AIMS Mathematics, 9(5): 10511–10535.
DOI: 10.3934/math.2024514
Received: 07 January 2024
Revised: 25 February 2024
Accepted: 05 March 2024
Published: 18 March 2024

http://www.aimspress.com/journal/Math

Research article

Assessing the impact of human behavior towards preventative measures on COVID-19 dynamics for Gauteng, South Africa: a simulation and forecasting approach

CW Chukwu[1,*], S. Y. Tchoumi[2,3], Z. Chazuka[4], M. L. Juga[5] and G. Obaido[6]

[1] Department of Mathematics, Wake Forest University, Winston-Salem, NC 27109, USA

[2] Department of Mathematics and Computer Sciences ENSAI, University of NGaoundere, Cameroon

[3] Department of Mathematics and Applied Mathematics, University of Pretoria, South Africa

[4] Department of Decision Sciences, College of Economics and Management Sciences, University of South Africa

[5] Department of Mathematics and Applied Mathematics, University of Johannesburg, South Africa.

[6] Center for Human Artificial Intelligence (CHAI), Berkeley Institute for Data Science (BIDS), University of California, Berkeley, CA 94720, USA

* **Correspondence:** Email: wiliam.chukwu@gmail.com.

Abstract: Globally, the COVID-19 pandemic has claimed millions of lives. In this study, we develop a mathematical model to investigate the impact of human behavior on the dynamics of COVID-19 infection in South Africa. Specifically, our model examined the effects of positive versus negative human behavior. We parameterize the model using data from the COVID-19 fifth wave of Gauteng province, South Africa, from May 01, 2022, to July 23, 2022. To forecast new cases of COVID-19 infections, we compared three forecasting methods: exponential smoothing (ETS), long short-term memory (LSTM), and gated recurrent units (GRUs), using the dataset. Results from the time series analysis showed that the LSTM model has better performance and is well-suited for predicting the dynamics of COVID-19 compared to the other models. Sensitivity analysis and numerical simulations were also performed, revealing that noncompliant infected individuals contribute more to new infections than those who comply. It is envisaged that the insights from this work can better inform public health policy and enable better projections of disease spread.

Keywords: human behavior; COVID-19; compliance; noncompliance; simulation; time series
Mathematics Subject Classification: 34A08, 92B05

1. Introduction

The coronavirus disease (COVID-19), first discovered in Wuhan, China, in 2019 [1], has been a worldwide pandemic since its emergence and has resulted in the loss of millions of lives. This infectious disease, caused by the pathogen coronavirus, SARS-CoV-2, is transmitted through contact with droplets from the respiratory tract of an infected person. Consequently, slowing down the transmission of COVID-19 requires large-scale behavioral change [2]. Such changes can influence the disease status of individuals, such as the infection rate, the recovery rate, and the contact network structure associated with the spread of the disease [3]. Generally, stopping a disease outbreak relies on the technical implementation of control strategies, such as vaccination, treatment, or educational campaigns. One key objective of these controls is to induce a change in human behavior that will either halt or eradicate the infection spread.

Human behavior has been critical in shaping the COVID-19 pandemic, with the actions of individuals, groups, nation-states, and international bodies all playing roles in curbing its spread. Although people in specific geographic settings are expected to behave in certain ways to help curb the spread of COVID-19, not all individuals adhere to these expected preventative measures. Some individuals neglect to wear face masks in public places, and others disregard social distancing, hand sanitization, or covering their mouths and nose while coughing and sneezing. Additionally, due to misinformation about vaccines [4] and healthcare facilities, particularly in Africa, individuals are reluctant to visit vaccination centers or seek hospital treatment when experiencing symptoms of the infection. Misinformation has also led to the use of informal or nonintegrated forms of healthcare, such as traditional and complementary medicine (TCM) [5]. Furthermore, the fear of stigmatization during or after infection can drive individuals to conceal their illness and avoid formal healthcare assistance. Alcohol abuse is also another significant factor. Alcohol has a suppressive effect on the immune system, particularly impairing the lungs' ability to fight off infectious diseases like COVID-19 [6]. A weakened immune system not only increases the risk of contracting the virus but may also exacerbate the severity of the illness and potentially slow recovery.

Normally, where treatment is available, health workers expect patients to behave in a prescribed manner, such as consistently taking medication. In the case of COVID-19, individuals are expected to wear masks in public places, maintain personal hygiene, comply with vaccination recommendations, seek medical help promptly when experiencing symptoms, remain quarantined for at least 14 days after infection, frequently clean their hands with an alcohol-based hand rub or soap and water, and cover their mouths and noses when coughing or sneezing [7]. Due to variations in human behavior, many of these control measures are not uniformly adhered to in communities affected by the disease. Poor adherence to control measures during a pandemic significantly affects transmission dynamics, particularly with diseases like COVID-19 with high inactivity and case fatality rates [8]. Therefore, human behavior is at the core of any disease eradication strategy [9].

Mathematical modeling has proven instrumental in examining the influence of human behavior on the evolution and dynamics of various diseases, as evidenced by the following literature: [3, 10–17]. However, quantifying human behavior is very challenging, which complicates the development and analysis of mathematical models that include human behavior [3]. This viewpoint is shared by Funk, et al. [10], who outlined some challenges faced in integrating the dynamics of human behavior into infectious disease models. These challenges include setting a baseline and establishing the effects of

deviating from it, determining the manner and extent to which human behavior should be explicitly modeled, ascertaining the minimum level of detail needed to model behavioral differences, estimating changes in reporting behavior, forecasting responses to interventions and health campaigns, identifying the role of movement and travel, constructing models that can be validated against real-time data, and facilitating communication across disciplines [10].

Zhang et al. [18] explored how the COVID-19 pandemic influenced human behaviors in Hong Kong and the subsequent impact on the transmission of other respiratory diseases (e.g., influenza). Their work focused on the spread of COVID-19 and influenza infections using COVID-19-reported cases and influenza surveillance data. They investigated changes in human behavior using mass transit railway data and telephone survey results. The study indicated that human behavior was significantly influenced by the COVID-19 pandemic in Hong Kong and that close contact control contributed to more than 47% of the reduction in COVID-19 infection risk. Umakrishnan et al. [19] analyzed COVID-19 preventive behavior and the socio-economic drivers behind health responses in South Africa. Through employing various statistical and econometric techniques, they found that preventative behavior evolved, with the most widely adopted measure being effective mask usage, surpassing hand washing. Additionally, they observed that other measures like social distancing and avoiding large groups had declined over time. Furthermore, various other mathematical models have explored COVID-19 infection dynamics globally, predicting the importance of several control measures, including [20–32] among many others. The aforementioned human behaviors may have serious effects on the spread of COVID-19, which we investigate in this work. Using a mathematical model that incorporates individuals who exhibit positive behavior (compliance to all control measures) and those with negative behavior (noncompliance to some or all of the control measures), we evaluate the impact of various human behaviors on the spread and dynamics of COVID-19 within South Africa.

In addition to the analysis of the deterministic model for COVID-19 in Gauteng Province, we construct time series models, namely the exponential smoothing (ETS) [33], long short-term memory (LSTM) [34], and the gated recurrent unit network (GRU) [35]. The reason behind adopting this approach is to rigorously assess and compare the efficacy of different modeling techniques in predicting the daily new cases of COVID-19 in Gauteng Province. Also, our exploration extends beyond conventional forecasting methods by incorporating considerations of positive and negative behaviors in the mathematical model. By so doing, we aim to capture human dynamics and societal responses, thereby providing a more comprehensive understanding of the factors influencing the spread of COVID-19 in Gauteng. To achieve optimal predictive accuracy, the performance of the ETS, LSTM, and GRU models will be critically evaluated. To the best of our knowledge, we present the first COVID-19 model for Gauteng Province that assesses the impact of both positive and negative human behaviors, coupled with the comparison of ETS, LSTM, and GRU models. We envisage that the insights from this work can better inform public health policy and enable better projections of disease spread.

The rest of the paper is organized as follows: Section 2 is devoted to the model formulation and the analysis of the basic reproduction number. Sections 3 and 4 present the time series data forecasting and numerical simulations of the mathematical model, respectively. The paper is finalized in Section 5 with a conclusion and discussion of the results.

2. The mathematical model

2.1. Model description

This study uses a mathematical model with 9 compartments to assess the impact of human behavior on COVID-19 transmission dynamics. The population under investigation is split into the population of individuals who are of positive behavior (individuals who adhere to all the COVID-19 preventative measures such as social distancing, wearing masks in crowded places, the washing and sanitizing of hands, quarantining when experiencing symptoms, those on treatment, etc.). These individuals are assumed to willingly show up at vaccination centers to be vaccinated against COVID-19. They also show up at the hospital when they start having symptoms and follow treatment diligently if they are infected. The model also considers individuals who are of negative behavior (those who do not adhere to all or some of the preventive measures, refuse to be vaccinated for various reasons, hide their infection status when they start experiencing symptoms, and seek informal health care from TCM practitioners rather than formal healthcare providers).

Our compartmental model consists of two suppopulations described as follows: susceptible individuals with positive (negative) behavior, S_p (S_n); asymptomatic individuals with positive (negative) behavior, A_p (A_n); infectious individuals with positive (negative) behavior, I_p (I_n); hospitalized individuals, H, and recovered individuals with positive (negative) behavior, R_p (R_n). Those in I_n are assumed to be hiding their infection and seeking informal rather than formal health care. Additionally, they are also assumed to be the most infectious as compared to individuals in I_p. Those in I_p are willing to seek hospital care, and so they are assumed to be less infectious than those in I_n. A susceptible individual within the population is either recruited into the class, S_p of individuals who are susceptible with positive behavior at a constant rate of π_p, or into the class, S_n of individuals who are susceptible with negative behavior at a rate of, π_n. Individuals in S_p contact COVID-19 at an infection rate of λ_p where

$$\lambda_p = \beta(I_n + \eta_1 A_p + \eta_2 I_p + \eta_3 A_n + \eta_4 H),$$

and move into the class of asymptomatic individuals with positive behavior, A_p. The model assumes that a proportion, p, moves into A_p while the remainder moves into the infectious compartment with positive behavior, I_p, at a rate, $(1 - p)$. The modification parameters within the force of infection are such that $0 < \eta_1 < \eta_2 < \eta_3 < \eta_4 < 1$, and the effective contact rate is given by β. Those in the asymptomatic class, (A_p) may recover at a rate of $\rho_p\sigma_p$ while a fraction $(1 - \rho_p)\sigma_p$ may progress to the infectious compartment with positive behavior, I_p. Now, a proportion, $\kappa_p\gamma_p$ may recover naturally and proceed to class, R_p, while the remaining proportion, $(1 - \kappa_p)\gamma_p$, due to full symptoms of infection, may progress to the hospitalized class, H at a later stage. We note that the individuals in the class S_n contract the disease at a relatively higher infection rate since they are of negative behavior, and, hence, the force of infection is given by

$$\lambda_n = k\lambda_p,$$

and $k > 1$. Upon interaction with an infectious individual, those in class, S_n can become actively infectious at a rate of $q\lambda_n$ and move into class I_n, while the remainder $(1 - q)\lambda_n$ becomes asymptomatically infected and move into class, A_n. Asymptomatic individuals with negative behavior may fully develop COVID-19 symptoms and move into class I_n at a rate of, $(1 - \rho_n)\sigma_n$, while some may recover naturally at a rate $\rho_n\sigma_n$. During the course of the infection, those in class I_n may change

their minds and seek proper health care, and, hence, progress to the hospitalized class at a rate of $(1-\kappa_n)\gamma_n$, while some may recover naturally at a rate of $\kappa_n\gamma_n$ and progress to class R_n. The model also assumes that hospitalized individuals recover at a rate of α and move into the R_p class. Individuals from all compartments are assumed to die from natural causes at a rate of μ, while those in I_p, I_n, and H have disease-induced death rates given by δ_p, δ_n, and δ_h, respectively. Table 1 gives a summarized epidemiological interpretation of our model parameters, and the model flow diagram is depicted in Figure 1.

Table 1. Model parameters and their descriptions.

Parameter	Description
μ	Natural mortality rate
β	Effective contact rate
$k, \eta_1, \eta_2, \eta_3, \eta_4$	Modification parameters
$\pi_p(\pi_n)$	Rate of recruitment into $S_p(S_n)$
$\lambda_p(\lambda_n)$	Infection rates of people in $S_p(S_n)$
$\delta_p, \delta_n, \delta_h$	Disease-induced death rate in I_p, I_n, and H, respectively
$\kappa_p(\kappa_n)$	Proportion of I_p (I_n), respectively, who recover naturally
$\rho_p(\rho_n)$	Proportion of A_p (A_n), respectively, who recover naturally
ω_p (ω_n)	Rates at which R_p (R_n), respectively, becomes susceptible again
γ_p (γ_n)	Exit rate from infectious with positive (negative) behavior, respectively
σ_p (σ_n)	Exit rate from asymptomatic with positive (negative) behavior, respectively
p (q)	Proportion of $S_p(S_n)$ who become asymptomatic A_p (A_n), respectively

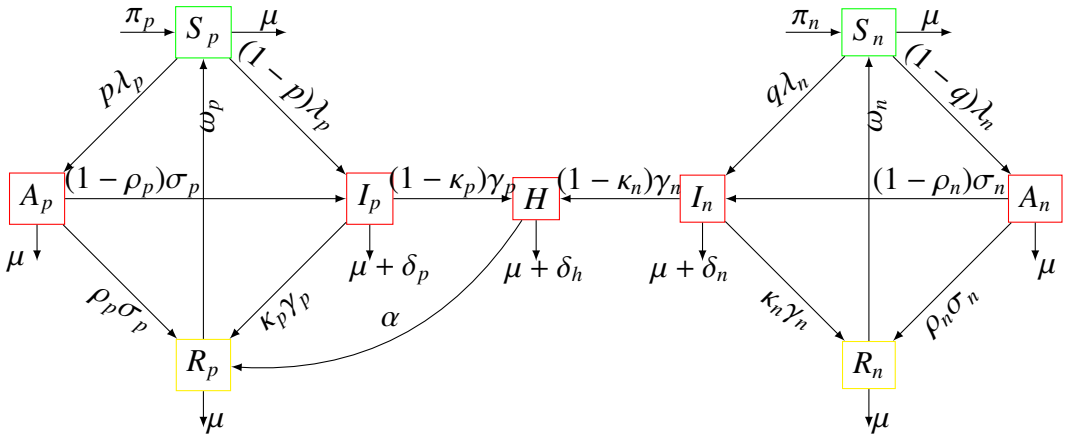


Figure 1. Model flow diagram for COVID-19 dynamics incorporating behavior change.

The model equations are given as

$$\begin{cases} S_p' &= \pi_p + \omega_p R_p - (\lambda_p + \mu) S_p, \\ A_p' &= p\lambda_p S_p - (\sigma_p + \mu) A_p, \\ I_p' &= (1-p)\lambda_p S_p + (1-\rho_p)\sigma_p A_p - (\gamma_p + \mu + \delta_p) I_p, \\ R_p' &= \rho_p A_p + \kappa_p I_p + \alpha H - (\omega_p + \mu) R_p, \\ S_n' &= \pi_n + \omega_n R_n - (\lambda_n + \mu) S_n, \\ A_n' &= q\lambda_n S_n - (\sigma_n + \mu) A_n, \\ I_n' &= (1-q)\lambda_n S_n + (1-\rho_n)\sigma_n A_n - (\gamma_n + \mu + \delta_n) I_n, \\ R_n' &= \rho_n A_n + \kappa_n I_n - (\omega_n + \mu) R_p, \\ H' &= (1-\kappa_p)\gamma_p I_p + (1-\kappa_n)\gamma_n I_n - (\mu + \alpha + \delta_p) H, \end{cases} \tag{2.1}$$

with positive initial conditions

$$S_p(0) > 0,\ S_n(0) > 0,\ I_p(0) \geq 0,\ I_n(0) \geq 0,\ A_p(0) \geq 0,\ A_n(0) \geq 0,\ H(0) \geq 0,\ R_p(0) \geq 0,\ R_n(0) \geq 0.$$

For the existence, uniqueness, and positivity of solutions, see Appendix 5.

2.2. Basic reproduction number

To compute the reproduction number, we first calculate the disease-free equilibrium point. This is found when we set system (2.1) to zero and solve it for the case where

$$S_p \neq 0, A_p = 0, I_p = 0, R_p = 0, S_n \neq 0, A_n = 0, I_n = 0, R_n = 0.$$

Therefore, this yields

$$\mathcal{E}_0 = (S_p^\star, 0, 0, 0, S_n^\star, 0, 0, 0, 0),$$

where $S_p^\star = \dfrac{\pi_p}{\mu}$ and $S_n^\star = \dfrac{\pi_n}{\mu}$.

The reproduction number is a critical threshold that measures how many secondary infections are produced by one infectious individual in a wholly susceptible population. To compute this threshold, we employ the next generation matrix approach by [36–38] and, therefore, $\mathcal{R}_0 = \mathcal{R}_p + \mathcal{R}_n$, where,

$$\mathcal{R}_p = \beta S_p^\star \left(p\frac{\eta_1}{Q_1} + (1-p)\frac{\eta_2}{Q_3} + p\frac{\eta_2 Q_2}{Q_1 Q_3} + (1-p)\frac{\eta_4 Q_7}{Q_3 Q_9} + p\frac{\eta_4 Q_2 Q_7}{Q_1 Q_3 Q_9} \right), \tag{2.2}$$

$$\mathcal{R}_n = k\beta S_n^\star \left(q\frac{\eta_3}{Q_4} + (1-q)\frac{1}{Q_6} + q\frac{Q_5}{Q_4 Q_6} + (1-q)\frac{\eta_4 Q_8}{Q_6 Q_9} + q\frac{\eta_4 Q_5 Q_8}{Q_4 Q_6 Q_9} \right) \tag{2.3}$$

and $Q_1 = \sigma_p + \mu$, $Q_2 = (1-\rho_p)\sigma_p$, $Q_3 = \gamma_p + \mu + \delta_p$, $Q_4 = \sigma_n + \mu$, $Q_5 = (1-\rho_n)\sigma_n$, $Q_6 = \gamma_n + \mu + \delta_n$, $Q_7 = (1-\kappa_p)\gamma_p$, $Q_8 = (1-\kappa_n)\gamma_n$ and $Q_9 = \alpha + \mu + \delta_h$.

The computed threshold quantity, $\mathcal{R}_0$, is a measure of the average number of new COVID-19 cases generated by a typical COVID-19 infected individual (living or dead but not buried) introduced into a purely susceptible human population [36–38]. The disease can thus be effectively controlled in the community if $\mathcal{R}_0$ can be decreased to (and sustained at) a value less than one (i.e., $\mathcal{R}_0 < 1$).

2.3. Interpretation of $\mathcal{R}_0$

When an infectious human is introduced into a population consisting only of susceptible individuals with positive behavior, individuals may acquire infection, and the probability of becoming asymptomatic is p, while that of being symptomatic is $(1-p)$. During the asymptomatic period, $\frac{1}{Q_1}$, an infectious individual can infect a susceptible individual with a probability of $p\beta\eta_1$. After this period, a fraction, $\frac{Q_2}{Q_1}$ will become symptomatic. During the period, $\frac{1}{Q_3}$, symptomatic individuals can infect a susceptible individual with a probability of $p\eta_2\beta$. Now, after this symptomatic period, a fraction, $\frac{Q_2Q_7}{Q_1Q_3}$ will become hospitalized, and during the hospitalization period, $\frac{1}{Q_9}$, they may infect a susceptible individual with a probability of, $p\eta_4\beta$. So, the total number of susceptible individuals infected by an asymptomatic individual will be,

$$p\beta S_p^*\left(\frac{\eta_1}{Q_1}+\eta_2\frac{Q_2}{Q_1Q_3}+\eta_4\frac{Q_2Q_7}{Q_1Q_3Q_9}\right).$$

On the other hand, if an individual is symptomatic, then during their symptomatic period $\frac{1}{Q_3}$, they will infect a susceptible individual with a probability of $(1-p)\eta_2\beta$. After the symptomatic period, a fraction, $\frac{Q_7}{Q_3}$ will become hospitalized, and during the hospitalization period $\frac{1}{Q_9}$, they can infect a susceptible individual with a probability of $(1-p)\eta_4\beta$. Thus, the total number of susceptible individuals infected by a symptomatic person will be

$$(1-p)\beta S_p^*\left(\frac{\eta_2}{Q_3}+\eta_4\frac{Q_7}{Q_3Q_9}\right).$$

Hence, the overall total number of susceptible individuals with positive behavior that will be infected is:

$$\mathcal{R}_p=\beta S_p^\star\left(p\frac{\eta_1}{Q_1}+(1-p)\frac{\eta_2}{Q_3}+p\frac{\eta_2Q_2}{Q_1Q_3}+(1-p)\frac{\eta_4Q_7}{Q_3Q_9}+p\frac{\eta_4Q_2Q_7}{Q_1Q_3Q_9}\right).$$

The same approach can be done for the susceptible individuals with negative behavior, and this will yield $\mathcal{R}_n$. Hence, we have $\mathcal{R}_0=\mathcal{R}_p+\mathcal{R}_n$.

3. Time series data analysis

Time series analysis is crucial for identifying trends in epidemics, such as COVID-19 [39–41]. This analysis is useful in providing valuable information that aids government and public health authorities in their preparedness and response strategies for current and future outbreaks. Some of the forecasting models chosen for this study are as follows: ETS, LSTM, and GRU.

3.1. Data overview

The dataset used for the simulations in this study was obtained from the COVID-19 dynamics in Gauteng, South Africa. The Gauteng Province has the highest population density in South Africa

compared to other provinces. We utilized the daily COVID-19 case data for the fifth wave in South Africa, from May 1, 2022, to July 23, 2022 [42]. The primary reason for choosing this specific dataset was its representation of the higher rate of COVID-19 transmission during the selected period compared to other provinces in the country, reflecting shifts in human behavioral patterns.

The descriptive statistics for daily confirmed cases are presented in Table 2, offering a comprehensive overview of the epidemiological trend throughout the study period, with 79 data points. As observed in Figure 2, the mean daily case count is 990.2, representing the average number of cases per day. However, the standard deviation of 1089.01 indicates significant variability in daily case numbers, reflecting substantial fluctuations over time. The positive skewness of 1.48 indicates an asymmetrical distribution of daily cases, with a tail extending toward higher case counts. This skewness suggests that, although most days recorded relatively lower numbers of new cases, there were occasional days with exceptionally high numbers of new cases, likely due to outbreaks or enhancements in testing and reporting measures.

Table 2. Descriptive statistics of daily COVID-19 cases in Gauteng, South Africa.

Statistics	Values
Count	79
Mean (Average)	990.2
Standard Deviation	1089.01
Minimum	0
25th Percentile (Q1)	205
Median (50th Percentile)	439
75th Percentile (Q3)	1352.5
Maximum	4332
Skewness	1.48

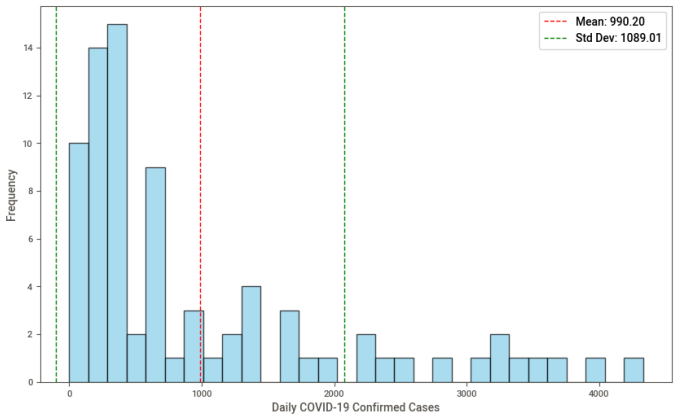


Figure 2. The histogram showing the daily COVID-19 cases in Gauteng, South Africa.

We observe considerable fluctuation in daily COVID-19 cases over time (Figure 3). There are peaks where the cases rise sharply, followed by troughs where the cases fall. The highest peaks appear to occur in early May, with the number of cases reaching the maximum observed on the graph. After the initial peaks, there is a downward trend with some variability, but the overall movement is toward

fewer daily cases as time progresses.

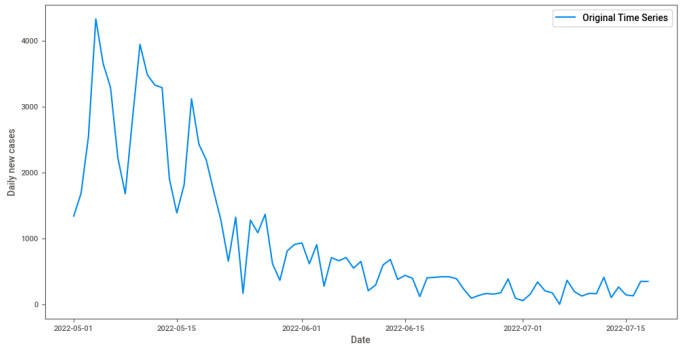


Figure 3. The sequence showing the trend of the daily COVID-19 cases in Gauteng over time.

3.2. Forecasting models

We applied three forecasting models, that is, the ETS, LSTM, and GRU, to forecast future COVID-19 cases. The models were trained and tested on the dataset using an 80/20 split, and future cases were forecasted to determine potential trends and patterns. We chose the models for their proven capabilities in handling the complexities of epidemic time series data. Each model offers unique advantages that is, the ETS for its simplicity and effectiveness in capturing trends, the LSTM for its ability to understand long-term dependencies, and the GRU for its efficiency in modeling time-dependent sequences.

3.2.1. ETS

ETS is a commonly employed technique in time series forecasting aimed at smoothing out noise in data and extracting underlying trends or patterns. Mathematically, ETS is defined as follows:

$$\hat{Y}_{t+1} = \alpha Y_t + (1 - \alpha)\hat{Y}_t \tag{3.1}$$

where: $\hat{Y}_{t+1}$ is the forecasted value for the next time period $t + 1$. Y_t is the observed value at time t. $\hat{Y}_t$ is the smoothed value (forecast) at time t. α is the smoothing parameter, often referred to as the smoothing factor or smoothing coefficient, with $0 < \alpha < 1$.

As shown in Figure 4, we applied the ETS model to forecast daily COVID-19 cases. This method is particularly suited for data with significant levels of trend and seasonality, as it assigns exponentially decreasing weights to past observations. However, the choice of the model's smoothing parameters is critical, as they determine how quickly the influence of previous observations decline. The training data, depicted in blue, demonstrates the initial volatile behavior of the epidemic, with a clear downward trend following the initial surge. The test data, in orange, is the subset of the data not used in the training of the model but is crucial for validating the model's predictive power. The forecast, represented by the green dashed line, projects a stabilization in the number of new cases, oscillating around a constant mean. This suggests that the model expects the incidence rate to plateau in the next 45 days, without significant peaks or troughs.

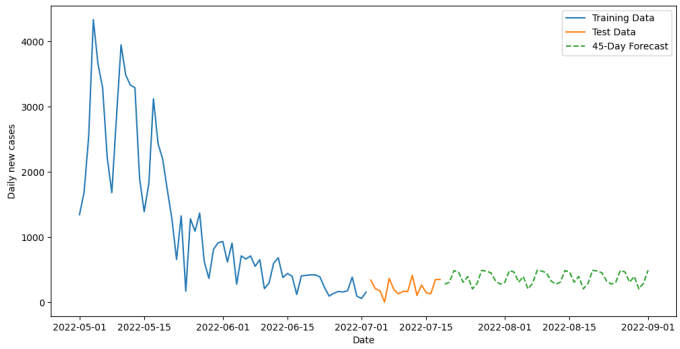


Figure 4. ETS result showing the training, testing, and 45-day forecast.

3.2.2. LSTM

The LSTM model is well-suited for time series prediction due to its ability to remember long-term dependencies [43, 44]. It is particularly beneficial when dealing with fluctuating data such as daily infection rates. The model was trained using data that was normalized to ensure effective learning. An Adam optimizer was employed for its efficiency in handling sparse gradients and adaptive learning rate capabilities, along with a mean squared error metric for the loss function to quantify the model's accuracy.

The model was fitted over 100 epochs, indicative of the number of complete passes through the training dataset, with a batch size of 1, ensuring that the model weights were updated after each sample. Also, the verbosity level of 2 was set ensuring that there was a per-epoch output to track the training progress. The training data, shown in blue, illustrates initial fluctuations in COVID-19 cases and a subsequent downward trend, as depicted in Figure 5. The test data, in orange, is used to evaluate the model's accuracy on new, unseen data and follows a similar downward pattern. The model's 45-day forecast, depicted by the green dashed line, suggests a stable trend of new cases in the near future, providing a basis for pandemic response planning.

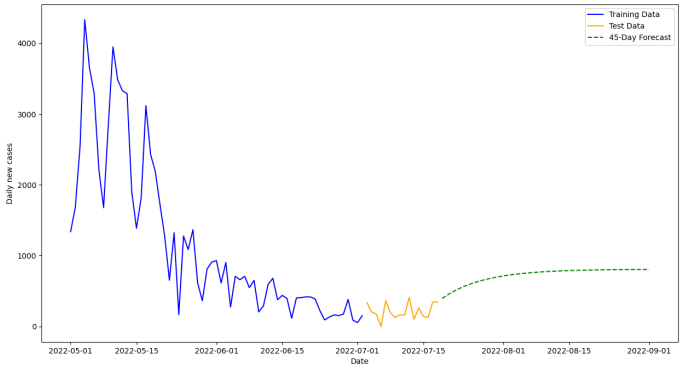


Figure 5. LSTM result showing the training, testing, and 45-day forecast.

3.2.3. GRU

Similar to the LSTM model, GRUs are a type of recurrent neural network that excel in capturing temporal dependencies in time series data. GRUs are known for their efficiency and simplicity as they use fewer parameters than LSTMs while achieving comparable performance, especially in cases

where the time series is not excessively long [45, 46]. A similar configuration to that used for the LSTM was employed for the GRU model to forecast daily COVID-19 cases. The GRU model, trained on data characterized by sharp initial fluctuations and a subsequent decline, has seemingly grasped the outbreak's dynamics. The test phase, essential for model validation, showed that the GRU's predictions mostly followed the trend set by the training data, albeit with some variability. The 45-day forecast by the GRU model, illustrated with a green dashed line, suggests a stable pattern in new cases with a slight uptick toward the end of the period, hinting at a possible rise in cases that could necessitate preemptive public health measures.

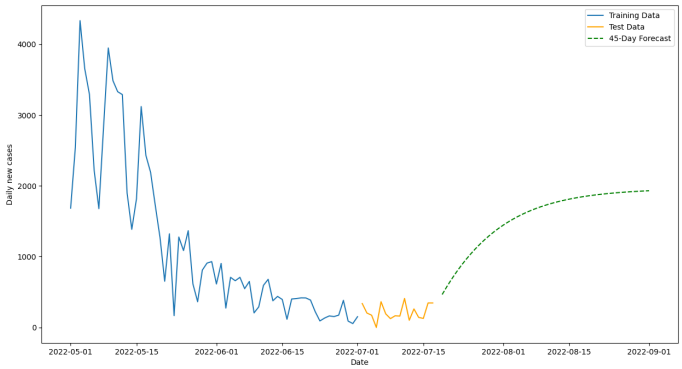


Figure 6. GRU result showing the training, testing, and 45-day forecast.

The analyses using the models revealed a trend toward short-term stabilization in the number of new COVID-19 cases, providing critical data for informed public health decision-making. The ETS model indicated a consistent trend with minimal variations, the LSTM model showed a continued decrease and a stable plateau of new cases, and the GRU model forecasted stability with a potential slight increase in cases as the forecast period concluded. These insights are instrumental for strategic planning in the public health domain.

3.3. Evaluation of forecasting models

The metrics used to evaluate model performance include Mean Average Error (MAE), the Mean Squared Error (MSE), and the Root Mean Squared Error (RMSE). The metrics MAE, MSE, and RMSE were calculated in each scenario by using Eqs (5.1)–(5.3) (see Appendix 5). The calculated result is summarized in Table 3, which reflects the metrics used to compare the predicted performance of the ETS, LSTM, and GRU models.

Table 3. Model performance evaluation.

Model	MAE	MSE	RMSE
ETS	183.19	50425.55	224.56
LSTM	156.66	31643.93	177.89
GRU	165.16	35078.70	187.29

The ETS model had an MAE of 183.19, which is the highest among the three models, indicating that, on average, the ETS model's forecasts were about 183.19 units off from the actual values. The MSE, which is the average squared difference between the estimated values and actual value,

was 50425.55, showing the highest among the three models, suggesting less accuracy in the predictions. The RMSE was 224.56, indicating that the typical deviation from the actual data was approximately 224.56 units. The LSTM model performed better on all three metrics, with an MAE of 156.66, indicating its forecasts were closer to the actual values on average. The MSE was significantly lower at 31643.93, suggesting greater accuracy, and the RMSE was 177.89, indicating a smaller spread in the forecast errors compared to the ETS model. The GRU model showed an intermediate performance between the ETS and LSTM models. The model reported an MAE of 165.16, an MSE of 35078.70, and an RMSE of 187.29. This suggests that the GRU model's predictions were less accurate than those of the LSTM model, but more accurate than those of the ETS model. The analysis underscores the LSTM model as the best-performing model, suggesting its potential for predicting trends in daily COVID-19 cases. The relative performance of these models highlights their importance in epidemiological forecasting, which can significantly influence the effectiveness of public health strategies. These insights are crucial for public health decision-making, enabling more useful projections of disease spread and facilitating timely and effective responses.

Our forecast predicts an increase in new daily infections. As a result, it is imperative to understand the likely causes of this increase. To gain deeper insights, we implement a mathematical model to understand the underlying dynamics driving the rise in infections and to identify potential intervention strategies. These results will indicate which risk factors should be targeted for effective disease control to reduce the number of people infected with COVID-19 in Gauteng.

4. Numerical simulations of the mathematical model

In this section, we present the results from the numerical simulation of the Model (2.1).

4.1. Model calibration

The epidemic has been ongoing in South Africa for over three years; hence, the model incorporates the demographic effects. Model (2.1) was fitted to the daily COVID-19 data from May 2022 to July 2022 for the Gauteng Province in South Africa [42], as shown in Figure 7, while Figure 8 gives the impact of increasing contact rates due to behavior change on the fitted model.

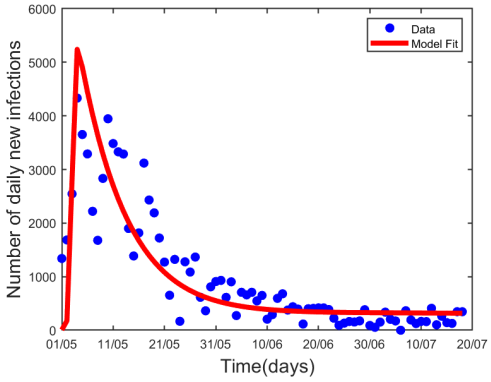


Figure 7. Model fitting for the Gauteng Province COVID-19 data from May 01, 2022 to July 23, 2022, using the fminsearch algorithm.

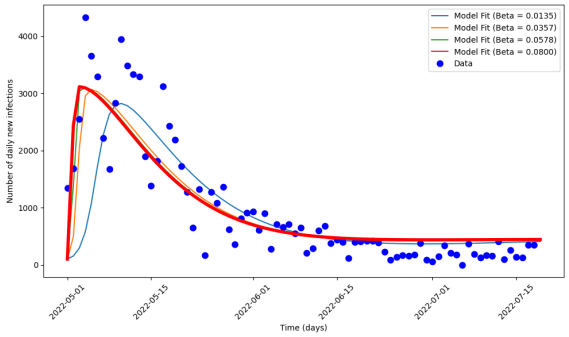


Figure 8. Model fitting for the Gauteng Province COVID-19 data using the fminsearch algorithm varying the infection rate $\beta = 0.0135, 0.0357, 0.0800$ varied while other parameters are kept constant as given in Table 4.

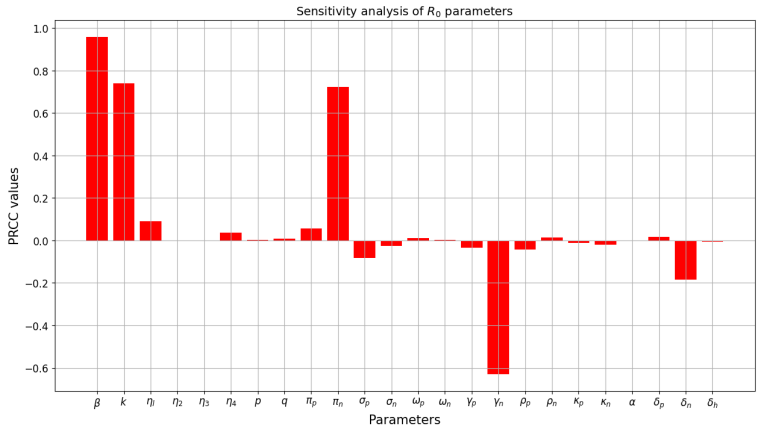
The model was calibrated in Matlab using the fminsearch algorithm. The data from [42] used for this calibration consists of clinical data collected after individuals tested positive for COVID-19. This model considers the infected individuals with negative behavior, I_n, as those who, because of stigma and other factors, never showed up at any hospital or testing center to be tested for COVID-19, but remained within the community or sought health care from TCM practitioners. However, it is important to note that there is no existing data for COVID-19-infected individuals with negative behavior. Therefore, we used the infectious with positive behavior, I_p, class to carry out our model fitting. It is observed from Figure 7 that the data fits well, and, hence, we estimated some of the parameters for the model. Table 4 gives the numerical values (or ranges) of the model parameters used in the simulations. While some parameter values were obtained from existing literature, others were fitted. The rest of the parameters associated with negative behavior are estimated similarly and presented in Table 4.

Table 4. Model parameters values used for numerical simulations obtained from literature and data fitting Figure 7.

Parameter	Value	Unit	Reference
π_p, π_n	$\frac{10000}{64 \times 365}$	day^{-1}	[47]
p, q	0.7	unit-less	[48]
γ_p,γ_n	0.0833	day^{-1}	[49]
σ_p,σ_n	0.13978	day^{-1}	[48,49]
κ_p,κ_n	0.05	day^{-1}	[50]
ρ_p,ρ_n	0.14	day^{-1}	[51]
α	0.0701	day^{-1}	[52]
μ	$\frac{1}{64 \times 365}$	day^{-1}	[53]
$\delta_p, \delta_n, \delta_h$	0.018	day^{-1}	[49,52]
ω_p, ω_n	0.011	day^{-1}	[54]
k	4.4463	day^{-1}	Fitted
β	0.0335	day^{-1}	Fitted
$\eta_1, \eta_2, \eta_3, \eta_4$	0.1296, 0.0168, 0.0251, 0.0279	unit-less	Fitted

4.2. Global sensitivity analysis

To determine which of the parameters have the highest impact on the outcome of the numerical simulations of the model, sensitivity analysis [55] is carried out on the entire parameter space of the model. Figure 9 presents a tornado plot for the partial rank correlation coefficient (PRCC) values for each parameter of the model, using the baseline values given in Table 4. From Figure 9, it can be seen that parameters with positive PRCC values (> 0.5) are β, k, π_n. Increasing these parameters consequently increases the reproduction number; hence, it becomes difficult to eliminate COVID-19 within the population, and the model will be driven to an endemic equilibrium. Identifying these key parameters is crucial in establishing effective control measures for fighting the spread of the COVID-19 disease. On the other hand, the parameter with a negative PRCC value (< 0.5) is γ_n. Increasing the rate at which individuals recover naturally or get hospitalized will significantly decrease the reproduction number; hence, the model will be driven to a disease-free equilibrium. Parameter, δ_n, though having an effect on the reproduction number when increased, is not feasible as it is a death-induced rate. The other parameters in Figure 9 have very small PRCC values and are, therefore, not discussed.



(a)

Figure 9. Tornado plot for the PRCC values of model system (2.1). Parameter values used are as given in Table 4.

4.3. Impact of human behaviors on the disease dynamics

The dynamics of infection are depicted in the time series plots shown in Figure 10. As seen in Figure 10(a) initially, an increase in the modification parameter, k, which relates to the force of infection for individuals exhibiting negative behavior, leads to a rise in the number of infected individuals with such behavior. However, over time, the dynamics shift. The shaded green area diminishes, suggesting that when $k = 10$, there are more infected individuals with negative behavior as compared to scenarios with smaller k values. The absence of the shaded region over time simply indicates that a reduction in k, does not reduce the number of infected humans with negative behavior. On the other hand, a similar situation is observed in Figure 10(b). However, it is observed that over time, there are more infections when $k = 10$ and fewer infections are reduced through the reduction of k, as indicated by the smaller shaded region. Figures 10(c) and 10(d) present time series plots for the effects of the exit rate out of classes I_n and I_p over time. It is observed from Figure 10(c) that when there are no infected individuals

who are either recovering from an infection or getting hospitalized, ($\gamma_n = 0$), then there is an influx of infected individuals with negative behavior. On the other hand, an increase in the exit rate, $\gamma_n > 0$ presents a significant reduction in the infected population with negative behavior. The green-shaded region represents the number of infections reduced due to increasing the exit rate. The same pattern can be observed in Figure 10(d).

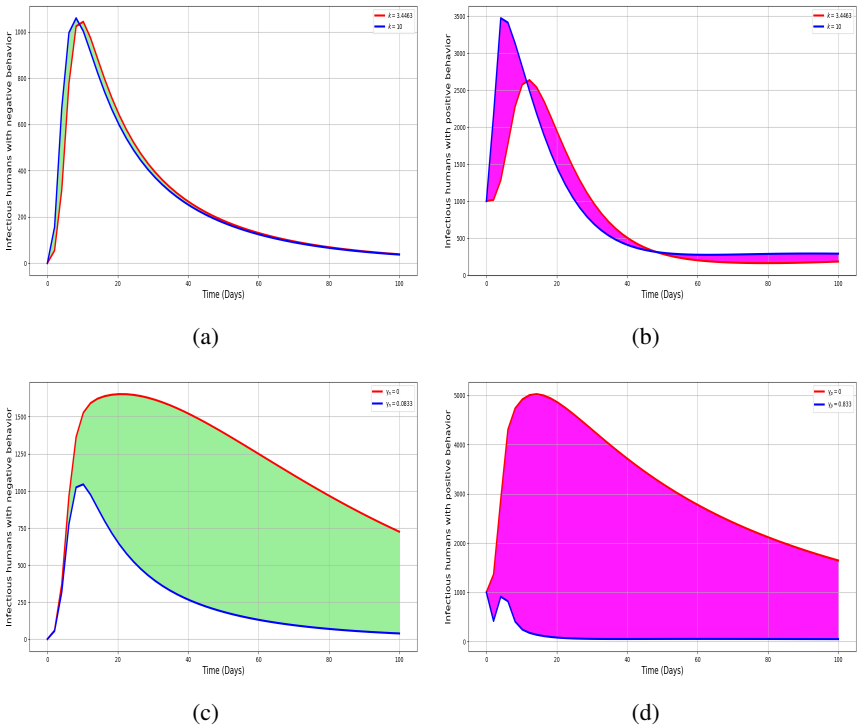
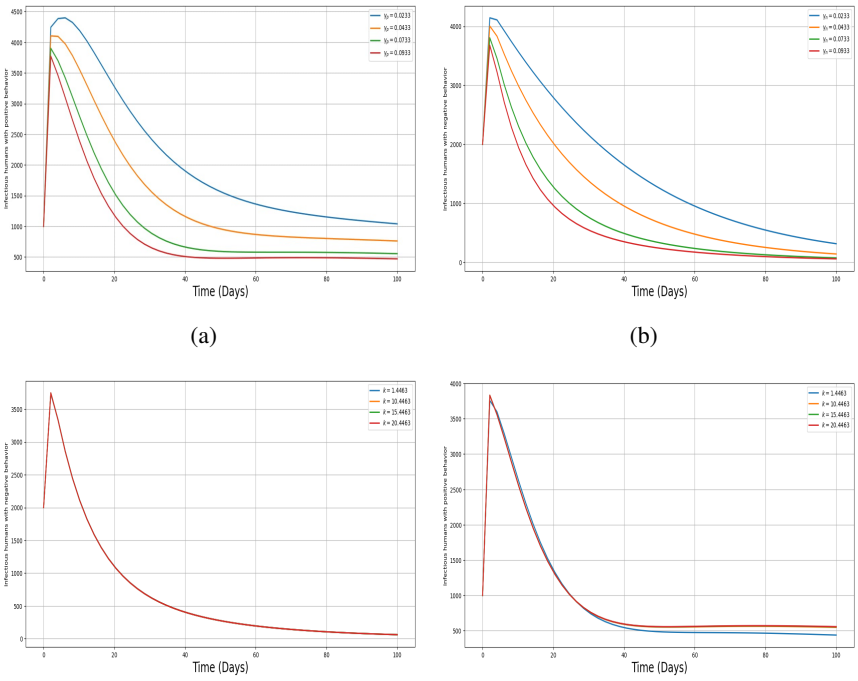


Figure 10. Impact of the behavior parameters for: (a) Infectious population I_n for $k = 3.4463$ and when $k = 10$, (b) infectious population I_p for $k = 3.4463$ and when $k = 10$, (c) infectious population I_n for $\gamma_n = 0$ and when $\gamma_n = 0.0833$, and (d) infectious population I_p for $\gamma_n = 0$ and when $\gamma_n = 0.0833$. The shaded area with green and magenta colors represents the impact of negative and positive behaviors while varying k and γ_n on I_n/I_p, respectively. The remaining parameter values are kept constant, as given in Table 4.

4.4. Effects of parameters on I_n and I_p

In this section, we explore the impact of parameters η_1, η_2, η_3, k, γ_n, and γ_p on the dynamics of our system. The simulations presented focus specifically on the infectious classes I_n and I_p. To simulate, we present the model solutions for varying values of these parameters. As illustrated in Figures 11(a) and 11(b), an increase in the exit parameters, γ_n, γ_p, consequently leads to a decrease in the number of infected individuals over time for both classes, I_n and I_p. This is due to more people either getting hospitalized or recovering from infection. In Figures 11(c) and 11(d), it can be seen that adjusting the modification parameter, k, results in an increase in the number of infected individuals over time. Given the presence of hospitalization and recovery processes, we notice that the infected classes eventually stabilize at a constant value as individuals transition out of these classes. Figure 12(a) and 12(b) present the dynamics of infection in the presence of varying parameters η_1, η_2 and k, η_3,

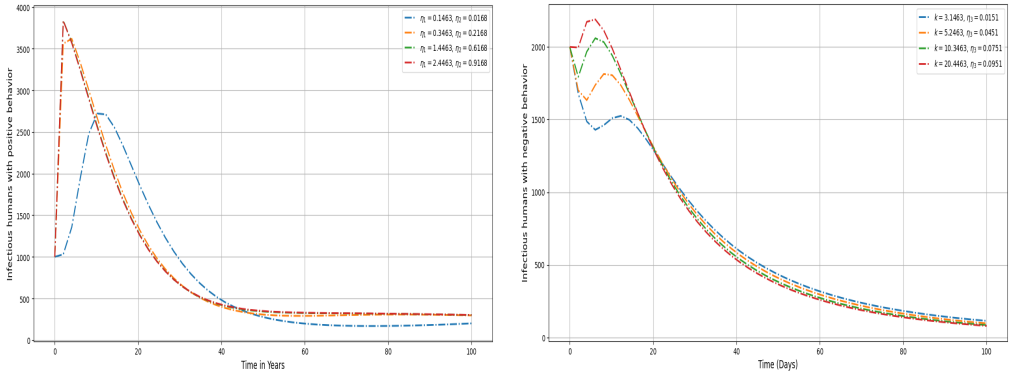
respectively. In Figure 12(a), we observe that an initial rise in infected individuals with positive behavior is evident when η_1 and η_2 are increased. However, as time progresses, the number of infected individuals stabilizes around a constant value after approximately 100 days, reflecting transitions out of I_p. Similar dynamics are observed in Figure 12(b) where parameters k, and η_3 are varied in the time series for I_n versus time. We initially observe an increase in the number of infected individuals when both parameters are increased. Now, due to progression out of class, I_n as a result of hospitalization or recovery, the number of infected individuals approaches a constant value over time.



(a) (b)

(c) (d)

Figure 11. Modeling the dynamics of infection through varying the parameters on I_n and I_p.



(a) (b)

Figure 12. Modeling the dynamics of infection through varying the parameters on I_n and I_p.

4.5. *Variations of $\mathcal{R}_0$ on the infected population*

The graphs in Figure 13 show the changes within the population of infected individuals with negative behavior and those with positive behavior for different values of the threshold quantity $\mathcal{R}_0$. We observe that as the value of $\mathcal{R}_0$ increases from low through to average, the population of infected individuals with negative behavior is higher than that of individuals with positive behavior over time. This indicates that those with negative behavior have a higher disease-spreading capacity and, therefore, generate more new infections than those with positive behavior. This points to the importance of the promotion of control strategies that decrease the practice of various negative behaviors to achieve disease containment/eradication. An interesting dynamic is also noticed at approximately 90 days, indicating a switch in the time series trajectories for the infectious individuals with positive and negative behavior. The epidemiological interpretation links to the onset of the COVID-19 spread. Most people are hesitant when it comes to taking vaccines and do not adhere to both pharmaceutical/non-pharmaceutical interventions, leading to more negatively behaved individuals. However, as the disease progresses and more people get infected and are hospitalized or die due to the disease, we see a change in human behaviors, as can be seen in Figure 13.

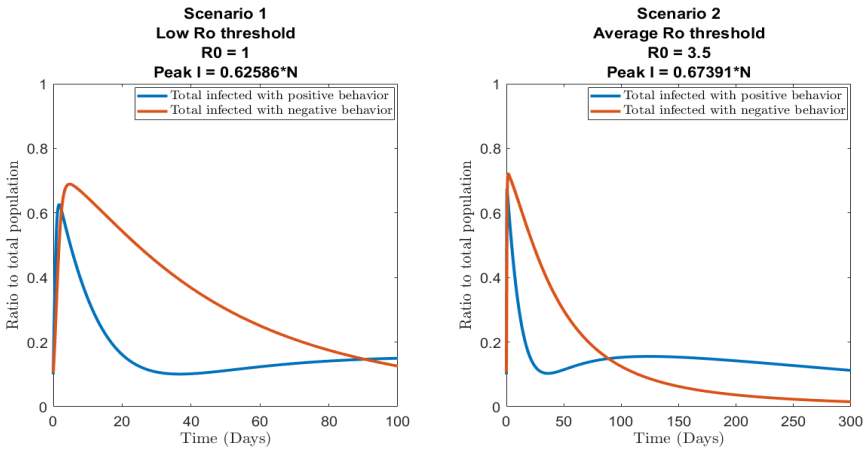
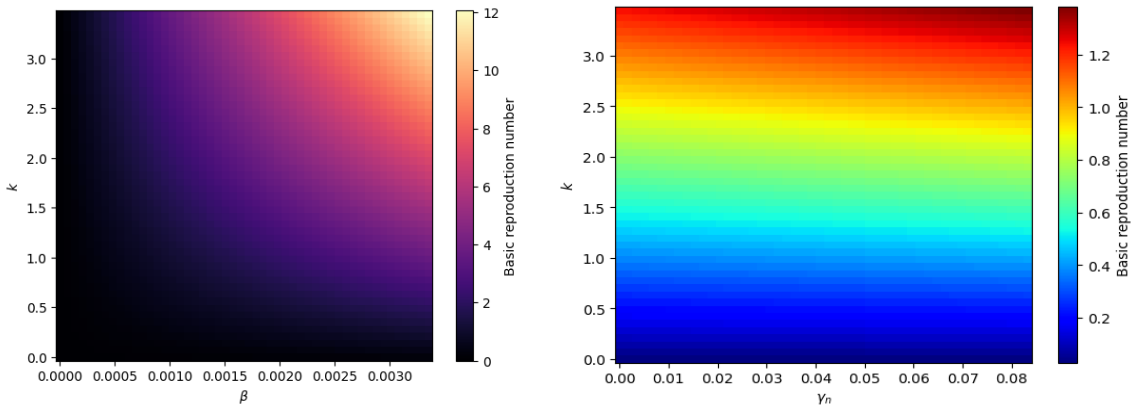


Figure 13. Time series trajectories of the infected population for different values of $\mathcal{R}_0$.

4.6. *Heat maps for parameters in $\mathcal{R}_0$.*

This subsection is devoted to using heat maps to numerically simulate some of the key parameters within the model (2.1) that are responsible for driving infection rates. We simulate the effective contact rate, (β), and the modification parameters (γ_n, k) as a function of the basic reproduction number $\mathcal{R}_0$, as shown in Figure 14. It is observed in 14(a) that increasing the modification parameter for negative behavior k, and the effective contact rate, β, leads to an increase $\mathcal{R}_0$. This means that as the basic reproduction number increases, the population will experience a greater number of COVID-19 infections, leading to an increased likelihood of individuals either contracting the virus or requiring hospitalization due to its heightened infectiousness. A similar interpretation applies to Figure 14(b), portraying k versus γ_n in relation to $\mathcal{R}_0$. The heat maps show that effective control measures are essential in reducing these key parameters to curb disease transmission.



(a) (b)

Figure 14. Heat maps of parameters β, γ_n, and k as a function of $\mathcal{R}_0$.

5. Discussion and conclusions

Human behavior plays a crucial role in the transmission and mitigation of COVID-19. This study has provided valuable insights into the relationship between human behavior and the dynamics of COVID-19 transmission. By leveraging time series analysis and mathematical modeling, we have unraveled the intricate dynamics shaping the spread of the virus in relation to human actions and interventions. Our time series analysis reveals that the LSTM model outperforms other models in terms of predictive accuracy and suitability for capturing the complex temporal patterns inherent in COVID-19 data. This finding underscores the effectiveness of deep learning approaches, particularly LSTM, in modeling and forecasting infectious disease dynamics. The superior performance of the LSTM model suggests its potential utility in informing public health interventions and decision-making processes aimed at controlling the spread of COVID-19.

A mathematical model was developed and used to differentiate the total population into individuals with positive and negative behaviors. This model was used to study COVID-19 infection in the Gauteng province of South Africa. It incorporated key features associated with disease transmission, such as the interaction between infectious individuals exhibiting both positive and negative behaviors. Our modeling framework assessed the effect of human behavior on COVID-19 transmission. The study revealed that, in the absence of individuals with negative behaviors, COVID-19 outbreaks could impose a much bigger public health burden in Gauteng and South Africa nationally. Variations of the basic reproduction number ($\mathcal{R}_0$), an epidemiological threshold quantity measuring the disease's spread capacity, indicate that the reproduction number is high in the presence of a large number of infectious individuals with negative behavior and decreases as their number decreases. Additionally, some isolated infected individuals chose to escape isolation, further contributing to the spread of infections. Furthermore, this study identifies the main negative human behaviors driving the COVID-19 outbreaks during the disease's early (pre-intervention) phase. These behaviors include disbelief in the existence of COVID-19, hiding infections, seeking help from TCM practitioners, and conducting unsafe burials of deceased individuals. Identifying these critical parameters aids in constructing an effective control strategy, such as one that minimizes the impact of disbelief in COVID-19 and

reduces stigmatization and the practice of unsafe burials. This could lead to effective community-wide control strategies through public health education campaigns. The study demonstrates that COVID-19 outbreaks are controllable using basic and affordable public health control measures. Specifically, a strategy that convinces community members that COVID-19 is a genuine viral infection and not merely government propaganda or a spiritual problem and that individuals who recover from it can no longer transmit the disease could dramatically reduce the infection burden in affected communities.

Other works in literature have shown that medical mistrust and other behaviors exhibited by humans across the globe significantly impacted the COVID-19 pandemic infection. To substantiate our claim, see, for instance, mentions of proof of human behaviors in the following countries: USA [56], South Africa [57], United Kingdom [58], and other country-level analyses [59, 60] responsible for increasing COVID-19 infection rates. The World Health Organization [60], classified these negative or positive behaviors exhibited by humans as an infodemic, a rapid and far-reaching spread of false or misleading information in digital and physical environments during a disease outbreak. This leads to confusion and risk-taking behaviors that harm health policies and mistrust in health authorities and undermine public health response [60]. In particular, Mphahlele [57] found that in many African countries, including South Africa, an infodemic of misinformation propagated by popular figures, trusted community members/leaders, and fake news and conspiracy theories surrounding vaccine hesitancy fueled the distrust toward the COVID-19 vaccine and, therefore, increased the risk of COVID-19 infection. Negative human behavior significantly fueled the COVID-19 outbreaks in South Africa [57]. People in affected communities generally did not adhere to preventive measures, often due to mistrust of authorities and the belief that COVID-19 did not exist; hence, they considered it government propaganda [56, 57].

However, the results obtained in this study were limited by the lack of specific data on COVID-19 for infectious individuals with negative behavior as well as the failure to consider vaccination, which is a widely used means of control against COVID-19. Despite this weakness, our model remains highly applicable to real-case scenarios of COVID-19 transmission pathways. In future work, we will take vaccination into account and extend the mathematical model to study the coinfection of COVID-19 with streptococcus pneumoniae.

Use of AI tools declaration

The authors declare that they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This work is supported by funding from Wake Forest University. The authors would like to thank their respective universities for the production of this manuscript. SYT expresses acknowledgment of DSI/NRF SARChI M3B2 grant N 82770 for postdoctoral funding.

Conflict of interest

The authors declare that they have no competing interests.

References

1. *Coronavirus Disease 2019 (COVID-19) Situation Report-40*, World Health Organization, 2020. Available from: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200229-sitrep-40-covid-19.
2. *Coronavirus Disease (COVID-19): How is it Transmitted?* World Health Organization, 2021. Available from: https://www.who.int/news-room/questions-and-answers/item/coronavirus-disease-covid-19-how-is-it-transmitted.
3. S. Funk, M. Salathé, V. A. A. Jansen, Modelling the influence of human behaviour on the spread of infectious diseases: a review, *J. Royal Soc. Interf.*, **7** (2010), 1247–1256. http://dx.doi.org/10.1098/rsif.2010.0142
4. *COVID-19 Vaccine: What You Need to Know*, Johns Hopkins Medicine, 2022. Available from: https://www.hopkinsmedicine.org/health/conditions-and-diseases/coronavirus/covid-19-vaccines-myth-versus-fact.
5. *Bringing Traditional Healing Under the Microscope in South Africa*, Medscape, 2020. Available from: https://www.medscape.com/viewarticle/943429.
6. S. L. Canham, P. M. Mauro, C. N. Kaufmann, A. Sixsmith, Association of alcohol use and loneliness frequency among middle-aged and older adult drinkers, *J. Aging Health*, **28** (2016), 267–284. https://doi.org/10.1177/0898264315589579
7. *Advice for the Public: Coronavirus Disease (COVID-19)*, World Health Organization, 2022. Available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public.
8. Y. Li, D. Ji, W. Cai, Y. Hu, Y. Bai, J. Wu, et al., Clinical characteristics, cause analysis and infectivity of COVID-19 nucleic acid repositive patients: a literature review, *J. Med. Virol.*, **93** (2021), 1288–1295. http://dx.doi.org/10.1002/jmv.26491
9. W. McNeill, *Plagues and Peoples*, New York: Anchor Press, 2010.
10. S. Funk, S. Bansal, C. T. Bauch, K. T. D. Eames, W. John Edmunds, A. P. Galvani, et al., Nine challenges in incorporating the dynamics of behaviour in infectious diseases models, *Epidemics*, **10** (2015), 21–25. https://doi.org/10.1016/j.epidem.2014.09.005
11. M. Salathé, S. Bonhoeffer, The effect of opinion clustering on disease outbreaks, *J. Royal Soc. Interf.*, **5** (2008), 1505–1508. http://dx.doi.org/10.1098/rsif.2008.0271
12. J. M. Epstein, J. Parker, D. Cummings, R. A. Hammond, Coupled contagion dynamics of fear and disease: mathematical and computational explorations, *PloS One*, **16** (2008), e3955. http://dx.doi.org/ 10.1371/journal.pone.0003955
13. D. H. Zanette, S. Risau-Gusmán, Infection spreading in a population with evolving contacts, *J. Biol. Phys.*, **34** (2008), 135–148. http://dx.doi.org/10.1007/ s10867-008-9060-9
14. T. Gross, C. J. Dommar D'Lima, B. Blasius, Epidemic dynamics on an adaptive network, *Phys. Rev. Lett.*, **96** (2006), 208701. https://doi.org/10.1103/PhysRevLett.96.208701
15. L. B. Shaw, I. B. Schwartz, Fluctuating epidemics on adaptive networks, *Phys. Rev. E*, **77** (2008), 066101. https://doi.org/10.1103/PhysRevE.77.066101
16. C. T. Bauch, Imitation dynamics predict vaccinating behaviour, *Proc. Royal Soc. B Biol. Sci.*, **272** (2005), 1669–1675. https://doi.org/10.1098/rspb.2005.3153

17. M. Juga, F. Nyabadza, F. Chirove, An Ebola virus disease model with fear and environmental transmission dynamics, *Infect. Disease Model.*, **6** (2021), 545–559. https://doi.org/10.1016/j.idm.2021.03.002
18. N. Zhang, W. Jia, H. Lei, P. Wang, P. Zhao, Y. Guo, et al., Effects of human behavior changes during the coronavirus disease 2019 (COVID-19) pandemic on influenza spread in Hong Kong, *Clin. Infect. Dis.*, **73** (2021), e1142–e1150. https://doi.org/10.1093/cid/ciaa1818
19. U. Kollamparambil, A.Oyenubi, Behavioural response to the Covid-19 pandemic in South Africa, *PloS One*, **16** (2021), e0250269. https://doi.org/10.1371/journal.pone.0250269
20. F. Nyabadza, F. Chirove, C. Chukwu, M. V. Visaya, Modelling the potential impact of social distancing on the COVID-19 epidemic in South Africa, *Comput. Math. Meth. Medic.*, **2020** (2020), 5379278. https://doi.org/10.1155/2020/5379278
21. S. P. Gatyeni, C. W. Chukwu, F. Chirove, Fatmawati, F. Nyabadza, Application of optimal control to the dynamics of COVID-19 disease in South Africa, *Sci. Afr.*, **16** (2022), e01268. https://doi.org/10.1016/j.sciaf.2022.e01268
22. C. J. Edholm, B. Levy, L. Spence, F. B. Agusto, F. Chirove, C. W. Chukwu, et al., A vaccination model for COVID-19 in Gauteng, South Africa, *Infect. Disease Model.*, **7** (2022), 333–345. https://doi.org/10.1016/j.idm.2022.06.002
23. C. W. Chukwu, Fatmawati, Modelling fractional-order dynamics of COVID-19 with environmental transmission and vaccination: a case study of Indonesia, *AIMS Math.*, **7** (2022), 4416–4438. https://doi.org/10.3934/math.2022246
24. J. Mushanyu, W. Chukwu, F. Nyabadza, G. Muchatibaya, Modelling the potential role of super spreaders on COVID-19 transmission dynamics, *Int. J. Math. Model. Numer. Optim.*, **12** (2022), 191–209. https://dx.doi.org/10.1504/IJMMNO.2022.122123
25. J. Mushanyu, C. W. Chukwu, C. E. Madubueze, Z. Chazuka, C. P. Ogbogbo, A deterministic compartmental model for investigating the impact of escapees on the transmission dynamics of COVID-19, *Healthc. Anal.*, **4** (2023), 100275. https://doi.org/10.1016/j.health.2023.100275
26. S. Gao, P. Binod, C. W. Chukwu, T. Kwofie, S. Safdar, L. Newman, et al., A mathematical model to assess the impact of testing and isolation compliance on the transmission of COVID-19, *Infect. Disease Model.*, **8** (2023), 427–444. https://doi.org/10.1016/j.idm.2023.04.005
27. S. M. Simelane, P. G. Dlamini, F. J. Osaye, G. Obaido, B. Ogbukiri, K. Aruleba, et al., Modeling the impact of public health education on tungiasis dynamics with saturated treatment: Insight through the Caputo fractional derivative, *Math. Biosci. Eng.*, **20** (2023), 7696–7720. http://dx.doi.org/10.3934/mbe.2023332
28. C. Chukwu, R. Alqahtani, C. Alfiniyah, F. Herdicho, Tasmi, A Pontryagin's maximum principle and optimal control model with cost-effectiveness analysis of the COVID-19 epidemic, *Decis. Anal. J.,* **8** (2023), 100273. https://doi.org/10.1016/j.dajour.2023.100273
29. Fatmawati, E. Yuliani, C. Alfiniyah, M. L. Juga, C. W. Chukwu, On the modeling of COVID-19 transmission dynamics with two strains: insight through caputo fractional derivative, *Fractal Fract.*, **6** (2022), 346. https://doi.org/10.3390/fractalfract6070346
30. E. Bonyah, M. Juga, L. Matsebula, C. Chukwu, On the modeling of COVID-19 spread via fractional derivative: a stochastic approach, *Math. Models Comput. Simul.*, **15** (2023), 338–356. https://doi.org/10.1134/S2070048223020023

31. T. Li, Y. Guo, Modeling and optimal control of mutated COVID-19 (Delta strain) with imperfect vaccination, *Chaos Solitons Fract.*, **156** (2022), 111825. https://doi.org/10.1016/j.chaos.2022.111825
32. Y. Guo, T. Li, Modeling the competitive transmission of the Omicron strain and Delta strain of COVID-19, *J. Math. Anal. Appl.*, **526** (2023), 127283. https://doi.org/10.1016/j.jmaa.2023.127283
33. G. Perone, Comparison of ARIMA, ETS, NNAR, TBATS and hybrid models to forecast the second wave of COVID-19 hospitalizations in Italy, *Eur. J. Health Econ.*, **23** (2022), 917–940. https://doi.org/10.1007/s10198-021-01347-4
34. A. A. Ismail, T. Wood, H. C. Bravo, Improving long-horizon forecasts with expectation-biased LSTM networks, preprint paper, 2018. https://doi.org/10.48550/arXiv.1804.06776
35. Z. Tarek, M. Y. Shams, S. K. Towfek, H. K. Alkahtani, A. Ibrahim, A. A. Abdelhamid, et al., An optimized model based on deep learning and gated recurrent unit for COVID-19 death prediction, *Biomimetics*, **8** (2023), 552. https://doi.org/10.3390/biomimetics8070552
36. O. Diekmann, J. A. P. Heesterbeek, J. A. Metz, On the definition and the computation of the basic reproduction ratio R0 in models for infectious diseases in heterogeneous populations, *J. Math. Biol.*, **28** (1990), 365–382. https://doi.org/10.1007/bf00178324
37. H. W. Hethcote, The mathematics of infectious diseases, *SIAM Rev.*, **42** (2000), 599–653. https://doi.org/10.1137/S0036144500371907
38. P. Van den Driessche, J. Watmough, Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission, *Math. Biosci.*, **180** (2002), 29–48. https://doi.org/10.1016/S0025-5564(02)00108-6
39. R. Gupta, S. K. Pal, Trend analysis and forecasting of COVID-19 outbreak in India, *MedRxiv*, 2020. https://doi.org/10.1101/2020.03.26.20044511
40. P. Wang, X. Zheng, G. Ai, D. Liu, B. Zhu, Time series prediction for the epidemic trends of COVID-19 using the improved LSTM deep learning method: case studies in Russia, Peru and Iran, *Chaos Solitons Fract.*, **140** (2020), 110214. https://doi.org/10.1016/j.chaos.2020.110214
41. S. Dash, C. Chakraborty, S. K. Giri, S. K. Pani, Intelligent computing on time-series data analysis and prediction of COVID-19 pandemics, *Pattern Recogn. Lett.*, **151** (2021), 69–75. https://doi.org/10.1016/j.patrec.2021.07.027
42. *COVID-19 ZA South Africa Dashboard*, Department of health, South Africa, 2022. Available from: `https://dsfsi.github.io/covid19za-dash/`.
43. R. T. Aruleba, T. A. Adekiya, N. Ayawei, G. Obaido, K. Aruleba, I. D. Mienye, et al., COVID-19 diagnosis: a review of rapid antigen, RT-PCR and artificial intelligence methods, *Bioeng.*, **9** (2022), 153. https://doi.org/10.3390/bioengineering9040153
44. E. Esenogho, I. D. Mienye, T. G. Swart, K. Aruleba, G. Obaido, A neural network ensemble with feature engineering for improved credit card fraud detection, *IEEE Access*, **10** (2022), 16400–16407. https://doi.org/10.1109/ACCESS.2022.3148298
45. K. Cho, B. Van Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, et al., Learning phrase representations using RNN encoder-decoder for statistical machine translation, preprint paper, 2014. https://doi.org/10.48550/arXiv.1406.1078
46. R. Dey, F. M. Salem, Gate-variants of gated recurrent unit (GRU) neural networks, In: *2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS), IEEE*, 2017, 1597–1600. https://doi.org/10.1109/MWSCAS.2017.8053243

47. M. Diagne, H. Rwezaura, S. Tchoumi, J. Tchuenche, A mathematical model of COVID-19 with vaccination and treatment, *Comput. Math. Meth. Medic.*, **2021** (2021), 1250129. https://doi.org/10.1155/2021/1250129

48. M. O. Adewole, A. A. Onifade, F. A. Abdullah, F. Kasali, A. I. Ismail, Modeling the dynamics of COVID-19 in Nigeria, *Int. J. Appl. Comput. Math.*, **7** (2021), 67. https://doi.org/10.1007/s40819-021-01014-5

49. A. Babaei, H. Jafari, S. Banihashemi, M. Ahmadi, Mathematical analysis of a stochastic model for spread of Coronavirus, *Chaos Solitons Fract.*, **145** (2021), 110788. https://doi.org/10.1016/j.chaos.2021.110788

50. C. T. Deressa, Y. O. Mussa, G. F. Duressa, Optimal control and sensitivity analysis for transmission dynamics of Coronavirus, *Res. Phys.*, **19** (2020), 103642. https://doi.org/10.1016/j.rinp.2020.103642

51. B. Tang, X. Wang, Q. Li, N. L. Bragazzi, S. Tang, Y. Xiao, et al., Estimation of the transmission risk of the 2019-nCoV and its implication for public health interventions, *J. Clin. Medic.*, **9** (2020), 462. https://doi.org/10.3390/jcm9020462

52. S. M. Garba, L. M. S. Lubuma, B. Tsanou, Modeling the transmission dynamics of the COVID-19 Pandemic in South Africa, *Math. Biosci.*, **328** (2020), 108441. https://doi.org/10.1016/j.mbs.2020.108441

53. *South Africa Life Expectancy 1950–2023*, Macrotrends, 2023. Available from: `https://www.macrotrends.net/countries/ZAF/south-africa/life-expectancy`

54. M. Q. Shakhany, K. Salimifard, Predicting the dynamical behavior of COVID-19 epidemic and the effect of control strategies, *Chaos Solitons Fract.*, **146** (2021), 110823. https://doi.org/10.1016/j.chaos.2021.110823

55. S. M. Blower, H. Dowlatabadi, Sensitivity and uncertainty analysis of complex models of disease transmission: an HIV model, as an example, *Int. Statist. Rev.*, **62** (1994), 229–243. https://doi.org/10.2307/1403510

56. L. M. Bogart, B. O. Ojikutu, K. Tyagi, D. J. Klein, M. G. Mutchler, L. Dong, et al., COVID-19 related medical mistrust, health impacts, and potential vaccine hesitancy among Black Americans living with HIV, *J. Acq. Imm. Def. Synd.*, **86** (2021), 200–207. https://doi.org/10.1097/QAI.0000000000002570

57. J. Mphahlele, Conspiracy theories on COVID-19 vaccine can be as deadly as virus itself, *South African Med. Res. Council: Cape Town*, 2021.

58. C. Jacob, P. Hausemer, A. Zagoni-bogsch, A. Diers-lawson, The effect of communication and disinformation during the Covid-19 pandemic, *European Parliament*, 2023. https://doi.org/10.2861/501274

59. O. F. Norheim, J. M. Abi-Rached, L. K. Bright, K. Bærøe, O. L. Ferraz, S. Gloppen, et al., Difficult tradeoffs in response to COVID-19: the case for open and inclusive decision making, *Nature Medic.*, **27** (2021), 10–13. https://doi.org/10.1038/s41591-020-01204-6

60. *Infodemic*, World Health Organization, 2024. Available from: `https://www.who.int/health-topics/infodemic#tab=tab_1`.

Appendix

Appendix A: Existence, uniqueness, and positivity of solutions

The righthand side of the system (2.1) consists of Lipschitz continuous functions. Picard's existence theorem, therefore, guarantees that Model (2.1) has a unique solution.

Theorem 1. *The solutions of the system (2.1) are all nonnegative, given that the initial conditions are nonnegative.*

Proof. We assume that there exists a time $\hat{t}$ such that $S_p(\hat{t}) = 0$, $S_p'(\hat{t}) < 0$, $S_p(t) > 0$, $S_n(t) > 0$, $I_p(t) > 0$, $I_n(t) > 0$, $A_p(t) > 0$, $A_n(t) > 0$, $H(t) > 0$, $R_p(t) > 0$, $R_n(t) > 0$ for $0 < t < \hat{t}$. From the first equation of (2.1), we have

$$\dot{S_p}(\hat{t}) = \pi_p + \omega_p R_p > 0.$$

This contradicts the assumption that $\dot{S_p}(\hat{t}) < 0$. Therefore $S(t)$ is positive. Similarly, $S_n(t) > 0$, $I_p(t) > 0$, $I_n(t) > 0$, $A_p(t) > 0$, $A_n(t) > 0$, $H(t) > 0$, $R_p(t) > 0$, $R_n(t) > 0$. □

Theorem 2. *The region*

$$\Theta = \left\{(S_p, S_n, A_p, A_n, I_p, I_n, R_p, R_n, H) \in \mathbb{R}_+^9 : N_p(t) \leq \frac{\pi_p}{\mu},\ N_n(t) \leq \frac{\pi_n}{\mu}\right\}$$

is the invariant region in which the solutions of system (2.1) are bounded.

Proof. Given that $N_p(t) = S_p(t) + A_p(t) + I_p(t) + R_p(t) + H(t)$ and $N_n(t) = S_n(t) + A_n(t) + I_n(t) + R_n(t)$ then

$$\begin{aligned} N_p' &= (\pi_p - \mu N_p) - (\delta_p I_p + (1-\kappa_n)\gamma_n I_n + \delta_h H), \\ &\leq \pi_p - \mu N_p. \end{aligned}$$

Solving the above differential inequality, we obtain

$$N_p(t) \leq \frac{\pi_p}{\mu} - \left(\frac{\pi_p}{\mu} - N_{p_0}\right)\exp(-\mu t).$$

Therefore,

$$\limsup_{t\longrightarrow\infty} N_p(t) = \frac{\pi_p}{\mu}.$$

Since $N_p(t) = S_p(t) + A_p(t) + I_p(t) + H(t) + R_p(t)$, it follows that $S_p(t) \leq \frac{\pi_p}{\mu}$, $A_p(t) \leq \frac{\pi_p}{\mu}$, $I_p(t) \leq \frac{\pi_p}{\mu}$, $H(t) \leq \frac{\pi_p}{\mu}$, $R_p(t) \leq \frac{\pi_p}{\mu}$.

Also,

$$\dot{N_n}(t) \leq \pi_n - \mu N_n.$$

Hence, we solve the differential inequality and obtain

$$N_n(t) \leq \frac{\pi_n}{\mu} - \left(\frac{\pi_n}{\mu} - N_{n_0}\right)\exp(-\mu t).$$

Therefore,

$$\limsup_{t\longrightarrow\infty} N_n(t) = \frac{\pi_n}{\mu},$$

and it follows that $S_n(t) \leq \frac{\pi_n}{\mu}$, $A_n(t) \leq \frac{\pi_n}{\mu}$, $I_n(t) \leq \frac{\pi_n}{\mu}$, $R_n(t) \leq \frac{\pi_n}{\mu}$. This clearly shows that all the solutions of the system (2.1) are bounded, and hence, this completes the proof. □

Appendix B: Forecast error measures

B1: MAE

The MAE measures the mean of the absolute differences between actual and predicted values, assigning equal significance to errors regardless of their direction or magnitude.

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |Y_i - \hat{Y}_i|. \tag{5.1}$$

B2:MSE

The MSE is the mean of the squared differences between predicted and actual values, magnifying larger errors due to the squaring process. This metric is valuable for evaluating a model's precision comprehensively and is frequently employed in optimization procedures.

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2. \tag{5.2}$$

B3: RMSE

The RMSE measures the square root of the mean of the squared differences between the actual and predicted values. RMSE proves valuable for assessing error magnitude across various models and gauging the scale of discrepancies relative to predicted values.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2}. \tag{5.3}$$