STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **15**, May 2026, pp 3430–3467.
Published online in International Academic Press (www.IAPress.org)

Stochastic Modeling of a Vaccine-Structured Epidemic Model Using Data From South Africa

Aubrey Ndovie*, Claris Shoko, Olusegun S. Ewemooje, Sivasamy Ramasamy

Department of Statistics, University of Botswana, plot 4775 Notwane Rd Gaborone Botswana

Abstract
Introduction: Understanding the initial dynamics of an epidemic, especially whether it will establish itself or die out, is critical for public health policy. Deterministic models provide insight into average population behavior but cannot capture the random chance, or demographic stochasticity, that governs the fate of an outbreak when infectious case numbers are low. This is particularly relevant for COVID-19, where population heterogeneity due to vaccination significantly influences transmission. In this paper, we develop and analyze a vaccine-structured epidemic model to quantify the probability of disease extinction and understand how vaccination status impacts these early, uncertain dynamics.**Materials and Methods**: We formulated a deterministic model using a system of eight ordinary differential equations (ODEs) to represent non-vaccinated and vaccinated populations, incorporating waning immunity. A corresponding Continuous-Time Markov Chain (CTMC) model was developed to capture stochastic effects. The basic reproduction number, $\mathscr{R}_0$, was derived using the next-generation matrix method. We applied multitype branching process theory to analytically calculate the probability of disease extinction (P_0) and used Gillespie's Stochastic Simulation Algorithm to run 10,000 CTMC sample paths to numerically approximate this probability (P_A) and the finite time to extinction. The model was grounded using parameters fitted to COVID-19 data from South Africa during the period from March 5, 2020 to March 21, 2022. **Results**: The basic reproduction number was calculated as $\mathscr{R}_0 \approx 1.4130$, indicating the potential for sustained transmission. The extinction probability derived from the branching process (P_0) showed excellent agreement with the simulated approximation (P_A). A key finding is that an infection introduced by a vaccinated individual has a significantly higher chance of extinction ($P_A \approx 0.90 - 0.93$) compared to one from a non-vaccinated individual ($P_A \approx 0.75 - 0.78$). Furthermore, outbreaks initiated by an infectious vaccinated person that do go extinct resolve the fastest ($T \approx 35$ days), while those from an infectious non-vaccinated person persist the longest ($T \approx 62$ days).**Conclusion**: This study demonstrates that vaccination provides a dual benefit in containing new disease introductions: it substantially increases the probability of stochastic fade-out and shortens the duration of abortive outbreaks. These findings highlight the limitations of relying solely on deterministic thresholds like $\mathscr{R}_0$ and underscore the importance of stochastic models in providing a more nuanced risk assessment for public health planning, emphasizing that vaccination is a powerful tool for preventing new sparks from becoming major epidemics.

Keywords COVID-19, Deterministic, Threshold, Markov Chain, Multi-type Branching Process.

DOI: 10.19139/soic-2310-5070-2707

1. Introduction

In the last months of 2019, a new virus, called the 2019 Coronavirus (COVID-19), emerged in Wuhan, China [1].The virus began to spread rapidly, resulting in a worldwide pandemic that affected every country and region of the world. In November 2023, there were 771,820,937 confirmed cases worldwide, resulting in 6,978,175 deaths [2]. The African region also experienced the wrath of this pandemic, although it occurred later than in other regions. The first case was reported in Egypt on February 14, 2020, followed by Algeria on February 25, 2020 [3]. In March 2020, COVID-19 entered Southern Africa via South Africa [4]. Botswana, Namibia, and South Africa exhibited the highest prevalence rates in Southern Africa. Botswana's prevalence rate was nearly 5238 cases per

*Correspondence to: Aubrey Ndovie (Email: 201300078@ub.ac.bw)

ISSN 2310-5070 (online) ISSN 2311-004X (print)

100,000, Namibia's was 4746 cases per 100,000, and South Africa's prevalence rate stood at 4212 per 100,000 [5].

COVID-19 arises from a novel identified coronavirus called Severe Acute Respiratory Syndrome Corona Virus 2 (SARS-CoV-2) [1]. This virus is part of the coronavirus family, which comprises Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV) and Middle East Respiratory Syndrome Coronavirus (MERS-CoV). However, SARS-CoV-2 exhibits a lower mortality rate [6]. Despite this fact, SARS-CoV-2 caused a more severe pandemic compared to those caused by SARS-CoV and MERS-CoV. This was primarily due to the high level of contagiousness of the virus, which enabled its rapid and extensive global spread. The increased transmissibility of the COVID-19 virus can be attributed to various factors, such as its ability to spread through respiratory droplets and aerosols, an extended period of infectivity, especially in asymptomatic carriers and the emergence of multiple variants with greater infectiousness than the original strain [7].

COVID-19, an RNA virus with high mutation rates, led to the emergence of variants like Alpha, Beta, Delta and Omicron, complicating containment efforts [8, 9]. Countries implemented stringent non-pharmaceutical interventions, which significantly affected sectors like tourism and travel [10], highlighting the need for vaccines. Various pharmaceutical companies, such as AstraZeneca, Johnson & Johnson, Pfizer, and Moderna, were instrumental in the development of vaccines [11]. However, research has indicated that the immunity provided by these vaccines is not enduring and diminishes over time, typically persisting for approximately six months [12].The importance of ongoing non-pharmaceutical interventions was highlighted. Consequently, the mathematical/statistical modeling community played a crucial role as well. Models were created to assess the effects of both pharmaceutical and non-pharmaceutical interventions, aiding in making informed decisions based on transmission rates, reproduction numbers, and recovery rates derived from their analyses.

Most infectious disease models use the concept of compartmental modeling frameworks [13]. In this kind of modeling, a population is divided into compartments such as susceptible individuals, infectious individuals and recovered individuals [13]. These models can be either deterministic or stochastic. Deterministic models, generally structured as ordinary differential equations (ODEs), are straightforward, well-suited for well-understood systems, and effectively average randomness over large populations [14, 15]. However, they lack the ability to incorporate unpredictable fluctuations (randomness). In contrast, stochastic models cater to the randomness inherent in probabilistic systems, such as disease outbreaks, using frameworks such as discrete-time or continuous-time Markov chains (CTMCs) and stochastic differential equations (SDEs) [16, 17]. These models offer the ability to simulate random fluctuations and heterogeneity in individual responses [18]. In particular, CTMCs prove advantageous because they handle discrete state spaces and individual variability efficiently, making them suitable for analyzing dynamics such as those of COVID-19 [18, 19].

1.1. Literature review

In the case of SARS-CoV, deterministic models have been instrumental in understanding disease spread and control measures, as demonstrated in a study by [20] that analyzed the impact of quarantine strategies. In a similar study by [21], a deterministic model was used showing how isolation measures, social distancing and knowledge of diffusion conditions helped contain the pandemic. [22] developed a stochastic Susceptible-Infected-Recovered-Cross-Immune (SIRC) epidemic model with time delay for COVID-19, incorporating environmental white noise to enhance the model realism and provide more accurate insights into the disease dynamics. The analytical and numerical results of the study highlighted the impact of different parameters on the spread of the virus, emphasizing the importance of considering environmental noise in epidemic models. In a study by [23], the CTMC model was applied to calculate the probability of the COVID-19 outbreak, to better predict the disease outbreak and control its spread. The authors used the mathematical model of the 4 compartment Susceptible-Exposed-Infectious-Removal (SEIR) to estimate the current trend and the upcoming infection status of COVID-19 at county levels of Texas.

While these studies provided vital early insights, recent literature has pivoted towards the long-term management of COVID-19, where waning immunity and vaccination heterogeneity are paramount. Deterministic models alone often fail to capture the probabilistic risk of resurgence when case numbers are low, a critical factor in post-peak pandemic phases. For instance, [24] highlighted that the trajectory of COVID-19 in the coming years heavily depends on the duration of immunity and the impact of vaccination, complex dynamics that simple SIR models may oversimplify. Furthermore, [25] highlighted that population heterogeneity can significantly alter the herd immunity threshold, leading to estimates that differ from those derived from homogeneous models. Recent work by [16] further demonstrated that stochastic models often predict lower peaks and different extinction timelines compared to their deterministic counterparts, particularly when accounting for heterogeneous transmission rates. This gap in the literature necessitates a comparative approach using both deterministic and stochastic frameworks, specifically tailored to vaccine-structured populations with waning immunity, to fully understand the risks of resurgence from small introduction events.

1.2. Contribution and Research highlights

The primary motivation for this study addresses the limitation of relying solely on deterministic models to predict early epidemic dynamics. While useful for large-scale average trends, deterministic frameworks cannot accurately assess the probability of a disease failing to take hold when introduced by only a few infectious individuals. This is particularly crucial when considering a heterogeneous population where some individuals possess partial, waning immunity from vaccination. Understanding whether a 'spark' from a vaccinated individual is as likely to cause a major fire as one from a non-vaccinated individual is vital for accurate public health risk assessment in the current phase of the pandemic. This study aims to bridge this gap by exploring both deterministic and stochastic compartmental models for COVID-19, explicitly incorporating the heterogeneity introduced by the decline in vaccine-derived immunity.

1.2.1. Research Highlights The main contributions of this work are:

- the development of a unified vaccine-structured framework designed to track disease progression and waning immunity across vaccinated and non-vaccinated sub-populations, grounded by parameters fitted to South African data.
- to conduct a comparative threshold analysis to contrast disease behaviors, specifically extinction versus outbreak scenarios, between deterministic approaches that use the basic reproduction number ($\mathcal{R}_0$) and stochastic approaches based on the spectral radius of the expectation matrix.
- to quantify the specific probability of extinction for scenarios where the disease is introduced by different index cases (e.g., a single infected vaccinated individual versus an infected non-vaccinated one), employing both analytical multi-type branching process theory and numerical Gillespie stochastic simulations.

1.3. Organization

The remaining sections of this paper are structured as follows: initially, the deterministic and stochastic models are formulated in Section 2. In Section 3 the results from the data fitting, parameter estimation and sensitivity analysis are discussed. Furthermore, numerical simulations are carried out in the same section. Finally, in Section 4 the results are discussed and the paper concludes with Section 5.

2. Methods

This section formulates a deterministic model that represents the transmission dynamics of COVID-19 in South Africa. The model considers the heterogeneity in the population due to vaccination by separating the population into non-vaccinated and vaccinated subpopulations. These subpopulations are then further divided into compartments representing the disease states of individuals. From this framework, a corresponding stochastic model is developed,

which reflects the randomness inherent in disease transmission during the early phase of an epidemic, when incidence is still low.

2.1. Formulation of the Deterministic Model

The formulation of the model is based on the Kermack-McKendrick epidemic model [26], which introduces a compartmental division of the population into susceptible, infected (or infectious), and recovered groups, known as the SIR model. This current study broadens the model to incorporate additional characteristics found in individuals affected by COVID-19, including those who have received vaccinations. Initially, individuals are categorized into non-vaccinated and vaccinated groups, indicated by the subscripts n and v respectively. Consequently, the entire population $N(t)$ can be expressed as $N(t) = N_n(t) + N_v(t)$, where $N_n(t)$ is the total number of non-vaccinated individuals and $N_v(t)$ represents the total vaccinated count. Within the non-vaccinated group, individuals are further classified as susceptible $(S_n(t))$, exposed $(E_n(t))$, infectious $(I_n(t))$, and recovered $(R_n(t))$. Similarly, the vaccinated population is divided into parallel categories, using the subscript v instead of n. Thus, we have the total numbers of the non-vaccinated and vaccinated populations as follows:

$$N_n(t) = S_n(t) + E_n(t) + I_n(t) + R_n(t) \quad \text{and} \quad N_v(t) = S_v(t) + E_v(t) + I_v(t) + R_v(t).$$

In the above equation, t represents a function that depends on time. However, for convenience, S_n will be used instead of $S_n(t)$, S_v instead of $S_v(t)$, and so forth from this point forward.

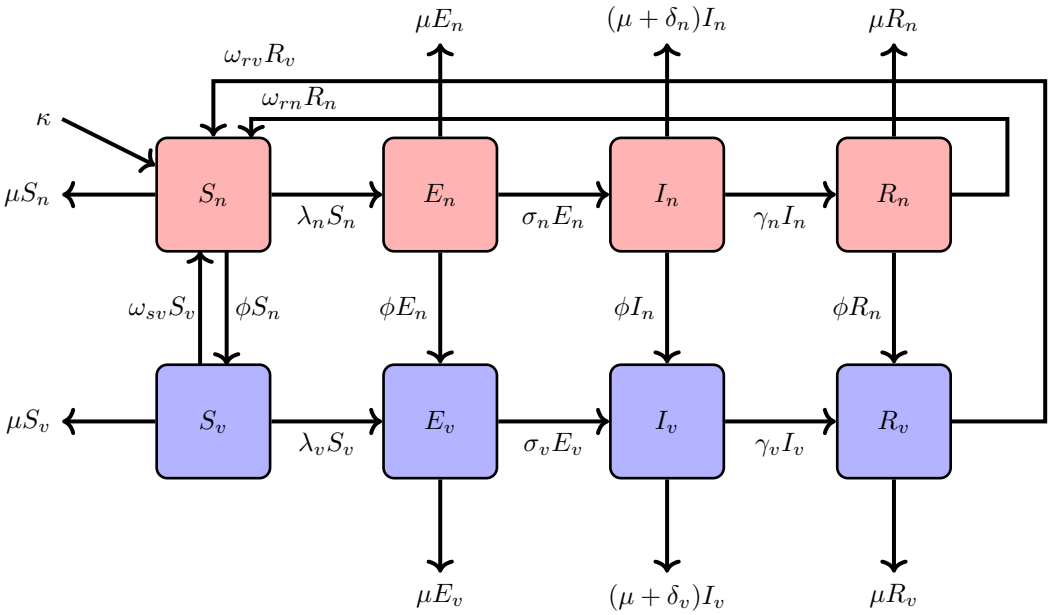

Figure 1. Schematic Diagram of the Epidemiological Model.

The model is structured to capture the key epidemiological features of COVID-19, accounting for vaccination status and waning immunity. New individuals are introduced into the population as non-vaccinated susceptibles (S_n) at a rate κ. These individuals can then be vaccinated, transitioning to the vaccinated susceptible class (S_v) at a rate ϕ, or they can become infected. Upon infection (at rates λ_n for S_n and λ_v for S_v), individuals move to their respective exposed compartments, E_n or E_v. These states represent the latency period of COVID-19, which is well-documented in clinical studies [12]. During this time, individuals are infected but not yet capable of transmitting the virus. Following the latency period, individuals progress at rates σ_n and σ_v to the infectious compartments, I_n and I_v. The term 'infectious' is explicitly used for these classes to differentiate them from the 'exposed' classes, which are infected but not yet transmitting the virus. After an infectious period, individuals

recover at rates γ_n and γ_v, moving to the recovered compartments R_n and R_v. A critical feature of COVID-19 epidemiology is temporary immunity [12]. This is modeled by having recovered individuals return to their respective susceptible classes, with natural immunity waning at a rate ω_{rn} ($R_n \to S_n$) and hybrid immunity waning at a rate ω_{rv} ($R_v \to S_v$). Therefore, 'recovered' is used rather than 'removed' to emphasize that immunity is not permanent. Finally, all compartments are subject to a natural death rate μ.

In recognition of the therapeutic benefits of vaccination, the rates of disease progression (σ), recovery (γ), and disease-induced mortality differ between non-vaccinated and vaccinated populations. The model applies the vaccination rate ϕ to all non-vaccinated compartments (S_n, E_n, I_n, and R_n). This structure is a deliberate simplification intended to capture the operational reality of a mass vaccination campaign operating with imperfect information, where it is not feasible to screen for prior infection or current asymptomatic status. The $S_n \to S_v$ transition is the primary pathway, representing the vaccination of known-susceptible or healthy-identifying individuals. The $R_n \to R_v$ pathway accounts for the vaccination of recovered individuals, a process supported by public health policy and the known immunological advantages of vaccinating previously infected persons. Vaccination post-recovery has been shown to elicit robust 'hybrid immunity', leading to stronger and more durable protection than infection-acquired immunity alone [27]. This pathway is a standard inclusion in many COVID-19 vaccination models [28]. The transitions $E_n \to E_v$ and $I_n \to I_v$ represent the vaccination of individuals who are in the incubation period (exposed) or are asymptomatically infectious. In a real-world campaign, these individuals are typically unidentifiable from the susceptible population. During the initial vaccine rollout, official public health guidance from bodies like the U.S. Advisory Committee on Immunization Practices (ACIP) explicitly recommended against virological screening for current infection prior to vaccination, as this would create unnecessary barriers and critically slow the campaign [29]. This modeling structure therefore assumes that these individuals are vaccinated at the same rate as the general population. The transition to a vaccinated compartment (e.g., E_v or I_v) is a structural one, reflecting that the individual has received the vaccine. The potential benefits of vaccination, such as a faster recovery (a higher γ_v) or reduced mortality (a lower δ_v), are then captured by the distinct parameters governing the vaccinated track.

The summary above, combined with the schematic diagram in Figure 1 and the variable and parameter explanations in Table 1, results in the following system of ordinary differential equations:

$$\frac{dS_n(t)}{dt} = \kappa + \omega_{sv}S_v + \omega_{rn}R_n + \omega_{rv}R_v - (\lambda_n + \phi + \mu)S_n, \tag{1a}$$

$$\frac{dS_v(t)}{dt} = \phi S_n - (\lambda_v + \mu + \omega_{sv})S_v, \tag{1b}$$

$$\frac{dE_n(t)}{dt} = \lambda_n S_n - (\sigma_n + \mu + \phi)E_n, \tag{1c}$$

$$\frac{dE_v(t)}{dt} = \lambda_v S_v - (\sigma_v + \mu)E_v + \phi E_n, \tag{1d}$$

$$\frac{dI_n(t)}{dt} = \sigma_n E_n - (\mu + \gamma_n + \delta_n + \phi)I_n, \tag{1e}$$

$$\frac{dI_v(t)}{dt} = \sigma_v E_v - (\mu + \gamma_v + \delta_v)I_v + \phi I_n, \tag{1f}$$

$$\frac{dR_n(t)}{dt} = \gamma_n I_n - (\mu + \phi + \omega_{rn})R_n, \tag{1g}$$

$$\frac{dR_v(t)}{dt} = \gamma_v I_v + \phi R_n - (\mu + \omega_{rv})R_v. \tag{1h}$$

The vaccine-structured model is described by equations (1a)- (1h), with the forces of infection λ_n and λ_v, given by:

$$\lambda_n = \frac{(\beta_{nn} I_n + \beta_{nv} I_v)}{N} \quad \text{and} \quad \lambda_v = \frac{(\beta_{vv} I_v + \beta_{vn} I_n)}{N}.$$

All parameters are positive and the initial conditions of the system are as follows;

$$\begin{aligned} &S_n(0) = S_{n_0} > 0, \quad S_v(0) = S_{v_0} \geq 0, \quad E_n(0) = E_{n_0} \geq 0, \quad E_v(0) = E_{v_0} \geq 0, \\ &I_n(0) = I_{n_0} \geq 0, \quad I_v(0) = I_{v_0} \geq 0, \quad R_n(0) = R_{n_0} \geq 0 \quad \text{and} \quad R_v(0) = R_{v_0} \geq 0. \end{aligned}$$

To determine the rate of change in the total human population, we take the derivative of N along the systems of equations (1a)- (1h),

$$\begin{aligned} N &= S_n + S_v + E_n + E_v + I_n + I_v + R_n + R_v, \\ \implies \frac{dN}{dt} &= \frac{dS_n}{dt} + \frac{dS_v}{dt} + \frac{dE_n}{dt} + \frac{dE_v}{dt} + \frac{dI_n}{dt} + \frac{dI_v}{dt} + \frac{dR_n}{dt} + \frac{dR_v}{dt}, \\ \implies \frac{dN}{dt} &= (\kappa + \omega_{sv} S_v + \omega_{rn} R_n + \omega_{rv} R_v - \lambda_n S_n - (\phi + \mu) S_n) \\ &+ (\phi S_n - \lambda_v S_v - (\mu + \omega_{sv}) S_v) + (\lambda_n S_n - (\sigma_n + \mu + \phi) E_n) \\ &+ (\lambda_v S_v - (\sigma_v + \mu) E_v + \phi E_n) + (\sigma_n E_n - (\mu + \gamma_n + \delta_n + \phi) I_n) \\ &+ (\sigma_v E_v - (\mu + \gamma_v + \delta_v) I_v + \phi I_n) + (\gamma_n I_n - (\mu + \phi + \omega_{rn}) R_n) \\ &+ (\gamma_v I_v + \phi R_n - (\mu + \omega_{rv}) R_v). \end{aligned}$$

Simplifying, the result follows,

$$\frac{dN}{dt} = \kappa - \mu N - \delta_n I_n - \delta_v I_v. \tag{2}$$

Table 1. Variables and Parameters in the vaccine-structured SEIR Model.

Notation	Description
S_n	Non-vaccinated susceptible individuals.
S_v	Vaccinated susceptible individuals.
E_n	Non-vaccinated exposed individuals.
E_v	Vaccinated exposed individuals.
I_n	Non-vaccinated infectious individuals.
I_v	Vaccinated infectious individuals.
R_n	Non-vaccinated recovered individuals.
R_v	Vaccinated recovered individuals.
κ	Recruitment rate.
σ_n, σ_v	Progression rate for E_n and E_v respectively.
$\beta_{nn}, \beta_{nv}, \beta_{vv}, \beta_{vn}$	Infection rates for S_n and S_v by I_n and I_v and their interactions.
γ_n, γ_v	Recovery rates of I_n and I_v respectively.
μ	Natural death rate.
δ_n, δ_v	Disease induced death rate of I_n and I_v respectively.
$\omega_{sv}, \omega_{rn}, \omega_{rv}$	Vaccine derived, natural and hybrid immunity waning rates respectively.
ϕ	Vaccination rate

2.1.1. Basic Model Properties

Positivity of Solutions The positivity of solutions refers to the non-negativity of the solutions of the model (1). This property is crucial for the equation to be epidemiologically meaningful, as it ensures that all state variables remain non-negative for all time t. Consider the following theorem:

Theorem 2.1
Given the initial conditions $S_n(0) = S_{n_0} > 0, S_v(0) = S_{v_0} \geq 0, E_n(0) = E_{n_0} \geq 0, E_v(0) = E_{v_0} \geq 0, I_n(0) = I_{n_0} \geq 0, I_v(0) = I_{v_0} \geq 0, R_n(0) = R_{n_0} \geq 0$ and $R_v(0) = R_{v_0} \geq 0$. The solutions $(S_n(t), S_v(t), E_n(t), E_v(t), I_n(t), I_v(t), R_n(t), R_v(t))$ of the system 1 are all positive for all $t \geq 0$.

Proof
From equation (1a) in the system of equations (1), it is assumed that

$$\frac{dS_n}{dt} = \kappa + \omega_{sv}S_v + \omega_{rn}R_n + \omega_{rv}R_v - (\lambda_n + \phi + \mu)S_n \geq -(\lambda_n + \mu + \phi)S_n.$$

Therefore,

$$\frac{dS_n}{dt} \geq -(\lambda_n + \mu + \phi)S_n. \tag{3}$$

Integrating (3) by separating the variables gives

$$S_n(t) \geq S(0)e^{-(\lambda_n+\mu+\phi)t} \geq 0, \forall t. \tag{4}$$

Using the same method for the remaining differential equations (1b) to (1h), we have, respectively:

$$S_v(t) \geq S_v(0)e^{-(\lambda_v+\mu+\omega_{sv})t} \geq 0, \tag{5}$$
$$E_n(t) \geq E_n(0)e^{-(\mu+\sigma_n+\phi)t} \geq 0, \tag{6}$$
$$E_v(t) \geq E_v(0)e^{-(\mu+\sigma_v)t} \geq 0, \tag{7}$$
$$I_n(t) \geq I_n(0)e^{-(\mu+\gamma_n+\delta_n+\phi)t} \geq 0, \tag{8}$$
$$I_v(t) \geq I_v(0)e^{-(\mu+\gamma_v+\delta_v)t} \geq 0, \tag{9}$$
$$R_n(t) \geq R_n(0)e^{-(\mu+\phi+\omega_{rn})t} \geq 0, \tag{10}$$
$$R_v(t) \geq R_v(0)e^{-(\mu+\omega_{rv})t} \geq 0, \tag{11}$$
$$\tag{12}$$

for all t.

□

Feasibility of the Model (Invariance) The feasibility of the model describes the region in which the solutions of the system of equations is biologically meaningful.

Theorem 2.2
Suppose equation (2) holds, every solution of the model in the system of equations (1) with initial conditions in $\mathbb{R}^8_+$ approaches and stays in the compact set $\mathcal{U}$ as $t \to \infty$. Then the feasible solution, which is a positively invariant set of the model, is given by

$$\mathcal{U} = \{(S_n, S_v, E_n, E_v, I_n, I_v, R_n, R_v) \in \mathbb{R}^8_+ : N(t) \leq \frac{\kappa}{\mu}\}.$$

Proof
From equation (2) we have

$$\frac{dN(t)}{dt} = \kappa - \mu N(t) - \delta_n I_n - \delta_v I_v,$$

thus

$$\frac{dN(t)}{dt} \leq \kappa - \mu N(t). \tag{13}$$

Multiplying by the integrating factor and using Birkoff and Rota's theorem [30] follows

$$\begin{aligned} e^{ut}\frac{dN(t)}{dt} &\leq e^{ut}\kappa - \mu N(t)e^{ut}, \\ \implies N(t) &\leq \frac{\kappa}{\mu} + Ce^{-ut}. \end{aligned} \tag{14}$$

Hence, as $t \to \infty$,

$$N(t) \leq \frac{\kappa}{\mu},$$

which implies $0 \leq N \leq \frac{\kappa}{\mu}$. Then the trajectories of the model (1) are bounded in the region $\mathcal{U}$. This completes the proof. □

2.1.2. *Equilibria and Stability Analysis of the Model*

Disease Free Equilibrium (DFE) To determine the steady states of model (1), the right-hand side of the system is set to zero:

$$\begin{aligned} &\kappa + \omega_{sv}S_v^0 + \omega_{rn}R_n^0 + \omega_{rv}R_v^0 - (\lambda_n + \phi + \mu)S_n^0 = 0, \\ &\phi S_n^0 - (\lambda_v + \mu + \omega_{sv})S_v^0 = 0, \\ &\lambda_n S_n^0 - (\sigma_n + \mu + \phi)E_n^0 = 0, \\ &\lambda_v S_v^0 - (\sigma_v + \mu)E_v^0 + \phi E_n^0 = 0, \\ &\sigma_n E_n^0 - (\mu + \gamma_n + \delta_n + \phi)I_n^0 = 0, \\ &\sigma_v E_v^0 - (\mu + \gamma_v + \delta_v)I_v^0 + \phi I_n^0 = 0, \\ &\gamma_n I_n^0 - (\mu + \phi + \omega_{rn})R_n^0 = 0, \\ &\gamma_v I_v^0 + \phi R_n^0 - (\mu + \omega_{rv})R_v^0 = 0. \end{aligned} \tag{15}$$

Substituting $E_n^0 = E_v^0 = I_n^0 = I_v^0 = R_n^0 = R_v^0 = 0$ into Equation (15) gives:

$$\begin{aligned} &\kappa + \omega_{sv}S_v^0 - (\phi + \mu)S_n^0 = 0, \\ &\phi S_n^0 - (\mu + \omega_{sv})S_v^0 = 0, \end{aligned} \tag{16}$$

hence,

$$S_n^0 = \frac{\kappa(\omega_{sv} + \mu)}{\mu(\omega_{sv} + \phi + \mu)} \quad \text{and} \quad S_v^0 = \frac{\phi\kappa}{\mu(\omega_{sv} + \phi + \mu)}.$$

Thus, the disease-free equilibrium (DFE) is given by

$$DFE = \left(\frac{\kappa(\omega_{sv} + \mu)}{\mu(\omega_{sv} + \phi + \mu)}, \frac{\phi\kappa}{\mu(\omega_{sv} + \phi + \mu)}, 0, 0, 0, 0, 0, 0 \right). \tag{17}$$

Further the initial total population size is $N^0 = S_n^0 + S_v^0$.

Basic Reproduction Number The basic reproduction number, $\mathscr{R}_0$, represents the average number of secondary infections generated by a single infectious individual in a wholly susceptible population. Following the next generation matrix approach described by [31], $\mathscr{R}_0$ is determined for model (1).

Let $\mathscr{F}$ and $\mathscr{V}$ denote the rate of appearance of new infections and the rate of transfer among compartments, respectively, evaluated at the disease-free equilibrium $Y^0 = (0, 0, 0, 0, S_n^0, S_v^0, 0, 0)$. The Jacobian matrices F and V, derived from $\mathscr{F}$ and $\mathscr{V}$, are used to form the next generation matrix FV^{-1} (see Appendix A.1). The spectral radius of this matrix provides the basic reproduction number:

$$\mathscr{R}_0 = \rho(FV^{-1}) = \tfrac{1}{2}(g_{11} + g_{22}) + \tfrac{1}{2}\sqrt{g_{11}^2 + 4g_{12}g_{21} - 2g_{11}g_{22} + g_{22}^2},$$

where all positive parameters g_{ij} $(i, j = 1, 2, 3, 4)$ are defined in Appendix A.1. The first term in the expression represents within-group transmission among the non-vaccinated and vaccinated populations, while the second term captures cross-group transmission effects. The detailed derivation of the Jacobian matrices, their components, and the full expressions for g_{ij} are provided in Appendix A.1.

Disease free Equilibrium and Stability The stability analysis of the disease-free equilibrium assists in understanding the long-term dynamics of a disease. This analysis is dependent on the basic reproduction number, as explained by the theorem below.

Theorem 2.3
The disease-free equilibrium point X^0 is locally asymptotically stable when $\mathscr{R}_0 < 1$, while X^0 is unstable when $\mathscr{R}_0 > 1$.

Proof
The Jacobian matrix J evaluated at the disease-free equilibrium point $X = X^0$ is

$$J_0 = J_{X=X^0} = \begin{pmatrix} M_1 & M_2 \\ M_3 & M_4 \end{pmatrix} \tag{18}$$

with

$$M_1 = \begin{pmatrix} J_{EE} & J_{EI} \\ J_{IE} & J_{II} \end{pmatrix}, M_2 = \begin{pmatrix} J_{ES} & J_{ER} \\ J_{IS} & J_{IR} \end{pmatrix}$$

$$M_3 = \begin{pmatrix} J_{SE} & J_{SI} \\ J_{RE} & J_{RI} \end{pmatrix}, M_4 = \begin{pmatrix} J_{SS} & J_{SR} \\ J_{RS} & J_{RR} \end{pmatrix}$$

where the detailed calculations for each component sub-matrix (e.g., $J_{EE}, J_{EI}, \ldots, J_{RR}$) at the DFE are presented in Appendix A.2.
Therefore, with $\mathscr{C}_n = (\beta_{nn} I_n + \beta_{nv} I_v)$ and $\mathscr{C}_v = (\beta_{vv} I_v + \beta_{vn} I_n)$, the matrices M_1 and M_4 at the DFE simplify to:

$$M_1 = \begin{pmatrix} -(\sigma_n + \mu + \phi) & 0 & \frac{\beta_{nn} S_n^0}{N^0} & \frac{\beta_{nv} S_n^0}{N^0} \\ \phi & -(\sigma_v + \mu) & \frac{\beta_{vn} S_v^0}{N^0} & \frac{\beta_{vv} S_v^0}{N^0} \\ \sigma_n & 0 & -(\mu + \gamma_n + \delta_n + \phi) & 0 \\ 0 & \sigma_v & \phi & -(\mu + \gamma_v + \delta_v) \end{pmatrix}$$

$$= \begin{pmatrix} A_n - \phi & 0 & B_{nn} & B_{nv} \\ \phi & -A_v & B_{vn} & B_{vv} \\ C_n & 0 & -D_n - \phi & 0 \\ 0 & C_v & \phi & -D_v \end{pmatrix}$$

and

$$M_4 = \begin{pmatrix} \mu + \phi & \omega_{sv} & 0 & 0 \\ -\phi & \mu + \omega_{sv} & 0 & 0 \\ 0 & 0 & \mu + \omega_{rn} + \phi & 0 \\ 0 & 0 & -\phi & \mu + \omega_{rv} \end{pmatrix}$$

The Jacobian Matrix J_0 has a characteristic equation that is the product of the characteristic equations of M_1 and M_4. Since the eigenvalues of M_4 are negative, the remaining four eigenvalues depend on the characteristic equation of M_1, which is

$$\lambda^4 + H_1\lambda^3 + H_2\lambda^2 + H_3\lambda + H^4 = 0 \tag{19}$$

where

$$\begin{aligned} H_1 &= (A_n + A_v + D_n + D_v)\,, \\ H_2 &= (A_nA_v - B_{nn}C_n - B_{vv}C_v + A_nD_n + A_vD_n + A_nD_v + A_vD_v + D_nD_v)\,, \\ &= A_nA_v + A_nD_v + A_vD_n + (A_nD_n - B_{nn}C_n) + (A_vD_v - B_{vv}C_v) + D_nD_v, \\ H_3 &= (-A_vB_{nn}C_n - A_nB_{vv}C_v + A_nA_vD_n - B_{vv}C_vD_n + A_nA_vD_v - B_{nn}C_nD_v + A_nD_nD_v + A_vD_nD_v)\,, \\ &= (A_n + D_n)(A_vD_v - B_{vv}C_v) + (A_v + D_v)(A_nD_n - B_{nn}C_n), \\ H_4 &= B_{nv}B_{vn}C_1C_v + B_{nn}B_{vv}C_nC_v - A_nB_{vv}C_vD_n - A_vB_{nn}C_nD_v + A_nA_vD_nD_v, \\ &= -A_nA_vD_nD_v + (B_{nn}B_{vv} - B_{nv}B_{vn})C_nC_v + A_vD_v(A_nD_n - B_{nn}C_n) + A_nD_n(A_vD_v - B_{vv}C_v). \end{aligned} \tag{20}$$

According to Descartes' rule of signs [32, 33, 34], the number of negative roots for the characteristic equation (19) is equal to the number of sign changes in its coefficients. Consequently, equation (19) contains four negative roots if

$$H_1 > 0, \quad H_2 > 0, \quad H_3 > 0, \quad H_4 > 0. \tag{21}$$

The conclusions presented in (21) remain valid provided that the parameters in (20) meet the following conditions:

$$\begin{aligned} &(A_nD_n - B_{nn}C_n) > 0, \quad (A_vD_v - B_{vv}C_v) > 0, \\ &-A_nA_vD_nD_v + (B_{nn}B_{vv} - B_{nv}B_{vn})C_nC_v > 0. \end{aligned} \tag{22}$$

□

Endemic Equilibrium and Stability The endemic equilibrium is a state in which a disease is present in the population; however, the disease does not die out or increase exponentially. This occurs when the reproduction number is greater than one. Considering model (1), by setting the right-hand side to zero, the endemic equilibrium satisfies the following equalities:

$$\kappa + \omega_{sv}S_v^* + \omega_{rn}R_n^* + \omega_{rv}R_v^* - (\beta_{nn}I_n^* + \beta_{nv}I_v^*)\frac{S_n^*}{N^*} - (\phi+\mu)S_n^* = 0, \tag{23a}$$

$$\phi S_n^* - (\beta_{vv}I_v^* + \beta_{vn}I_n^*)\frac{S_v^*}{N^*} - (\mu+\omega_{sv})S_v^* = 0, \tag{23b}$$

$$(\beta_{nn}I_n^* + \beta_{nv}I_v^*)\frac{S_n^*}{N^*} - (\sigma_n+\mu+\phi)E_n^* = 0, \tag{23c}$$

$$(\beta_{vv}I_v^* + \beta_{vn}I_n^*)\frac{S_v^*}{N^*} - (\sigma_v+\mu)E_v^* + \phi E_n^* = 0, \tag{23d}$$

$$\sigma_n E_n^* - (\mu+\gamma_n+\delta_n+\phi)I_n^* = 0, \tag{23e}$$

$$\sigma_v E_v^* - (\mu+\gamma_v+\delta_v)I_v^* + \phi I_n^* = 0, \tag{23f}$$

$$\gamma_n I_n^* - (\mu+\phi+\omega_{rn})R_n^* = 0, \tag{23g}$$

$$\gamma_v I_v^* + \phi R_n^* - (\mu+\omega_{rv})R_v^* = 0, \tag{23h}$$

Combining (23a) with (23c) and (23b) with (23d) results in;

$$\begin{aligned} &\kappa + \omega_{sv}S_v^* + \omega_{rn}R_n^* + \omega_{rv}R_v^* - (\phi+\mu)S_n^* - (\sigma_n+\mu+\phi)E_n^* = 0,\\ &\phi S_n^* - (\mu+\omega_{sv})S_v^* - (\sigma_v+\mu)E_v^* + \phi E_n^* = 0. \end{aligned} \tag{24}$$

The aim is to represent S_n, S_v, I_n, I_v, R_n and R_v as linear functions of E_n and E_v. Through algebraic manipulation (detailed in Appendix A.3), we obtain expressions for the state variables (Eq. 77) and the total population N^* (Eq. 78).
Together with these expressions, equations (23c) and (23d) can be rewritten as:

$$f_n(E_n^*, E_v^*) = a_{nn}E_n^{*2} + 2(a_n + a_{nv}E_v^*)E_n^* + (2a_vE_v^* + a_{vv}E_v^{*2}) = 0, \tag{25}$$

$$f_v(E_n^*, E_v^*) = b_{nn}E_n^{*2} + 2(b_n + b_{nv}E_v^*)E_n^* + (2b_vE_v^* + b_{vv}E_v^{*2}) = 0, \tag{26}$$

with the coefficients $(a_{nn}, a_n, \ldots, b_{vv})$ being complex combinations of the model parameters, as detailed in Appendix A.3.
The endemic equilibrium point X^* is located at the intersection of $f_n(E_n^*, E_v^*) = 0$ and $f_v(E_n^*, E_v^*) = 0$.

Theorem 2.4 (Existence and Uniqueness of the Endemic Equilibrium points.)
If one of the following conditions hold, then model (1) has a unique endemic equilibrium point X^*.

1. $a_{nn} \neq 0$ and $b_{nn} \neq 0$,
2. $a_{nn} = 0$ and $b_{nn} \neq 0$,
3. $a_{nn} \neq 0$ and $b_{nn} = 0$,
4. $a_{nn} = b_{nn} = 0$.

Proof
Case 1: $a_{nn} \neq 0$ **and** $b_{nn} \neq 0$
Let E_n^* be considered as a function of E_v^*. From equation $f_n(E_n^*, E_v^*) = 0$, which is $a_{nn}E_n^{*2} + 2(a_n + a_{nv}E_v^*)E_n^* + (2a_vE_v^* + a_{vv}E_v^{*2}) = 0$, for a unique solution for E_n^* (given E_v^*), we require the discriminant Δ_n of this quadratic in E_n^* to be zero. Let $A = a_{nn}$, $B = 2(a_n + a_{nv}E_v^*)$, and $C = (2a_vE_v^* + a_{vv}E_v^{*2})$. The discriminant is $\Delta_n = B^2 - 4AC$. Setting $\Delta_n = 0$:

$$(2(a_n + a_{nv}E_v^*))^2 - 4a_{nn}(2a_vE_v^* + a_{vv}E_v^{*2}) = 0. \tag{27}$$

If $\Delta_n = 0$, the unique solution for E_n^* is $E_n^* = -\frac{B}{2A}$:

$$E_n^* = -\frac{2(a_n + a_{nv}E_v^*)}{2a_{nn}} = -\frac{a_n + a_{nv}E_v^*}{a_{nn}}. \tag{28}$$

Similarly, from equation $f_v(E_n^*, E_v^*) = 0$, which is $b_{nn}E_n^{*2} + 2(b_n + b_{nv}E_v^*)E_n^* + (2b_vE_v^* + b_{vv}E_v^{*2}) = 0$, for a unique solution for E_n^* (given E_v^*), we require its discriminant Δ_v to be zero. Let $A' = b_{nn}$, $B' = 2(b_n + b_{nv}E_v^*)$, and $C' = (2b_vE_v^* + b_{vv}E_v^{*2})$. Setting $\Delta_v = (B')^2 - 4A'C' = 0$:

$$(2(b_n + b_{nv}E_v^*))^2 - 4b_{nn}(2b_vE_v^* + b_{vv}E_v^{*2}) = 0. \tag{29}$$

If $\Delta_v = 0$, the unique solution for E_n^* is $E_n^* = -\frac{B'}{2A'}$:

$$E_n^* = -\frac{2(b_n + b_{nv}E_v^*)}{2b_{nn}} = -\frac{b_n + b_{nv}E_v^*}{b_{nn}}. \tag{30}$$

Equating the expressions for E_n^* from (28) and (30):

$$-\frac{a_n + a_{nv}E_v^*}{a_{nn}} = -\frac{b_n + b_{nv}E_v^*}{b_{nn}}$$

Thus,

$$E_v^* = \frac{a_{nn}b_n - b_{nn}a_n}{a_{nv}b_{nn} - a_{nn}b_{nv}}, \tag{31}$$

provided that the denominator $D_{ab} = a_{nv}b_{nn} - a_{nn}b_{nv} \neq 0$. Substituting this E_v^* back into equation (28):

$$E_n^* = -\frac{a_n + a_{nv}E_v^*}{a_{nn}} = -\frac{1}{a_{nn}}\left(a_n + a_{nv}\frac{a_{nn}b_n - b_{nn}a_n}{a_{nv}b_{nn} - a_{nn}b_{nv}}\right)$$

Thus,

$$E_n^* = \frac{a_nb_{nv} - a_{nv}b_n}{a_{nv}b_{nn} - a_{nn}b_{nv}}. \tag{32}$$

For these unique values of E_n^* and E_v^* to be the endemic equilibrium, they must satisfy the conditions $\Delta_n = 0$ and $\Delta_v = 0$. Substituting E_v^* from (31) into (27) yields the first condition on the coefficients:

$$(a_n + a_{nv}E_v^*)^2 = a_{nn}(2a_vE_v^* + a_{vv}E_v^{*2}).$$

Noting from (28) that $a_n + a_{nv}E_v^* = -a_{nn}E_n^*$, this condition becomes:

$$(-a_{nn}E_n^*)^2 = a_{nn}(2a_vE_v^* + a_{vv}E_v^{*2})$$

Since $a_{nn} \neq 0$:

$$a_{nn}(E_n^*)^2 = 2a_vE_v^* + a_{vv}(E_v^*)^2. \tag{33}$$

Similarly, substituting E_v^* from (31) into (29) yields the second condition. Noting from (30) that $b_n + b_{nv}E_v^* = -b_{nn}E_n^*$, this condition becomes:

$$(-b_{nn}E_n^*)^2 = b_{nn}(2b_vE_v^* + b_{vv}E_v^{*2})$$

Since $b_{nn} \neq 0$:

$$b_{nn}(E_n^*)^2 = 2b_vE_v^* + b_{vv}(E_v^*)^2. \tag{34}$$

Therefore, for a unique endemic equilibrium (E_n^*, E_v^*) given by (32) and (31) to exist under the assumption that E_n^* is a unique root of both quadratics (for a given E_v^*), the coefficients a_{ij}, b_{ij} must satisfy conditions (33) and (34), where E_n^* and E_v^* are themselves expressions in terms of these coefficients as defined in (32) and (31). Substituting the expressions for E_n^* and E_v^* into (33) and (34): Let $D_{ab} = a_{nv}b_{nn} - a_{nn}b_{nv}$. Condition 1 becomes:

$$a_{nn}\left(\frac{a_nb_{nv} - a_{nv}b_n}{D_{ab}}\right)^2 = 2a_v\left(\frac{a_{nn}b_n - b_{nn}a_n}{D_{ab}}\right) + a_{vv}\left(\frac{a_{nn}b_n - b_{nn}a_n}{D_{ab}}\right)^2$$

Multiplying by D_{ab}^2 (assuming $D_{ab} \neq 0$ for E_n^*, E_v^* to be well-defined):

$$a_{nn}(a_nb_{nv} - a_{nv}b_n)^2 = 2a_v(a_{nn}b_n - b_{nn}a_n)D_{ab} + a_{vv}(a_{nn}b_n - b_{nn}a_n)^2. \tag{35}$$

Condition 2 becomes:

$$b_{nn}\left(\frac{a_n b_{nv} - a_{nv} b_n}{D_{ab}}\right)^2 = 2b_v\left(\frac{a_{nn} b_n - b_{nn} a_n}{D_{ab}}\right) + b_{vv}\left(\frac{a_{nn} b_n - b_{nn} a_n}{D_{ab}}\right)^2$$

Multiplying by D_{ab}^2:

$$b_{nn}(a_n b_{nv} - a_{nv} b_n)^2 = 2b_v(a_{nn} b_n - b_{nn} a_n)D_{ab} + b_{vv}(a_{nn} b_n - b_{nn} a_n)^2. \tag{36}$$

These two equations, (35) and (36), along with $a_{nn} \neq 0$, $b_{nn} \neq 0$, and $D_{ab} \neq 0$, define the conditions for this specific type of unique endemic equilibrium. Additionally, E_n^* and E_v^* must be positive for epidemiological relevance.

Case 2: $a_{nn} = 0$ **and** $b_{nn} \neq 0$
With $a_{nn} = 0$, equation (25) simplifies to:

$$2(a_n + a_{nv}E_v^*)E_n^* + (2a_v E_v^* + a_{vv}E_v^{*2}) = 0. \tag{37}$$

Equation (26) is $b_{nn}E_n^{*2} + 2(b_n + b_{nv}E_v^*)E_n^* + (2b_v E_v^* + b_{vv}E_v^{*2}) = 0$. For a unique solution for E_n^* from this quadratic (given E_v^*), its discriminant Δ_v must be zero:

$$\Delta_v = (2(b_n + b_{nv}E_v^*))^2 - 4b_{nn}(2b_v E_v^* + b_{vv}E_v^{*2}) = 0. \tag{38}$$

If $\Delta_v = 0$, the unique solution for E_n^* is:

$$E_n^* = -\frac{b_n + b_{nv}E_v^*}{b_{nn}}. \tag{39}$$

We substitute this expression for E_n^* into the simplified equation (37):

$$2(a_n + a_{nv}E_v^*)\left(-\frac{b_n + b_{nv}E_v^*}{b_{nn}}\right) + (2a_v E_v^* + a_{vv}E_v^{*2}) = 0.$$

Multiplying by $-b_{nn}$ (noting $b_{nn} \neq 0$):

$$2(a_n + a_{nv}E_v^*)(b_n + b_{nv}E_v^*) - b_{nn}(2a_v E_v^* + a_{vv}E_v^{*2}) = 0.$$

Expanding and rearranging terms yields a quadratic equation in E_v^*:

$$(2a_{nv}b_{nv} - a_{vv}b_{nn})E_v^{*2} + 2(a_n b_{nv} + a_{nv}b_n - a_v b_{nn})E_v^* + 2a_n b_n = 0. \tag{40}$$

Let $P_v = 2a_{nv}b_{nv} - a_{vv}b_{nn}$, $Q_v = 2(a_n b_{nv} + a_{nv}b_n - a_v b_{nn})$, and $R_v = 2a_n b_n$. So, equation (40) is $P_v E_v^{*2} + Q_v E_v^* + R_v = 0$.
Subcase 2.1: $P_v \neq 0$ For a unique solution for E_v^* from $P_v E_v^{*2} + Q_v E_v^* + R_v = 0$, its discriminant must be zero:

$$Q_v^2 - 4P_v R_v = 0. \tag{41}$$

If this condition holds, then E_v^* is uniquely given by $E_v^* = -Q_v/(2P_v)$. This value of E_v^* must also satisfy the condition (38). Substituting $E_v^* = -Q_v/(2P_v)$ into (38) gives the second condition:

$$\left(2\left(b_n + b_{nv}\left(-\frac{Q_v}{2P_v}\right)\right)\right)^2 - 4b_{nn}\left(2b_v\left(-\frac{Q_v}{2P_v}\right) + b_{vv}\left(-\frac{Q_v}{2P_v}\right)^2\right) = 0. \tag{42}$$

The conditions for a unique endemic equilibrium are $a_{nn} = 0$, $b_{nn} \neq 0$, $P_v \neq 0$, equation (41), and equation (42). The resulting $E_v^* = -Q_v/(2P_v)$ and $E_n^* = -(b_n + b_{nv}E_v^*)/b_{nn}$ must be positive.

Subcase 2.2: $P_v = 0$ Equation (40) becomes $Q_v E_v^* + R_v = 0$. If $Q_v \neq 0$, then E_v^* is uniquely given by $E_v^* = -R_v/Q_v$. This value of E_v^* must satisfy the condition (38). Substituting $E_v^* = -R_v/Q_v$ into (38) gives the condition:

$$\left(2\left(b_n + b_{nv}\left(-\frac{R_v}{Q_v}\right)\right)\right)^2 - 4b_{nn}\left(2b_v\left(-\frac{R_v}{Q_v}\right) + b_{vv}\left(-\frac{R_v}{Q_v}\right)^2\right) = 0. \tag{43}$$

The conditions for a unique endemic equilibrium are $a_{nn} = 0$, $b_{nn} \neq 0$, $P_v = 0$, $Q_v \neq 0$, and equation (43). The resulting $E_v^* = -R_v/Q_v$ and $E_n^* = -(b_n + b_{nv}E_v^*)/b_{nn}$ must be positive.

Case 3: $a_{nn} \neq 0$ **and** $b_{nn} = 0$
With $b_{nn} = 0$, equation (26) becomes linear in E_n^*: $2(b_n + b_{nv}E_v^*)E_n^* + (2b_vE_v^* + b_{vv}E_v^{*2}) = 0$. Assuming $2(b_n + b_{nv}E_v^*) \neq 0$, $E_n^* = -\frac{2b_vE_v^* + b_{vv}E_v^{*2}}{2(b_n + b_{nv}E_v^*)}$. Equation (25) is $a_{nn}E_n^{*2} + 2(a_n + a_{nv}E_v^*)E_n^* + (2a_vE_v^* + a_{vv}E_v^{*2}) = 0$. For a unique E_n^* (given E_v^*), its discriminant Δ_n must be zero:

$$(2(a_n + a_{nv}E_v^*))^2 - 4a_{nn}(2a_vE_v^* + a_{vv}E_v^{*2}) = 0. \tag{44}$$

If $\Delta_n = 0$, then $E_n^* = -\frac{a_n + a_{nv}E_v^*}{a_{nn}}$. Equating the two expressions for E_n^* leads to a quadratic in E_v^*: $P_v'E_v^{*2} + Q_v'E_v^* + R_v' = 0$, where $P_v' = a_{nn}b_{vv} - 2a_{nv}b_{nv}$, $Q_v' = 2(a_{nn}b_v - a_nb_{nv} - a_{nv}b_n)$, and $R_v' = -2a_nb_n$.
Subcase 3.1: $P_v' \neq 0$ For a unique E_v^*, the discriminant of $P_v'E_v^{*2} + Q_v'E_v^* + R_v' = 0$ must be zero:

$$(Q_v')^2 - 4P_v'R_v' = 0. \tag{45}$$

Then $E_v^* = -Q_v'/(2P_v')$. This E_v^* must satisfy (44):

$$\left(2\left(a_n + a_{nv}\left(-\frac{Q_v'}{2P_v'}\right)\right)\right)^2 - 4a_{nn}\left(2a_v\left(-\frac{Q_v'}{2P_v'}\right) + a_{vv}\left(-\frac{Q_v'}{2P_v'}\right)^2\right) = 0. \tag{46}$$

Conditions for unique positive endemic equilibrium: $a_{nn} \neq 0, b_{nn} = 0, P_v' \neq 0$, (45), (46), and resulting $E_n^*, E_v^* > 0$.
Subcase 3.2: $P_v' = 0$ Then $Q_v'E_v^* + R_v' = 0$. If $Q_v' \neq 0$, then $E_v^* = -R_v'/Q_v'$. This E_v^* must satisfy (44):

$$\left(2\left(a_n + a_{nv}\left(-\frac{R_v'}{Q_v'}\right)\right)\right)^2 - 4a_{nn}\left(2a_v\left(-\frac{R_v'}{Q_v'}\right) + a_{vv}\left(-\frac{R_v'}{Q_v'}\right)^2\right) = 0. \tag{47}$$

Conditions for unique positive endemic equilibrium: $a_{nn} \neq 0, b_{nn} = 0, P_v' = 0, Q_v' \neq 0$, (47), and resulting $E_n^*, E_v^* > 0$.

Case 4: $a_{nn} = 0$ **and** $b_{nn} = 0$
Both equations (25) and (26) become linear in E_n^*:

$$2(a_n + a_{nv}E_v^*)E_n^* + (2a_vE_v^* + a_{vv}E_v^{*2}) = 0, \tag{48}$$

$$2(b_n + b_{nv}E_v^*)E_n^* + (2b_vE_v^* + b_{vv}E_v^{*2}) = 0. \tag{49}$$

Assuming denominators $2(a_n + a_{nv}E_v^*)$ and $2(b_n + b_{nv}E_v^*)$ are non-zero, we can express E_n^* from both: $E_n^* = -\frac{2a_vE_v^* + a_{vv}E_v^{*2}}{2(a_n + a_{nv}E_v^*)}$ and $E_n^* = -\frac{2b_vE_v^* + b_{vv}E_v^{*2}}{2(b_n + b_{nv}E_v^*)}$. Equating these and simplifying (assuming $E_v^* \neq 0$ to divide by $2E_v^*$ in an intermediate step) leads to a quadratic in E_v^*: $P_v''E_v^{*2} + Q_v''E_v^* + R_v'' = 0$, where $P_v'' = a_{vv}b_{nv} - b_{vv}a_{nv}$, $Q_v'' = 2a_vb_{nv} + a_{vv}b_n - 2b_va_{nv} - b_{vv}a_n$, and $R_v'' = 2(a_vb_n - b_va_n)$.
Subcase 4.1: $P_v'' \neq 0$ For a unique E_v^*, the discriminant of $P_v''E_v^{*2} + Q_v''E_v^* + R_v'' = 0$ must be zero:

$$(Q_v'')^2 - 4P_v''R_v'' = 0. \tag{50}$$

Then $E_v^* = -Q_v''/(2P_v'')$. Conditions for unique positive endemic equilibrium: $a_{nn} = 0, b_{nn} = 0, P_v'' \neq 0$, equation (50), and the resulting $E_v^* > 0$ and $E_n^* > 0$ (where E_n^* is found by substituting this E_v^* into (48) or (49), ensuring denominators are non-zero and consistent).
Subcase 4.2: $P_v'' = 0$ Then $Q_v'' E_v^* + R_v'' = 0$. If $Q_v'' \neq 0$, then $E_v^* = -R_v''/Q_v''$. Conditions for unique positive endemic equilibrium: $a_{nn} = 0, b_{nn} = 0, P_v'' = 0, Q_v'' \neq 0$, and the resulting $E_v^* > 0$ and $E_n^* > 0$ (with non-zero and consistent denominators for E_n^*). □

Theorem 2.5 (Local Asymptotic Stability of the Endemic Equilibrium point)
The endemic equilibrium point X^* is locally asymptotically stable when $\mathscr{R}_1 < 0$ and unstable when $\mathscr{R}_0 > 1$.

Proof
The Jacobian matrices for the non-vaccinated and vaccinated groups are respectively constructed as follows

$$J_n = \begin{pmatrix} -\frac{\mathscr{C}_n S_n^*}{(N^*)^2} - (\sigma_n + \mu + \phi) & \frac{\beta_{nn} S_n^*}{N^*} - \frac{\mathscr{C}_n S_n^*}{(N^*)^2} & -\frac{\mathscr{C}_n S_n^*}{(N^*)^2} + \frac{\mathscr{C}_n}{N^*} & -\frac{\mathscr{C}_n}{(N^*)^2} \\ -\sigma_n & -(\mu + \gamma_n + \delta_n + \phi) & 0 & 0 \\ -\frac{\mathscr{C}_n S_n^*}{N^*} & -\frac{B_{nn}}{N^*} + \frac{\mathscr{C}_n S_n^*}{(N^*)^2} & \frac{\mathscr{C}_n S_n^*}{(N^*)^2} - \frac{\mathscr{C}_n}{N^*} - (\phi + \mu) & -\frac{\mathscr{C}_n}{(N^*)^2} \\ 0 & \gamma_n & 0 & -(\mu + \phi_v + \omega_{rn}) \end{pmatrix}, \tag{51}$$

$$J_v = \begin{pmatrix} -\frac{\mathscr{C}_v S_v^*}{(N^*)^2} - (\sigma_v + \mu) & \frac{\beta_{vv} S_v^*}{N^*} - \frac{\mathscr{C}_v S_v^*}{(N^*)^2} & -\frac{\mathscr{C}_v S_v^*}{(N^*)^2} + \frac{\mathscr{C}_v}{N^*} & -\frac{\mathscr{C}_v}{(N^*)^2} \\ -\sigma_v & \phi - (\mu + \gamma_v + \delta_v) & 0 & 0 \\ -\frac{\mathscr{C}_v S_v^*}{N^*} & -\frac{B_{vv}}{N^*} + \frac{\mathscr{C}_v S_v^*}{(N^*)^2} & \frac{\mathscr{C}_v S_v^*}{(N^*)^2} - \frac{\mathscr{C}_v}{N^*} - (\mu + \omega_{sv}) & -\frac{\mathscr{C}_v}{(N^*)^2} \\ 0 & \gamma_v & 0 & -(\mu + \omega_{rv}) \end{pmatrix}, \tag{52}$$

with $\mathscr{C}_n = (\beta_{nn} I_n^* + \beta_{nv} I_v^*)$ and $\mathscr{C}_v = (\beta_{vv} I_v^* + \beta_{vn} I_n^*)$. J_n and J_v have the same structure so we use M to denote them as follows

$$M = \begin{pmatrix} -k_{nn} & k_{nv} & k_{nq} & -k_{nr} \\ k_{vn} & -k_{vv} & 0 & 0 \\ -k_{qn} & -k_{nv} & -k_{qq} & k_{nr} \\ 0 & k_{rv} & 0 & -k_{rr} \end{pmatrix}. \tag{53}$$

The corresponding characteristic equation of M follows

$$\lambda^4 + H_1 \lambda^3 + H_2 \lambda^2 + H_3 \lambda + H_4 = 0, \tag{54}$$

with

$$H_1 = k_{nn} + k_{qq} + k_{rr} + k_{vv}, \tag{55a}$$
$$H_2 = (k_{nn}k_{vv} - k_{nv}k_{vn}) + k_{nq}k_{qn} + k_{nn}k_{qq} + k_{nn}k_{rr} + k_{qq}k_{rr} + k_{qq}k_{vv} + k_{rr}k_{vv}, \tag{55b}$$
$$H_3 = k_{qq}(k_{nn}k_{vv} - k_{nv}k_{vn}) + k_{rr}(k_{nn}k_{vv} - k_{nv}k_{vn}) + k_{nq}k_{qn}k_{rr} + k_{nn}k_{qq}k_{rr} + k_{nq}k_{nv}k_{vn} + k_{nr}k_{rv}k_{vn} + k_{nq}k_{qn}k_{vv} + k_{qq}k_{rr}k_{vv}, \tag{55c}$$
$$H_4 = k_{qq}k_{rr}(k_{nn}k_{vv} - k_{nv}k_{vv}) + k_{nr}k_{rv}(k_{qq}k_{vn} - k_{vn}k_{nq}) + k_{nq}k_{nv}k_{rr}k_{vn} + k_{nq}k_{qn}k_{rr}k_{vv}. \tag{55d}$$

Based on the Discartes rule of signs [32, 33, 34], the number of negative roots of the characteristic equation (54) is equal to the number of variations in the change in the coefficient signs, so (54) has 4 negative signs if

$$H_1 > 0, \quad H_2 > 0, \quad H_3 > 0, \quad H_4 > 0. \tag{56}$$

The conclusions in (56) hold true if the parameters meet:

$$(k_{nn}k_{vv} - k_{nv}k_{vn}) > 0, \quad (k_{qq}k_{vn} - k_{vn}k_{nq}) > 0. \tag{57}$$

To be precise, the characteristic values of the equation system within this model are entirely negative under constraint (57), thereby indicating that the endemic equilibrium point is locally asymptotically stable. □

2.2. Formulation of the Stochastic Model

Deterministic models, such as the system of ordinary differential equations presented in Section 2, describe the average behavior of a large population [14, 15]. However, the transmission of infectious diseases is an inherently stochastic process, especially when the number of infected individuals is small, as is the case at the beginning of an outbreak or when the disease is near elimination. Random events, or demographic stochasticity, related to individual births, deaths, disease transmission, progression, and recovery can lead to outcomes that differ significantly from deterministic predictions [16]. For instance, an infection may die out by chance even if the basic reproduction number ($\mathscr{R}_0$) is greater than one, an event not captured by deterministic frameworks [16, 17]. To account for these random effects, which are crucial for understanding phenomena like disease extinction and the probability of major outbreaks, a Continuous-Time Markov Chain (CTMC) model based on the assumptions of the deterministic vaccine-structured model is developed.

2.2.1. CTMC Model Development A CTMC model where time $t \in [0, \infty)$ is continuous, and the state variables representing the number of individuals in each compartment are discrete, non-negative integers is formulated. Let $S_n(t), S_v(t), E_n(t), E_v(t), I_n(t), I_v(t), R_n(t), R_v(t)$ denote these discrete-valued random variables for the non-vaccinated susceptible, vaccinated susceptible, non-vaccinated exposed, vaccinated exposed, non-vaccinated infectious, vaccinated infectious, non-vaccinated recovered, and vaccinated recovered individuals at time t, respectively. The total non-vaccinated population is $N_n(t) = S_n(t) + E_n(t) + I_n(t) + R_n(t)$, the total vaccinated population is $N_v(t) = S_v(t) + E_v(t) + I_v(t) + R_v(t)$, and the total population is $N(t) = N_n(t) + N_v(t)$.

Let

$$\vec{Y}(t) = [S_n(t), S_v(t), E_n(t), E_v(t), I_n(t), I_v(t), R_n(t), R_v(t)]^T$$

be the random vector representing the state of the system at time t. The CTMC model is assumed to be time-homogeneous and to satisfy the Markov property, meaning that the future state of the process depends only on the current state and not on the past history. Consequently, the time until the next event occurring is exponentially distributed.

The possible state transitions and their corresponding rates for the CTMC model are defined based on the events described in the deterministic model. These are summarized in Table 2. Each row in the table describes an event, the change in the state vector, and the rate at which that event occurs. For a sufficiently small time interval Δt, the probability of a particular transition (event j with rate r_j) occurring is $r_j \Delta t + o(\Delta t)$, and it is assumed that at most one event can take place during such an infinitesimal interval.

Table 2. Events, Transitions and Rates of the CTMC Model.

Event	**Population components at** t	**Population components at** $t+\Delta t$	r_j
Recruitment of S_n	$(S_n, S_v, E_n, E_v, I_n, I_v, R_n, R_v)$	$(S_n+1, S_v, E_n, E_v, I_n, I_v, R_n, R_v)$	κ
Infection of S_n	$(S_n, S_v, E_n, E_v, I_n, I_v, R_n, R_v)$	$(S_n-1, S_v, E_n+1, E_v, I_n, I_v, R_n, R_v)$	$\lambda_n S_n$
Vaccination of S_n	$(S_n, S_v, E_n, E_v, I_n, I_v, R_n, R_v)$	$(S_n-1, S_v+1, E_n, E_v, I_n, I_v, R_n, R_v)$	ϕS_n
Natural death of S_n	$(S_n, S_v, E_n, E_v, I_n, I_v, R_n, R_v)$	$(S_n-1, S_v, E_n, E_v, I_n, I_v, R_n, R_v)$	μS_n
Waning immunity of S_v	$(S_n, S_v, E_n, E_v, I_n, I_v, R_n, R_v)$	$(S_n+1, S_v-1, E_n, E_v, I_n, I_v, R_n, R_v)$	$\omega_{sv} S_v$
Infection of S_v	$(S_n, S_v, E_n, E_v, I_n, I_v, R_n, R_v)$	$(S_n, S_v-1, E_n, E_v+1, I_n, I_v, R_n, R_v)$	$\lambda_v S_v$
Natural Death of S_v	$(S_n, S_v, E_n, E_v, I_n, I_v, R_n, R_v)$	$(S_n, S_v-1, E_n, E_v, I_n, I_v, R_n, R_v)$	μS_v
Progression of E_n to I_n	$(S_n, S_v, E_n, E_v, I_n, I_v, R_n, R_v)$	$(S_n, S_v, E_n-1, E_v, I_n+1, I_v, R_n, R_v)$	$\sigma_n E_n$
Vaccination of E_n	$(S_n, S_v, E_n, E_v, I_n, I_v, R_n, R_v)$	$(S_n, S_v, E_n-1, E_v+1, I_n, I_v, R_n, R_v)$	ϕE_n
Natural death of E_n	$(S_n, S_v, E_n, E_v, I_n, I_v, R_n, R_v)$	$(S_n, S_v, E_n-1, E_v, I_n, I_v, R_n, R_v)$	μE_n
Progression of E_v to I_v	$(S_n, S_v, E_n, E_v, I_n, I_v, R_n, R_v)$	$(S_n, S_v, E_n, E_v-1, I_n, I_v+1, R_n, R_v)$	$\sigma_v E_v$
Natural death of E_v	$(S_n, S_v, E_n, E_v, I_n, I_v, R_n, R_v)$	$(S_n, S_v, E_n, E_v-1, I_n, I_v, R_n, R_v)$	μE_v
Recovery of I_n	$(S_n, S_v, E_n, E_v, I_n, I_v, R_n, R_v)$	$(S_n, S_v, E_n, E_v, I_n-1, I_v, R_n+1, R_v)$	$\gamma_n I_n$
Natural death of I_n	$(S_n, S_v, E_n, E_v, I_n, I_v, R_n, R_v)$	$(S_n, S_v, E_n, E_v, I_n-1, I_v, R_n, R_v)$	μI_n
Disease induced death of I_n	$(S_n, S_v, E_n, E_v, I_n, I_v, R_n, R_v)$	$(S_n, S_v, E_n, E_v, I_n-1, I_v, R_n, R_v)$	$\delta_n I_n$
Vaccination of I_n	$(S_n, S_v, E_n, E_v, I_n, I_v, R_n, R_v)$	$(S_n, S_v, E_n, E_v, I_n-1, I_v+1, R_n, R_v)$	ϕI_n
Recovery of I_v	$(S_n, S_v, E_n, E_v, I_n, I_v, R_n, R_v)$	$(S_n, S_v, E_n, E_v, I_n, I_v-1, R_n, R_v+1)$	$\gamma_v I_v$
Natural death of I_v	$(S_n, S_v, E_n, E_v, I_n, I_v, R_n, R_v)$	$(S_n, S_v, E_n, E_v, I_n, I_v-1, R_n, R_v)$	μI_v
Disease induced death of I_v	$(S_n, S_v, E_n, E_v, I_n, I_v, R_n, R_v)$	$(S_n, S_v, E_n, E_v, I_n, I_v-1, R_n, R_v)$	$\delta_v I_v$
Natural immunity waning of R_n	$(S_n, S_v, E_n, E_v, I_n, I_v, R_n, R_v)$	$(S_n+1, S_v, E_n, E_v, I_n, I_v, R_n-1, R_v)$	$\omega_{rn} R_n$
Vaccination of R_n	$(S_n, S_v, E_n, E_v, I_n, I_v, R_n, R_v)$	$(S_n, S_v, E_n, E_v, I_n, I_v, R_n-1, R_v+1)$	ϕR_n
Natural death of R_n	$(S_n, S_v, E_n, E_v, I_n, I_v, R_n, R_v)$	$(S_n, S_v, E_n, E_v, I_n, I_v, R_n-1, R_v)$	μR_n
Vaccine-induced immunity waning of R_v	$(S_n, S_v, E_n, E_v, I_n, I_v, R_n, R_v)$	$(S_n+1, S_v, E_n, E_v, I_n, I_v, R_n, R_v-1)$	$\omega_{rv} R_v$
Natural death of R_v	$(S_n, S_v, E_n, E_v, I_n, I_v, R_n, R_v)$	$(S_n, S_v, E_n, E_v, I_n, I_v, R_n, R_v-1)$	μR_v

The sum of all possible transition rates from state $\vec{Y}(t)$ is given by $\psi(\vec{Y}(t))$, where:

$$
\begin{aligned}
\psi(\vec{Y}(t)) = {} & \underbrace{\kappa}_{\text{Recruitment } S_n} + \underbrace{\lambda_n(t)S_n(t)}_{\text{Infection of } S_n} + \underbrace{\phi S_n(t)}_{\text{Vaccination of } S_n} + \underbrace{\mu S_n(t)}_{\text{Death of } S_n} \\
& + \underbrace{\omega_{sv} S_v(t)}_{\text{Waning } S_v} + \underbrace{\lambda_v(t)S_v(t)}_{\text{Infection of } S_v} + \underbrace{\mu S_v(t)}_{\text{Death of } S_v} \\
& + \underbrace{\sigma_n E_n(t)}_{\text{Progression } E_n \to I_n} + \underbrace{\phi E_n(t)}_{\text{Vaccination of } E_n} + \underbrace{\mu E_n(t)}_{\text{Death of } E_n} \\
& + \underbrace{\sigma_v E_v(t)}_{\text{Progression } E_v \to I_v} + \underbrace{\mu E_v(t)}_{\text{Death of } E_v} \\
& + \underbrace{\gamma_n I_n(t)}_{\text{Recovery of } I_n} + \underbrace{\mu I_n(t)}_{\text{Death of } I_n} + \underbrace{\delta_n I_n(t)}_{\text{Disease death } I_n} + \underbrace{\phi I_n(t)}_{\text{Vaccination of } I_n} \\
& + \underbrace{\gamma_v I_v(t)}_{\text{Recovery of } I_v} + \underbrace{\mu I_v(t)}_{\text{Death of } I_v} + \underbrace{\delta_v I_v(t)}_{\text{Disease death } I_v} \\
& + \underbrace{\omega_{rn} R_n(t)}_{\text{Waning } R_n} + \underbrace{\phi R_n(t)}_{\text{Vaccination of } R_n} + \underbrace{\mu R_n(t)}_{\text{Death of } R_n} \\
& + \underbrace{\omega_{rv} R_v(t)}_{\text{Waning } R_v} + \underbrace{\mu R_v(t)}_{\text{Death of } R_v}
\end{aligned}
$$

where $\lambda_n(t) = \frac{(\beta_{nn} I_n(t) + \beta_{nv} I_v(t))}{N(t)}$ and $\lambda_v(t) = \frac{(\beta_{vv} I_v(t) + \beta_{vn} I_n(t))}{N(t)}$, with $N(t) > 0$.

The CTMC framework allows for the direct simulation of sample paths of the epidemic process using algorithms such as Gillespie's Stochastic Simulation Algorithm (SSA). This approach enables the estimation of quantities like the probability of disease extinction or outbreak, and the expected time to extinction, which are critical for assessing public health risks and intervention strategies. During the early stages of an epidemic, when the number of infected individuals is small and the number of susceptibles is large (close to the disease-free equilibrium values S_n^0 and S_v^0), the dynamics of the infected classes can be approximated by a multitype branching process. This approximation will be used in subsection 4.2 to derive an analytical expression for the probability of disease extinction.

2.2.2. Branching Process Approximation The theory of multitype Galton-Watson branching processes boasts a lengthy history. It has been employed to estimate the dynamics of the CTMC model in proximity to the disease-free equilibrium (DFE) and the stochastic threshold for a disease outbreak [17, 16]. The stochastic threshold has a direct link to the basic reproduction number, $\mathscr{R}_0$, as defined in the related deterministic model [14]. Specifically, when $\mathscr{R}_0 < 1$, disease extinction is certain, and the branching process is termed subcritical. Conversely, if $\mathscr{R}_0 > 1$, the probability of disease extinction is less than one (though still positive), and this process is identified as supercritical.

The subsequent methodology applies a multitype branching process approximation of the CTMC model at the DFE (where susceptible populations are S_n^0 and S_v^0, and $N^0 = S_n^0 + S_v^0$) to estimate the probability of disease extinction. The infected classes are of particular interest: non-vaccinated exposed (E_n), vaccinated exposed (E_v), non-vaccinated infectious (I_n), and vaccinated infectious (I_v). Let $X = (X_1, X_2, X_3, X_4) = (E_n, E_v, I_n, I_v)$ represent a vector of integer random variables denoting the number of individuals in these respective states. Let δ_{ij} denote the Kronecker delta. The offspring probability generating function (PGF) for an initial individual of type i, given $X_j = \delta_{ij}$ (i.e., starting with one individual of type i and zero of other types), is a function $f_i : [0,1]^4 \to [0,1]$, expressed as:

$$f_i(x_1, x_2, x_3, x_4) = \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} \sum_{k_3=0}^{\infty} \sum_{k_4=0}^{\infty} P_i(k_1, k_2, k_3, k_4) x_1^{k_1} x_2^{k_2} x_3^{k_3} x_4^{k_4} \tag{58}$$

In this context, $P_i(k_1, k_2, k_3, k_4)$ represents the probability that an individual of type i generates k_j individuals of type j (for $j = 1, 2, 3, 4$) over its lifetime before it is removed from its current state. The PGFs for the four infectious types are defined as follows, based on the possible transitions from each state at the DFE:

$$\begin{aligned}
f_1(x_1, x_2, x_3, x_4) &= \frac{\sigma_n x_3 + \phi x_2 + \mu}{\sigma_n + \phi + \mu} \\
f_2(x_1, x_2, x_3, x_4) &= \frac{\sigma_v x_4 + \mu}{\sigma_v + \mu} \\
f_3(x_1, x_2, x_3, x_4) &= \frac{\frac{\beta_{nn} S_n^0}{N^0} x_1 x_3 + \frac{\beta_{vn} S_v^0}{N^0} x_2 x_3 + \phi x_4 + \delta_n + \gamma_n + \mu}{\frac{\beta_{nn} S_n^0}{N^0} + \frac{\beta_{vn} S_v^0}{N^0} + \phi + \delta_n + \gamma_n + \mu} \\
f_4(x_1, x_2, x_3, x_4) &= \frac{\frac{\beta_{nv} S_n^0}{N^0} x_1 x_4 + \frac{\beta_{vv} S_v^0}{N^0} x_2 x_4 + \delta_v + \gamma_v + \mu}{\frac{\beta_{nv} S_n^0}{N^0} + \frac{\beta_{vv} S_v^0}{N^0} + \delta_v + \gamma_v + \mu}
\end{aligned} \tag{59}$$

Here, x_1, x_2, x_3, x_4 are dummy variables corresponding to the number of E_n, E_v, I_n, I_v offspring, respectively. For instance, in f_1, an E_n individual can progress to I_n (producing one x_3 offspring) with rate σ_n, be vaccinated to become E_v (producing one x_2 offspring) with rate ϕ, or be removed (e.g., die) with rate μ (producing zero offspring of these types). Similar interpretations apply to f_2, f_3, and f_4. Note that when an infectious individual (I_n or I_v) successfully infects a susceptible, it produces one new exposed individual (E_n or E_v) and the original infectious individual itself is considered to survive that specific interaction event, hence terms like $x_1 x_3$.

As suggested by the theory of multitype branching processes [35, 36], the fixed points of the offspring PGFs provide an estimate of the probability of disease extinction. Let (q_1, q_2, q_3, q_4) denote the minimal non-negative fixed points

of the PGFs in $[0,1]^4$, which satisfy $f_i(q_1,q_2,q_3,q_4)=q_i$ for $i=1,2,3,4$. An estimate of the extinction probability, given an initial state $X(0)=(e_n(0),e_v(0),i_n(0),i_v(0))$, is:

$$\boldsymbol{P}_{\text{ext}}=\lim_{t\to\infty}\boldsymbol{P}(\boldsymbol{X}(t)=\boldsymbol{0})=q_1^{e_n(0)}q_2^{e_v(0)}q_3^{i_n(0)}q_4^{i_v(0)}. \tag{60}$$

Thus, the likelihood of a major outbreak is:

$$\boldsymbol{P}_{\text{out}}=1-\boldsymbol{P}_{\text{ext}}=1-q_1^{e_n(0)}q_2^{e_v(0)}q_3^{i_n(0)}q_4^{i_v(0)}. \tag{61}$$

The expectation matrix, $\mathscr{M}=[m_{ji}]$, is a 4×4 matrix where the element m_{ji} signifies the expected number of new individuals of type j produced by a single infectious individual of type i during its infectious period (or before transitioning out of its current state). Elements of $\mathscr{M}$ are obtained by differentiating f_i with respect to x_j and then evaluating all x_k variables at 1:

$$m_{ji}=\left.\frac{\partial f_i(x_1,x_2,x_3,x_4)}{\partial x_j}\right|_{x_1=x_2=x_3=x_4=1}.$$

The expectation matrix $\mathscr{M}$ (where rows indicate offspring type and columns indicate parent type, i.e., $\mathscr{M}_{jk}=$ expected offspring of type j from parent type k) is:

$$\mathscr{M}=\begin{pmatrix}\frac{\partial f_1}{\partial x_1} & \frac{\partial f_2}{\partial x_1} & \frac{\partial f_3}{\partial x_1} & \frac{\partial f_4}{\partial x_1}\\ \frac{\partial f_1}{\partial x_2} & \frac{\partial f_2}{\partial x_2} & \frac{\partial f_3}{\partial x_2} & \frac{\partial f_4}{\partial x_2}\\ \frac{\partial f_1}{\partial x_3} & \frac{\partial f_2}{\partial x_3} & \frac{\partial f_3}{\partial x_3} & \frac{\partial f_4}{\partial x_3}\\ \frac{\partial f_1}{\partial x_4} & \frac{\partial f_2}{\partial x_4} & \frac{\partial f_3}{\partial x_4} & \frac{\partial f_4}{\partial x_4}\end{pmatrix}_{x_k=1}$$

$$=\begin{pmatrix}0 & 0 & \frac{\frac{\beta_{nn}S_n^0}{N^0}}{\frac{\beta_{nn}S_n^0}{N^0}+\frac{\beta_{vn}S_v^0}{N^0}+\phi+\delta_n+\gamma_n+\mu} & \frac{\frac{\beta_{nv}S_n^0}{N^0}}{\frac{\beta_{nv}S_n^0}{N^0}+\frac{\beta_{vv}S_v^0}{N^0}+\delta_v+\gamma_v+\mu}\\ \frac{\phi}{\sigma_n+\phi+\mu} & 0 & \frac{\frac{\beta_{vn}S_v^0}{N^0}}{\frac{\beta_{nn}S_n^0}{N^0}+\frac{\beta_{vn}S_v^0}{N^0}+\phi+\delta_n+\gamma_n+\mu} & \frac{\frac{\beta_{vv}S_v^0}{N^0}}{\frac{\beta_{nv}S_n^0}{N^0}+\frac{\beta_{vv}S_v^0}{N^0}+\delta_v+\gamma_v+\mu}\\ \frac{\sigma_n}{\sigma_n+\phi+\mu} & 0 & \frac{\frac{\beta_{nn}S_n^0}{N^0}+\frac{\beta_{vn}S_v^0}{N^0}}{\frac{\beta_{nn}S_n^0}{N^0}+\frac{\beta_{vn}S_v^0}{N^0}+\phi+\delta_n+\gamma_n+\mu} & 0\\ 0 & \frac{\sigma_v}{\sigma_v+\mu} & \frac{\phi}{\frac{\beta_{nn}S_n^0}{N^0}+\frac{\beta_{vn}S_v^0}{N^0}+\phi+\delta_n+\gamma_n+\mu} & \frac{\frac{\beta_{nv}S_n^0}{N^0}+\frac{\beta_{vv}S_v^0}{N^0}}{\frac{\beta_{nv}S_n^0}{N^0}+\frac{\beta_{vv}S_v^0}{N^0}+\delta_v+\gamma_v+\mu}\end{pmatrix} \tag{62}$$

The spectral radius of the expectation matrix, $\rho(\mathscr{M})$, indicates whether the disease will fade away or persist [14]. A condition where $\rho(\mathscr{M})>1$ implies that $q_j<1$ for at least one j, offering a non-zero probability for disease persistence (major outbreak). Conversely, if $\rho(\mathscr{M})\le 1$, the probability of the disease ultimately becoming extinct is 1 [37, 14]. The characteristic equation for matrix $\mathscr{M}$ is $P(\lambda)=\det(\mathscr{M}-\lambda I)=0$. For a 4×4 matrix, this is a quartic polynomial:

$$\lambda^4+\Theta_3\lambda^3+\Theta_2\lambda^2+\Theta_1\lambda+\Theta_0=0, \tag{63}$$

where the coefficients Θ_k are functions of the elements m_{ji} of $\mathscr{M}$. In the CTMC model, the threshold for disease extinction or persistence is identified by $\rho(\mathscr{M})$. This threshold in stochastic models mirrors the function of $\mathscr{R}_0$ in deterministic models. According to [14], there is a proven relationship between the stochastic and deterministic thresholds for disease extinction, specifically $\rho(\mathscr{M})$ and $\mathscr{R}_0$:

$$\begin{aligned}&\mathscr{R}_0<1 \quad \text{if and only if} \quad \rho(\mathscr{M})<1\\ &\mathscr{R}_0>1 \quad \text{if and only if} \quad \rho(\mathscr{M})>1\\ &\mathscr{R}_0=1 \quad \text{if and only if} \quad \rho(\mathscr{M})=1.\end{aligned} \tag{64}$$

While handling the fourth-order polynomial (63) to find a direct closed-form expression for $\rho(\mathscr{M})$ presents some analytical difficulty, this is not necessary. The established relationship in (64) shows that the stochastic threshold $\rho(\mathscr{M}) = 1$ is identical to the deterministic threshold $\mathscr{R}_0 = \rho(FV^{-1}) = 1$, where $\mathscr{R}_0$ is derived using the next-generation matrix (NGM) method. This $\mathscr{R}_0$ provides the required interpretable, analytical threshold. In Appendix A.4 a full derivation is provided, which demonstrates the formal equivalence $F - V = (\mathscr{M} - I)W$ for the model and explicitly derives the F (new infections) and V (transitions) matrices. Therefore, the threshold for disease persistence is determined by the analytical expression for $\mathscr{R}_0$, while the matrix $\mathscr{M}$ is used to numerically solve for the extinction probabilities (q_1, q_2, q_3, q_4) when the threshold is exceeded.

3. Results

Before data fitting, paremeter estimation and sensitivity analysis, it is necessary to perform exploratory data analysis (EDA) on the COVID-19 data set to be used. The data set used for this study was retrieved from a repository by [38]. The data set has daily and cumulative data on the number of COVID-19 infected patients, deaths, and recoveries; however, we focus more on incidence data (infected patient data).

3.1. Data Fitting and Parameter Estimation Using the Deterministic Model

The fundamental characteristics of COVID-19 in South Africa are shown in table 3, where Q_1 represents the first quartile and Q_2 is the second quartile. In the COVID-19 data set, data from March 5, 2020 to March 21, 2022 are presented. The minimum recorded infection of 1 and a maximum of over 3.7 million illustrate the broad range of values, spanning nearly 3.7 million cases. The first quartile (around 640,000) and third quartile (about 2.87 million) indicate that half of the data lies between these values, highlighting substantial variability in infection levels across different time points. The median of approximately 1.53 million suggests that half of the recorded infection counts fall below this figure, while the mean (around 1.62 million) is slightly higher, reflecting a slight skew toward larger values. The relatively high standard deviation (about 1.20 million) also confirms that the data are widely dispersed. A skewness of 0.27 indicates a mild right skew, meaning most infection counts cluster toward the lower end but with a tail extending toward higher values. Meanwhile, a kurtosis of 1.77 suggests the distribution has moderately less pronounced tails compared to a normal distribution.

Table 3. Characteristics of South African COVID-19 Data.

Statistics	**Infected**
Minimum	1.00
Maximum	3704785.00
Q_1	639901.50
Median	1528414.00
Q_3	2871309.00
Mean	1616999.64
Range	3704784.00
Standard Deviation	1203633.35
Skewness	0.27
Kurtosis	1.77

The plot in Figure 2 represents daily new COVID-19 cases in South Africa (spanning March 2020 to March 2022) and reveals a clear pattern of multiple waves of infection. Each wave rises sharply to a peak and then gradually declines before the next wave begins. The first noticeable surge occurs from week 24 of 2020 to week 34 of 2020 (June 7, 2020 to August 22, 2020). It was followed by a second, larger wave in week 47 of 2020 to week 5 of 2021 (November 15, 2020 to February 6, 2021). The third wave emerges around week 19 of 2021 to week 37 of 2021 (May 9, 2021 to September 18,2021), showing a similarly sharp increase and subsequent decline. Notably,

the largest spike appears toward the end of 2021, aligning with the emergence of the Omicron variant [39]. This was during week 47 of 2021 to week 3 of 2022 (November 21, 2021 to January 22, 2022). Between these peaks, the number of daily new cases dips significantly, indicating periods of lower transmission. The oscillations suggest that factors such as public health interventions including vaccination campaigns and viral mutations played a role in driving infection levels.

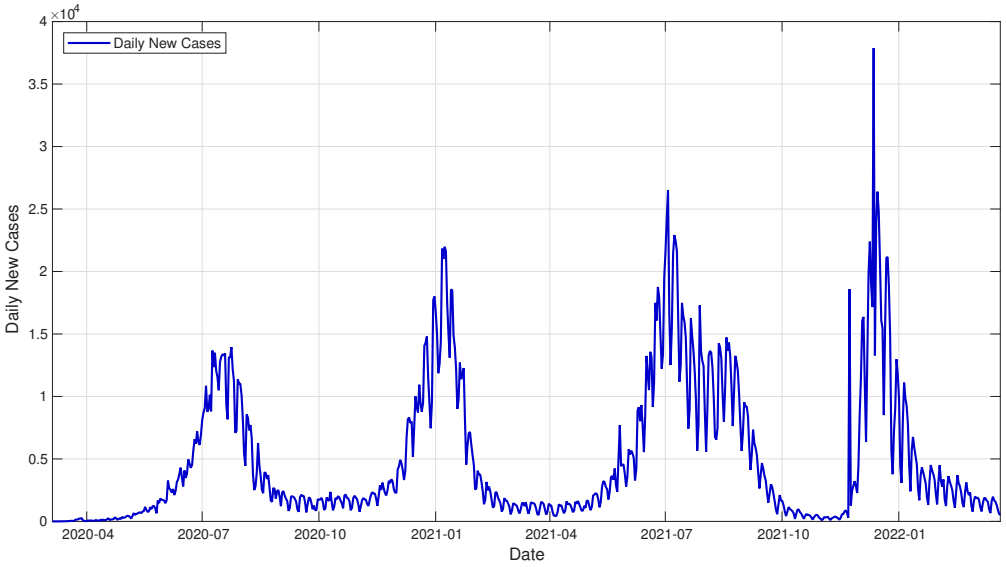

Figure 2. Daily Infection Cases of COVID-19 in South Africa.

To validate the proposed vaccine-structured model,a parameter estimation process was performed by fitting the deterministic system of differential equations (1) to the daily reported COVID-19 case data from South Africa. The specific focus was on wave three, spanning from May 9, 2021, to September 18, 2021, as this period coincided with active vaccination efforts [40]. The estimation process employed the Ordinary Least Squares (OLS) method to minimize the discrepancy between the observed data and the model's predicted incidence.

Let Y_j denote the observed number of daily new confirmed cases at time point t_j for $j = 1, 2, \ldots, M$, where M is the total number of data points in the time series. Correspondingly, let $Z(t_j, \Theta)$ represent the model-predicted daily incidence at time t_j, which depends on the vector of unknown parameters $\Theta = (\beta_{nn}, \beta_{nv}, \beta_{vv}, \ldots, \phi)$. In the context of the model structure, the theoretical daily incidence is defined as the rate at which individuals transition from the exposed compartments to the infectious compartments, given by:

$$Z(t, \Theta) = \sigma_n E_n(t) + \sigma_v E_v(t). \tag{65}$$

According to [41] the parameter estimation problem is formulated as an optimization problem designed to find the optimal parameter set $\hat{\Theta}$ that minimizes the sum of squared errors (SSE) objective function, $J(\Theta)$:

$$J(\Theta) = \sum_{j=1}^{M} \left(Y_j - Z(t_j, \Theta)\right)^2. \tag{66}$$

Thus, the optimal parameter estimator is obtained by solving:

$$\hat{\Theta} = \arg\min_{\Theta \in \Omega} J(\Theta) \tag{67}$$

where Ω denotes the feasible parameter space defined by the lower and upper bounds listed in Table 4. The numerical solution to this non-linear least squares optimization problem was implemented using the

`lsqcurvefit` function in the MATLAB Optimization Toolbox, which utilizes a trust-region-reflective algorithm to iteratively converge upon the 'local minimizer' subject to the specified boundary constraints.

Table 4. Bounds Imposed on the Parameters in the Optimization Problem.

Parameter	Bounds	Notes / Source
β	0.19 – 2.14	Derived from $R_0 \in [1.9, 6.5]$ and $\gamma \in [0.10, 0.33]$; see [42] for R_0 meta-analysis.
β_{nn}	0.19 – 2.14	As above.
β_{nv}, β_{vn}, β_{vv}	$\beta_{nn}\times$[0.17–0.78]	Multiplier = (1 – Vaccine Efficacy(VE)_infection); VE against infection range 22 %–83 %. [43]
ω_{sv}	0.0020 – 0.0200	Corresponds to immunity duration ≈50 – 500 days. (1/ω) ; Vaccine-induced immunity wanes substantially by 3–6 months for infection. [12]
ω_{rn}	0.0010 – 0.0100	Natural immunity after infection may last 100 – 1000 days before significant waning. [44]
ω_{rv}	0.000004 – 0.0050	Hybrid immunity appears more durable; thus slower waning. [45]
δ_n	1×10^{-6} – 5.5×10^{-3}	Infection Fatality Rate (IFR) ≈0.53%–0.82% . [46]
δ_v	1×10^{-7} – 5.5×10^{-4}	Assume vaccinated severe outcome reduction ≈50–95%.
σ_n, σ_v	0.07 – 0.50	Incubation period ≈2–14 days, so $\sigma = 1/$incubation. [12]
γ_n, γ_v	0.10 – 0.36	Infectious period ≈3–10 days so $\gamma = 1/$infectious period.
ϕ	0 – 0.20	Depends on rollout speed; e.g. 0.01/day ≈1% of susceptibles vaccinated per day.

Table 5 presents the initial conditions for each compartment during the third wave. The parameter estimation simulation spans from May 9, 2021 to September 18, 2021. The initial values for the simulation include $S_n(0)$, calculated by deducting $S_v(0)$ and $I_n(0)$ from the total population of South Africa. As daily data for $I_n(0)$ were not available, the exact value of the smoothed new case variable in the data set is used. The remaining initial values, namely $E_n(0)$, $E_v(0)$, $I_v(0)$, $R_n(0)$, and $R_v(0)$, are set to zero based on the assumption that non-vaccinated infectious individuals are beginning to infect the population during the third wave, thus focusing on values relevant to that specific wave. Using these initial conditions, the data from [38] is fitted.

Table 5. Initial Conditions for COVID-19 Wave 3.

Variable	Parameter Estimate	Source
$S_n(0)$	59509694	$S_n(0) = N(0) - S_v(0) - I_n(0)$
$S_v(0)$	382480	[38]
$E_n(0)$	0	Assumed
$E_v(0)$	0	Assumed
$I_n(0)$	1710	[38]
$I_v(0)$	0	Assumed
$R_n(0)$	0	Assumed
$R_v(0)$	0	Assumed

As observed in plot 3, the model (1) fits well with the new daily COVID-19 cases in South Africa.

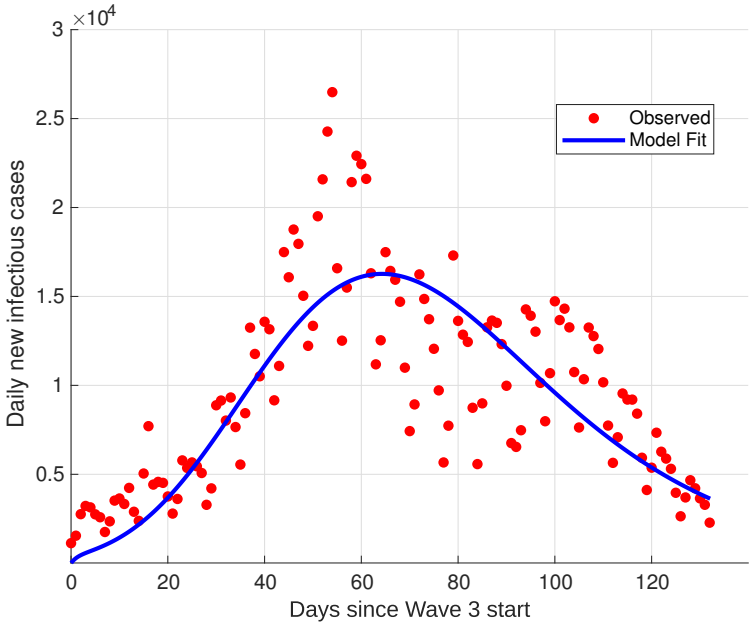

Figure 3. South African COVID-19 Wave 3 Model Fit to Data.

In table 6, multiple parameter values are generated throughout the whole curve fitting process based on the data from the third wave. These parameter values will be used to perform sensitivity analysis.

Table 6. Parameter Values for COVID-19.

Variable	**Value / Initial Condition**	**Source**	**Units**
κ	20	[47]	persons/day
μ	0.0161	[47]	day^{-1}
β_{nn}	0.96198	Fitted	day^{-1}
β_{nv}	0.16159	Fitted	day^{-1}
β_{vv}	0.016788	Fitted	day^{-1}
β_{vn}	0.026805	Fitted	day^{-1}
ω_{sv}	0.010343	Fitted	day^{-1}
ω_{rn}	0.0053871	Fitted	day^{-1}
ω_{rv}	4.4706e^{-06}	Fitted	day^{-1}
σ_n	0.23195	Fitted	day^{-1}
σ_v	0.062307	Fitted	day^{-1}
δ_n	3.7319e^{-05}	Fitted	day^{-1}
δ_v	4.3188e^{-07}	Fitted	day^{-1}
γ_n	0.35511	Fitted	day^{-1}
γ_v	0.18351	Fitted	day^{-1}
ϕ	0.018191	Fitted	day^{-1}

3.2. Sensitivity Analysis

To assess the robustness of the model predictions and identify the most influential parameters driving the epidemic dynamics, a global sensitivity analysis was conducted using the Latin Hypercube Sampling (LHS) method coupled with Partial Rank Correlation Coefficients (PRCC). This approach is particularly effective for non-linear dynamical systems where the relationship between inputs and outputs is monotonic but not necessarily linear.
The LHS method stratifies the probability distributions of the k input parameters into N equiprobable intervals, ensuring a comprehensive exploration of the multi-dimensional parameter space. Let $\Theta = [\theta_1, \theta_2, \ldots, \theta_k]$ be the vector of input parameters. $N = 1000$ samples were generated, constructing a parameter matrix X of size $N \times k$. The model system (1) was evaluated for each row of X to produce the corresponding output vector Y.

To quantify the sensitivity, the PRCC was calculated for each parameter, which measures the strength of the linear relationship between the rank-transformed parameter and the rank-transformed output, while linearly accounting for the effects of all other parameters. Let R_{θ_i} and R_Y denote the rank-transformed vectors of the i-th parameter and the output, respectively. Linear regressions of R_{θ_i} and R_Y on the remaining ranked parameters are performed to obtain the residuals ε_{θ_i} and ε_Y. The PRCC is defined as the Pearson correlation coefficient between these residuals:

$$\text{PRCC}(\theta_i, Y) = \frac{\text{Cov}(\varepsilon_{\theta_i}, \varepsilon_Y)}{\sqrt{\text{Var}(\varepsilon_{\theta_i})\text{Var}(\varepsilon_Y)}} \tag{68}$$

A PRCC value close to $+1$ or -1 indicates a strong positive or negative influence, respectively, while a value near 0 implies negligible impact. The sampling design was executed using the `lhsdesign` function, and the correlation coefficients were computed via the `partialcorr` function in the MATLAB Statistics and Machine Learning Toolbox. For more information on this method, refer to [48].

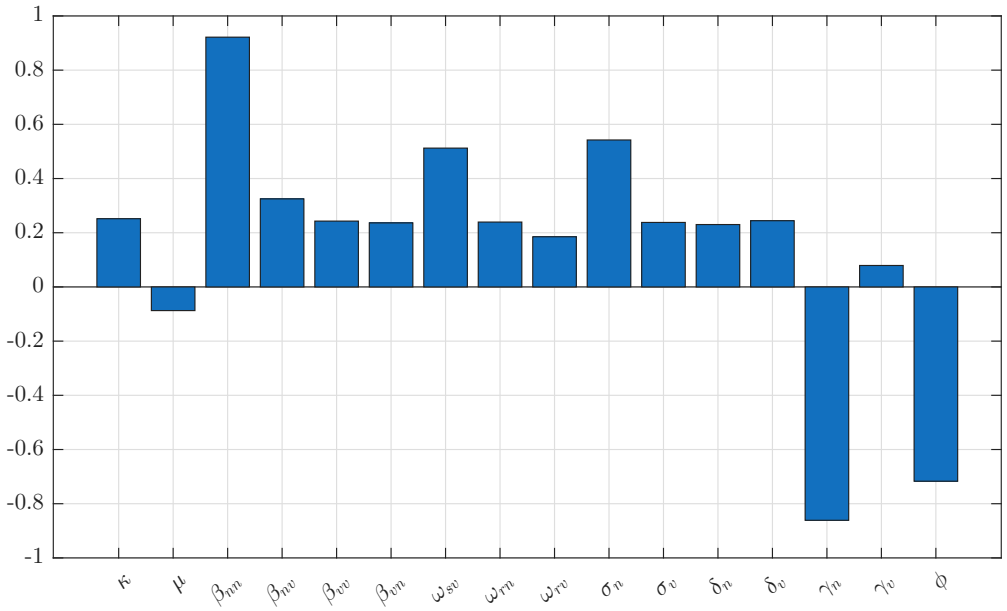

Figure 4. PRCC Analysis for COVID-19 Third Wave.

According to figure 4, the parameters with significant influence on the model are the infection rates between non-vaccinated individuals β_{nn}, the progression rate for the non-vaccinated individuals σ_n, the recovery rates of non-vaccinated individuals γ_n, and the vaccination rate ϕ. β_{nn}, which is the infection rate between non-vaccinated individuals, has a high positive PRCC suggesting that an increase in transmission rates among non-vaccinated individuals is strongly associated with higher infections. The progression rate, σ_n for the non-vaccinated exposed

also contributes to higher infections. This is because it contributes to a higher number of infectious individuals. The recovery rate of the non-vaccinated individuals γ_n, is inversely proportional to the model output, meaning that if the non-vaccinated recover quickly, the overall number of new cases will drop. The vaccination rate ϕ, is also inversely proportional, suggesting that increased vaccination rates will generally reduce infections.

3.3. Numerical Simulations for the Stochastic Model

Numerical illustrations will be used to depict the dynamics of both the deterministic and CTMC epidemic models. The probability of disease extinction or outbreak is approximated, along with the finite time to extinction. All the results in this section rely on the parameter values presented in Table 6. For these parameters, the basic reproduction number $\mathcal{R}_0 = 1.4130$, and the extinction threshold for the CTMC model $\rho(M) = 1.0880$. $\rho(M)$ is supercritical implying a positive probability of disease persistence in the population.

3.3.1. Probability of Disease Extinction The probability of extinction, $\mathbb{P}_0$, is derived from the fixed point of the multitype branching process, specifically equation (58). $\mathbb{P}_0$ is compared to the approximated probability of disease extinction, $\mathbb{P}_A$, obtained from a proportion of 10 000 sample paths of the CTMC model, for which the total number of infected individuals, that is $(E_n + E_v + I_n + I_v)$ equals zero before reaching a minimum outbreak size. According to table 7, $\mathbb{P}_0$ and $\mathbb{P}_A$ are quite close signifying a good agreement level. The fixed point, $(q_1, q_2, q_3, q_4) \in (0, 1)^4$ of the offspring pgfs (59) is $(q_1, q_2, q_3, q_4) = (0.7736, 0.9204, 0.7464, 0.8998)$

Table 7. Probability of Extinction $\mathbb{P}_0$ calculated from the fixed points of the branching process and numerical approximation $\mathbb{P}_A$ based on 10 000 sample paths of the CTMC model.

e_n	e_v	i_n	i_v	$\mathbb{P}_0$	$\mathbb{P}_A$
1	0	0	0	0.7736	0.7819
0	1	0	0	0.9204	0.9260
0	0	1	0	0.7464	0.7486
0	0	0	1	0.8998	0.9055
1	1	1	1	0.4782	0.4979

The results in table 7 show that an infection introduced by a single vaccinated individual, whether exposed (E_v) or infectious (I_v) has a considerably higher chance of fading out compared to an infection starting from a non-vaccinated individual, whether exposed (E_n) or infectious (I_n). This suggests that while vaccination may not entirely stop individuals from being infected or infectious, it substantially reduces the likelihood of an infection. As expected, if one of each individual in the four exposed/infectious categories $(E_n = 1, E_v = 1, I_n = 1, I_v = 1)$ introduces the disease, the probability of extinction becomes very low. This implies that the risk of an outbreak becomes higher if multiple people in different categories are infected at the onset of the pandemic. These results underscores that even if vaccines do not have 100% efficacy, they play an important role in containing an epidemic at the start [49]. In contrast to deterministic models, extinction can occur even if the reproduction number $\mathcal{R}_0 > 0$ due to demographic stochasticity that is present in populations. Figure 5 demonstrates disease persistence and extinction in the CTMC model when $\mathcal{R}_0 = 1.4130$. For easy comparison the deterministic and stochastic paths are drawn on the same graph.

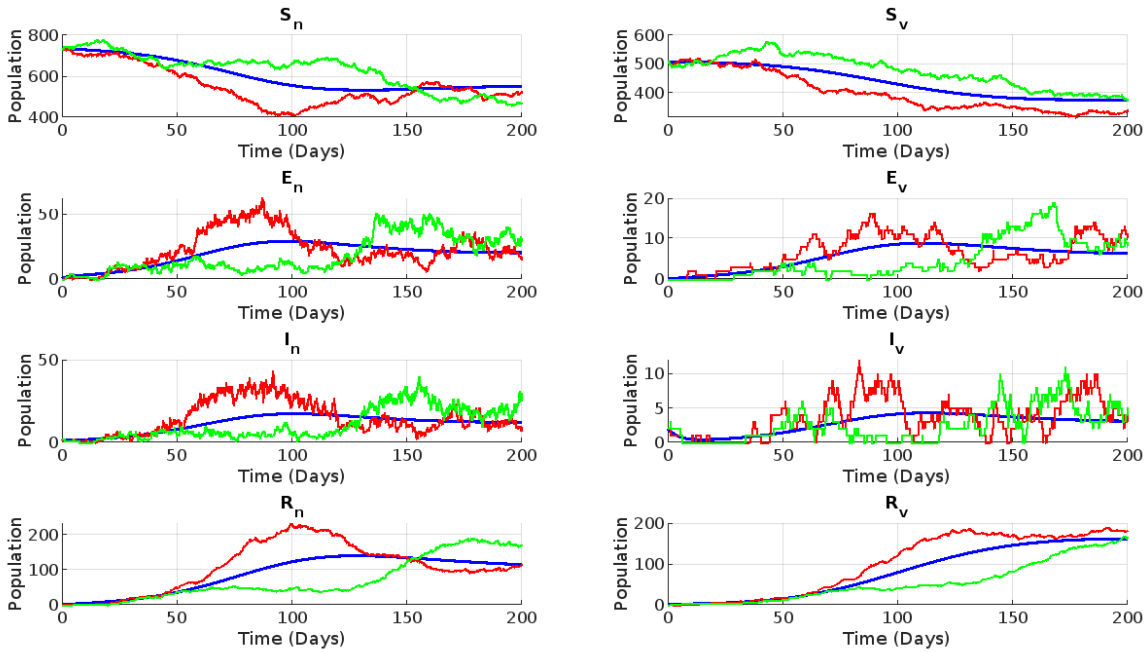

Figure 5. CTMC (Red and Green lines) vs Deterministic trajectories(Blue Lines) for each compartment in the Vaccine-Structured Model. Parameter values are as in Table 6 with initial conditions $S_n(0) = 735, S_v(0) = 506, E_n(0) = E_v(0) = R_n(0) = R_v(0) = 0, I_n(0) = 2, I_v(0) = 2$.

3.3.2. Finite Time to Extinction Stochastic models allow the estimation of finite time to extinction T. This finite time to extinction also known as expected epidemic duration cannot be estimated by deterministic models, so this is unique to stochastic models. Table 8 shows a significantly shorter duration when the infection originates from an infectious vaccinated individual I_v at $T \approx 35.38$ days. This is considerably faster than if the infection starts with an infectious non-vaccinated individual I_n, where $T \approx 61.61$ days. This suggests that while infections in vaccinated individuals are more likely to go extinct, those that do are also resolved more rapidly. The corresponding histograms in Figure 6 visually corroborates this, with the distribution for I_v extinctions being more concentrated at earlier time points.

Table 8. Finite Time to Extinction for the CTMC model when the disease is introduced by a non-vaccinated exposed, a vaccinated exposed, a non-vaccinated infectious and a vaccinated infectious individual. Parameter values are in table 6 with $\mathcal{R}_0 = 1.4130$ and $\rho(M) = 1.0880$.

Group(s) introducing the disease	P_A	T **days**
E_n	0.7819	56.00
E_v	0.9260	54.16
I_n	0.7486	61.61
I_v	0.9055	35.38

Considering infections initiated by exposed individuals, an exposed vaccinated individual E_v leads to an extinction time of $T \approx 54.16$ days, while an exposed non-vaccinated individual E_n results in a marginally longer $T = 56.00$ days. The higher the probability of extinction aligns with a slightly shorter extinction duration.

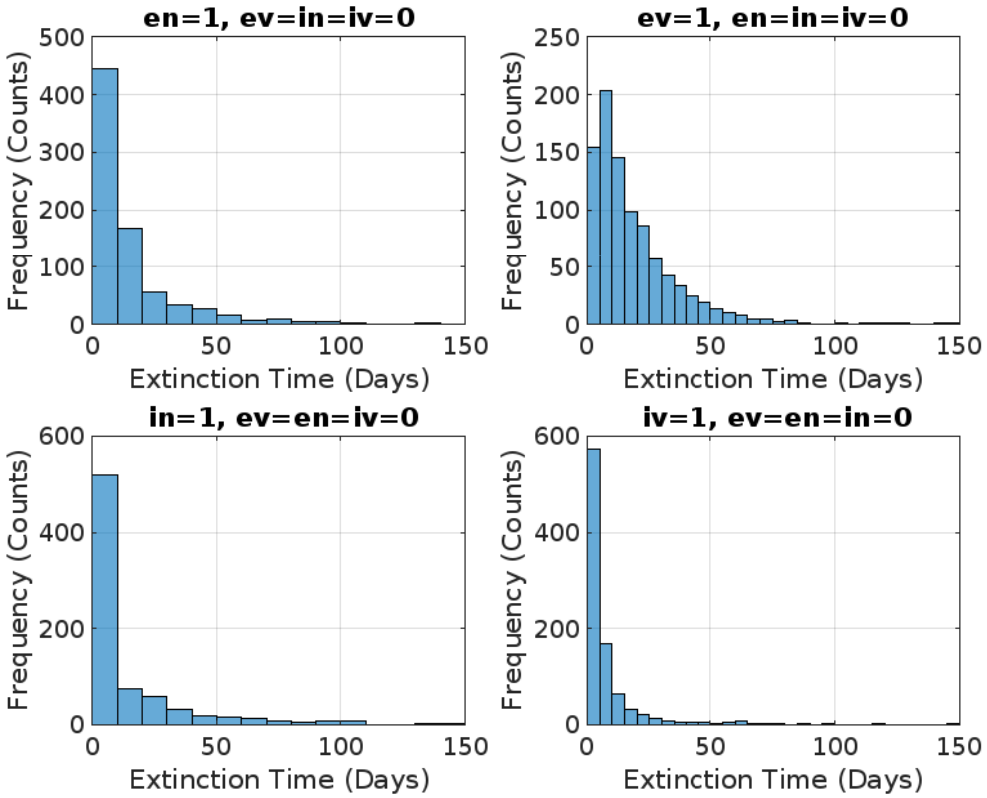

Figure 6. Approximate frequency distribution for the time of extinction for the CTMC model. Histograms are plotted based on 10 000 sample paths. Parameter values are as in table 6 and initial conditions are specified above each graph.

Collectively, these results highlight that vaccination not only increases the overall chance of disease fade-out from an initial case but can also dramatically shorten the duration of these outbreaks. This underscores the double benefit of vaccination in curtailing early epidemic spread. The histograms consistently show that most extinctions, regardless of initial state, occur relatively early, with a tail of less frequent, longer duration fade-outs.

4. Discussion

4.1. General discussion

In this study, the transmission dynamics of COVID-19 were investigated within a population characterized by both vaccinated and non-vaccinated individuals. The model was developed and fitted using data from the third wave of COVID-19 in South Africa, a period marked by active vaccination campaigns. The deterministic analysis yielded a basic reproduction number of $\mathscr{R}_0 \approx 1.4130$, indicating that, on average, the conditions were sufficient for a sustained epidemic. However, deterministic thresholds often mask the crucial role of random chance in the initial phase of an outbreak. This prompted an in-depth exploration of a stochastic analysis through the use of a Continuous-Time Markov Chain (CTMC) model.

The stochastic analysis focused on quantifying the probability of disease extinction, a critical factor for public health planning. A multitype branching process was employed to analytically derive the extinction probability (P_0) and these findings were validated against simulations (P_A) using Gillespie's algorithm [14]. The excellent

agreement between P_0 and P_A (as seen in Table 7) confirms the robustness of the branching process approximation for this complex, eight-state system. A central finding of this study is that the initial state of the infection profoundly influences the likelihood of it fading out. An infection introduced by a vaccinated individual, whether exposed (E_v, $P_A \approx 0.93$) or infectious (I_v, $P_A \approx 0.91$), has a substantially higher chance of going extinct than an infection originating in a non-vaccinated counterpart (E_n, $P_A \approx 0.78$; I_n, $P_A \approx 0.75$). This highlights a key community-level benefit of vaccination: even if breakthrough infections occur, they are significantly more likely to represent dead-end transmission chains that self-terminate before seeding a wider outbreak.

The findings align with and extend recent literature on vaccine-structured stochastic models. Previous theoretical work by [28] demonstrated that in heterogeneous populations, the probability of a minor outbreak is highly sensitive to the sub-group in which the infection is introduced. The results of this study confirm this in the specific context of COVID-19 vaccination in South Africa, showing quantitatively how the 'leaky' nature of these vaccines, where infection is possible but less likely to transmit, creates a distinct stochastic buffer. Furthermore, while recent stochastic studies such as [50] have emphasized the role of non-pharmaceutical interventions (like isolation) in increasing extinction probabilities early in a COVID-19 outbreak, this work complements this by isolating the specific contribution of vaccine-derived immunity to this stochastic fade-out. Unlike simpler stochastic models that may aggregate vaccinated and non-vaccinated susceptibles, the stratified approach allows for a more granular risk assessment of breakthrough introductions.

This study further confirms that stochastic models provide insights unattainable through deterministic methods alone [16, 14]. While the deterministic model predicted a sustained outbreak for $\mathscr{R}_0 > 1$, the stochastic simulations (Figure 5) revealed a spectrum of possibilities, including both major outbreaks and stochastic fade-out, which more accurately reflects the uncertainty inherent in real-world epidemics. The analysis of the finite time to extinction (T) further elucidated the dual impact of vaccination. Notably, outbreaks initiated by an infectious vaccinated individual (I_v) that are destined for extinction resolve the fastest ($T \approx 35$ days). This suggests that not only are such transmission chains more likely to break, but they also persist for a shorter duration, limiting the window for further spread.

4.2. Novel Insights and Practical Implications

Many models have focused on capturing key heterogeneities to guide policy, with age structure being a primary consideration. For instance, the work by [51] developed a detailed age-structured model for West Africa to specifically investigate the impact of vaccine prioritization strategies among different age groups. Similarly, [28] constructed an age-structured model with vaccination to analyze epidemic data from China, focusing on the stability of equilibria and sensitivity of the reproduction number to various parameters. While these age-structured models provide critical insights into targeted vaccination campaigns, the model in this work concentrates on a different, equally vital form of heterogeneity: the stochastic dynamics arising from vaccination status alone. The model shares some similarities with the work of [52], which also analyzed COVID-19 vaccination in South Africa (specifically Gauteng province) using a compartmental framework with vaccinated compartments. However, their model's primary focus was on exploring optimal control strategies through time-dependent parameters for non-pharmaceutical interventions. In contrast, our model introduces a more complex internal structure, notably the feature where vaccination is not just for the susceptible class but can occur in all compartments (E_n, I_n, R_n), moving individuals to their vaccinated counterparts. This novel feature reflects a real-world scenario where individuals at any stage of the natural infection cycle might receive a vaccine, altering their subsequent disease trajectory.

While the aforementioned studies provide valuable deterministic and, in some cases, optimal control analyses, this work is fundamentally centered on the stochastic nature of disease transmission. By formulating a Continuous-Time Markov Chain (CTMC) model and employing multitype branching process theory, the analysis moves beyond average-case predictions based on $\mathscr{R}_0$. The study analytically and numerically quantifies the probability of disease extinction (P_0), a critical metric for the assessment of public health risk that deterministic models cannot provide.

It is demonstrated how vaccination status dramatically alters this probability and also influences the finite time to extinction for abortive outbreaks. This focus on the probabilistic fate of an outbreak, especially how it is shaped by the interplay between non-vaccinated and vaccinated individuals at the very beginning of a transmission chain, offers a unique and complementary perspective to the existing literature. The study therefore fills an important gap by providing a stochastic analysis of a complex vaccine-structured model, yielding insights into the probabilistic, rather than purely deterministic, impact of vaccination on epidemic containment.

4.3. Limitations

The main limitation of this study is its reliance on data from the third wave of COVID-19 in South Africa, which may not be readily generalizable to other contexts. The unique socio-economic landscape and public health infrastructure of South Africa may influence parameters in ways that are specific to the region. Thus, one immediate next step for this line of research would be to apply the model to data from other countries to assess the robustness of the findings and improve generalizability. Another limitation is the assumption that model parameters remain fixed over time. This simplification may not fully capture the reality of an evolving pandemic marked by dynamic shifts in population behavior, the emergence of new viral variants with different characteristics, and changes in public health policy. Future studies could enhance the accuracy and adaptability of the model by incorporating time-varying parameters or non-stationary time series techniques to better reflect these real-world complexities. The model further presumes equal interaction between vaccinated and non-vaccinated groups and does not account for age differentiation. This simplification does not accurately represent actual contact patterns, and the lack of age stratification prevents the development of age-specific vaccination strategies. Adding contact matrices and age structures would significantly enhance future research.

Furthermore, the model's comprehensive nature, encompassing eight compartments and four infectious states, introduces significant analytical complexity. While this provides more detailed insights, it renders analytical expressions for the basic reproduction number ($\mathcal{R}_0$) and endemic equilibria extremely complex and difficult to interpret intuitively. The reliance on numerical methods to solve for key thresholds, such as the spectral radius of the expectation matrix, may present a barrier to rapid replication or application by public health bodies without access to advanced computational tools. To address this limitation, researchers can simplify the model by reducing the number of classes or excluding certain demographic components, like death rates from specific compartments, tailored to their research scenarios. Finally, the model adopts a simplified vaccination mechanism where the rate ϕ is applied uniformly to non-vaccinated susceptible (S_n), exposed (E_n), infectious (I_n), and recovered (R_n) individuals. This uniform structure might overestimate the flow from non-vaccinated compartments into the vaccinated classes, representing an area where future models could incorporate more targeted vaccination protocols.

5. Conclusion

This paper developed and analyzed a vaccine-structured stochastic epidemic model to provide a deeper understanding of COVID-19 transmission dynamics. By comparing analytical results from branching process theory with large-scale CTMC simulations, we demonstrated that while the deterministic threshold ($\mathscr{R}_0 \approx 1.4130$) predicted a sustained outbreak, the probability of stochastic extinction for an initial case was high, particularly when the infection originated in a vaccinated individual. Our findings quantitatively show that vaccination provides a dual public health benefit: it dramatically increases the probability that a new infection will fade out by chance, and for those that do, it significantly shortens the time to extinction. This work underscores the critical importance of stochastic modeling in providing a more refined and realistic risk assessment for infectious disease control, highlighting that vaccination is a powerful tool for preventing new sparks from igniting into major epidemics.

A. Appendix

A.1. Derivation of the Basic Reproduction Number

To compute $\mathscr{R}_0$, the next generation matrix approach as outlined in [31] was employed. The appearance of new infective individuals in compartments I and E is written as;

$$\mathscr{F}_{1\times 8} = \left((\beta_{nn} I_n + \beta_{nv} I_v)\frac{S_n}{N}, (\beta_{vv} I_v + \beta_{vn} I_n)\frac{S_v}{N}, 0, 0, ..., 0 \right).$$

Define $\mathscr{V} = \mathscr{V}^- - \mathscr{V}^+ = (Q_1, Q_2, Q_3, Q_4)^T - (P_1, P_2, P_3, P_4)^T$, where $\mathscr{V}^-$ represents individuals leaving compartment I, and $\mathscr{V}^+$ denotes individuals entering compartment I by any means other than the emergence of new infective individuals. The components are, respectively,

$$\begin{aligned} P_1 &= ((\sigma_n + \mu + \phi)E_n, (\sigma_v + \mu)E_v), \quad & P_2 &= ((\mu + \gamma_n + \delta_n + \phi)I_n, (\mu + \gamma_v + \delta_v)I_v), \\ P_3 &= ((\phi + \mu)S_n, (\mu + \omega_{sv})S_v), & P_4 &= ((\mu + \phi + \omega_{rn})R_n, (\mu + \omega_{rv})R_v), \\ Q_1 &= (0, \phi E_n), & Q_2 &= (\sigma_n E_n, \sigma_v E_v + \phi I_n), \\ Q_3 &= (\kappa, \phi S_n), & Q_4 &= (\gamma_n I_n, \gamma_v I_v + \phi R_n). \end{aligned} \tag{69}$$

Let $Y = (E_n, E_v, I_n, I_v, S_n, S_v, R_n, R_v)$ represent the variables for $i, j = 1, 2, 3, 4$ and define $Y^0 = (0, 0, 0, 0, S_n^0, S_v^0, 0, 0)$. According to Lemma 1 in [31], this yields

$$F = \left(\frac{\partial \mathscr{F}_i}{\partial Y_j}(Y^0) \right)_{4\times 4} = \begin{pmatrix} 0 & 0 & \frac{\beta_{nn} S_n^0}{N^0} & \frac{\beta_{nv} S_n^0}{N^0} \\ 0 & 0 & \frac{\beta_{vn} S_v^0}{N^0} & \frac{\beta_{vv} S_v^0}{N^0} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{70}$$

and

$$V = \left(\frac{\partial \mathscr{V}_i}{\partial Y_j}(Y^0) \right)_{4\times 4} = \begin{pmatrix} (\sigma_n + \mu + \phi) & 0 & 0 & 0 \\ -\phi & (\sigma_v + \mu) & 0 & 0 \\ -\sigma_n & 0 & \mu + \gamma_n + \delta_n + \phi & 0 \\ 0 & -\sigma_v & -\phi & (\mu + \gamma_v + \delta_v) \end{pmatrix} \tag{71}$$

which subsequently results in

$$FV^{-1} = \begin{pmatrix} g_{11} & g_{12} & g_{13} & g_{14} \\ g_{21} & g_{22} & g_{23} & g_{24} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{72}$$

where

$$
\begin{aligned}
g_{11} &= \frac{S_n^0\beta_{nn}\sigma_n}{N^0(\mu+\phi+\sigma_n)(\delta_n+\gamma_n+\mu+\phi)} + \frac{S_n^0\beta_{nv}\phi\sigma_n}{N^0(\mu+\phi+\sigma_n)(\delta_n+\gamma_n+\mu+\phi)(\delta_v+\gamma_v+\mu)},\\
&\quad + \frac{S_n^0\beta_{nv}\phi\sigma_v}{N^0(\mu+\phi+\sigma_n)(\mu+\sigma_v)(\delta_v+\gamma_v+\mu)},\\
g_{12} &= \frac{S_n^0\beta_{nv}\sigma_v}{N^0(\mu+\sigma_v)(\delta_v+\gamma_v+\mu)},\\
g_{13} &= \frac{S_n^0}{N^0}\left[\frac{\beta_{nn}}{\delta_n+\gamma_n+\mu+\phi} + \frac{\beta_{nv}\phi}{(\delta_n+\gamma_n+\mu+\phi)(\delta_v+\gamma_v+\mu)}\right],\\
g_{14} &= \frac{S_n^0\beta_{nv}}{N^0(\delta_v+\gamma_v+\mu)},\\
g_{21} &= \frac{S_v^0\beta_{vn}\sigma_n}{N^0(\mu+\phi+\sigma_n)(\delta_n+\gamma_n+\mu+\phi)} + \frac{S_v^0\beta_{vv}\phi\sigma_n}{N^0(\mu+\phi+\sigma_n)(\delta_n+\gamma_n+\mu+\phi)(\delta_v+\gamma_v+\mu)},\\
&\quad + \frac{S_v^0\beta_{vv}\phi\sigma_v}{N^0(\mu+\phi+\sigma_n)(\mu+\sigma_v)(\delta_v+\gamma_v+\mu)},\\
g_{22} &= \frac{S_v^0\beta_{vv}\sigma_v}{N^0(\mu+\sigma_v)(\delta_v+\gamma_v+\mu)},\\
g_{23} &= \frac{S_v^0}{N^0}\left[\frac{\beta_{vn}}{\delta_n+\gamma_n+\mu+\phi} + \frac{\beta_{vv}\phi}{(\delta_n+\gamma_n+\mu+\phi)(\delta_v+\gamma_v+\mu)}\right],\\
g_{24} &= \frac{S_v^0\beta_{vv}}{N^0(\delta_v+\gamma_v+\mu).}
\end{aligned}
\tag{73}
$$

The characteristic equation for the matrix FV^{-1} can be expressed as follows:

$$
\lambda^4 + (-g_{11} - g_{22})\,\lambda^3 + (-g_{12}g_{21} + g_{11}g_{22})\,\lambda^2. \tag{74}
$$

Following this, the eigenvalues of (74) appear sequentially as follows:

$$
\begin{aligned}
\lambda_1 &= 0,\\
\lambda_2 &= 0,\\
\lambda_3 &= \frac{1}{2}\,g_{11} + \frac{1}{2}\,g_{22} - \frac{1}{2}\,\sqrt{g_{11}^2 + 4\,g_{12}g_{21} - 2\,g_{11}g_{22} + g_{22}^2},\\
\lambda_4 &= \frac{1}{2}\,g_{11} + \frac{1}{2}\,g_{22} + \frac{1}{2}\,\sqrt{g_{11}^2 + 4\,g_{12}g_{21} - 2\,g_{11}g_{22} + g_{22}^2}.
\end{aligned}
\tag{75}
$$

with

$$
\sqrt{g_{11}^2 + 4\,g_{12}g_{21} - 2\,g_{11}g_{22} + g_{22}^2} = (g_{11} - g_{22})^2 + 4g_{12}g_{21} > 0.
$$

In the next generation method approach, the basic reproduction number $\mathscr{R}_0$ is determined as the spectral radius of FV^{-1}, specifically:

$$
\begin{aligned}
\mathscr{R}_0 &= \rho(FV^{-1}),\\
&= \frac{1}{2}\,g_{11} + \frac{1}{2}\,g_{22} + \frac{1}{2}\,\sqrt{g_{11}^2 + 4\,g_{12}g_{21} - 2\,g_{11}g_{22} + g_{22}^2},
\end{aligned}
\tag{76}
$$

where all positive parameters g_{ij} for $(i, j = 1, 2, 3, 4)$ are specified in (73).

A.2. Jacobian Sub-Matrix Components for DFE Stability

The following are the detailed calculations of the component sub-matrices for the Jacobian matrix J_0 evaluated at the disease-free equilibrium (DFE), as referenced in the proof of Theorem 1.

$$J_{EE} = \begin{pmatrix} \frac{\partial E_n}{\partial E_n} & \frac{\partial E_n}{\partial E_v} \\ \frac{\partial E_v}{\partial E_n} & \frac{\partial E_v}{\partial E_v} \end{pmatrix} = \begin{pmatrix} -\frac{(\beta_{nn}I_n+\beta_{nv}I_v)S_n^0}{(N^0)^2} - (\sigma_n+\mu+\phi) & -\frac{(\beta_{nn}I_n+\beta_{nv}I_v)S_n^0}{(N^0)^2} \\ -\frac{(\beta_{vv}I_v+\beta_{vn}I_n)S_v^0}{(N^0)^2} + \phi & -\frac{(\beta_{vv}I_v+\beta_{vn}I_n)S_v^0}{(N^0)^2} - (\sigma_v+\mu) \end{pmatrix}$$

$$J_{EI} = \begin{pmatrix} \frac{\partial E_n}{\partial I_n} & \frac{\partial E_n}{\partial I_v} \\ \frac{\partial E_v}{\partial I_n} & \frac{\partial E_v}{\partial I_v} \end{pmatrix} = \begin{pmatrix} \frac{B_{nn}S_n^0}{N^0} - \frac{(\beta_{nn}I_n+\beta_{nv}I_v)S_n^0}{(N^0)^2} & \frac{\beta_{nv}S_n^0}{N^0} - \frac{(\beta_{nn}I_n+\beta_{nv}I_v)S_n^0}{(N^0)^2} \\ \frac{\beta_{vn}S_v^0}{N^0} - \frac{(\beta_{vv}I_v+\beta_{vn}I_n)S_v^0}{(N^0)^2} & \frac{\beta_{vv}S_v^0}{N^0} - \frac{(\beta_{vv}I_v+\beta_{vn}I_n)S_v^0}{(N^0)^2} \end{pmatrix}$$

$$J_{IE} = \begin{pmatrix} \frac{\partial I_n}{\partial E_n} & \frac{\partial I_n}{\partial E_v} \\ \frac{\partial I_v}{\partial E_n} & \frac{\partial I_v}{\partial E_v} \end{pmatrix} = \begin{pmatrix} \sigma_n & 0 \\ 0 & \sigma_v \end{pmatrix}$$

$$J_{II} = \begin{pmatrix} \frac{\partial I_n}{\partial I_n} & \frac{\partial I_n}{\partial I_v} \\ \frac{\partial I_v}{\partial I_n} & \frac{\partial I_v}{\partial I_v} \end{pmatrix} = \begin{pmatrix} -(\mu+\gamma_n+\delta_n+\phi) & 0 \\ \phi & -(\mu+\gamma_v+\delta_v) \end{pmatrix}$$

$$J_{ES} = \begin{pmatrix} \frac{\partial E_n}{\partial S_n^0} & \frac{\partial E_n}{\partial S_v^0} \\ \frac{\partial E_v}{\partial S_n^0} & \frac{\partial E_v}{\partial S_v^0} \end{pmatrix} = -(\beta_{nn}I_n+\beta_{nv}I_v) \begin{pmatrix} \frac{S_n^0}{(N^0)^2} - \frac{1}{N^0} & \frac{S_n^0}{(N^0)^2} \\ \frac{S_v^0}{(N^0)^2} & \frac{S_v^0}{(N^0)^2} - \frac{1}{N^0} \end{pmatrix}$$

$$J_{ER} = \begin{pmatrix} \frac{\partial E_n}{\partial R_n} & \frac{\partial E_n}{\partial R_v} \\ \frac{\partial E_v}{\partial R_n} & \frac{\partial E_v}{\partial R_v} \end{pmatrix} = -\frac{1}{(N^0)^2} \begin{pmatrix} (\beta_{nn}I_n^0+\beta_{nv}I_v^0) & (\beta_{nn}I_n^0+\beta_{nv}I_v^0) \\ (\beta_{vv}I_v^0+\beta_{vn}I_n^0) & (\beta_{vv}I_v^0+\beta_{vn}I_n^0) \end{pmatrix}$$

$$J_{IS} = \begin{pmatrix} \frac{\partial I_n}{\partial S_n} & \frac{\partial I_n}{\partial S_v} \\ \frac{\partial I_v}{\partial S_n} & \frac{\partial I_v}{\partial S_v} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$J_{IR} = \begin{pmatrix} \frac{\partial I_n}{\partial R_n} & \frac{\partial I_n}{\partial R_v} \\ \frac{\partial I_v}{\partial R_n} & \frac{\partial I_v}{\partial R_v} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$J_{SE} = \begin{pmatrix} \frac{\partial S_n}{\partial E_n} & \frac{\partial S_n}{\partial E_v} \\ \frac{\partial S_v}{\partial E_n} & \frac{\partial S_v}{\partial E_n} \end{pmatrix} = -\frac{1}{N^0} \begin{pmatrix} (\beta_{nn}I_n^0+\beta_{nv}I_v^0)S_n^0 & (\beta_{nn}I_n^0+\beta_{nv}I_v^0)S_n^0 \\ (\beta_{vv}I_v^0+\beta_{vn}I_n^0)S_v^0 & (\beta_{vv}I_v^0+\beta_{vn}I_n^0)S_v^0 \end{pmatrix}$$

$$J_{SI} = -J_{EI}$$

$$J_{RE} = \begin{pmatrix} \frac{\partial R_n}{\partial E_n} & \frac{\partial R_n}{\partial E_v} \\ \frac{\partial R_v}{\partial E_n} & \frac{\partial R_v}{\partial E_n} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$J_{RI} = \begin{pmatrix} \frac{\partial R_n}{\partial I_n} & \frac{\partial R_n}{\partial I_v} \\ \frac{\partial R_v}{\partial I_n} & \frac{\partial R_v}{\partial I_n} \end{pmatrix} = \begin{pmatrix} \gamma_n & 0 \\ 0 & \gamma_v \end{pmatrix}$$

$$J_{SS} = \begin{pmatrix} \frac{\partial S_n}{\partial S_n} & \frac{\partial S_n}{\partial S_v} \\ \frac{\partial S_v}{\partial S_n} & \frac{\partial S_v}{\partial S_n} \end{pmatrix} = \begin{pmatrix} \frac{(\beta_{nn}I_n^0+\beta_{nv}I_v^0)S_n^0}{(N^0)^2} - \frac{(\beta_{nn}I_n^0+\beta_{nv}I_v^0)}{N^0} - (\phi+\mu) & \omega_{sv} \\ \phi & \frac{(\beta_{vv}I_v^0+\beta_{vn}I_n^0)S_v^0}{(N^0)^2} - \frac{C_v}{N^0} - (\mu+\omega_{sv}) \end{pmatrix}$$

$$J_{SR} = J_{ER}$$

$$J_{RS} = \begin{pmatrix} \frac{\partial R_n}{\partial S_n} & \frac{\partial R_n}{\partial S_v} \\ \frac{\partial R_v}{\partial S_n} & \frac{\partial R_v}{\partial S_n} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$J_{RR} = \begin{pmatrix} \frac{\partial R_n}{\partial R_n} & \frac{\partial R_n}{\partial R_v} \\ \frac{\partial R_v}{\partial R_n} & \frac{\partial R_v}{\partial R_n} \end{pmatrix} = \begin{pmatrix} -(\mu+\phi+\omega_{rn}) & 0 \\ \phi & -(\mu+\omega_{rv}) \end{pmatrix}$$

A.3. Endemic Equilibrium Variable Derivations

The state variables at endemic equilibrium can be expressed as linear functions of E_n^* and E_v^*. The derivations proceed as follows:

$$\begin{aligned}
S_n^* &= \frac{k+\omega_{sv}S_v^* - \omega_{rn}R_n^* - \omega_{rv}R_v^* - (\sigma_n+\mu+\phi)E_n^*}{(\phi+\mu)} \coloneqq u_{no} + u_{nn}E_n^* + u_{nv}E_v^*, \\
S_v^* &= \frac{\phi S_n^* + (\sigma_v+\mu)E_v^* + \phi E_n^*}{(\mu+\omega_{sv})} \coloneqq u_{vo} + u_{vn}E_n^* + u_{vv}E_v^*, \\
I_n^* &= \frac{\sigma_n E_n^*}{(\mu+\gamma_n+\delta_n+\phi)} \coloneqq v_{nn}E_n^*, \\
I_v^* &= \frac{\phi\sigma_n E_n^*}{(\mu+\gamma_v+\delta_v)(\mu+\gamma_n+\delta_n+\phi)} + \frac{\sigma_v E_v^*}{(\mu+\gamma_v+\delta_v)} \coloneqq v_{vn}E_n^* + v_{vv}E_v^*, \\
R_n^* &= \frac{\gamma_n}{(\mu+\phi+\omega_r n)} \cdot \frac{\sigma_n E_n^*}{(\mu+\gamma_n+\delta_n+\phi)} \coloneqq x_{nn}E_n^*, \\
R_v^* &= \frac{\gamma_v\phi\sigma_n E_n^*}{(\mu+\omega_{rv}))(\mu+\gamma_v+\delta_v)(\mu+\gamma_n+\delta+\phi} + \frac{\phi\gamma_n\sigma_n E_n^*}{(\mu+\omega_{rv})(\mu+\phi+\omega_{rn})(\mu+\gamma_n+\delta_n+\phi)}, \\
&+ \frac{\gamma_v\sigma_v E_v^*}{(\mu+\omega_{rv})(\mu+\gamma_v+\delta_v)} \coloneqq x_{vn}E_n^* + x_{vv}E_v^*.
\end{aligned} \tag{77}$$

The total population size N^* is given by

$$\begin{aligned}
N^* &= S_n^* + S_v^* + E_n^* + E_v^* + I_n^* + I_v^* + R_n^* + R_v^*, \\
&= u_{no} + u_{nn}E_n^* + u_{nv}E_v^* + u_{vo} + u_{vn}E_n^* + u_{vv}E_v^* + v_{nn}E_n^* + v_{vn}E_n^* + v_{vv}E_v^* + x_{nn}E_n^* + x_{vn}E_n^* + x_{vv}E_v^*, \\
&= u_{no} + u_{vo} + u_{nn}E_n^* + u_{vn}E_n^* + v_{nn}E_n^* + v_{vn}E_n^* + x_{nn}E_n^* + x_{vn}E_n^* + u_{nv}E_v^* + u_{vv}E_v^* + v_{vv}E_v^* + x_{vv}E_v^*, \\
&\coloneqq w_o + w_n E_n^* + w_v E_n^*.
\end{aligned} \tag{78}$$

Substituting these into (23c) and (23d) yields the coefficients for equations (25) and (26) as follows:

$$\begin{aligned}
a_{nn} &= (\beta_{nn}v_{nn} + \beta_{nv}v_{vn})u_{nn} - (\sigma_n + \mu + \phi)\omega_n,\\
2a_n &= (\beta_{nn}v_{nn} + \beta_{nv}v_{vn})u_{no} - (\sigma_n + \mu + \phi)\omega_o,\\
2a_{nv} &= (\beta_{nv}v_{vv})u_{nn} - (\sigma_n + \mu + \phi)\omega_v,\\
2a_v &= (\beta_{nv}v_{vv})u_{no},\\
a_{vv} &= (\beta_{nv}v_{vv})u_{nv},\\
b_{nn} &= (\beta_{vn}v_{nn} + \beta_{vv}v_{vn}u_{vn} - \phi\omega_n),\\
2b_n &= (\beta_{vn}v_{nn} + \beta_{vv}v_{vn})u_{vo} - \phi\omega_o,\\
2b_{nv} &= (\beta_{vn}v_{vn} + \beta_{vv}v_{vn})u_{uv} + \beta_{vv}v_{vv}u_{vn} - (\sigma_v + \mu)\omega_n - \phi\omega_v,\\
2b_v &= (\beta_{vv}v_{vv}u_{vo} - (\sigma_v + \mu)\omega_o),\\
b_{vv} &= (\beta_{vv}v_{vv} - (\sigma_v + \mu)\omega_v).
\end{aligned} \tag{79}$$

A.4. Relationship Between the Stochastic and Deterministic Epidemic Threshold

In this Appendix section, the formal mathematical relationship between the stochastic expectation matrix $\mathscr{M}$ and the deterministic next-generation matrices F and V is demonstrated. The equation $F - V = (\mathscr{M} - I)W$ enables the calculation of the basic reproduction number $\mathscr{R}_0 = \rho(FV^{-1})$, which is the same as the stochastic threshold condition $\rho(\mathscr{M}) = 1$.
The 4×4 expectation matrix $\mathscr{M} = [m_{ji}]$ is given by:

$$\mathscr{M} = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14}\\ m_{21} & m_{22} & m_{23} & m_{24}\\ m_{31} & m_{32} & m_{33} & m_{34}\\ m_{41} & m_{42} & m_{43} & m_{44} \end{pmatrix}$$

Substituting the parameters from the branching process, this matrix is:

$$\mathscr{M} = \begin{pmatrix} 0 & 0 & \frac{\beta_{nn}S_n^0}{N^0} & \frac{\beta_{nv}S_n^0}{N^0}\\ & & \overline{\frac{\beta_{nn}S_n^0}{N^0} + \frac{\beta_{vn}S_v^0}{N^0} + \phi + \delta_n + \gamma_n + \mu} & \overline{\frac{\beta_{nv}S_n^0}{N^0} + \frac{\beta_{vv}S_v^0}{N^0} + \delta_v + \gamma_v + \mu}\\ \frac{\phi}{\sigma_n + \phi + \mu} & 0 & \frac{\beta_{vn}S_v^0}{N^0} & \frac{\beta_{vv}S_v^0}{N^0}\\ & & \overline{\frac{\beta_{nn}S_n^0}{N^0} + \frac{\beta_{vn}S_v^0}{N^0} + \phi + \delta_n + \gamma_n + \mu} & \overline{\frac{\beta_{nv}S_n^0}{N^0} + \frac{\beta_{vv}S_v^0}{N^0} + \delta_v + \gamma_v + \mu}\\ \frac{\sigma_n}{\sigma_n + \phi + \mu} & 0 & \frac{\beta_{nn}S_n^0}{N^0} + \frac{\beta_{vn}S_v^0}{N^0} & 0\\ & & \overline{\frac{\beta_{nn}S_n^0}{N^0} + \frac{\beta_{vn}S_v^0}{N^0} + \phi + \delta_n + \gamma_n + \mu} & \\ 0 & \frac{\sigma_v}{\sigma_v + \mu} & \phi & \frac{\beta_{nv}S_n^0}{N^0} + \frac{\beta_{vv}S_v^0}{N^0}\\ & & \overline{\frac{\beta_{nn}S_n^0}{N^0} + \frac{\beta_{vn}S_v^0}{N^0} + \phi + \delta_n + \gamma_n + \mu} & \overline{\frac{\beta_{nv}S_n^0}{N^0} + \frac{\beta_{vv}S_v^0}{N^0} + \delta_v + \gamma_v + \mu} \end{pmatrix}$$

First, the diagonal matrix W is defined, which contains the total transition rates w_i (the denominators from $\mathscr{M}$) out of each of the four infected states (E_n, E_v, I_n, I_v).

$$\begin{aligned}
w_1 &= \sigma_n + \phi + \mu\\
w_2 &= \sigma_v + \mu\\
w_3 &= \frac{\beta_{nn}S_n^0}{N^0} + \frac{\beta_{vn}S_v^0}{N^0} + \phi + \delta_n + \gamma_n + \mu\\
w_4 &= \frac{\beta_{nv}S_n^0}{N^0} + \frac{\beta_{vv}S_v^0}{N^0} + \delta_v + \gamma_v + \mu
\end{aligned}$$

The matrix W is explicitly:

$$W = \begin{pmatrix} w_1 & 0 & 0 & 0 \\ 0 & w_2 & 0 & 0 \\ 0 & 0 & w_3 & 0 \\ 0 & 0 & 0 & w_4 \end{pmatrix}$$

Next, the product $(\mathscr{M} - I)W$ is computed, where I is the identity matrix. This product represents the Jacobian of the infected compartments.

$$(\mathscr{M} - I)W = \begin{pmatrix} (m_{11}-1)w_1 & m_{12}w_2 & m_{13}w_3 & m_{14}w_4 \\ m_{21}w_1 & (m_{22}-1)w_2 & m_{23}w_3 & m_{24}w_4 \\ m_{31}w_1 & m_{32}w_2 & (m_{33}-1)w_3 & m_{34}w_4 \\ m_{41}w_1 & m_{42}w_2 & m_{43}w_3 & (m_{44}-1)w_4 \end{pmatrix}$$

By substituting the definitions of m_{ji} (from $\mathscr{M}$) and w_i, each term can be simplified. For example, the term $(1,1)$ is $(m_{11}-1)w_1 = (0-1)w_1 = -w_1 = -(\sigma_n + \phi + \mu)$. The non-diagonal term $(1,3)$ becomes $m_{13}w_3 = \left(\frac{\beta_{nn}S_n^0/N^0}{w_3}\right) w_3 = \frac{\beta_{nn}S_n^0}{N^0}$. The diagonal term $(3,3)$ simplifies as follows:

$$\begin{aligned} (m_{33}-1)w_3 &= \left(\frac{\frac{\beta_{nn}S_n^0}{N^0} + \frac{\beta_{vn}S_v^0}{N^0}}{w_3} - 1\right) w_3 \\ &= \left(\frac{\beta_{nn}S_n^0}{N^0} + \frac{\beta_{vn}S_v^0}{N^0}\right) - w_3 \\ &= \left(\frac{\beta_{nn}S_n^0}{N^0} + \frac{\beta_{vn}S_v^0}{N^0}\right) - \left(\frac{\beta_{nn}S_n^0}{N^0} + \frac{\beta_{vn}S_v^0}{N^0} + \phi + \delta_n + \gamma_n + \mu\right) \\ &= -(\phi + \delta_n + \gamma_n + \mu) \end{aligned}$$

Simplifying all terms in this manner yields the matrix $(\mathscr{M} - I)W$:

$$(\mathscr{M} - I)W = \begin{pmatrix} -(\sigma_n + \phi + \mu) & 0 & \frac{\beta_{nn}S_n^0}{N^0} & \frac{\beta_{nv}S_n^0}{N^0} \\ \phi & -(\sigma_v + \mu) & \frac{\beta_{vn}S_v^0}{N^0} & \frac{\beta_{vv}S_v^0}{N^0} \\ \sigma_n & 0 & -(\phi + \delta_n + \gamma_n + \mu) & 0 \\ 0 & \sigma_v & \phi & -(\delta_v + \gamma_v + \mu) \end{pmatrix}$$

From the theory of deterministic epidemic models, the Jacobian of the infected compartments evaluated at the Disease-Free Equilibrium is $F - V$, where F is the matrix of new infections and V is the matrix of transitions. For system (1), this matrix is:

$$F - V = \begin{pmatrix} -(\sigma_n + \phi + \mu) & 0 & \frac{\beta_{nn}S_n^0}{N^0} & \frac{\beta_{nv}S_n^0}{N^0} \\ \phi & -(\sigma_v + \mu) & \frac{\beta_{vn}S_v^0}{N^0} & \frac{\beta_{vv}S_v^0}{N^0} \\ \sigma_n & 0 & -(\phi + \delta_n + \gamma_n + \mu) & 0 \\ 0 & \sigma_v & \phi & -(\delta_v + \gamma_v + \mu) \end{pmatrix}$$

By inspection, it is clear that $(\mathscr{M} - I)W = F - V$. This confirms the formal equivalence between the stochastic and deterministic formulations. Additionally, this matrix can be broken down into F, which includes positive off-diagonal terms that depict new infections, and V, representing the negative of the remaining terms that account for transitions and removals.

$$F = \begin{pmatrix} 0 & 0 & \frac{\beta_{nn}S_n^0}{N^0} & \frac{\beta_{nv}S_n^0}{N^0} \\ 0 & 0 & \frac{\beta_{vn}S_v^0}{N^0} & \frac{\beta_{vv}S_v^0}{N^0} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$V = \begin{pmatrix} \sigma_n + \phi + \mu & 0 & 0 & 0 \\ -\phi & \sigma_v + \mu & 0 & 0 \\ -\sigma_n & 0 & \phi + \delta_n + \gamma_n + \mu & 0 \\ 0 & -\sigma_v & -\phi & \delta_v + \gamma_v + \mu \end{pmatrix}$$

The condition for a major outbreak in the stochastic model, $\rho(\mathscr{M}) > 1$, is identical to the condition for an epidemic in the deterministic model, $\mathscr{R}_0 > 1$, where $\mathscr{R}_0 = \rho(FV^{-1})$. Therefore, a closed-form analytical solution for the eigenvalues of $\mathscr{M}$ is not required. The epidemic threshold is fully described by the interpretable, analytical expression for $\mathscr{R}_0 = \rho(FV^{-1})$, which can be derived from the matrices F and V above.

Data Availability Statement

The data that supporting the findings of this study are openly available at [38].

Authorship Contribution

Aubrey Ndovie: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Visualization. **Claris Shoko:** Conceptualization, Methodology, Supervision, Writing - Review & Editing. **Olusegun S. Ewemooje:** Supervision, Validation, Writing - Review & Editing. **Sivasamy Ramasamy:** Supervision, Validation, Writing - Review & Editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to acknowledge the University of Botswana for providing the platform and environment conducive to this research.

References

1. W. jie Guan, Z. yi Ni, Y. Hu, W. hua Liang, C. quan Ou, J. xing He, L. Liu, H. Shan, C. liang Lei, D. S. Hui, B. Du, L. juan Li, G. Zeng, K.-Y. Yuen, R. chong Chen, C. li Tang, T. Wang, P. yan Chen, J. Xiang, S. yue Li, J. lin Wang, Z. jing Liang, Y. xiang Peng, L. Wei, Y. Liu, Y. hua Hu, P. Peng, J. ming Wang, J. yang Liu, Z. Chen, G. Li, Z. jian Zheng, S. qin Qiu, J. Luo, C. jiang Ye, S. yong Zhu, and N. shan Zhong, "Clinical characteristics of coronavirus disease 2019 in china," *New England Journal of Medicine*, vol. 382, no. 18, pp. 1708–1720, 2020.
2. WHO, "Covid-19 epidemiological update - 2023-11-15." Accessed on 2023-11-15.
3. S. A. Lone and A. Ahmad, "Covid-19 pandemic – an african perspective," *Emerging Microbes and Infections*, vol. 9, no. 1, pp. 1300–1308, 2020.
4. E. Yingi and P. Hlungwani, "The covid-19 pandemic and regional integration in africa: Implications of the responses from ecowas and sadc," *South African Journal of International Affairs*, vol. 29, no. 2, pp. 227–241, 2022.
5. J. E. Joseph, "The south african development community (sadc) and covid-19: revisiting security community in sadc," *EUREKA: Social and Humanities*, vol. 5, pp. 69–81, 2021.
6. Y. Shi, G. Wang, X. Cai, J. Deng, L. Zheng, H. Zhu, M. Zheng, B. Yang, and Z. Chen, "An overview of COVID-19," *Journal of Zhejiang University. Science. B*, vol. 21, no. 5, pp. 343–360, 2020.
7. A. Adebowale, R. Afolabi, S. Bello, *et al.*, "Spread and seasonality of covid-19 pandemic confirmed cases in sub-saharan africa: Experience from democratic republic of congo, nigeria, senegal, and uganda," *BMC Infectious Diseases*, vol. 23, p. 187, 2023.
8. M. Betti, N. Bragazzi, J. Heffernan, J. Kong, and A. Raad, "Could a new covid-19 mutant strain undermine vaccination efforts? a mathematical modelling approach for estimating the spread of b.1.1.7 using ontario, canada, as a case study," *Vaccines*, vol. 9, no. 6, 2021.
9. O. Yagan, A. Sridhar, R. Eletreby, S. Levin, J. Plotkin, and H. V. Poor, "Modeling and analysis of the spread of covid-19 under a multiple-strain model with mutations," *Harvard Data Science Review*, 03 2021.
10. C. J. Makgala and I. S. Malila, "Challenges of constitutional reform, economic transformation and covid-19 in botswana," *Review of African Political Economy*, vol. 49, no. 172, pp. 303–314, 2022.
11. S. Prakash, "Development of covid 19 vaccine: A summarized review on global trials, efficacy, and effectiveness on variants," *Diabetes and Metabolic Syndrome*, vol. 16, no. 4, p. 102482, 2022.
12. D. R. Feikin, M. M. Higdon, L. J. Abu-Raddad, N. Andrews, R. Araos, Y. Goldberg, M. J. Groome, A. Huppert, K. L. O'Brien, P. G. Smith, A. Wilder-Smith, S. Zeger, M. Deloria Knoll, and M. K. Patel, "Duration of effectiveness of vaccines against sars-cov-2

infection and covid-19 disease: results of a systematic review and meta-regression," *The Lancet*, vol. 399, no. 10328, pp. 924–944, 2022.
13. R. M. Anderson and R. M. May, *Infectious diseases of humans: dynamics and control*. Oxford university press, 1991.
14. L. J. Allen and P. van den Driessche, "Relations between deterministic and stochastic thresholds for disease extinction in continuous- and discrete-time infectious disease models," *Mathematical biosciences*, vol. 243, no. 1, pp. 99–108, 2013.
15. D.-E. Gratie, B. Iancu, and I. Petre, "Ode analysis of biological systems," *Formal Methods for Dynamical Systems: 13th International School on Formal Methods for the Design of Computer, Communication, and Software Systems, SFM 2013, Bertinoro, Italy, June 17-22, 2013. Advanced Lectures*, pp. 29–62, 2013.
16. L. J. Allen and G. E. Lahodny Jr, "Extinction thresholds in deterministic and stochastic epidemic models," *Journal of biological dynamics*, vol. 6, no. 2, pp. 590–611, 2012.
17. L. J. Allen, "Stochastic population and epidemic models," *Mathematical biosciences lecture series, stochastics in biological systems*, vol. 1, pp. 120–128, 2015.
18. W. J. Anderson, *Continuous-time Markov chains: An applications-oriented approach*. Springer Science & Business Media, 2012.
19. L. J. Allen, "An introduction to stochastic epidemic models," in *Mathematical epidemiology*, pp. 81–130, Springer, 2008.
20. L. Monteiro, "An epidemiological model for sars-cov-2," *Ecological complexity*, vol. 43, p. 100836, 2020.
21. J. M. Carcione, J. E. Santos, C. Bagaini, and J. Ba, "A simulation of a covid-19 epidemic based on a deterministic seir model," *Frontiers in public health*, vol. 8, p. 230, 2020.
22. F. A. Rihan, H. J. Alsakaji, and C. Rajivganthi, "Stochastic sirc epidemic model with time-delay for covid-19," *Advances in difference equations*, vol. 2020, no. 1, p. 502, 2020.
23. M. N. Hassan, M. S. Mahmud, K. F. Nipa, and M. Kamrujjaman, "Mathematical modeling and covid-19 forecast in texas, usa: a prediction model analysis and the probability of disease outbreak," *Disaster medicine and public health preparedness*, vol. 17, p. e19, 2023.
24. C. M. Saad-Roy, C. E. Wagner, R. E. Baker, S. E. Morris, J. Farrar, A. L. Graham, S. A. Levin, M. J. Mina, C. J. E. Metcalf, and B. T. Grenfell, "Immune life history, vaccination, and the dynamics of sars-cov-2 over the next 5 years," *Science*, vol. 370, no. 6518, pp. 811–818, 2020.
25. T. Britton, F. Ball, and P. Trapman, "A mathematical model reveals the influence of population heterogeneity on herd immunity to sars-cov-2," *Science*, vol. 369, no. 6505, pp. 846–849, 2020.
26. W. O. Kermack and A. G. McKendrick, "A contribution to the mathematical theory of epidemics," *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, vol. 115, no. 772, pp. 700–721, 1927.
27. T. M. León, V. Dorabawila, L. Nelson, E. Lutterloh, U. E. Bauer, B. Backenson, M. T. Bassett, H. Henry, B. Bregman, C. M. Midgley, J. F. Myers, I. D. Plumb, H. E. Reese, R. Zhao, M. Briggs-Hagen, D. Hoefer, J. P. Watt, B. J. Silk, S. Jain, and E. S. Rosenberg, "Covid-19 cases and hospitalizations by covid-19 vaccination status and previous covid-19 diagnosis — california and new york, may–november 2021," *MMWR. Morbidity and Mortality Weekly Report*, vol. 71, no. 4, pp. 125–131, 2022.
28. R. Zhou and F. Wei, "An age-structured epidemic model with vaccination," *arXiv preprint arXiv:2205.03912*, 2022.
29. S. E. Oliver, J. W. Gargano, M. Marin, M. Wallace, and A. T, Mba-Jonas, "The advisory committee on immunization practices' interim recommendation for use of mrna covid-19 vaccines — united states, december 2020," *MMWR. Morbidity and Mortality Weekly Report*, vol. 69, no. 50, pp. 1922–1924, 2020.
30. G. Birkhoff and G. Rota, *Ordinary Differential Equations*. Introductions to higher mathematics, Ginn, 1962.
31. P. Van den Driessche and J. Watmough, "Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission," *Mathematical biosciences*, vol. 180, no. 1-2, pp. 29–48, 2002.
32. B. E. Meserve, *Fundamental concepts of algebra*. Courier Corporation, 1982.
33. D. Curtiss, "Recent extentions of descartes' rule of signs," *Annals of Mathematics*, vol. 19, no. 4, pp. 251–278, 1918.
34. R. Descartes, *The Geometry of Rene Descartes: With a facsimile of the first edition*. Courier Corporation, 2012.
35. K. B. Athreya, P. E. Ney, and P. Ney, *Branching processes*. Courier Corporation, 2004.
36. K. S. Dorman, J. S. Sinsheimer, and K. Lange, "In the garden of branching processes," *SIAM review*, vol. 46, no. 2, pp. 202–229, 2004.
37. L. J. Allen, *An introduction to stochastic processes with applications to biology*. CRC press, 2010.
38. V. Marivate, R. Arbi, H. Combrink, A. de Waal, H. Dryza, D. Egersdorfer, S. Garnett, B. Gordon, L. Greyling, O. Lebogo, D. Mackie, B. Merry, S. Mkhondwane, M. Mokoatle, S. Moodley, J. Mtsweni, N. Mtsweni, P. Myburgh, J. Richter, V. Rikhotso, S. Rosen, J. Sefara, A. van der Walt, S. van Heerden, J. Welsh, S. Hazelhurst, C. Petersen, R. Mbuvha, N. Dhlamini, and V. James, "Coronavirus disease (COVID-19) case data - South Africa," Mar. 2020.
39. W. Yang and J. L. Shaman, "Covid-19 pandemic dynamics in south africa and epidemiological characteristics of three variants of concern (beta, delta, and omicron)," *eLife*, vol. 11, p. e78933, aug 2022.
40. W. Jassat, S. S. Abdool Karim, C. Mudara, R. Welch, L. Ozougwu, M. J. Groome, N. Govender, A. von Gottberg, N. Wolter, M. Wolmarans, P. Rousseau, L. Blumberg, and C. Cohen, "Clinical severity of covid-19 in patients admitted to hospital during the omicron wave in south africa: a retrospective observational study," *The Lancet Global Health*, vol. 10, no. 7, pp. e961–e969, 2022.
41. G. Chowell, "Fitting dynamic models to epidemic outbreaks with quantified uncertainty: A primer for parameter uncertainty, identifiability, and forecasts," *Infectious Disease Modelling*, vol. 2, no. 3, pp. 379–398, 2017.
42. M. A. Billah, M. M. Miah, and M. N. Khan, "Reproductive number of coronavirus: A systematic review and meta-analysis based on global level evidence," *PLoS ONE*, vol. 15, no. 11, p. e0242128, 2020.
43. P. Ssentongo, A. E. Ssentongo, E. S. Heilbrunn, D. M. Ba, and V. M. Chinchilli, "Sars-cov-2 vaccine effectiveness against infection, symptomatic and severe covid-19: a meta-analysis," *BMC Infectious Diseases*, vol. 22, no. 1, p. 439, 2022.
44. K. Talaat, Y. Goldberg, A. Freedman, *et al.*, "Had covid recently? here's what to know about how long immunity lasts," *AAMC News*, 2023. Online article.
45. C. G. Smith, M. K. Jones, *et al.*, "Long-term data reveals sars-cov-2 infection- and vaccine-induced antibody responses are long-lasting," *Immunity*, vol. 57, no. 2, pp. 318–330, 2024.

46. G. Meyerowitz-Katz and L. Merone, "A systematic review and meta-analysis of published research data on covid-19 infection fatality rates," *International Journal of Infectious Diseases*, vol. 101, pp. 138–148, 2020.
47. Statistics South Africa, "Covid-19 epidemic reduces life expectancy in 2021," jul 2021.
48. S. Marino, I. B. Hogue, C. J. Ray, and D. E. Kirschner, "A methodology for performing global uncertainty and sensitivity analysis in systems biology," *Journal of Theoretical Biology*, vol. 254, no. 1, pp. 178–196, 2008.
49. M. Zevika, A. Triska, N. Nuraini, and G. Lahodny Jr., "On the study of covid-19 transmission using deterministic and stochastic models with vaccination treatment and quarantine," *Communication in Biomathematical Sciences*, vol. 5, no. 1, pp. 1–19, 2022.
50. J. Hellewell, S. Abbott, A. Gimma, N. I. Bosse, C. I. Jarvis, T. W. Russell, J. D. Munday, A. J. Kucharski, W. J. Edmunds, C. for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, *et al.*, "Feasibility of controlling covid-19 outbreaks by isolation of cases and contacts," *The Lancet Global Health*, vol. 8, no. 4, pp. e488–e496, 2020.
51. H. B. Taboe, M. Asare-Baah, A. Yesmin, and C. N. Ngonghala, "Impact of age structure and vaccine prioritization on covid-19 in west africa," *Infectious Disease Modelling*, vol. 7, no. 4, pp. 151–173, 2022.
52. C. J. Edholm, B. Levy, L. Spence, F. B. Agusto, F. Chirove, C. W. Chukwu, D. Goldsman, M. Kgosimore, I. Maposa, K. White, and S. Lenhart, "A vaccination model for covid-19 in gauteng, south africa," *Infectious Disease Modelling*, vol. 7, no. 3, pp. 333–345, 2022.