaccine, and Sinovac's CoronaVac vaccine [1–5]. Because the virus originated in China, many vaccines were developed there using different platforms, although the research data and information regarding the development of vaccines is still limited. In this section, we will discuss the development of the vaccine and the mathematics involved in the development process. We will be focusing primarily on the Moderna and Pfizer mRNA vaccines.

2.1 Structure of the Virus and How It Enters Human Cells

The COVID-19 virus has a lipid membrane, on top of which there are various proteins. One of the most important proteins is the Spike protein. The image most commonly presented in media of the COVID-19 virus is a red ball with spikes on it. This is a simplified image of the virus and the spikes represent the Spike protein. An important characteristic of any virus is that it cannot multiply by itself. In order to do so, the virus has to merge with a living cell and use the cell to create more of the virus. However, entering a human cell is no simple task—in order for the COVID-19 virus to penetrate a human cell membrane, it must first go through multiple layers of protection such as proteins, sugars, etc. In the case of COVID-19, the Spike protein (located on the virus) attaches itself to an angiotensin-converting enzyme (ACE) receptor/protein located on the outer membrane of human cells. Once the virus attaches itself to the ACE receptor on the cell, it goes through a mechanical process in which it lifts the membrane of the cell like a lever and merges with the human cell. This way the mRNA of the virus is able to enter the cell where it can multiply.

Spike protein exists on virus surfaces in two conformationally different forms: (1) pre-fusion, which is less stable, and (2) post fusion, which is more stable. A pre-fusion Spike protein is able to attach itself to ACE receptors on the human cell. A post-fusion Spike protein is not able to attach itself to ACE receptors on the human cell.

For the development of the vaccine, most researcher teams focused on the pre-fusion Spike protein.

2.2 Immuno-Response to the Virus

Immuno-response is the body's ability to neutralize potentially pathogenic organisms, agents, and substances, including the SARS-CoV-2 Virus. When SARS-CoV-2 infects the body, humans develop antibodies for the Spike proteins. It is important to note that the antibodies that correspond to post-fusion Spike proteins do *not* neutralize the virus, meaning they do not stop pre-fusion Spike protein from attaching to ACE receptors on a human cell and entering the cell. The function of antibodies is not to kill the virus, but to attach themselves to their corresponding proteins on the surface of the virus, similar to a cork in a bottle, effectively blocking the virus's ability to penetrate the cell and multiply.

2.3 The Vaccine

The goal of an effective vaccine is to create a large or significant number of antibodies that could attach themselves to the pre-fusion Spike proteins that have entered the human body. By doing so, the human body is prepared to attack any pre-fusion Spike proteins it may encounter in the future. This means if the human body is exposed to the actual SARS-CoV-2 virus, which uses pre-fusion Spike proteins as the "key" to penetrate a human cell, it will already have defensive antibodies designed to protect itself from this specific characteristic of the virus. Therefore, the human cells will be protected from the virus itself.

In the development of the mRNA-1273 vaccine for the COVID-19 pandemic, the researchers coded the mRNA by studying the 3D structure of Spike protein and identifying its nucleic acid genetic sequence that supports the protein in pre-fusion form. Researchers at the National Institute of Health (NIH) coded, modified, and stabilized pre-fusion Spike protein by changing only two amino acids. To deliver the vaccine into a human body, they enclosed the newly developed nuclear acid sequence in a saline solution enclosed within liposomes. These fatty spheres that contain a stable, modified mRNA vaccine within them enters a human cell and prompts the body to produce modified viral protein and, as a result of immune response, massive amounts of antibodies specifically designed to fight pre-fusion Spike proteins.

Both the Moderna and Pfizer vaccines used this technology to develop their product. The difference between them is that of the temperature in which they need to be stored and the timeline allowed for the injection. The Moderna vaccine can be stored in a regular freezer while the Pfizer vaccine requires special freezers with temperatures lower than $-70\ ^{\circ}\mathrm{C}$. This difference is due to the endurance of these vaccines against Ribonuclease (RNase)—a type of nuclease that catalyzes the degradation of RNA into smaller components. It is our body's defense mechanism against all foreign genetic code. The nucleotides in the Pfizer Vaccine are more susceptive to RNase, whereas the nucleotides in the Moderna Vaccine are modified

to be less susceptive to RNase and therefore more stable at reasonable temperatures. This difference causes different requirements and challenges for the transportation and storage of the vaccine.

AstraZeneca produced a different type of vaccine, which was developed in the University of Oxford. The difference between this vaccine in comparison to the Moderna Vaccine by NIH or the Pfizer Vaccine is that it is a DNA based vaccine instead of mRNA based and uses an adeno-viral vector as the DNA delivery tool. In this case, there is an extra step involved because in order to produce a protein from DNA, there needs to be a transition to RNA. This requires additional time and energy from the human cell. A big advantage of DNA-based vaccines is that it is substantially more stable than those that are RNA-based and doesn't require deep freezing for storage and transportation.

Yet another approach to the development of a SARS-CoV-2 vaccine is being used by the French pharmaceutical company, Sanofi. They are producing a protein based vaccine, which created significant challenges with isolating the protein during its production process. This approach is also more expensive and time-consuming. The research team has delayed on reporting their success.

A few Chinese teams, including Sinovac and Sinopharm companies that worked on producing a COVID-19 vaccine, chose a traditional approach for developing a vaccine, which entails injecting samples of the dead virus into a vaccine with the expectation that it will illicit an immune response to create antibodies that will neutralize the virus. The problem with this approach in the case of SARS-CoV-2 is that, based on previous research on other corona viruses, this type of vaccine would mostly generate antibodies to the post-fusion Spike protein instead of the pre-fusion Spike protein form, because the Spike protein on a dead virus exists in its post-fusion form only. This means that when the live virus is introduced to a human body, the post-fusion Spike protein antibodies present do not protect the body against the pre-fusion proteins that are used by the virus to penetrate the cell. This hypothesis is in agreement with recently published efficiency data for Sinovac's vaccine clinical trial phase 3, that was carried out in Brazil with an efficiency of 50.4% compared to Oxford's vaccine's about 70% and about 95% for the two mRNA vaccines [1, 6].

3 The Basic SIR Model

Compartment models for epidemics have their origins in the work of [7]. Brief explanations of this model and other more complicated models may be found in [8–12]. Let N be the fixed total population of a community experiencing the outbreak. A community can be a country, but it could also be a smaller community such as a rural town, a school, an assisted living community, or a prison.

In the simplest compartment model, the Susceptible-Infectious-Recovered (SIR) model, at any time t, the population of the community is split among three groups or "compartments":

Fig. 1 A diagram of the two processes—infection and recovery—by which people move from one compartment to another

- Susceptible S: People who do not have the disease and have no immunity to it.
- Infectious I: People who have an active infection that can spread to susceptible people.
- Recovered R: People who have been infected but are no longer at risk of infecting others and no longer susceptible to becoming infected. This group is sometimes called "resistant" because the previous infection provides immunity, but it is also sometimes called "removed" because it would also include people who have died from the disease.

Starting with an initial small fraction of infected people, initially no one is recovered, and everyone else susceptible, a dynamical system then describes how people move from one compartment to another at different rates:

$$\frac{dS}{dt} = -\beta\frac{IS}{N}, \tag{1}$$

$$\frac{dI}{dt} = \beta\frac{IS}{N} - \gamma I, \tag{2}$$

$$\frac{dR}{dt} = \gamma I. \tag{3}$$

This simple model includes only two driving processes—infection and recovery—as depicted in Fig. 1. In the original work of Kermack and McKendrick, the rates of these two processes depend on the distribution of ages of the susceptible or infectious populations, but in practice, most often models rely on the simplifying assumptions that these rates are the same for all ages.

The recovery rate γ is the reciprocal of the average time required for an infected person to recover, or equivalently, the number of people per time period who cease to be infectious. It appears twice in the system to account for the transfer of people from infectious to recovered: the $-\gamma I$ term in Eq. (2) represents the rate at which infected people cease to be infectious, and its counterpart γI in Eq. (3) adds these people to those who have already recovered.

The second process—infection—is more complicated to model. Recovery happens on an individual basis. Some may recover faster than others due their health conditions prior to infection and access to available medical treatments, but the recovery rate is assumed to be unaffected by the number of people who are either infectious or recovered. The same is not true for the rate of new infections, which

must be proportional to both the number of infectious people and also the number of susceptible people because if everyone has been infected, no one is left to become infected. Thus, we have the nonlinear term $-\beta\frac{IS}{N}$ in Eq. (1), which can be seen from the susceptible perspective as the number of susceptible people S times the transmission rate or contact rate β, which measures the number of encounters with another person close enough to transmit an infection per time, times the probability $\frac{I}{N}$ that a person encountered is infectious. These new cases are then added to the infectious total by the $\beta\frac{IS}{N}$ term in Eq. (2).

The values of these rates are affected by biological aspects of the particular disease pathogen, but they also depend on properties of the population prior to the introduction of the pathogen and the response of public health officials. For example, if the general underlying health of a population (or a certain segment of the population, such as the elderly) was poor prior to the outbreak, γ for this community could be lower than a generally healthy community facing an outbreak of the same disease. Conversely, effective treatments and a robust health care system accessible to everyone could help the infected recover faster, thereby increasing γ. The transmission rate β is affected by whether or not the pathogen can be aerosolized and how long it survives on surfaces, but it could be reduced if the population took preventative measures such as mask wearing or hand washing that would reduce the likelihood that a single encounter between an infectious person and a susceptible person would result in a new infection. This could also change with time to reflect a reduction in the rate at which such encounters occur, such as increased social distancing and restrictions on the size of public gatherings.

4 R_0 and Herd Immunity

If any aspect of the modeling of an epidemic appears in media written for a general audience, such as when a deadly virus is first diagnosed in the United States [13] or the 2011 film *Contagion* [14] or news articles fact-checking *Contagion* [15], it is likely to be the *reproduction number* R_0. For a general audience, it has a practical, common-sense definition and an intuitional effect: R_0 measures the average number of new infections in a population with no immunity that can be traced to a single infectious person. That is, if a student returns from spring break infected with a new disease with $R_0 = 3$, he will infect his roommate, his chemistry lab partner, and his math professor when he makes up that test he missed leaving early for spring break. Each of these three people will in turn infect three more people, for a total of nine new cases at this stage of the outbreak. These nine would lead to 27 new cases in the third generation, and so the outbreak grows rapidly.

Given this loose definition, we see three different kinds of behavior based on the value of R_0:

- If $R_0 < 1$, on average, an infected person recovers before infecting another person. Thus, transmission chains are easily broken, and the outbreak will die out before spreading rapidly throughout the population.
- If $R_0 = 1$, the outbreak neither grows nor declines, and the number of infected people remains a stable proportion of the population. This kind of disease is known as *endemic*.
- If $R_0 > 1$, the outbreak exhibits sustained growth, given a sufficient number of susceptible people. This is considered the definition of an *epidemic*.

Estimates of R_0 for COVID-19 are in the range 2-to-4 [16].

Many factors contribute to the value of R_0, but they are essentially the same as biological aspects and public health measures we discussed previously for the transmission rate β and the recovery rate γ. It should therefore not be surprising that we can give a more precise mathematical definition of R_0 involving these two parameters by nondimensionalizing system (1)–(3). This can make it easier to examine the behavior of the system due to reducing the number of its parameters to just one, R_0.

The first step is to switch the unknown functions from the total people in the three compartments S, I, and R to the proportions of the total population in each category. That is, we want to rewrite the system as equations for $s = \frac{S}{N}$, $i = \frac{I}{N}$, and $r = \frac{R}{N}$. This removes the total population N as an explicit parameter in the system and can be done whether or not we proceed to the next step.

The key to the reducing the number of parameters lies in carefully choosing the timescale of the problem. Instead of measuring time in units of hours or days, we measure time by a timescale defined by the problem itself: the reciprocal of the recovery rate, or $\frac{1}{\gamma}$. In this case, we can define a new variable $\tau = \gamma t$, which functions in the role of time but is a pure number with no units, and we think of our unknown proportions $s(\tau)$, $i(\tau)$, and $r(\tau)$ as functions of this new nondimensional time.

The complicating factor is that we must now also think of differentiating these with respect to this new variable τ instead of t. According to the Chain Rule, $\frac{dI}{dt} = \frac{dI}{d\tau}\frac{d\tau}{dt} = \gamma\frac{dI}{d\tau}$. Plugging this into Eq. (2) and dividing by N yields

$$\frac{\gamma}{N}\frac{dI}{d\tau} = \frac{\beta}{N}\frac{IS}{N} - \gamma\frac{I}{N}$$

$$\frac{di}{d\tau} = \frac{\beta}{\gamma}is - i$$

Now, there is effectively only a single parameter in this equation, the ratio $\frac{\beta}{\gamma}$, which is dimensionless because it is a measure of the rate of encounters per unit of time divided by the number of recoveries per unit of time. That is, it measures almost exactly what we described R_0 as above—the rate of new infections per resolution of a single infection. This transforms the system (1)–(3) into

$$\frac{ds}{d\tau} = -R_0 i s, \qquad \frac{di}{d\tau} = R_0 i s - i, \qquad \frac{dr}{d\tau} = i, \tag{4}$$

where $R_0 = \frac{\beta}{\gamma}$.

We can also see that the earlier intuitive effects that come from R_0 being less than, equal to, or greater than 1 can be now predicted from this model and with greater insight. As a proportion, we must have $0 \leq s \leq 1$. Therefore, if $R_0 < 1$, the righthand side of $\frac{di}{d\tau} = (R_0 s - 1)\, i$ is always negative, and so the number of infected will decay to 0. The endemic case with $R_0 = 1$ will also have this property, but the decay rate will be slow enough that it may not be observed.

However, if $R_0 > 1$, with the initial susceptible proportion $s(0) > \frac{1}{R_0}$ (which is likely when a new disease is introduced into a population), then $\frac{di}{d\tau}$ will initially be positive leading to rapid growth phase at the beginning of the outbreak. As the infected proportion increases, s decreases, which causes i to grow at a slower rate, and eventually when $s = \frac{1}{R_0}$, the infected proportion reaches its maximum value, and the outbreak begins to recede because it becomes increasingly more difficult to find people who still remain susceptible.

This concept is often referred to as "herd immunity." While there are still susceptible people in the community, their number is small enough that the introduction of the disease will die out because it cannot sustain the chain of transitions necessary to grow into an epidemic.

This is often seen from the recovered point of view when there are no active cases so that $r = 1 - s$. In this context, the herd immunity threshold is $r > 1 - \frac{1}{R_0}$. This can be achieved through natural immunity from prior infections, but it can also give us a target proportion of the population to build up herd immunity prior to an outbreak. In the midst of combating an epidemic—that is, when $r = 1 - s - i < 1 - s$ and the context of our simulations below—it will be more appropriate to compare s to $s^* = \frac{1}{R_0}$. While people will still become infected when $s < s^*$, the epidemic phase has ended and the number of infected will begin to decrease.

4.1 Numerical Solution with Euler's Method

We approximately solve these systems with Euler's method. While there are higher-order methods for systems of ordinary differential equations (ODEs), we feel the simplicity of Euler's method allows us to make a more direct connection with the reasoning on which the model is built. It also permits us to more easily let the parameters be functions of time or of the values of the compartments themselves allowing us to model dynamic changes in the response of public health efforts.

Euler's method starts by discretizing the time variable. We give the details for the nondimensional system (4), and so we will discretize τ, but for (1)–(3) or the models below, the dimensional t could be used. An appropriate time step $\Delta\tau$ is chosen. Generally, the smaller $\Delta\tau$ is, the better the result for an ODE system, but

if data is only known at a given frequency such as daily reports of new confirmed cases, deaths, or hospital capacities, it may be useful to choose that period to be the time step. Once the step length is chosen, we define a discrete series of times $\tau_k = k\Delta\tau$ at which we will estimate the values of the unknowns, which we will denote with $s_k \approx s(\tau_k)$, $i_k \approx i(\tau_k)$, and $r_k \approx r(\tau_k)$.

We start with given initial conditions such as $s_0 = 0.999$, $i_0 = 0.001$, and $r_0 = 0$ so that the outbreak (or perhaps more precisely, the modeling of the outbreak) starts with 0.1% of the population infected, no one recovered, and everyone else susceptible. Euler's method takes one step forward in time for each compartment by approximating the derivatives in each equation in (4). For example, if $\Delta\tau$ is "small," then when $\tau = \tau_k$, we may we may use the definition of the derivative approximate the equation $\frac{ds}{d\tau} = -R_0 is$ by $\frac{s_{k+1}-s_k}{\Delta\tau} = -R_0 i_k s_k$. When $k = 0$, the only unknown value is s_1 since s_0 and i_0 are given by the initial conditions, and so we may solve for s_1. We may also do this for the other two equations and then proceed to the next value of k, repeating this process to take the next step forward in time to get the update equations

$$s_{k+1} = s_k + \Delta\tau\,(-R_0 i_k s_k)\,, \tag{5a}$$

$$i_{k+1} = i_k + \Delta\tau\,(R_0 i_k s_k - i_k)\,, \tag{5b}$$

$$r_{k+1} = r_k + \Delta\tau i_k. \tag{5c}$$

Figure 2 shows solutions of (5a)–(5c) for four different values of $R_0 > 1$ and an initial infection rate of 0.1%. All exhibit the same qualitative behavior:

- i increases to a maximum at τ^* when s passes through the herd immunity threshold $s^* = \frac{1}{R_0}$, and then i decays to 0,
- s decreases to a stable plateau s_∞
- r increases to a stable plateau r_∞.

Increasing R_0 steepens the initial rapid growth of i, increases the value of the maximum infected proportion, and as shown in Fig. 3, it shortens the time τ^* to reach the maximum infected proportion. Figure 3 also shows that s_∞ decreases as R_0 increases, and r_∞ increases.

For the figures in the rest of this chapter, we will frame the problem in terms of the number of people in each compartment, but when we solve the systems, we will consider the systems of the proportions of the population in each compartment as functions of the dimensional time t. When giving test values to solve the following systems, we will pick values of γ and R_0 and then define $\beta = R_0\gamma$.

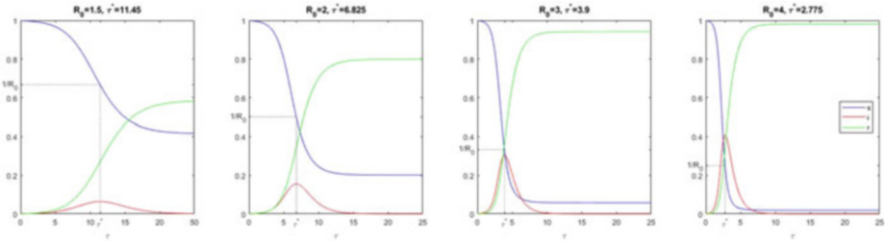


Fig. 2 Solutions to system (5a)–(5c) with initial conditions $s(0) = 0.999$, $i(0) = 0.001$, and $r(0) = 0$ for $R_0 = 1.5$, 2, 3, and 4. The corresponding (nondimensional) times τ^* of the maximum infectious proportion are 11.45, 6.825, 3.900, and 2.775. These are also the times when s passes through the herd immunity threshold $s^* = \frac{1}{R_0}$. The larger R_0 is, the steeper the initial increase of the infectious proportion (red curves), the larger its maximum, and the earlier the time τ^* that the maximum occurs. The dotted black lines indicate that τ^* also corresponds to the time at which the susceptible proportion passes through the herd immunity threshold s^*

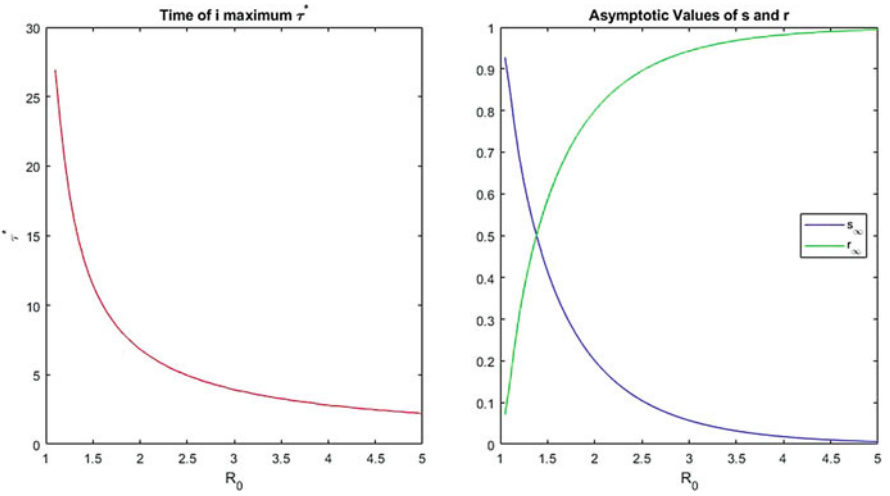


Fig. 3 Left: τ^*, the time of the maximum of i and when $s = s^* = \frac{1}{R_0}$, as a function of R_0. Right: The asymptotic values of s and r as functions of R_0

5 Modifications of the Basic Model

The basic SIR framework (1)–(3) can be modified in many ways to account for more complicated aspects of disease propagation and efforts to combat it. We will briefly discuss the effect of changing the number of compartments and the connections between them, with a focus on how the rollout of a vaccination strategy can reduce the number of deaths from the disease.

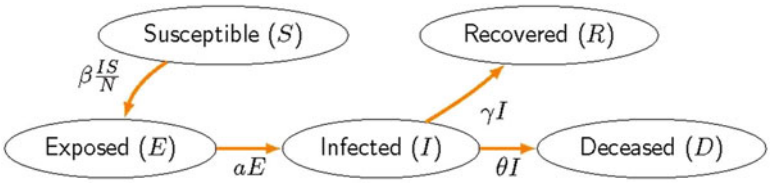


Fig. 4 A diagram of the SEIRD model. It differs from the SID model in Fig. 1 with the presence of the compartments exposed E and deceased D

5.1 Susceptible-Exposed-Infectious-Recovered-Deceased (SEIRD) Model

We begin by adding two compartments to help us better follow the natural disease progression. First, when susceptible people contract the disease, they do not immediately become infectious. There is a latent period first (the time until infectiousness). To account for this, we add an exposed compartment E that people move to before moving into the infectious compartment I. The rate at which people move from E to I is α, which is the reciprocal of the incubation period.

The second modification is to separate out of the removed compartment the number of people D who have died from the disease, which allows R to now represent only the resistant, those who have been infected but who have recovered. This expanded progression is depicted in Fig. 4, and the system of Eqs. (1)–(3) is expanded to the system of five equations

$$\frac{dS}{dt} = -\beta\frac{IS}{N}, \tag{6a}$$

$$\frac{dE}{dt} = \beta\frac{IS}{N} - \alpha E, \tag{6b}$$

$$\frac{dI}{dt} = \alpha E - \gamma I - \theta I, \tag{6c}$$

$$\frac{dR}{dt} = \gamma I, \tag{6d}$$

$$\frac{dD}{dt} = \theta I. \tag{6e}$$

The death rate θ in Eqs. (6c) and (6e) is not the case fatality rate, which is the proportion of infections that result in death instead of recovery which we will denote by δ. The "recovery rate" γ from Eqs. (2) and (3) was strictly speaking the removal

rate since it accounted for both recoveries and deaths together. We now recast that rate as $\widetilde{\gamma}$ and use the case fatality rate to define the recovery rate as $\gamma = (1-\delta)\widetilde{\gamma}$ and the death rate as $\theta = \delta\widetilde{\gamma}$ so that $\gamma + \theta = \widetilde{\gamma}$. We also now use this removal rate to define the reproduction number $R_0 = \frac{\beta}{\widetilde{\gamma}}$. An alternative way to model deaths due to the disease is the SEIR model, which is Eqs. (6a)–(6d) with $\theta = 0$ and $\gamma = \widetilde{\gamma}$. The total deaths would then be δR.

The SEIR model makes the same predictions the SIR model for the long-term values s_∞ and r_∞, which would imply the same total number of infections. However, the effect of separating the pre-infectious exposed from the infectious spreads out the infectious spike so that the maximum infectious proportion occurs later and has a smaller value compared to the SIR model. The timing of S passing through the herd immunity occurs slightly before the maximum of I because it now corresponds to the time of the maximum of the sum $E + I$. That is, the total proportion who have acquired the virus but have not yet recovered.

5.2 SEIRD Model with Vaccination

We now further modify this model by adding an important new pathway connecting compartments–vaccination. We call this new model the SEIRDv model.

If we let the susceptible population be vaccinated at a rate ν, then a group of vaccinated people νS would be transferred directly from susceptible to resistant without passing through the exposed and infectious compartments. A diagram of this model is in Fig. 5, and the system of Eqs. (6a)–(6e) now becomes

$$\frac{dS}{dt} = -\beta\frac{IS}{N} - \nu S, \tag{7a}$$

$$\frac{dE}{dt} = \beta\frac{IS}{N} - \alpha E, \tag{7b}$$

$$\frac{dI}{dt} = \alpha E - \gamma I - \theta I, \tag{7c}$$

$$\frac{dR}{dt} = \gamma I + \nu S, \tag{7d}$$

$$\frac{dD}{dt} = \theta I. \tag{7e}$$

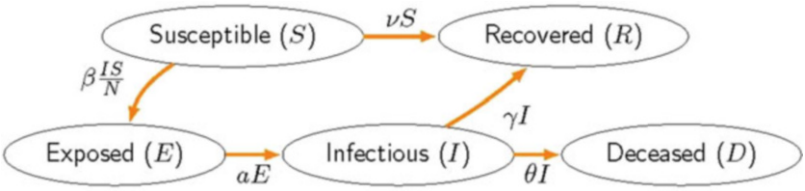


Fig. 5 A diagram of the pathway of the SEIRDv model adding the vaccination νS pathway

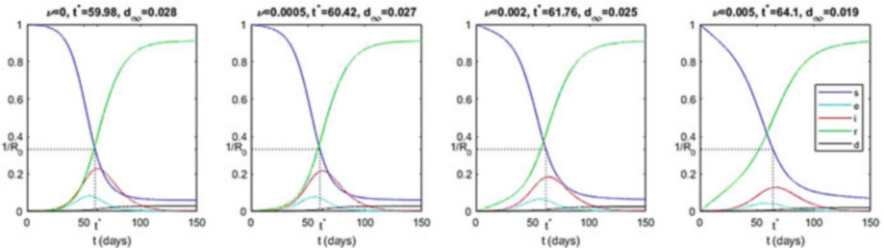


Fig. 6 The effect of the vaccination rate on the SEIRDv model with $\beta = 0.3$, $\gamma = 0.097$, ($R_0 = 3$), $\alpha = 0.333$, $\theta = 0.003$, (all have units of day^{-1}). On the left is the model with no vaccination, resulting in 2.8% total deaths. In the other three, vaccination rates $\nu = 0.0005$, 0.002, and 0.005 day^{-1} cut the total deaths to 2.7%, 2.7%, and 1.9% respectively

Figure 6 shows solutions for four different values of ν with $R_0 = 3$ and a death rate of $\theta = 0.003$ from a case fatality rate of $\delta = 3\%$. With no vaccination plan, the long-term prediction is that 2.8% of the population would die from the disease. However, modest increases in the vaccination rate can reduce this. When $\nu = 0.0005$—that is, 0.05% of the susceptible population is vaccinated per day—the total deaths are reduced to 2.7%. More ambitious rollouts of $\nu = 0.002$ and 0.005 reduce the total deaths to 2.6% and 1.9%, respectively, which is a reduction of deaths by nearly one-third compared to the case without vaccination.

Figure 7 shows the effect of the vaccination rate ν on two different aspects of the solution. On the left, we see how the total proportion of deaths decreases as the speed of the vaccination effort increases. If 1% of the remaining susceptible population were vaccinated each day, the prediction is that the number of deaths would be cut in half. On the right is the time that herd immunity is reached, which is also the time of the maximum of the sum of the exposed and infected. We see that a slow vaccination rollout with ν less than approximately 0.0086 lengthens the time to reach the herd immunity threshold, even though decline in the total number of deaths is observed. For rollouts with $\nu > 0.0086$, the time to herd immunity decreases.

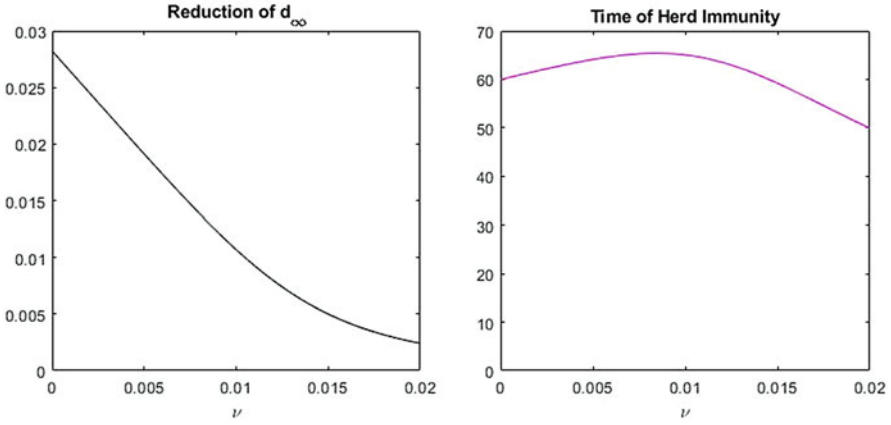


Fig. 7 Left: The proportion of total deaths as a function of ν. Right: The time of herd immunity as a function of ν. The initial increase for very small values of ν indicates that the main driving force in creating immunity is still active infections, but the vaccine can delay the maximum. However, above a threshold ν, the vaccination rate is fast enough to significantly reduce the total number of infections

5.3 SEIRDv Model with Reinfection

We have included R and D separately to give us the freedom to add another disease pathway that cannot be otherwise modeled—reinfection. This may not play a significant role in the overall predictions of the model, but we present it here briefly as another example of how effects may be added to the model by connecting compartments.

Someone who has died from the disease cannot become infected a second time. However, some of those who recover and become "resistant" may not have perfect immunity and may become re-infected, especially if they had experienced the disease in a mild form during the first round of infection. Similarly, a vaccine does not provide 100% immunity, and so some of the vaccinated in the R compartment of the SEIRDv model may still contract the disease. For example, Moderna and Pfizer-BioNTech reported about 95% effectiveness of the vaccine, which means that about five percent of people who received the vaccine got sick but experienced a mild form of the disease. The model can incorporate this effect into Eqs. (7a)–(7e) in the same manner that the susceptible become infected with resistant transferring to exposed at a rate of $\chi \frac{IR}{N}$ added in the righthand side of Eq. (7b) and subtracted in Eq. (7d). The reinfection transmission rate χ similar to β but have a significantly smaller value.

In this chapter we focused on exploring mathematics in the context of the COVID-19 pandemic. We discussed the effects of the virus on human cells, the development of the vaccines by several research teams, and mathematical models that describe and predict the spread of the Severe Acute Respiratory Syndrome Corona Virus 2. We introduced the general framework of "compartment models" used to model epidemics as a first-order system of nonlinear ordinary differential

equations in the context of the COVID-19 pandemic. We derived the simplest version of the SIR model, we adapted this model to account for the rollout of a vaccination plan, and examined its effect on the expected deaths from the virus. In this chapter, we used COVID-19 as an example that foretells nature.

References

1. Bastian, H. (2020). *Why Two Vaccines Passed the Finishing Line In a Year and Others Didn't, and a Month 12 Roundup*. PLOS. https://absolutelymaybe.plos.org/2020/12/20/why-two-vaccines-passed-the-finishing-line-in-a-year-and-others-didnt-and-a-month-12-roundup/
2. Butantan Institute. (2020). *Clinical Trial of Efficacy and Safety of Sinovac's Adsorbed COVID-19 (Inactivated) Vaccine in Healthcare Professionals (PROFISCOV)*. ClinicalTrials.Gov. https://clinicaltrials.gov/ct2/show/NCT04456595?cond=Sinovac+vaccine&draw=2&rank=1
3. National Institute of Health (2021, January). U.S. National Library of Medicine, Clinical Trials. https://www.clinicaltrials.gov
4. World Health Organization. (2021a). *The COVID-19 candidate vaccine landscape*. https://www.who.int/publications/m/item/draft-landscape-of-covid-19-candidate-vaccines
5. World Health Organization. (2021b). *Status of COVID-19 Vaccines within WHO EUL/PQ evaluation process* [PDF]. https://clinicaltrials.gov/ct2/show/NCT04456595?cond=Sinovac+vaccine&draw=2&rank=1
6. Science Media Center. (2021). *Expert reaction to reported results of the Sinovac COVID-19 vaccine from a trial by the Butantan Institute in Brazil*. https://www.sciencemediacentre.org/expert-reaction-to-reported-results-of-the-sinovac-covid-19-vaccine-from-a-trial-by-the-butantan-institute-in-brazil/
7. Kermack, W.O., and A.G. McKendrick, "A contribution to the mathematical theory of epidemics." *Proc. Roy. Soc. A* 115, 700-721 (1927)
8. Abou-Ismail A. (2020). "Compartmental Models of the COVID-19 Pandemic for Physicians and Physician-Scientists." *SN comprehensive clinical medicine*, 1–7. Advance online publication. https://doi.org/10.1007/s42399-020-00330-z
9. Bertozzi, Andrea, Elisa Franco, George Mohler, Martin B. Short, and Daniel Sledge, "The challenges of modeling and forecasting the spread of COVID-19." *PNAS*, July, 2020. https://doi.org/10.1073/pnas.2006520117
10. Blackwood, J. & Childs, L. (2018). An introduction to compartmental modeling for the budding infectious disease modeler. *Letters in Biomathematics, 5.1*. 195–221.
11. Giordano, G., Blanchini, F., Bruno R., "Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy." *Nat Med* 26, 855-860 (2020). https://doi.org/10.1038/s41591-020-0883-7
12. Hethcote, Herbert, "The Mathematics of Infectious Diseases," *SIAM Review* 42, 599-652 (2000).
13. Doucleff, Michaeleen, "No, Seriously, How Contagious is Ebola?" *Shots*, October 2, 2014, National Public Radio www.npr.org/sections/health-shots/2014/10/02/352983774/no-seriously-how-contagious-is-ebola. Accessed December 29, 2020.
14. Soderbergh, Steven, dir. *Contagion*, Warner Bros. Pictures, 2011.
15. Kritz, Fran, "Fact-Checking 'Contagion'—In Wake of Coronavirus, The 2011 Movie is Trending." *Goats and Soda*, February 16, 2020, National Public Radio www.npr.org/sections/goatsandsoda/2020/02/16/802704825/fact-checking-contagion-in-wake-of-coronavirus-the-2011-movie-is-trending. Accessed December 29, 2020.
16. Centers for Disease Control and Prevention, "COVID-19 Pandemic Planning Scenarios," https://www.cdc.gov/coronavirus/2019-ncov/hcp/planning-scenarios.html Accessed January 11, 2021.

17. World Health Organization. (2020, October 12). *Coronavirus disease (COVID-19)*. https://www.who.int/news-room/q-a-detail/coronavirus-disease-covid-19
18. Center for Disease Control and Prevention. (2021, March 15). *COVID-19 Science and research*. https://www.cdc.gov/coronavirus/2019-ncov/science/science-and-research.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fcoronavirus%2F2019-ncov%2Fmore%2Fscience-and-research.html

Application and Modeling of a Tick-Killing Robot, TickBot

Alexis L. White and Holly D. Gaff

1 Introduction

Effective tick management strategies are needed to reduce tick populations and protect humans and animals from the rising number of tick-borne diseases, but a comprehensive method has yet to be developed [6, 23]. Knowing the phenology of the local tick species can facilitate more effective timing of treatment. Research is needed in tick management to develop novel methods that account for varying phenologies, the type of habitat, and the corresponding relationship with ticks and the pathogens they carry. Additionally, many tick control technologies have the potential for ecological consequences beyond the desired tick-control. For example, host-targeted approaches that lure the hosts with food can potentially increase reproductive success in food-limited species such as rodents. Similarly, broadcast acaricides affect all invertebrates, and so are not permitted in many wetland areas. There is a need to ensure balance between the benefits of tick control and the ecological consequences of those interventions.

Most tick control efforts have focused on reducing blacklegged tick, *Ixodes scapularis*, populations, but here we focus on the lone star tick, *Amblyomma americanum*. *Amblyomma americanum* are aggressive hard-bodied ticks known to spread a variety of pathogens to humans, pets, and wildlife [3]. More recently, it has been shown that exposure to the saliva of *A. americanum* can trigger a

A. L. White (✉)
University of Florida, Gainesville, FL, USA
e-mail: alexis.white@ufl.edu

H. D. Gaff
Old Dominion University, Norfolk, VA, USA
e-mail: hgaff@odu.edu

B. Toni (ed.), *The Mathematics of Patterns, Symmetries, and Beauties in Nature*,
STEAM-H: Science, Technology, Engineering, Agriculture, Mathematics & Health,
https://doi.org/10.1007/978-3-030-84596-4_4

potentially life-threatening red-meat allergy [4]. The life history of *A. americanum* consists of four life stages: egg, larva, nymph, and adult. After emerging from eggs as larvae, ticks need a complete bloodmeal from a host to progress from one stage to the next [25]. In the Mid-Atlantic, *A. americanum* adults and nymphs are active from late spring through mid-summer, while larvae are active in late summer through autumn. *Amblyomma americanum* can also be found questing throughout the day, unlike the crepuscular dynamics of *I. scapularis* [24]. White-tailed deer, *Odocoileus virginianus*, are commonly parasitized by *A. americanum*. Movement of *O. virginianus* and other hosts provide the means to disperse ticks to novel areas, concomitant with a variety of pathogens. Other hosts of *A. americanum* include ground-nesting birds and medium-to-large-sized mammals [22].

One of the greatest challenges with tick control is teasing apart the natural variability of tick populations and the impact of the control measures. Intervention programs are very expensive and labor intensive, and so mathematical models can provide a useful tool for quickly and cheaply exploring control options. Tick population control has been modeled through optimal control [10] and age-structured difference equations [18, 19]. Optimal control theory is a method to find the best set of parameters to achieve a given outcome. Gaff et al. [10] studied optimal ways to control tick-borne diseases, but not tick populations explicitly. Mount and Haile [18] developed a lone star tick simulation (LSTSIM) using an age-structured difference equation model to understand the effects of integrated tick management methods on *A. americanum* affecting cattle. More models are needed to study the management of ticks for a wider variety of intervention methods.

Agent-based modeling (ABM) can be used as a tool to study the individual behaviors of hosts and ticks so as to gain insight into management regimes. Agent-based modeling is defined as a type of simulation where individuals' interactions with each other and their environment can be programmed and monitored. In the past ten years, the use of agent-based modeling in tick research has become popular [8, 9, 14, 20, 26–28]. TICKSIM was developed in 2011 to compare tick, host, and disease dynamics using differential equation-based models [8]. In 2013, TICKSIM was modified to look at the invasion of a tick species [9]. Wang et al. [26] developed an agent-based model that examined the interaction of ticks, hosts, and the landscape. This model was later expanded to address how the prevalence of infected ticks changes with varying host densities in different landscapes [28]. Halsey and Miller [14] built an agent-based model based on *I. scapularis* population dynamics and the efficacy of host grooming was studied. None of these agent-based models have addressed tick management.

A recently introduced method of tick management that provides immediate reduction of questing ticks without affecting the environment is use of the TickBot. TickBot is a tick-killing robot created jointly by engineers at the Virginia Military Institute and Daniel Sonenshine at Old Dominion University. It functions as a semi-autonomous, 4-wheeled robot that kills ticks by luring them to its permethrin-treated denim cloth. In 2013, TickBot was tested in a wildlife preserve in Portsmouth, Virginia. The robot uses sensors to detect a magnetic field from a guide wire placed on a given transect. In the original design, the TickBot attracted ticks with carbon

dioxide (CO_2) distributed as a gas through tubing fitted with pores and control valves that followed the same path as the guide wire. Ticks (the majority of which were *A. americanum*) were flagged in the treatment area as well as adjacent untreated areas prior to treatment. Any ticks captured were marked and returned to the same location. In control runs using flags without permethrin, Gaff et al. [11] found the TickBot reduced tick numbers on the transect to undetectable levels within 1-hour of treatment; tick densities remained zero for approximately 24 h in the treatment area. No marked ticks were recaptured in any treated area. The control results supported the use of CO_2 and permethrin.

Since the 2013 study, other trials of TickBot and other projects have been completed to provide initial insights and parameters needed to build a TickBot agent-based model. With this model we explore application frequency of the TickBot, as well as the use of CO_2 on reducing questing tick populations. In this paper, I discuss additional related studies in Sect. 2. This is followed by the model description in Sect. 3, model results in Sect. 4, and discussion in Sect. 5.

2 Parameter Estimation Studies

A series of research projects were completed to better understand the usefulness of the TickBot. These projects provide the initial insights and parameters needed to build the TickBot agent-based model.

2.1 Permethrin Potency

Ticks that encounter the TickBot's permethrin treated cloth were assumed dead in the initial field trial because ticks in the area were not recaptured [11], but little was known how much contact with the treated cloth was needed in order to kill a tick. A study was completed to gather information about how often and how fast *A. americanum* dies after being brushed with a permethrin treated cloth, and how much contact, in terms of surface area of the cloth, was required. One control and four different treatment groups were established as 2% permethrin with 2 brushes of the cloth, 2% permethrin with 1 brush, 0.5% permethrin with 2 brushes, and 0.5% permethrin with 1 brush. Miniature flags were created by attaching the cloth onto a pencil, one for each permethrin concentration, and one for control ticks. The control ticks were brushed twice with an untreated cloth.

Eleven *A. americanum* nymphs and 21 adults were collected from a wildlife preserve and divided into the four treated groups and one untreated group. All ticks in this study exposed to permethrin died within 36 min of exposure. None of the control ticks died. This supports the assumption that any ticks that come in contact with the TickBot permethrin-treated cloth will die within an hour of exposure.

2.2 TickBot 2016

Building from the initial 2013 study of the TickBot [11], a second year of testing was done. In the summer of 2016, TickBot was implemented at two sites in the same wildlife preserve as was used in 2013, and a nearby residential site in Portsmouth, Virginia. In 2013, TickBot was applied for an hour using gaseous CO_2 distributed through an elaborate set of plastic tubing to lure ticks to the robot. The TickBot ran on a guide wire that was laid out for each run with the tubing along the prescribed path. The set-up process exposed the researchers to ticks while laying out the wire and CO_2 tubing. Also, the canisterized CO_2, along with tubing, was cumbersome and impractical in a field setting. To improve the set-up of the TickBot, we compared the efficacy of the TickBot with CO_2 sublimating from dry ice as an attractant versus the original gaseous CO_2 from a canister. To compare TickBot efficacy relative to the original study, the same procedures were used: ticks were captured, marked with nail polish, then returned to the habitat the day before treatment. After each TickBot run, treated and untreated areas were flagged for ticks at 1 h, 4 h, next morning, and next afternoon. Each trial location had an adjacent control transect within five meters of the treated transect. Eight trials were completed during May-August 2016. Efficacy of the TickBot was measured by density of ticks before and after treatment, as well as the absence of marked ticks in treated areas. No differences in tick density were found between the use of CO_2 from a canister or dry ice as attractants. Efficacy of the TickBot trials did not vary between the wildlife preserve and the residential site. No marked ticks were collected in treated areas, and no ticks were found on treated transects within 24-hours of treatment. This supports the study by Gaff et al. [11], confirming the ability of the TickBot to reduce ticks on the treated area to undetectable for 2 h.

2.3 Dry Ice Attraction Study

We completed a supporting study to quantify the mobility of *A. americanum* toward a dry ice attractant by measuring how many ticks, from varying distances, came within the range of the TickBot during treatment. Previous studies have been conducted to test the movement of *A. americanum* toward a stationary source of CO_2 [7], but distance to a moving attractant has not been studied. The TickBot attracts ticks through the use of a piece of dry ice in a cup (with holes) attached to the TickBot. The dry ice sublimates as the TickBot drives along the prescribed circuit. A total of 18 *A. americanum* were used during each trial experiment, with six ticks placed into plastic tubing at 1 m, 3 m, and 5 m from the dry ice source (Fig. 1). All of the ticks were painted with nail polish according to the location of initial placement. Dry ice was used as an attractant at one end of the tubing, and the opposite end was sealed as were the holes in the tubing where the ticks were introduced. The ticks were observed for 15 min, and movement of each tick was

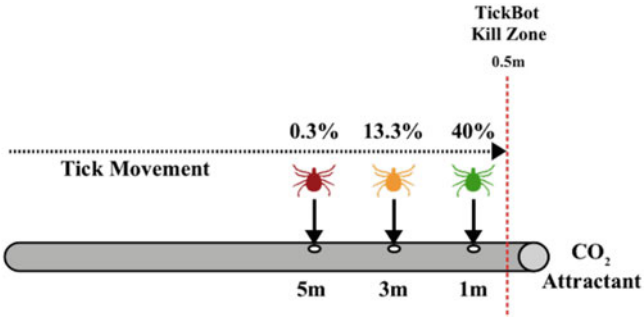


Fig. 1 Study design for each trial is displayed here. Painted ticks were placed at the three distances and movement toward the attractant was documented. The percentage of ticks that moved within 0.5 m of the CO_2 were considered to reach the kill zone of the TickBot

documented. The percentages of ticks that moved within 0.5 m of the open end of the tube, the presumed kill zone of the TickBot, were calculated for each distance (Fig. 1). For example, a tick initially placed farthest away from the attractant at 5 m, would have to move at least 4.5 m to be killed by the TickBot. This study found that attraction can be coded into a model as 40% of ticks are lured from one meter away, 13.3% at three meters, 0.3% at five meters into the path of the TickBot.

2.4 TickBot 2017

In the studies of 2013 and 2016, TickBot was tested at a wildlife preserve and a residential area; the TickBot study of 2017 was the first semi-permanent installation of the robot. The study ran from May 9 to August 9, 2017, at NASA Langley Child Development Center in Hampton, Virginia. The Center already utilized two forms of tick control by maintaining the landscape through mowing, and lining the fence with a mulch barrier; but they would still encounter ticks in the playground. The goal of the study was to protect the fenced area of the playground by creating a tick-free zone between the fence and the nearby long grass. The TickBot guide wire was buried approximately 0.1 m deep in the area just outside the fence of the playground to create the TickBot's course. During the study period, the TickBot was run on 23 occasions. For the safety of the children and facility employees, ticks were not marked and returned as in previous studies. An untreated site was established on the edge of the grass line adjacent to the fence to provide data on ticks that inhabit the area. The untreated site and the area within the fence were flagged on each visit. Staff at the Center also passively collected ticks from themselves or children during the time of the study. All ticks collected were identified morphologically in the laboratory. A total of 74 ticks were collected (71 by flagging, 3 found on humans). The most abundant species of tick collected were American dog ticks, *Dermacentor variabilis* (39 ticks), followed by 26 lone star ticks, *A. americanum*, and 3 rabbit

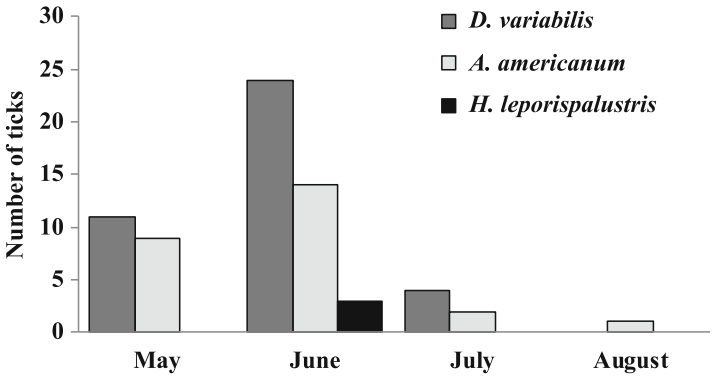


Fig. 2 Results from application of the TickBot at NASA Langley. Bars represent the different tick species collected in the untreated areas

ticks, *Haemaphysalis leporispalustris*. Of the 71 flagged ticks, 68 were collected in the untreated area. Three *A. americanum* were collected from flagging in the playground area. No ticks were collected in the playground area in July or August (Fig. 2). This study supports the efficacy of the TickBot as part of an integrated tick management program.

3 The Model

This model is based on the previously published TICKSIM [8, 9]. Data from field and lab experiments were used to parameterize, initialize, and validate the model. The model description follows the ODD (Overview, Design concepts, Details) protocol [12, 13]. The ODD protocol was created to standardize agent-based model descriptions and increase reproducibility. This protocol consists of seven major sections: (1) Purpose, (2) Entities, state variables, and scales (3) Process overview and scheduling, (4) Design concepts, (5) Initialization, (6) Input data, and (7) Submodels.

3.1 Purpose

The purpose of this model is to simulate the efficacy of the TickBot as a method of controlling lone star ticks, *A. americanum*, within a peridomestic setting from May-June in southeastern Virginia. TickBot is a tick-killing robot that lures ticks to the predetermined path of the robot and kills them using an attached permethrin treated cloth. Frequency of application with or without an attractant of CO_2 is also explored within this model.

3.2 Entities, State Variables, and Scales

3.2.1 Agents/Individuals

The model consists of two populations of agents, ticks and hosts. To initialize the model 10,000 ticks agents are introduced (10% adults and 90% nymphs) [8]. Tick agents, (Fig. 3A), are characterized by the following state variables: identification number, sex, life stage, activity, list of hosts, and closest patch that is on the TickBot path. Ticks were assigned sex upon birth and moved throughout the four life stages: egg, larva, nymph, and adult. Ticks move through three different activities for each life stage (resting, questing, feeding). Adult female ticks that have fed will complete an additional activity of laying eggs. The tick population was not held constant and included the phenology of *A. americanum* [21]. If the model included attraction of ticks to TickBot treated patches, ticks were able to detect the closest TickBot patch to determine if movement was possible. Beyond movement toward the TickBot, ticks are unable to move except by host because of the spatial resolution of the model.

To initialize the model 50 deer are introduced as generalized host agents, this would include all natural hosts of *A. americanum* in a backyard forest including: domesticated animals, ground-dwelling birds, and small and large wild mammals [2, 5, 15, 17]. Host agents (Fig. 3B), were characterized by the following state variables: identification number, mortality rate, number of ticks of each life stage currently feeding on the host, and the maximum number of ticks able to feed on each host of each life stage. The host population remained constant, so if one host died another was born into a random cell. If a host dies, all ticks on that host also die.

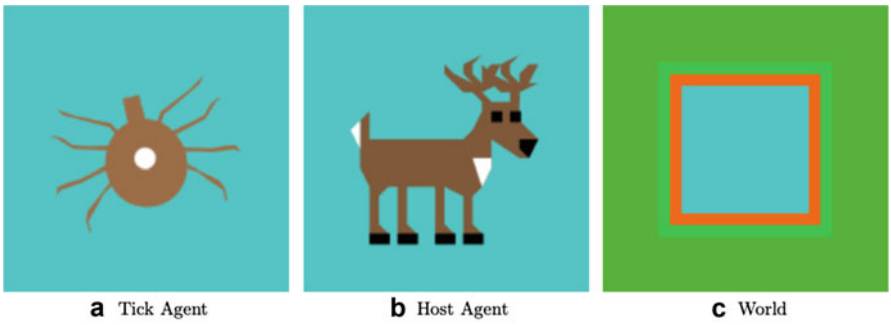


Fig. 3 Here NetLogo agents and the environment are described, with (**a**) the tick agents, (**b**) host agents, and (**c**) the patches with dark green as forest habitat, light green as forest edge, orange as areas treated with TickBot, and cyan as the lawn habitat

3.2.2 Spatial Units

The model environment was a 25×25 patch grid assumed to be $100\,\text{m} \times 100\,\text{m}$. Patch colors were used to denote various habitats in a generic peridomestic habitat (lawn, TickBot area, forest edge, and forest). At the center of the grid was the lawn habitat designated by a 12x12 cyan square (Fig. 3C). At the edge of the lawn habitat was a ring of orange TickBot patches, followed by a ring of forest edge lime green patches, and the remaining patches on the edge of the model environment were green to represent forested habitat.

3.2.3 Environment

The original environment is set up as a grid of 25×25 patches of equal quality with wrapping boundaries and assumed to represent approximately 100 m by 100 m. Lawn colored patches were assumed to have higher mortality based on field observations of *A. americanum* in residential mowed habitats (Sects. 2.2 and 2.4).

Time of year also effects tick mortality, with increased mortality during the winter months. Each time step in the model represents one hour because the TickBot is only used for one hour on treatment days. Each simulation was run for 5760 hourly time steps, or 8 months, to allow for the questing adults and nymphs to peak in early summer and questing larvae to peak late summer into early fall [21].

3.3 Process Overview and Scheduling

The model proceeds in hourly time steps. Four modules happen in each time step: set the day of the year, tick changes, host changes, and running TickBot if it is time. Within each module, individuals are processed randomly. The sequence of processes is detailed in Fig. 4.

3.4 Design Concepts

3.4.1 Basic Principles

Similar to [8, 9] and [20], the current model design is based on host-tick interaction as different agents. Hosts move randomly over patches. When a tick encounters a suitable host on a patch, the tick uses the host as a bloodmeal. Adult female ticks will reproduce after feeding. Mortality of ticks is caused by habitat (lawn patches are less favorable), time of year, host availability, and TickBot application. TickBot is implemented as patches around a lawn that will kill all ticks present during the hour of application. Ticks only move by host or through being lured to the TickBot

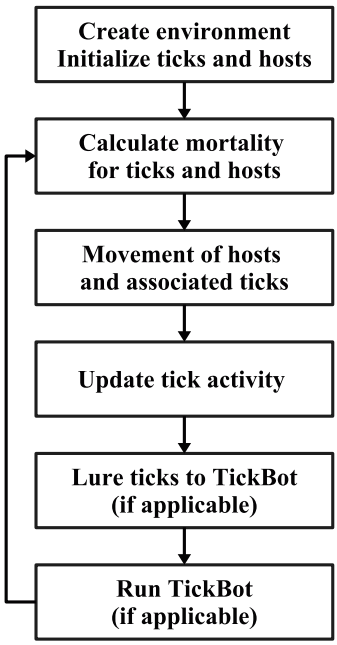


Fig. 4 Flow diagram for agent-based model

patches an hour before application of TickBot. By varying application frequency, with or without attraction, the efficacy of the TickBot as a tick control method can be studied.

3.4.2 Sensing

Ticks sense hosts within the same patch and have a probability of successful attachment on the host (Table 1). Ticks could only attach if there was space to feed. This is modeled using a maximum numbers of ticks per host at a particular life stage: 100 larvae, 25 nymphs, 10 adults. These numbers have been adapted for each life stage from previous models which have allowed 200 ticks per host [8]. Ticks did not sense other ticks, and hosts did not sense other hosts. During TickBot treatment with CO_2, ticks will sense the closest TickBot patch. If the closest patch is one patch (or less) away, the ticks will move to the TickBot patches.

3.4.3 Interaction

As in [8, 9] and [20], ticks interact with hosts in their cell and after successful attachment move from "questing" activity to "feeding" activity. After feeding for three days, a tick would drop off in the patch the host was in, and the tick would switch to "resting". After 60 days (1,440 h), the "resting" ticks will switch life stages

Table 1 Baseline parameters used in the model are adapted from either [8]† or [9]‡. Remaining values are from field observation or biological understanding of the system*

Entities	Parameter	Category or value	Reference
Environment	Simulation extent	10,000m^2	‡
	Number of cells	625 patches (25 x 25 Grid)	‡
	Patch size	16m^2	‡
	Monthly mortality (μ)	0.1 in Jan., Feb., Mar., Sept., Oct., Nov., and Dec.; 0.01 in Apr., May, June, July, and Aug.	†
	Lawn desiccation parameter	0.01	†
Hosts	Initial host population	50 hosts	*
	Host movement rate	1 patch per time step	†
	Host mortality	0.002/24hr	†
	Max adults per host	10 ticks	†
	Max nymphs per host	25 ticks	†
	Max larvae per host	100 ticks	†
Ticks	Initial tick population	10,000 ticks (90% nymphs 10% adults)	†
	Prob. successful attachment	0.003	†
	Eggs laid per female	250 eggs	†
	Time from egg to hatching	60 days	*
	Molt time from larvae to nymph	270 days	†
	Molt time from nymph to adult	360 days	†
	Maximum questing time	42 days	†
	Length of bloodmeal	3 days	†

and begin to "quest" again. When an adult female tick successfully feeds, she will lay 250 eggs, this is not reliant on the interaction with a male tick in this model. The timing of phenology was adjusted from the previous models to incorporate known phenological peaks of field collected data in Virginia.

3.4.4 Stochasticity

Stochasticity was used in this model for host movement, host mortality, tick mortality, successful tick attachment to hosts, and the chance of a tick being lured to the TickBot. These probabilities are included in (Table 1).

3.4.5 Observation

The following metrics were monitored throughout each time step: number of ticks in each life stage in each habitat, number of ticks lured to the TickBot patches, and

the number of ticks killed by the TickBot. These metrics were used to determine the following four outputs: number of ticks killed by TickBot, maximum density of each life stage of tick throughout the simulation in each habitat, total sum of all questing ticks throughout the simulation of each life stage in each habitat, and total sum of all ticks in any activity of each life stage in each habitat throughout the simulation.

3.5 Initialization

Each simulation began with a fixed landscape of habitats: lawn, TickBot area, forest edge, and forest (Fig. 3C). Hosts and ticks were randomly distributed across all habitat types. The simulations ran for 5760 hourly time steps, equating to May 1 to December 30 to encompass the phenological peaks of *A. americanum*. The initial conditions for each simulation are outlined in Table 1, which were derived from values in the literature or based on field observation [8, 9, 20].

3.6 Input Data

This model did not use input data.

3.7 Submodels

3.7.1 Process Passage of Time

Time steps were measured in one hour intervals based on the field tested hour long application of the TickBot on treatment days. Time of year influences tick mortality and activity. Each month was assumed to be 30 days, and each month had a corresponding mortality since ticks are more likely to die in the winter months, than in the summer (Table 1).

3.7.2 Process Tick Life Cycle

For every time step, each tick had a given probability of dying [8, 9, 20]; rates were adjusted to be hourly for our model. Tick mortality (M) is dependent on monthly mortality (μ), total tick population size (N), the carrying capacity (K), and was calculated as:

$$M = \mu * \frac{N}{K}$$

Parameters for monthly mortality (μ) and carrying capacity (K) are found in Table 1. N was counted after each time step to recalculate M. During each time step, a random number between 0 and 1 was generated for each tick; if the random number was less than M, the tick would die.

Ticks moved through three activities resting, questing, and feeding. At the beginning of the model, ticks are resting fed nymphs and larvae, waiting to emerge as adults and nymphs respectively. Each tick had a "time in activity" timer that was set to allow them to switch activity to questing when the timer reached 0. Fed nymphs that become adults, begin to quest at a random time in May or June to reflect *A. americanum* adult phenology. Fed larvae that become nymphs, begin to quest at a random time between May and mid-July to reflect the nymph phenology of *A. americanum*. Questing ticks had 42 days to find a host or they die. If ticks found a host, there was a probability of successful attachment. A random number was selected from 0 to 1; if the number was less than the probability of successful attachment (0.003) then the tick would switch activity to feeding. All ticks that successfully attached to a host, would take a bloodmeal for three days and move with the host agent across the patches. After feeding, nymphal ticks would drop-off their host, and switch activity to resting as they remained quiescent and molted into adults the following year. Adult male ticks die after the feeding activity. Adult female ticks, drop-off their host and lay 250 eggs and die. Eggs hatched as new tick agents and remain in the resting activity for 60 days before switching to questing and becoming larval ticks. All values for the life cycle of the ticks are described in Table 1.

Additional mortality was applied to ticks in the lawn patches. As observed in TickBot field studies in 2016 and 2017 (Sects. 2.2 and 2.4) very few ticks are found in lawn habitat; this increased mortality was adjusted to match numbers found in our field studies.

3.7.3 Process Host Mortality and Movement

In each time step, each individual host had a given probability of dying. As in [20], a random number was generated between 0 and 1; if the random number was less than the host mortality value, then the host would die. This allowed for some population turn over, but also allowed the host population to remain constant as each host that died was replaced by another randomly in the environment. When hosts die, all ticks on that host were assumed to die and were not placed on the new host.

3.7.4 Process TickBot

TickBot application frequency was varied but was assumed to begin the second week of May and continue for four weeks. All TickBot applications were done on treatment days at noon. Once a week application ran Mondays, twice a week was Mondays and Wednesdays, three times a week was Mondays, Wednesdays,

and Fridays, and daily was all seven days for the four weeks. During this time all ticks in the questing activity on TickBot patches were killed. Ticks were also killed if they switched into questing activity within the next 12 h of the simulation; this additional death comes from field collected data from the TickBot in 2013 [11] and 2016 (Sect. 2.2), which supports the lack of questing ticks in a treated area for 24 h.

Within the model, all ticks within the TickBot habitat were killed. This is based on the study in Sect. 2.1. All ticks in the model die instantly during TickBot treatment times.

Ticks were also lured, in the model, during scenarios that included attraction to CO_2. This is based on the study described in Sect. 2.3. Within the model, ticks could only be lured one hour before treatment time (as was done in field studies [11], Sects. 2.2, and 2.4), during this time a random number in the model was generated between 0 and 1 for each tick that was one patch away; if the random number was less than 0.27 (chance of being lured) then the tick would move onto a TickBot patch. The luring probability of 0.27 was calculated using data from Sect. 2.3. Each patch is considered to be 4 m by 4 m, therefore one patch away from the TickBot should be approximately the average of what we found at 3 m (13.3%) and 1 m (40%)(Fig. 5).

The model was programmed using NetLogo version 6.0.3 [29]. This software was written by Uri Wilensky in 1999 and is freely available. The model was run on the Turing High Performance Computing cluster at Old Dominion University using BehaviorSpace to run experiments "headless", from the command line. All statistical analyses on model results were completed in MATLAB [16].

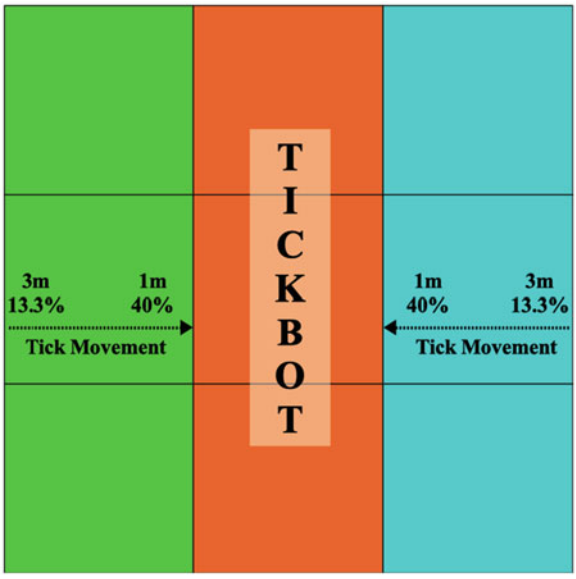


Fig. 5 Diagram of luring ticks to TickBot patches in the model. Green patches are forest edge, orange are TickBot, and cyan are lawn patches. Each patch is 4 m by 4 m and 27% of ticks were lured from one patch away averaging the values of 13.3% and 40%

3.8 Scenarios

To evaluate the use of TickBot as a tick control method for *A. americanum*, eight scenarios were created by varying the frequency of when the TickBot was used and whether or not CO_2 was used to attract the ticks. The scenarios include: once a week (1X), twice a week (2X), three times a week (3X), or daily, and each scenario ran with and without the use of CO_2 to lure nearby ticks to the treatment area. A "no control" scenario with no treatment used was also run.

Four metrics were used to compare the efficacy of each scenario for each habitat type: number of ticks killed by the TickBot, maximum density of questing ticks, total number of questing ticks, and the total number of ticks. The latter three metrics were all compared to a null scenario, where no treatment was used, to calculate a percent reduction of ticks from the null [1]. The TickBot in the model kills 100% of ticks on the day it runs, and the maximum number of treatment days is 10% of the total number of simulation days. Therefore a percent reduction of more than 10% was defined here as effective. Tick populations in some habitats such as lawn are small and therefore percent reductions are paired with the numerical differences in populations to reflect whether a 50% reduction is from a single tick or thousands of ticks. An additional habitat of "yard" was created by combining lawn, TickBot, and forest edge patches to better describe areas where the greatest risk of tick encounters would occur.

4 Results

Agent-based modeling allows for stochasticity therefore 30 replicates of each scenario were run to account for this variation. Replicates were then averaged together to reflect the average output for the scenarios. In Fig. 6 the standard deviation is shown around the mean values for the density of questing ticks for the no control scenario. All scenarios had approximately the same level of variation and therefore only the no control is presented. Adult questing ticks had a standard deviation at ~15% of the mean (Fig. 6B). Nymphs had the least variation out of the different life stages, with standard deviation at ~5% of the mean (Fig. 6C), while larvae had the most variation with standard deviation at ~30% of the mean (Fig. 6D).

Densities of questing ticks per patch varied with the different scenarios but the standard pattern with no control can be found in Fig. 7A. For scenarios without the use of luring ticks (no CO_2), tick density in TickBot habitat increased similarly to the forest and forest edge, until treatment days when it would be reduced to zero and increase as ticks began to quest in those areas (Fig. 7B). When ticks were lured to TickBot patches before treatment, the density of questing ticks would increase in TickBot patches and decrease in forest edge and lawn as ticks were lured from these

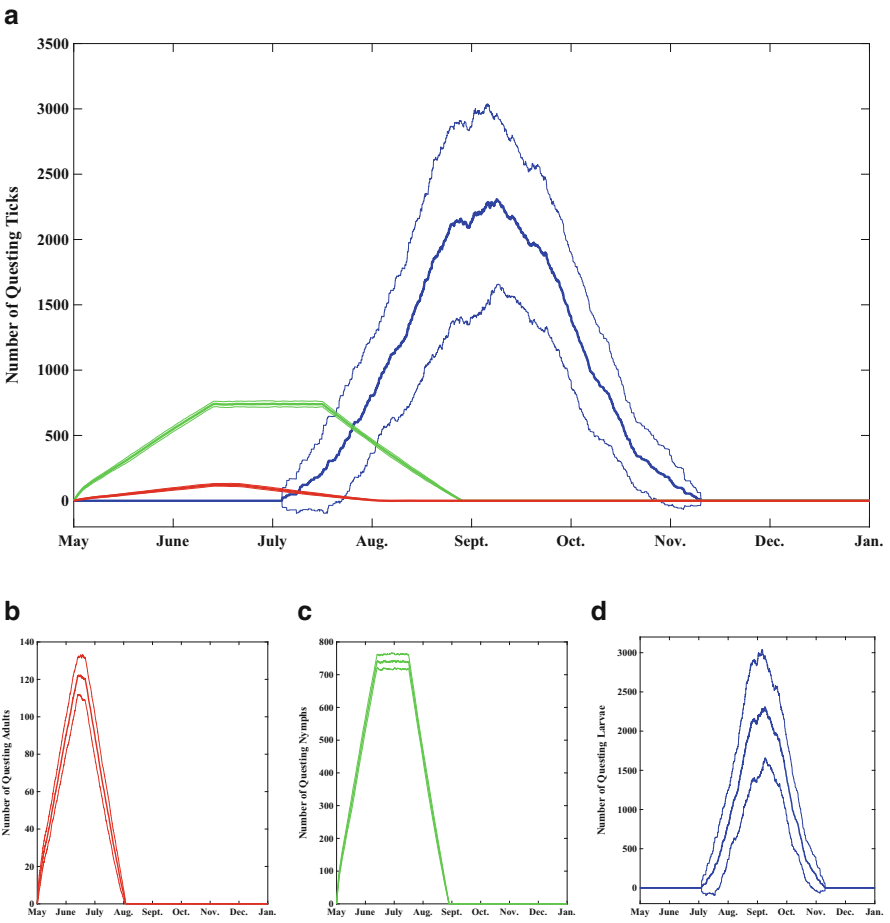


Fig. 6 Standard deviation is presented as fine lines above and below the mean bolded line for populations of questing ticks in the yard habitat. (**a**) All life stages of ticks on the same scale with adult ticks in red, nymphs in green, and larvae in blue (**b**) adult ticks (**c**) nymphs, and (**d**) larvae

areas (Fig. 7C). The greatest decline in density of questing ticks was seen from daily treatment of the TickBot with CO_2 (Fig. 7D).

4.1 Number of Ticks Killed

The number of ticks killed in each scenario from the TickBot was compared in Fig. 8 and Table 2. All ticks lured to TickBot patches, or that were currently on TickBot patches during treatment, were assumed to all die. Scenarios with the use of CO_2 killed more ticks then just the TickBot alone. Frequency of application

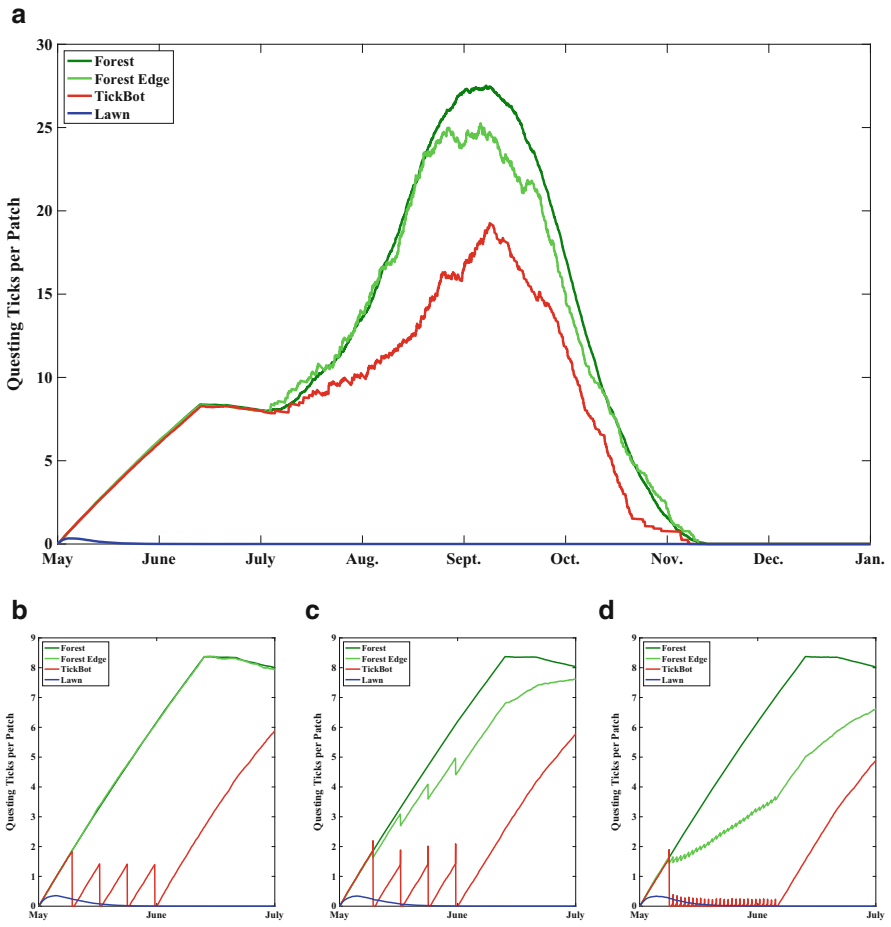


Fig. 7 Four scenarios were selected to compare the resulting tick densities (ticks per patch) for each. Forest habitat is represented in dark green, forest edge in light green, TickBot in red, and lawn in blue. Fig. (**a**) shows no control densities over the entire simulation May-December, and (**b**-**d**) focus on the months when treatment occurred May-July

also increased the number of ticks killed. Daily treatment without CO_2 and once a week treatment with CO_2 killed nearly the same number of ticks until the final treatment (Fig. 8 and Table 2). In each scenario hundreds of ticks are killed, but provide perspective on total risk reduction in the yard habitat the three times a week scenarios with and without CO_2 were compared with the no control scenario (Fig. 9). Although the TickBot kills all ticks on the TickBot patches, this does not include all ticks present in areas of human encounter, here defined as yard habitat.

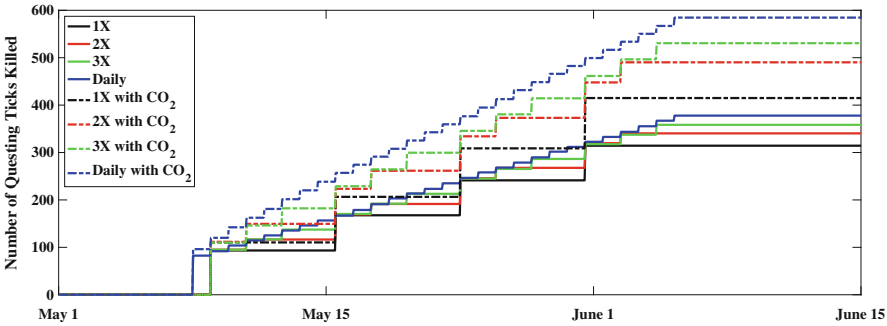


Fig. 8 The number of questing ticks killed by the TickBot per scenario is compared. TickBot scenarios with the use of CO_2 are represented by dashed lines. Colors are matched by frequency of application: 1X in black, 2X in red, 3X in green, and daily in blue

Table 2 The number of ticks killed in each scenario is compared by life stage: "L" for larvae, "N" for nymphs, "A" for adults, "All" for all life stages. Scenarios without the use of TickBot are marked with a dash (-)

	Ticks killed by TickBot			
Scenarios	L	N	A	All
No control	–	–	–	–
1X	19	254	41	314
2X	38	261	42	340
3X	55	262	42	358
Daily	134	212	32	378
1X with CO_2	18	341	56	415
2X with CO_2	38	389	64	490
3X with CO_2	52	411	67	531
Daily with CO_2	128	395	62	585

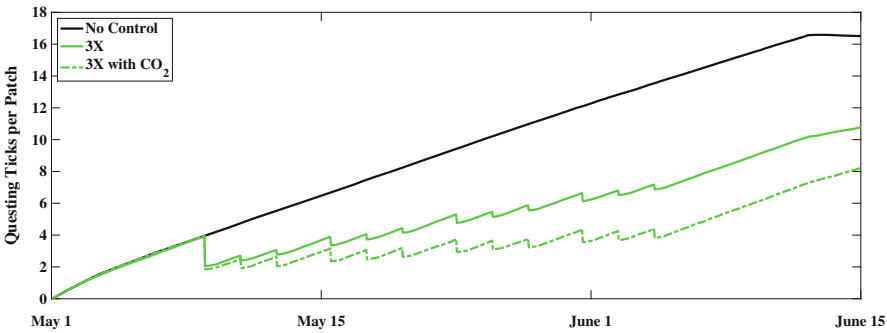


Fig. 9 Density of questing ticks in yard habitat during three times a week treatment with and without CO_2 compared to the untreated no control scenario. No control is a black line, three times in green, and three times with CO_2 as dashed green line

4.2 Maximum Density of Questing Ticks

Percent reduction of the maximum density of questing ticks was compared to the difference in ticks in each scenario and each life stage of tick. In the lawn habitat, the density of ticks was very low because of the high risk of desiccation. None of the scenarios were statistically different from the control scenario, and only the three times a week scenario without CO_2 was effective (>10% reduction) for adults (Table 3). The differences for these reductions are less than 0.03 ticks per lawn patch (Table 3).

In the yard habitat all scenarios were effective at reducing the maximum density of larvae, adults, and all ticks combined, but no scenario was effective in controlling nymphs (Table 4). The effective percent reduction in maximum density of all tick life stages combined is caused by the increased density of larvae in both habitats. For larvae and all ticks combined only the scenarios one time with CO_2, two times with CO_2, and three times with CO_2 were statistically significant from the no control scenario. For nymphs, all scenarios were significantly different from the no control except the one a week and twice a week scenarios. All scenarios were statistically significant from the control for adults (Table 4).

In the forest habitat, no scenario was effective or significantly different from the no control scenario Table 5). The forest has the greatest density of ticks in the model and the greatest difference in the mean scenario values from the no control scenario is less than 1.11 ticks per forest patch (Table 5).

Table 3 Percent reduction and difference in maximum density of questing ticks in the lawn habitat at each life stage for each scenario. Numbers in bold represent an effective reduction of >10%. Italicized numbers are the densities of ticks in the scenario with no control. Life stages of ticks are represented by "L" for larvae, "N" for nymphs, "A" for adults, "All" for all life stages. Scenarios without the use of TickBot are marked with a dash (-). A one-way ANOVA was used to compare the results of the scenarios. None of the scenarios are significantly different from the no control scenario

	Percent reduction				Difference			
Scenarios	L	N	A	All	L	N	A	All
No control	–	–	–	–	*0.00*	*0.29*	*0.05*	*0.34*
1X	0.00%	−3.95%	−6.70%	−5.97%	0.00	−0.01	0.00	−0.02
2X	0.00%	9.77%	−10.61%	5.40%	0.00	0.03	−0.01	0.02
3X	0.00%	−1.97%	**11.17%**	−0.98%	0.00	−0.01	**0.01**	0.00
Daily	0.00%	0.85%	2.79%	0.00%	0.00	0.00	0.00	0.00
1X with CO_2	0.00%	−1.22%	7.26%	−1.23%	0.00	0.00	0.00	0.00
2X with CO_2	0.00%	−6.30%	−5.59%	−7.20%	0.00	−0.02	0.00	−0.02
3X with CO_2	0.00%	1.69%	−5.03%	−0.33%	0.00	0.00	0.00	0.00
Daily with CO_2	0.00%	3.01%	−5.59%	0.41%	0.00	0.01	0.00	0.00

Table 4 Percent reduction and difference in maximum density of questing ticks in the yard habitat at each life stage for each scenario. Numbers in bold represent an effective reduction of >10%. Italicized numbers are the densities of ticks per yard patch in the scenario with no control. Life stages of ticks are represented by "L" for larvae, "N" for nymphs, "A" for adults, "All" for all life stages. Scenarios without the use of TickBot are marked with a dash (-). A one-way ANOVA was used to compare the results of the scenarios. Results significantly different from the no control scenario are noted by an asterisk (*)

	Percent reduction				Difference			
Scenarios	L	N	A	All	L	N	A	All
No control	–	–	–	–	*10.26*	*3.30*	*0.54*	*10.26*
1X	**19.93%**	0.39%	**19.72%***	**19.93%**	**2.05**	0.01	**0.11***	**2.05**
2X	**19.92%**	−0.66%	**26.43%***	**19.92%**	**2.04**	−0.02	**0.14***	**2.04**
3X	**24.61%**	0.47%*	**27.50%***	**24.61%**	**2.53**	0.02*	**0.15***	**2.53**
Daily	**19.99%**	2.59%*	**28.61%***	**19.99%**	**2.05**	0.09*	**0.16***	**2.05**
1X with CO_2	**14.34%**	0.65%*	**26.70%***	**14.34%**	**1.47**	0.02*	**0.15***	**1.47**
2X with CO_2	**26.49%***	0.22%*	**32.87%***	**26.49%***	**2.72***	0.01*	**0.18***	**2.72***
3X with CO_2	**23.84%***	0.85%*	**41.00%***	**23.84%***	**2.45***	0.03*	**0.22***	**2.45***
Daily with CO_2	**32.79%***	2.11%*	**43.40%***	**32.79%***	**3.37***	0.07*	**0.24***	**3.37***

Table 5 Percent reduction and difference in maximum density of questing ticks in the forest habitat at each life stage for each scenario. Numbers in bold represent an effective reduction of >10%. Italicized numbers are the densities of ticks per forest patch in the scenario with no control. Life stages of ticks are represented by "L" for larvae, "N" for nymphs, "A" for adults, and "All" for all life stages combined. Scenarios without the use of TickBot are marked with a dash (-). A one-way ANOVA was used to compare the results of the scenarios. None of the scenarios are significantly different from the no control scenario

	Percent reduction				Difference			
Scenarios	L	N	A	All	L	N	A	All
No control	–	–	–	–	*27.50*	*7.17*	*1.21*	*27.50*
1X	0.24%	−0.08%	1.28%	0.24%	0.07	−0.01	0.02	0.07
2X	−0.51%	−0.06%	0.67%	−0.51%	−0.14	0.00	0.01	−0.14
3X	2.08%	0.37%	0.78%	2.08%	0.57	0.03	0.01	0.57
Daily	−0.41%	0.60%	0.89%	−0.41%	−0.11	0.04	0.01	−0.11
1X with CO_2	4.02%	−0.14%	1.30%	4.02%	1.11	−0.01	0.02	1.11
2X with CO_2	1.54%	0.73%	1.62%	1.54%	0.42	0.05	0.02	0.42
3X with CO_2	3.10%	0.13%	1.20%	3.10%	0.85	0.01	0.01	0.85
Daily with CO_2	−0.73%	−0.21%	1.66%	−0.73%	−0.20	−0.01	0.02	−0.20

4.3 Sum of Questing Ticks

Percent reduction in the sum of questing ticks for each scenario are compared for nymphs and adults combined (N+A) and all life stages, in each habitat. No scenario was statistically significant from the no control scenario and no scenario had an effect on the lawn or forest habitat (Tables 6 and 8). The values for the sum of

Table 6 Percent reduction and difference of the sum of questing ticks in the lawn habitat at each life stage for each scenario. Numbers in bold represent an effective reduction of >10%. Italicized numbers are the sum of questing ticks in the scenario with no control. Values for "N+A", nymphs and adults, are the same as all life stages, "All", in this habitat therefore only the value for "All" is reported. Scenarios without the use of TickBot are marked with a dash (-). A one-way ANOVA was used to compare the results of the scenarios. Results significantly different from the no control scenario are noted by an asterisk (*)

	Percent reduction		Difference	
Scenarios	N+A	All	N+A	All
No Control	–	–	*92.38*	*92.38*
1X	2.59%	2.59%	2.39	2.39
2X	7.20%	7.20%	6.65	6.65
3X	4.69%	4.69%	4.33	4.33
Daily	6.28%	6.28%	5.80	5.80
1X with CO_2	4.52%	4.52%	4.18	4.18
2X with CO_2	0.92%	0.92%	0.85	0.85
3X with CO_2	5.88%	5.88%	5.43	5.43
Daily with CO_2	9.53%	9.53%	8.80	8.80

Table 7 Percent reduction and difference of the sum of questing ticks in the yard habitat at each life stage for each scenario. Numbers in bold represent an effective reduction of >10%. Italicized numbers are the sum of questing ticks in the scenario with no control. Life stages of ticks are represented by "N+A" for nymphs and adults, and "All" for all life stages. Scenarios without the use of TickBot are marked with a dash (-). A one-way ANOVA was used to compare the results of the scenarios. Results significantly different from the no control scenario are noted by an asterisk (*)

	Percent reduction		Difference	
Scenarios	N+A	All	N+A	All
No Control	–	–	*6613.87*	*21126.49*
1X	**17.18%***	**21.97%***	**1136.31***	**4641.45***
2X	**18.91%***	**18.85%***	**1250.80***	**3981.88***
3X	**20.08%***	**23.58%***	**1328.23***	**4981.57***
Daily	**21.93%***	**19.21%***	**1450.09***	**4058.49***
1X with CO_2	**22.39%***	**19.16%***	**1481.04***	**4048.12***
2X with CO_2	**26.60%***	**26.07%***	**1759.41***	**5507.99***
3X with CO_2	**29.85%***	**25.90%***	**1973.99***	**5471.67***
Daily with CO_2	**33.38%***	**31.60%***	**2207.96***	**6674.99***

questing ticks in the lawn as nymphs and adults or all life stages are the same for this metric because of low densities of ticks in the lawn habitat (Table 7).

All scenarios were effective in reducing the sum of questing ticks in the yard for nymphs and adults, as well as all life stages combined and all scenarios were significantly different from the no control scenario. The most effective scenario

Table 8 Percent reduction and difference of the sum of questing ticks in the forest habitat at each life stage for each scenario. Numbers in bold represent an effective reduction of >10% and italicized numbers are the sum of questing ticks in the scenario with no control. Life stages of ticks are represented by "N+A" for nymphs and adults, and "All" for all life stages. Scenarios without the use of TickBot are marked with a dash (-). A one-way ANOVA was used to compare the results of the scenarios. Results significantly different from the no control scenario are noted by an asterisk (*)

	Percent reduction		Difference	
Scenarios	N+A	All	N+A	All
No Control	–	–	*14264.62*	*53781.14*
1X	0.05%	-0.20%	7.49	−108.05
2X	0.28%	2.14%	39.32	1152.35
3X	-0.02%	2.95%	−2.36	1588.05
Daily	0.32%	1.84%	45.00	988.79
1X with CO_2	0.11%	3.51%	15.87	1886.08
2X with CO_2	0.53%	3.45%	75.31	1857.56
3X with CO_2	0.00%	3.17%	−0.28	1703.97
Daily with CO_2	-0.06%	0.01%	−7.89	6.44

was daily treatment with the use of CO_2 with 33.38% reduction for nymphs and adults, and 31.60% reduction for all life stages combined. Differences in the sum of questing ticks also showed a reduction of thousands of ticks less with daily treatment with CO_2 (Table 7).

4.4 Sum of All Ticks

Percent reduction in the sum of all ticks, in any activity, in each habitat, for nymphs and adults combined (N+A) and all life stages, were calculated for each habitat. No scenario was significantly different from the no control scenario and no scenario was effective in reducing the sum of ticks in either the lawn or forest habitat (Tables 9 and 11). However, all scenarios were effective for the yard habitat and all scenarios were significantly different than the no control scenario (Table 10). The greatest percent reduction in the sum of all ticks resulted from daily application with the use of CO_2. Differences also support that the greatest reduction of ticks with the daily treatment scenario and CO_2 (Table 11).

Table 9 Percent reduction and difference of the sum of ticks in the lawn habitat at each life stage for each scenario. Numbers in bold represent an effective reduction of >10%. Italicized numbers are the sum of ticks in the scenario with no control. Life stages of ticks are represented by "N+A" for nymphs and adults, and "All" for all life stages. Scenarios without the use of TickBot are marked with a dash (-). Results significantly different from the no control scenario are noted by an asterisk (*)

	Percent reduction		Difference	
Scenarios	N+A	All	N+A	All
No Control	–	–	*92.38*	*1795.96*
1X	5.67%	2.82%	16.77	50.60
2X	7.25%	2.16%	21.42	38.79
3X	5.82%	3.47%	17.21	62.41
Daily	5.99%	2.56%	17.71	45.90
1X with CO_2	4.67%	1.68%	13.80	30.25
2X with CO_2	4.95%	2.71%	14.63	48.67
3X with CO_2	6.60%	2.60%	19.50	46.67
Daily with CO_2	8.99%	4.53%	26.59	81.39

Table 10 Percent reduction and difference of the sum of ticks in the yard habitat at each life stage for each scenario. Numbers in bold represent an effective reduction of >10%. Italicized numbers are the sum of ticks in the scenario with no control. Life stages of ticks are represented by "N+A" for nymphs and adults, and "All" for all life stages. Scenarios without the use of TickBot are marked with a dash (-). Results significantly different from the no control scenario are noted by an asterisk (*)

	Percent reduction		Difference	
Scenarios	N+A	All	N+A	All
No Control	–	–	*12464.65*	*43593.84*
1X	**14.44%***	**15.81%***	**1799.99***	**6891.95***
2X	**15.12%***	**13.54%***	**1885.03***	**5901.21***
3X	**17.60%***	**16.96%***	**2193.51***	**7392.99***
Daily	**18.24%***	**14.29%***	**2273.06***	**6228.27***
1X with CO_2	**18.07%***	**14.05%***	**2252.97***	**6125.32***
2X with CO_2	**21.54%***	**19.06%***	**2684.68***	**8307.78***
3X with CO_2	**23.99%***	**20.01%***	**2990.36***	**8725.03***
Daily with CO_2	**27.74%***	**23.63%***	**3457.27***	**10302.58***

4.5 Cost of Effort

To take into account the effort required to implement the TickBot, metrics for the yard habitat were adjusted by frequency of TickBot runs in each scenario (Tables 12, 13, and 14). Maximum density, sum of questing ticks, and sum of all ticks, were all

Table 11 Percent reduction and difference of the sum of ticks in the forest habitat at each life stage for each scenario. Numbers in bold represent an effective reduction of >10%. Italicized numbers are the sum of ticks in the scenario with no control. Life stages of ticks are represented by "N+A" for nymphs and adults, and "All" for all life stages. Scenarios without the use of TickBot are marked with a dash (-). Results significantly different from the no control scenario are noted by an asterisk (*)

	Percent reduction		Difference	
Scenarios	N+A	All	N+A	All
No control	–	–	*29295.69*	*113030.50*
1X	-0.44%	0.70%	−127.48	794.52
2X	-0.20%	2.25%	−58.12	2540.74
3X	-0.11%	3.04%	−31.24	3438.88
Daily	0.31%	1.89%	90.32	2131.43
1X with CO_2	-0.13%	3.35%	−39.35	3783.11
2X with CO_2	-0.40%	2.75%	−115.99	3103.50
3X with CO_2	0.77%	3.29%	225.49	3719.66
Daily with CO_2	0.45%	0.74%	132.61	832.24

Table 12 Percent reduction and difference in maximum density of questing ticks in the yard habitat adjusted by the cost of effort for each treatment. Numbers in bold represent an effective reduction of >10%. Italicized numbers are the densities of ticks in the scenario with no control. Life stages of ticks are represented by "L" for larvae, "N" for nymphs, "A" for adults, "All" for all life stages. Scenarios without the TickBot are marked with a dash (-)

	Percent reduction				Difference			
Scenarios	L	N	A	All	L	N	A	All
No Control	–	–	–	–	*10.26*	*3.30*	*0.54*	*10.26*
1X	**19.93%**	0.39%	**19.72%**	**19.93%**	**2.05**	0.01	**0.11**	**2.05**
2X	9.96%	−0.33%	**13.22%**	9.96%	1.02	−0.01	**0.07**	1.02
3X	8.20%	0.16%	9.17%	8.20%	0.84	0.01	0.05	0.84
Daily	2.86%	0.37%	4.09%	2.86%	0.29	0.01	0.02	0.29
1X with CO_2	**14.34%**	0.65%	**26.70%**	**14.34%**	**1.47**	0.02	**0.15**	**1.47**
2X with CO_2	**13.24%**	0.11%	**16.43%**	**13.24%**	**1.36**	0.00	**0.09**	**1.36**
3X with CO_2	7.95%	0.28%	**13.67%**	7.95%	0.82	0.01	**0.07**	0.82
Daily with CO_2	4.68%	0.30%	6.20%	4.68%	0.48	0.01	0.03	0.48

effectively reduced by the one time a week scenario with and without CO_2 scenarios. Some efficacy was seen for various life stages with twice a week, twice a week with CO_2, and three times a week with CO_2.

Table 13 Percent reduction and difference in the sum of questing ticks of questing ticks in the yard habitat adjusted by the cost of effort for each treatment. Numbers in bold represent an effective reduction of >10%. Italicized numbers are the densities of ticks in the scenario with no control. Life stages of ticks are represented by "L" for larvae, "N" for nymphs, "A" for adults, "All" for all life stages. Scenarios without the TickBot are marked with a dash (-)

	Percent reduction		Difference	
Scenarios	N+A	All	N+A	All
No Control	–	–	*6613.87*	*21126.49*
1X	**17.18%**	**21.97%**	**1136.31**	**4641.45**
2X	9.46%	9.42%	625.40	1990.94
3X	6.69%	7.86%	442.74	1660.52
Daily	3.13%	2.74%	207.16	579.78
1X with CO_2	**22.39%**	**19.16%**	**1481.04**	**4048.12**
2X with CO_2	**13.30%**	**13.04%**	**879.71**	**2754.00**
3X with CO_2	9.95%	8.63%	658.00	1823.89
Daily with CO_2	4.77%	4.51%	315.42	953.57

Table 14 Percent reduction and difference in the sum of ticks in the yard habitat adjusted by the cost of effort for each treatment. Numbers in bold represent an effective reduction of >10%. Italicized numbers are the densities of ticks in the scenario with no control. Life stages of ticks are represented by "L" for larvae, "N" for nymphs, "A" for adults, "All" for all life stages. Scenarios without the TickBot are marked with a dash (-)

	Percent reduction		Difference	
Scenarios	N+A	All	N+A	All
No control	–	–	*12464.65*	*43593.84*
1X	**14.44%**	**15.81%**	**1799.99**	**6891.95**
2X	7.56%	6.77%	942.52	2950.60
3X	5.87%	5.65%	731.17	2464.33
Daily	2.61%	2.04%	324.72	889.75
1X with CO_2	**18.07%**	**14.05%**	**2252.97**	**6125.32**
2X with CO_2	**10.77%**	9.53%	**1342.34**	4153.89
3X with CO_2	8.00%	6.67%	996.79	2908.34
Daily with CO_2	3.96%	3.38%	493.90	1471.80

5 Conclusions

The overall goal in tick control is to reduce human encounters with ticks. In this model, the yard habitat (consisting of the lawn, TickBot, and forest edge patches) represents the area of greatest risk for humans. Using the metrics explored in this model, there is unanimous support that the optimal scenario is daily treatment with the TickBot and CO_2. Logically, the more the TickBot is applied, the more ticks are killed, but the cost of running the TickBot can be cumbersome and time consuming if applied daily. When adjusted for effort the percent reduction metrics by the number of TickBot applications a week, the results highlight once a week

scenarios. Once a week treatments may provide some reduction in tick populations while limiting costs, but more work is needed to identify the specific application frequency based on cost effectiveness.

How effectiveness of a control method is measured can also change the interpretation of results. The metric of maximum density of questing ticks for this model resulted in effective population reductions of larvae and adults but not nymphs (Tables 4 and 5). The timing of the TickBot in this model targeted the adult phenology early in the summer. By decreasing adult populations, fewer eggs were laid which resulted in less larvae. However, nymphs typically are active much longer than adults in the summer, therefore the timing of TickBot would need to be adjusted to have effective at reducing nymph abundance.

Like all models, this one depends heavily on the parameters used. Future work is needed to complete a full sensitivity analysis of all possible parameters. Additional information is required for application of the TickBot to a specific area with additional host types, varying environmental conditions, etc. Finally, this model can be extended to run over multiple years to assess the long-term impact, if any, on the tick population.

Determining the efficacy of the TickBot at varying frequencies with or without CO_2 has several layers to deciding what scenario is best. If the focus of a control effort is to target a specific life stage, all life stages, or provide cost effective treatment, different answers will be required. Overall, this study found that the TickBot should be run at least three times a week, with CO_2, to have the maximum protection for the lowest cost.

Acknowledgments The parameter estimation studies would not have been possible without the help of several ODU tick team members including Amanda DeVleeschower, Armin Bahrani, Linda Carlsen, Sean Campbell, Alex Suarez, Allyn Bryan, Caitlin Murphy, Lindsey Bidder, and Jason McCain. Thank you to Hoffler Creek Wildlife Preserve, Eric Walters, and Peter Van Dyke at NASA Langley for the use of their sites as test areas for the TickBot. Thank you to James Squire at Virginia Military Institute for his ability to provide TickBot technical support. This research was also supported by the Turing High Performance Computing cluster at Old Dominion University.

References

1. Walter S Abbott. A method of computing the effectiveness of an insecticide. *Journal of Economic Entomology*, 18(2):265–267, 1925.
2. FC Bishopp and Helen Louise Trembley. Distribution and hosts of certain North American ticks. *The Journal of Parasitology*, 31(1):1–54, 1945.
3. James E Childs and Christopher D Paddock. The ascendancy of *Amblyomma americanum* as a vector of pathogens affecting humans in the United States. *Annual Review of Entomology*, 48(1):307–337, 2003.
4. Scott P Commins and Thomas AE Platts-Mills. Tick bites and red meat allergy. *Current Opinion in Allergy and Clinical Immunology*, 13(4):354, 2013.
5. Robert Allen Cooley and Glen M Kohls. The genus *Amblyomma* (Ixodidae) in the United States. *The Journal of Parasitology*, 30(2):77–111, 1944.

6. Lars Eisen and Marc C Dolan. Evidence for personal protective measures to reduce human contact with blacklegged ticks and for environmentally based control methods to suppress host-seeking blacklegged ticks and reduce infection with Lyme disease spirochetes in tick vectors and rodent reservoirs. *Journal of Medical Entomology*, 53(5):1063–1092, 2016.
7. Richard C Falco and Durland Fish. Horizontal movement of adult *Ixodes dammini* (Acari: Ixodidae) attracted to CO_2-baited traps. *Journal of Medical Entomology*, 28(5):726–729, 1991.
8. Holly Gaff. Preliminary analysis of an agent-based model for a tick-borne disease. *Mathematical Biosciences and Engineering: MBE*, 8(2):463–473, 2011.
9. Holly Gaff and Robyn Nadolny. Identifying requirements for the invasion of a tick species and tick-borne pathogen through TICKSIM. *Mathematical Biosciences and Engineering: MBE*, 10:625–635, 2013.
10. Holly D Gaff, Elsa Schaefer, and Suzanne Lenhart. Use of optimal control models to predict treatment time for managing tick-borne disease. *Journal of Biological Dynamics*, 5(5):517–530, 2011.
11. Holly D Gaff, Alexis White, Kyle Leas, Pamela Kelman, James C Squire, David L Livingston, Gerald A Sullivan, Elizabeth W Baker, and Daniel E Sonenshine. TickBot: a novel robotic device for controlling tick populations in the natural environment. *Ticks and Tick-borne Diseases*, 6(2):146–151, 2015.
12. Volker Grimm, Uta Berger, Finn Bastiansen, Sigrunn Eliassen, Vincent Ginot, Jarl Giske, John Goss-Custard, Tamara Grand, Simone K Heinz, Geir Huse, Andreas Huth, Jane U Jepsen, Christian Jrgensenc, Wolf M Mooij, Birgit Muller, Guy Pe'er, Cyril Piou, Steven F Railsback, Andrew M Robbins, Martha M Robbins, Eva Rossmanith, Nadja Ruger, Espen Strand, Sami Souissi, Richard A Stillman, Rune Vab, Ute Vissera, and Donald L DeAngelis. A standard protocol for describing individual-based and agent-based models. *Ecological Modelling*, 198(1):115–126, 2006.
13. Volker Grimm, Uta Berger, Donald L DeAngelis, J Gary Polhill, Jarl Giske, and Steven F Railsback. The ODD protocol: a review and first update. *Ecological Modelling*, 221(23):2760–2768, 2010.
14. Samniqueka J Halsey and James R Miller. A spatial agent-based model of the disease vector *Ixodes scapularis* to explore host-tick associations. *Ecological modelling*, 387:96–106, 2018.
15. Thomas M Kollars, James H Oliver, Edwin J Masters, Peggy G Kollars, and Lance A Durden. Host utilization and seasonal occurrence of *Dermacentor*. *Experimental and Applied Acarology*, 24(8):631–643, 2000.
16. MATLAB. *version 9.4.0.813654 (R2018a)*. The MathWorks Inc., Natick, Massachusetts, 2018.
17. Donald E Mock, Roger D Applegate, and Lloyd B Fox. Preliminary survey of ticks (Acari: Ixodidae) parasitizing wild turkeys (Aves: Phasianidae) in eastern Kansas. *Journal of Medical Entomology*, 38(1):118–121, 2001.
18. GA Mount and DG Haile. Computer simulation of area-wide management strategies for the lone star tick, *Amblyomma americanum* (Acari: Ixodidae). *Journal of Medical Entomology*, 24(5):523–531, 1987.
19. GA Mount, DG Haile, DR Barnard, and E Daniels. New version of LSTSIM for computer simulation of *Amblyomma americanum* (Acari: Ixodidae) population dynamics. *Journal of Medical Entomology*, 30(5):843–857, 1993.
20. RM Nadolny and HD Gaff. Modelling the effects of habitat and hosts on tick invasions. *Letters in Biomathematics*, 5(1):2–29, 2018.
21. Robyn M Nadolny, Chelsea L Wright, Daniel E Sonenshine, Wayne L Hynes, and Holly D Gaff. Ticks and spotted fever group rickettsiae of southeastern Virginia. *Ticks and Tick-borne Diseases*, 5(1):53–57, 2014.
22. CD Paddock and MJ Yabsley. Ecological havoc, the rise of white-tailed deer, and the emergence of *Amblyomma americanum*-associated zoonoses in the United States. In *Wildlife and Emerging Zoonotic Diseases: The Biology, Circumstances and Consequences of Cross-species Transmission*, pages 289–324. Springer, 2007.

23. Christopher D Paddock, RS Lane, E Staples, and MB Labruna. Changing paradigms for tick-borne diseases in the Americas. In *Global health impacts of vector-borne diseases: workshop summary, 21 September 2016, Washington D.C.*, pages 221–257. National Academies Press, 2016.
24. Terry L Schulze and Robert A Jordan. Meteorologically mediated diurnal questing of *Ixodes scapularis* and *Amblyomma americanum* (Acari: Ixodidae) nymphs. *Journal of Medical Entomology*, 40(4):395–402, 2003.
25. Daniel Sonenshine. *Biology of Ticks: Vol. 1*. Oxford University Press, New York, New York, 1991.
26. Hsiao-Hsuan Wang, WE Grant, and PD Teel. Simulation of climate–host–parasite–landscape interactions: a spatially explicit model for ticks (Acari: Ixodidae). *Ecological Modelling*, 243:42–62, 2012.
27. Hsiao-Hsuan Wang, WE Grant, PD Teel, and SA Hamer. Simulation of climate-tick-host-landscape interactions: effects of shifts in the seasonality of host population fluctuations on tick densities. *Journal of Vector Ecology*, 40(2):247–255, 2015.
28. Hsiao-Hsuan Wang, WE Grant, PD Teel, and SA Hamer. Tick-borne infectious agents in nature: Simulated effects of changes in host density on spatial-temporal prevalence of infected ticks. *Ecological Modelling*, 323:77–86, 2016.
29. Uri Wilensky. *NetLogo*. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL, 1999.

Simulations of Social Distancing Scenarios and Analysis of Strategies to Predict the Spread of COVID-19

Fran Sérgio Lobato, Gustavo B. Libotte, Gustavo M. Platt, Regina C. Almeida, Renato S. Silva, and Sandra M. C. Malta

1 Introduction

Computational models are crucial in an attempt to predict the dynamics of epidemics. In general, governmental policies aiming at mitigating the spread of infectious diseases may also be based on simulations from such models. By means of robust models, it is possible to estimate the reproduction capacity of a virus in a given location, in order to determine actions such as closing non-essential businesses, imposing restrictive measures in relation to the mobility of people across borders, or even total shut down. In view of the complexity of an epidemic and all the variables that can affect the course of the virus, creating models that incorporate all of these characteristics may be unfeasible. Despite this, it is essential that new strategies are continually developed in order to provide increasingly appropriate apparatus to properly analyze an epidemic.

In the early stages of the disease outbreak, Dehning et al. [4] proposed an approach to infer potential change points in the dynamics of spreading the virus, in order to simulate short-term forecasts associated with different levels of social distancing in Germany. Change points are meant to represent the moment at which

F. S. Lobato (✉)
Chemical Engineering Faculty, Federal University of Uberlândia, Uberlândia, Brazil
e-mail: fslobato@ufu.br

G. B. Libotte · R. C. Almeida · R. S. Silva · S. M. C. Malta
National Laboratory for Scientific Computing, Petrópolis, Brazil
e-mail: glibotte@lncc.br; rcca@lncc.br; rssr@lncc.br; smcm@lncc.br

G. M. Platt
Graduate Program in Agroindustrial Systems and Processes, School of Chemistry and Food, Federal University of Rio Grande, Santo Antônio da Patrulha, Brazil
e-mail: gmplatt@furg.br

B. Toni (ed.), *The Mathematics of Patterns, Symmetries, and Beauties in Nature*, STEAM-H: Science, Technology, Engineering, Agriculture, Mathematics & Health,
https://doi.org/10.1007/978-3-030-84596-4_5

governmental interventions are put into practice and change the behavior of the population in the face of the epidemic. They assume different degrees of stringency in relation to social distancing measures, being effectively initiated on particular days, and analyze the impact of such variations regarding the number of infected individuals. Following this framework, our objective is to analyze a delay observed in COVID-19 due to the interplay between the incubation period of the disease and testing. The delay scheme is formulated as an alternative to simulate the influence of disease latency in infected individuals using the SIR model, and to compare its effect to the parameter that represents the incubation period in the SEIR model. In addition, we provide detailed descriptions of the incorporation of a delay scheme into the statistical inference procedure and the strategy for simulating social distancing measures. We analyze possible scenarios of the epidemic in Italy, associating the influence of policy interventions with computer simulations and comparing how decisions under different conditions would impact the number of infected individuals.

Since the first cases of COVID-19 diagnosed in Italy, on January 31, 2020 [23], the country has been one of the most affected by the pandemic. After the disease initially spread across Asia, Italy was the first country in Europe to have an officially identified case [18]—although there is evidence that the virus was already circulating in the country at the end of 2019 [1]. Several measures were adopted by the Italian government in an attempt to curb the spread of the disease. More stringent measures, referred to as "red zone", were initially imposed on locations in the north of the country, affecting areas of Lombardy, Veneto, and Emilia-Romagna. In view of the gradual and alarming spread of the disease, such measures were extended to the entire national territory on March 9, with total lockdown. Figure 1 shows the timeline of some important events and policy interventions in the period between February 21 and March 22, as well as the number of new cases [6] reported on the corresponding time point. The events listed in this work follow the comprehensive study of Berardi et al. [2], which analyzes such policies in terms of their social impact regarding economic and health aspects, taking into account factors such as strictness and response time to determine interventions.

The effects of social distancing measures have been widely studied in the context of computer simulations. Predictions and analysis on possible scenarios can provide a basis for decision making, in addition to insights about eventual incidents. Given the dynamics of the disease in Italy, a great deal of effort has been devoted to analyze the effect of contingency measures and their effectiveness in mitigating the disease. Gatto et al. [7] analyzed the effect of such interventions by modeling in time and space the unfolding epidemic, considering mobility restrictions and contact reduction. Mobility data between provinces were adopted, and changes in social behavior that lead to increased social distancing were simulated taking into account the reduction in the transmission rate. Liu et al. [14] adopted an extension of the SEIR model, including symptomatic/asymptomatic and quarantined individuals. The authors proposed formulations subject to exponential decay in order to evaluate the reduction in the rate of contact and diagnostics over time, in the period in which containment measures are adopted. Reno et al. [18] and Supino et al. [21]

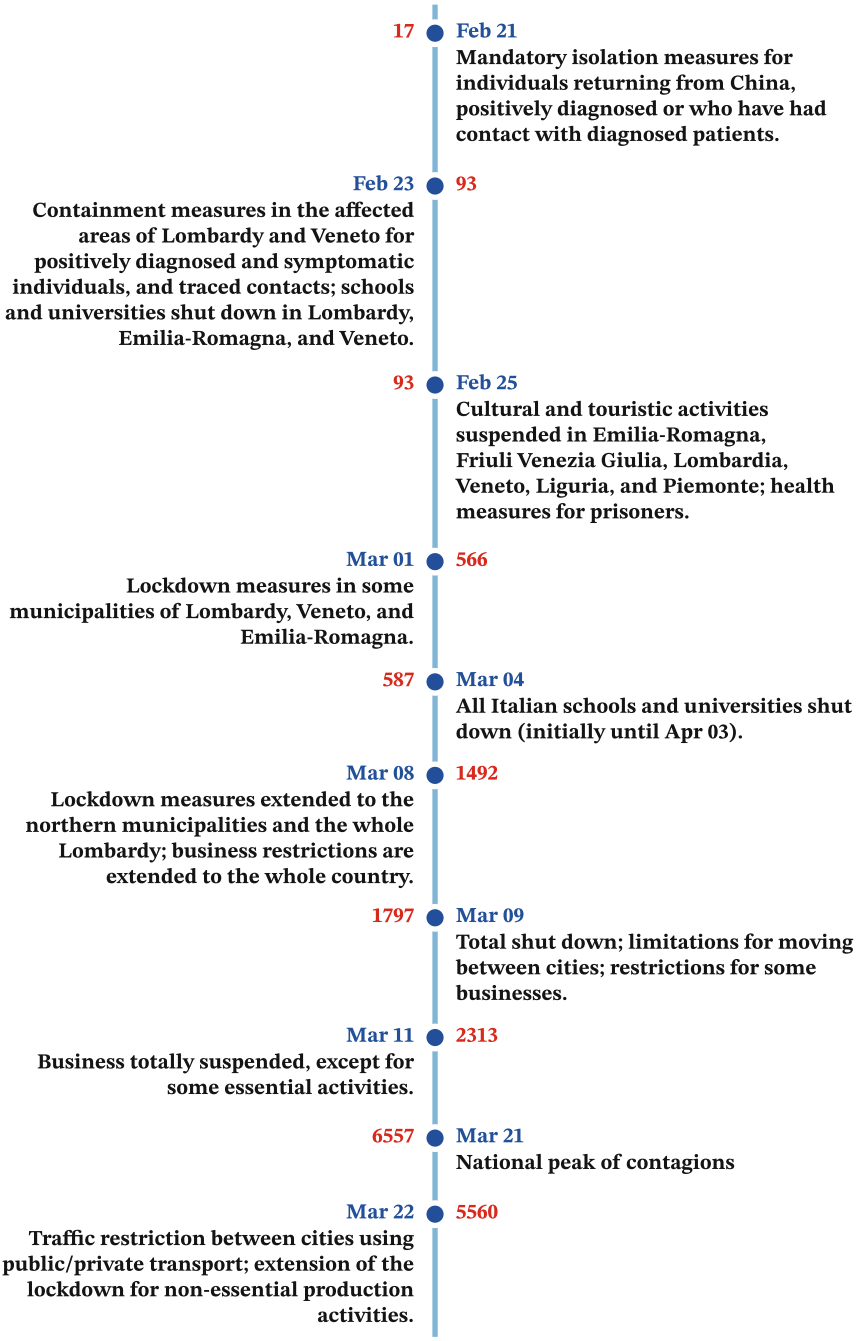


Fig. 1 Some relevant facts and policy interventions related to the spread of the SARS-COV-2 virus in Italy, from February 21 to March 22, 2020 [2]. The values shown in red are the number of new infections in the corresponding day [6]

assessed the effect of social distancing measures in the context of the impact on public health. The former analyzed the burden of hospitalizations under different scenarios regarding social distancing and the latter proposed a strategy for predicting the saturation of intensive care units. Traini et al. [22] combined the SEIR model with quarantine characteristics and time-dependent contact and diagnosis rates to estimate long-term scenarios taking into account a reproduction ratio threshold. Other analyzes in the same context were also discussed in Refs [5, 8, 9, 13, 20].

2 Materials and Methods

Below we present details about the compartmental models used in the simulations, as well as the description of the methodology for estimating parameters, analyzing scenarios related to the spread of the disease considering social distancing measures, and on the delay scheme to simulate the effect of the virus incubation period. We also provide information about the data.

2.1 Data

We adopt data on the number of new infections in Italy per day, between February 22 and March 8, 2020. The time frame was chosen taking into account the first day that the country had a significant number of infections, until the day before the national lockdown (which is the event that we are mainly concerned). Data are obtained from the COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University [6].

2.2 Models Description

The following two subsections present the compartmental models used to simulate the spreading dynamics of infectious diseases, as well as the framework for estimating parameters of the corresponding models.

2.2.1 Discrete-Time SIR Model

In one of the approaches to simulate the behavior of the epidemic, a discrete-time SIR model [11] is used, following Dehning et al. [4]. In this model, the rate of transmission per contact between susceptible (S) and infected (I) individuals, that is, the rate at which the disease spreads in the population, is given by β. In turn, the rate at which individuals are removed (R), for having recovered or died from the

disease, is given by γ. We disregard the birth rate and mortality from natural causes because, as the focus of the work is to analyze short-term scenarios, it is reasonable to consider that such events do not substantially affect the population. As data is reported daily, it is plausible to take a time step $\Delta t = 1$ day. With this assumption, the description of the model is given by

$$\begin{aligned} \Delta S &= -\beta \frac{S_{t-1} I_{t-1}}{N} \\ \Delta I &= \beta \frac{S_{t-1} I_{t-1}}{N} - \gamma I_{t-1} \\ \Delta R &= \gamma I_{t-1} \ , \end{aligned} \tag{1}$$

where ΔS represents the estimate of the number of susceptible individuals in two consecutive time steps, that is, $\Delta S = S_t - S_{t-1}$. The same notation holds equivalently to the other compartments. In turn, N stands for the total population size.

A useful indicator of the transmission of infectious diseases is the basic reproduction number, denoted by $\mathcal{R}_0$. In the case of the SIR model, $\mathcal{R}_0 = \beta/\gamma$.

2.2.2 Continuous SEIR Model

The main difference between this model and the one described in the previous section is the explicit representation of the exposed individuals. During the early phase after being infected, an individual may not present typical symptoms, perhaps still not being able to transmit the disease due to the low load of the pathogen in the organism. The latent period between the infected and infectious (once the level of parasite is sufficiently large within the host) states is called the incubation period. After this period, the exposed individual is moved to the infected compartment. Thus, let E be the compartment of exposed individuals and $1/\epsilon$ denote the incubation period. The continuous SEIR model is given by

$$\begin{aligned} \frac{\mathrm{d}S}{\mathrm{d}t} &= -\beta \frac{SI}{N} \\ \frac{\mathrm{d}E}{\mathrm{d}t} &= \beta \frac{SI}{N} - \epsilon E \\ \frac{\mathrm{d}I}{\mathrm{d}t} &= \epsilon E - \gamma I \\ \frac{\mathrm{d}R}{\mathrm{d}t} &= \gamma I \ , \end{aligned} \tag{2}$$

so that the other compartments and parameters are similar to those of the SIR model previously presented. Since this model does not explicitly include the compartment of dead individuals, the basic reproduction number is the same as the SIR model, $\mathcal{R}_0 = \beta/\gamma$.

2.3 Parameter Estimation Framework

Let $y\,(t_i,\ \boldsymbol{\theta})$ denote the model responses at different times t_i, for $i = 1,\ \ldots,\ p$, given the parameter vector $\boldsymbol{\theta}$. For now, assume that it represents an arbitrary vector of design parameters, which will be precisely defined later on. We may refer to the model responses simply as y_i and, in our case, they specifically represent the estimates of the number of infected individuals at time t_i. Likewise, we express an observable quantity as $\mathcal{D}_i$, which is equivalent to $\mathcal{D}\,(t_i)$, and $\boldsymbol{\mathcal{D}} = \{\mathcal{D}_1,\ \ldots,\ \mathcal{D}_p\}$ collectively assembles the time series of the number of infected individuals.

Estimators for the design parameters are calculated using a Bayesian approach. We are interested in obtaining the posterior probability distribution of the parameters $p_{\text{post}}\,(\boldsymbol{\theta} \mid \boldsymbol{\mathcal{D}})$, that is, the distribution that reflects the updated knowledge from the prior current knowledge $p_{\text{prior}}\,(\boldsymbol{\theta})$ and the likelihood function $p_{\text{like}}\,(\boldsymbol{\mathcal{D}} \mid \boldsymbol{\theta})$, which in turn measures how well the model responses fit the data available, for a given estimator. The posterior distribution is numerically obtained by using sampling methods according to $p_{\text{post}}\,(\boldsymbol{\theta} \mid \boldsymbol{\mathcal{D}}) \propto p_{\text{like}}\,(\boldsymbol{\mathcal{D}} \mid \boldsymbol{\theta})\, p_{\text{prior}}\,(\boldsymbol{\theta})$. In this work we assume a Gaussian likelihood, which is given by

$$p_{\text{like}}\,(\boldsymbol{\mathcal{D}} \mid \boldsymbol{\theta}) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\sum_{i=1}^{p} \frac{(\mathcal{D}\,(t_i) - y\,(t_i, \boldsymbol{\theta}))^2}{2\sigma^2}\right), \tag{3}$$

where σ is the scale factor and measures how far the model outcomes spread out from the data.

2.4 Model and Parameter Setups

This section presents specific information about the parameters of the models and the choices adopted for the Bayesian inference procedure. Details on the delay scheme and the strategy for simulating social distancing measures are also provided.

2.4.1 Model Parameters

The parameters to be estimated considering Eq. (1) are β and γ. For both parameters, Dehning et al. [4] propose the use of informative priors, following some

assumptions. In the early stages of the spread of the disease, it is reasonable to assume that $\mathcal{R}_0 > 1$. Particularly in the period analyzed in this work, D'Arienzo and Coniglio [3] estimate that $\mathcal{R}_0 \approx 3.1$ in Italy. Therefore, we assume prior distributions for such parameters given by $\beta \sim \text{Lognormal}\,(\log(0.4),\ 0.5)\ \text{day}^{-1}$ and $\gamma \sim \text{Lognormal}\,(\log(1/8),\ 0.2)\ \text{day}^{-1}$, respectively, whose medians are consistent with the corresponding estimate of the basic reproduction number. In the case of Eq. (2), the additional parameter to be estimated is the incubation period which, according to McAloon et al. [15], may be modeled with $\epsilon \sim \text{Lognormal}\,(1.63,\ 0.5)\ \text{day}^{-1}$.

2.4.2 Delay Scheme

The disease incubation period is artificially introduced into the SIR model through a delay strategy, as proposed by Dehning et al. [4]. In this way, the well-known effect of the exposed individuals, who are infected individuals but not yet infectious, is introduced by a delay time on the reported infected individuals, denoted by D. This new parameter is also capable of incorporating the time between an individual start to manifest symptoms of the disease until it is effectively tested and diagnosed. To combine an incubation period of approximately 5 days and an arbitrary choice of 3 days for the time until an individual is diagnosed, since the onset of symptoms, we choose the prior distribution $D \sim \text{Lognormal}\,(\log(8),\ 0.2)$ days.

Essentially, the effect of such a delay is directly considered in the model simulation. In general, the model is simulated for as many days as the amount of data available, for calibration purposes. However, the delay scheme inserts the need to simulate the model for a few more days beyond the day corresponding to the last data available in $\boldsymbol{\mathcal{D}}$. To describe the implementation issues, let $\boldsymbol{\mathcal{P}}$ be the vector containing the model responses for t_s days, considering a particular $\boldsymbol{\theta}$. In this case, $t_s = p + \xi$, where p is the cardinality of $\boldsymbol{\mathcal{D}}$ and ξ is the number of supplementary days required to evaluate the appropriate delay.

Simulating a delay D requires selecting p adjacent columns of the vector $\boldsymbol{\mathcal{P}}$, starting from column $D + 1$ but also considering that the maximum allowed delay is ξ. Of note, the definition of ξ avoids including low probability D values, thereby reducing the computational cost of the inference procedure. In order to select such simulations that take into account the value of D, consider a $t_s \times p$ binary matrix $\boldsymbol{\mathcal{Q}}$. This matrix is composed only of zeros in the first D lines. In turn, the block composed of the following p lines must be equal to the identity matrix with corresponding dimension. If $t_s > D + p$ or, equivalently, if $D < \xi$, the last $\xi - D$ lines remaining are also composed of zeros. Calculating $\boldsymbol{\mathcal{C}} = \boldsymbol{\mathcal{P}} \times \boldsymbol{\mathcal{Q}}$ consists of obtaining the vector composed of the p adjacent columns of $\boldsymbol{\mathcal{P}}$, which are subsequent to the first D columns. Therefore, a model response (considering a possible delay) that, in fact, is evaluated in the statistical inference procedure is given by $y_i = \mathcal{C}_i$. This procedure is repeated for all samples generated in the parameter estimation framework (see Sect. 2.3), considering different values of D

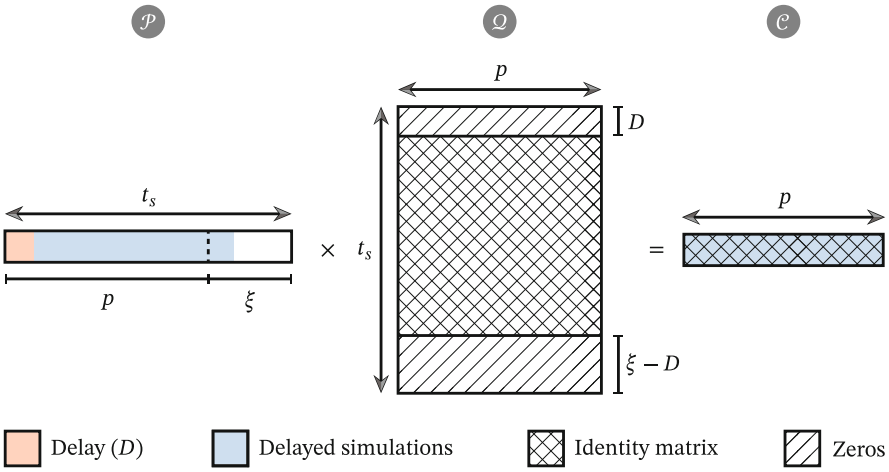


Fig. 2 Graphical representation of the delay scheme

and the other model parameters. Figure 2 shows a graphical representation of the delay scheme.

2.4.3 Social Distancing Scheme

Social distancing measures aim to reduce the rate of transmission of infectious diseases, whose main indicator is the basic reproduction number. Assuming a constant removal rate (given a relatively short period), scenarios for different levels of social distancing can be established by varying the spreading rate. Therefore, measures to contain the spread of the virus as a result of social distancing are incorporated into the model considering two factors [4]: (*i*) level of social distancing and; (*ii*) the number of days since such measures are implemented until they take effect.

Let ω be the parameter that simulates the level of social distancing on the population, with $0 \leq \omega < 1$, that is, high values within the feasible range represent more restrictive levels of social distancing, whereas $\omega = 0$ represents the scenario where there is no restriction on social distancing. In turn, let $\tau\,(t)$ be an expression that relates the number of days until social distancing measures start to take effect, that is, the number of days until the spreading rate starts to be reduced, in addition to the time needed so that the restriction has a gradual effect until reaching the maximum estimated effectiveness. Such expression is given by

$$\tau\,(t) = \min\left(\max\left(0,\ \frac{t-\delta}{\ell}\right),\ 1\right)\ , \tag{4}$$

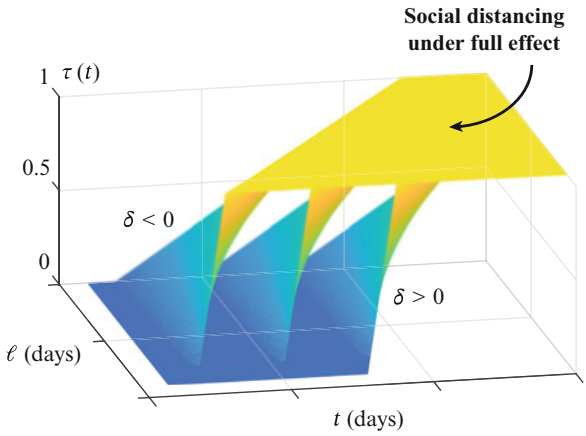


Fig. 3 Progressive effect of social distancing for different days on which measures are put into practice

where δ is the number of days until the transmission rate starts to decrease and ℓ is the time required for the measure to be gradually effective, as shown in Fig. 3. One may notice that Eq. (4) ensures that $\tau\,(t) \in [0,\ 1]$, for all t. The spreading rate considering such measures of social distancing varies with time according to

$$\beta_{\mathrm{d}}\,(t) = (1 - \omega\,\tau\,(t))\,\beta\ . \tag{5}$$

Therefore, scenarios considering measures of social distancing can be simulated by varying the value of ω, depending on how restrictive these actions must be, and also for different values of δ, according to the moment when the measures start to take effect.

This framework can be incorporated into any compartmental model only by replacing the parameter corresponding to the spreading rate of the disease (in this work, denoted as β) by the parameter $\beta_{\mathrm{d}}\,(t)$ in Eq. (5). For instance, the expression for the number of susceptible individuals in the SIR model given by Eq. (1) would become $\Delta S = -\beta_{\mathrm{d}}\,(t)\,S_{t-1}\,I_{t-1}/N$, and the rest of the equations would be modified accordingly.

2.4.4 Parameter Setting

In order to determine the initial conditions for both models, first consider the Italian population, $N = 60.4$ million inhabitants [24]. In the case of the SIR model, to reduce possible uncertainties associated with the number of individuals infected at the beginning of the epidemic, we follow Dehning et al. [4] and estimate the value of $I\,(0)$, instead of selecting the first value in set $\mathcal{D}$, corresponding to $t = 0$. The prior distribution is given by $I\,(0) \sim \mathrm{HalfCauchy}\,(100)$ individuals. In turn, since $R\,(0) \to 0$ at the onset of the pandemic, and $N \gg R\,(0)$, it is reasonable to define the initial number of susceptible individuals $S\,(0) = N - I\,(0)$. Additionally, we

Table 1 Variables associated with the parameters of the inference problem and corresponding prior distributions, in appropriate units (a.u.)

Variable	Parameter	Prior distribution		
		SIR	SEIR	
β	Spreading rate	Lognormal (log (0.4) , 0.5)		
γ	Removal rate	Lognormal (log (1/8) , 0.2)		
D	Simulation delay	Lognormal (log (8) , 0.2)	—	
ϵ	Incubation period	—	Lognormal (1.63, 0.5)	1/5.8
$E\,(0)$	Initially exposed	—	HalfCauchy (100)	
$I\,(0)$	Initially infected	HalfCauchy (100)	—	
σ	Scale factor	HalfCauchy (10)		

also estimate the scale factor σ in Eq. (3) as an hyperparameter of the model, in such a way that $\sigma \sim$ HalfCauchy (10) individuals. Therefore, the vector of variables to be estimated in the SIR model is $\boldsymbol{\theta} = (\beta,\ \gamma,\ D,\ I\,(0)\,,\ \sigma)$.

Regarding the SEIR model, as the number of exposed individuals on the date corresponding to the beginning of the simulations is not known, $E\,(0)$ must be estimated, with prior distribution given by $E\,(0) \sim$ HalfCauchy (100) individuals. In turn, aiming to avoid parameter identification problems [12, 16, 17] in this case, we adopt $I\,(0) = \mathcal{D}_1$. Since it also holds that $R\,(0) = 0$, the number of susceptible individuals initially is $S\,(0) = N - (E\,(0) + I\,(0))$. For the incubation period, we consider two possibilities: in the first, we take ϵ as a biological parameter and simulate the model considering $1/\epsilon = 5.8$ days [15]; in the second, we estimate the value of the incubation period according to the prior distribution defined in Sect. 2.4.1. Here we also estimate the scale factor σ. Table 1 summarizes the parameters to be estimated for each problem, as well as the associated prior distributions.

3 Results and Discussion

The PyMC3 library [19] is employed to perform the sampling procedure and parameter inference in all cases analyzed here. Parameter sampling is conducted by NUTS (No-U-Turn Sampler) [10], an efficient extension of the Hamiltonian Monte Carlo (HMC) algorithm that makes use of the function gradient information and eliminates the need to define the number of steps for HMC. For all simulations, we choose the number of supplementary simulations $\xi = 20$ days, in order to support the choice of the prior distribution of D. Additionally, the progressive effect of social distancing measures is calculated using $\ell = 7$ days.

3.1 Analysis on the Effect of Social Distancing

Initially we are interested in analyzing the effect of the social distancing scheme shown in Sect. 2.4.3, for different levels of stringency using the SIR model. Therefore, it is required to estimate the model parameters, so that it is possible to carry out simulations considering different values of ω. Figure 4 shows the posterior distribution of the parameters, as well as the maximum a posterior (MAP) value and the 95% credible interval. The MAP value of each parameter is used to simulate the corresponding model in all cases. For this case, note that $D = 8.28$, which is consistent with the physical meaning of the parameter [4], covering approximately five days corresponding to the virus incubation period, and three more regarding the time from the first symptoms to diagnosis. In addition, β and γ values lead to $\mathcal{R}_0 = 2.9109$, agreeing with D'Arienzo and Coniglio [3].

The effects of social distancing are analyzed considering three scenarios related to the strictness of the measures: extensive, moderate, and mild. These scenarios are different from each other by the particular value of ω in Eq. (4), which are taken as 0.9, 0.7, and 0.5, respectively. Of note, we do not consider the scenario in which $\omega = 0$ because, during an epidemic in which individuals are aware of the existence of the virus, a portion of the population tends to adopt social distancing measures on their own, even if they have not been imposed by health authorities. Therefore, this scenario could be unrealistic. Figure 5 shows the simulations of the SIR model

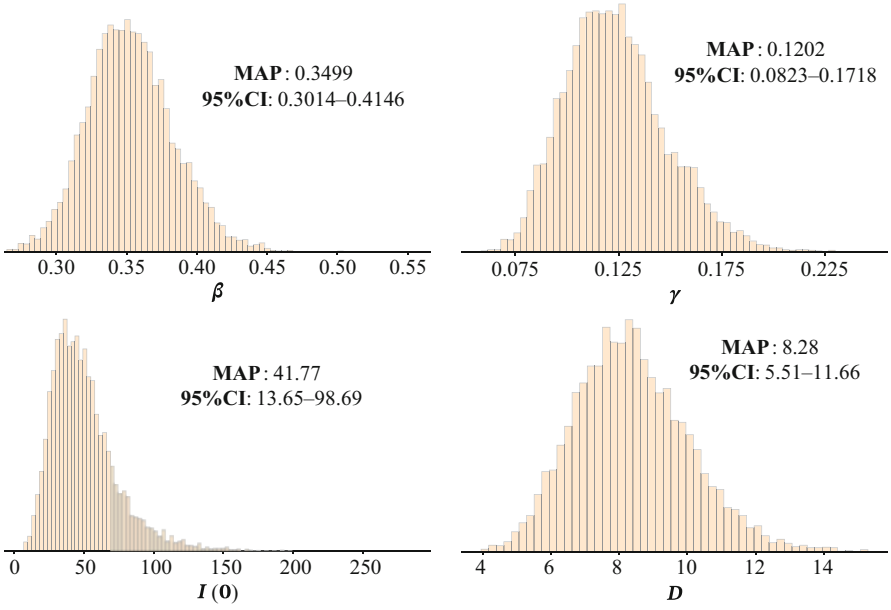


Fig. 4 Posterior distribution histograms, maximum a posterior (MAP), and credible interval of the parameters of SIR model (in a.u.)

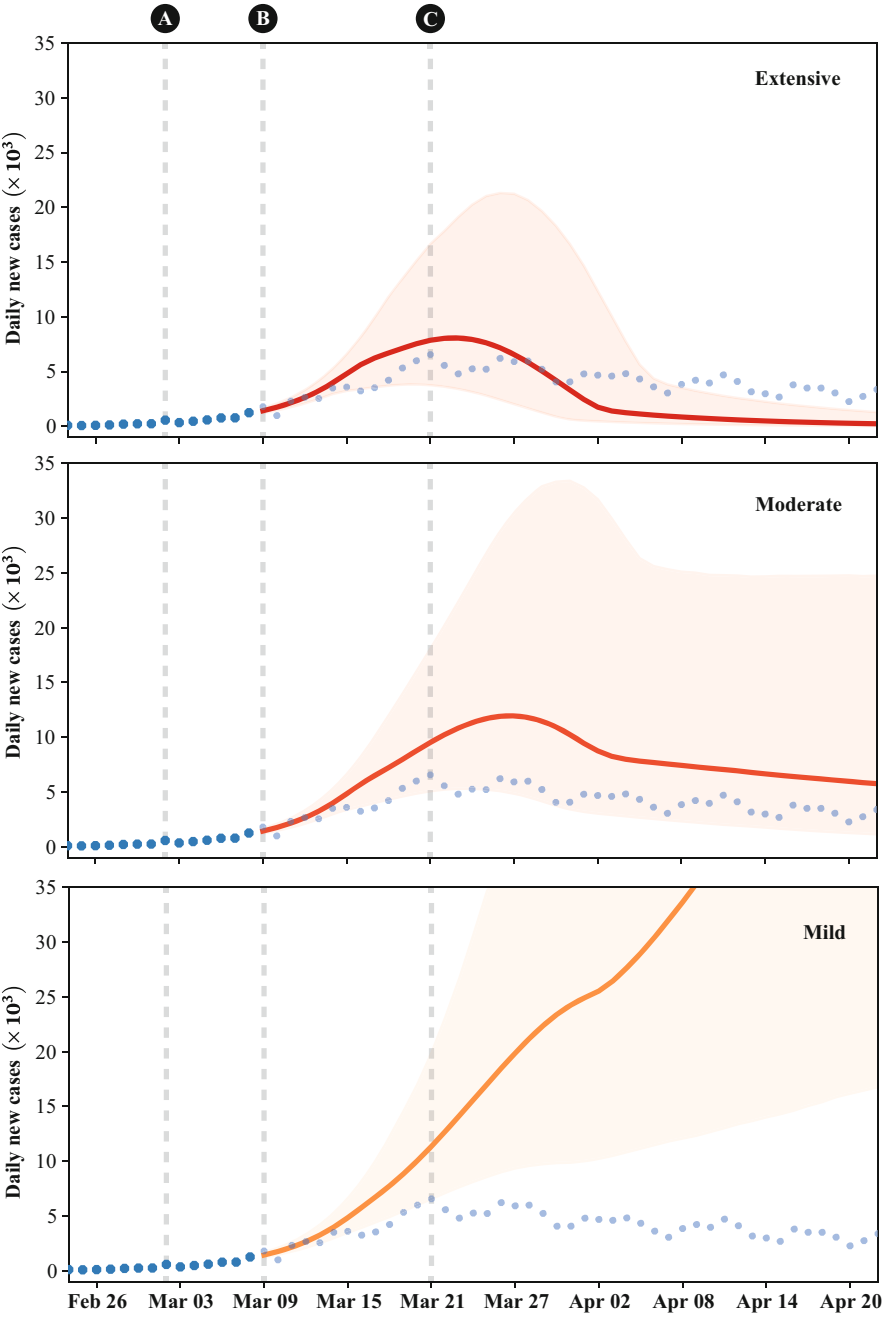


Fig. 5 Scenarios related to daily new cases of COVID-19 in Italy for different levels of social distancing. The upper frame shows the case in which extensive social distancing is adopted, in the middle frame a moderate social distancing is adopted, and in the lower frame a mild social distancing is adopted. We highlight three moments (dashed vertical lines) related to the timeline shown in Fig. 1: A (Mar 01) is when the lockdown measures in some municipalities of Lombardy, Veneto and Emilia-Romagna; B (Mar 09) indicates when it was experienced total shut down and; C (Mar 21) corresponds to the day of the national peak of contagions. The dark blue points are data used for calibrating the model, and the light points are plotted for comparison purposes only

in terms of the number of daily infected individuals and considering the parameters presented in Fig. 4, for the three scenarios analyzed. We associate three facts related to the policies adopted in Italy with the corresponding dates (see Fig. 1), which are indicated by the letters A to C, and we also show the data from March 9, represented by light blue dots, just for the sake of comparison.

The comparison of the two upper frames in Fig. 5, which show the simulations considering the effect of extensive and moderate social distancing, indicates that these adopted measures had a intermediate effect between these two thresholds, in view of the known data. However, it is noteworthy the reduction of the credible interval when extensive social distancing measures are implemented. In turn, the bottom frame illustrates how mild measures can be inefficient to contain the spread of the disease. In this case, the spread rate of the disease is greater than the effective capacity of the social distancing measures to contain the spread of the virus. The increase in the number of cases could persist until more restrictive measures are adopted, or until the curve reaches saturation according to the total population size. This obviously would have a direct effect on the number of deaths and the capacity of the country's health services.

Note that the simulation for extensive social distancing measures is in good agreement with the peak of contagion in the country (marker C). In the other considered scenarios, in addition to the number of cases being greater, the peak of the curve is also shifted to later dates, causing the contagion to continue for a longer time. Despite the analysis considering levels of social distancing, quantified by ω, being equally spaced among themselves, in addition to the fact that $\tau\,(t)$ in Eq. (4) has a linear relation with δ and ℓ, milder social distancing measures do not mean a linear increasing in the number of cases, given the exponential nature of the contagion dynamics between individuals. Note also that social distancing measures restricted to specific locations, as in the case of the intervention indicated by marker A, may not be as effective as expected, given the scenario illustrated by the bottom frame in Fig. 5.

We also analyze the effects of social distancing measures implemented on different days. Taking the more restrictive scenario in which we consider here ($\omega = 0.9$), the model is simulated with $\delta = -5$ days and $\delta = 5$ days, that is, the measures hypothetically start to take effect five days before or after the day when, in fact, they have been established, respectively. Figure 6 shows the simulations for each of these scenarios, where the dashed vertical line corresponds to the value of δ in each frame, in relation to the day the total shut down was decreed (March 9).

Even with a difference of relatively few days, the late implementation of social distancing measures can have a very harmful impact. This is because social distancing has a gradual effect and until the state in which the population effectively changes its behavior according to the proposed measures is not reached, new infections continue to occur. The late implementation of the social distancing measures makes the peak of the curve of daily new cases reach approximately 19,815 infections on March 28. On the other hand, the more agile decision reduces

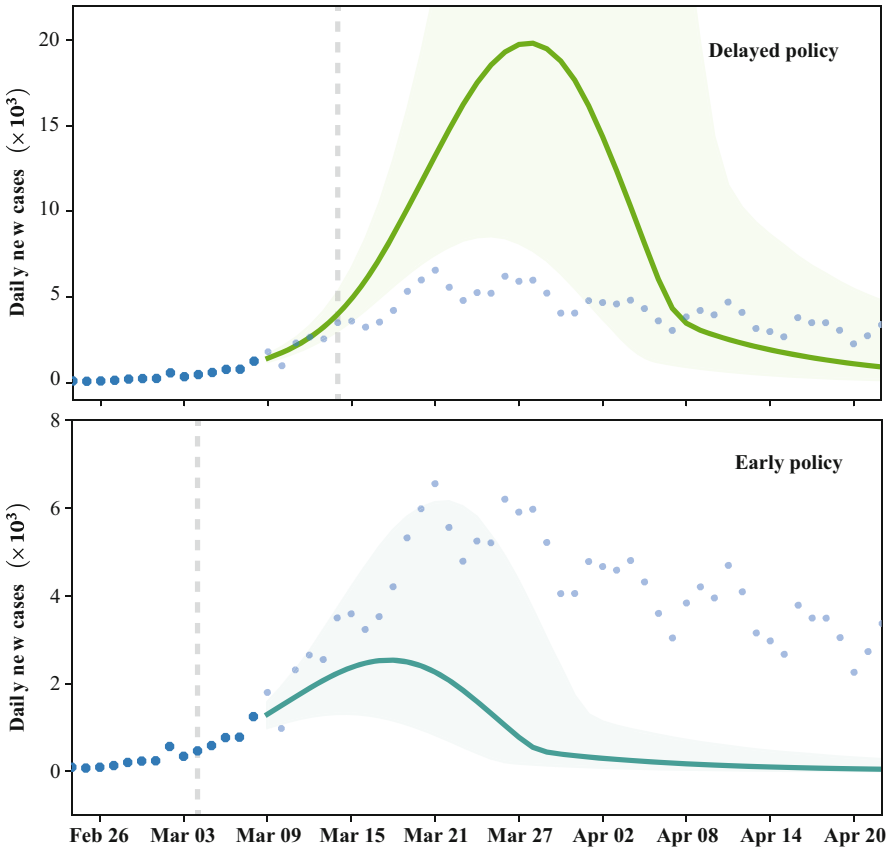


Fig. 6 Hypothetical scenarios related to the moment at which extensive social distancing measures start to take effect. In the upper frame, the measures are put into practice five days after the date when, in fact, the total shut down was decreed (March 9); in the lower frame, the measures are put into practice five days before this date. The dashed vertical lines indicate such moments

the peak of the curve to 2,534 on March 18. In the range of days analyzed, the total number of cases varies between 49,016 and 368,740 for such scenarios.

3.2 Comparison of Strategies to Simulate the Effect of Latency

Now we present an analysis of the scenarios proposed in this work in terms of the levels of stringency of social distancing, using the continuous SEIR model. The purpose is to assess the difference in effect caused by the incubation period parameter,

compared to simulations conducted using the delay scheme. As mentioned before, we consider both the cases in which the incubation period must be estimated and being taken as a biological parameter. The inference methodology and the other parameters of the model are the same adopted in the previous section.

First consider the problem in which ϵ must be estimated. Figure 7 shows the posterior distribution of the parameters, as well as the simulations for some ω values. By visual inspection, one may notice that the β and γ distributions are quite similar with those shown in Fig. 4, in terms of the respective mean and standard deviation. However, $1/\epsilon = 3.2481$ (days) is less than the delay inferred in the previous problem ($D = 8.28$ days). Therefore, comparing the simulations of Figs. 5 and 7, the effect of each of the latency approaches is clear: since ϵ does not consider the days from the appearance of the symptoms to the positive diagnosis, but only the virus incubation period, the tendency is for individuals to move from the exposed to the infected compartment more quickly. Taking into account that the spreading and removal rates are approximately equal in both cases, simulations whose latency is lower tend to produce flatter curves. This behavior occurs because, as the social distancing measures affect the spreading rate (β), which, in turn, directly interferes in the exposed (SEIR) and infected (SIR) compartments, the curves of daily new cases tend to undergo such a change. This is in line with the spreading dynamics of the virus, since symptomatic individuals before being positively diagnosed—when they tend to comply with isolation measures more restrictively—are more likely to infect other people.

In turn, consider the case where ϵ is taken as a biological parameter, such that $1/\epsilon = 5.8$ days [15]. In this case, the only parameters to be inferred are the spreading and removal rates, in addition to the initial number of exposed individuals (see Table 1). As ϵ is fixed, the parameter estimation procedure has to adjust the model to the observable data, considering such effect. This is evidently reflected in the posterior distribution of the adjustable parameters, as can be seen in Fig. 8. The MAP values of β and γ produce $\mathcal{R}_0 \approx 6.74$. Since the basic reproduction number is greater than that inferred by D'Arienzo and Coniglio [3], social distancing measures have to be even more strict to achieve a result similar to the previous case. When analyzing the simulations in Fig. 8, it is clear that social distancing measures (especially moderate and mild) do not have the expected effect in the analyzed period. Since $\mathcal{R}_0$ is very large, the effect of reducing the spread rate by means of Eq. (5) is more effective in the long run.

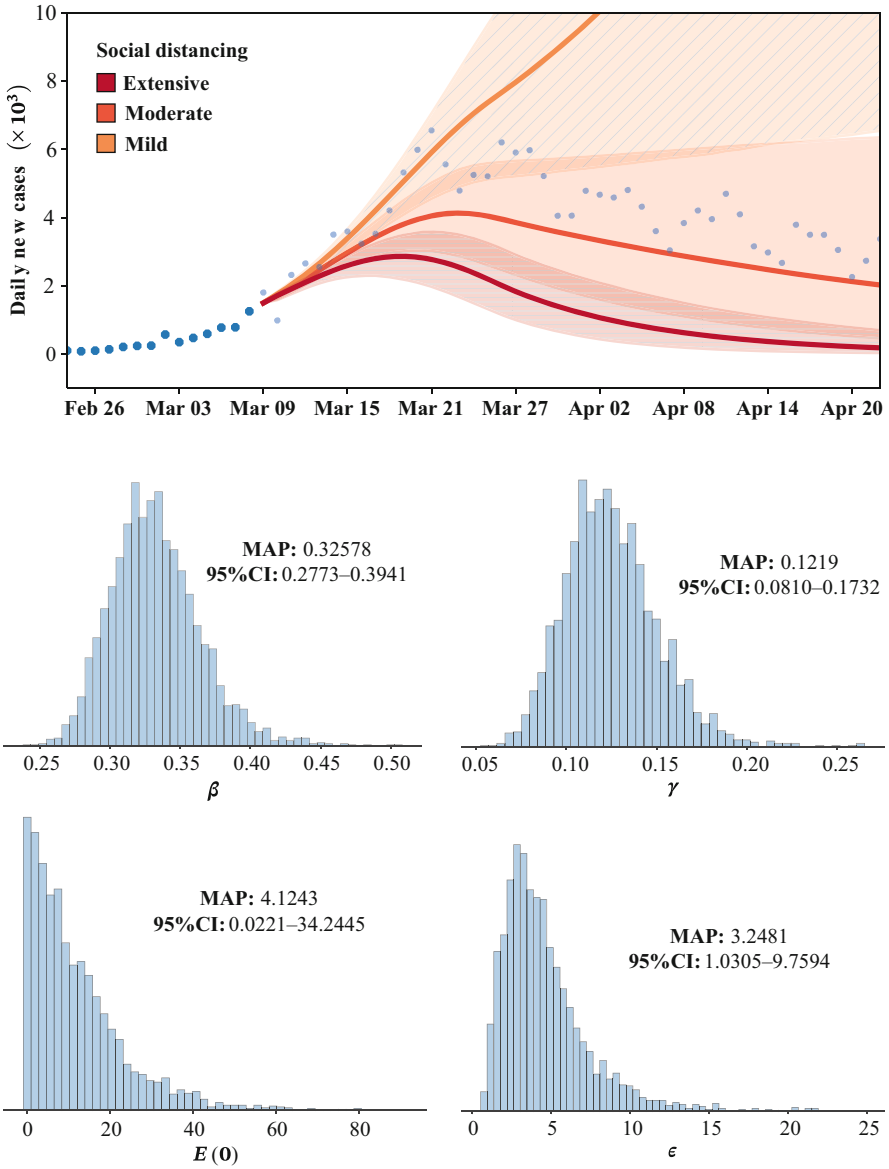


Fig. 7 Effect of social distancing measures considering some levels of stringency (extensive, moderate, and mild), simulated using the continuous SEIR model. The upper frame shows the daily new cases in each scenario. The textures in the shaded areas are for viewing purposes of better identifying the credible interval. In the lower frame, the posterior distribution histograms of the estimated parameters are shown

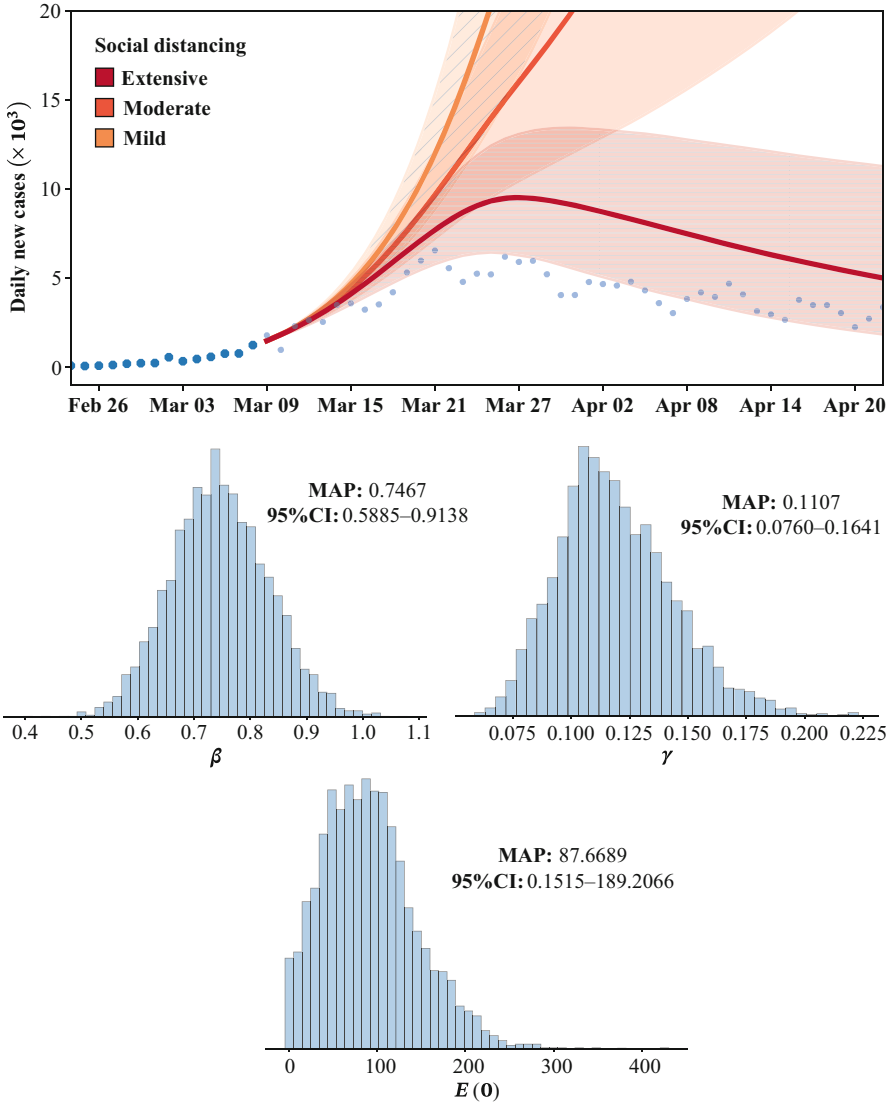


Fig. 8 Effect of social distancing measures considering some levels of stringency (extensive, moderate, and mild), simulated using the continuous SEIR model. The upper frame shows the daily new cases in each scenario. The textures in the shaded areas are for viewing purposes of better identifying the credible interval. In the lower frame, the posterior distribution of the estimated parameters are shown. In these simulations, the incubation period parameter is taken as a biological parameter, such that $1/\epsilon = 5.8$ days [15]

4 Conclusions

The strategy for simulating social distancing and delay scheme proposed by Dehning et al. [4], applied in the analysis of scenarios regarding the epidemic in Italy, show that both are capable of providing satisfactory prior knowledge of the dynamics of epidemics. Both are simple techniques to be incorporated into any simulation with compartmental models and, if associated with other more complex models, can provide important insights for decision making policies. Particularly on the delay scheme, its use can be upheld by the fact that it is possible to incorporate into any compartmental model, however simple it may be, the disease latency without having to explicitly include the compartment of exposed individuals. This tends to make the procedure of parameter inference more simple, since it eliminates the need to calibrate the initial condition associated with the exposed individuals.

Acknowledgments The authors would like to thank the Ministry of Science, Technology, Innovation and Communication (MCTIC) of Brazil. Gustavo Libotte is supported by a postdoctoral fellowship from the Institutional Training Program (PCI) of the Brazilian National Council for Scientific and Technological Development (CNPq), grant number 303185/2020-1.

References

1. Amendola, A., Bianchi, S., Gori, M., Colzani, D., Canuti, M., Borghi, E., Raviglione, M.C., Zuccotti, G.V., Tanzi, E.: Evidence of SARS-CoV-2 RNA in an Oropharyngeal Swab Specimen, Milan, Italy, early December 2019. Emerging Infectious Diseases **27**(2) (2021). https://doi.org/10.3201/eid2702.204632
2. Berardi, C., Antonini, M., Genie, M.G., Cotugno, G., Lanteri, A., Melia, A., Paolucci, F.: The COVID-19 pandemic in Italy: Policy and technology impact on health and non-health outcomes. Health Policy and Technology **9**(4), 454–487 (2020). https://doi.org/10.1016/j.hlpt.2020.08.019
3. D'Arienzo, M., Coniglio, A.: Assessment of the SARS-CoV-2 basic reproduction number, $\mathcal{R}_0$, based on the early phase of COVID-19 outbreak in Italy. Biosafety and Health **2**(2), 57–59 (2020). https://doi.org/10.1016/j.bsheal.2020.03.004
4. Dehning, J., Zierenberg, J., Spitzner, F.P., Wibral, M., Neto, J.P., Wilczek, M., Priesemann, V.: Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. Science **369**(6500), eabb9789 (2020). https://doi.org/10.1126/science.abb9789
5. Dickson, M.M., Espa, G., Giuliani, D., Santi, F., Savadori, L.: Assessing the effect of containment measures on the spatio-temporal dynamic of COVID-19 in Italy. Nonlinear Dynamics **101**(3), 1833–1846 (2020). https://doi.org/10.1007/s11071-020-05853-7
6. Dong, E., Du, H., Gardner, L.: An interactive web-based dashboard to track COVID-19 in real time. The Lancet Infectious Diseases **20**(5), 533–534 (2020). https://doi.org/10.1016/S1473-3099(20)30120-1
7. Gatto, M., Bertuzzo, E., Mari, L., Miccoli, S., Carraro, L., Casagrandi, R., Rinaldo, A.: Spread and dynamics of the COVID-19 epidemic in Italy: Effects of emergency containment measures. Proceedings of the National Academy of Sciences **117**(19), 10484–10491 (2020). https://doi.org/10.1073/pnas.2004978117
8. Gregori, D., Azzolina, D., Lanera, C., Prosepe, I., Destro, N., Lorenzoni, G., Berchialla, P.: A first estimation of the impact of public health actions against COVID-19 in Veneto (Italy).

Journal of Epidemiology and Community Health pp. jech–2020–214209 (2020). https://doi.org/10.1136/jech-2020-214209
9. Guzzetta, G., Riccardo, F., Marziano, V., Poletti, P., Trentini, F., Bella, A., Andrianou, X., Del Manso, M., Fabiani, M., Bellino, S., Boros, S., Urdiales, A.M., Vescio, M.F., Piccioli, A., Brusaferro, S., Rezza, G., Pezzotti, P., Ajelli, M., Merler, S.: Impact of a nationwide lockdown on SARS-CoV-2 transmissibility, Italy. Emerging Infectious Diseases **27**(1), 267–270 (2021). https://doi.org/10.3201/eid2701.202114
10. Hoffman, M.D., Gelman, A.: The No-U-Turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. Journal of Machine Learning Research **15**(47), 1593–1623 (2014). https://doi.org/10.5555/2627435.2638586
11. Keeling, M.J., Rohani, P.: Modeling Infectious Diseases in Humans and Animals, 1 edn. Princeton University Press, Princeton (2008)
12. Kreutz, C., Raue, A., Kaschek, D., Timmer, J.: Profile likelihood in systems biology. FEBS Journal **280**(11), 2564–2571 (2013). https://doi.org/10.1111/febs.12276
13. Lilleri, D., Zavaglio, F., Gabanti, E., Gerna, G., Arbustini, E.: Analysis of the SARS-CoV-2 epidemic in Italy: The role of local and interventional factors in the control of the epidemic. PLOS ONE **15**(11), e0242305 (2020). https://doi.org/10.1371/journal.pone.0242305
14. Liu, P.Y., He, S., Rong, L.B., Tang, S.Y.: The effect of control measures on COVID-19 transmission in Italy: Comparison with Guangdong province in China. Infectious Diseases of Poverty **9**(1), 130 (2020). https://doi.org/10.1186/s40249-020-00730-2
15. McAloon, C., Collins, Á., Hunt, K., Barber, A., Byrne, A.W., Butler, F., Casey, M., Griffin, J., Lane, E., McEvoy, D., Wall, P., Green, M., O'Grady, L., More, S.J.: Incubation period of COVID-19: a rapid systematic review and meta-analysis of observational research. BMJ Open **10**(8), e039652 (2020). https://doi.org/10.1136/bmjopen-2020-039652
16. Raue, A., Kreutz, C., Maiwald, T., Bachmann, J., Schilling, M., Klingmüller, U., Timmer, J.: Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. Bioinformatics **25**(15), 1923–1929 (2009). https://doi.org/10.1093/bioinformatics/btp358
17. Raue, A., Kreutz, C., Theis, F.J., Timmer, J.: Joining forces of Bayesian and frequentist methodology: a study for inference in the presence of non-identifiability. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **371**(1984), 20110544 (2013). https://doi.org/10.1098/rsta.2011.0544
18. Reno, C., Lenzi, J., Navarra, A., Barelli, E., Gori, D., Lanza, A., Valentini, R., Tang, B., Fantini, M.P.: Forecasting COVID-19-associated hospitalizations under different levels of social distancing in Lombardy and Emilia-Romagna, Northern Italy: Results from an extended SEIR compartmental model. Journal of Clinical Medicine **9**(5), 1492 (2020). https://doi.org/10.3390/jcm9051492
19. Salvatier, J., Wiecki, T.V., Fonnesbeck, C.: Probabilistic programming in Python using PyMC3. PeerJ Computer Science **2**, e55 (2016). https://doi.org/10.7717/peerj-cs.55
20. Sebastiani, G., Massa, M., Riboli, E.: COVID-19 epidemic in Italy: evolution, projections and impact of government measures. European Journal of Epidemiology **35**(4), 341–345 (2020). https://doi.org/10.1007/s10654-020-00631-6
21. Supino, M., D'Onofrio, A., Luongo, F., Occhipinti, G., Dal Co, A.: The effects of containment measures in the Italian outbreak of COVID-19. BMC Public Health **20**(1), 1806 (2020). https://doi.org/10.1186/s12889-020-09913-w
22. Traini, M.C., Caponi, C., Ferrari, R., De Socio, G.V.: A study of SARS-CoV-2 epidemiology in Italy: from early days to secondary effects after social distancing. Infectious Diseases **52**(12), 866–876 (2020). https://doi.org/10.1080/23744235.2020.1797157
23. Vicentini, C., Bordino, V., Gardois, P., Zotti, C.: Early assessment of the impact of mitigation measures on the COVID-19 outbreak in Italy. Public Health **185**, 99–101 (2020). https://doi.org/10.1016/j.puhe.2020.06.028
24. Worldometer: World population. URL https://www.worldometers.info/world-population/italy-population/. Visited on 18 Jan 2020

Mathematical Modelling of the Evolution Dynamics of the Coronavirus Disease 2019 (COVID-19) in Burkina Faso

Aboudramane Guiro, Blaise Koné, and Stanislas Ouaro

1 Introduction

Since its appearance in China in December 2019, the coronavirus disease 2019 (COVID-19) has been subject to intensive activities in the field of mathematical modelling [1–5]. Modelling is done to allow a better understanding of the evolution dynamics of the disease. Several models have been proposed, some to describe the dynamics, other to estimate the parameters, but all aim at allowing decision makers to take appropriate measures in dealing with the epidemic. In fact, mathematical models play a very important role in the understanding of the spread of several diseases. The advent of COVID-19 is so another opportunity for mathematical modelers to translate the results of their modelling into clear terms for decision makers. The majority of Western countries have relied on mathematical models to predict the spread of the disease in their countries [1, 5–8]. This has allowed them to take measures ranging from hand washing to general containment. Similarly, on the African continent, many leaders rely on mathematical models for the management of epidemic such as COVID-19.

A. Guiro
Département de Mathématiques, UFR Sciences et Techniques, Université Nazi Boni, Bobo-Dioulasso, Burkina Faso
e-mail: abouguiro@yahoo.fr

B. Koné · S. Ouaro (✉)
LAboratoire de Mathématiques et Informatique (LAMI), Université Joseph Ki-Zerbo, Ouagadougou, Burkina Faso
e-mail: leizon71@yahoo.fr; ouaro@yahoo.fr

B. Toni (ed.), *The Mathematics of Patterns, Symmetries, and Beauties in Nature*, STEAM-H: Science, Technology, Engineering, Agriculture, Mathematics & Health,
https://doi.org/10.1007/978-3-030-84596-4_6

In Burkina Faso, a commitee of experts comprising various competences including mathematicians was set up as soon as the first case of coronavirus disease was reported in the country in March 2020. Since this date, the authorities' concern was to know the dynamics of the pandemic and to see how to act to eliminate it. In doing so, several exchanges took place with public health doctors. It was at the end of these discussions that it was decided to develop a model which highlights the compartments of infected persons reported and not reported. Also it was a question of putting the accent on the high-risk hospitalized and quarantine or low-risk hospitalized. This explains the adoption of such a model study in this paper. The model concerning this paper make it possible to do projections and especially to see the effects of confinements, quarantines and cover fires in the country. Finally, the parameters used in this model for the simulations are parameters specific to the health context of Burkina Faso.

The difference between the former model (see [2]) and this one is that the new model takes into account some concerns of hospital practitioners and health epidemiologists from the Burkina Faso National Health Commission against COVID-19, people, of people hospitalized with serious health situation and death cases. We consider that dead people can contaminate health care workers or their loved ones when handling the bodies.

The paper is organized along 6 Sections. In the Sect. 2, we show the mathematical model, Sect. 3 is devoted to the basic properties. We present real daily data given by Burkina Faso National Health Commission against Coronavirus Disease 2019 (COVID-19) in Sect. 4 and we make simulations according to these data in Sect. 5. In Sect. 6 we end by a conclusion.

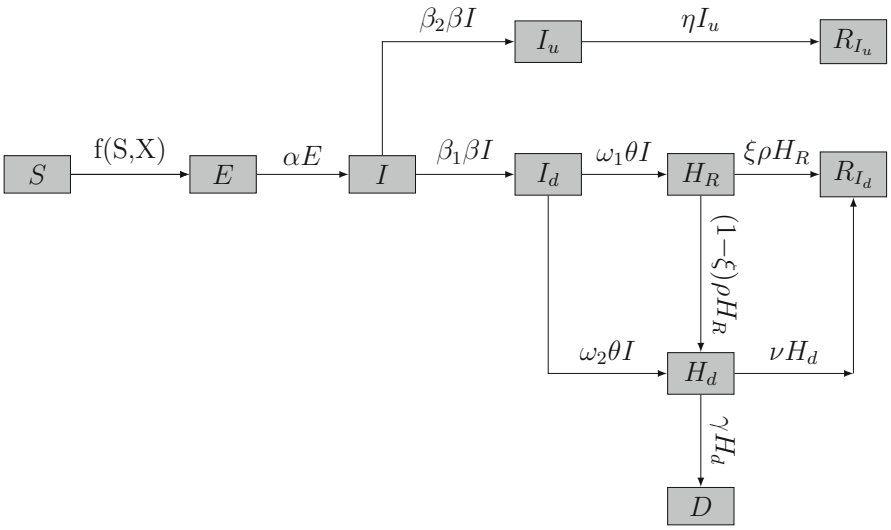
2 Mathematical Model

The model built for the coronavirus disease 2019 (COVID-19) in this paper is a $SEIRD$ model, taking into account Susceptibles cases(S), Exposed persons (E), Infected individual (I), Recovered patients (R) and the Dead patients (D). Based on the epidemiological characteristics of COVID-19, the $SEIRD$ model and its variants are more appropriate to study the dynamics of this disease, which is caused by the SARS-Cov-2. The output of the mathematical model gives the following transfer diagram (Fig. 1):

According to Fig. 1, we obtain the following system of ten differential equations describing the dynamics of the disease.

Here, we define the variables as follow:

- $S(t)$ (Susceptible) represents persons not infected by the disease pathogen.
- $E(t)$ (Exposed) refers to persons who are in the incubation period after being infected by the disease pathogen, and haves no visible clinical signs. These individuals could infect other people but with a lower probability than people



where $X = (E, I, I_u, I_d, H_R, H_D, R, D)$ and

$$f(S,X) = \frac{S}{N}(\beta_E E + \beta_I I + \beta_{I_u} I_u + \beta_{I_d} I_d + \beta_{H_R} H_R + \beta_{H_d} H_d + \beta_D D)$$

Fig. 1 Transfer diagram for the mathematical model of COVID-19

in the infectious compartments. After the incubation period, these persons move to one of the Infectious states. In this model, it can be seen has contact persons

- $I(t)$ (Infectious) refers the number of persons who are beginning to develop clinical signs, these persons are symptomatic infectious,
- $I_d(t)$ refers to Infectious; that is persons who can infect other people, are developing clinical signs and therefore will be detected and reported by authorities (when arriving in the H_R or the H_d compartments). After this period, the people in this compartment are taken in charge by sanitary authorities and we classify them as Hospitalized patients. It is the reported infectious cases.
- $I_u(t)$ is the number of unreported symptomatic infectious individuals (i.e., symptomatic infectious with mild symptoms) at time t.
- Persons hospitalized or quarantined at home (but detected and reported by the authorities) who will recover (denoted by H_R): These persons are in hospital (or quarantined at home) and can still infect other people. At the end of this state, these persons move to the Recovered state. It is the low-risk hospitalized cases.
- Hospitalized people who are going to die (denoted by H_d): These persons are hospitalized and can still infect other people. At the end of this state, these persons move to the death state. It is the high-risk hospitalized cases.

- R_{I_u} Recovered persons from unreported infectious cases (denoted by R_{I_u}): These persons have survived the disease, are no longer infectious and have developed a natural immunity to the disease pathogen.
- R_{I_d} Recovered persons from reported infectious cases (denoted by R_{I_d}): These persons have survived the disease, are no longer infectious and have developed a natural immunity to the disease pathogen.
- D Dead persons (denoted by D): These persons have not survived the disease.
- N is the number of people within the territory before the start of the pandemic.

The transcription of the transfer diagram gives the following mathematical system:

$$\begin{cases} \dot{S} = -\dfrac{S(t)}{N}\Big(m_E(t)\beta_E E(t) + m_I(t)\beta_I I(t) + \beta_{I_u} m_{I_u}(t) I_u(t) + \beta_{I_d} m_{I_d}(t) I_d(t)\Big) \\ \qquad -\dfrac{S(t)}{N}\Big(\beta_{H_R} m_{H_R}(t) H_R(t) + \beta_{H_d} m_{H_d}(t) H_d(t)\Big), \\ \dot{E} = \dfrac{S(t)}{N}\Big(m_E(t)\beta_E E(t) + m_I(t)\beta_I I(t) + \beta_{I_u} m_{I_u}(t) I_u(t) + \beta_{I_d} m_{I_d}(t) I_d(t)\Big) \\ \qquad +\dfrac{S(t)}{N}\Big(\beta_{H_R} m_{H_R}(t) H_R(t) + \beta_{H_d} m_{H_d}(t) H_d(t)\Big) - \alpha E(t), \\ \dot{I} = \alpha E(t) - \beta I(t), \\ \dot{I}_d = \beta_1 \beta I(t) - \theta I_d(t), \\ \dot{I}_u = \beta_2 \beta I(t) - \eta I_u(t), \\ \dot{H_R} = \omega_1 \theta I_d(t) - \rho H_R(t), \\ \dot{H_d} = \omega_2 \theta I_d(t) + (1-\xi)\rho H_R - \gamma H_d(t) - \nu H_d, \\ \dot{R_{I_d}} = \xi \rho H_R(t) + \nu H_d, \\ \dot{R_{I_u}} = \eta I_u(t), \\ \dot{D} = \gamma H_d(t). \end{cases} \tag{1}$$

The initial data of the system is supplemented by

$$S(t_0) = S_0 > 0, \quad E(t_0) = E_0 > 0, \quad I(t_0) = I_0 > 0, \quad I_d(t_0) = 0 \quad I_u(t_0) \geq 0, \tag{2}$$

$$H_R(0) = 0, \quad H_d(0) = 0, \quad R_{I_u}(0) = 0, \quad R_{I_d}(0) = 0 \quad \text{and} \quad D(0) = 0.$$

The time t is in day, the asymptomatic infectious individuals $I(t)$ are infectious for an average time period of $1/\alpha$ days. We suppose that the population is constant i.e $N = S(t) + E(t) + I(t) + R(t)$.

We notice that 8th, 9th and 10th equations of system (1) are not coupled with the other equations. So we can solve the first seven equations of the system and compute them after the other three equations. Thus, the system taken into account is the following with seven equations:

$$\begin{cases} \dot{S} = -\dfrac{S(t)}{N}\left(m_E(t)\beta_E E(t) + m_I(t)\beta_I I(t) + \beta_{I_u} m_{I_u}(t) I_u(t) + \beta_{I_d} m_{I_d}(t) I_d(t)\right) \\ \qquad -\dfrac{S(t)}{N}\left(\beta_{H_R} m_{H_R}(t) H_R(t) + \beta_{H_d} m_{H_d}(t) H_d(t)\right), \\ \dot{E} = \dfrac{S(t)}{N}\left(m_E(t)\beta_E E(t) + m_I(t)\beta_I I(t) + \beta_{I_u} m_{I_u}(t) I_u(t) + \beta_{I_d} m_{I_d}(t) I_d(t)\right) \\ \qquad +\dfrac{S(t)}{N}\left(\beta_{H_R} m_{H_R}(t) H_R(t) + \beta_{H_d} m_{H_d}(t) H_d(t)\right) - \alpha E(t), \\ \dot{I} = \alpha E(t) - \beta I(t), \\ \dot{I}_d = \beta_1 \beta I(t) - \theta I_d(t), \\ \dot{I}_u = \beta_2 \beta I(t) - \eta I_u(t), \\ \dot{H}_R = \omega_1 \theta I_d(t) - \rho H_R(t), \\ \dot{H}_d = \omega_2 \theta I_d(t) + (1-\xi)\rho H_R - \gamma H_d(t) - \nu H_d. \end{cases} \tag{3}$$

3 Basic Properties

In this section, we prove the positivity and the boundedness of the solution of the system (1) with initial conditions defined in (2).

For that, let us state the following lemma.

Lemma 3.1 ([9]) *Suppose* $\Omega \subset \mathbb{R} \times \mathbb{C}^n$ *is open and* $f_i \in C(\Omega, \mathbb{R})$, $i = 1, 2 \cdots, n$. *If* $f_i|_{x_i(t)=0,\, X_t \in C^n_{+0}} \geq 0$, $X_t = (x_{1t}, x_{2t}, \cdots, x_{nt})$, *then,* $\phi = (\phi_1, \phi_2, \cdots, \phi_n) \in C([-\tau, 0], \mathbb{R}^n_{+0})$ *is the invariant domain of the following equations.*

$$\frac{dx_i(t)}{dt} = f_i(t, X_t),\ t \geq \sigma,\ i = 1, 2 \cdots, n;$$

where $\mathbb{R}^n_{+0} = \{(x_1, x_2, \cdots, x_n) : x_i \geq,\ i = 1, 2 \cdots, n.$

Proposition 3.1 *The system (1) is invariant in* $\mathbb{R}^{10}_+$.

Proof Let us write system *(1)* as

$$\frac{dX}{dt} = f(X(t)),\ X(0) = X_0 \geq 0,$$

so, note that

$$\frac{dS}{dt}|_{S=0} = 0,$$
$$\frac{dE}{dt}|_{E=0} = \frac{S}{N}\left(m_I\beta_I I + \beta_{I_u}m_{I_u}I_u + \beta_{I_d}m_{I_d}I_d + \beta_{H_R}m_{H_R}H_R + \beta_{H_d}m_{H_d}H_d\right) \geq 0,$$
$$\frac{dI}{dt}|_{I=0} = \alpha E \geq 0, \quad \frac{dI_d}{dt}|_{I_d=0} = \beta_1\beta I \geq 0,$$
$$\frac{dI_u}{dt}|_{I_u=0} = \beta_2\beta I \geq 0, \quad \frac{dH_R}{dt}|_{H_R=0} = \omega_1\theta I_d \geq 0,$$
$$\frac{dH_d}{dt}|_{H_d=0} = \omega_2\theta I_d + (1-\xi)\rho H_R \geq 0, \text{ since } \zeta \in]0,1[.$$

From Lemma 3.1, we conclude that system (1) is invariant in $\mathbb{R}^{10}_+$.

Lemma 3.2 *The solution of system* (1) *is bounded in the region*

$$\Omega = \{(S, E, I, I_d, I_u, H_R, H_d, R_{I_d}, R_{I_u}, D) \in \mathbb{R}^{10}_+ | S+E+I+I_d+I_u+H_R+H_d+R_{I_d}+R_{I_u}+D = N\}.$$

Proof For system (1), we observe that
$\dot{N} = 0, \Rightarrow N(t) = Cst,\ \forall t \geq 0$, thus all the solution of system (1) are bounded.

The only equilibrium of the model is the disease free equilibrium (DFE) defined by $E_{DFE} = (N, 0, 0, 0, 0, 0, 0, 0, 0, 0) = (S_0, 0, 0, 0, 0, 0, 0, 0, 0, 0)$.

Proposition 3.2 *The basic reproduction number for the model system* (1) *is defined by*

$$\mathcal{R}_0 = \frac{1}{\alpha}m_E(t)\beta_E + \frac{1}{\beta}m_I(t)\beta_I + \frac{\beta_2}{\eta}m_{I_u}(t)\beta_{I_u} + \frac{\beta_1}{\theta}m_{I_d}(t)\beta_{I_d} + \frac{\omega_1\beta_1}{\rho}m_{H_R}(t)\beta_{H_R} \tag{4}$$

$$+\frac{\beta_1}{\gamma+\nu}(\omega_2 + \omega_1(1-\xi))m_{H_d}(t)\beta_{H_d}. \tag{5}$$

Proof We use the method of next generation matrix in [10] to compute the reproduction number $\mathcal{R}_0$. So we get:

$$\mathcal{F} = \begin{pmatrix} \mathcal{F}_1 \\ \mathcal{F}_2 \\ \mathcal{F}_3 \\ \mathcal{F}_4 \\ \mathcal{F}_5 \\ \mathcal{F}_6 \end{pmatrix} = \begin{pmatrix} \frac{S(t)}{N}\left(m_E(t)\beta_E E(t) + m_I(t)\beta_I I(t) + m_{I_u}(t)\beta_{I_u} I_u(t) + m_{I_d}(t)\beta_{I_d} I_d(t)\right) \\ + \frac{S(t)}{N}\left(m_{H_R}(t)\beta_{H_R} H_R(t) + m_{H_d}(t)\beta_{H_d} H_d(t)\right) \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

and

$$\mathcal{V} = \begin{pmatrix} \mathcal{V}_1 \\ \mathcal{V}_2 \\ \mathcal{V}_3 \\ \mathcal{V}_4 \\ \mathcal{V}_5 \\ \mathcal{V}_6 \end{pmatrix} = \begin{pmatrix} -\alpha E(t) \\ \alpha E(t) - \beta I(t) \\ \beta_1 \beta I(t) - \theta I_d(t) \\ \beta_2 \beta I(t) - \eta I_u(t) \\ \omega_1 \theta I_d(t) - \rho H_R(t) \\ \omega_2 \theta I_d(t) + (1-\xi)\rho H_R(t) - \gamma H_d(t) - \nu H_d(t) \end{pmatrix}.$$

We get the Jacobian matrixes of these two functions $\mathcal{F}$ and $\mathcal{V}$ as follow.

$$F = \mathcal{D}\mathcal{F}_{(E,I)} = \begin{pmatrix} \frac{\partial \mathcal{F}_1}{\partial E} & \frac{\partial \mathcal{F}_1}{\partial I} & \frac{\partial \mathcal{F}_1}{\partial I_u} & \frac{\partial \mathcal{F}_1}{\partial I_d} & \frac{\partial \mathcal{F}_1}{\partial H_R} & \frac{\partial \mathcal{F}_1}{\partial H_d} \\ \frac{\partial \mathcal{F}_2}{\partial E} & \frac{\partial \mathcal{F}_2}{\partial I} & \frac{\partial \mathcal{F}_2}{\partial I_u} & \frac{\partial \mathcal{F}_2}{\partial I_d} & \frac{\partial \mathcal{F}_2}{\partial H_R} & \frac{\partial \mathcal{F}_2}{\partial H_d} \\ \frac{\partial \mathcal{F}_3}{\partial E} & \frac{\partial \mathcal{F}_3}{\partial I} & \frac{\partial \mathcal{F}_3}{\partial I_u} & \frac{\partial \mathcal{F}_3}{\partial I_d} & \frac{\partial \mathcal{F}_3}{\partial H_R} & \frac{\partial \mathcal{F}_3}{\partial H_d} \\ \frac{\partial \mathcal{F}_4}{\partial E} & \frac{\partial \mathcal{F}_4}{\partial I} & \frac{\partial \mathcal{F}_4}{\partial I_u} & \frac{\partial \mathcal{F}_4}{\partial I_d} & \frac{\partial \mathcal{F}_4}{\partial H_R} & \frac{\partial \mathcal{F}_4}{\partial H_d} \\ \frac{\partial \mathcal{F}_5}{\partial E} & \frac{\partial \mathcal{F}_5}{\partial I} & \frac{\partial \mathcal{F}_5}{\partial I_u} & \frac{\partial \mathcal{F}_5}{\partial I_d} & \frac{\partial \mathcal{F}_5}{\partial H_R} & \frac{\partial \mathcal{F}_5}{\partial H_d} \\ \frac{\partial \mathcal{F}_6}{\partial E} & \frac{\partial \mathcal{F}_6}{\partial I} & \frac{\partial \mathcal{F}_6}{\partial I_u} & \frac{\partial \mathcal{F}_6}{\partial I_d} & \frac{\partial \mathcal{F}_6}{\partial H_R} & \frac{\partial \mathcal{F}_6}{\partial H_d} \end{pmatrix}$$

and at the point E_{DFE}, we have

$$F = \begin{pmatrix} m_E(t)\beta_E & m_I(t)\beta_I & m_{I_u}(t)\beta_{I_u} & m_{I_d}(t)\beta_{I_d} & m_{H_R}(t)\beta_{H_R} & m_{H_d}(t)\beta_{H_d} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

$$\mathcal{DV}_{(E,I)} = \begin{pmatrix} \frac{\partial \mathcal{V}_1}{\partial E} & \frac{\partial \mathcal{V}_1}{\partial I} & \frac{\partial \mathcal{V}_1}{\partial I_u} & \frac{\partial \mathcal{V}_1}{\partial I_d} & \frac{\partial \mathcal{V}_1}{\partial H_R} & \frac{\partial \mathcal{V}_1}{\partial H_d} \\ \frac{\partial \mathcal{V}_2}{\partial E} & \frac{\partial \mathcal{V}_2}{\partial I} & \frac{\partial \mathcal{V}_2}{\partial I_u} & \frac{\partial \mathcal{V}_2}{\partial I_d} & \frac{\partial \mathcal{V}_2}{\partial H_R} & \frac{\partial \mathcal{V}_2}{\partial H_d} \\ \frac{\partial \mathcal{V}_3}{\partial E} & \frac{\partial \mathcal{V}_3}{\partial I} & \frac{\partial \mathcal{V}_3}{\partial I_u} & \frac{\partial \mathcal{V}_3}{\partial I_d} & \frac{\partial \mathcal{V}_3}{\partial H_R} & \frac{\partial \mathcal{V}_3}{\partial H_d} \\ \frac{\partial \mathcal{V}_4}{\partial E} & \frac{\partial \mathcal{V}_4}{\partial I} & \frac{\partial \mathcal{V}_4}{\partial I_u} & \frac{\partial \mathcal{V}_4}{\partial I_d} & \frac{\partial \mathcal{V}_4}{\partial H_R} & \frac{\partial \mathcal{V}_4}{\partial H_d} \\ \frac{\partial \mathcal{V}_5}{\partial E} & \frac{\partial \mathcal{V}_5}{\partial I} & \frac{\partial \mathcal{V}_5}{\partial I_u} & \frac{\partial \mathcal{V}_5}{\partial I_d} & \frac{\partial \mathcal{V}_5}{\partial H_R} & \frac{\partial \mathcal{V}_5}{\partial H_d} \\ \frac{\partial \mathcal{V}_6}{\partial E} & \frac{\partial \mathcal{V}_6}{\partial I} & \frac{\partial \mathcal{V}_6}{\partial I_u} & \frac{\partial \mathcal{V}_6}{\partial I_d} & \frac{\partial \mathcal{V}_6}{\partial H_R} & \frac{\partial \mathcal{V}_6}{\partial H_d} \end{pmatrix}$$

and at the point E_{DFE}, we have

$$V = \mathcal{DV}_{(E_{DFE})} = \begin{pmatrix} -\alpha & 0 & 0 & 0 & 0 & 0 \\ \alpha & -\beta & 0 & 0 & 0 & 0 \\ 0 & \beta_1\beta & 0 & -\theta & 0 & 0 \\ 0 & \beta_2\beta & -\eta & 0 & 0 & 0 \\ 0 & 0 & 0 & -\omega_1\theta & -\rho & 0 \\ 0 & 0 & 0 & -\omega_2\theta & (1-\xi)\rho & -\gamma-\nu \end{pmatrix}.$$

Thus, we obtain

$$-FV^{-1} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

where

$$A = -\begin{pmatrix} \frac{-1}{\alpha} m_E(t)\beta_E - \frac{1}{\beta} m_I(t)\beta_I - \frac{\beta_2}{\eta} m_{I_u}(t)\beta_{I_u} - \frac{\beta_1}{\theta} m_{I_d}(t)\beta_{I_d} - \frac{\omega_1\beta_1}{\rho} m_{H_R}(t)\beta_{H_R} \\ -\frac{\beta_1}{\gamma+\nu}(\omega_2+\omega_1(1-\xi))m_{H_d}(t)\beta_{H_d} \end{pmatrix} \in \mathcal{M}_{1,1}(\mathbb{R}),$$

$$B = 0_{1,6} \in \mathcal{M}_{1,6}(\mathbb{R}), \quad C = 0_{5,1} \in \mathcal{M}_{5,1}(\mathbb{R}) \quad \text{and} \quad D = 0_{5,5} \in \mathcal{M}_{5,5}(\mathbb{R}).$$

The basic reproduction number is defined as the dominant eigeinvalue of the matrix $-FV^{-1}$.

Therefore,

$$\mathcal{R}_0 = \frac{1}{\alpha} m_E(t)\beta_E + \frac{1}{\beta} m_I(t)\beta_I + \frac{\beta_2}{\eta} m_{I_u}(t)\beta_{I_u} + \frac{\beta_1}{\theta} m_{I_d}(t)\beta_{I_d} + \frac{\omega_1\beta_1}{\rho} m_{H_R}(t)\beta_{H_R} + \frac{\beta_1}{\gamma+\nu}(\omega_2+\omega_1(1-\xi))m_{H_d}(t)\beta_{H_d}$$

□

The basic reproduction number $\mathcal{R}_0$ is defined as the number of cases that one infected person generates on average during his infectious period, within an uninfected population and without any special control measures. This number does not change during the spread of the disease.

The effective reproduction number $\mathcal{R}_e(t)$ is defined as the number of cases that one infected person generates during his infectious period. This effective reproduction number depends on time, therefore, on public policies (it changes during the spread of the disease). Furthermore, $\mathcal{R}_e(t) = \mathcal{R}_e - \mathcal{R}_0$ at $t = 0$ and the spread of the disease slows when $\mathcal{R}_e(t) < 1$.

Without any public policies,

$$m_E(t) = m_I(t) = m_{I_u}(t) = m_{I_d}(t) = m_{H_R}(t) = m_{H_d}(t) = 1,$$

so

$$\mathcal{R}_0 = \frac{1}{\alpha}\beta_E + \frac{1}{\beta}\beta_I + \frac{\beta_2}{\eta}(t)\beta_{I_u} + \frac{\beta_1}{\theta}\beta_{I_d} + \frac{\omega_1\beta_1}{\rho}\beta_{H_R} + \frac{\beta_1}{\gamma+\nu}(\omega_2+\omega_1(1-\xi))\beta_{H_d}$$

and, therefore,

$$\mathcal{R}_e(t) = \left(\frac{1}{\alpha} m_E(t)\beta_E + \frac{1}{\beta} m_I(t)\beta_I + \frac{\beta_2}{\eta} m_{I_u}(t)\beta_{I_u} + \frac{\beta_1}{\theta} m_{I_d}(t)\beta_{I_d} + \frac{\omega_1\beta_1}{\rho} m_{H_R}(t)\beta_{H_R} + \frac{\beta_1}{\gamma+\nu}(\omega_2+\omega_1(1-\xi))m_{H_d}(t)\beta_{H_d}\right)\frac{S(t)}{N}. \tag{6}$$

Using Theorem 2 in [10], the following result is established.

Proposition 3.3 *The disease free equilibrium E_{DFE} of he model (3) is locally asymptotically stable if $\mathcal{R}_0 < 1$ and unstable if $\mathcal{R}_0 > 1$.*

4 Data

In this part, we show real data generated by CORUS, the Burkina Faso National Health Commission against the Coronavirus Disease 2019 (COVID-19). Recall that the first cases were reported in Burkina Faso on March 9, 2020.

In **Appendix A**, Table 1 represents the daily cumulative number of reported infected cases and Table 2 the daily cumulative reported dead cases.

5 Numerical Simulations

For the numerical simulations, we focus on the public policies represented in the model (1) by functions m_E, m_I, m_{I_u}, m_{I_d}, m_{H_R} and m_{H_d}. From the beginning date of the public policies, we decrease the contact rate, as a result of the various types of measures taken by the public authorities. The effect of these measures can be seen on the spikes of the different types of infectious cases; the spikes decrease according to the degrees of the measures taken. The curve of the contact rate (contact function) is represented in the Fig. 4.

$$\gamma(t) = m_E(t) = m_I(t) = m_{I_u}(t) = m_{I_d}(t) = m_{H_R}(t) = m_{H_d}(t) \tag{7}$$

$$= \begin{cases} 1, \quad 0 \leq t < t_0, \\ \gamma_{min0} + (\gamma_0 - \gamma_{min0}) \exp(-\mu_1(t - t_1)), \quad t_0 \leq t < t_1, \\ \dots \\ \gamma_{minq} + (\gamma_q - \gamma_{minq}) \exp(-\mu_q(t - t_{q-1})), \quad t_{q-1} \leq t < \infty, \end{cases}$$

where γ_{minq} is the minimum contact rate and μ_q is the rate at which the contact decreases.

Depending on the public measures taken, μ_i increases, so the contact rate $\gamma(t)$ decreases and it is possible to limit the number of infected persons (see Fig. 4). μ_i is chosen in such a way that the simulation for the time interval aligns with the cumulative reported case data. Also, the effective reproduction number $\mathcal{R}_e(t)$ is strongly dependent on the contact function which decreases exponentially by interval as shown in Fig. 5. Thus we are able to predict the future values of the epidemic from the early cumulative reported data. The earlier public policy decisions are made, the better the management of the epidemic.

For Burkina Faso, we present the model, first, without any public policy intervention and, secondly with public policies taken by the local government. It is important to observe that the basic reproduction number with the parameters taken for this simulation: $\mathcal{R}_0 = \mathcal{R}_e(0) = 4.5$; this mean that 10 persons infect 45 persons, which is very high. The role of public policies is to reduce this $\mathcal{R}_e$ down to less than one, which will ensure the extinction of the epidemic.

5.1 Situation Without Public Policies

Without any public policies, the situation of the pandemic would be dramatic for the African countries and for Burkina Faso in particular. In fact, as shown in the simulation (see Figs. 2 and 3), 9 millions persons would be infected by the COVID-19 and around 850 thousand would die, that is 9% of infected persons.

5.2 Situation with Public Policies

Fortunately, in Burkina Faso, as in almost all countries in the world, from the first cases of the disease, a succession of barrier measures were taken, ranging from hand washing, to the closing of schools, and, places of worship and markets, and the establishment of curfews. Consequently, in the following simulations, we have chosen our contact function in such a way that the cumulative number of infected

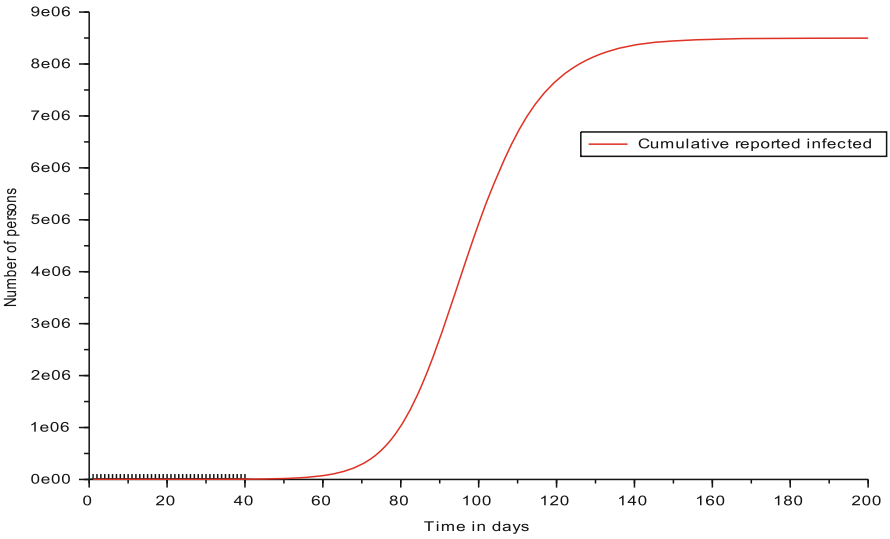


Fig. 2 Cumulative Infected persons generated by the model

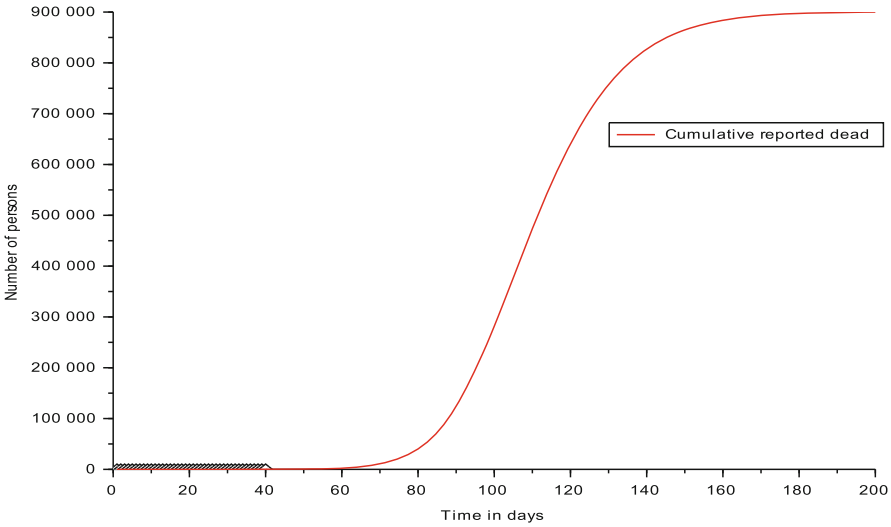


Fig. 3 Cumulative dead persons generated by the model

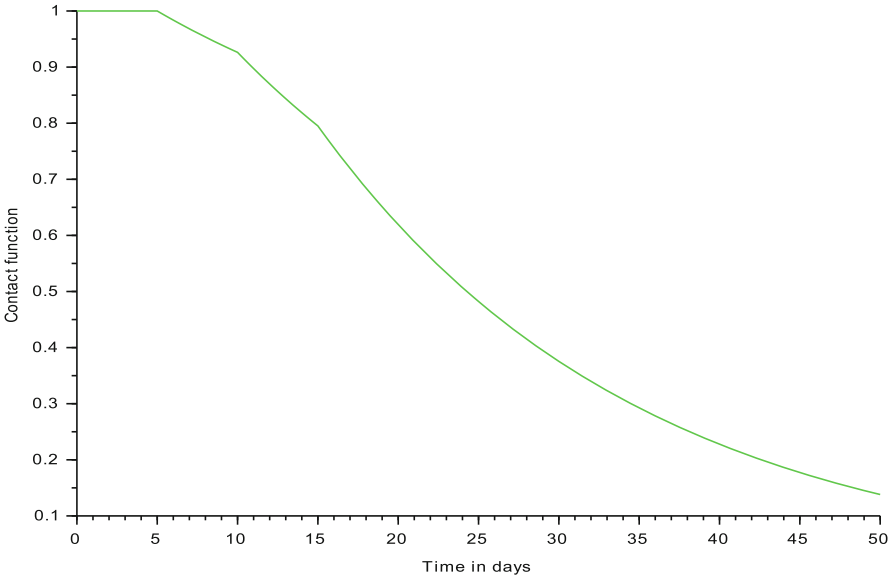


Fig. 4 Contact function

persons (model (1)) fits the data produced by the Burkina Faso National Health Commission against COVID-19.

Figure 5 represents the dynamic of the effective reproduction number and Fig. 4 represents the dynamic of the contact function. These two curves clearly show the effects of three measurements according to their implementation time.

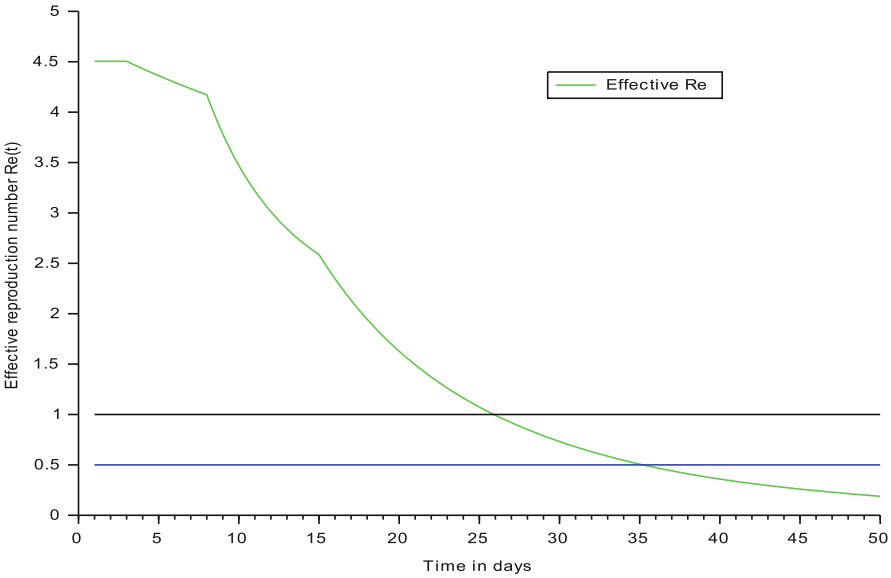


Fig. 5 Effective reproduction number $R_e(t)$ through the time

Figure 6 presents the histogram of the daily reported infected cases and the infected and the dead given by the model 1.

6 Conclusion

In this paper, we are proposing a model for the transmission of the coronavirus disease 2019. We calculated the $\mathcal{R}_0$ which is very essential in understanding the disease and we have showed the local stability of the disease free equilibrium DFE. We subsequently adjusted the model to the actual data of the National Health Commission against the coronavirus disease 2019. From this adjustment, we have been able to draw a number of consequences for the further management of the pandemic.

Firstly, from Fig. 5, we can say that the peak of the epidemic was reached in Burkina Faso around April 5, 2020 (Fig. 6).

Secondly, the data collected are not sufficiently homogenous, which allow for some reservation on the reliability of the data (Fig. 7). Nevertheless these data constitute a basis to make prescriptions for the dynamics of the disease and especially for removing the barrier measures. We achieved this by fitting the model with the data in Figs. 8, 9 and 10.

Finally, we can notice that the effective reproduction number $\mathcal{R}_e$ would be less than 0.5 at April 15, 2020, naturally according to the data fit with the model (1).

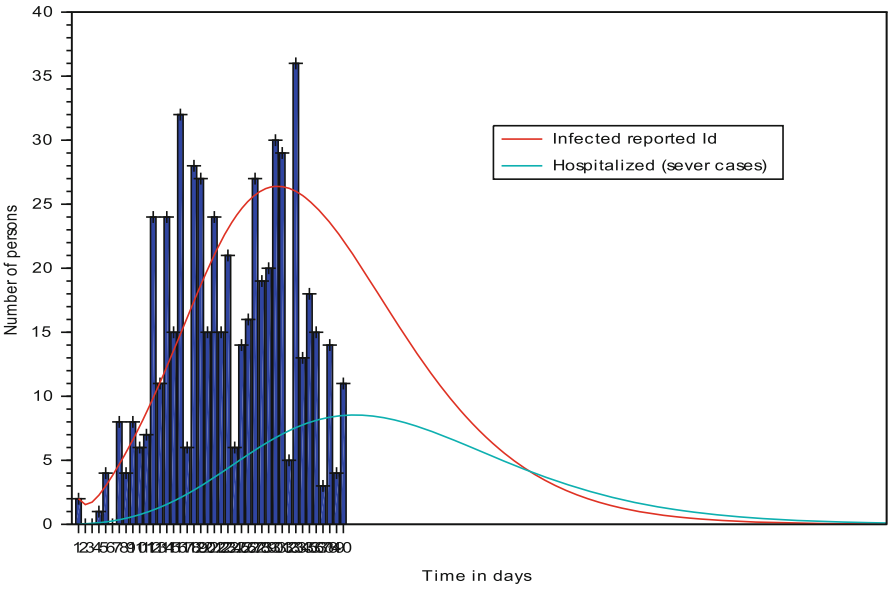


Fig. 6 Daily evolution of reported cases. The histogram is from real data and the curves from the model

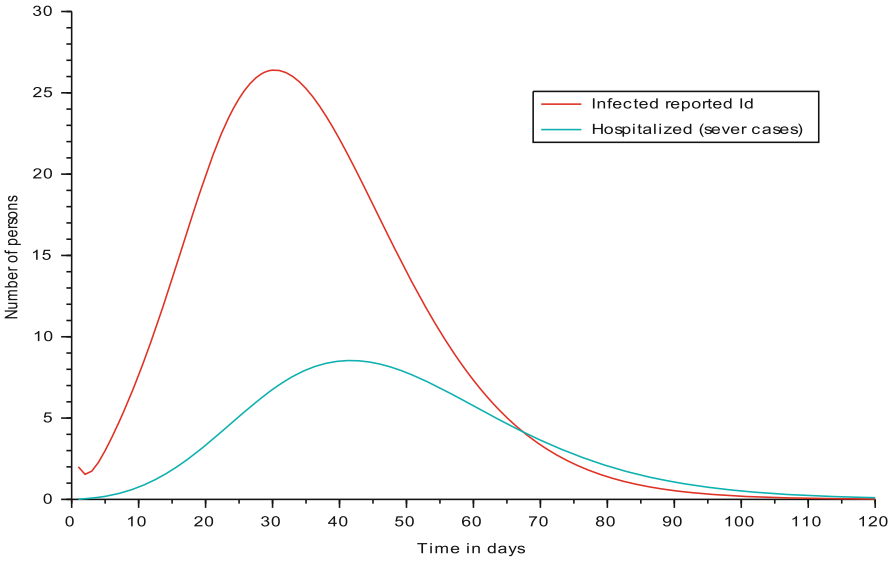


Fig. 7 Daily evolution of reported cases generated by the model

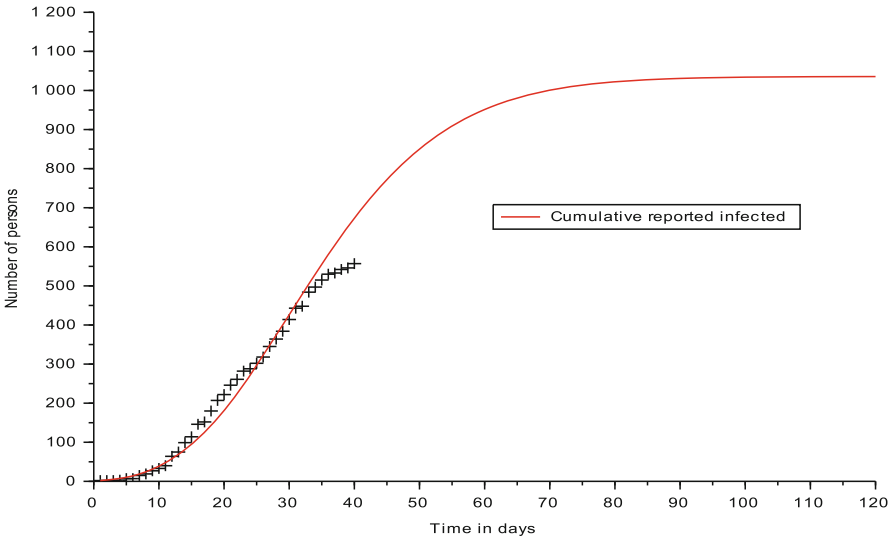


Fig. 8 Cumulative reported infected cases, real data (black point) and red curve (from the model)

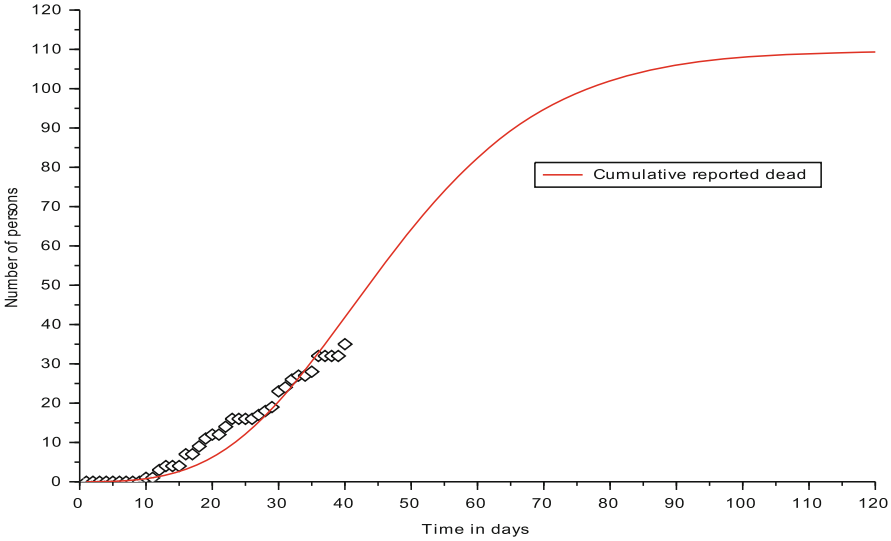


Fig. 9 Cumulative reported dead cases. real data (black point) and red curve (from the model)

We therefore think that one week after this date, certain barrier measures could be reviewed, for instance to allow for the opening of markets but maintaining the rule of distance, the opening of schools under conditions which do not allow the spread of the disease, and authorization for many other sectors of active life to be reopen.

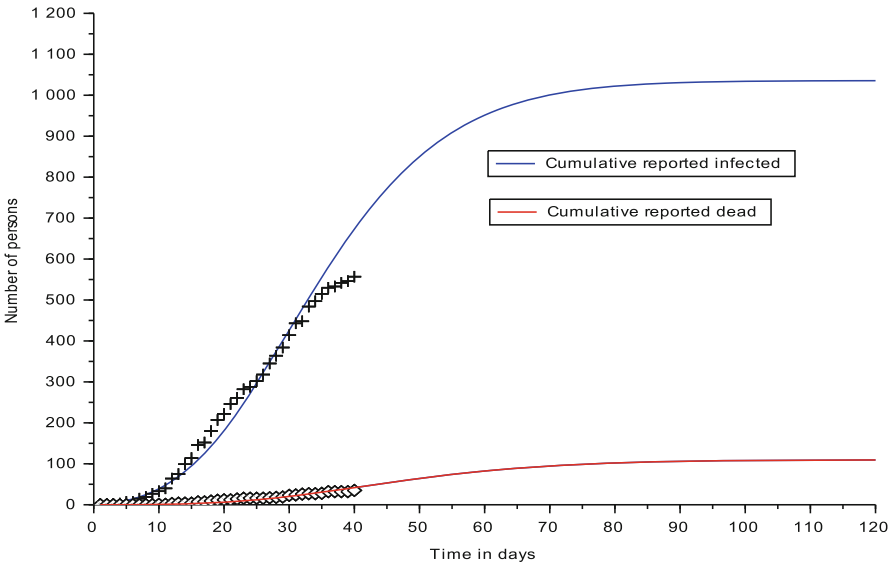


Fig. 10 Cumulative reported infected and dead cases. The blacks point are generated by the real data and the curves from the model

Appendix A. Tables of Data

Table 1 Daily cumulative reported cases data from March 9, 2020 to April 20, 2020 for Burkina Faso by Burkina Faso National Health Commission against the Coronavirus Disease 2019 (COVID-19)

March	9	12	13	15	16	17	18	19	20	21	22	23
	2	3	7	15	19	27	33	40	64	75	99	114

March	24	25	26	27	28	29	30	31
	146	152	180	207	222	246	261	282

April	1	2	3	4	5	6	7	8	9	10
	288	302	318	345	364	384	414	443	448	484

April	11	12	13	14	15	16	17	18	19	20
	497	515	530	533	542	546	557	568	573	592

Table 2 Daily cumulative reported dead cases data from March 9, 2020 to April 17, 2020 for Burkina Faso by Burkina Faso National Health Commission against the Coronavirus Disease 2019 (COVID-19)

March	9	18	20	21	24	26	27	28	30	31
	0	1	3	4	7	9	11	12	14	16
April	4	5	6	7	8	9	10	12	13	17
	17	18	19	23	24	26	27	28	32	35

References

1. B. Ivorra, M. R. Ferrández, M. Vela-Pérez, A. M. Ramos; *Mathematical modeling of the spread of the coronavirus disease 2019 (COVID-19) considering its particular characteristics. The case of China*; Preprint, March 2020.
2. A. Guiro, B. Koné and S. Ouaro; *Mathematical model of the spread of the coronavirus disease 2019 (COVID-19) in Burkina Faso*. Applied Mathematics, 2020, 11, 1204-1218. https://www.scirp.org/journal/am.
3. Q.Y. Lin, S. Zhao, D. Z. Gao, et coll; (2020) *A conceptual model for the coronavirus disease 2019 (COVID-19) outbreak in Wuhan, China with individual reaction and governmental action.*Int. J. Infect. Dis. 93 :211–216.
4. Z. Liu, P. Magal, O. Seydi and G. Webb; *Predicting the cumulative number of cases for the COVID-19 epidemic in China from early data*, medRxiv preprint, available under a CC-BY-NC-ND 4.0 International license.
5. Z. Liu, P. Magal, O. Seydi and G. Webb; *Understanding Unreported Cases in the COVID-19 Epidemic Outbreak in Wuhan, China, and the Importance of Major Public Health Interventions*, Biology 2020, 9, 50; https://doi.org/10.3390/biology9030050.
6. European Centre for Disease Prevention and Control. Discharge criteria for confirmed COVID-19 cases When is it safe to discharge COVID-19 cases from the hospital or end home isolation, https://www.ecdc.europa.eu/sites/default/files/documents/COVID-19-Discharge-criteria.pdf, March 2020.
7. Johns Hopkins University (JHU). *Coronavirus COVID-19 Global Cases by the Center for Systems Scienceand Engineering (CSSE)*. https://gisanddata.maps.arcgis.com/apps/opsdashboard/index.html/bda7594740fd40299423467b48e9ecf6, March 2020.
8. R. Li, S. Pei, B. Chen, Y. Song, T. Zhang, W. Yang, and J. Shaman; *Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2)*. Science, 2020.
9. X. Yang, L. Chen, and J. Chen; *Permanence and positive periodic solution for the single-species nonautonomous delay diffusive models*. Computers & Mathematics with Applications, 32(4):109116, 1996.
10. P. Van den Driesche and J. Watmough; *"Reproduction Numbers and Substhreshold Endemic Equilibria for the Compartmental Models of Disease Transmission"*, Mathematical Biosciences, Vol. 180, No. 1–2, 2002, pp. 29–48.
11. B. Ivorra, A. M. Ramos, and D. Ngom; *Be-CoDiS: A mathematical model to predict the risk of human diseases spread between countries. Validation and application to the 2014 Ebola Virus Disease epidemic*. Bulletin of Mathematical Biology, 77(9): 1668–1704, 2015.
12. T. Liu, J. Hu, M. Kang, L. Lin, H. Zhong, J. Xiao, and et al.; *Transmission dynamics of 2019 novel coronavirus* (2019-ncov). bioRxiv, 2020.
13. World Health Organization. *Report of the WHO-China Joint Mission on Coronavirus Disease 2019*. https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf, February 2020.

Spatio-Temporal Modelling of Progression of the COVID−19 Pandemic

Dilini Gamage, Jennifer L. Matthews, Norou Diawara, and Hueiwang Anna Jeng

1 Introduction

Coronavirus (COVID-19) is very transmissible disease and can cause estimated fatality rates of up to 2–3% [1]. The COVID-19 has created the largest pandemic in recent history and has affected almost all countries in the world. Although all countries were affected, the different rate differs and some countries experienced the higher burden of COVID-19 new cases of infection and death. Thus, given the severity of the situation, our goal and attention has been focused on utilizing statistical models to ascertain the spread of COVID-19. In doing so, we bring consideration on important risk factors and surveillance efforts to control progression and transmission of the disease.

The inherent nature of the COVID-19 pandemic calls for the need to use features that are specific to locations and time. The features also account for the spatial dependency between neighbouring or communicating countries along with the temporal dimension of its transmission. The communicable nature of the data and the statistical procedure used to analyze COVID-19 must account for spatial and temporal characteristics in the process.

D. Gamage · N. Diawara (✉)
Mathematics and Statistics Department, Old Dominion University, Norfolk, VA, USA
e-mail: dkatu001@odu.edu; ndiawara@odu.edu

J. L. Matthews
United States Naval Academy, Annapolis, MD, USA
e-mail: jmatthew@usna.edu

H. A. Jeng
Community and Environmental Health, Old Dominion University, Norfolk, VA, USA
e-mail: hjeng@odu.edu

B. Toni (ed.), *The Mathematics of Patterns, Symmetries, and Beauties in Nature*,
STEAM-H: Science, Technology, Engineering, Agriculture, Mathematics & Health,
https://doi.org/10.1007/978-3-030-84596-4_7

As mentioned in Dallatomasina et al. [1], epidemiologic characteristics of the outbreak using different countries are vital. McKinley et al. [2] used approximate likelihood-based inference for epidemic models and expressed the challenges due to the evaluation of the likelihood. Those authors proposed Markov Chain Monte Carlo and Sequential Monte Carlo algorithms for parameter estimates. Statistical inferences under a stochastic discrete-time based susceptible-exposed-infectious-recovered (SEIR) epidemic and under Bayesian change point models have been proposed by Lekone and Finkenstadt [3] and in Milletich et al. [4], respectively. Extension of that model to the continuous time type has been developed by McKinley et al. [2]. Those authors suggested a time point at which the transmission parameter changes. We proposed a Bayesian modelling of the weekly count of new COVID-19 cases and built models at two levels: in the spatial framework and in the case where the counts are modelled over time.

The dynamic behavior of the COVID-19 pandemic has been proposed by Dewhurst et al. [5] and Gencoglu and Gruber [6]. These authors looked at the time series of the word used from September 2019 to April 2020 from different languages in the Twitter 10% random sample of messages. They aggregated the data by considering the dominant languages (24 of them) in several countries. While such research is needed, it does not uncover or model causal relationships, and the ability to describe cluster for public health understanding remains a quest. Our goal of the current study is to provide a bridge by presenting the Bayesian probabilistic models from the statistical tools that add to the understanding of COVID-19 control and prevention.

Dealing with heterogeneous and spatial populations, the hierarchical Bayesian model or the conditional autoregressive model (CAR), as proposed by Millar et al. [7] or Hossain et al. [8], could serve as an attractive methodology. Millar et al. [7] stated that the use of hierarchical Markov Chain Monte Carlo (MCMC) methods are computationally viable methods. We propose a spatio-temporal CAR models using the full data and then used data based on disjoint time intervals for the COVID-19 pandemic. That pandemic can be tracked by looking at the number of new cases. As, numbers, such count data are often modelled using the Poisson distribution, which assumes that the mean and variance are equal. However, the existence of overdispersion is frequent in modelling count data, especially data measured over time, and that is the case for the COVID-19 data. Ignoring overdispersion can lead to underestimating standard errors in regression coefficients and result in biased statistical inferences (e.g., Type I Error; Gardner et al. [9]). As recommended by Navarro et al. [10], the negative binomial distribution can be used to model the part of the variance, which the Poisson distribution is unable to identify. We constructed generalized linear mixed models to model the new COVID-19 cases using Gaussian distribution.

The goals of this paper are to use data on the recent COVID-19 pandemic to build statistical models to understand the development of its transmission, identify time points spread of COVID-19 as well as growth periods, and identify/detect statistical covariates associated to the pandemic. Specifically, data on occurrences and new COVID-19 cases were collected by the World Health Organization (WHO) [28]

from January 27, 2020 to August 10, 2020. In this study, the models presented included the measure of spatial correlation, Bayesian disease mapping models, Moran's measure of spatial correlations, and generalized linear mixed models. One major role of the paper is to build a stochastic model in order to capture some of the disease characteristics. Doing so, we can pointers as to how to assist public health leaders in their efforts to manage the pandemic, as suggested in Dallatomasina et al. [1].

The paper is organized as follows. Basic data description and preliminary analyses as presented in Sect. 2. The space time-mixture of the Poisson regression with the use of the Moran and CAR model along with the Bayesian model are described in Sect. 3. The model formulations, estimation processes and prior distributions are presented in that section. Next, we introduce the generalized linear mixed model with both a Moran's spatial correlations, under spatio-temporal framework. Section 3 also discusses the results and comparisons from all the countries and block areas. Lastly, we end with a discussion and conclusion in Sect. 4.

2 The Dataset

Data on counts of new COVID-19 cases from January 27, 2020 to August 10, 2020 were collected via Situation Reports from the WHO's website (http://apps.who.int/situation-report). New cases were aggregated into weekly counts, for a total of 29 weeks per country. The aggregate data are also available under the US Center of Disease Controls and Prevention (US CDC). Given the association between population size and COVID-19, data on time in weeks and area sizes were collected as covariates. In particular, weekly counts based on a particular day (starting Monday January 27, and every week from there) over the time period were collected, aggregated into weeks, for a total of 29 weekly counts for each country.

2.1 Preliminary Analyses

A preliminary study shows basic information about the disease. Our goal is to fit model of the COVID-19 evolution, with the spatial-temporal association of locations and neighbors to obtain a physical description based on the scale of counts that may misalign the spatial estimates and interpretations. The cluster/block neighbors are selected targeted countries (US, Senegal, Spain, Malaysia) and the countries surrounding each of them. The blocks are referred to as North America Block, Africa Block, Europe Block and Asia Block. When testing for blocks, the most collections of cases defined within spatial proximity is key. However, the computational complexities cannot be avoided. Table 1 presents a summary of these statistics by country and aggregated across blocks. Figure 1 presents the distribution of number of new cases of COVID-19 aggregated across countries. The purpose of

Table 1 Descriptive Statistics of New COVID-19 Cases by Country within Block

	New COVID-19 cases by countries over 29 weeks' time period					
Countries	Mean	*SD*	Min	Median	(25%, 75%)	Max
USA	26,830	21883.0	0	25,870	(3, 33,510)	66,963
Canada	621	781.1	0	336	(60, 966)	3793
Mexico	2540	2883.0	0	1349	(12, 4410)	9556
Spain	1634	2046.0	0	647	(159, 2913)	6549
France	481	635.0	0	288	(28, 556)	2497
Portugal	242	221.5	0	209	(92, 320)	792
Senegal	59.97	57.7	0	57	(2, 110)	172
Guinea	37.1	37.2	0	40	(0, 64)	104
Mali	10.86	17.9	0	3	(0, 15)	87
Mauritania	36.1	67.7	0	0	(0, 22)	210
Guinea-Bissau	8.138	21.4	0	0	(0, 0)	77
Indonesia	581.1	635.2	0	349	(2, 862)	1893
Malaysia	51.93	73.5	0	16	(5, 67)	315
Thailand	24.45	63.0	0	3	(0, 9)	310
Combined	2368.4	8984.5	0	39.5	(0, 345.8)	66,963

Note. *SD* standard deviation, *25%* 25th percentile, *75%* 75th percentile. Estimates were obtained from weekly counts based on 29 weeks of observed data

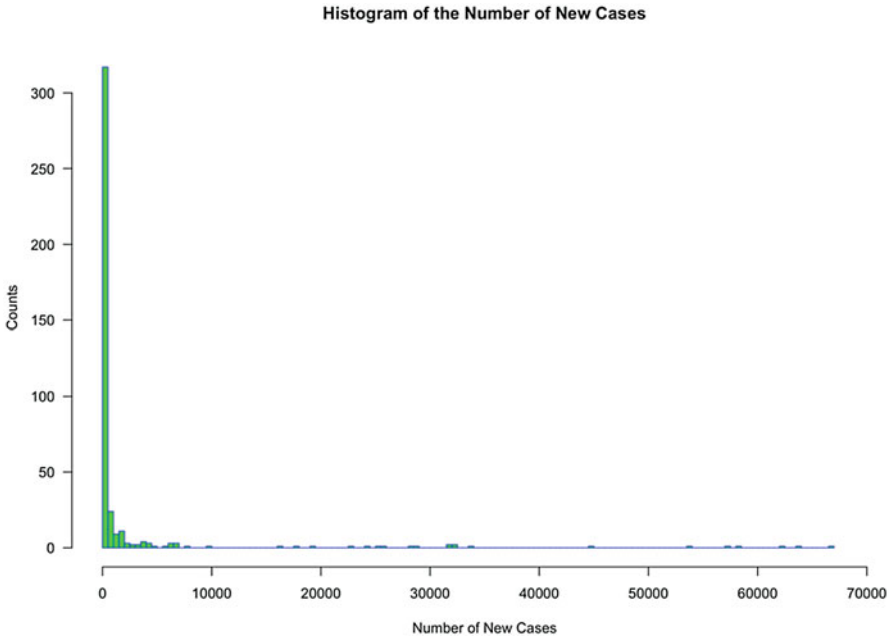


Fig. 1 Histogram of the Number of New Cases Due to COVID-19 Aggregated Across 14 Countries

this study is to investigate the new cases using inference under the Bayesian model in both space and time effects.

It is increasingly observed that the data collected in the locations and time differ. They have different patterns and behaviors as dependence is not just on the predictors and responses, but also in the space and time. As for the COVID-19, the interpolation across scales must be linked to the location and time. For that reason, connecting the progression of the pandemic over space and time blends the usual aggregation/generalization of its impacts, hence minimizing biases. The changes that are reflected under spatial varying area problem differences are compared at the cluster levels of contiguous areal units as in Anderson et al. [11, 12], and Lee et al. [13]. Other variables (such as population of the countries and the population density) were also considered, but as we analyzed them, they were removed if they were not found significant in the models. In fact, the population density can also be found in the country specific area size, and the time variable could underline other variables as well. So, other covariates may be added when available.

To illustrate the approach, we apply the spatial collections with neighbourhood countries. The time pattern of infections defines the response under the first order conditional auto-regressive (CAR) model. The geographical precisions suggest separate focus to build change among the set of locations and time statistical tools in epidemiology for the spatial analysis.

Figure 1 displays the histogram of the number of cases for the 14 countries. In Fig. 2, the data was further divided into the four area blocks. The data was further divided into country and block. Figure 3 shows the number of new COVID-19 cases for each country that appears to be generally decreasing over time. The pattern shown in Fig. 1 can be also seen in Fig. 3. However, there are fluctuations in the disease new case counts, indicative of a change in the behavior of the disease growth. As reported by the CDC and its various offices, China was the first place to report cases. Due to following travel restrictions on China, the number of cased was highly related to neighbouring countries and most cases of suspected, probable and confirmed COVID-19 cases were reported in CDC https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/cases-in-us.html or even from https://www.usaid.gov/coronavirus. Disease spread was not contained, however, and the basic characteristics of spread associated with space and time cannot be ignored in modelling.

The segregated analysis per block is displayed on Fig. 4. The trend cannot be used to make any general recommendation because there is no common trend evident among the countries and the blocks. Moreover, such descriptions do not take into account the variability associated with counts. As we will propose later, the Bayesian alternative offers a better understanding of the data.

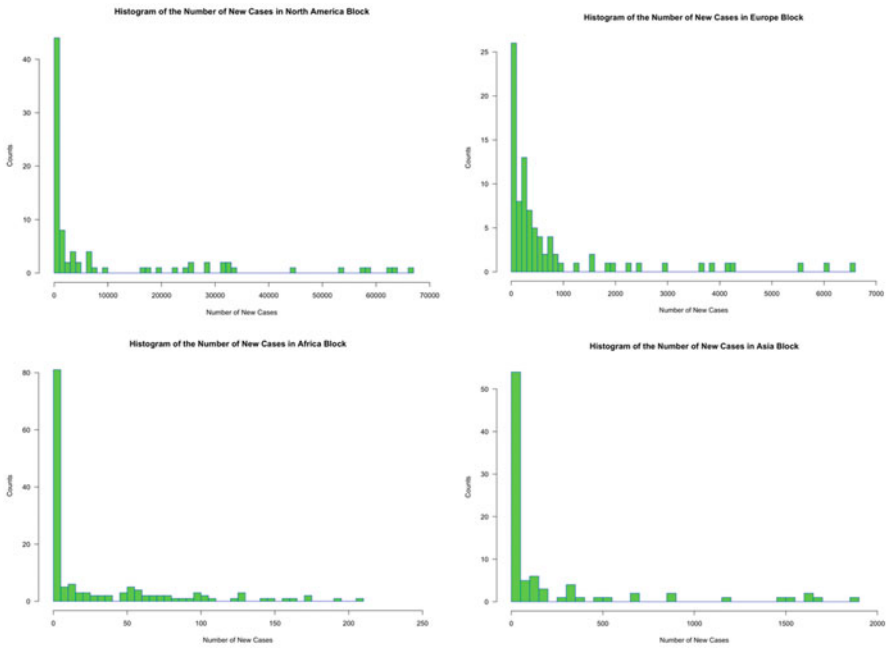


Fig. 2 Histogram of New cases by blocks

3 Statistical Models and Results

Analyzing the data under Poisson model did not lead to convergence. So, the statistical methods used is the Bayesian under normal approximation. The study area includes a set of $i = 1, \ldots, I$ non-overlapping areal units $S = \{S_1, \ldots S_I\}$, and data are recorded for each unit for $t = 1, \ldots, N$ consecutive time periods. Thus, data are available for a $I \times N$ rectangular array with I rows (spatial units) and N columns (time periods). The response data denoted by $\mathbf{Y} = (\mathbf{Y}_1, \ldots, \mathbf{Y}_N)_{I \times N}$, where $\mathbf{Y}_t = (Y_{1t}, \ldots, Y_{It})$ denotes the $I \times 1$ column vector of observations for all I spatial units for time period t. Finally, $X_{it} = (x_{it1}, \ldots, x_{itp})$ is a vector of p known covariates for areal unit i and time period t, and can include factors or continuous variables and a column of ones for the intercept term. CARBayesST can fit the following generalized linear mixed model to these data. The model is written as:

$$Y_{it} \mid \mu_{it} \sim f\left(y_{it} | \mu_{it}, \nu^2\right) \quad \text{for} \quad i = 1, \ldots, I, \quad t = 1, \ldots, N, \tag{1}$$

$$g\left(\mu_{it}\right) = \mathrm{X}_{it}^{\mathrm{T}} \beta + \psi_{it},$$

$$\boldsymbol{\beta} \sim N\left(\boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta\right).$$

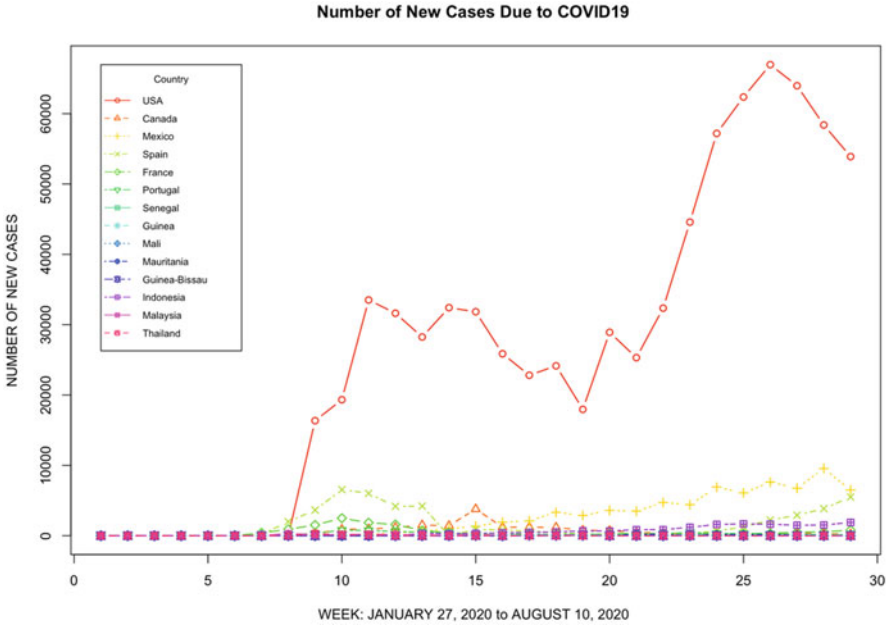


Fig. 3 Line plots of new cases over time

The vector of covariate regression parameters are denoted by $\beta = (\beta_1, \ldots, \beta_p)$, and a multivariate Gaussian or t-distributed prior is assumed with mean $\boldsymbol{\mu}_\beta$ and diagonal variance matrix $\boldsymbol{\Sigma}_\beta$ that can be chosen by the user. The ψ_{it} term is a latent component for areal unit i and time period t about one or more sets of spatio-temporally autocorrelated random effects, and the complete set are denoted by $\psi = (\psi_1, \ldots, \psi_N)$, where $\psi_t = (\psi_{1t}, \ldots, \psi_{It})$ [13]. Here, ν^2 represents the observed variance of the data.

Given that the distribution of new COVID-19 case counts is skewed (see Fig. 1), the CAR Bayes approach allows for local smoothing of estimates over neighboring areas, with approximation with Gaussian parameters. We use the approximation of the Poisson data with normal as in Peccati et al. [14]. The model convergence of the chosen distribution will be checked as it is an important part of the structure.

The software can implement Eq. (1) and the exact specification of Gaussian data model is:

$$Y_{it} \sim N\left(\mu_{it}, \nu^2\right) \quad \text{and} \quad \mu_{it} = X_{it}^T \beta + \psi_{it}.$$

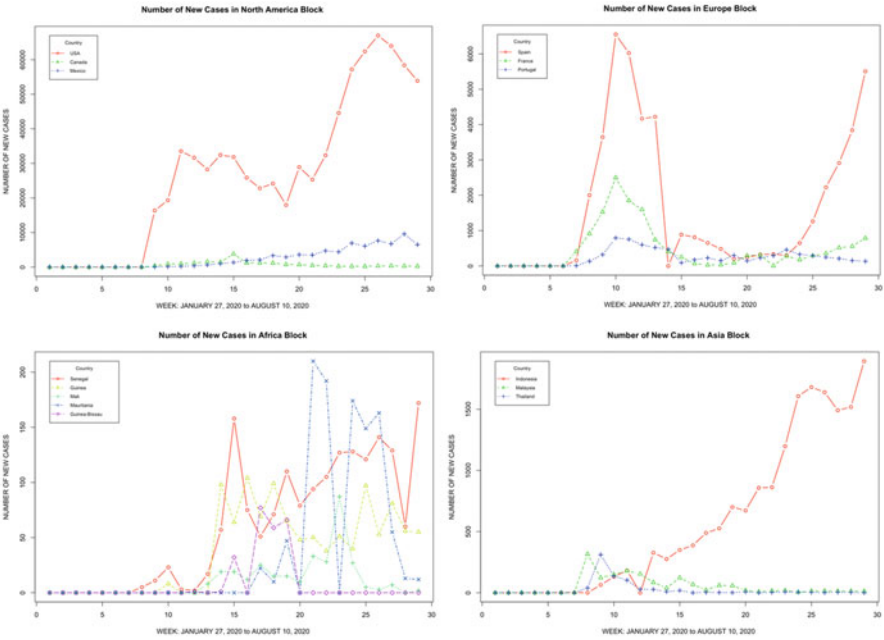


Fig. 4 Line plots of new cases per block

3.1 Conditional Autoregressive Bayesian Disease Mapping Models for Full Data

Based on the block locations, hypotheses about the progression of the COVID-19 pandemic may be misaligned based one simple interpolation. A CAR model is considered, and we assume that the adjacent countries that share a border linked. This model will be appropriate to estimate the growth of the spatial fitted value surface over time. The inclusion of the latent spatial effect follows then the Leroux prior (Cramb et al. [29]). The model specification is given below [13] as:

$$\psi_{it} = \phi_{it}, \quad \text{random effect in spatial location } i \text{ at time } t, \tag{2}$$

$$\phi_t \mid \phi_{t-1} \sim N\left(\rho_T \phi_{t-1}, \tau^2 \mathbf{Q}(\mathbf{W}, \rho_s)^{-1}\right), \quad t = 2, \ldots, N,$$

$$\phi_1 \sim N\left(0, \tau^2 \mathbf{Q}(\mathbf{W}, \rho_s)^{-1}\right),$$

$$\tau^2 \sim \text{Inverse} - \text{Gamma}\,(a, b)\,,$$

$$\rho_s, \rho_T \sim \text{Uniform}\,(0, 1)\,.$$

In this model, $\phi_t = (\phi_{1t}, \ldots, \phi_{It})$ is the vector of random effects for time period t, which change over time via a multivariate first order autoregressive process with temporal autoregressive parameter ρ_T, that is indicative of the temporal autocorrelation. Thus, temporal autocorrelation is induced via the mean $\rho_T \phi_{t-1}$, while spatial autocorrelation is induced by the variance $\tau^2 \boldsymbol{Q}(\boldsymbol{W}, \rho_s)^{-1}$. The corresponding precision matrix $\boldsymbol{Q}(\boldsymbol{W}, \rho_s)$ was proposed by Leroux et al. [30] and corresponds to the CAR models used in the other models above. The algebraic form of this matrix is given by:

$$\mathbf{Q}(\mathbf{W}, \rho_s) = \rho_s \left[\text{diag}(\mathbf{W}1) - \mathbf{W}\right] + (1 - \rho_s)\mathbf{K},$$

where $\mathbf{1}$ is the $I \times 1$ vector of ones while $\mathbf{K}$ is the $I \times I$ identity matrix. In common, with all other models the random effects are zero-mean centered, while flat and conjugate priors are specified for (ρ_s, ρ_T) and τ^2 respectively, with $(a = 1, b = 0.01)$ being the default values for the latter. The dependence parameters (ρ_s, ρ_T) can be fixed at values in the unit interval [0,1] rather than being estimated in the model.

Spatial autocorrelation is controlled by a symmetric non-negative $I \times I$ neighbourhood or adjacency matrix $\mathbf{W} = (w_{ij})$, where w_{ij} represents the spatial closeness between areal units (S_i, S_j). Larger values represent spatial closeness between the two areas in question, where as smaller or zero values correspond to areas that are not spatially close. Most often $\mathbf{W}$ is assumed to be binary, where $w_{ij} = 1$ if areal units (S_i, S_j) share a common border (i.e., are spatially close) and is zero otherwise. Additionally, $w_{ii} = 0$. Under this binary specification the values of (ψ_{it}, ψ_{jt}) for spatially adjacent areal units (where $w_{ij} = 1$) are spatially autocorrelated, where as values for non-neighbouring areal units (where $w_{ij} = 0$) are conditionally independent given the remaining $\{\psi_{kt}\}$ values. This binary specification of $\mathbf{W}$ based on sharing a common border is the most commonly used for areal data, but the only requirement by CARBayesST is for $\mathbf{W}$ to be symmetric, non-negative, and for each row sum to be greater than zero [13].

The Bayesian estimation model is implemented in the four spatial blocks as described in Eq. (2). The ideas are also mentioned in Finley et al. [15], Banerjee et al. [16], and Finley et al. [17].

3.2 Output and Results

Disease mapping models are often applied in epidemiological settings to understand the incidence or prevalence of a specific disease. The data are counts of observed cases within multiple regions coupled with potentially relevant background information [18]. One popular disease mapping model is the conditional autoregressive (CAR) model. This model gets its name from the type of random effect specified (i.e., conditional autoregressive) used to account for the spatial dependency among the observed data. Fitting hierarchical model using Markov Chain Monte Carlo (MCMC) methods is possible even for overdispersion of count data.

Table 2 North America Block

	Estimate	Std. error	T value	Pr(>|t|)
(Intercept)	−1.318e+04	4.556e+03	−2.893	0.00486
Week	8.856e+02	1.927e+02	4.595	1.51e−05
Area	3.527e−03	1.116e−03	3.160	0.00219

AIC: 1925.5

Table 3 Europe Block

	Estimate	Std. Error	T value	Pr(>|t|)
(Intercept)	6.151e+01	3.934e+02	0.156	0.8761
Week	1.763e+01	1.736e+01	1.016	0.3128
Area	2.875e−03	1.605e−03	1.791	0.0768

AIC: 1506.7

Bayesian statistics have gained great attention due to the computing power available and the flexibility that the models offer. Applications can be found in many fields and in medical disease progression [19]. Computations are commonly carried out using R. In the present study, for the COVID-19 new cases count data, with the characteristics as time expressed in weeks and the area, the following CAR model is specified as,

$$y_{it} = \beta_{0t} + \beta_{1t}\text{week}_{it} + \beta_{2t}\text{area}_{it} + \psi_{it},$$

for $i = 1, \ldots, I$ countries, $t = 1, 2, \ldots, 18$ weeks, and $\psi_{it} \sim \text{CAR}\left(\sigma_u^2\right)$. Here I is 3, 3, 3, and 5, representing the number of countries in the North America, Europe, Africa, and Asia blocks, respectively.

The CAR models were estimated in WinBUGS [20] using a Metropolis-Hastings algorithm. Now, with R version obtained from Lee et al. [13], we will use large iterations and a burn-in period, and thinning set to 100 to reduce the autocorrelation in parameter draws from the posterior distribution.

Results of the CAR model show that both covariates, week (time) and area, are significantly associated with relative new cases due to COVID-19. Tables 2, 3, 4, and 5 show the outputs of the results for the four blocks. Specifically, the effects are all positive indicating that an increase in time (week) and area, except for Africa block where only time is significant. These effects are associated with *higher* new cases from COVID-19. For Europe block, neither of the effects is significant. The epidemic started at different instants in time in each country, therefore, interpretation of coefficients is linked to the corresponding block. Parameterized as such, the intercept represents the unadjusted relative new case rate at each week (i.e., the beginning of the observed data) [21]. The Schengen agreement of 1985 that largely abolished internal border checks may help explain the lack of significance for the Europe Block.

The chains converge after 220,000 iterations with a burn-in period of 20,000. Three chains are simulated. The trace plots of informal diagnostic for each parameter (from multiple chains) show a stationary distribution. The Geweke

Table 4 Africa Block

	Estimate	Std. error	T value	Pr(>|t|)
(Intercept)	−7.167e+00	8.133e+00	−0.881	0.380
Week	2.870e+00	4.150e−01	6.917	1.46e−10
Area	−2.570e−05	1.845e−05	−1.393	0.166

AIC: 1499.1

Table 5 Asia Block

	Estimate	Std. error	T value	Pr(>|t|)
(Intercept)	−4.413e+02	8.403e+01	−5.252	1.12e−06
Week	2.204e+01	4.131e+00	5.335	7.94e−07
Area	9.329e−04	1.273e−04	7.329	1.31e−10

AIC: 1256.8

Table 6 Posterior quantities for selected parameters and DIC for North America Block

	Median	2.5%	97.5%	n.effective	Geweke.diag
(Intercept)	−3.481150e+02	−2047.1068	1.376670e+03	3658.3	0.2
Week	3.397180e+01	−88.2454	1.505980e+02	4000.0	−0.8
Area	1.000000e−04	−0.0003	5.000000e−04	4000.0	0.6
tau2	8.300000e−03	0.0022	8.150000e−02	2269.9	−0.7
nu2	1.308691e+08	91482053.7702	1.952245e+08	4000.0	−0.9
rho.S	3.678000e−01	0.0164	9.055000e−01	3822.7	−0.5
rho.T	3.029000e−01	0.0119	8.909000e−01	4000.0	0.0

DIC = 1165.661, p.d = 1.071452, LMPL = −583.7464

and Gelman-Rubin diagnostics (gelman.diag) compare the variability between and within the chain (See Tables 6, 7, 8, and 9). If values are closed to one, there is evidence that the chains have converged. The results show that the estimates of the potential scale reduction factor (PSRF) or R values are close to 1. The summary for the model parameters has the p-values of the predictors and the posterior quantiles. The posterior probabilities change from one block to another. The variable week (time) is significant for all Blocks, except Europe Block. Area is only significant in Asia and North America Blocks. According to the AIC, the best model fit is reported in Asia then followed by Africa, then Europe and North America, respectively. From Tables 6, 7, 8, and 9, all spatial and temporal correlations are around 0.38 and 0.30, respectively.

Further descriptive boxplots showing the temporal trends in the new cases between week 1 and week 29 are displayed in Fig. 5. They illustrate our vision that temporal trend must be block specific tied up with the variability of the measurements, in contrasts with the plots suggested in Figs. 3 and 4.

Figure 6a–d reveals that the algorithms converge for the intercept and slope parameters. Further examination reveals that the density plots (see right hand side of Fig. 6a–d) for the intercept and slope parameters are normally distributed. However, the density plot for the variance component is severely positively skewed as shown

Table 7 Posterior quantities for selected parameters and DIC for Europe Block

	Median	2.5%	97.5%	n.effective	Geweke.diag
(Intercept)	−128.0889	−1027.8663	759.0198	4000.0	−1.4
Week	36.3589	−24.7818	98.4758	4000.0	0.0
Area	0.0024	−0.0013	0.0058	4277.0	1.5
tau2	0.0083	0.0021	0.0919	3304.2	1.0
nu2	2094850.3225	1451687.7198	3180953.8029	4000.0	0.4
rho.S	0.3773	0.0176	0.9215	3503.6	0.5
rho.T	0.3048	0.0129	0.8953	3821.3	−1.2

DIC = 944.4763, p.d = 3.030266, LMPL = −474.2304

Table 8 Posterior quantities for selected parameters and DIC for Africa Block

	Median	2.5%	97.5%	n.effective	Geweke.diag
(Intercept)	−12.7454	−23.6519	−1.2798	2000.0	0.0
Week	3.1331	2.1474	4.0895	2000.0	−0.1
Area	0.0000	0.0000	0.0000	2000.0	0.9
tau2	0.0083	0.0022	0.0898	1056.3	−2.0
nu2	598.6476	453.0848	819.0348	2000.0	−0.6
rho.S	0.3800	0.0155	0.9164	1790.8	−1.5
rho.T	0.3133	0.0147	0.9024	2000.0	1.5

DIC = 836.4537, p.d = 3.921353, LMPL = −420.2662

Table 9 Posterior quantities for selected parameters and DIC for Asia Block

	Median	2.5%	97.5%	n.effective	Geweke.diag
(Intercept)	−88.7806	−158.0813	−17.3196	2000.0	−0.5
Week	12.3856	6.9171	17.9507	2000.0	−0.2
Area	0.0002	0.0001	0.0003	2000.0	1.2
tau2	0.0083	0.0022	0.0979	1330.3	−1.4
nu2	12194.6086	8364.9625	18165.3167	2000.0	1.5
rho.S	0.3682	0.0163	0.9135	1770.3	0.2
rho.T	0.3038	0.0145	0.8918	2000.0	−1.6

DIC = 667.2592, p.d = 3.824253, LMPL = −334.2956

in Fig. 7a–d. When variance components are approximately zero, there are two primary interpretations from a statistical perspective: (1) after controlling for all other parameters in the model, there is not enough variation in the response to attribute any variation to the random effect, and (2) despite the near zero random effect, the random effect should be retained because it is essential to the dependent structure in the data [22].

The different results in Fig. 8a–d different aspects of the underlying progression. Different blocks evidence the different significant predictors. Temporal correlation (week) was formally found significant in Africa, America and Asia. The results indicate that the spatially CAR model with spatially varying parameter highlights

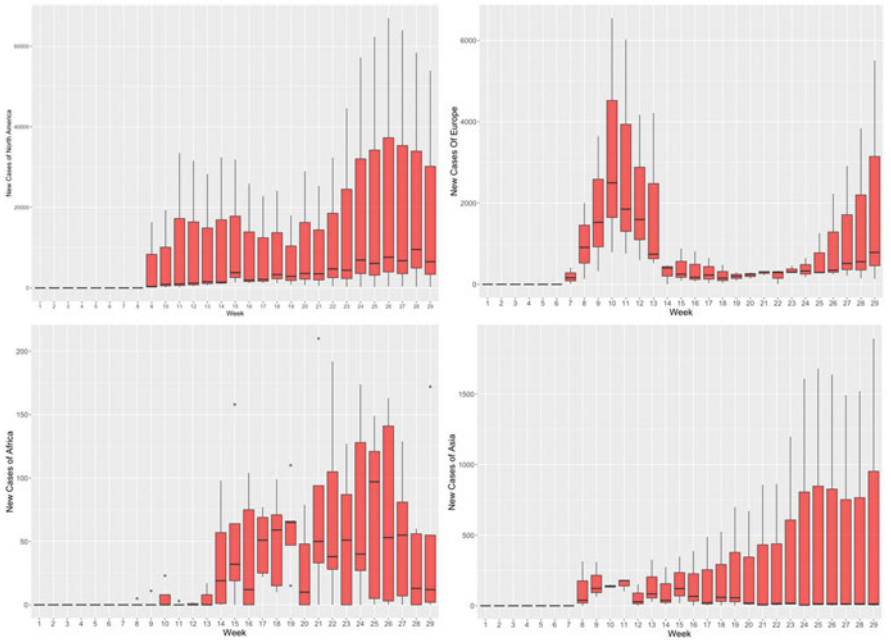


Fig. 5 Box plots of new cases under Bayesian models for the four blocks

the discrepancies among the four clusters and the spatial varying coefficients of the model.

The analysis also called for the investigation of the spatial autocorrelation. The latter captures the spatial variation of the new counts for each block. The spatial autocorrelation values indicate core areas of new cases incidences. To quantify the presence of spatial autocorrelation in the residuals from this model, we compute Moran's I statistic [23] and conduct a permutation test for each week of data separately. From Fig. 9, only negative Moran's spatial autocorrelation values are obtained. They then relate to the dissimilar values of cluster within a block, or spatial heterogeneity. The targeted locations are different from their neighbors. They imply that the data are dispersed. Also, all the blocks converge to a lower Moran's autocorrelation values, whereas the values are higher for the North America Block in absolute values.

It is possible to improve the estimation and the adaptive Bayesian model by integrating other predictors that may be more readily available. However, in such a simple context of the COVID-19, the observed scale of the clustering suggests infection operating processes are more localized and one single measure cannot describe this phenomenon.

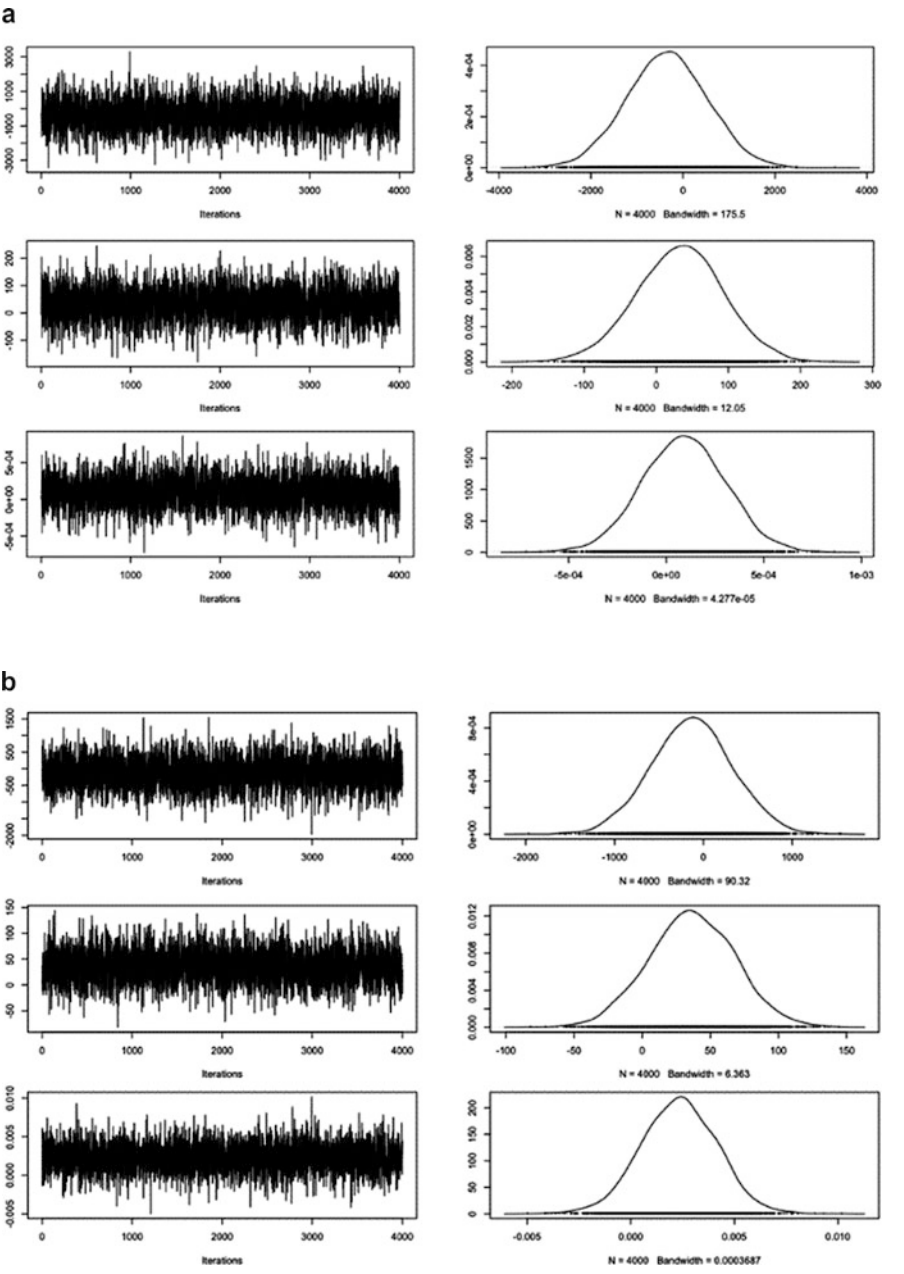


Fig. 6 (**a**) North America block. (**b**) Europe block. (**c**) Africa Block. (**d**) Asia Block. (**a**–**d**) Density plots and kernel densities of the three parameters (intercept, week and area) for North America, Europe, Africa and Asia Blocks

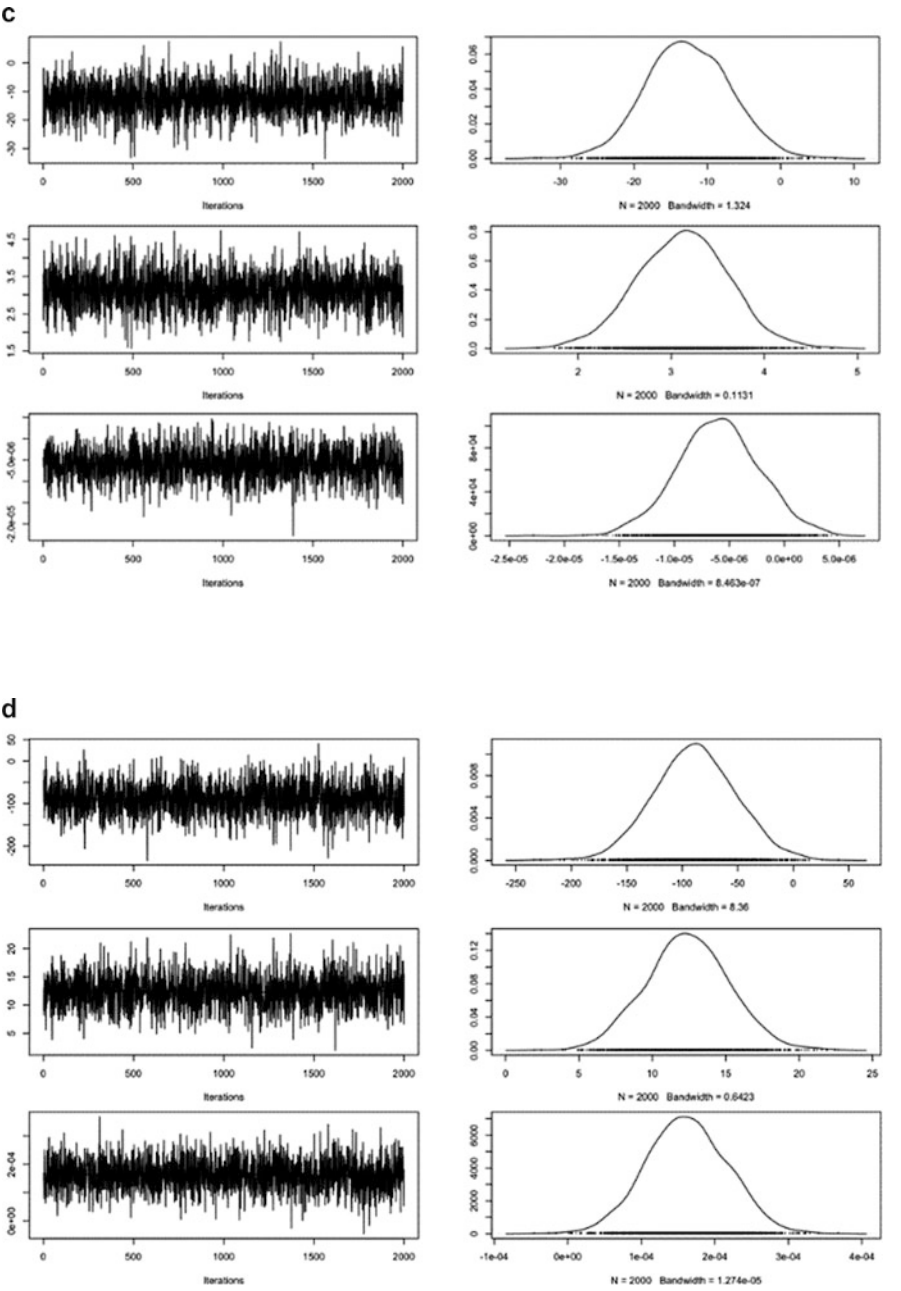


Fig. 6 (continued)

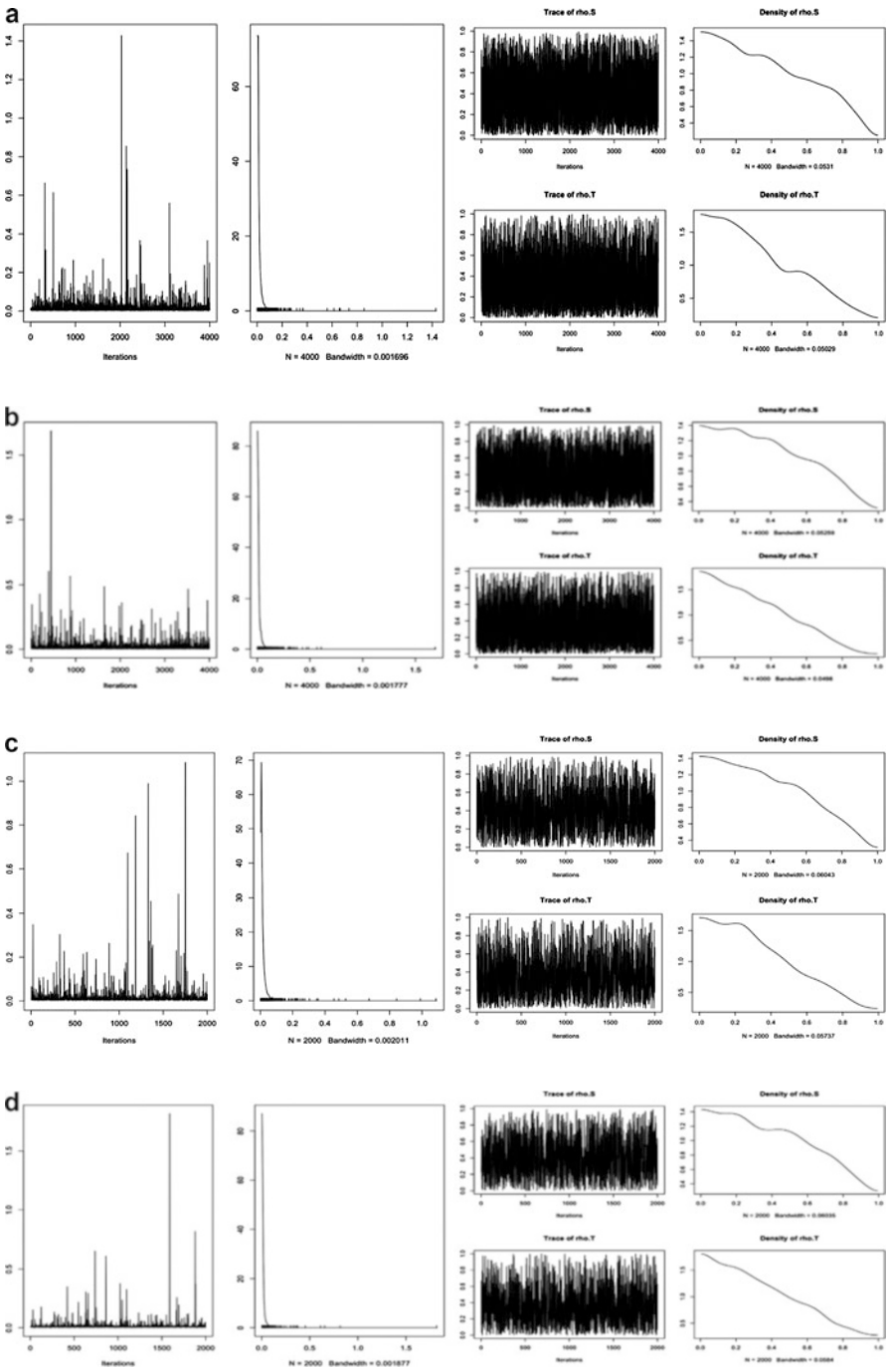
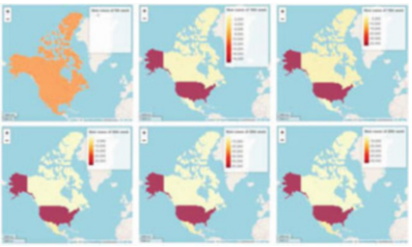


Fig. 7 (**a**) North America block. (**b**) Europe block. (**c**) Africa block. (**d**) Asia block. (**a**–**d**) Plots of the variances for the North America, Europe, Africa and Asia Blocks

a

NORTH AMERICA BLOCK

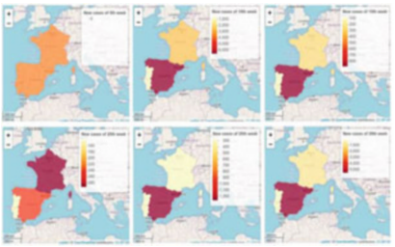


Point est.	Upper C.I.
0.9999116	1.0002020
0.9997652	0.9998028
0.9999339	1.0002634

Multivariate psrf: 0.999893

b

EUROPE BLOCK

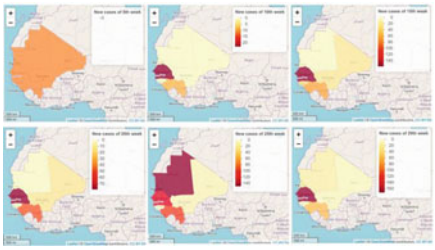


Point est.	Upper C.I.
0.9999692	1.000431
1.0000057	1.000072
1.0001241	1.001023

Multivariate psrf: 1.00012

c

AFRICA BLOCK

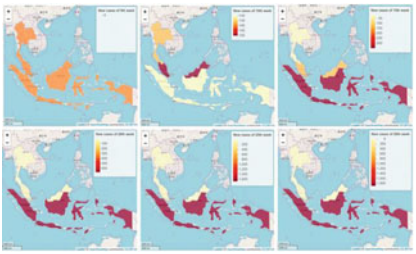


Point est.	Upper C.I.
1.000899	1.002185
1.001561	1.003450
1.000690	1.002788

Multivariate psrf: 1.000929

d

ASIA BLOCK



Point est.	Upper C.I.
1.001177	1.003723
1.000812	1.002740
1.002007	1.003601

Multivariate psrf: 1.00162

Fig. 8 (**a**–**d**) Maps depicting the pattern of the COVID-19 over time in North America, Europe, Africa and Asia Blocks

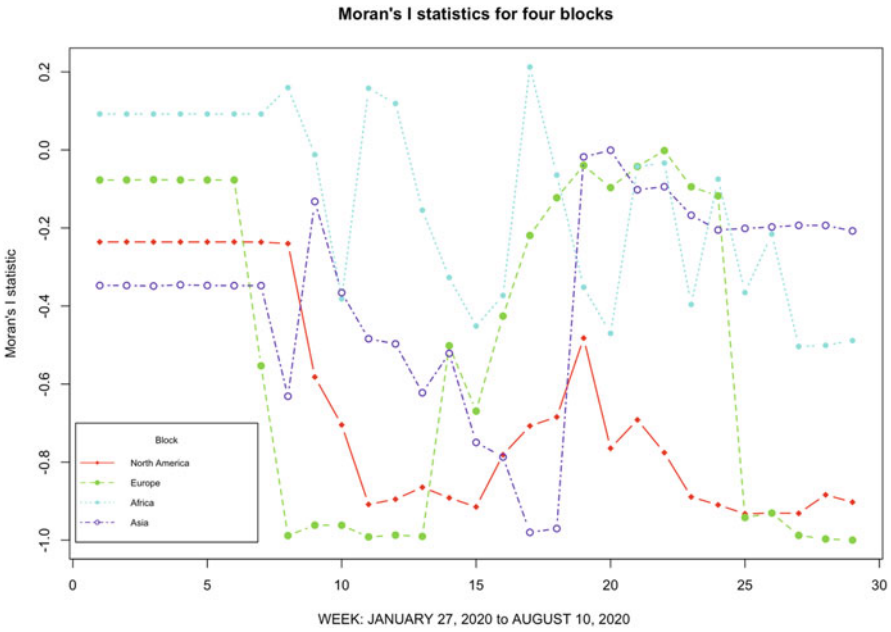


Fig. 9 Plots of the Moran's autocorrelation values by block

4 Discussion: Conclusion

This study sought to explore the understanding of the progression of the COVID-19 new cases. The spatio-temporal statistical models were applied to data on new COVID-19 cases in USA, Spain, Senegal and Malaysia over a 29-week period with consideration of neighboring countries. The factors influencing the pandemic based on the data available from the WHO website have been used to fit longitudinal structure models adding the neighborhood clustering into blocks. When all the counts were aggregated across blocks, results of a CAR model demonstrated that both week and area had statistically significant effects on new COVID-19 cases.

Other factors, such as occupations, race, age and location, could be taken into consideration, particularly for developing strategies and policy. In fact, the time and area factors that were selected may also underline other covariates. The findings enhance the understanding of the presumption that new cases are subordinate factor in understanding this pandemic. The country specific factors (such as income, population level of education and policy cultures) could be included along with time. Statistical investigations will increasingly shed light on the steps needed in minimizing the spread and the effort to control the disease.

Results of the current study demonstrate that CAR models are a viable modelling framework for any country and area in time. The CAR models may be applied to contact tracing to help slow down the spread of COVID-19. Also, they can enhance

the understanding of the pandemic (see Baum et al. [24]), which will help public health officials develop strategies and policies to slow down the spread based on occupations (restaurants, education, hospitals), race, age, and location (ZIP code). Nowadays, mass scale-up of testing and contact tracing are needed in order to contain the outbreak (see Larremore et al. [25]).

As more data become available at the city-level, rather than the country-level, CAR models and other disease mapping models can be used to identify high risk clusters or areas where disease incidence is highest. That is, Bayesian models can be applied to model the nested structure of the data with time nested within city, which is nested within country. These models will help identify key covariates at different levels (e.g., cost of delayed care, socio-economic measures, city and country) that can ultimately help reduce the development of current and future COVID-19 outbreaks. Such information is critical for public health and health care professionals to take effective remediation and prevention measure to limit or stop the spread of this disease. It helps to support prepare for potential outbreaks, and project occurrences of new cases of COVID-19.

5 Data Availability Statement

The dataset and code for this study can be obtained upon request.

Acknowledgements The research team is very thankful for the Editor and the Reviewers for comments and editing suggestions.

References

1. Bretschger, L., Grieg, E., Welfens, P. J., & Xiong, T. (2020). COVID-19 infections and fatalities developments: Empirical evidence for OECD countries and newly industrialized economies. International Economics and Economic Policy, 1–47.
2. McKinley, T., Cook, A. R., & Deardon, R. (2009). Inference in epidemic models without likelihoods. The International Journal of Biostatistics, 5, 1-40.
3. Lekone, P. E., & Finkenstadt, B. F. (2006). Statistical inference in a stochastic epidemic SEIR model with control intervention: Ebola as a case study. Biometrics, 62, 1170-1177.
4. Milletich R.J., N. Diawara, Jeng, H.A., (2015), "Modelling of the Deaths Due to Ebola Virus Disease Outbreak in Western Africa," Int. J. of Statistics in Medical Research, Vol. 4, pp. 306-321, https://doi.org/10.6000/1929-6029.2015.04.04.1.
5. Dewhurst, D.R.; Alshaabi, T.; Arnold, M.V.; Minot, J.R.; Danforth, C.M.; Dodds, P.S. (2020) Divergent Modes of Online Collective Attention to the COVID-19 Pandemic are Associated with Future Caseload Variance, arXiv:2004.03516
6. Gencoglu, O. and Gruber, M. (2020). Causal Modeling of Twitter Activity during COVID-19, Computation 2020, 8, 85; doi:https://doi.org/10.3390/computation8040085.
7. Millar, R. B. (2009). Comparison of hierarchical Bayesian models for overdispersed count data using DIC and Bayes' factors. Biometrics, 65, 962-969.

8. Hossain, M. M., Lawson, A. B., Cai, B., Choi, J., Liu, J., & Kirby, R. S. (2014). Space-time areal mixture modelling: Relabeling algorithm and model selection issues. Environmetrics, 25, 84-96.
9. Gardner, W., Mulvey, E. P., & Shaw, E. C. (1995). Regression analyses of counts and rates: Poisson, overdispersed Poisson, and negative binomial models. Psych Bulletin, 118, 392-404.
10. Navarro, A., Utzet, F., Puig, P., Caminal, J., & Martin, M. (2001). Negative binomial distribution versus Poisson in the analysis of recurrent phenomena. Gaceta Sanitaria, 15, 447-452.
11. Anderson, C., Lee, D., & Dean, N. (2014). Identifying clusters in Bayesian disease mapping. Biostatistics, 15(3), 457-469.
12. Anderson, C., Lee, D., & Dean, N. (2016). Bayesian cluster detection via adjacency modelling. Spatial and Spatio-Temporal Epidemiology, 16, 11-20.
13. Lee, D., Rushworth, A., & Napier, G. (2018). Spatio-Temporal Areal Unit Modelling in R with Conditional Autoregressive Priors Using the CARBayesST Package. Journal of Statistical Software, 84, 1-39.
14. Peccati, G., Solé, J., Taqqu, M., & Utzet, F. (2010). Stein's Method and Normal Approximation of Poisson Functionals. The Annals of Probability, 38(2), 443-478.
15. Finley, A. O., Banerjee, S., & Carlin, B. P. (2007). spBayes: An R Package for Univariate and Multivariate Hierarchical Point-Referenced Spatial Models. Journal of Statistical Software, 19(4), 1-24.
16. Banerjee, S., Gelfand, A. E., Finley, A. O., & Sang, H. (2008). Gaussian Predictive Process Models for Large Spatial Data Sets. Journal of the Royal Statistical Society B, 70(4), 825-848.
17. Finley, A. O., Banerjee, S., & Gelfand, A. (2015). spBayes for Large Univariate and Multivariate Point-Referenced Spatio-Temporal Data Models. Journal of Statistical Software, 63(13), 1–28.
18. Lawson, A. B. (2013). Bayesian disease mapping: Hierarchical modelling in spatial epidemiology. 2nd ed. Boca Raton, FL: CRC Press.
19. Gelman, A., Carlin, J., Stern, H., & Rubin, D. B. (2013). Bayesian data analysis. 3rd ed. London: Chapman & Hall.
20. Lunn, D. J., Thomas, A., Best, N., & Spiegelhalter, D. (2000). WinBUGS - A Bayesian modelling framework: Concepts, structure, and extensibility. Stats Comp, 10, 325-337.
21. Raudenbush, S. W., & Bryk, A. S. (2002). Hierarchical linear models. 2nd ed. Thousand Oaks: Sage Publications.
22. Kiernan, K., Tao, J., & Gibbs, P. (2012). Tips and strategies for mixed modelling with SAS/STAT procedures. SAS 2012 Global Forum: Orlando, FL.
23. Moran, P. A. P. (1950), Notes on Continuous Stochastic Phenomena," Biometrika, 37(1-2), 17-23.
24. Baum, A. (2020). Antibody cocktail to SARS-CoV-2 spike protein prevents rapid mutational escape seen with individual antibodies. Science, 1-7.
25. Larremore, D. B. (2020). Test sensitivity is secondary to frequency and turnaround time for COVID-19 surveillance. medRxiv. https://doi.org/10.1101/2020.06.22.20136309
26. Hawkins, D. M. (2001). Fitting multiple change-point models to data. Comp Stats Data Analysis, 37, 323-341.
27. Ng, S., Basta, N. E., & Cowling, B. J. (2014). Association between temperature, humidity and ebolavirus disease outbreaks in Africa, 1976 to 2014. Euro Surveill, 19, 1-11.
28. World Health Organization [homepage on the Internet]. Coronavirus disease 2019. Available from: http://www.who.int/mediacentre/factsheets/fs103/en/.
28. Cramb, S. M., Duncan, E. W., Baade, P. D., & Mengersen, K. L. (2017). Investigation of Bayesian spatial models. Brisbane: Cancer Council Queensland and Queensland University of Technology (QUT).
28. Leroux, B. G., Lei, X., & Breslow, N. (2000). Estimation of disease rates in small areas: a new mixed model for spatial dependence. In: M. E. Halloran and D. Berry (eds.), Statistical models in epidemiology, the environment, and clinical trials. Springer, New York.

Archimedean and Non-Archimedean Approaches to Mathematical Modeling

Bourama Toni

1 Introduction

1.1 Cultural Approaches to Mathematics

Have you ever wondered where mathematical ideas/concepts are coming from? Some think they are *invented* but others think they are actually *discovered*. We briefly present here two cultural approaches to mathematical creativity.

It is known that some of the leading French mathematician (France is the country of Descartes) held for some time a secular, rational perspective of mathematics, for instance, initially doubting the legitimacy of *infinite set*. These include René Baire (1874–1932), Henri Lebesgue (1875–1944), and Emile Borel (1875–1956). Borel was later captivated by Cantor's ideas. On the other hand Russian mathematician from the Moscow school of the so-called *Name-Worshipers* held dear human absolute freedom to *invent* mathematics, which results, among other *inventions* in the *descriptive set theory*. These Russian mathematician include Dmitrii Egorov (1869–1931), Nikolai Luzin (1883–1950), and their friend Pavel Florenskii (1882–1937). Some of these French mathematician even thought that Cantor's Alephs might prompt some mental disturbance: Cantor's own depression and Baire' suicide were linked to such concept. Indeed a particular environment, socio-economic-cultural, could influence mathematician working on the same math problems at the same time period but would come up with different conclusions/solutions.

One interesting example has been Set Theory, mathematical construct that involves the concept of Infinity: indeed the cardinality aleph $\aleph$ has been given to any

B. Toni (✉)
Department of Mathematics, Howard University, Washington, DC, USA
e-mail: bourama.toni@howard.edu

B. Toni (ed.), *The Mathematics of Patterns, Symmetries, and Beauties in Nature*,
STEAM-H: Science, Technology, Engineering, Agriculture, Mathematics & Health,
https://doi.org/10.1007/978-3-030-84596-4_8

set in bijective correspondence with the set of integers, then so-called *denumerable set*. The concept was introduced by the German mathematician Georg Cantor in 1882, coining the term *actual infinite* as opposed to the Aristotle *potential infinite* (thought of as a limit process). Cantor was inspired by Bernard Bolzano a proponent of the "actual infinite" who actually first introduced the word "set" [1].

Mathematician have considered Philosophy as soft, not rigorous thus non-mathematics. Emile Picard rejected Discontinuity for functions on a philosophical ground, adding that "Nature does not make jumps (natura non facit saltus): in nature there is no place for discontinuity". Henri Poincaré, the French mathematician who mostly relied on "Intuition" as opposed to 'logic" defended by Bertrand Russel and Giuseppe Peano, also rejected the idea of "actual infinity". That is, the "intuitionist" vs the "logicist" current in mathematics. The closed, centralized French educational system made the influence of Descartes even greater on French thinkers, scientists and mathematician; For Descartes mathematics is the universal and least biased form of knowledge, based of the simple principle of rigorous analysis and expression. Nicolas Boileau followed up with saying that "Anything that is understood well can be expressed clearly, and the words then come easily' (*Ce que l'on concoit bien s'énonce clairement, Et les mots pour le dire arrivent aisément*). Indeed for French mathematician through Cartesianism and Auguste Comte's Positivism, philosophical and mathematical questions should be totally separated, whereas for Russian mathematician, philosophical/religious and mathematical issues should be integrated. Geometry as well became very important in the French mathematical tradition due to the influence of Blaise Pascal's *there is no absolute truth just geometrical clarity*.

Religion, philosophy and mathematics were combined in the mind of the Russian mathematician Luzin, Egorov and Florenskii, to the point of seeing 'mathematical concept of 'continuity" as responsible for the "ethical decline' of the time, and of promoting "discontinuity" as morally and religiously superior. Proponents of the so-called *Name-Worshipping*, they also believe that "naming" is creating as in "let there be light: and there was light".

We conclude by paraphrasing Hermann Weyl [2] *Mathematics is the necessary instrument of natural sciences, and by its special character, certainty and stringency, it lifts the human mind into closer proximity with the divine than is attainable through any other medium; mathematics is the science of infinite, its goal, the symbolic comprehension of the infinite with human/finite means.* Following the "invention" of Descriptive Set Theory by Luzin and its Moscow School of Mathematics, Henri Lebesgue conceded in 1930 that *mathematical exigencies and philosophical exigencies are constantly associated, one can even say fused*, adding that *M. Luzin examine les questions d'un point de vue philosophique et aboutit ainsi a des résultats mathématiques: originalité sans précédent* [3].

We next present a note on the expanding computer assisted mathematics.

1.2 Artificial Mathematics

In these times of machine/deep learning and artificial intelligence, we should ask ourselves whether mathematical creativity is uniquely human. Von Neumann in his book *The computer and the brain* [4] has considered computers' potential ability to think, a thinking of a different nature, keeping in mind that both the computer and the brain are "information processing centers': the brain is perceived as slow, prone to error, with limited memory and highly parallel, whereas the computer is very fast, reliable, with possibly large memory, not very highly parallel.

Therefore a follow up concern is the nature of the mathematics to be created by artificial mathematical creativity; computer/artificial intelligence is indeed quite different from human intelligence. Nowadays computers perform heavy logical or numerical tasks beyond human capabilities leading sometimes to rigorous proofs such as the proof of the four-color theorem by Kenneth Appel and Wolfgang Haken. The so-called computer verified proofs or formal proofs are more reliable than human verified proofs. As noted by Hadamard and Poincaré, oftentimes doing mathematics is a combinatorial task, resulting in a theorem from various pieces put together with accepted rules of logic.

The next question is to whether there is *a structure to mathematics which is independent of the human brain*, to quote David Ruelle in his work on "Post-Human Mathematics" [5], which indeed inspired this section. We again paraphrase Ruelle as saying *There are two sides to mathematics: one is non-human logical necessity, the other is human brain activity*.

We should thus acknowledge there is a strong possibility that computer/artificial mathematical prowess will soon surpass human mathematical creativity, providing in the process some advanced mathematical modeling, possibly in a language impenetrable to the human brain. Should we be concerned/worried by such an eventuality?

1.3 Qualitative Mathematics

Interest in qualitative mathematical methods has been expanding during the last decades, mostly due to the wide applications in biosciences, social sciences including behavioral sciences and economics. In these mathematical models variables define parts of the modeled system, parameters designate factors that influence the system dynamic but are not usually influenced by it; the system dynamics are defined as relationships among system parts and between the parts and some extra systematic factors; the dynamics are described by the model equations which could be in the differential or difference format. For models representing complex systems in biological and behavioral sciences, it is usually impossible or infeasible to determine the quantitative value or the precise functional form of most of the interactions between system parts. However, it is often possible to determine the

qualitative properties of these interactions; sometimes what can only be ascertained is that there is or there is not interaction between variables, which could be translated by yes or no, 0 or 1, e.g., in *Boolean* models, making qualitative modeling more appropriate in these sciences. For example within ecology, qualitative models are more easily ascertained in the attempt to estimate intrinsic growth rate, carrying capacity, competition coefficients. Economists trust more the sign and direction of interactions between major parts of the economy but doubt their functional form can be determined more precisely. In psychology, there is little expectation for a precise mathematical function to accurately represent human behavior as reflected in imprecise belief states or preferences of typical real-world agents. Indeed, in biosciences, physical-chemistry, economics and behavioral sciences, informations about the underlying dynamics often reside in the rules of construct of the system and not in the absolute quantitative values. The data and phenomena being studied are essentially qualitative. Therefore, absent the precise quantitative, qualitative modeling concerns what properties, in particular dynamical properties, can be derived from these qualitative relations between the model variables.

Results established by qualitative models, with less commitment to details, tend to achieve a greater generality. In addition this type of modeling allows an understanding of phenomena less susceptible to the drawbacks of the quantitative usual idealization methods. Simplifications are inherent to both quantitative and qualitative models; in the former, they are realized by decreasing specificity, whereas, in the latter, they usually involve unrealistic assumptions in order to use some precise and tractable mathematical equations with fewer or more easily estimable parameters, in the hope that these intentional misrepresentations will not distort the salient features of the system. Several qualitative methods have been proposed. See Levins, Puccia and Levins, Orzack and Sober [6, 7]. However as claimed by Levins *Scientific modeling can maximize at most two of three virtues: generality, realism, and precision:*

(1) Sacrifice generality for precise quantitative predictions about specific systems and maximize realism by representing as many system details as possible.
(2) Sacrifice realism to make unrealistic assumptions so systems can be described with general mathematically tractable equations producing precise quantitative predictions.
(3) Sacrifice precision to abandon quantitative accuracy for qualitative relations between variables for maximum generality and realism.

We present in the last section a model of qualitative mathematics based on the Jacobian Feedback loops methodology.

2 The Non-Archimedean or Ultrametric/p-adic Approach

2.1 p-adic Mathematical Physics

Hermann Weyl is quoted with saying in Philosophie der Mathematik und Naturwissenschaft 1927, p. 36: *As a matter of fact, it is by no means impossible to build up a consistent "non-Archimedean" theory of magnitudes in which the axiom of Eudoxus (usually named after Archimedes) does not hold* [8].

Indeed the Axiom of Eudoxus/Archimedes is the main difference between the real and p-adic/ultrametric space; however the axiom is more of a physical one which concerns the process of measurement: for instance, exchanging the real numbers field with the p-adic number field is tantamount to exchanging axiomatics in quantum physics.

Ultrametricity in physics means the emergence of ultrametric spaces in physical models, as first used in the 1980s by Parisi and others in the theory of spin glasses and by Frauenfelder and others in physics of proteins. In both cases, the space of states of a complex system has a hierarchical structure which play a central role in the physical behavior of the system. On the other hand, in the 1930s, Bronstein showed that general relativity and quantum mechanics imply that the uncertainty Δx of any length measurement satisfies $\Delta x \geq L_{\text{Planck}} := \sqrt{\frac{\hbar G}{c^3}}$, where L_{Planck} is the Planck length ($L_{\text{Planck}} \approx 10^{-33}$ cm). This implies that spacetime is not an infinitely divisible continuum. Mathematically speaking, spacetime must be a completely disconnected topological space. The ultrametric spaces are naturally completely disconnected. There are several possible interpretations of the Bronstein inequality. One of them drives the loop quantum gravity. Another interpretation of Bronstein's inequality was given by Volovich in the 1980s. The inequality mentioned implies that real numbers cannot be used in models at the level of Planck's length, because the Archimedean axiom, which appears naturally if we use real numbers, implies that lengths can be measured with arbitrary precision. Volovich [9] proposed using p-adic numbers in physical models at the Planck scale.

These ideas have propelled development of a very large number of areas in mathematics and theoretical physics, which in turn has led to applications in computer science, biology, etc. For instance, the p-adic dynamics (discrete time flows) has been proved to be effective in a variety of areas: computer science (p-adic matrix processors; parallel p-adic linear solver); cryptography (stream ciphers, T-functions); automata theory and formal languages, genetics, and data mining, among other areas, see e.g. [10].

To paraphrase Murtagh [11], ultrametricity is a pervasive property of observational data. It offers the theoretical framework and processing tools to handle the "big data" and high dimensional data sets. One could say that "the p-adic methodology" consists in representing/interpreting phenomena/data in ultrametric spaces to in turn produce models that can be studied using non-Archimedean mathematical techniques. In protein physics [12], it is regarded as one of the most

profound ideas put forward to explain the nature of distinctive life attributes. As a consequence of this, the stochastic processes on ultrametric spaces and their connections with models of complex systems have received a lot of attention in recent years, see e.g. [10], and the references therein.

On the other hand, the relevance of constructing p-adic quantum field theory was stressed in [13] and [9]. In the last 35 years p-adic QFT has attracted a lot of attention of physicists and mathematicians, see e.g. [14] and the references therein.

2.2 Mathematically Thinking p-adically

A noted weakness of the current trends in mathematical modeling is that the classical mathematical framework involves mostly, if not exclusively, Archimedean spaces (e.g. Euclidean, Banach or Hilbert spaces). These are spaces that satisfy the *Archimedean Principle* in that there are all endowed with the usual/standard Euclidean norm and its induced metric which satisfied the triangle inequality; much of our visual and mental perception is based on the standard Euclidean space with its perfectly straight lines and planes; we came to see and represent the physical universe as actually Euclidean in its geometry, resulting oftentimes in biased and not so realistic mathematic model.

Indeed the physical and natural systems, described by the so-called Evolution Equations, are inherently non-Euclidean, with a natural geometrical ordering that is not the usual real line, but the more adequate *hierarchical generating tree.* For instance the usual geometrical Archimedean/Euclidean distance "suitable" for measuring spatial location separation between human beings is less effective for the genetic distance measuring the hierarchical kinship relations.

Fortunately, thanks to the pioneering work by Kurt Hensel in the late nineteenth century, a new trend has been emerging in recent years, a differential analysis, in Non-Archimedean or *p-adic* spaces, endowed with the so-called *ultrametrics;* ultrametrics could be induced by *p-adic absolute values* $|.|_p$, for p prime, that satisfy a more stringent inequality than the usual triangle inequality; specifically in a non-Archimedean space $\mathbb{K}$, the *ultrametric* distance d (or denoted $d_p(x, y)$ when induced by the *p-adic norm* $|.|_p$) satisfies the inequality

$$d(x, z) \leq max(d(x, y), d(y, z)) \quad \forall \mathrm{x}, \mathrm{y}, \mathrm{z} \in \mathbb{K}. \tag{2.1}$$

Recall that the set $\mathbb{R}$ of reals is the completion of the set $\mathbb{Q}$ of rationals by the usual Euclidean infinite norm denoted $|.|_\infty$, which amounts to creating new numbers as limits of Cauchy sequences that do not have rational limits. The same construction replacing the norm $|.|_\infty$ by the p-adic norm $|.|_p$ yields a new complete field, analogous to $\mathbb{R}$ denoted $\mathbb{Q}_p$, for every prime p, and called the field of *p-adic numbers*. Inside this field $\mathbb{Q}_p$ of p-adic numbers lies the ring $\mathbb{Z}_p$ of *p-adic integers* whose geometry is similar to the Cantor set. The set of rationals $\mathbb{Q}$ is densely contained in every $\mathbb{Q}_p$ as it is in the set of reals $\mathbb{R}$.

In short mathematical modelers of natural and physical have at their disposal a toolbox of various fields in the so-called *book structure* with the rationals $\mathbb{Q}$ as the book spine and every page represents a field in which to carry mathematical modeling, to include $\mathbb{R} = \mathbb{Q}_\infty$, and $\mathbb{Q}_p$, for every p prime.

Non-Archimedean spaces, equivalently called *p-adic spaces* or *ultrametric spaces,* have some peculiar and important features described as follows:

- One immediate interesting fact about *p-adic integers*, is that these integers are bounded in the norm by 1; that is, $|n|_p \leq 1$, i.e., as the number tends to infinity, its p-adic size remains less than one and tends to zero, in violation of the *Archimedean Principle*.
- The topology has a basis of *clopen* sets, i.e., sets that are simultaneously open and closed; so phrases such "open ball" and "closed balls" become meaningless
- Every point in a p-adic ball is also its center. For instance if *modeled p-adically* the center of the universe could be found at the tip of the nose, as once alluded to by the physicist Hawkins.
- Two balls are either disjoint or one within the other. This could hint towards the notion of parallel universes.
- The p-adic spaces are *totally disconnected*: the connected component of every point is the point itself. Consequently the principle of *analytic continuation* is lost, as well as the *Intermediate value theorem*.
- The ultramatric geometry allows only *isosceles triangles*.
- The geometrical ordering is not along the real line but rather on a *hierarchical generating tree.* Consequently the notion of time as we know seems to be a purely statistical construct. $\mathbb{Z}_3$ for instance is homeomorphic to the fractal-like Sierpinski Gasket.
- Convergence of p-adic series: $\sum_{n=1}^{\infty} a_n < \infty$ if and only $a_n \to 0$. (A calculus student dream!)
- The most consequential outcome and challenge is the use of differential representations with a *p-adic time* instead of the usual real time. That is, how to model a system that evolves *p-adically.* This is still an hard open problem.
- Consider the following simple looking differential equations modeling so many natural systems:

$$\frac{dy}{dt} = \dot{y} = \lambda y, \tag{2.2}$$

and

$$\frac{d^2y}{dt^2} + \omega^2 y = 0, \quad \text{(Harmonic Oscillator)} \tag{2.3}$$

Respectively the solutions are in the form $Ae^{\lambda x}$, and as level curves $H^{-1}(c) \quad c \in \mathbb{R}_{\geq 0}$ of a quadratic Hamiltonian H (i.e. concentric circles in the Euclidean plane $\mathbb{R}^2$).

However the *p-adic exponential function* $exp_p(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ converges only for $x \in B(0,r), \quad r = p^{1/(1-p)}$; that is, a very small radius of convergence in contract of the standard exponential series that converges on the whole real line $\mathbb{R}$. Therefore the series $\sum_{n\geq 1} \frac{(\lambda x)^n}{n!}$ converges if and only if $|x| < |\lambda|^{-1} p^{1/(1-p)}$.

- Existence of *singular function* or Devil's Staircase or Cantor function: in the classic real analysis a function differentiable with vanishing derivative is a constant function. In contrast there exist in the p-adic space, functions that are *non-constant, continuous and differentiable with vanishing derivatives*; these functions are called *pseudo-constants.* They are used to describe situation in which an object is known to move from point A to point B, but whenever it is observed it appears to be at rest. "Spooky motion" Einstein would say!
- The next most consequential notion is that of periodicity. Indeed, periodicity, an important concept in differential mathematical modeling, is non-existent in non-Archimedean spaces; functions cannot be non-trivially periodic; however almost periodicity is allowed and seems to be the norm in the p-adic world. In other words, systems modeled do not have a time-periodic behavior in their p-adic dynamics. For instance it has been noted (e.g., Corduneanu) that almost periodic oscillations are much more common than the periodic ones. There has been a tremendous amount of work in mathematical sciences involving periodicity, to include the famous Hilbert's 16th problem, still unsolved to this day. Therefore all these results on periodicity are made meaningless by just changing to the non-Archimedean/p-adic spaces. Such a realization is humbling and should be kept in mind when drawing conclusions based on periodicity.

In non-Archimedean/ultrametric spaces, almost periodicity is defined and understood in the following sense: consider a non-trivially valued complete algebraically closed non-Archimedean field $\mathbb{K}$, and a $\mathbb{K}-$valued function f over an additive semigroup $\mathbb{G}$.

The function f is said to be *p-adic/non-archimedean almost periodic function* if the set $f_{\mathbb{G}}$ of all the translates of f is a *compactoid* in the $\mathbb{K}-$Banach algebra $\mathscr{B}$ of all bounded continuous functions $\mathbb{G} \longrightarrow \mathbb{K}$ with respect to pointwise operations and the supremum norm $||f||_{\infty} = \sup_{\mathbb{G}} |f(x)|$.

That is, $\forall \epsilon > 0, \ \exists$ a finite subset $F_{\epsilon} \subset \mathscr{B}$ such that

$$f_{\mathbb{G}} \subset B_{\epsilon} + aco(F_{\epsilon}) \tag{2.4}$$

where $aco(F_{\epsilon})$ is the *smallest absolutely convex* subset containing F_{ϵ}. In the classic notion of almost periodicity the set F_{ϵ} is only required to be *relatively compact/pre-compact.*

In the natural sciences ultrametricity is emerging as a consequence of randomness and the law of large numbers; exact in the limit $N \longrightarrow \infty$ for systems with a large number N of degrees of freedom, ultrametricity provides a more natural type of organization. It has proven effective for studying Evolution Equations describing neutral evolution of *pseudogenes*, stochastic branching processes in large space, energy landscape of *disordered frustrated systems* (spin glasses, problems in

engineering and biology of combinatorial optimization), and in *taxonomy* where representation is given by dendrogram of hierarchy pictured in inverted tree. Ultrametricity serves also to better model Mental spaces, and the emerging and evolution of languages.

The richness of non-Archimedean space allows more realistic mathematical models, e.g., for evolution equations, making it a state-of-art tool in the arsenal of researchers; however still few math modelers have expertise in Non-Archimedean analysis. That is, the classical, Archimedean mathematical modeling must be redirected towards *p-adic/ultrametric mathematical modeling*.

2.3 p-adic Mental Spaces

The development of a mathematicall rigorous model for the real physical space spread over three hundred years and has been based on the real continuum, assumed to be a continuous infinitely indivisible space that is also homogeneous (all points have equal rights in geometric representations). It has involved the Cartesian product of real lines to embed physical objects into a mathematical space. This model has been very successful in creating today modernity. Researchers have been trying to find similar successful mathematical model for the life and behavioral/cognitive sciences, including an adequate model for the *mental space*. Khrennikov and his collaborators have made significant contributions to the latter; they proposed a model of mental space that is discontinuous, hierarchical, and p-adic. They have extended their mathematical p-adic approach to Freud's theory of interaction between conscious and subconscious mental domains, with the subconscious domain described by complex dynamical systems acting as *thinking processors*; this is the starting of the *p-adic cognitive science*, leading to the mathematical ultrametric classification for neurotic behaviors, idées fixes, hysteria, etc . . . [15].

3 The Archimedean or Euclidean Approach

An ordered field $< F, +, \cdot, \leq >$ has the *Archimedean Property* if, given any positive x and y in F there is an integer $n > 0$ so that $nx > y$.

Equivalently, $\forall \epsilon \in F, (\epsilon > 0 \rightarrow \exists n \in \mathbb{N} \mid \frac{1}{n} < \epsilon)$. An Archimedean algebraic structure has no pair (x, y) such that x is infinitesimal with respect to y.

3.1 Signed Qualitative Modeling: An Example

We give here an overview of a recent and progressive methodology in qualitative modeling to showcase how much can be achieved about the structure and behavior of systems partially specified by using the sign of an interspecific interaction [16]. In a series of studies we have developed an efficient tool in the qualitative study of systems described by differential or difference equations, namely, a tool based on the dynamical roles of Jacobian Feedback Loops. Such a tool intends to survey the dynamical response of models simulating physico-chemical, biological and economical systems by stressing qualitative understanding as the primary goal rather than numerical prediction [17–22].

Dynamical systems theory is mostly based on quantitative values of the Jacobian entries. But for some systems, mainly in biosciences such as biology and biochemistry, in economics and behavioral sciences, the relevant informations are of qualitative nature. Quantitative results are rare in studying interactions in a system of biochemical compounds: A gene X could be shown to be an activator (or a repressor) of the expression of a gene Y, but usually without knowing the strength of the interaction, the concentrations and their kinetics. In a conflict resolution model featuring the variables of Attitude (A), Behavior (B) and Contradiction (C), one can "accurately" determine if the variables mutually influence each other positively (+), negatively (−) or no influence (0) with no need to quantify the strength of the influences.

The theory of Jacobian loops is therefore the analysis of the dynamics (simple and complex) using solely the loop-pattern Jacobian matrix, that is, even when only the signs, not the magnitudes of the Jacobian terms, are known.

Consider the autonomous differential system

$$\begin{aligned} \dot{x}_i(t) &= \frac{dx_i(t)}{dt} = F_i(x,a) \\ x &= (x_1, \cdots, x_n) \in \mathbb{R}^n, \quad a = (a_1, a_2, \cdots, a_N) \in \mathbb{R}^N, \end{aligned} \tag{3.1}$$

describing a dynamical system with phase space in $\mathbb{R}^n$, and the parameter/control space in $\mathbb{R}^N$. The component functions F_i, $i = 1, 2, \cdots, n$ of $F(x, a)$ are assumed to be at least $C^1(U)$, that is, differentiable along with their first partial derivatives on U an open set of $\mathbb{R}^n$. The partial order relation $x \leq y \Longleftrightarrow x_i \leq y_i, \quad i = 1, ..., n$ defines the vector order in $\mathbb{R}^n$. The Jacobian matrix at $\bar{x} = (\bar{x}_1, \ldots, \bar{x}_n)$ is given by

$$J(\bar{x}) = DF(\bar{x}) = \left[\frac{\partial(F_1, \cdots, F_n)}{\partial(x_1, \cdots, x_n)}(\bar{x})\right] = \left[\frac{\partial F_i}{\partial x_j}(\bar{x})\right]_{1 \leq i,j \leq n} = [J_{ij}]_{1 \leq i,j \leq n}, \tag{3.2}$$

and in general depends on the state variables, except for linear systems [23].

Remark 3.1 Actually the relation $\dot{x}_i(t) = F_i(x, a)$ shows how the rate of change in variable x_i is dependent on changes in any given variable x_j. Therefore the Jacobian entry $\frac{\partial F_i}{\partial x_j}(x) = J_{ij}$, for $1 \leq i, j \leq n$ describes the interaction between the variables x_i and x_j, as positive (respectively negative, no) interaction for $J_{ij} > 0$ (respectively $J_{i,j} < 0$, $J_{i,j} = 0$).

3.2 Jacobian Feedback Loops

Let $\mathscr{I}_n$ be the set of indices $1, \cdots, n$ and denote by $I_k = \{i_1, \cdots, i_k\}$ an ordered subset of k different elements of $\mathscr{I}_n$ and by $\tilde{I}_k = \pi_k(I_k) = \{j_1, \cdots, j_k\}$, with $\pi_k \in \Xi_k$ a permutation of I_k. Recall $Card(\Xi_k) = k!$, i.e., there are $k!$ permutations. Every permutation π_k may be factored into ν disjoint circular (cyclic) permutations $\sigma_i, i = 1, \cdots, \nu$, that is, $\pi_k = \sigma_1 \sigma_2 \cdots \sigma_\nu$. The signature of π_k, denoted $sg(\pi_k)$, is $(-1)^\eta$, η the number of inversions in π_k, that is, the number of pairs (j_m, j_n) with $j_m > j_n$ while $i_m < i_n$, for $j_m = \pi_k(i_m)$, and $j_n = \pi_k(i_n)$. The permutation π_k is even (resp. odd) for an even (resp. odd) η. There are exactly $\frac{k!}{2}$ even and exactly $\frac{k!}{2}$ odd permutations in Ξ_k. We denote Ξ_k^c (resp. Ξ_k^e, Ξ_k^o) the subset of circular (resp. even, odd) permutations. The set Ξ_n is the classic symmetric group of permutations on the set of indices $\mathscr{I}_n$.

We have the following defining concepts [19, 20].

Definition 3.2

1. The set of nonzero terms J_{ij}, $i \in I_k$, and $j \in \tilde{I}_k$, describes a *Jacobian loop* associated with the nonzero product

$$P(\pi_k, J) := \prod_{l=1}^{l=k} J_{i_l \pi_k(i_l)} = J_{i_1 \pi_k(i_1)} J_{i_2 \pi_k(i_2)} \ldots J_{i_k \pi_k(i_k)} \tag{3.3}$$

   called a *loop product*.
2. The loop is called a $k-$order *simple Jacobian loop* L_k when the permutation π_k is a $k-$cycle$= (i_1, i_2, \cdots, i_k)$ with the loop product

$$p_k = J_{i_1 i_2} J_{i_2 i_3} \cdots J_{i_{k-1} i_k} J_{i_k i_1}. \tag{3.4}$$

3. Its sign $sgn(L_k)$ is that of the loop product $p_k = P(L_k) := P(\pi_k, J)$. Its length or dimension $l(L_k) = k$ is the number of loop factors $J_{i_l \pi_k(i_l)}$ involved, as well as the number of related system variables x_i.

Remark 3.3

1. A simple Jacobian loop L_k is positive (resp. negative) for an even (resp. odd) number of negative loop factors $J_{i_l \pi_k(i_l)}$ in the loop product $P(L_k)$.

2. The loop L_k has the following representation called *loop graph* or *interaction graph* and denoted by $\mathbb{L}_k$: it consists of k distinct vertices given by the system variables $x_i, i = 1, \ldots, k$ and k edges $E_{ij} = (x_i, x_j, s_{ij})$ directed from j to i where $s_{ij} = sign(J_{ij})$ denotes the nature of the interaction between the variable x_j and x_i.
3. A positive (resp. negative) loop involving the variables $x_1, x_2, \ldots, x_k$ is also conveniently denoted $L^+_{x_1 x_2 \cdots x_k}$ (resp. $L^-_{x_1 x_2 \cdots x_k}$).
4. A loop graph $\mathbb{L}_k$ is *complete* if for every $i \neq j$ there is a directed polygonal line connecting x_i to x_j, that is, $\overline{x_i x_{k_1}}, \overline{x_{k_1} x_{k_2}}, \cdots, \overline{x_{k_r} x_j}$.

Definition 3.4

1. A non-circular permutation π_k yields a *union* of simple Jacobian loops, called a *composite loop* $\mathscr{L}^\nu_k = \cup_{i=1}^{i=\nu} L_i = (L_1, \ldots, L_\nu)$ of dimension $l(\mathscr{L}^\nu_k) = k$ given by the sum of the lengths of its ν simple component loops, i.e., $k = \sum_{i=1}^{i=\nu} d(L_i) = 1 + \cdots + \nu$.
2. A proper composite loop $\mathscr{L}^\nu_k$ of *resonance* (ν, k) is a disjoint union of ν simple loops of total length k, that is, the component loops do not share a vertex.

Remark 3.5 We denote P^ν_k the loop product of a composite loop $\mathscr{L}^\nu_k$. The sign of a composite loop $\mathscr{L}^\nu_k$ is the sign of P^ν_k, or equivalently, $sign(\mathscr{L}^\nu_k) = \prod_{l=1}^{l=\nu} sign(L_i) = (-1)^{\nu_-}$, where ν_- is the number of negative simple loops in $\mathscr{L}^\nu_k$.

$\chi^\nu_k = (-1)^{\nu+1}$ is the characteristic of the proper composite loop of resonance (ν, k). Therefore a $k-$order proper composite loop has a negative (resp. positive) resonance, i.e., a negative (resp. positive) characteristic for ν even (resp. odd).

Definition 3.6 A $k-$order Feedback F_k is defined by

$$F_k = \sum_{all\nu} (-1)^{\nu+1} P^\nu_k, \tag{3.5}$$

where P^ν_k is the loop product of the proper composite loop $\mathscr{L}^\nu_k$.

Consequently we have the following

Remark 3.7

1. A composite loop $\mathscr{L}^\nu_k$ is positive (resp. negative) for an even (resp. odd) number of its negative simple loops.
2. A proper composite loop $\mathscr{L}^\nu_k$ with all component simple loops negative has a negative resonance in the Feedback F_k as defined above.

Definition 3.8

1. We call *qualitative matrix* S a matrix consisting exclusively of the signed entries $s_{ij} \in \{+, -, 0\}$, that is, $S := [s_{ij}]_{1 \leq i,j \leq n}$. We denote S by A_q for a qualitative matrix associated with a matrix $A = [a_{ij}]$, that is, $s_{ij} = sign(a_{ij})$, for $1 \leq i, j \leq n$.

2. The *loop structure (or qualitative structure),* denoted $\mathbb{L}_{\mathscr{R}}$, corresponding to the region $\mathscr{R}$ in the phase space or to a sign-pattern is the set of all Jacobian loops (simple and proper composite) along with their signs.

Remark 3.9 The Jacobian loop analysis requires first the determination of the loop structure associated with the system in a given region of the phase space, either from the signed entries of a Jacobian matrix evaluated at equilibria or constant at some parameter values or solely from the qualitative evaluation of the interaction between the variables in terms of positive, negative or zero. Determining and analyzing the qualitative structure anywhere in the phase space, including around the steady states, if any, yields some understanding of the local and global dynamics of the system.

3.3 Loops and Jacobian Spectrum

For a matrix A, Jacobian or otherwise, given by $A = [A_{ij}]_{1\leq i,j\leq n}$ the characteristic polynomial is defined by the monic polynomial

$$\mathscr{C}_A(\lambda) = |\lambda I - A| = \lambda^n + c_1\lambda^{n-1} + \cdots + c_k\lambda^{n-k} + \cdots + c_{n-1}\lambda + c_n. \tag{3.6}$$

From Linear Algebra [24] the coefficients may be expressed as

$$c_k = coefficient(\lambda^{n-k}) = \sum(-1)^k m_k, \quad k = 0, \cdots, n-1, \tag{3.7}$$

where the sum extends over all kth order principal minors m_k of A. For instance we have

$$\begin{aligned} c_n &= (-1)^n det(A) = (-1)^n |A|, \quad \text{for } k = n, \\ c_1 &= -\sum A_{ii} = -Tr(A), \quad \text{where Tr(A) is the trace of } A. \end{aligned} \tag{3.8}$$

From the theory of determinant and permutations we may write

$$m_k = \sum_{\pi_k \in \Xi_k} (-1)^{\eta} \prod_{i_l \in \mathscr{I}_k} A_{i_l \pi_k(i_l)} = \sum_{all\, v} (-1)^{k-v} \prod_{i=1}^{i=v} P(\sigma_i, A) = \sum_{all\, v} (-1)^{k-v} P(\mathscr{L}_k^v), \tag{3.9}$$

where the permutation $\pi_k \in \Xi_k$ of the indices $1 \leq i_1 < i_2 < \cdots < i_k \leq n$ factors into the cyclic permutations $(\sigma_1, \cdots, \sigma_v)$ yielding the proper composite loop $\mathscr{L}_k^v = (\sigma_1, \cdots, \sigma_v)$ with loop product $P(\mathscr{L}_k^v) = P_k^v$ as defined above. Consequently we obtain an expression of the characteristic coefficients c_k in terms of the proper composite loops. Importantly we obtain [25].

Lemma 3.10 *The kth order Feedback $\mathscr{F}_k$ involving all the proper composite loops $\mathscr{L}_k^v$ with $v = 1, \cdots, k$ may be expressed in terms of the coefficients of the characteristic polynomial by*

$$\mathscr{F}_k = c_k = \sum_{all\, v} (-1)^{v+1} P_k^v, \quad k = 1, \cdots, n \tag{3.10}$$

Remark 3.11

1. First recall the zeros of the characteristic polynomial are the eigenvalues of the matrix A, that is, they are the elements of the spectrum Λ_A. They are of multiplicity m if $(z - \lambda)^m$ factorizes $\mathscr{C}_A(z)$. For $m = 1$ the corresponding eigenvalue is said to be simple, such as when A has n distinct eigenvalues.
2. Importantly the $k-$ order Feedback $\mathscr{F}_k$ being the $k-$order coefficient of the characteristic polynomial entails that the loop factors $A_{i\sigma(i)}$ defined above are the only Jacobian entries contributing to the characteristic equation, and therefore, influence directly the eigenvalues of the matrix, and consequently the dynamics.
3. From the standard theory of equations it is also known that the coefficients $c_k = \mathscr{F}_k$ are related to the eigenvalues λ_i in a systematic way by the following Viete formulas: See [24].

$$\begin{aligned}
c_1 = F_1 &= -(\lambda_1 + \cdots + \lambda_n) = -\sum_{i=1}^{i=1} A_{ii}. \\
c_2 = F_2 &= \sum_{i,j=1, i<j} \lambda_i \lambda_j = \sum_{i,j=1, i<j} (A_{ii}A_{jj} - A_{ij}A_{ji}) \\
&= \lambda_1\lambda_2 + \lambda_1\lambda_3 + \cdots + \lambda_{n-1}\lambda_n. \\
c_3 = F_3 &= -\sum_{i,j,k=1, i<j<k} \lambda_i \lambda_j \lambda_k \\
&= -(\lambda_1\lambda_2\lambda_3 + \lambda_1\lambda_2\lambda_4 + \cdots + \lambda_{n-2}\lambda_{n-1}\lambda_n). \\
&\cdots \\
c_n = F_n &= (-1)^n \lambda_1\lambda_2 \cdots \lambda_n.
\end{aligned} \tag{3.11}$$

Immediate from the above definitions and the theory of determinants, [24], we have

Lemma 3.12

$$\begin{aligned}
det(A) = |A| &= \sum_{\pi_n \in \Xi_n} sg(\pi_n) P(\pi_n, A) \\
&= \sum_{\pi_n^e \in \Xi_n^e} P(\pi_n^e, A) - \sum_{\pi_n^o \in \Xi_n^o} P(\pi_n^o, A).
\end{aligned} \tag{3.12}$$

A classic result therefore leads to the following results proved in [16]:

Theorem 3.13 *A necessary condition to have all eigenvalues with negative real parts $\mathscr{R}_e < 0$ is that all k*th *order Feedback $\mathscr{F}_k$ must be positive.*

Lemma 3.14 *A proper composite loop $\mathscr{L}_k^v$ with all v components simple loops negative has a negative resonance in the k*th *order Feedback, that is, its characteristic χ_k^v is negative.*

Theorem 3.15 *If there is no proper composite loop $\mathscr{L}_k$ of dimension $k \leq n$, then the characteristic coefficient $c_k = 0$.*

Moreover at least one proper composite loop $\mathscr{L}_n$ of the system dimension is necessary to have a nonsingular jacobian matrix.

Definition 3.16 We say that the qualitative equivalent class $\langle A \rangle$ or the loop equivalence class $\rangle A \langle$ is *qualitatively nondegenerate* if every matrix in the class is nonsingular in the sense $|A|$ is nonzero.

We prove in [16].

Theorem 3.17 *If the loop structure $\mathbb{L}$ does contains a composite loop $\mathscr{L}_n$ of the dimension of the system, and all such loop $\mathscr{L}_n$ have the same sign, then the corresponding Jacobian determinant $|A|$ is nonzero.*

Theorem 3.18 *A positive simple loop in the loop equivalence class is a necessary condition for the Jacobian matrix to have a positive real eigenvalue.*

3.4 Qualitative existence of Multiple Equilibria

Equilibria or steady states of system (1.1) are solutions of the equations $F(x, a_0) = 0$ at the parameter value a_0. Together with closed orbits they are the simple dynamics or limit sets of a system, and sometimes they are "essentially" all that can occur, e.g., for gradient systems and planar systems. Variants of the qualitative study of the existence of multiple equilibria may also be found in other literatures. See for instance [17, 18, 26–28].

Assume that the elements of the Jacobian matrix J are constant in a region $\mathscr{D}$ (open convex) of the phase space not necessarily a neighborhood of a steady state, and that the equivalence class $\langle J \rangle$ is qualitatively nondegenerate. Set $\mathbb{L}_{\mathscr{R}}$ to be the corresponding loop structure. We prove in [16].

Theorem 3.19 *Assume the loop structure $\mathbb{L}$ in a region $\mathscr{D}$ contains a composite loop $\mathscr{L}_n$ of the dimension n of the system, and that all such loops have the same sign. Then there is a subregion $\bar{\mathscr{D}} \subset \mathscr{D}$ where the dynamical system cannot have more than one fixed point.*

3.5 Applications and Examples

3.5.1 Thomas Conjectures

We restate Thomas's conjectures in the loop formalism [17, 18, 29, 30].

1. Conjecture 1 [1981] The presence of a positive feedback loop (somewhere in the phase space) is necessary condition for multistationarity.
2. Conjecture 2 [1981] The presence of a negative feedback loop of length at least two (somewhere in pahse space) is a necessary condition for stable periodicity.
3. Conjecture 3 [1999] Chaotic dynamics require both a positive feedback (*to allow multistationarity*) and a negative feedback loop (*to allow for permanent periodicity*).

Remark 3.20

1. Soulé in 2003 presented a proof of conjecture 1. Some partial results were also given by Plahte et al (1995), Snoussi (1998), Gouzé (1998), Cinquin and Demongeot (2002). See [27, 28, 31–33]
2. Under additional assumptions Snoussi and Gouzé also proved Conjecture 2. See [27, 34]
3. Conjecture 3 is also included in a more general conjecture by Toni et al in 1999, yet to be settled [20–22].

3.5.2 Eisenfeld Qualitative Stability

Eisenfeld et al also studied in [35] qualitative stability, that is, strictly from the sign patterns. The results could be easily derived as well from the above analysis using the qualitative equivalence terminology. For instance

Lemma 3.21 *If the loop structure of the $n-$dimensional system contains a composite loop of length n and all such composite loops have same sign then the Jacobian is nondegenerate, i.e., $det(J) \neq 0$.*

See details in [35]. Lemma that entails

Theorem 3.22 (Eisenfeld stability)
Assume the loop structure has all composite $n-$loop of same sign. Then a necessary and sufficient condition for a stable sign equivalence class is that the loop structure has the following features

1. *There is at least a $1-$loop.*
2. *There are no positive $1-$loops*
3. *There are no positive $2-$loops*
4. *There are no $k-$loops with $k \geq 3$.*

And leads to

Corollary 3.23 *A necessary condition of the sign equivalence class to be unstable is that its loop structure contains no more than* $(n-1)$ *negative* $1-$*loops.*

For a system to undergo a Hopf bifurcation, ensuring the existence of limit cycles, its Jacobian must admit an exchange of stabilities. Therefore we have

Corollary 3.24 *A Hopf bifurcation requires a sign equivalence class that is neither stable nor unstable.*

Remark 3.25

1. An example of a sign equivalence class neither stable nor unstable is one associated with a loop structure containing n negative 1-loop $L^{-}_{x_i}, i = 1, \cdots, n$ i.e., at each vertex and at least one $k-$loop with $k \geq 3$.
2. Intuitively the negative loop may be seen as stabilizing whereas the positive loop could be seen as destabilizing.

In terms of composite loops and their sign as defined in the previous sections, we state

Theorem 3.26 *A sufficient condition for an unstable sign equivalence class (qualitatively unstable matrix) is that in the loop structure there is at least one integer k,* $(1 \leq k \leq n)$ *such that either there is no composite* $k-$ *loop or they are all positive (strong instability.)*

See details and proof in [35].

3.5.3 Loop Analysis in the Plane

Consider a square matrix A of order 2 given by

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \tag{3.13}$$

with entries constant with respect to the state variables, possibly depending on some parameters. The above stability criteria translate into

$$\begin{aligned} Tr(A) &= a_{11} + a_{22} \leq 0 \\ |A| &= a_{11}a_{22} - a_{12}a_{21}. \end{aligned} \tag{3.14}$$

We obtain the following loop interpretation. There is a parabolic boundary line at $Tr(A)^2 - 4|A| = 0$ between real and complex eigenvalues. Complex eigenvalues, i.e., oscillations are possible only when $a_{12}a_{21} < 0$, which corresponds to a negative 2-loop. Crossing the boundaries of the second quadrant is the fundamental way to lose stability, leading to the appearance of a positive feedback loop rendering positive the real part of the eigenvalues. Thereby the stable periodicity is destabi-

lized into a limit cycle for complex eigenvalues, or promoting multistationarity, i.e., saddle point for real eigenvalues.

Oscillations require the necessary condition $Tr(A)^2 - 4|A| < 0$, that is, $(a_{11} - a_{22})^2 < -4a_{12}a_{21}$. Therefore a negative 2-loop L^-_{xy} is necessary for any periodic behavior (center, stable focus, or limit cycle).

For a 2-dimensional system, we also have the following results.

Theorem 3.27 *Any loop structure in the plane consisting of two 1-loop L_1 of opposite signs and a negative 2-loop cannot be loop stable or loop unstable.*

The proof is based on the following lemma:

Lemma 3.28 *The qualitative equivalence class given by*

$$\langle\rangle = \begin{pmatrix} + & - \\ + & - \end{pmatrix}, \tag{3.15}$$

cannot be stable and cannot be unstable, that is, there is a matrix $A \in \langle\rangle$ such A is unstable or stable.

Based on the Linear Stability Theory, we can classify the loop equivalence classes in the plane as follows. See also [36].

Theorem 3.29

1. *The loop structure consists of only a negative 2-loop L^-_2: The dynamic is that of a linear center, that is, a family of periodic orbits surrounding the origin.*
2. *The loop structure $\mathbb{L}$ consists of a negative 2-loop L^-_2 and two 1-loop of same sign; the dynamic is that of a focus point, that is, the presence of sustained oscillations around the origin. If the system is bounded then there exists a limit cycle surrounding the origin.*
3. *The loop structure consists of only two 1-loops of opposite sign (resp. same sign). Then the origin is a saddle point (resp. a node stable for negative 1-loops, and unstable for positive 1-loops).*
4. *These loop equivalence classes are the only loop equivalence classes in the plane.*

Therefore we have formally, as indicated above.

Corollary 3.30 *A negative 2-loop L^-_2 is a necessary condition for any periodic behavior such as a center, a focus, or a limit cycle, i.e., an isolated periodic orbit.*

3.5.4 Biochemical Application: Two-Component Oscillators

Consider a two-component network of a chemical reaction system given by

$$\begin{aligned} \dot{x}_1 &= f_1(x_1, x_2) \\ \dot{x}_2 &= f_2(x_1, x_2), \end{aligned} \tag{3.16}$$

possibly with some dependence of f_1 and f_2 on the kinetic parameters. Recall that the Bendixon's negative criterion claims that if the divergence $div(f_1, f_2)$ is of constant sign in a region of the plane, then there can be no periodic solution in that region. For chemical reaction systems the diagonal entries of the Jacobian matrix are usually negative. If both a_{11} and a_{22} are always negative, that is, the existence of a constant negative $1-$loop in the loop structure, then the trace never changes sign, and Hopf bifurcation cannot occur in such a system. So at least one of them must be positive, indicating autocatalysis. With the diagonal elements of opposite sign, in order to have the determinant positive, the off-diagonal elements must also be of opposite sign. The typical sign patterns for a Hopf bifurcation are given by the following.

$$\langle\rangle = \begin{pmatrix} + & + \\ - & - \end{pmatrix}, \tag{3.17}$$

representing the so-called substrate-depletion oscillator. [Tyson]. The production of x_1 is autocatalytic, and the reaction speeds up as x_1 increases, until the substrate x_2 is depleted to the extend that the reaction ceases. The matrix

$$\langle\rangle = \begin{pmatrix} + & - \\ + & - \end{pmatrix}, \tag{3.18}$$

corresponds to the *activator-inhibitor* models. Intuitively, when x_2 is rare, x_1 increases autocatalytically. The degradation of the x_2 is inhibited with their accumulation stimulated by abundant x_1, which feeds back to inhibit the production of x_1. After x_1 disappears, x_2 is also destroyed, and then x_1 can make a comeback.

Therefore a two-component biochemical reaction system can oscillate if there exists in its loop structure at least one positive $1-$ loop, that is, autocatalysis along with a negative $2-$ loop. *Autocalysis* represented by $a_{ii} > 0$ has a major role in biochemical oscillations, and usually occurs when a chemical decelerates the rate of its own destruction. $a_{ij} > 0$ together with $a_{ji} > 0$ indicates the x_i activates the production of x_j and vice versa, leading to a *feedback loop* generating an indirect autocatalysis. When $a_{ij}a_{ji} < 0$, there exist a negative $2-$ feedback loop indicating that x_i activates the production of x_j but x_j inhibits the production of x_i [21].

3.5.5 Two-dimensional Model for Electrochemical Corrosion

We present how the Jacobian feedback loop methodology is applied to the electrochemical corrosion model initially developed by Talbot and Oriani. See also [22]. That is, a metal M is dissolving in an electrolyte solution in such a way that any given point of the metal surface at any given time is either bare or covered with adsorbed MOH to passivate the underlying metal. The model system reproduces the dynamics observed during potentiostatic dissolution of copper in an acetate buffer.

In terms of dimensionless variables the system is given by

$$\begin{aligned} \dot{x} &= p(1-y) - qx \\ \dot{y} &= x(1-y) - ye^{-\beta y} \end{aligned} \tag{3.19}$$

with the state variables confined to interval [0, 1], for the positive parameters values p, q, and β. In the region of steady states, the Jacobian matrix is

$$J = \begin{pmatrix} -q & -p \\ 1-y & a_{22} \end{pmatrix}, \tag{3.20}$$

where a_{22} is a function of the parameters p, q, and β given by $a_{22} = \frac{p(y-1)}{qy}(\beta y^2 - \beta y + 1)$. The associated qualitative matrix is

$$J = \begin{pmatrix} - & - \\ + & sign(a_{22}) \end{pmatrix}.$$

Therefore the loop structure has a negative 1-loop L_x^- and a negative 2-loop L_{xy}^-, satisfying the necessary condition for periodic behavior. Moreover the only way to secure a positive loop to promote multistationarity and sustained oscillations is to get $a_{22} > 0$. This is realized for $\beta > 4$ and $y \in [Y_1, Y_2]$ with $Y_{1,2} = \frac{1}{2}(1 \mp \sqrt{1 - \frac{4}{\beta}})$. This implies there exists a region in the parameter space with a positive loop necessary for multistationarity. There is also the possibility of sustained oscillations due to the presence of the negative 2-loop L_{xy}^-.

Therefore the loop analysis allows to predict a region of coexistence of multistationarity and sustained oscillations, actually limit cycle. See more in details in [22]. We note that the sign of the single Jacobian term a_{22} was crucial for both multistationarity and limit cycle. We were able to predict the global dynamics without resorting to actual integration of the system.

3.5.6 A Loop Analysis of the Lorenz System

The Lorenz system [19, 37] is described by

$$\begin{aligned} \dot{x} &= \sigma(y - x) \\ \dot{y} &= \rho x - y - xz \\ \dot{z} &= -\beta z + xy, \end{aligned} \tag{3.21}$$

where $(x, y, z) \in \mathbb{R}^3$, and $\sigma, \rho, \beta > 0$. The Lorenz equations can also be found in simplified models for lasers, dynamos, electric circuits, chemical reactions. The variable x is proportional to the intensity of convective motion, y is proportional to the temperature difference between ascending and descending currents and z is

proportional to the distortion from linearity of the vertical temperature profile. The Lorenz flow is dissipative, contracting volume with a negative divergence $-(1+\beta+\sigma)$. The Jacobian matrix at a steady state $p* = (x*, y*, z*)$ is

$$A = \begin{pmatrix} -\sigma & \sigma & 0 \\ \rho - z* & -1 & -x* \\ y* & x* & -\beta \end{pmatrix}, \tag{3.22}$$

with the corresponding qualitative Jacobian

$$A_q = \begin{pmatrix} - & + & 0 \\ sgn(\rho - z*) & - & -sgn(x*) \\ sgn(y*) & sgn(x*) & - \end{pmatrix}, \tag{3.23}$$

The main characteristics of the Lorenz system are

1. The system is invariant under reflection in the z-axis.
2. Its equilibria are the origin $O = (0, 0, 0)$ for all values of the parameter ρ, and $E_{\pm} = (\pm\sqrt{\beta(\rho-1)}, \pm\sqrt{\beta(\rho-1)}, \rho - 1)$ appearing at $\rho > 1$ and symmetric with respect to the z-axis.

A complete loop interpretation of this most celebrated system is found in [16]. First the stability parameters given by the k-order feedback loops lead to

$$\begin{aligned} \mathscr{F}_1 &= -(\sigma + \beta + 1) < 0 \\ \mathscr{F}_2 &= \beta + \sigma(1+\beta) - \sigma\rho + \sigma z* + x* \\ \mathscr{F}_3 &= -\sigma(\beta(1-\rho) + \beta z* + x* y* + x*) \end{aligned} \tag{3.24}$$

Hence we have

1. $\mathscr{F}_1$ is always negative, ensuring the constant presence of at least one negative 1-loop $L^{-}_{x/y/z}$.
2. $\mathscr{F}_2 = 0$ for $\rho = \rho_c = \frac{\beta+\sigma(1+\beta)-\sigma\rho+\sigma z*+x*}{\sigma}$. This entails $\mathscr{F}_2 > 0$ for $\rho < \rho_c$ to promote stability and $\mathscr{F}_2 > 0$ for $\rho < \rho_c$ to promote instability.
3. $\mathscr{F}_3 = 0$ for $\rho = \rho_1 = \frac{\sigma(\beta(1-\rho)+\beta z*+x*y*+x*}{\sigma}\beta$

3.5.7 A Loop Analysis of the Rossler System

The Rossler system [38] is given by

$$\begin{aligned} \dot{x} &= -y - z \\ \dot{y} &= x + ay \\ \dot{z} &= b + xz - cz \end{aligned} \tag{3.25}$$

The system has two steady states or equilibria $E\pm$ located at

$$(x\pm, y\pm, z\pm) = (\frac{c\sqrt{c^2-4ab}}{2}, -\frac{c\sqrt{c^2-4ab}}{2a}, \frac{c\sqrt{c^2-4ab}}{2a}) \tag{3.26}$$

associated with the Jacobian matrix

$$\begin{pmatrix} 0 & -1 & -1 \\ 1 & a & 0 \\ z & 0 & -c \end{pmatrix}, \tag{3.27}$$

and its qualitative equivalence class

$$\begin{pmatrix} 0 & - & - \\ + & + & 0 \\ + & 0 & - \end{pmatrix}. \tag{3.28}$$

For a wide range of the parameters a, b, c the system exhibits two unstable equilibria of type *saddle-focus* periodically repulsive (resp. attractive) in a plane while attractive (resp. repulsive) along a normal direction.

The loop structure contains two negative 2-loops L^-_{xy} and L^-_{xz} along with two 1-loops of opposite signs L^+_y and L^-_z. Note, as conjectured, the presence of two proper 3-dimensional composite loops $\mathscr{L}^2_3$ given by the negative (L^+_y, L^-_{xz}) and the positive (L^-_z, L^-_{xy}).

A complete loop analysis as in the Lorenz case above leads to predict the chaotic behavior of the system.

3.6 Summary

Given its fundamental qualitative nature, the methodology of Jacobian feedback loops allows only necessary conditions, not sufficient ones. However such necessary conditions are powerful enough to predict the effect to different structural or parameter changes. Indeed changes in the equations leading to a violation of any necessary conditions in terms of loops should yield the change as well in the dynamic behavior. And changes in the equations which do not affect the feedback loop structure of the Jacobian should preserve the dynamic behavior.

The above definitions, properties and theorems are the fundamentals of the theory of the Jacobian Loops analysis. They show that Jacobian loops and their combinations play an important dynamical role in a system, even when only the signs, not the magnitudes of the Jacobian terms, are known. Using stability analysis one can attain only the local dynamics of the system, and hence the need to use, for instance, numerical integration in conjunction to obtain global dynamics which

can be predicted by the feedback loop methodology, as for the Lorenz and Rossler systems. This is a great advantage over the classical approach. Indeed it allows one to assert whether sustained oscillations, multistationarity, or chaotic dynamics are possible. As such this analysis is certainly an efficient tool in the qualitative modeling of complex systems. It allows to:

1. Stress qualitative understanding as the primary goal rather than numerical prediction.
2. Supplement the more familiar large scale quantitative methods made possible by improved computer technology.
3. Include variables difficult or even impossible to measure, e.g., a diabetes model should include measurable variables such as glucose, insulin and other chemicals but also "real variables" such as anxiety or stress but any attempt to measure stress is itself stress inducing.

In economics [39, 40], behavioral and social sciences, as well as in complex physical sciences relevant informations about the underlying dynamics reside in the rules of construct of the system and not in the absolute values.

The Jacobian loops technique, easy to implement, intends to quickly demarcate both parameter and phase spaces into exciting regions (limit cycles, multiple equilibria, chaotic behavior), non-exciting regions (single stable fixed point), and hard-instance regions (ergodic behavior). Hence it could prove useful in surveying dynamical response of models simulating physico-chemical, biological and bio-chemical, and economical systems, as well as in game theory.

3.6.1 Research Directions

Here are some directions to improve the effectiveness of the qualitative analysis of systems based on the Jacobian feedback loops. Recall a loop is determined by a set of nonzero terms $a_{ij} = \frac{\partial f_i}{\partial x_j}$ of the Jacobian matrix whose i (row) and j (column) indices are in cyclic permutation. Its oriented edges (arrows) are the a_{ij} elements considered with their signs to indicate positive, negative or no interaction. A loop is usually symbolized by the product of its elements: for example, a 3-loop L_{xyz} is given by $a_{12}a_{23}a_{31}$. A loop is positive or negative depending on the sign of this product, this is, depending on whether it comprises an even or an odd number of negative elements. A positive feedback loop is destabilizing, whereas a negative feedback loop is stabilizing. For instance a minimal requirement for oscillations is the existence of at least one positive and a negative feedback loop, e.g. in chemical reactions systems [41, 42].

1. We want to emphasize that qualitative modeling should consist of strictly qualitative relations and assumptions, as opposed to studying the qualitative structure of models consisting of some quantitative features which actually describes the classic qualitative analysis.

2. In qualitative modeling the rate of change $\dot{x}$ should be defined as a qualitative rate of change $\frac{dx}{dt}$, that is, rather than just taking the sign of the entries of the Jacobian, one should consider directly the qualitative interaction per se of the variables $x = (x_i, i = 1, \cdots, n)$.
3. A qualitative algebraic structure should be developed independently, not just by "signing" the current algebras. The study of the qualitative and loop equivalence classes is a first step.
4. There is a need of strong qualitative or loop stability theorem similar to the Lyapunov's theorem; this could be done only through a direct qualitative study, rather than translating from the signs of quantitative values. That is, proving the existence of a Lyapunov function using only qualitative properties and relations of the variables and/or parameters interactions.
5. Complexity is the keyword in the evolution of systems. And its main tool of analysis is a qualitative one, which could achieve a greater generality and realism than does the usual quantitative idealization of most mathematical models.

References

1. Graham, Loren and Kantor, Jean-Michel, A Comparison of Two Cultural Approaches to Mathematics: France and Russia, 1890–1930. Isis, Vol 97(1) (2006), 56–74.
2. Weyl, Hermann, God and the universe: the open world. New Haven, Conn.: yale University Press, 1932
3. Lebesgue, Henri, "Préface" in Nicolas Lusin *Lecons sue les ensembles analytiques et leurs applications* Paris Gauthier-Villars, 1930.
4. von Neumann, J., The Computer and the Brain. Yale U.P., New Haven, 1958
5. Ruelle, David, Post-Human Maathematics. arXiv: 1308.4678v[math.HO] 21 Aug 2013.
6. Justus, J. *Loop Analysis and Qualitative Modeling: Limitations and Merits. Biology and Philosophy* 21 (2006): 647–666.
7. Levins, R., 1974 *The qualitative analysis of partially specified systems, Ann.N.Y.Acad.Sci.,* 231, 123–138
8. Ehrlich, Philip, The Rise on non-Archimedean Mathematics and the Roots of a misconception I: The Emergence of non-Archimedean Systems of Magnitudes. Arch.Hist.Exact Sci. 60 (2006) 1-121
9. Volovich I. V., p-adic string, Classical Quantum Gravity 4(4), L83–L87 (1987).
10. Dragovich B., Khrennikov A. Yu., Kozyrev S. V., Volovich I. V., Zelenov E. I., p-Adic mathematical physics: the first 30 years, p-Adic Numbers Ultrametric Anal. Appl. 9 (2017), no. 2, 87–121.
11. Murtagh Fionn, Thinking ultrametrically, thinking p-adically. Clusters, orders, and trees: methods and applications, 249–272, Springer Optim. Appl., 92, Springer, New York, 2014.
12. Frauenfelder H, Chan S. S., Chan W. S. (eds), The Physics of Proteins. Springer-Verlag, 2010.
13. Vladimirov V. S., Volovich I. V., Zelenov E. I., p-adic analysis and mathematical physics. Series on Soviet and East European Mathematics, 1. World Scientific Publishing Co., Inc., River Edge, NJ, 1994.
14. Khrennikov Andrei, Non-Archimedean analysis: quantum paradoxes, dynamical systems and biological models. Mathematics and its Applications, 427. Kluwer Academic Publishers, Dordrecht, 1997.
15. Khrennikov, A., Information Dynamics in Cognitive, Psychological and social Phenomena (Fundamental Theories of Physics), Kluwer, Dordrecht, 2004.

16. Toni, B., Dynamical Roles of Jacobian Feedback Loops and Qualitative Modeling, Springer PROMS vol 90(2014), 205–240
17. Thomas, R., 1994 *The role of Feedback Circuits: positive feedback circuits are a necessary condition for positive eigenvalues in the feedback matrix*, *Ber. Bunzenges. Phys. Chem. vol. 98 1148–1151*
18. Thomas, R., 1996 *Analyse et synthèse de systèmes à dynamique chaotique en terme de loops de rétroaction Académie Royale de Belgique* 6^e *série,* Tome VII, 101–124
19. Toni, B., Thieffry, D., Bulajich, R., *Feedback Loops analysis for chaotic dynamics with an application to Lorenz system*, *Fields Inst.Commun.* 21,473–483 (1999)
20. Toni, B., *Jacobian Feedback Loops Analysis I. International Journal of Evolution Equations,* 1(4), 415–428 (2005)
21. Toni, B. *Jacobian Feedback Loops Analysis II: Stabilty and Instability*, *International Journal of Evolution Equations,* 2(4), 355–366 (2008)
22. Toni, B., Parmananda, P., Bulajich, R., and Thieffry, D., *Dynamics of a two-dimensional Modelfor Electrochemical Corrosion Using Feedback Circuit and Nullcline Analysis, J.Phys.Chem.B.* 1998, 102, 4118–4122.
23. Verhulst, F., 1990 *Nonlinear Differential Equations and Dynamical Systems*, Springer-Verlag Universitext
24. Bellman, R., 1997 *Introduction to Matrix Analysis*, *Classics in Applied Matematics* 19, 2nd Ed. SIAM
25. Puccia, Charles, and Richard Levins (1985), *Qualitative Modeling of Complex Systems. Cambridge, MA: Harvard University Press.*
26. Kaufman, M., Soulé, C., Thomas, R., *A new necessary condition on interaction graphs for multistationarity Journal of Theoretical Biology, (2007) 248, 675–685*
27. Snoussi, E.H., 1998 *Necessary condition for Multistationarity and stable periodicity J.Biol.Syst.*, 6, 3–9
28. Soulé, C., 2003. *Graphic requirements for multistationarity. ComplexUs* 1, 123–133.
29. Thomas, R. and D'Ari, R. *Biological Feedback CRC Press Inc., Boca Raton*, 1990
30. Thomas, R., Thieffry, D. and Kaufman, M., 1995 *Dynamical behaviour of biological regulatory networks. I. Biological role of feedback loops and practical use of the concept of the loop-characteristic state Bull. Math. Biol.* 57, 247–276
31. Cinquin, P., Demongeot, J., 2002. *Positive and negative feedback: striking a balance between necessary antagonists. J. Theor. Biol.* 216, 229–241.
32. Gouzé, J.L., 1998 *Positive and negative circuits in Dynamical Systems*, *J. Biol. Syst.* Vol. 6 pp. 11–15
33. Plahte, E. Mestl, T. and Omholt, W., 1995, *Feedback Loop, Stability and Multistationarity in Dynamical Systems, J.Biol.Syst.,*3(2), 409–413
34. Demongeot, J., Kaufman, M., Thomas, R., 2000. *Positive feedback circuits and memory. C.R.Acad.Sci. Paris Life Sci.* 323, 69–79.
35. Eisenfeld, J. and De Lisi, C., 1994 *On conditions for qualitative instability of regulatory loops with applications to immunoqualitative control loops*, *Mathematics and Computers in Biomathematical Applications*, Elsevier, New York pp. 39–53
36. Kirkland, S.J., McDonald, J.J., Tsatsomeros, M.J., *Sign-patterns which require a positive eigenvalue, Linear and Multilinear Algebra* 41(3), 199–210, 1996.
37. Lorenz E. N., *Deterministic nonperiodic flow, J. Atmos. Sci.* 20 (1963), 130–141.
38. Rossler, O. E. *Continuous Chaos – Four prototype equations, Ann. N.Y. Acad. Sci.* (1979), 316, 376–392.
39. Quirk, R. and Ruppert, R., 1965 *Qualitative economics and the stability of equilibrium, Rev. Econ. Studies,* 32, 311–326
40. Samuelson, P., *Foundations of economic analysis. Harvard University Press, Cambridge*, 1947.
41. Sensse, A., Hauser, M., Eiswirth, M., *Feedback loops for Shil'nikov chaos: The peroxidase-oxidase reaction*, *J.Chem.Phys.* 125, 014901–12 (2006)
42. Tyson, J., 1975 *Classification of Instabilities in Chemical reaction systems, J. Chem. Phys.* 62, 1010–1015

43. Kochubei Anatoly N., Pseudo-differential equations and stochastics over non-Archimedean fields. Marcel Dekker, Inc., New York, 2001.
44. Abdelmalek Abdesselam, Ajay Chandra and Gianluca Guadagni, Rigorous quantum field theory functional integrals over the p-adics I: anomalous dimensions. arXiv:1302.5971.
45. Soulé, C., *Mathematical approaches to differentiation and gene regulation*, *C.R.Biologies* 329 (2006) 13–20.