



OPEN

Adaptive SIR model with vaccination: simultaneous identification of rates and functions illustrated with COVID-19

Tchavdar T. Marinov^{1✉} & Rossitza S. Marinova^{2,3}

An Adaptive Susceptible-Infected-Removed-Vaccinated (A-SIRV) epidemic model with time-dependent transmission and removal rates is constructed for investigating the dynamics of an epidemic disease such as the COVID-19 pandemic. Real data of COVID-19 spread is used for the simultaneous identification of the unknown time-dependent rates and functions participating in the A-SIRV system. The inverse problem is formulated and solved numerically using the Method of Variational Imbedding, which reduces the inverse problem to a problem for minimizing a properly constructed functional for obtaining the sought values. To illustrate and validate the proposed solution approach, the present study used available public data for several countries with diverse population and vaccination dynamics—the World, Israel, The United States of America, and Japan.

Infectious diseases modelling attracted great deal of attention by scientists and people around the world during the COVID-19 pandemic. The novel coronavirus (SARS-CoV-2) started to quickly spread in the early months of 2020 and was announced as a global pandemic by the World Health Organization¹ in March 2020. Population-wide vaccination is critical for achieving herd immunity and for controlling the COVID-19 pandemic while combined with effective testing and preventive measures.

Inevitably, the development of vaccines became the highest priority of governments and pharmaceutical companies². Several vaccines were available in the last months of 2020. As of 30 July 2021, over 28% of the world population is partly or fully vaccinated³. Not knowing much about the coronavirus disease in the early months, mathematical models have played an important role in shedding some light on the disease dynamics.

The SIR model categorizes individuals as Susceptible, Infectious, and Recovered. Mathematical infectious disease models based on the classical SIR model⁴ are widely used to examine the spread of a disease. These models display compelling results especially during the early period of the pandemic^{5–19}. In a recent article²⁰, the authors studied the immune response of recovered from COVID-19 individuals up to 8 months patients and found that they have considerable immune memory.

Vaccination is a common method of reducing infectious diseases spread^{21–25} because it reduces the number of susceptible, from where the reproduction number naturally also decreases. The reproductive number is an indicator of the transmissibility of the virus caused by infectious individuals. It is affected by the population density, vaccinations, quarantines, social distancing, mask wearing and other measures^{25,26}.

In a recent work²⁷, the authors present a study of the temporal evolution of epidemic outbreaks accounting for vaccinations with monitored real time COVID-19 data, using the SIRV model, V denoting the relative fractions of currently vaccinated. They make certain assumptions and reduce the time-dependent general SIRV equations to an analytical model. SIR and SEIR models with vaccination are used to simulate and predict the development of the COVID-19 spread, e.g.^{25,28–30}.

In several very recent publications^{31–35} applied to the COVID-19 epidemic, researchers have developed and used SIR and SEIR based models with vaccination to overcome the limitations of the conventional SIR model. The work in³¹ presents an investigation of the dynamics of a stochastic SIRV epidemic model with general non-linear incidence and vaccination. The introduced random fluctuations controls the disease outbreak. Zhao et al.³²

¹Department of Natural Sciences, Southern University at New Orleans, 6801 Press Drive, New Orleans, LA 70126, USA. ²Department of Mathematical and Physical Sciences, Concordia University of Edmonton, 7128 Ada Boulevard, Edmonton, AB T5B 4E4, Canada. ³Department of Computer Science, Varna Free University, Varna, Bulgaria. ✉email: tmarinov@suno.edu

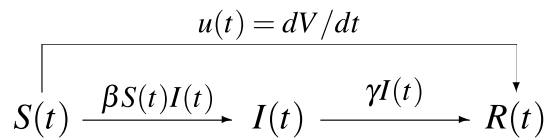


Figure 1. The A-SIRV epidemic model.

use improved SIRV to evaluate the performance of non-pharmaceutical interventions in reducing the number of daily new cases of COVID-19 in South and Southeast Asia. They apply statistical methods to estimate parameters.

The research presented in³³ proposes an SIRV evolutionary game model for infectious disease vaccination strategies based on the scale-free networks with tunable clustering. Their model analyzes the vaccination strategies taking into account factors such as vaccination effectiveness, vaccination cost, treatment cost after illness, government subsidy rate and treatment discount rate. Other notable work is about modelling infectious diseases with herd immunity in a randomly mixed population³⁴. The authors formulate two new SIR models to mimic the declining transmission rate of infectious diseases at different stages of transmission. They found that natural herd immunity might not be sufficiently effective in infectious diseases with high reproduction numbers.

Researchers implemented a modification of a SIR model³⁵ to study the role of the rate of vaccination, rate of transmission and the likelihood of emergence of resistant strains. They used parameters realistically resembling SARS-CoV-2 transmission to run simulations for a total time of three years, with vaccination starting one year into the model.

The goal of the present work is to develop and to demonstrate the effectiveness of an inverse method for identifying the time-dependent functions and parameters of the SIRV model simultaneously (Sect. 2). We apply the method to the adaptive SIRV (A-SIRV) epidemic model using publicly available COVID-19 data. In contrast to other works which use statistical approaches to estimate parameters, we apply an inverse problem approach to identify these parameters. The dynamically estimated rates can be particularly useful in running other simulations, such as in³⁵, to study the epidemic.

The modified SIR model for the spread of an infectious disease and vaccination (SIRV)

The standard notations in the SIR model are: $S(t)$ denotes the number of *susceptible*, $I(t)$ – *infectives*, and $R(t)$ – *removed* individuals. Assume that the time-dependent function $u = u(t)$ represents the *vaccination rate*. Then the total number of vaccinated individuals is given by

$$V(t) = \int_0^t u(t)dt. \quad (1)$$

Every vaccine has different level of efficacy. We assume that the vaccine efficacy impacts disease spread and prevents transmission at the same rate, which is a reasonable assumption according to the study³⁵. If people are vaccinated with the same vaccine type / brand, it is possible to introduce the efficacy of the vaccine in the model. Since there is lack of information about the vaccine types and other details concerning the vaccinated individuals, we make the following assumptions, which may not be true for all

- Individuals being vaccinated belong to the class S (susceptible) before the vaccine. There are recovered people who vaccinate themselves; this is not included in the model because of lack of data.
- Individuals move to the class R (removed) after vaccination. In other words, the model assumes the vaccine is 100% effective against the disease, namely vaccinated people become fully immune.

Due to unavailability of data about COVID-19 variants and vaccine details on a country level, at this stage, the model does not include important assumptions, such as:

- Vaccinated individuals could be infected and be infectious if the infection is caused by other variants, known as the vaccine breakthrough problem; there might be multiple variants circulated in the same country, not only a single strain.
- Infected individuals could be vaccinated again to improve their immune level; vaccinated individuals could be infected again due to waning of immunity for COVID-19.

The *transmission rate* $\beta > 0$ gives the probability that a random infective person infects a random susceptible person. A major approximation here is the assumption that the population under study is well mixed so that every person has equal probability of coming into contact with every other person. The *removal rate* $\gamma > 0$ gives the probability that an infective person recovers. In the classical SIR model β and γ are constants, while they are functions of time in¹⁴. The diagram in Fig. 1 describes the adaptive SIR (A-SIR) model with vaccinations (A-SIRV) and time-dependent coefficients:



Figure 2. The time nodes v_i where the data values are given.

$$L_1(S, I, \beta) = \frac{dS(t)}{dt} + \beta(t)S(t)I(t) + u(t) = 0 \tag{2}$$

$$L_2(S, I, \beta, \gamma) = \frac{dI(t)}{dt} - \beta(t)S(t)I(t) + \gamma(t)I(t) = 0 \tag{3}$$

$$L_3(I, R, \gamma) = \frac{dR(t)}{dt} - \gamma(t)I(t) - u(t) = 0. \tag{4}$$

The total population $N = S(t) + I(t) + R(t)$ is considered constant in the equations. Fig. 1 shows the diagram for the A-SIRV model, corresponding to the system (2)–(4).

The A-SIRV model used in this study assumes that the removed individuals are no longer susceptible nor infectious. The number of cases for recovered from COVID-19 individuals who are re-infected at the present moment is very limited and the rate cannot be estimated; thus, this possibility is not taken into account.

Effective and basic reproductive ratios R_e and R_0 . An epidemic occurs if an infective individual introduced into a population of susceptible individuals infects on average more than one other person, namely $I(t)$ is increasing in time. The original SIR model assumes that the transmission and removal rates are constants. Equations (2)–(4), with proper initial conditions, allow the determination of $I(t)$, and $S(t)$, and $R(t)$, if the coefficients β and γ are known constants. However, in the case of a pandemic, the rates may vary in time; hence, $\beta = \beta(t)$ and $\gamma = \gamma(t)$.

The so-called *effective reproduction number* (also *effective reproduction rate or ratio*) for a given epidemic is the parameter $R_e(t)$ given by

$$R_e(t) = \frac{\beta(t)S(t)}{\gamma(t)}. \tag{5}$$

An epidemic occurs when $R_e(t) > 1$. Then the fraction of the population that is immune increases (because of vaccination or because of recovering from the disease) so much that $R_e(t) < 1$, “herd immunity” is achieved. Hence, the number of new cases occurring in the population will decrease to zero.

Another important characteristic of an epidemic, the *basic reproduction number (or ratio, or rate)* R_0 , is defined as:

$$R_0 = \frac{\beta(t)N}{\gamma(t)}, \tag{6}$$

where N is the size of the total population.

Multiple factors may cause the rates to change over time. In the case of COVID-19, examples include social distancing, restrictions imposed by governments, and preventive treatments. Therefore, we define the effective reproduction number and the basic reproduction number to be functions of time, defined in (5) and (6), respectively.

The main goal of the present work is to identify the time-dependent reproduction rates directly using the SIRV model.

Inverse problem formulation. The initial-value problem consisting of the system of equations (2)–(4), with coefficients $\beta(t)$ and $\gamma(t)$ known, along with proper initial conditions derived from the given data, constitutes the direct problem. Note that the vaccination rate function $u(t)$ in equations (2) and (4) can be obtained from the function $V(t)$, which is known from the reported public data on a daily basis.

In reality, the values of the time-dependent parameters $\beta(t)$ and $\gamma(t)$ are unknown for a new epidemic disease. Hence, the simultaneous determination of the coefficients and functions from the available data is an *inverse problem*.

Let $[0, P]$, where P is a number of days, be a time sub-interval (Fig. 2), where approximate values of the susceptible $S(t)$ and currently infectious $I(t)$ are known at specified time moments v_0, v_1, \dots, v_P ,

$$S(v_i) \approx \sigma_i, \quad I(v_i) \approx \lambda_i, \quad i = 0, 1, \dots, P. \tag{7}$$

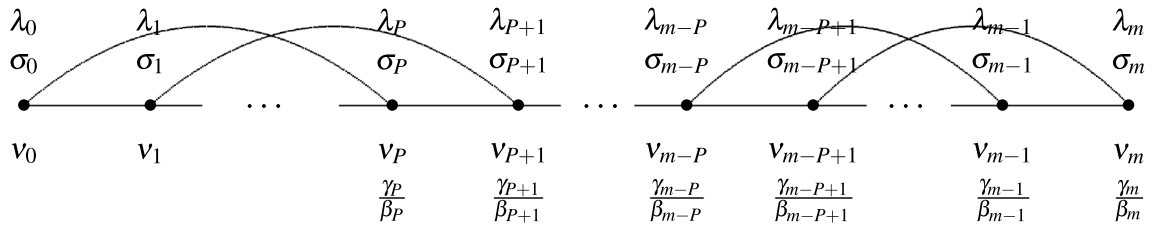


Figure 3. The time nodes and the subsets of fixed length of $P + 1$ days for identifying $\beta(t)$ and $\gamma(t)$.

At the same time, the values of $V(v_i)$ are given exactly since the number of vaccinated people is available, as reported in³. Therefore, we can approximate the data values of $V(t)$ using a function with a continuous second derivative. From this function, the values $u(t) = \frac{dV}{dt}$ can be estimated.

Since the values of the vaccination rate function $u(t)$ can be found from the available data of vaccinated individuals, we can find $I(t)$, $S(t)$, $\beta(t)$, and $\gamma(t)$ from equations (2), (3), and (7). Note that this problem is inverse and require a special treatment. The approach used here is based on the Method of Variational Imbedding (MVI)^{14,36}. Following the idea of MVI, we construct a functional using the original equations (2), (3), and the available data values (7). We define the functional over the sub-interval $[0, P]$ as

$$\mathcal{F} = \int_0^P \left[L_1^2 + L_2^2 + \sum_{i=1}^{P-1} \delta(t - v_i) \mu_i ((S(t) - \sigma_i)^2 + (I(t) - \lambda_i)^2) \right] dt, \tag{8}$$

where μ_i are the weights prescribed for the i -th node and $\delta(t)$ is the Dirac delta function $\delta(t - v_i)$ defined as:

$$\delta(t - v_i) = \begin{cases} \infty, & t = v_i \\ 0, & t \neq v_i \end{cases} \quad \text{and} \quad \int_{-\infty}^{\infty} \delta(t - v_i) dt = 1.$$

In other words, we substitute the problem for finding the unknown functions $I(t)$, $S(t)$ and the coefficients $\beta(t)$, $\gamma(t)$ in equations (2), (3), and (7) with a problem for minimization of the functional \mathcal{F} defined by (8). Finding $R(t)$ is straightforward after knowing the values of $I(t)$, $S(t)$, $\beta(t)$, and $\gamma(t)$.

The absolute minimum of \mathcal{F} is equal to zero with the functional becoming zero if and only if equations (2), (3) and conditions (7) are satisfied. We want to emphasize here that the functions (S, I, V) and the parameters (β, γ, u) are unknown and one have to identify them simultaneously by solving a minimization problem.

Method for solving the inverse problem with time-dependent rates

In order to solve the inverse problem with time-dependent rates, we first find the minimum of the functional \mathcal{F} , defined in (8), numerically over a sub-interval of the entire period assuming constant transmission and recovery rates. The details of solving the minimization sub-problem are given in ‘‘Appendix A’’.

Let the values of $S(t)$ and $I(t)$ be known at some time moments v_0, v_1, \dots, v_m , shown in Fig. 3, namely

$$S(v_l) = \sigma_l, \quad I(v_l) = \lambda_l, \quad \text{for } l = 0, 1, \dots, m. \tag{9}$$

We consider the following two approaches for estimating the coefficients β_l and γ_l in the system of equations (2), (3):

- i. Solve the inverse problem for the system of equations (2), (3) under the boundary conditions

$$S(v_{l-1}) = \sigma_{l-1}, \quad I(v_{l-1}) = \lambda_{l-1}, \quad S(v_l) = \sigma_l, \quad I(v_l) = \lambda_l, \tag{10}$$

for $l = 1, 2, \dots, m$. This approach is similar to the method developed in³⁷ for estimating the coefficient in Euler-Bernoulli equation, later modified in³⁶ for the SIR equations. It works well if the data represent the exact values of the functions $S(t)$ and $I(t)$. In a real-world data, e.g. the available public data for the COVID-19 pandemic, there usually exists random noise causing oscillations in the numerical results.

- ii. Use the solution method, proposed in Sect. 3, for estimating the parameters β and γ as constants on every sub-interval $[v_l, v_{l+P}]$. Then, we use the obtained constant values to approximate the non-constant values β_{l+P} and γ_{l+P} for $l = 0, 1, \dots, m - P$. This approach is smoothing the data automatically. Knowing the approximate values of the transmission and recovery rates, we obtain the reproduction rates

$$R_{0,k} = \frac{\beta_k}{\gamma_k} N \quad \text{and} \quad R_{e,k} = \frac{\beta_k}{\gamma_k} \sigma_l, \quad k = P, P + 1, \dots, m.$$

Results

The numerical simulations have been performed using the available public data from³ and³⁸ websites. The number of currently infected persons, λ_k , is reported daily on³⁸. Both websites³ and³⁸, report the total number of infected individuals, $T_k = I_k + R_k$ from the beginning of the COVID-19 pandemic.

	Fully vaccinated (%)	Recovered (%)	Susceptible (%)
World	14.78	2.56	82.66
Israel	57.70	9.43	32.87
United States	49.54	10.79	39.67
Japan	30.10	0.75	69.15

Table 1. Vaccinated, recovered, and susceptible for the World and selected countries as of August 3, 2021.

The daily number of vaccinated people, V_k , is taken from³. We assume that the reported data is correct. In order to obtain a smooth approximation of the vaccinated persons $V(t)$ and the vaccination rate $u(t) = dV/dt$, we use cubic spline approximation based on the values V_k . Consequently, the number of susceptible individuals, σ_k , for a given country can be found in the following way: $\sigma_k = N - T_k - V_k$, where N is the total population.

We do not pretend that the available data are accurate. For sure, the so-called "hidden cases" (asymptomatic cases or cases without official tests) are not included in the reported data. Moreover, the posted data contain random (human) mistakes which can be treated as random noise. Here, we use the posted data to illustrate the method described in "Method for solving the inverse problem with time-dependent rates" section.

The presented results for the rates $N\beta$, $\gamma(t)$, $R_0(t)$, and $R_e(t)$ are based on 28-day, 35-day, and 42-day sub-periods, over the entire multi-month period, from August 2020 until August 3, 2021.

The selected countries (Israel, United States, and Japan) for this study represent population with different vaccination dynamics. Israel performed early aggressive vaccination during the first half of 2021, United States is catching up, whereas Japan is behind compared to them. The World is a good example of aggregated global data. Table 1.

The World. Estimated rates $N\beta$, γ , R_0 , and R_e for the World are shown at Fig. 4. According to the reported data as of August 3, 2021: 14.78% of the population was fully vaccinated; about 2.56% of the population met the virus. Thus, 82.66% of the population remained susceptible. The basic and effective reproduction rates have been near 1 since January 2021.

Israel. During the first half of 2021, Israel was the COVID-19 vaccination champion – over 57.7% of the population of the country was vaccinated by the end of July. At the same time, according to the reports, approximately 9.43% of the population met the virus; hence, only 32.87% of the population was susceptible. According to²¹, early mass vaccination programs predict a reduction of the effective reproduction rate of infection within communities. The estimated values of the basic and effective reproduction ratios R_0 and R_e for Israel are shown in Fig. 5. The effective reproduction rate is practically constant, slightly less than 1 from January to middle June 2021. In June 2021, the rates increased significantly for a short period of time. Then, the reproduction rates decreased but they still remained above 1.

The United States of America. The reported data for The United States of America state: 49.54% were fully vaccinated, about 10.79% of the population met the virus; hence, 39.67% of the population were susceptible. The estimated values of the transmission rate $N\beta$, γ , the basic and effective reproduction ratios R_0 and R_e are shown at Fig. 6. The transmission rate was the highest in fall 2020. It started to increase again in June–July 2021. COVID-19 is on the rise in many countries, causing a new wave. This surge is due to widespread resumption of normal activities.

Japan. Japan is with about 30.10% fully vaccinated individuals, approximately 0.75% reported to be totally infected; this, about 69.15% are susceptible; The estimated values of the transmission rate $N\beta$, γ , the basic and effective reproduction rates R_0 and R_e are shown in Fig. 7. While the rates had been relatively reasonable and low until recently, they started to increase lately. This growth can be explained with the 2020 Summer Olympics held from 23 July to 8 August 2021 in Japan.

Discussion

We performed numerical simulations with the developed method for the A-SIRV model for identifying the transmission, removal and reproduction rates globally and for three countries (Israel, United States, and Japan) using publicly available data. Estimated rates for the World show the presence of three waves and a fourth wave being formed since July 2021.

The values for the identified rates for the World do not oscillate, regardless of the number of days P . By contrast, the obtained rates of the selected individual countries have fluctuations during 2021, which are smaller for $P = 42$ compared to $P = 28$ and $P = 35$. This is because a larger time period is having a smoothing effect of the oscillating functions.

We observe oscillations during time periods that include vaccination data. The vaccination rate $u(t)$ is approximated from the given data for the number of vaccinated individuals $V(t)$. It clearly affects the computed rates in the A-SIRV equations. The oscillations in the rates for Israel, Japan, and to some extent United States can be explained with fluctuations in the vaccination rate $u(t)$ due to irregularities in the data for vaccinated people $V(t)$. We assume that the problem is posed correctly, namely the data have "physical meaning" and, therefore,

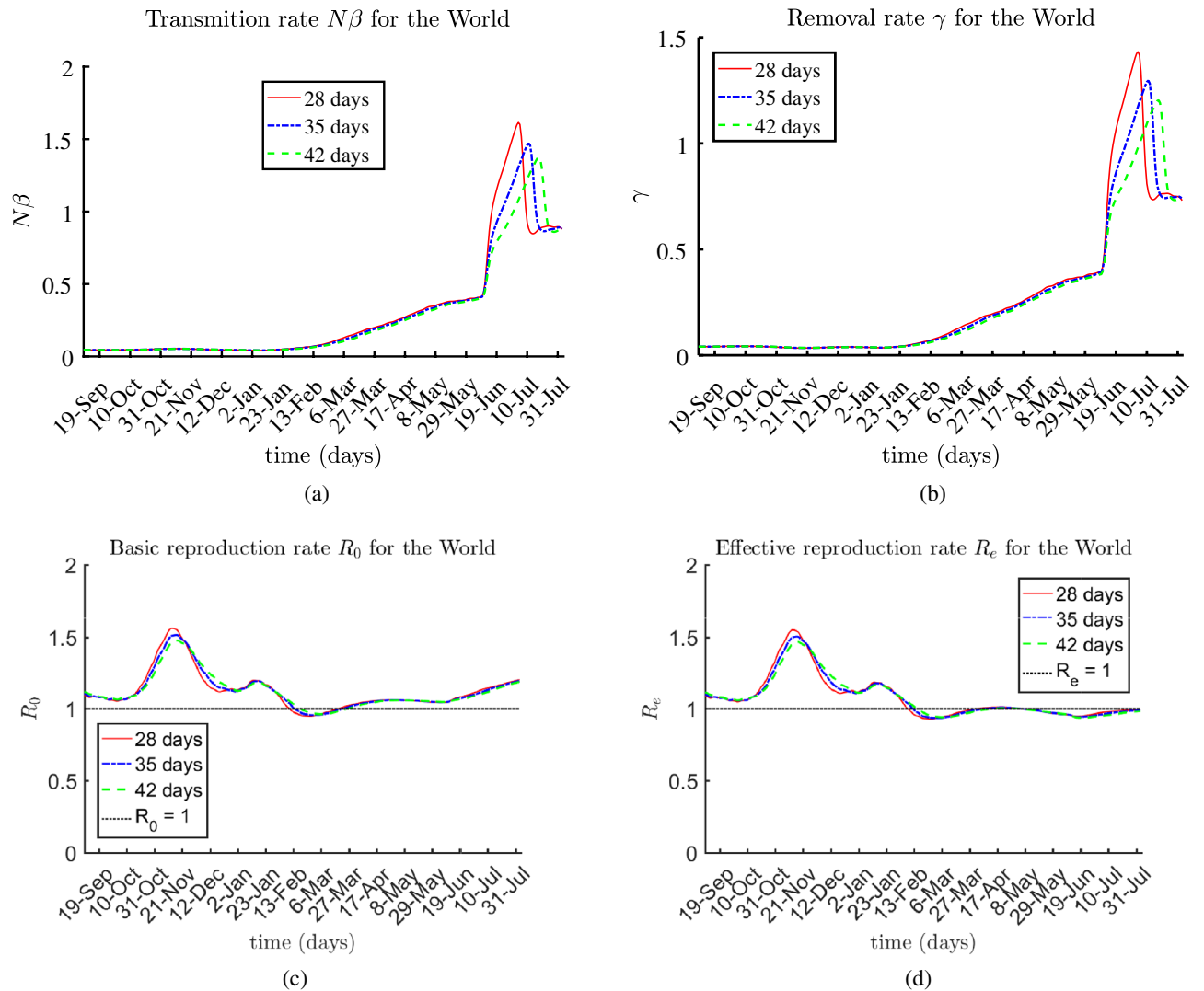


Figure 4. Estimated rates $N\beta$, γ , R_0 , and R_e for the World.

a solution of the problem exists, see³⁹. The randomness of the vaccination data introduces oscillations in the obtained numerical values of the rates.

Finally, it is important to mention that the A-SIRV model considers only one homogeneous population for the selected countries: thus, giving aggregated results for the estimated rates and the unknown functions. The populations of The World and The United State of America are obviously not homogeneous. Such results can still help see trends of potential future growth of the epidemics and, if needed, guide the design of alternative interventions.

Conclusions

Mathematical models can help with visualizing and predicting the long-term behaviour of an infectious disease, despite of the fact that there are many limitations to using them. For instance, the SIR type models contain many assumptions such as: accuracy and completeness of reported data; mixing of the population; no reinfection; constant population; and so forth. The present work studies the performance of a method for an epidemic based on an inverse problem approach for estimating the time-dependent transmission and removal rates in the A-SIRV epidemic model. The inverse problem is solved by defining a minimization problem using the entire dataset for the examined population, with available COVID-19 data. The work utilizes an inverse problem approach to the time-dependent transmission and removal rates identification as well as the unknown functions in the A-SIRV (Adaptive SIRV) model. This can give insight into how well the method identifies the parameters that can be used to predict the infectious disease spread. If conditions change, then the predictions may no longer be accurate; hence, adjustments will be required based on the existing conditions, for obtaining new predictions.

Data availability

The COVID-19 data is publicly available at <https://ourworldindata.org/coronavirus-source-data> and <https://www.worldometers.info/coronavirus/>.

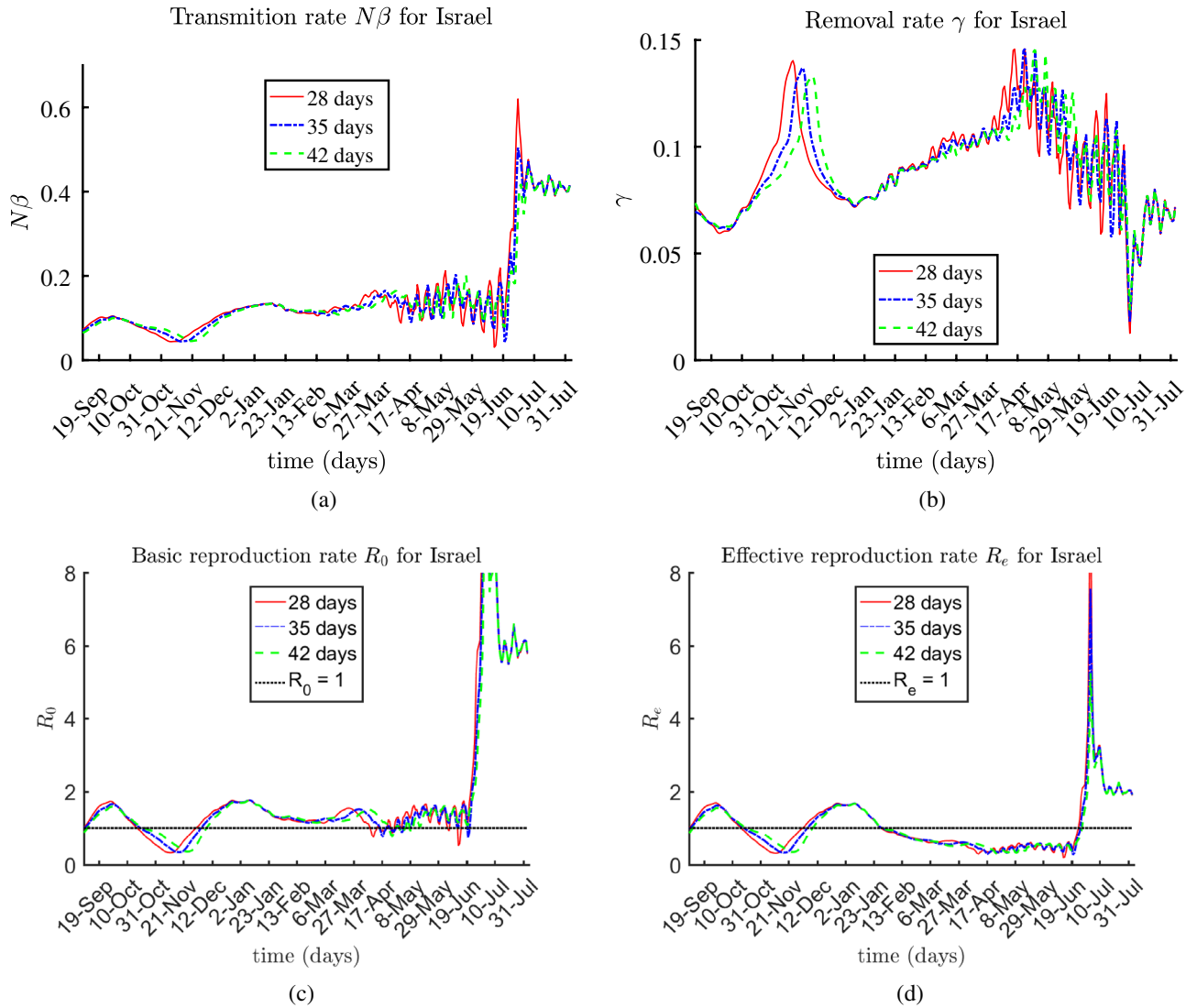


Figure 5. Estimated rates βN , γ , R_0 , and R_e for Israel.

Appendix A: Solving the minimization sub-problem

In order to solve the minimization sub-problem used in Sect. 3, we first approximate the derivatives and integrals in (8).

A.1 Discretization of the minimization sub-problem. Let $\tau = \frac{P}{n}$ be the time-step of a uniform grid on the finite interval $[0, P]$, where n is the total number of grid nodes. The grid nodes are defined as: $t_k = k\tau$, $k = 0, 1, \dots, n$. It is important that τ (respectively n) is chosen to ensure that every time moment v_1, v_2, \dots, v_P coincides with one grid node. Let S_k, I_k , and u_k be notations for the corresponding grid values of the functions $S(t_k), I(t_k)$, and $u(t_k)$.

Since the problem for minimization of the functional \mathcal{F} is non-linear, it requires linearization at some stage of the solution process. Let \hat{S}_k, \hat{I}_k be known approximate values (say from the previous iteration), used in the non-linear term of the differential equations (2), (3). Such linearization has the advantage that it allows building an iterative procedure with a time-independent matrix of the resulting linear system. Hence, we invert this matrix once, at the start of the iterations.

In order to secure $O(\tau^2)$ errors of approximation of the operators in (8), we discretize the derivatives in L_1 and L_2 at the grid nodes $t_{k-1/2}$. We introduce the notations for the approximations of $S(t)$ and $I(t)$ at the midpoints of $[t_{k-1}, t_k]$, where $k = 1, 2, \dots, n$:

$$\hat{S}_{k-1/2} = 0.5(\hat{S}_{k-1} + \hat{S}_k), \quad \hat{I}_{k-1/2} = 0.5(\hat{I}_{k-1} + \hat{I}_k), \quad u_{k-1/2} = 0.5(u_{k-1} + u_k).$$

Consequently, the differential operators L_1 and L_2 are approximated by the following linear difference operators Λ_1 and Λ_2

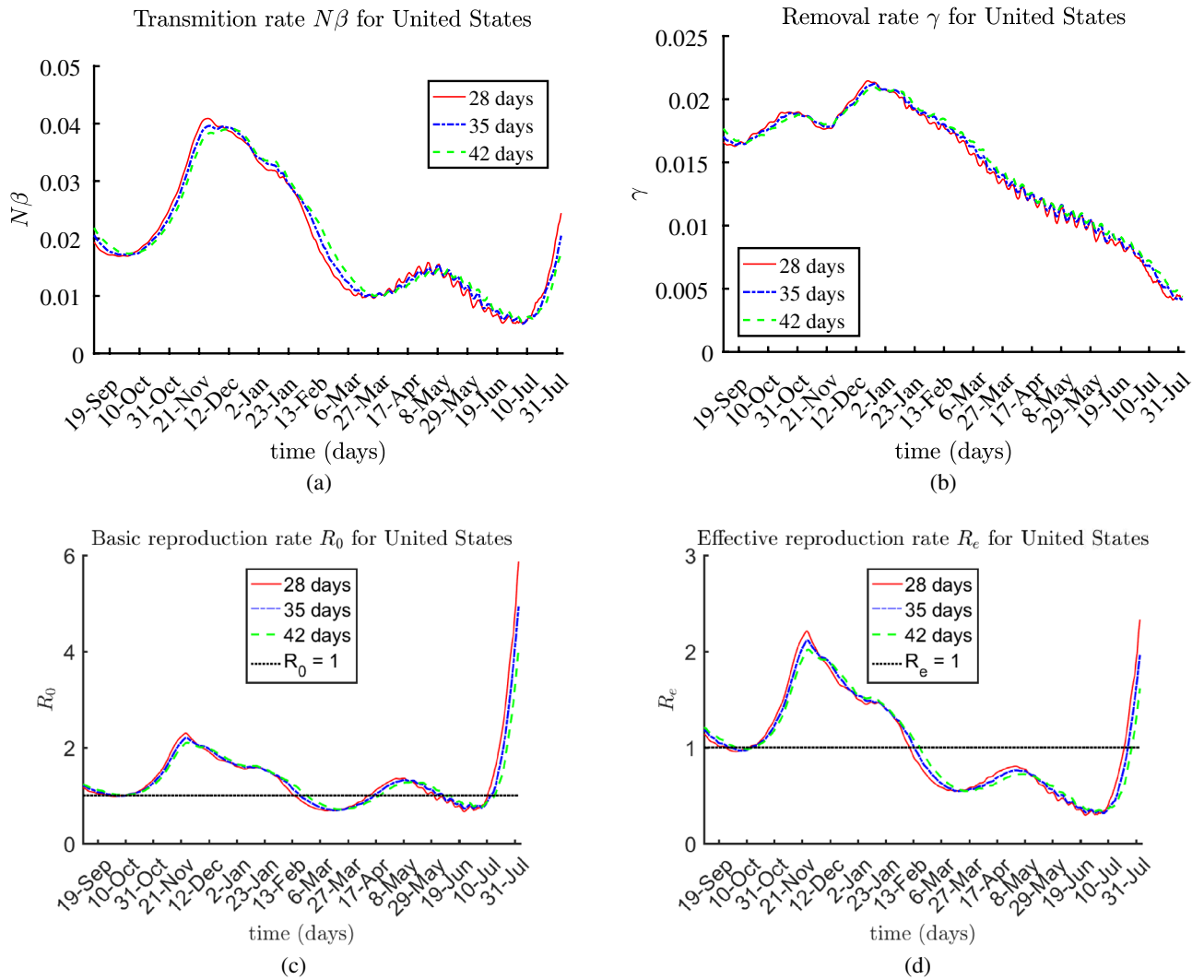


Figure 6. Estimated rates $N\beta$, γ , R_0 , and R_e for United States.

$$\Lambda_{1,k} = \frac{S_k - S_{k-1}}{\tau} + \beta \hat{S}_{k-1/2} \hat{I}_{k-1/2} + u_{k-1/2} \tag{11}$$

$$\Lambda_{2,k} = \frac{I_k - I_{k-1}}{\tau} - \beta \hat{S}_{k-1/2} \hat{I}_{k-1/2} + \gamma \hat{I}_{k-1/2}, \tag{12}$$

for $k = 1, 2, \dots, n$.

We approximate $\delta(t_k - v_i)$ as

$$\delta(t_k - v_i) = \begin{cases} \frac{1}{\tau}, & t_k = v_i \\ 0, & t_k \neq v_i \end{cases} \text{ for } k = 1, 2, \dots, n, \quad i = 1, 2, \dots, P - 1. \tag{13}$$

As already stated, it is important that for every $1 \leq i \leq P - 1$, there exists index k_i , such that $v_i = t_{k_i}$, i.e., the set of time moments $\{v_0, v_1, \dots, v_P\}$, is a subset of the set of grid nodes $\{t_0, t_1, \dots, t_n\}$. To simplify the presentation, let us introduce the notations

$$\bar{\mu}_k = \begin{cases} \mu_i, & t_k = v_i, \\ 0, & \text{otherwise} \end{cases} \quad \bar{\sigma}_k = \begin{cases} \sigma_i, & t_k = v_i, \\ 0, & \text{otherwise} \end{cases} \quad \bar{\lambda}_k = \begin{cases} \lambda_i, & t_k = v_i, \\ 0, & \text{otherwise} \end{cases}.$$

In other words, the values $\bar{\mu}$, $\bar{\sigma}$, and $\bar{\lambda}$ on the grid $\{t_0, t_1, t_2, \dots, t_n\}$ are equal to μ , σ , and λ , respectively, if there exists $k \in \{0, 1, \dots, n\}$ such that $t_k = v_i, 1 \leq i \leq P - 1$. Otherwise, for $t_k \neq v_i$ the values $\bar{\mu}$, $\bar{\sigma}$, and $\bar{\lambda}$ are zero.

This way, the discretized version of the functional \mathcal{F} becomes

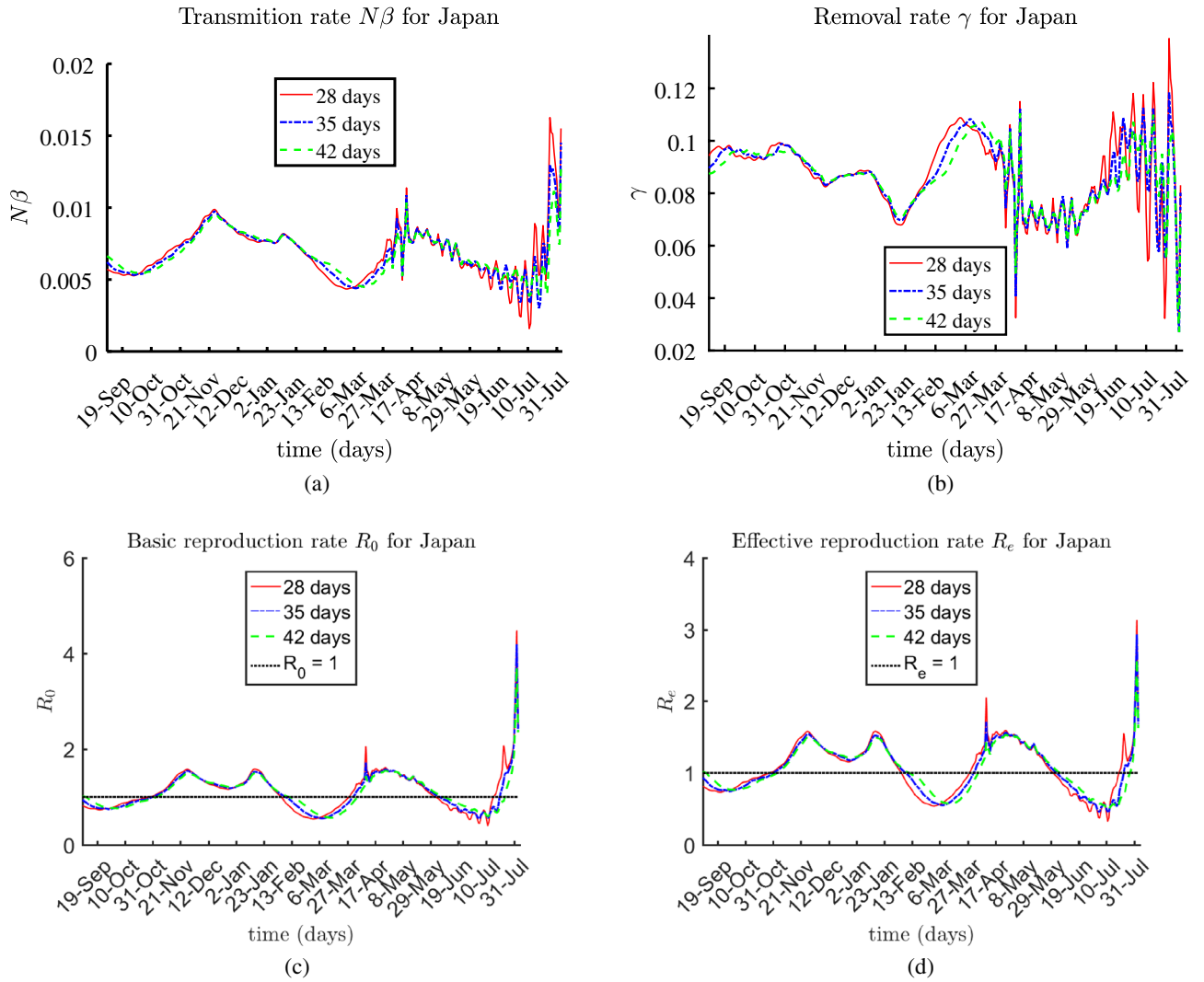


Figure 7. Estimated rates βN , γ , R_0 , and R_e for Japan.

$$\Phi = \sum_{k=1}^n \left[(\Lambda_{1,k}^2 + \Lambda_{2,k}^2) \tau + \bar{\mu}_k \left((S_k - \bar{\sigma}_k)^2 + (I_k - \bar{\lambda}_k)^2 \right) \right]. \tag{14}$$

A.2 Equations for S and I . The necessary conditions for minimization of the function Φ with respect to its arguments S_k and I_k are

$$\frac{\partial \Phi}{\partial S_k} = 0, \quad \frac{\partial \Phi}{\partial I_k} = 0. \tag{15}$$

For the case under consideration, we obtain the following equations in the internal grid nodes t_k , where $k = 1, 2, \dots, n - 1$

$$\begin{aligned} S_{k-1} - (2 + \tau \mu_k) S_k + S_{k+1} \\ = -\tau \beta (\hat{S}_{k+1/2} \hat{I}_{k+1/2} - \hat{S}_{k-1/2} \hat{I}_{k-1/2}) - \tau (u_{k+1/2} - u_{k-1/2}) - \tau \mu_k \sigma_k, \end{aligned} \tag{16}$$

$$\begin{aligned} I_{k-1} - (2 + \tau \mu_k) I_k + I_{k+1} \\ = \tau \beta (\hat{S}_{k+1/2} \hat{I}_{k+1/2} - \hat{S}_{k-1/2} \hat{I}_{k-1/2}) - \tau \gamma (\hat{I}_{k+1/2} - \hat{I}_{k-1/2}) - \tau \mu_k \lambda_k. \end{aligned} \tag{17}$$

The equations at the start grid node t_0 and the end grid node t_n are

$$S_1 - S_0 = \tau (-\beta \hat{S}_{1/2} \hat{I}_{1/2} - u_{1/2}), \tag{18}$$

$$I_1 - I_0 = \tau(\beta \hat{S}_{1/2} \hat{I}_{1/2} - \gamma \hat{I}_{1/2}), \quad (19)$$

$$S_n - S_{n-1} = \tau(-\beta \hat{S}_{n-1/2} \hat{I}_{n-1/2} - u_{n-1/2}), \quad (20)$$

$$I_n - I_{n-1} = \tau(\beta \hat{S}_{n-1/2} \hat{I}_{n-1/2} - \gamma \hat{I}_{n-1/2}), \quad (21)$$

This way, given that approximate values of β and γ are known from the previous iteration, there are two well-posed systems of $(n+1)$ linear equations: (16), (18), and (20) for the unknown set of new values $(S_0, S_1, S_2, \dots, S_n)$; and (17), (19), and (21) for $(I_0, I_1, I_2, \dots, I_n)$. Due to the early linearization, these two systems are with constant matrices. Hence, we need to invert them once and use the inverse matrices during the entire iterative process.

A.3 Equations for β and γ . We rewrite the function Φ in the form

$$\Phi = \alpha_{00} + \alpha_{10}\beta + \alpha_{01}\gamma + \alpha_{20}\beta^2 + \alpha_{11}\beta\gamma + \alpha_{02}\gamma^2, \quad (22)$$

where

$$\begin{aligned} \alpha_{00} &= \sum_{k=1}^n \left[(I_k - I_{k-1})^2 + (S_k - S_{k-1})^2 + \tau^2 u_{k-1/2}^2 \right], \\ &\quad + \bar{\mu}_k \left[(S_k - \bar{\sigma}_k)^2 + (I_k - \bar{\lambda}_k)^2 \right], \\ \alpha_{10} &= \sum_{k=1}^n -2\hat{I}_{k-1/2} \hat{S}_{k-1/2} (I_k - I_{k-1} - S_k + S_{k-1} - \tau u_{k-1/2}) \tau, \\ \alpha_{01} &= \sum_{k=1}^n 2\hat{I}_{k-1/2} (I_k - I_{k-1}) \tau, \\ \alpha_{20} &= \sum_{k=1}^n 2\hat{I}_{k-1/2}^2 \hat{S}_{k-1/2}^2 \tau^2, \quad \alpha_{11} = \sum_{k=1}^n -2\hat{I}_{k-1/2}^2 \hat{S}_{k-1/2} \tau^2, \quad \alpha_{02} = \sum_{k=1}^n \hat{I}_{k-1/2}^2 \tau^2. \end{aligned}$$

The necessary conditions for minimization of the function Φ with respect to β and γ are given by

$$\frac{\partial \Phi}{\partial \beta} = \alpha_{10} + 2\alpha_{20}\beta + \alpha_{11}\gamma = 0, \quad (23)$$

$$\frac{\partial \Phi}{\partial \gamma} = \alpha_{01} + \alpha_{11}\beta + 2\alpha_{02}\gamma = 0. \quad (24)$$

The solution of the system (23), (24) is

$$\beta = -\frac{2\alpha_{02}\alpha_{10} - \alpha_{01}\alpha_{11}}{-\alpha_{11}^2 + 4\alpha_{02}\alpha_{20}}, \quad (25)$$

$$\gamma = -\frac{\alpha_{10}\alpha_{11} - 2\alpha_{01}\alpha_{20}}{\alpha_{11}^2 - 4\alpha_{02}\alpha_{20}}. \quad (26)$$

Up to this point, we have constructed two systems of linear equations, one system for (S, I) under the condition that (β, γ) are given, and another system for (β, γ) under the condition that (S, I) are given. This allows us to build an algorithm for finding a solution to the entire non-linear problem, by means of an iterative procedure, replacing (β, γ) (when calculating (S, I)), or (S, I) (when calculating (β, γ)) with their values from the previous iteration. If the iterations converge, then they will give one of the possible solutions of the problem. Thus, the existence of the solution to the problem can be established *a-posteriori*. If the iterative process diverges, this will mean that there exists no solution to the problem.

A.4 Iterative algorithm for the inverse sub-problem. Algorithm 1 shows the iterative procedure for solving the system (16)–(21), along with obtaining the values of the transmission and removal rates from equations (25) and (26).

τ	$N\beta$	rate $_{\beta}$	γ	rate $_{\gamma}$
1/10	0.1005883672528720	–	0.0741311130168358	–
1/20	0.1005886480876050	–	0.0741311570649426	–
1/40	0.1005887182822550	2.0002884	0.0741311680719028	2.0006639
1/80	0.1005887358165650	2.0011804	0.0741311708184778	2.0027105
1/160	0.1005887402874910	1.9715351	0.0741311715365409	1.9354511

Table 2. Obtained values of the constants β and γ and the rate of convergence for four different values of the mesh spacing τ .

Algorithm 1: Iterative algorithm for solving the inverse sub-problem for S, I, β, γ .

Input: Dataset (S, I, V) , μ , ε_0 , initial values for $\bar{S}, \bar{I}, \beta, \gamma, \tau$.
Output: Approximate values of S, I, β, γ within tolerance ε_0 .

- 1 Find approximate values of u using the dataset V
- 2 **repeat**
- 3 $\bar{\beta} = \beta, \bar{\gamma} = \gamma$
- 4 **repeat**
- 5 Compute S and I from the system (16)–(21), using the current values of $\bar{S}, \bar{I}, \bar{\beta}, \bar{\gamma}$, and the values of u and μ .
- 6 **until** $\|S - \bar{S}\| / \|S\| < \varepsilon_0$ **and** $\|I - \bar{I}\| / \|I\| < \varepsilon_0$
- 7 $\bar{S} = S, \bar{I} = I$
- 8 Compute the new values of β and γ from equations (25) and (26).
- 9 **until** $|\beta - \bar{\beta}|N < \varepsilon_0$ **and** $|\gamma - \bar{\gamma}| < \varepsilon_0$
- 10 Compute R_0 and R_e using the approximate values of β and γ

A.5 Validation of the numerical solution method. The accuracy of the developed numerical method and the corresponding algorithm are verified with mandatory tests involving different values of the time step τ . We confirmed the practical convergence and the $O(\tau^2)$ approximation of the difference scheme. We chose Israel for these tests because this country started mass vaccination of the population relatively early compared to other countries.

The computed numerical values of the identified coefficients β and γ with four different values of τ for a time period $P = 28$ days for Israel are given in Table 2, from May 24, 2021 to June 21, 2021. The rates of convergence are computed using the formulas

$$\text{rate}_{\beta} = \log_2 \frac{\beta_{\tau} - \beta_{2\tau}}{\beta_{2\tau} - \beta_{4\tau}}, \quad \text{rate}_{\gamma} = \log_2 \frac{\gamma_{\tau} - \gamma_{2\tau}}{\gamma_{2\tau} - \gamma_{4\tau}}. \quad (27)$$

The tests confirm the second order of convergence of the numerical solution with respect to τ . The results for other countries and time periods P are similar.

Received: 3 September 2021; Accepted: 12 September 2022

Published online: 20 September 2022

References

1. WHO. World Health Organization, Novel Coronavirus (2019-nCoV) situation reports. <https://www.who.int/> (2020).
2. Callaway, E. The race for coronavirus vaccines: a graphical guide. *Nature* **580**, 576–577. <https://doi.org/10.1038/d41586-020-01221-y> (2020).
3. Ritchie, H. et al. Coronavirus Pandemic (COVID-19). <https://ourworldindata.org/coronavirus> (2020).
4. Kermack, W. & McKendrick, A. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. Ser. A* **115**, 700–721. <https://doi.org/10.1098/rspa.1927.0118> (1927).
5. Ajbar, A., Alqahtani, R. & Boumaza, M. Dynamics of an SIR-based COVID-19 model with linear incidence rate, nonlinear removal rate, and public awareness. *Front. Phys.* **9**, 634251. <https://doi.org/10.3389/fphy.2021.634251> (2021).
6. Griette, Q. & Magal, P. Clarifying predictions for COVID-19 from testing data: The example of New York State. *Infect. Dis. Model.* **6**, 273–283. <https://doi.org/10.1016/j.idm.2020.12.011> (2021).
7. Griette, Q., Magal, P. & Seydi, O. Unreported cases for age dependent COVID-19 outbreak in Japan. *Biology* **9**, 132. <https://doi.org/10.1016/j.jtbi.2020.110501> (2020).
8. Kucharski, A. et al. Early dynamics of transmission and control of COVID-19: A mathematical modelling study. *Lancet Infect. Dis.* **20**, 512–5013. [https://doi.org/10.1016/S1473-3099\(20\)30144-4](https://doi.org/10.1016/S1473-3099(20)30144-4) (2020).
9. Kurkina, E. S. & Koltsova, E. M. Mathematical modeling of the propagation of Covid-19 pandemic waves in the World. *Comput. Math. Model.* **32**, 147–170. <https://doi.org/10.1007/s10598-021-09523-0> (2021).

10. Li, Y. *et al.* Mathematical modeling and epidemic prediction of COVID-19 and its significance to epidemic prevention and control measures. *Ann. Infect. Dis. Epidemiol.* **5**, 1052 (2020).
11. Liao, Z., Lan, P., Liao, Z., Zhang, Y. & Liu, S. TW-SIR: time-window based SIR for COVID-19 forecasts. *Sci. Rep.* **10**, 22454. <https://doi.org/10.1038/s41598-020-80007-8> (2020).
12. Lin, Y. *et al.* Spread and impact of COVID-19 in china: A systematic review and synthesis of predictions from transmission-dynamic models. *Front. Med.* **7**, 321. <https://doi.org/10.3389/fmed.2020.00321> (2020).
13. Lobo, A. *et al.* COVID-19 epidemic in Brazil: Where are we at?. *Int. J. Infect. Dis.* **97**, 382–385. <https://doi.org/10.1016/j.ijid.2020.06.044> (2020).
14. Marinov, T. & Marinova, R. Dynamics of covid-19 using inverse problem for coefficient identification in sir epidemic models. *Chaos Solitons Fract.* **X 5**, 100041. <https://doi.org/10.1016/j.csfx.2020.100041> (2020).
15. Nishiura, H., Linton, N. & Akhmetzhanov, A. Serial interval of novel coronavirus (COVID-19) infections. *Int. J. Infect. Dis.* **93**, 284–286. <https://doi.org/10.1016/j.ijid.2020.02.060> (2020).
16. Pereira, F., Schimit, P. & Bezerra, F. A deep learning based surrogate model for the parameter identification problem in probabilistic cellular automaton epidemic models. *Comput. Methods Prog. Biomed.* **205**, 106078 (2021). <https://www.sciencedirect.com/science/article/pii/S016926072100153X>.
17. Roosa, K. *et al.* Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24th, 2020. *Infect. Dis. Model.* **5**, 256–263. <https://doi.org/10.1016/j.idm.2020.02.002> (2020).
18. Shereen, M., Khan, S., Kazmi, A., Bashir, N. & Siddique, R. COVID-19 infection: Origin, transmission, and characteristics of human coronaviruses. *J. Adv. Res.* **24**, 91–98. <https://doi.org/10.1016/j.jare.2020.03.005> (2020).
19. Wacker, B. & Schlüter, J. Time-continuous and time-discrete SIR models revisited: Theory and applications. *Adv. Differ. Equ.* **2020**, 556. <https://doi.org/10.1186/s13662-020-02995-1> (2020).
20. Cohen, K. *et al.* Longitudinal analysis shows durable and broad immune memory after SARS-CoV-2 infection with persisting antibody responses and memory B and T cells. *Cell Rep. Med.* **2**, 100354. <https://doi.org/10.1016/j.xcrim.2021.100354> (2021).
21. Murray, J. *Mathematical biology. I. An introduction* / J.D. Murray. 3rd ed. (Springer, New York, 1993).
22. Bärnighausen, T., Bloom, D., Cafiero-Fonseca, E. & O'Brien, J. Valuing vaccination. In *Proceedings of the National Academy of Sciences*, vol. 111(34) (ed Rappuoli, R.) 12313–12319 (2014). <https://doi.org/10.1073/pnas.1400475111>.
23. Widyaniingsih, P., Nugroho, A. & Saputro, D. Susceptible infected recovered model with vaccination, immunity loss, and relapse to study tuberculosis transmission in indonesia. In *AIP Conference Proceedings 2014*, 020121 (AIP Publishing, 2018). <https://doi.org/10.1063/1.5054525>.
24. Saha, S., Samanta, G. & Nieto, J. J. Impact of optimal vaccination and social distancing on COVID-19 pandemic. *Math. Comput. Simul.* **200**, 285–314. <https://doi.org/10.1016/j.matcom.2022.04.025> (2022).
25. Xu, Z., Wu, B. & Topcu, U. Control strategies for COVID-19 epidemic with vaccination, shield immunity and quarantine: A metric temporal logic approach. *PLoS ONE* **16**, e0247660. <https://doi.org/10.1371/journal.pone.0247660> (2020).
26. Sy, K. T. L., White, L. & Nichols, B. Population density and basic reproductive number of COVID-19 across United States counties. *PLoS ONE* **16**, e0249271. <https://doi.org/10.1371/journal.pone.0249271> (2021).
27. Schlickeiser, R. & Kröger, M. Analytical modeling of the temporal evolution of epidemics outbreaks accounting for vaccinations. *Physics* **3**, 386–426. <https://doi.org/10.3390/physics3020028> (2021).
28. Etxeberria-Etxaniz, M., Alonso-Quesada, S. & la Sen, M. D. On an SEIR epidemic model with vaccination of newborns and periodic impulsive vaccination with eventual on-line adapted vaccination strategies to the varying levels of the susceptible subpopulation. *Appl. Sci.* **10**, 8296. <https://doi.org/10.3390/app10228296> (2020).
29. Ghostine, R., Gharamti, M., Hassrouny, S. & Hoteit, I. An extended SEIR model with vaccination for forecasting the COVID-19 pandemic in Saudi Arabia using an Ensemble Kalman Filter. *Mathematics* <https://doi.org/10.3390/math9060636> (2021).
30. Sen, M. D. L., Alonso-Quesada, S., Ibeas, A. & Nistal, R. On a discrete SEIR epidemic model with two-doses delayed feedback vaccination control on the susceptible. *Vaccines* **9**, 398. <https://doi.org/10.3390/vaccines9040398> (2021).
31. Rifhat, R., Teng, Z. & Wang, C. Extinction and persistence of a stochastic SIRV epidemic model with nonlinear incidence rate. *Adv. Differ. Equ.* **2021**(1), 1–21. <https://doi.org/10.1186/s13662-021-03347-3> (2021).
32. Zhao, Z. *et al.* Stringent nonpharmaceutical interventions are crucial for curbing COVID-19 transmission in the course of vaccination: A case study of South and Southeast Asian countries. *Healthcare* **9** (2021). <https://www.mdpi.com/2227-9032/9/10/1292>.
33. Meng, X., Cai, Z., Dui, H. & Cao, H. Vaccination strategy analysis with SIRV epidemic model based on scale-free networks with tunable clustering. In *IOP Conference Series: Materials Science and Engineering*, vol. 1043, 032012 (IOP Publishing, 2021). <https://doi.org/10.1088/1757-899x/1043/3/032012>.
34. Law, K. B., Peariasamy, M., Ibrahim, H. M. & Abdullah, N. H. Modelling infectious diseases with herd immunity in a randomly mixed population. *Sci. Rep.* **11**, 20574. <https://doi.org/10.1038/s41598-021-00013-2> (2021).
35. Rella, S. A., Kulikova, Y. A., Dermitzakis, E. T. & Kondrashov, F. A. Rates of SARS-CoV-2 transmission and vaccination impact the fate of vaccine-resistant strains. *Sci. Rep.* **11**, 15729. <https://doi.org/10.1038/s41598-021-95025-3> (2021).
36. Marinov, T., Marinova, R., Omojola, J. & Jackson, M. Inverse problem for coefficient identification in SIR epidemic models. *Comput. Math. Appl.* **67**, 2218–2227. <https://doi.org/10.1016/j.camwa.2014.02.002> (2014).
37. Marinov, T. & Marinova, R. Coefficient identification in Euler-Bernoulli equation from over-posed data. *J. Comput. Appl. Math.* **235**, 450–459. <https://doi.org/10.1016/j.cam.2010.05.048> (2010).
38. Worldometer. COVID-19 Coronavirus Pandemic. <https://www.worldometers.info/coronavirus/> (2020).
39. Tikhonov, A. & Arsenin, V. *Methods for Solving Incorrect Problems* (Nauka, Moscow, 1974).

Acknowledgements

The work of T.M. was supported from the Grant LA Dept. of Health, Office of Public Health, Bureau of Community Preparedness of the State of LA (LDH), contract number LaGov PO#: 2000620243.

Author contributions

T.M. and R.M. contributed to the design and implementation of the research, to the analysis of the results and to the writing of the manuscript. Both authors have approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to T.T.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022