BIOMATH
https://biomath.math.bas.bg/biomath/index.php/biomath
Biomath Forum

ORIGINAL ARTICLE

A modified SIR model with dependent infection and hospitalization rates

Janice Moore, Meir Shillor

Department of Mathematics and Statistics,
Oakland University, Rochester, Michigan, USA
jhmoore@oakland.edu
shillor@oakland.edu

Received: May 13, 2024, Accepted: November 14, 2024, Published: December 2, 2024

Abstract: **This work constructs, analyzes and simulates the SIR-IH model, a modified SIR epidemiological model for the spread of a disease, in which the infection rate and hospitalization ratio are system variables. The motivation, in part, of making the infection rate a state variable comes from the observations that the infectivity of a disease, such as COVID-19, has been changing with the evolution of the disease, and not necessarily by the appearance of new variants. Moreover, it may change in time if more than one variant are present. The addition of a hospitalization rate is done to make the model more applied for those who need to make decisions on the preparedness of the health system, in particular the hospitals, in case of a pandemic. The model consists of a coupled system of differential equations, and its analysis shows the existence, positivity and boundedness of the solutions. Then, computer simulations depict some typical or interesting dynamic behaviors, and the way the system approaches the steady states.**

Keywords: **extended SIR model, variable hospitalization rates, variable infectivity, analysis, simulations**

MSC2020: **92D30, 34C23, 34C60, 92C60**

I. Introduction

We develop, study mathematically, and computationally simulate a new extended version of the SIR model for the spread of a communicable disease, in which the infected population is split into those individuals who have mild or medium symptoms, and those who have severe symptoms that make hospitalization necessary. The novelty in this work lies in the assumption that the infectivity β and the hospitalization ratio α^*, both of which measure the severity of the disease, are system dependent variables. For the sake of simplicity, the rate equations for both variables are linear in the populations, however, it is very likely that more accurate relationships are more complex, and will need to be determined for each disease separately, by some parameter identification method. This extension of the model is meant to open a way to take into account that some diseases change over time, and both variables are time dependent. In the standard epidemiological models both the infectivity rate and the hospitalization ratio are assumed to be constant, or change in time in known ways. Here, we allow them to evolve with the disease, as well.

The concept of system-dependent infectivity rate was already considered in [1]. Our approach here is quite general and we do not have, at this stage of the research, a definite disease to which the general model may be applied. We hope that this work will help find those diseases the dynamics of which would benefit from such an extended modeling.

Infectious diseases are a major cause of suffering, morbidity, and mortality worldwide. In an effort to minimize the impact of an infectious disease, epidemiologists must both understand the causes of the disease and be able to predict its progression. To this end, mathematical models were developed as early as 1906.

Citation: Janice Moore, Meir Shillor, A modified SIR model with dependent infection and hospitalization rates, Biomath 13 (2024), 2411146, https://doi.org/10.55630/j.biomath.2024.11.146

Public health physicians including William Hamer and Ronald Ross laid the foundations for compartmentalized models. Hamer proposed that the spread of infection should depend on the number of susceptible individuals and the number of infective individuals. This idea led to compartmental models, the most basic of which is the SIR model. Ross introduced the basic reproduction number in his study of eradicating malaria by reducing, rather than eliminating, the population of mosquitoes [2]. The number, variety, and complexity of compartmental models have grown extensively since then.

Herbert W. Hethcote was one of the major contributors to the development of these types of models from 1970 until 2009. His most cited work is "The Mathematics of Infectious Diseases," [3], which describes how to build a mathematical, compartmental, epidemiological model and analyzes how to obtain the basic reproduction number, R_0, the contact number, σ, and the replacement number R. Some compartmentalized models are very general and designed to understand the behavior of a disease or diseases qualitatively. Others are more specific and detailed, with the purpose of making quantitative predictions on the behavior of a particular disease.

Our model is of the general variety with the intent of understanding the impact of varying infection rates and hospitalization rates on the progression, over time, of a rather general disease. Such changes may be because the agent causing of the disease evolves, undergoes mutations, or just there are a few versions the relative ratio of which is changing.

Following this introduction, we describe the model in Section 2, where we present the compartmental structure and the flow chart. We provide explanations of the equations for the infectivity rate function β and the of hospitalization ratio function α^*. The model consists of six differential equations for the four populations: susceptibles, mildly infected, hospitalized and recovered, and the two rates. The section also provides a table with the description of the various system parameters. Section 3 analyzes the model. The populations are shown to be nonnegative and bounded, which leads to the existence of the unique solution to the model.

The stability of the Disease Free Equilibrium (DFE) is shown in Section 4, using the Jacobian matrix for the populations. Also, the basic stability number, $\mathcal{R}_0$, is discussed and shown to be below the value 1 in the baseline case. We use the ideas of Diekmann et al., [4] to compute $\mathcal{R}_0$, dealing with the reduced system that includes only the infection states. For more recent references, we refer the reader to [3, 5] and the host of references therein. The conditions for the existence of the Endemic Equilibrium (EE) are described shortly in Section 5.

The model simulations are presented in Section 6, where a simple explicit time stepping scheme is presented, and baseline simulations are depicted, and it is seen that in this case the DFE is stable and attracting. Simulations are also provided in which the system tends asymptotically to an EE state. Section 7 provides a short summary of the results, some conclusions, and suggests some topics for future research.

II. The model

This section presents a modified general SIR model of disease transmission that takes into account dependent rates of infection and hospitalization. To fix ideas, we use the terminology of a generic disease in a large city, in which the disease is spreading. The modifications of the model to specific diseases and locations are straightforward.

We deal with whole populations, do not consider their spatial spread, and assume that the populations are large enough to justify the use of a continuous description, that is, using ordinary differential equations (ODEs). As was noted in the introduction, we describe the dynamics of four populations: susceptibles S, those who are healthy and can become infected; I_h, infected who need hospitalization; I_m, mildly infected who do not need hospitalization; and R recovered individuals. These are functions of time, which we measure in days, and the rate of change of each population is described below. We also assume, and that is the novelty in this work, that the rate of infection constant β, which is the probability that one susceptible coming in contact with one infected becomes sick, measuring how rapidly the disease may spread; and the disease severity function α^* that is the hospitalization fraction, are both dependent variables.

We assume that the infectivity rate function β and the seriousness or severity rate function α^* are only mildly related, and we use a differential equation for each one that depends on the infected and recovered populations. Moreover, we assume that the fraction ρ of those who recover can loose their immunity and be reinfected.

The compartmental structure of the model is depicted in Fig. 1.

The susceptible population $S(t)$ consists of all the individuals who are alive and do not belong to any of the other groups. The sub-population $I_m(t)$ denotes individuals with mild clinical symptoms of the disease,

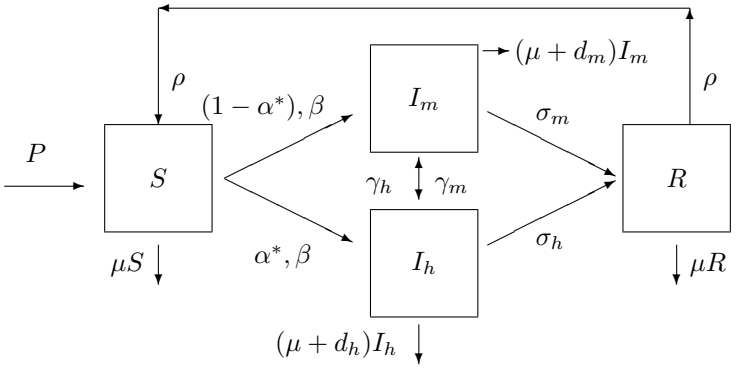

Fig. 1: Compartmental structure and flow chart, β and α^* are dependent variables, too.

or have no symptoms, who recover at home and do not need hospitalization. The sub-population $I_h(t)$ denotes the hospitalized individuals, and the sub-population $R(t)$ denotes those who recovered from the disease, and a fraction of which can become sick again.

We denote by $P(t)$ the net number of individuals that are added to the city per day by birth or arrive from the outside, and μ $(1/day)$ the death rate in absence of the disease. We assume that the disease increases the death rates $\mu + d_m$ $(1/day)$ for those with a mild condition, and $\mu + d_h$ $(1/day)$, for those hospitalized, with $0 \le d_m \le d_h$.

The probability that one contact between a susceptible and a mildly infected results in infection is $\beta(t)$ $(1/day)$, the probability that one contact between a susceptible and a hospitalized individual results in infection is $\epsilon_h\beta$ $(1/day)$, where $0 < \epsilon_h$, is a correction factor, since although interactions between susceptibles and hospitalized are limited, they may be more or less infectious. We denote by $\alpha^* = \alpha^*(t)$ $(0 \le \alpha^* \le 1)$ the fraction of those infected who need hospitalization, and then $1 - \alpha^*$ is the fraction of those with a mild condition.

The rates of interactions between the hospitalized and the mild ones are γ_h and γ_m, so that the number of mild individuals (per day) who become hospitalized is $\gamma_m I_m$ and the number hospitalized individuals (per day) who become mild and leave the hospital, is $\gamma_h I_h$. The disease recovery constants are σ_m and σ_h and, then, the number of recovered (per day) is $\sigma_m I_m$ for the mild and $\sigma_h I_h$ for the hospitalized. Next, we denote by $\gamma_m^+, \gamma_h^+, \gamma^-$ the rate constants for β and $\delta_m^+, \delta_h^+, \delta^-$ the rate constants for α^*. We let $\rho \in (0, 1]$ denote the fraction of the recovered that become susceptibles again.

The equations for β and α^* are likely to be of the general form

$$\begin{aligned} \frac{d\beta}{dt} &= \Psi_\beta(I_m, I_h, R, \beta, \alpha^*), \\ \frac{d\alpha^*}{dt} &= \Psi_*(I_m, I_h, R, \beta, \alpha^*), \end{aligned} \tag{1}$$

for given functions Φ and Ψ, which need to be found for each disease. Actually, these are likely to be set inclusions, and we comment on this in Remark 1 below.

Since β, α^* are dependent variables, we need to make sure that their values remain in reasonable intervals, no matter what the populations are. To that end, let $0 < \beta_- < \beta_+ \le 1$ and $0 \le \alpha_- < \alpha_+ \le 1$ be the limits for the allowed values of β and α^*, respectively, and we want to guarantee that for $0 \le t \le T$,

$$\beta(t) \in [\beta_-, \beta_+], \qquad \alpha^*(t) \in [\alpha_-, \alpha_+]. \tag{2}$$

We next introduce the restriction operators Φ_β, Φ_*, acting on $s \in \mathbb{R}$, as follows,

$$\Phi_\beta(s) = \begin{cases} \beta_-, & s \le \beta_-, \\ s, & \beta_- < s < \beta_+, \\ \beta_+ & s \ge \beta_+, \end{cases}$$

$$\Phi_*(s) = \begin{cases} \alpha_-, & s \le \alpha_-, \\ s, & \alpha_- < s < \alpha_+, \\ \alpha_+ & s \ge \alpha_+. \end{cases}$$

Both operators are piecewise linear and Lipschitz.

The total living population, at time t, is given by

$$N(t) = S(t) + I_m(t) + I_h(t) + R(t), \tag{3}$$

and is not assumed to be constant.

These model assumptions lead to the following *mathematical model for the dynamics of a disease with variable infection and severity or hospitalization rates*.

Model 1. *Find six functions* $(S, I_m, I_h, R, \widetilde{\beta}, \widetilde{\alpha^*})$*, defined on* $[0,T]$*, such that the following equations hold, for* $0 < t \leq T$*:*

$$\frac{dS}{dt} = P - (\beta I_m + \epsilon_h \beta I_h)\frac{S}{N} - \mu S + \rho R, \tag{4}$$

$$\frac{dI_m}{dt} = (1-\alpha^*)(\beta I_m + \epsilon_h \beta I_h)\frac{S}{N} + \gamma_h I_h - (\gamma_m + \sigma_m + \mu + d_m) I_m, \tag{5}$$

$$\frac{dI_h}{dt} = \alpha^*(\beta I_m + \epsilon_h \beta I_h)\frac{S}{N} + \gamma_m I_m - (\gamma_h + \sigma_h + \mu + d_h) I_h, \tag{6}$$

$$\frac{dR}{dt} = \sigma_m I_m + \sigma_h I_h - (\mu + \rho) R, \tag{7}$$

$$\frac{d\widetilde{\beta}}{dt} = \gamma_m^+ I_m + \gamma_h^+ I_h - \gamma^- \widetilde{\beta} R, \tag{8}$$

$$\frac{d\widetilde{\alpha^*}}{dt} = \delta_m^+ I_m + \delta_h^+ I_h - \delta^- \widetilde{\alpha^*} R, \tag{9}$$

where

$$\beta(t) = \Phi_\beta(\widetilde{\beta(t)}), \qquad \alpha^*(t) = \Phi_*(\widetilde{\alpha^*(t)}), \tag{10}$$

together with N *given in (3). The initial conditions are*

$$\begin{aligned} &S(0) = S_0, \qquad I_m(0) = I_{m0}, \quad I_h(0) = I_{h0}, \\ &R(0) = R_{init}, \quad \beta(0) = \beta_1, \quad \alpha^*(0) = \alpha_0^*. \end{aligned} \tag{11}$$

Here, $\widetilde{\beta}$ and $\widetilde{\alpha^*}$ are the unrestricted auxiliary dependent variables, and everywhere below we refer to the restricted variables β and α^* as the infection and hospitalization rates.

Moreover, $S_0 > 0$ is the initial population before the break-out of the disease, and $I_{m0}, I_{h0} \geq 0$, $R_{init} \geq 0$ are the initial populations in (11) at $t = 0$. Additionally, $\beta_1 \in [\beta_-, \beta_+]$ is the initial infectivity, and $\alpha^*(0) = \alpha_0^* \in [\alpha_-, \alpha_+]$.

In practice, one typically assumes that initially $S_0 = N(0) > 0$, so that at the start the population is only of susceptibles, and all the other populations vanish. However, for the sake of generality, we allow for them to be nonnegative but satisfying $N(0) = S_0 + I_{m0} + I_{h0} + R_{init}$, so that (3) holds at $t = 0$.

Next, we discuss equations (8) and (9). We assume that the infectivity of the disease increases with the number of infected, I_m, I_h, as the disease vector has more options to mutate, and decreases with the number of recovered, as more individuals are immune. As noted above, we assume that the relationship is linear, which is probably a severe restriction. We plan to look into the question of the form of Φ_β and Φ_* in the future, since it deserves a thorough investigation.

We note that (8) and (9) can be integrated, thus,

$$\widetilde{\beta(t)} = \beta_1 + \int_0^t \left(\gamma_m^+ I_m(s) + \gamma_h^+ I_h(s) - \gamma^- R(s)\right) ds,$$

$$\widetilde{\alpha^*(t)} = \alpha_0^* + \int_0^t \left(\delta_m^+ I_m(s) + \delta_h^+ I_h(s) - \delta^- R(s)\right) ds,$$

which makes the system (4)–(7) into a history dependent system of integral equations. We do not pursue further this form of the model here.

Remark 1. *We note that instead of using the restriction operators* Φ_β, Φ_**, we may use the notion of the subdifferential. To that end, we let* I_β, I_α *be the indicator functions of* $[\beta_-, \beta_+]$ *and* $[\alpha_-, \alpha_+]$*, that is* $I_\beta(\beta) = 0$ *if* $\beta \in [\beta_-, \beta_+]$ *and* $I_\beta = +\infty$ *otherwise, and similarly,* $I_\alpha(\alpha^*) = 0$ *if* $\alpha \in [\alpha_-, \alpha_+]$ *and* $I_\alpha = +\infty$ *otherwise. Let* $\partial I_\beta, \partial I_\alpha$ *be the subdifferentials,*

$$\partial I_\beta(s) = \begin{cases} (-\infty, 0], & s = \beta_-, \\ 0, & \beta_- < s < \beta_+, \\ [0, \infty)) & s = \beta_+, \end{cases}$$

and similarly for ∂I_α*, with* α *replacing* β*.*

Then, instead of using restrictions, we may use the set-valued inclusions for β *and* α^**, which are*

$$\frac{d\beta}{dt} \in \Psi_\beta(I_m, I_h, R, \beta, \alpha^*) - \partial I_\beta(\beta),$$

$$\frac{d\alpha^*}{dt} \in \Psi_*(I_m, I_h, R, \beta, \alpha^*) - \partial I_\beta(\alpha^*).$$

The subdifferentials guarantee that the variables cannot leave the designated intervals.

A summary of the definitions of the model parameters are given in Table 1.

III. Model analysis

This section establishes the non-negativity, boundedness, and then existence of solutions to Model 1. Indeed, for the model to makes sense biologically, the solutions must be non-negative and bounded. We have the following positivity and boundedness result.

Theorem 1. *(Positivity and estimates) Let* $S_0 > 0$ *and let* I_{m0}, I_{h0}, R_{init} *be nonnegative. If* $z(t) = (S(t), I_m(t), I_h(t), R(t))$ *is any solution to the system, then all of the components of* z *are nonnegative and uniformly bounded on each closed interval* $[0,T]$ *on which they exist. Then, for* $0 \leq t < T$*,*

$$0 \leq S, I_m, I_h, R \leq N_0 + \frac{P}{\mu}.$$

Parameter	Description
P	recruitment rate of susceptible individuals
β	restricted infection transmission rate (11)
α^*	restricted hospitalization rate (11)
ϵ_h	correction factor in transmission rate by hospitalized
μ	natural death rate
d_m	disease-induced death for mildly infected
d_h	disease-induced death for hospitalized
γ_m	hospitalization rate of mildly infected
γ_h	release rate of hospitalized
σ_m	recovery rate of mildly infected
σ_h	recovery rate of hospitalized
ρ	reinfection rate of recovered
$\gamma_m^+, \gamma_h^+, \gamma^-$	infectivity change rates
$\delta_m^+, \delta_h^+, \delta^-$	hospitalization change rates
$\beta_-, \beta_+, \alpha_-, \alpha_+$	restriction intervals values
$S_0, I_{m0}, I_{h0}, R_{init},$	initial populations
β_1, α_0^*	initial rates

Table 1: Symbols and description of parameters used in the model, the dimensional quantities are per day

Moreover,

$$0 \le \widetilde{\beta(t)} \le \beta_1 + \max\{\gamma_m^+, \gamma_h^+\}(N_0 + P/\mu)T.$$

and

$$0 \le \widetilde{\alpha^*(t)} \le \alpha_0^* + \max\{\delta_m^+, \delta_h^+\}(N_0 + P/\mu)T.$$

It follows from (10) that $\beta(t) \in [\beta_-, \beta_+]$ and $\alpha^*(t) \in [\alpha_-, \alpha_+]$, for $0 \le t \le T$.

Proof: We first assume that $R_{init} > 0$ and because of the continuity of the solutions, there exists a, possibly very short, time interval with T^* such that $R(t) > 0$ for $0 \le t < T^*$. Let $0 < t < T < T^*$ and denote by λ the *force of infection*

$$\lambda = \lambda(t) = \frac{\beta I_m + \varepsilon_h \beta I_h}{N}.$$

Then, $0 \le \lambda \le \beta(1 + \varepsilon_h)$ and (4) yields,

$$\frac{dS}{dt} = P - \lambda S - \mu S + \rho R \ge -(\lambda + \mu)S.$$

Using separation of variables and integration on $[0, t]$, together with the initial conditions, yields

$$S(t) \ge S_0 \exp\left(-\int_0^t \lambda(s)ds - \mu t\right) > 0,$$

since $S_0 > 0$. Similarly, for (5)–(7), we have

$$\begin{aligned} I_m(t) &\ge I_{m_0} e^{-(\gamma_m + \sigma_m + d_m + \mu)t} \ge 0, \\ I_h(t) &\ge I_{h_0} e^{-(\gamma_h + \sigma_h + d_h + \mu)t} \ge 0, \\ R(t) &\ge R_{init} e^{-(\mu + \rho)t} \ge 0. \end{aligned}$$

Consequently, the solution is nonnegative on $[0, T]$. The last inequality allows us to remove the restriction $T < T^*$ and, also, we may assume that $R_{init} \ge 0$, and so the system is nonnegative on every interval on which it exists.

Next, we show the boundedness of the solutions. By adding the first four equations of the model and using (3), we find

$$\frac{dN}{dt} = P - \mu N - d_m I_m - d_h I_h.$$

Hence, $\frac{dN}{dt} \le P - \mu N$. By multiplying both sides of this inequality by $e^{\mu t}$, we obtain

$$\frac{d}{dt}\left(e^{\mu t} N\right) \le P e^{\mu t},$$

hence,

$$e^{\mu t} N \le N_0 + \int_0^t P e^{\mu \tau} d\tau.$$

Thus,

$$N \le N_0 e^{-\mu t} + \frac{P}{\mu}(1 - e^{-\mu t}) \le N_0 + \frac{P}{\mu}.$$

Since N is bounded, and each of its components, S, I_m, I_h, and R are nonnegative, we have that all of these components are bounded, as well.

Next, it follows from (8) that

$$\widetilde{\beta}' \ge -\gamma^- \widetilde{\beta} R \ge -\gamma^- \widetilde{\beta}(N_0 + P/\mu),$$

therefore,

$$\widetilde{\beta(t)} \ge \beta_1 \exp(-\gamma^-(N_0 + P/\mu)t) > 0.$$

Also,

$$\widetilde{\beta}' \le \gamma_m^+ I_m + \gamma_h^+ I_h \le \max(\gamma_m^+, \gamma_h^+)(N_0 + P/\mu),$$

hence,

$$\widetilde{\beta(t)} \le \beta_1 + \max(\gamma_m^+, \gamma_h^+)(N_0 + P/\mu)t.$$

Similarly, it follows from (9) that

$$\widetilde{\alpha^*(t)} \ge \alpha_0^* \exp(-\delta^-(N_0 + P/\mu)t) > 0,$$

and

$$\widetilde{\alpha^*(t)} \le \alpha_0^* + \max(\delta_m^+, \delta_h^+)(N_0 + P/\mu)t.$$

These complete the proofs, since $0 \le t \le T$. ■

To establish the existence of a solution, we use the estimates above and the fact that since all the parameters are bounded, the right-hand-sides of the system are Lipschitz, and therefore the existence of the unique solution follows from standard results, see, e.g., [6]. Let $\boldsymbol{x} = \boldsymbol{x}(t) = (S, I_m, I_h, R, \widetilde{\beta}, \widetilde{\alpha^*})^T$, then the model can be written in the condensed form

$$\frac{d\boldsymbol{x}}{dt} = \mathbf{F}(\boldsymbol{x}, t), \tag{12}$$

where

$$\mathbf{F} = \begin{bmatrix} P - (\beta I_m + \epsilon_h \beta I_h)\frac{S}{N} - \mu S + \rho R \\ (1-\alpha^*)(\beta I_m + \epsilon_h \beta I_h)\frac{S}{N} + \gamma_h I_h - \Lambda_m I_m \\ \alpha^*(\beta I_m + \epsilon_h \beta I_h)\frac{S}{N} + \gamma_m I_m - \Lambda_h I_h \\ \sigma_m I_m + \sigma_h I_h - (\mu + \rho)R \\ \gamma_m^+ I_m + \gamma_h^+ I_h - \gamma^- \widetilde{\beta} R \\ \delta_m^+ I_m + \delta_h^+ I_h - \delta^- \widetilde{\alpha^*} R \end{bmatrix}. \tag{13}$$

Here, $\beta = \Phi_\beta(\widetilde{\beta})$, $\alpha^* = \Phi_*(\widetilde{\alpha^*})$, and

$$\Lambda_m = \gamma_m + \sigma_m + \mu + d_m, \quad \Lambda_h = \gamma_h + \sigma_h + \mu + d_h.$$

It follows from the estimates in Theorem 1 that, for $0 < T < \infty$,

$$\|\boldsymbol{x}(t)\| \le L(N_0, P, T),$$

where L is a constant that depends only on N_0, P, T and the system parameters.

Theorem 2. *(Existence and uniqueness) Let $S_0 > 0$ and I_{m0}, I_{h0}, R_{init}, be nonnegative, $\beta_1 \in [\beta_-, \beta_+]$, $\alpha_0^* \in [\alpha_-, \alpha_+]$, and let all the coefficients be positive constants. Then, there exists a unique solution to Model 1 on every time interval $[0, T]$, for $T < \infty$.*

Proof: It is straightforward, but quite tedious, to show that since all the coefficients and the functions are bounded, the system is locally Lipschitz continuous. Then, it follows from the estimates above that it is Lipschitz on every finite time interval. Therefore, the result follows from standard results for dynamical systems, see, e.g., [6]. ■

IV. Stability of the DFE state

The disease-free equilibrium (DFE) is the state $S = S_0 = P/\mu$ and $I_m = I_h = R = 0$, obtained by setting the right-hand sides of the equations (4)–(7) equal to zero. We also have $\beta = \beta_1, \alpha^* = \alpha_0^*$, however, these are automatically satisfied at the $DFE = (P/\mu, 0, 0, 0)$.

We turn to finding the *basic reproduction number*, $\mathcal{R}_0$, which controls the stability of the disease-free equilibrium. We note that in what follows the total population is kept constant, $N = S_0$.

Let $\boldsymbol{x} = (S, I_m, I_h, R, \widetilde{\beta}, \widetilde{\alpha^*})^T$, then the model can be written in a condensed form as the system (12) with the right-hand side given in (13).

To compute the Jacobian J of the system with respect to the populations, we evaluate the partial derivatives $\partial F_i/\partial x_j$ for $i, j = 1, \ldots, 4$ of $\boldsymbol{F}$ with respect to (S, I_m, I_h, R). The matrix J is given in the appendix. Here, we evaluate it at the disease free steady state $DFE = (S_0, 0, 0, 0)$. Thus,

$$J(DFE) = \begin{bmatrix} -\mu & -\beta_1 & -\epsilon_h\beta_1 & \rho \\ 0 & J_{22} & J_{23} & 0 \\ 0 & \alpha_0^*\beta_1 + \gamma_m & J_{33} & 0 \\ 0 & \sigma_m & \sigma_h & -(\mu+\rho) \end{bmatrix}, \tag{14}$$

where

$$\begin{aligned} J_{22} &= (1-\alpha_0^*)\beta_1 - \Lambda_m, \\ J_{23} &= (1-\alpha_0^*)\epsilon_h\beta_1 + \gamma_h, \\ J_{33} &= \alpha_0^*\epsilon_h\beta_1 - \Lambda_h. \end{aligned}$$

The stability of the DFE is determined by the eigenvalues of $J(DFE)$. It is straightforward to see that two of the eigenvalues are

$$\lambda_1 = -\mu, \qquad \lambda_2 = -\mu - \rho.$$

Therefore, the DFE is stable and attracting if the other two eigenvalues have negative real parts. The other two eigenvalues are rather complicated, given by,

$$\lambda_{3,4} = \frac{1}{2}\Big((J_{22} + J_{33}) \pm \big(4(\alpha_0^*\beta_1 + \gamma_m)J_{23} + (J_{22} - J_{33})^2\big)^{1/2}\Big).$$

In the simulations we found the eigenvalues numerically. In the baseline case, we found that

$$\begin{aligned} \lambda_1 &= -0.000034, & \lambda_2 &= -0.350034, \\ \lambda_3 &= -0.140298, & \lambda_4 &= -1.299470. \end{aligned}$$

Therefore, in the baseline case the DFE is stable and attracting.

A different way to find this result is the following.

A. Basic stability number $\mathcal{R}_0$

We use the results in [4] to compute the 'basic stability number' $\mathcal{R}_0$. To that end, consider the DFE state $(N,0,0,0,0)$, and then $N = S = const.$, so that $P = \mu N$. We linearize the infected populations parts in $\mathbf{F}$, and following [4], we may write the reduced system as

$$\frac{dI_m}{dt} = (1-\alpha_0^*)(\beta_1 I_m + \epsilon_h\beta_1 I_h) + \gamma_h I_h - \Lambda_m I_m, \tag{15}$$

$$\frac{dI_h}{dt} = \alpha_0^*(\beta_1 I_m + \epsilon_h\beta_1 I_h) + \gamma_m I_m - \Lambda_h I_h, \tag{16}$$

where, we recall that,

$$\Lambda_m = \gamma_m + \sigma_m + \mu + d_m, \quad \Lambda_h = \gamma_h + \sigma_h + \mu + d_h,$$

are positive constants.

Let $\boldsymbol{x}_i = (I_m, I_h)^T$, then the linearized reduced system (15)–(16) can be written as

$$\frac{d\boldsymbol{x}_i}{dt} = \boldsymbol{G}\boldsymbol{x}_i,$$

where

$$\boldsymbol{G} = \begin{bmatrix} (1-\alpha_0^*)\beta_1 - \Lambda_m & (1-\alpha_0^*)\epsilon_h\beta_1 + \gamma_h \\ \alpha_0^*\beta_1 + \gamma_m & \alpha_0^*\epsilon_h\beta_1 - \Lambda_h \end{bmatrix}.$$

We may write $\boldsymbol{G} = \boldsymbol{T} + \boldsymbol{\Sigma}$, where $\boldsymbol{T}$ is the *transmission* matrix and $\boldsymbol{\Sigma}$ the *transition* matrix, as

$$\boldsymbol{T} = \begin{bmatrix} (1-\alpha_0^*)\beta_1 & (1-\alpha_0^*)\epsilon_h\beta_1 \\ \alpha_0^*\beta_1 & \alpha_0^*\epsilon_h\beta_1 \end{bmatrix}$$

and

$$\boldsymbol{\Sigma} = \begin{bmatrix} -\Lambda_m & \gamma_h \\ \gamma_m & -\Lambda_h \end{bmatrix}.$$

Then, the next generation matrix (NGM) $\boldsymbol{K}$ is given by

$$\boldsymbol{K} = -\boldsymbol{T}\boldsymbol{\Sigma}^{-1} = \begin{bmatrix} (1-\alpha_0^*)\beta_1 & (1-\alpha_0^*)\epsilon_h\beta_1 \\ \alpha_0^*\beta_1 & \alpha_0^*\epsilon_h\beta_1 \end{bmatrix} \times \\ \times \frac{1}{\Delta}\begin{bmatrix} \Lambda_h & \gamma_h \\ \gamma_m & \Lambda_m \end{bmatrix},$$

where $\Delta = \Lambda_m\Lambda_h - \gamma_m\gamma_h$ is the determinant of $\boldsymbol{\Sigma}$. Thus,

$$\boldsymbol{K} = \frac{1}{\Delta}\begin{bmatrix} (1-\alpha_0^*)\beta_1(\Lambda_h + \epsilon_h\gamma_m) & (1-\alpha_0^*)\beta_1(\gamma_h + \epsilon_h\Lambda_m) \\ \alpha_0^*\beta_1(\Lambda_h + \epsilon_h\gamma_m) & \alpha_0^*\epsilon_h\beta_1(\gamma_h + \Lambda_m) \end{bmatrix}.$$

It is seen that $\det\boldsymbol{K} = 0$, hence the largest eigenvalue, which is $\mathcal{R}_0$, is

$$\mathcal{R}_0 = \mathrm{tr}\boldsymbol{K} = \frac{\beta_1}{\Delta}\Big((1-\alpha_0^*)(\Lambda_h + \epsilon_h\gamma_m) + \alpha_0^*\epsilon_h(\gamma_h + \Lambda_m)\Big).$$

In the baseline case, we found that $\mathcal{R}_0 = 0.444$, which reinforces our conclusion that the DFE is stable and attracting, hence the disease will be eradicated, eventually.

Remark 2. *If one wishes to consider the case when asymptotically $P = P(t)$, then it is necessary to study the fibers of the global DFE attractor, see, e.g., [6]. This may be of interest in later stages of our study of such models.*

V. Endemic equilibrium (EE) states

The unusual structure of the model allows for various endemic equilibrium states (EEs), unlike the standard models. It relates to the various cases for the equilibrium values of β, α^*. We assume that an endemic equilibrium state exists and denote its populations by $\hat{S}, \hat{I}_m, \hat{I}_h, \hat{R}$, where either $\hat{I}_m > 0$, or $\hat{I}_h > 0$, or both. The populations in such a state may depend on the signs of the terms

$$L_\beta = \gamma_m^+\hat{I}_m + \gamma_h^+\hat{I}_h - \gamma^-\widetilde{\beta}\hat{R},$$
$$L_* = \delta_m^+\hat{I}_m + \delta_h^+\hat{I}_h - \delta^-\widetilde{\alpha^*}\hat{R}.$$

We assume that $\beta_s \in [\beta_-, \beta_+]$ and $\alpha_s^* \in [\alpha_-, \alpha_+]$ (where $0 < \alpha_- \le \alpha_+ < 1$, are given and fixed, within the allowed intervals. Then, at an EE, $\hat{\boldsymbol{z}} = (\hat{S}, \hat{I}_m, \hat{I}_h, \hat{R})$ is the solution of the system,

$$0 = P - \beta_s(\hat{I}_m + \epsilon_h\hat{I}_h)\frac{\hat{S}}{\hat{N}} - \mu\hat{S} + \rho\hat{R}, \tag{17}$$

$$0 = (1-\alpha_s^*)\beta_s(\hat{I}_m + \epsilon_h\hat{I}_h)\frac{\hat{S}}{\hat{N}} + \gamma_h\hat{I}_h - \Lambda_m\hat{I}_m, \tag{18}$$

$$0 = \alpha_s^*\beta_s(\hat{I}_m + \epsilon_h\hat{I}_h)\frac{\hat{S}}{\hat{N}} + \gamma_m\hat{I}_m - \Lambda_h\hat{I}_h, \tag{19}$$

$$0 = \sigma_m\hat{I}_m + \sigma_h\hat{I}_h - (\mu+\rho)\hat{R}. \tag{20}$$

Here,

$$\hat{N} = \hat{S} + \hat{I}_m + \hat{I}_h + \hat{R}.$$

We, next, introduce the notation

$$a_1 = \frac{(1-\alpha_s^*)\Lambda_h + \alpha_s^*\gamma_h}{\alpha_s^*\Lambda_m + (1-\alpha_s^*)\gamma_m},$$
$$a_2 = \frac{(1-\alpha_s^*)\beta_s(a_1 + \epsilon_h)}{\Lambda_m a_1 - \gamma_h},$$

$$a_3 = \frac{\rho(\sigma_m a_1 + \sigma_h)}{\mu(\mu+\rho)} - \frac{\Lambda_m a_1 - \gamma_h}{\mu(1-\alpha_s^*)},$$
$$a_4 = \frac{\sigma_m a_1 + \sigma_h}{\mu+\rho}.$$

We conclude, based on the results in the Appendix, that the solution of the endemic steady state, for the chosen β_s, α_s^* is as follows.

Proposition 1. *An endemic steady state* $\hat{\boldsymbol{z}} = (\hat{S}, \hat{I}_m, \hat{I}_h, \hat{R})$*, corresponding to* $\beta = \beta_s$*,* $\alpha^* = \alpha_s^*$*, is given by*

$$\hat{S} = \frac{P}{\mu} + a_3\hat{I}_h = \frac{P}{\mu} + \frac{a_3 P(a_2-1)}{\mu(1+a_1-a_2a_3+a_3+a_4)},$$
$$\hat{I}_m = a_1\hat{I}_h = \frac{a_1 P(a_2-1)}{\mu(1+a_1-a_2a_3+a_3+a_4)},$$
$$\hat{I}_h = \frac{P(a_2-1)}{\mu(1+a_1-a_2a_3+a_3+a_4)},$$
$$\hat{R} = \frac{\sigma_m a_1+\sigma_h}{\mu+\rho}\hat{I}_h = \frac{(\sigma_m a_1+\sigma_h)P(a_2-1)}{(\mu+\rho)\mu(1+a_1-a_2a_3+a_3+a_4)}.$$

Moreover,

$$\hat{N} = \frac{P}{\mu} + (1+a_1+a_3+a_4)\,\hat{I}_h.$$

The coefficients a_1, a_2, a_3, a_4 *are given above.*

Therefore, for each choice of $\beta_s \in [\beta_-, \beta_+]$ and $\alpha_s^* \in [\alpha_-, \alpha_+]$ there corresponds one steady state. Hence, the model has a *continuum of steady states*. This raises the question of the *reachability* of these steady states, that is, given a steady state, can we find initial conditions, different from the steady state, which guarantee that the time dependent solution converges to this state in the limit $t \to \infty$. The question is of some mathematical interest and my be studied further in the future.

It seems, and some of the simulations support the observation that when $L_\beta > 0$ then $\beta \to \beta_+$ as $t \to \infty$ and when $L_\beta < 0$ then $\beta \to \beta_-$ as $t \to \infty$. However, this is not the case for L_* and α^*. Mathematically, there are a continua of solutions, and this question may be of interest to address, too.

VI. Simulations

To generate the computer approximate solutions of the model, a Python code was written in MATLAB, based on simple explicit time stepping. It performed very well, and the runs typically took a few seconds on a laptop. Some of the simulations are presented in what follows.

The values of the various parameters for each simulation are given in Table 2, where the DFE column provides parameter values for the simulation depicting the system with a disease free equilibrium and the EE columns are the parameter values for the simulations depicting systems with endemic equilibiria.

First, we describe the baseline simulations, the results of which are depicted in Figure 2, where it is seen that the system converges to the disease-free steady state. The simulation is up to $T = 30$ and the dynamics of the four populations are depicted, as well as those of β and α^*. The Susceptibles (yellow) are in the top figure, while the dynamics of the mildly infected (green), the hospitalized (maroon), and the recovered (blue), are below it. It is seen that the populations converge, as $t \to \infty$, to the values $\hat{\boldsymbol{z}} = (100030, 0, 0, 0)$. The total population, eventually, consists of just the Susceptibles, since we allow Recovered to loose their immunity and become susceptible after a while.

The dynamics of β (purple) are depicted in the third figure, and it is seen that $\beta \to 0.36$ as $t \to \infty$, so it converges to the upper limit of its allowed interval. However, α^* converges to 0.0141 which is above its lower limit. This is of theoretical interest and may be worth further study. Here, $L_\beta = 0$ and $L_* = 0$.

Next, we consider two simulations when the system has an attracting endemic steady state (EE). The values of the parameters for both simulations are shown in Table 2.

The first simulation, Figure 3, shows the convergence of the solutions to the EE. Under these conditions, as $t \to \infty$, $\hat{\boldsymbol{z}} = (7763, 727, 678, 377)$, with $a_1 = 1.0729$, $a_2 = 1.2296$, $a_3 = -136.053$, $a_4 = 0.5558$. The plots in Figure 3 depict the behavior over $0 \le t \le T = 200$, which visually shows the endemic behavior. When we let $T = 26,000$ (not shown), we found that each of the populations reached our calculated values for the EE under these conditions, as provided in Section 5. Figure 3 has the same form as Figure 2. The Susceptibles (yellow), in the top of the figure, show relatively little changes. The mildly infected (green), the hospitalized (maroon), and the recovered (blue) all show evolution initially, before reaching equilibrium. The plot of β shows that it climbs to its maximum value monotonically, whereas α^* has more variability. Though it reaches its maximum very early, it oscillates irregularly before reaching its maximum value at equilibrium. As $t \to \infty$, $L_\beta = 0.000415$ and $L_* = 0.0891$.

The second EE simulation explores the case when at a later time ($T = 20$ in the simulation) a new variant of the pathogen appears with a higher infectivity rate in the equation for β. The initial state is the same as in the first EE case, with the addition of a change in γ_m^+

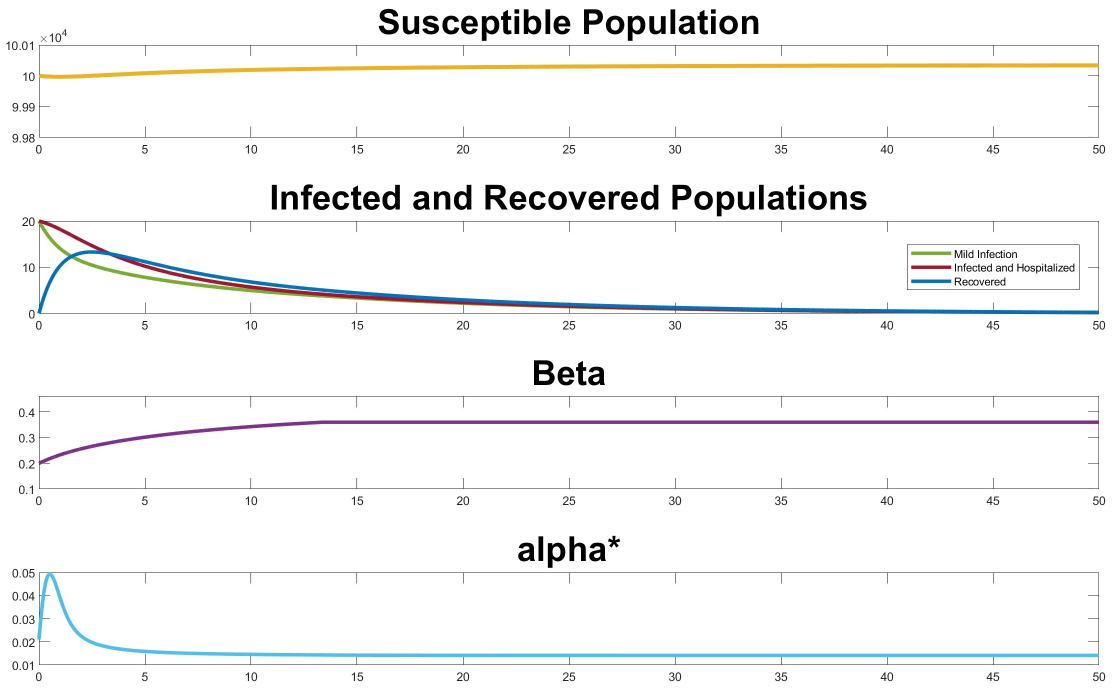

Fig. 2: Convergence to the Disease-Free Equilibrium

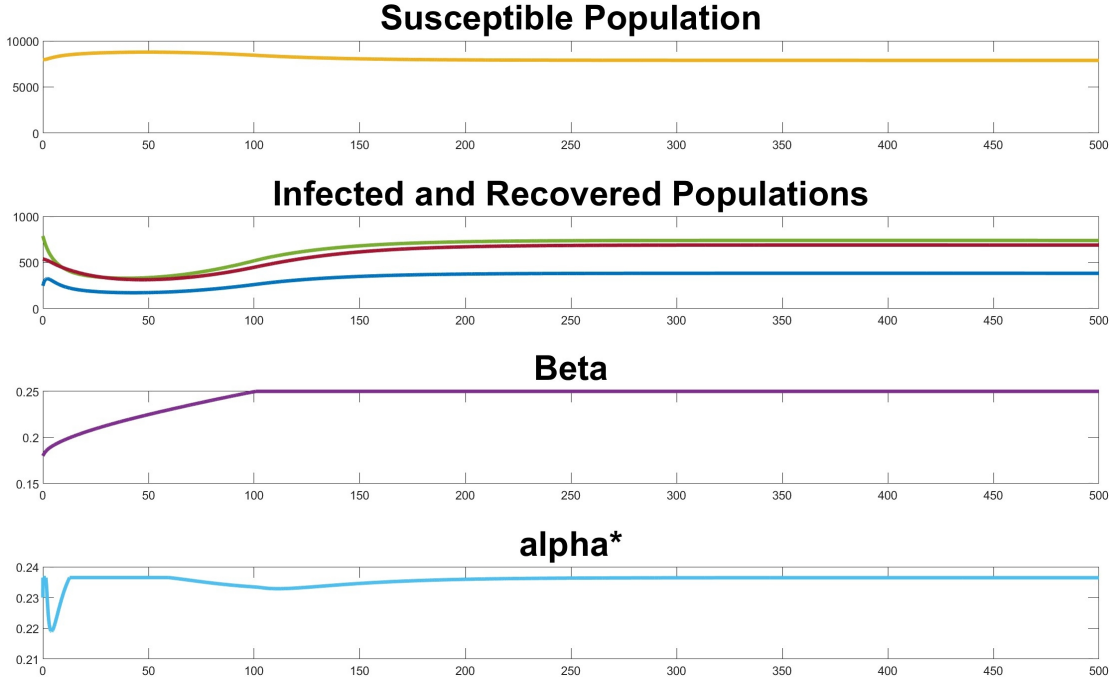

Fig. 3: Convergence to the Endemic Equilibrium (first case)

Parameter	DFE value	EE value
P	3.4	3.4
β_1	0.2	0.18
α_0^*	0.021	0.23
ϵ_h	1.000	1.000
ρ	0.75	0.75
μ	0.000034	0.000034
d_m	0.0012	0.0005
d_h	0.0294	0.004
γ_m	0.31	0.00501
γ_h	0.1501	0.006
σ_m	0.67	0.3
σ_h	0.201	0.095
$\gamma_m^+, \gamma_h^+, \gamma^-$	$10^{-3}, 10^{-3}, 2\times10^{-3}$	$10^{-5}, 10^{-6}, 8\times10^{-5}$
$\delta_m^+, \delta_h^+, \delta^-$	$5\times10^{-4}, 5\times10^{-3}, 0.313$	$4.5\times10^{-4}, 7\times10^{-4}, 9\times10^{-3}$
$\beta_-, \beta_+, \alpha_-, \alpha_+$	0.1, 0.36, 0.01, 0.05	0.02, 0.25, 0.1, 0.2365
$S_0, I_{m0}, I_{h0}, R_{init},$	100,000, 20, 20, 0	7960, 787, 538, 247

Table 2: Symbols and values of the parameters in the baseline simulation

and γ_h^+ at $T = 20$. The values of the parameters for the second endemic simulation are the same as those in the first endemic simulation, except for the change that at $T = 20$, γ_m^+ is manually changed to 10^{-4} and γ_h^+ is changed to 10^{-5} to see how the system responds. Under these conditions, as $t \to \infty$, again we have, $\hat{\boldsymbol{z}} = (7763, 727, 678, 377)$, with $a_1 = 1.0729$, $a_2 = 1.2296$, $a_3 = -136.053$, $a_4 = 0.5558$. We conclude that in this case the EE is attracting, and attracts the modified solutions, too.

It seems to be important to add randomness into the equations for β and α^* to include possible mutations of the pathogen and the appearance of new variants, as the case of COVID-19 demonstrates. However, in light of the above example, this has to done be judiciously.

VII. Conclusions and future work

This work develops and studies mathematically and computationally a new extended version of the SIR model for the spread of a communicable disease in which the infected population is divided into light or moderately sick individuals and those with severe form that are hospitalized. The novelty lies in the assumption that the infectivity rate β and the hospitalization fraction α^*, both of which measure the severity of the disease, are system dependent variables. For the sake of simplicity, the rate equations for both variables are linear in the populations, however, it is very likely that the relationships are more complex and interesting, and will need to be determined for each disease separately, by some parameter identification method. We return to this point below. For the model to make epidemiological sense both variables are limited to values in prescribed intervals, in particular, α^* is a fraction so it is restricted to the interval $0 < \alpha_- \le \alpha^* \le \alpha_+ \le 1$.

We establish the positivity and boundedness of the solutions, which lead to the proof of their existence. Then, we study the system steady states and show that in the baseline case the DFE is stable and attracting (asymptotically stable). We present two different ways to show this, the second is based on the work [4]. Then, we study cases with EE states which, because of the structure of the system, may form a continuum. These findings are supported by the computer simulations of the model. Moreover, this observation raises the question of reachability of the steady states, mentioned below.

A simple explicit time stepping algorithm was constructed and implemented, and the resulting simulations provide insights into the model predictions. The first simulation shows how the system, with stable and attracting DFE, approaches this equilibrium. In that state only Susceptibles remain, the other populations vanish, β approaches its upper limit and α^* its lower limit.

The second simulation involves the case of an EE state that is stable and attracting. Here, the populations approach nonzero limits, as can be seen in Figure 4. While β converges to its upper allowed limit, the fraction of hospitalized α^* approaches a limit within its allowed interval.

The third simulation has the same parameters as the second, with an EE state. However, to show the effect of the appearance of a more infective variant, at time $T = 5$ we changed the rate constants in the equation for β. it is seen that the change leads to a jump in β. However,

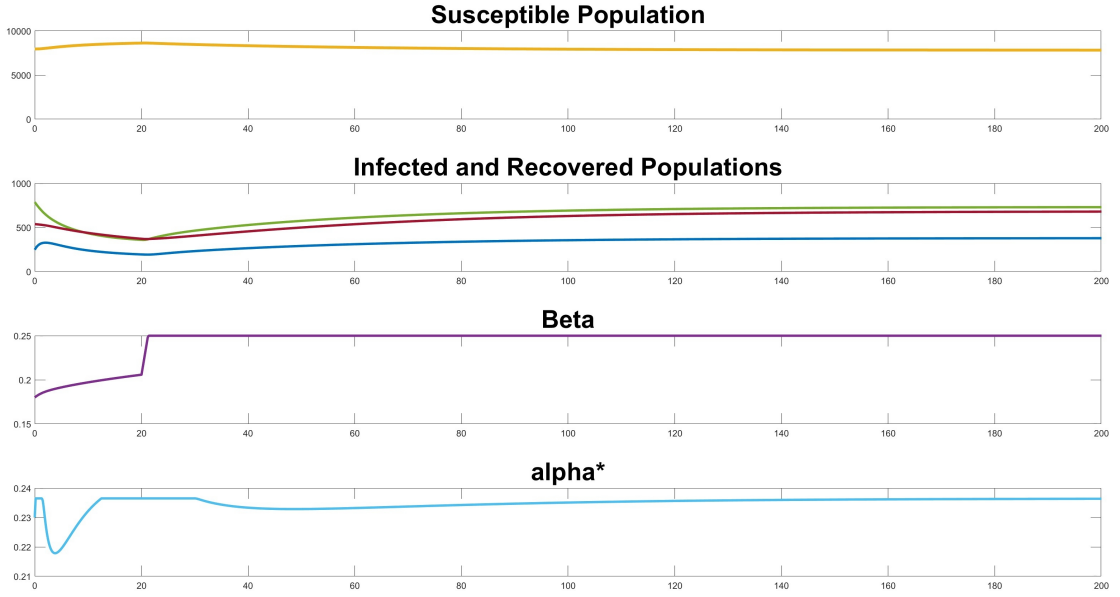

Fig. 4: Second Endemic Equilibrium simulation. The values of the γ coefficients in (8) were changed at $T = 20$.

in the long run the solution in this case converges to the same EE limit as in the second simulation.

Some related issues and topics that warrant further study are the following.

One of the main open questions is the form of the equations for β and α^*. Assuming the form (1) with the restrictions (2), how do we find the functional forms of Ψ_β and Ψ_*? It seems to us that some statistical input is needed, as well as curve fitting to field data. The latter does not exist yet in a form that may be used in this model, so it needs to be developed.

The result on the existence of a continuum of steady state solutions raises the question of the *reachability* of these steady states, that is, given a steady state, can we find initial conditions which guarantee that the time dependent solution converges to this state in the limit $t \to \infty$. The topic is of interest to understand better the model.

Furthermore, it is of interest to take into account the sudden appearance of new variants of the agents of the disease. This may be accomplished by adding randomness to some of the system coefficients. We have begun such a study, and it seems to be of importance to continue it. It is found that just adding random impulses to the two equations for β and α^* does not address the issue since the simulations show that such addition dies out over time.

The addition of periodicity into some of the coefficients to take into account climate variations may be of interest.

More simulations are of interest, together with comparison to known information about some current diseases.

ACKNOWLEDGEMENT

Both authors are very grateful to the anonymous reviewers for their very thorough reviews which considerably improved the presentation.

REFERENCES

[1] A. Cesmelioglu, K. L. Kuttler, M. Shillor, A. M. Spagnuolo, *A mathematical model of the COVID-19 pandemic dynamics with dependent variable infection rate: Application to the Republic of Korea*, 13 May 2021, arXiv:2008.03248.

[2] F. Brauer, Mathematical epidemiology: Past, present, and future, *Infectious Disease Modelling*, 2:113–127, 2017.

[3] H. W. Hethcote, The Mathematics of Infectious Diseases, *SIAM Review*, 42:599–653, 2000.

[4] O. Diekmann, J. A. P. Heesterbeek, M. G. Roberts, The construction of next-generation matrices for compartmental epidemic models, *Journal of the Royal Society Interface*, 7:873–885, 2010.

[5] Wikipedia, [26-Nov-2024], https://en.wikipedia.org/wiki/Compartmental_models_in_epidemiology.

[6] P. E. Kloeden, M. Rasmussen, *Nonautonomous Dynamical Systems*, Mathematical Surveys and Monographs, 176, American Mathematical Society, 2011.

VIII. APPENDIX

The Jacobian at the DFE. We compute the Jacobian matrix of the system. First, we note that

$$\frac{\partial(1/N)}{\partial S} = \frac{\partial(1/N)}{\partial I_m} = \frac{\partial(1/N)}{\partial I_h} = \frac{\partial(1/N)}{\partial R} = -\frac{1}{N^2}.$$

Then, the entries of J are:

$$J_{11} = \partial\mathcal{F}_1/\partial S = -(\beta I_m + \epsilon_h\beta I_h)\left(\frac{N-S}{N^2}\right) - \mu,$$
$$J_{12} = \partial\mathcal{F}_1/\partial I_m = -\beta\frac{S}{N} + (\beta I_m + \epsilon_h\beta I_h)\frac{S}{N^2},$$
$$J_{13} = \partial\mathcal{F}_1/\partial I_h = -\epsilon_h\beta\frac{S}{N} + (\beta I_m + \epsilon_h\beta I_h)\frac{S}{N^2},$$
$$J_{14} = \partial\mathcal{F}_1/\partial R = \rho,$$

$$J_{21} = \partial\mathcal{F}_2/\partial S = (1-\alpha^*)(\beta I_m + \epsilon_h\beta I_h)\frac{N-S}{N^2},$$
$$J_{22} = \partial\mathcal{F}_2/\partial I_m = (1-\alpha^*)\beta\frac{S}{N} - (1-\alpha^*)(\beta I_m + \epsilon_h\beta I_h)\frac{S}{N^2} - (\gamma_m + \sigma_m + \mu + d_m),$$
$$J_{23} = \partial\mathcal{F}_2/\partial I_h = (1-\alpha^*)\epsilon_h\beta\frac{S}{N} - (1-\alpha^*)(\beta I_m + \epsilon_h\beta I_h)\frac{S}{N^2} + \gamma_h,$$
$$J_{24} = \partial\mathcal{F}_2/\partial R = 0,$$

$$J_{31} = \partial\mathcal{F}_3/\partial S = \alpha^*(\beta I_m + \epsilon_h\beta I_h)\frac{N-S}{N^2},$$
$$J_{32} = \partial\mathcal{F}_3/\partial I_m = \alpha^*\beta\frac{S}{N} - \alpha^*(\beta I_m + \epsilon_h\beta I_h)\frac{S}{N^2} + \gamma_m,$$
$$J_{33} = \partial\mathcal{F}_3/\partial I_h = \alpha^*\epsilon_h\beta\frac{S}{N} - \alpha^*(\beta I_m + \epsilon_h\beta I_h)\frac{S}{N^2} - (\gamma_h + \sigma_h + \mu + d_h),$$
$$J_{34} = \partial\mathcal{F}_3/\partial R = 0,$$

$$J_{41} = \partial\mathcal{F}_4/\partial S = \rho,$$
$$J_{42} = \partial\mathcal{F}_4/\partial I_m = \sigma_m,$$
$$J_{43} = \partial\mathcal{F}_4/\partial I_h = \sigma_h,$$
$$J_{44} = \partial\mathcal{F}_4/\partial R = -\mu.$$

In (14) the entries are evaluated at the DFE.

Proof of Proposition 1: To find solutions for the EE, we assume that $\hat{I}_m, \hat{I}_h \neq 0$, since otherwise it corresponds to the DFE. Solving for $\hat{S}/\hat{N}$ in equations (18) and (19) and setting them equal yields

$$\frac{\Lambda_m\hat{I}_m - \gamma_h\hat{I}_h}{(1-\alpha_s^*)\beta_s(\hat{I}_m + \epsilon_h\hat{I}_h)} = \frac{\Lambda_h\hat{I}_h - \gamma_m\hat{I}_m}{\alpha_s^*\beta_s(\hat{I}_m + \epsilon_h\hat{I}_h)}, \quad (21)$$

and consequently,

$$\hat{I}_m = a_1\hat{I}_h,$$

where

$$a_1 = \frac{(1-\alpha_s^*)\Lambda_h + \alpha_s^*\gamma_h}{\alpha_s^*\Lambda_m + (1-\alpha_s^*)\gamma_m}.$$

Solving for R in equation (20) and substituting the above equation for I_m gives

$$\hat{R} = \frac{\sigma_m a_1 + \sigma_h}{\mu + \rho}\hat{I}_h.$$

We now have equations for $\hat{I}_m$ and $\hat{R}$ in terms of $\hat{I}_h$, which we can substitute into equation (17), and we can also substitute the left-hand side of equation (21) in for $\hat{S}/\hat{N}$, thus,

$$\hat{S} = \frac{P}{\mu} + \left[\frac{\rho(\sigma_m a_1 + \sigma_h)}{\mu(\mu+\rho)} - \frac{\Lambda_m a_1 - \gamma_h}{\mu(1-\alpha_s^*)}\right]\hat{I}_h.$$

Since the left-hand side of equation (21) equals $\hat{S}/\hat{N}$, we can solve for $\hat{N}$,

$$\hat{N} = \frac{(1-\alpha_s^*)\beta_s(a_1+\epsilon_h)}{\Lambda_m a_1 - \gamma_h}\left[\frac{P}{\mu} + \left[\frac{\rho(\sigma_m a_1 + \sigma_h)}{\mu(\mu+\rho)} - \frac{\Lambda_m a_1 - \gamma_h}{\mu(1-\alpha_s^*)}\right]\hat{I}_h\right],$$

and, since $\hat{N}$ is also the total population, we have

$$\hat{N} = \frac{P}{\mu} + \left[\frac{\rho(\sigma_m a_1 + \sigma_h)}{\mu(\mu+\rho)} - \frac{\Lambda_m a_1 - \gamma_h}{\mu(1-\alpha_s^*)} + 1 + a_1 + \frac{\sigma_m a_1 + \sigma_h}{\mu+\rho}\right]\hat{I}_h = \frac{P}{\mu}\left(1 + a_3 + a_1 + a_4\right)\hat{I}_h.$$

Here,

$$a_2 = \frac{(1-\alpha_s^*)\beta_s(a_1+\epsilon_h)}{\Lambda_m a_1 - \gamma_h},$$
$$a_3 = \frac{\rho(\sigma_m a_1 + \sigma_h)}{\mu(\mu+\rho)} - \frac{\Lambda_m a_1 - \gamma_h}{\mu(1-\alpha_s^*)},$$
$$a_4 = \frac{\sigma_m a_1 + \sigma_h}{\mu+\rho}.$$

Setting the two equations for $\hat{N}$ equal gives an equation entirely in terms of $\hat{I}_h$, thus

$$a_2\left(\frac{P}{\mu} + a_3\hat{I}_h\right) = \frac{P}{\mu} + (a_3 + 1 + a_1 + a_4)\hat{I}_h.$$

And solving for $\hat{I}_h$, we obtain,

$$\hat{I}_h = \frac{P(a_2 - 1)}{\mu(a_3 + 1 + a_1 + a_4 - a_2 a_3)}.$$

■