Article

Novel Approach for Identification of Basic and Effective Reproduction Numbers Illustrated with COVID-19

Tchavdar T. Marinov [1],*, Rossitza S. Marinova [2,3], Radoslav T. Marinov [4] and Nicci Shelby [1]

[1] Department of Natural Sciences, Southern University at New Orleans, 6801 Press Drive, New Orleans, LA 70126, USA; nicci.shelby@sus.edu

[2] Department of Mathematical & Physical Sciences, Concordia University of Edmonton, 7128 Ada Boulevard, Edmonton, AB T5B 4E4, Canada; rossitza.marinova@concordia.ab.ca

[3] Department Computer Science, Varna Free University, 9007 Varna, Bulgaria

[4] Rescale, 33 New Montgomery Street, Suite 950, San Francisco, CA 94105, USA; rmarinov@rescale.com

* Correspondence: tmarinov@suno.edu

Abstract: This paper presents a novel numerical technique for the identification of effective and basic reproduction numbers, R_e and R_0, for long-term epidemics, using an inverse problem approach. The method is based on the direct integration of the SIR (Susceptible–Infectious–Removed) system of ordinary differential equations and the least-squares method. Simulations were conducted using official COVID-19 data for the United States and Canada, and for the states of Georgia, Texas, and Louisiana, for a period of two years and ten months. The results demonstrate the applicability of the method in simulating the dynamics of the epidemic and reveal an interesting relationship between the number of currently infectious individuals and the effective reproduction number, which is a useful tool for predicting the epidemic dynamics. For all conducted experiments, the results show that the local maximum (and minimum) values of the time-dependent effective reproduction number occur approximately three weeks before the local maximum (and minimum) values of the number of currently infectious individuals. This work provides a novel and efficient approach for the identification of time-dependent epidemics parameters.

Keywords: SIR model; time-dependent parameters; reproduction numbers; inverse problem; infection and recovery rates; data mining; infectious disease modeling; epidemic dynamics

1. Introduction

Infectious diseases have become a particularly important subject of study in recent years due to the COVID-19 epidemic. Daily reports from across the world have enabled scientists to collect data for research into new approaches for studying infectious diseases. Early identification of outbreaks is of major concern to public health authorities [1–4].

Key parameters of epidemic spread are the basic reproduction number, R_0 and the effective reproduction number, R_e. These measure the severity of the infection.

- The *basic reproduction number* R_0 is defined in epidemiology as the average number of secondary infections caused by a single infected person in a susceptible population, as found in [5–7]. It is used to measure the transmission potential of a disease;
- The *effective reproduction number* R_e is defined as the product of the basic reproduction number and the fraction of the host population that is susceptible. It estimates the rate of spread of an epidemic.

In [8] and other works [6,7,9–12], the focus was on the basic reproduction number, whereas the authors of [13] estimated the effective reproduction number. They applied their method to the spread of COVID-19 in municipalities within Rio de Janeiro, Brazil. Alvarez et al. [14] proposed a variational technique for inverting the renewal equation, based on two formulations of the renewal equation presented in [15,16], to estimate the daily reproduction number.

The SIR model, one of the simplest models for the spread of an infectious disease, was proposed in [17] in 1927. It categorizes individuals as **S**usceptible, **I**nfectious, and **R**ecovered (SIR). The SIR model assumes that infectious individuals transition to recovered and become immune, and the recovered group includes deceased individuals. SIR models displayed compelling results, especially during the early stages of the pandemic [18–31].

Recent works have implemented various SIR-type models to investigate the COVID-19 infectious disease dynamics [10,32–34]. Variations of the basic SIR model applied to COVID-19 data include SIS (S for susceptible, no immunity upon recovery), SIRD (D for deceased—recovered and dead considered separately), SIRV (V for vaccinated), SEIR (E for exposed), MSEIR (M for maternal/passively immune) [35], or MSEIR (meso-scale SEIR) [36]. Researchers have used SIR and SEIR-based models with vaccination to overcome the limitations of the conventional SIR model when applied to the COVID-19 epidemic [37–39]. Zhao et al. [38] applied statistical methods to estimate parameters.

The present work proposes a novel approach for the identification of the time-dependent reproduction numbers. It extends the SIR-type model by assuming time-dependent infection and recovery rates [40,41], which are required for the estimation of the reproduction numbers. The method is based on direct integration of the SIR system and an inverse problem approach for solving the resulting equation for the ratio of the infection and recovery rates. The evaluation of the time-dependent reproduction numbers over long time periods gives a tool for investigating the wave formation of the infection curve for the COVID-19 epidemic. This study examines and shows how the time-dependent effective reproduction number compares to the number of infectious individuals over time and gives a novel approach for predicting epidemic waves.

2. The SIR Model for the Spread of an Infectious Disease

The notations in the SIR model include $S(t)$—the number of *susceptible*, $I(t)$—the number of *infectious*, $R(t)$—the number of *removed* individuals, and t—time. The total population $N = S(t) + I(t) + R(t)$ is considered constant in the above equations.

The *infection rate* β ($0 < \beta < 1$) is the probability that a random infectious 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 an equal probability of coming into contact with every other person. The *recovery rate* γ ($0 < \gamma < 1$) gives the probability that an infectious person recovers.
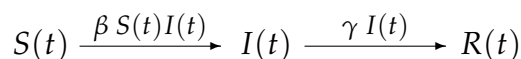
The SIR model consists of the following system of differential equations:

$$\frac{dS(t)}{dt} = -\beta S(t)I(t); \tag{1}$$

$$\frac{dI(t)}{dt} = \beta S(t)I(t) - \gamma I(t); \tag{2}$$

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

Figure 1 shows the diagram for the SIR model, corresponding to the system (1)–(3).

$$S(t) \xrightarrow{\beta\, S(t)I(t)} I(t) \xrightarrow{\gamma\, I(t)} R(t)$$

Figure 1. The SIR epidemic model.

The SIR model assumes that the removed individuals are no longer susceptible or infectious. The number of cases of people recovered from COVID-19 who are re-infected at the present moment is very limited, and the rate cannot be estimated; thus, the possibility of reinfection is not taken into account. Specific details on the assumptions and the adaptive SIR (A-SIR) model can be found in [40,41].

2.1. Integrating the SIR System

Equations (1) and (2) do not depend on the function $R(t)$. Therefore, the SIR system can be viewed as consisting of two parts: (i) Equations (1) and (2), and (ii) Equation (3). If the function $I(t)$ is known, the last Equation (3) can be solved by direct integration.

Let the coefficients β and γ be constant, and $S(t) > 0$, $I(t) > 0$, $\beta \neq 0$, $\gamma \neq 0$.

- From Equations (1) and (2) follows the ordinary differential equation below, connecting the functions I and S (see also [42]):

$$\frac{dI}{dS} = \frac{\beta SI - \gamma I}{-\beta SI}, \tag{4}$$

namely,

$$\frac{dI}{dS} = -1 + \frac{\gamma}{\beta S}. \tag{5}$$

Integrating with respect to S yields the following general solution to Equation (5):

$$I = -S + \frac{\gamma}{\beta} \ln S + C, \tag{6}$$

representing I as a function of S;

- Another way to integrate the system is by excluding dt from Equations (1) and (3). This gives the following ordinary differential equation connecting the functions $S(t)$ and $R(t)$:

$$\frac{dS}{dR} = -\frac{\beta}{\gamma}S. \tag{7}$$

The general solution to Equation (7) is then as follows:

$$\frac{\beta}{\gamma}R \ln S + C = 0, \tag{8}$$

representing R as a function of S.

2.2. The Effective and Basic Reproduction Numbers R_e and R_0

The conditions under which an epidemic occurs present an important question. An epidemic occurs if the number of infected individuals $I(t)$ is increasing [12,42]. This means that the rate of change of the infectious, $\frac{dI}{dt}$, is positive. From Equation (2), the following is implied:

$$\frac{dI(t)}{dt} = \beta S(t)I(t) - \gamma I(t) > 0,$$

which results in the following condition:

$$R_e(t) = \frac{\beta S(t)}{\gamma} > 1. \tag{9}$$

The parameter $R_e(t)$ is called the *effective reproduction number* (also referred to as *ratio* or *rate*) for a given disease. From a statistical point of view, the effective reproduction number is the expected number of secondary cases produced, in a completely susceptible population, by a typical infectious individual.

Another important characteristic of an infectious disease, *the basic reproduction number (ratio, rate)* R_0, can be computed as the ratio of the known rates β and γ over time:

$$R_0 = \frac{\beta N}{\gamma}, \tag{10}$$

where N is the size of the total population.

The basic reproduction number, R_0, is used for answering two important questions, the first being what fraction of the population would become ill should an epidemic occur, and the second being what is the minimum fraction of the population that must be vaccinated in order to prevent an epidemic.

Diseases with smaller values of $R_0 = \frac{\beta N}{\gamma}$ are easier to eradicate than those with larger values R_0, since a population can acquire herd immunity with a smaller fraction of the population vaccinated. For example, smallpox with $R_0 \approx 4$ has been eradicated worldwide, while measles with $R_0 \approx 14$ still has occasional outbreaks.

2.3. The Direct and the Inverse Problem

The initial value problem consisting of the system (1)–(3) with known constant coefficients β and γ, along with proper initial conditions derived from the given data, constitutes the direct problem. In reality, for a new disease, the values of the parameters β and γ are unknown. Finding these parameters involves solving an inverse problem for determining the coefficients and functions from the available data. The inverse problem for the estimation of the constants β and γ in the classical SIR model is solved in [43].

2.4. Time-Dependent Infection and Recovery Rates and Reproduction Numbers

The original SIR model assumes that the infection and recovery rates are constants. Equations (1)–(3), with proper initial conditions, allow the determination of $I(t)$, $S(t)$, and $R(t)$, if the coefficients β and γ are known constants.

However, in the case of a pandemic, the rates may vary over time; hence, $\beta = \beta(t)$ and $\gamma = \gamma(t)$. Multiple factors can cause the rates to change over time. In the case of COVID-19, examples include social distancing, government restrictions, and preventive treatments.

Therefore, the basic reproduction number can be assumed to be a function of time and can be defined:

$$R_0(t) = \frac{\beta(t)}{\gamma(t)} N, \tag{11}$$

with the effective reproduction number as follows:

$$R_e(t) = \frac{\beta(t)}{\gamma(t)} S(t). \tag{12}$$

An algorithm for solving the inverse problem (1) and (2), under the condition that $\beta = \beta(t)$ and $\gamma = \gamma(t)$ are functions of time, is given in [40]. Knowing $\beta = \beta(t)$ and $\gamma = \gamma(t)$ allows us to estimate the reproduction numbers. The present work focuses on alternative methods for directly identifying the reproduction numbers from a given dataset, as well as comparisons of the different approaches.

3. The Inverse Problem for the Time-Dependent Reproduction Numbers

Let the values of $S(t)$ and $I(t)$ at some time moments, $\nu_1, \nu_2, \ldots, \nu_m$, shown in Figure 2, be known and given by the following:

$$S(\nu_l) = \sigma_l, \qquad I(\nu_l) = \lambda_l, \quad l = 1, 2, \ldots, m. \tag{13}$$

If the population N is constant, the values of $R(t)$ at time moments ν_l can be obtained:

$$R(\nu_l) = \rho_l = N - \sigma_l - \lambda_l. \tag{14}$$

Assuming N and $S(t)$ are available, the time-dependent reproduction numbers can be determined after the estimation of the ratio $\gamma(t)/\beta(t)$.

The following two approaches for estimating the constant γ/β in the differential Equation (5), or β/γ in (7), at time moments $\nu_1, \nu_2, \ldots, \nu_m$ are used in this work:

(i)　　By solving the boundary value problem (BVP) for (5) with the following boundary conditions:

$$S(\nu_{l-1}) = \sigma_{l-1}, \qquad I(\nu_{l-1}) = \lambda_{l-1}, \tag{15}$$

$$S(\nu_l) = \sigma_l, \qquad I(\nu_l) = \lambda_l, \tag{16}$$
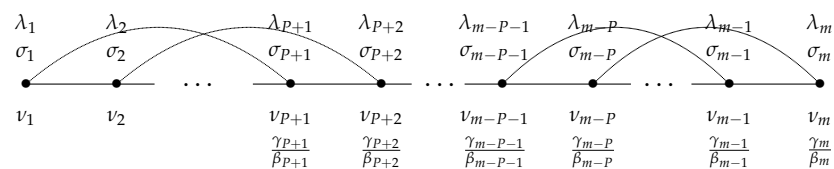
for $l = 2, 3, \ldots, m$;

(i′)　　By solving the BVP for (7) with the following boundary conditions:

$$S(\nu_{l-1}) = \sigma_{l-1}, \qquad R(\nu_{l-1}) = \rho_{l-1}, \tag{17}$$

$$S(\nu_l) = \sigma_l, \qquad R(\nu_l) = \rho_l, \tag{18}$$

for $l = 2, 3, \ldots, m$;

(ii)　　By using the least-squares method (LSM) for estimating the parameter γ/β in (6);

(ii′)　　By using the LSM for estimating the unknown parameter β/γ in (8).

Figure 2. The time nodes and the subsets of fixed length of P days for identifying $\gamma(t)/\beta(t)$.

3.1. Estimating γ/β as a Solution of BVP for Equation (5)

The solution (6) to Equation (5) can be written as shown:

$$\frac{\gamma}{\beta} \ln S + C = S + I. \tag{19}$$

Equation (19) and the conditions (15) and (16), written for the time interval $[\nu_{k-1}, \nu_k]$, result in the following system:

$$\frac{\gamma_k}{\beta_k} \ln(\sigma_{k-1}) + C_k = \sigma_{k-1} + \lambda_{k-1}, \tag{20}$$

$$\frac{\gamma_k}{\beta_k} \ln(\sigma_k) + C_k = \sigma_k + \lambda_k, \tag{21}$$

for $k = 2, 3, \ldots, m$.

The system of Equations (20) and (21), after excluding C_k, gives a formula for the ratio $\frac{\gamma_k}{\beta_k}$:

$$\frac{\gamma_k}{\beta_k} = \frac{\sigma_k - \sigma_{k-1} + \lambda_k - \lambda_{k-1}}{\ln\left(\frac{\sigma_k}{\sigma_{k-1}}\right)}. \tag{22}$$

Substituting the ratio (22) into Equations (11) and (12) gives the following expressions for the basic and effective reproduction numbers $R_{0,k}$ and $R_{e,k}$:

$$R_{0,k} = \frac{\beta_k}{\gamma_k} N = \frac{\ln\left(\frac{\sigma_k}{\sigma_{k-1}}\right)}{\sigma_k - \sigma_{k-1} + \lambda_k - \lambda_{k-1}} N, \quad R_{e,k} = \frac{\beta_k}{\gamma_k} \sigma_k = \frac{\ln\left(\frac{\sigma_k}{\sigma_{k-1}}\right)}{\sigma_k - \sigma_{k-1} + \lambda_k - \lambda_{k-1}} \sigma_k, \tag{23}$$

respectively, on the time interval $[\nu_{k-1}, \nu_k]$ for $k = 2, 3, \ldots, m$.

If the number of susceptible individuals at two consecutive time moments ν_{k-1} and ν_k is the same, i.e., $\sigma_{k-1} = \sigma_k$, then the number of newly infected individuals for this period is zero, and the reproduction number is zero. In other words, the infection is starting to decline.

3.2. Estimating β/γ as a Solution of BVP for Equation (7)

Equation (7) and the conditions (17) and (18), written for the time interval $[v_{k-1}, v_k]$, result in the following system for $k = 2, 3, \ldots, m$:

$$\frac{\beta_k}{\gamma_k}\rho_{k-1} + C_k = -\ln\sigma_{k-1}, \tag{24}$$

$$\frac{\beta_k}{\gamma_k}\rho_k + C_k = -\ln\sigma_k. \tag{25}$$

The system of Equations (24) and (25) gives a formula for the ratio $\frac{\beta_k}{\gamma_k}$:

$$\frac{\beta_k}{\gamma_k} = \frac{\ln\left(\frac{\sigma_{k-1}}{\sigma_k}\right)}{\rho_k - \rho_{k-1}}. \tag{26}$$

Substituting the ratio (26) into Equations (11) and (12) gives the expressions for the basic and effective reproduction numbers $R_{0,k}$ and $R_{e,k}$ on the time interval $[v_{k-1}, v_k]$ for $k = 2, 3, \ldots, m$:

$$R_{0,k} = \frac{\beta_k}{\gamma_k}N = \frac{\ln\left(\frac{\sigma_{k-1}}{\sigma_k}\right)}{\rho_k - \rho_{k-1}}N, \qquad R_{e,k} = \frac{\beta_k}{\gamma_k}\sigma_k = \frac{\ln\left(\frac{\sigma_{k-1}}{\sigma_k}\right)}{\rho_k - \rho_{k-1}}\sigma_k. \tag{27}$$

If the number of removed individuals, i.e., ρ_{k-1} and ρ_k, for two consecutive time moments v_{k-1} and v_k is the same ($\rho_{k-1} = \rho_k$), then the number of newly infected individuals for this period is zero, and the reproduction ratio is also zero. In other words, the infection is diminishing.

3.3. Identifying $R_0(t)$ and $R_e(t)$ from Equation (5) Using LSM

Let the dataset for the values of S and I (σ_k and λ_k, respectively) be divided into subsets of fixed length of P days, as shown in Figure 2.

Following the LSM, we construct the following functions:

$$\Psi\left(\frac{\gamma_k}{\beta_k}, C_k\right) = \sum_{l=k-P}^{k} \mu_l\left(\frac{\gamma_k}{\beta_k}\ln(\sigma_l) + C_k - \sigma_l - \lambda_l\right)^2 \tag{28}$$

for $k = P+1, P+2, \ldots, m$. Here, μ_l is the weight coefficient of the corresponding nodes. The necessary conditions for the minimization of the function Ψ with respect to $\frac{\gamma_k}{\beta_k}$ and C_k are as shown:

$$\frac{\partial\Psi}{\partial\frac{\gamma_k}{\beta_k}} = 0, \qquad \frac{\partial\Psi}{\partial C_k} = 0. \tag{29}$$

Let us introduce the following notations:

$$A_{11} = \sum_{l=k-P}^{k} \mu_l(\ln(\sigma_l))^2, \qquad A_{12} = \sum_{l=k-P}^{k} \mu_l\ln(\sigma_l), \qquad A_{22} = \sum_{l=k-P}^{k} \mu_l, \tag{30}$$

$$B_1 = \sum_{l=k-P}^{k} \mu_l(\sigma_l + \lambda_l)\ln(\sigma_l), \qquad B_2 = \sum_{l=k-P}^{k} \mu_l(\sigma_l + \lambda_l). \tag{31}$$

The necessary conditions (29) become a system of two linear equations for the unknowns $\frac{\gamma_k}{\beta_k}$ and C_k:

$$A_{11}\frac{\gamma_k}{\beta_k} + A_{12}C_k = B_1, \qquad A_{12}\frac{\gamma_k}{\beta_k} + A_{22}C_k = B_2. \tag{32}$$

The system (32) gives the values of the ratio $\frac{\gamma_k}{\beta_k}$ and C_k:

$$\frac{\gamma_k}{\beta_k} = \frac{A_{22}B_1 - A_{12}B_2}{A_{11}A_{22} - A_{12}^2}. \tag{33}$$

Then, the reproduction numbers become the following:

$$R_{0,k} = \frac{\beta_k}{\gamma_k}\,N = \frac{A_{11}A_{22} - A_{12}^2}{A_{22}B_1 - A_{12}B_2}\,N, \qquad R_{e,k} = \frac{\beta_k}{\gamma_k}\,\sigma_k = \frac{A_{11}A_{22} - A_{12}^2}{A_{22}B_1 - A_{12}B_2}\,\sigma_k. \tag{34}$$

Note that, in the case when $P = 1$, the formulas in (34) are equivalent to the corresponding formulas in (23).

3.4. Identifying $R_0(t)$ and $R_e(t)$ from Equation (7) Using LSM

Similarly, following the LSM, we construct the following functions:

$$\Phi\left(\frac{\beta_k}{\gamma_k}, C_k\right) = \sum_{l=k-P}^{k} \mu_l\left(\frac{\beta_k}{\gamma_k}\rho_l + \ln\sigma_l + C_k\right)^2, \tag{35}$$

for $k = P+1, P+2, \ldots, m$. Here, μ_l is the weight coefficient of the corresponding nodes. The necessary conditions for the minimization of the function Φ with respect to $\frac{\beta_k}{\gamma_k}$ and C_k are below:

$$\frac{\partial\Phi}{\partial\frac{\beta_k}{\gamma_k}} = 0, \qquad \frac{\partial\Phi}{\partial C_k} = 0. \tag{36}$$

After introducing the following notations:

$$D_{11} = \sum_{l=k-P}^{k} \mu_l\rho_l^2, \qquad D_{12} = \sum_{l=k-P}^{k} \mu_l\rho_l, \qquad D_{22} = \sum_{l=k-P}^{k} \mu_l, \tag{37}$$

$$E_1 = \sum_{l=k-P}^{k} \mu_l\rho_l\ln(\sigma_l), \qquad E_2 = \sum_{l=k-P}^{k} \mu_l\ln(\sigma_l), \tag{38}$$

the necessary conditions (36) become a system of two linear equations for the unknowns $\frac{\beta_k}{\gamma_k}$ and C_k:

$$D_{11}\frac{\beta_k}{\gamma_k} + D_{12}\,C_k = -E_1, \qquad D_{12}\frac{\beta_k}{\gamma_k} + D_{22}\,C_k = -E_2. \tag{39}$$

The system (39) yields the values of the ratio $\frac{\beta_k}{\gamma_k}$

$$\frac{\beta_k}{\gamma_k} = \frac{D_{12}E_2 - D_{22}E_2}{D_{12}^2 - D_{11}D_{22}}. \tag{40}$$

Then, the reproduction numbers become the following:

$$R_{0,k} = \frac{\beta_k}{\gamma_k}\,N = \frac{D_{12}E_2 - D_{22}E_2}{D_{12}^2 - D_{11}D_{22}}\,N, \qquad R_{e,k} = \frac{\beta_k}{\gamma_k}\,\sigma_k = \frac{D_{12}E_2 - D_{22}E_2}{D_{12}^2 - D_{11}D_{22}}\,\sigma_k. \tag{41}$$

Note that the LSM solution is equivalent to the BVP solution if $P = 1$.
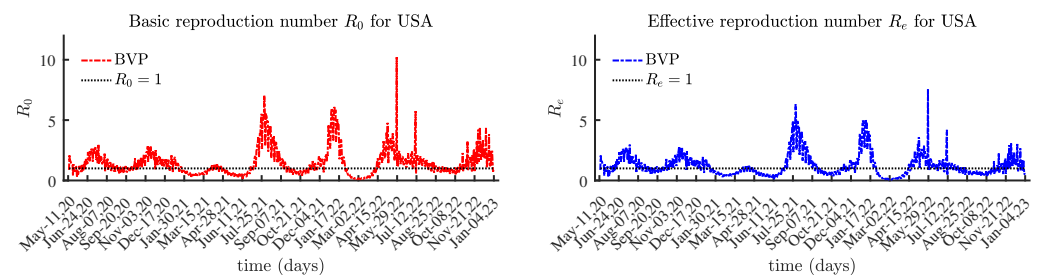
4. Results

We applied the method for reproduction number identification to real COVID-19 data, available in [44], for the countries of the United States and Canada, and the states of Georgia, Texas, and Louisiana, for a period ending on 4 January 2023.

We conducted simulations using the approaches described in Section 3. The solutions to the BVP are very sensitive to random perturbations in data values, as seen in Figure 3, while Figure 4 presents the identified values using the LSM method.

The numerically identified reproduction numbers using BVP and LSM follow the same trend, although the LSM method over 7 or more days helps reduce the effect of reported data containing errors, e.g., due to not being reported daily. This is why the basic and effective reproduction numbers are identified over 3 different time periods: 7 days; 14 days; and 21 days. The results using LSM for the solutions of (6) and (8) are identical. For this reason, herein we present the results obtained using Equation (6).

Figure 3. The basic and effective reproduction numbers for United States with BVP.
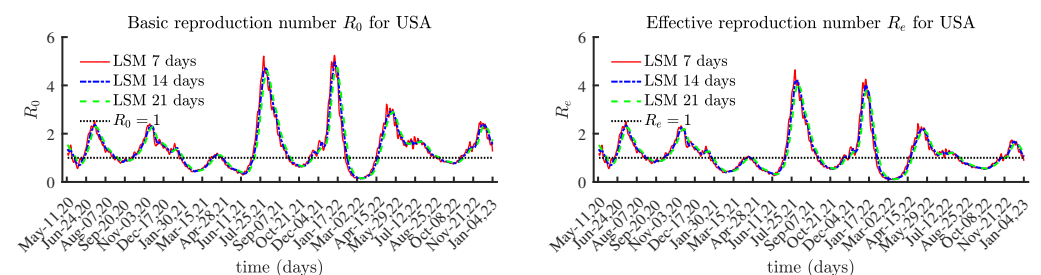
Figure 4. Basic and effective reproduction numbers for United States with LSM.

The results obtained from this work show an interesting relationship between the number of currently infectious individuals (active COVID-19 cases reported in [44]) and the effective reproduction number, namely:

The local maximum/minimum values of the effective reproduction number $R_e(t)$ occur approximately three weeks before the local maximum/minimum values of the number of currently infectious individuals $I(t)$.

This observed behavior is due to the fact that the reproduction number influences the rate of spread of an epidemic. This means that, if the reproduction number starts to increase/decrease, then this causes the infectious disease spread to rise/decline. The obtained numerical results are in agreement with this behavior. Finding mathematical proof of this statement is worth pursuing.

We illustrate the observations for the above relation by comparing the COVID-19 data for the daily infectious individuals $I(t)$ and the computed effective reproduction number $R_e(t)$, based on a 21-day period, for the selected countries and states.

4.1. United States

The obtained results for the basic reproduction number $R_0(t)$ and the effective reproduction number $R_e(t)$ are given in Figure 4. The values for the last few months are lower compared to those prior to February 2022.

Figure 5 compares the local maximum values of the number of currently infectious individuals and the corresponding local maxima of the identified effective reproduction number R_e. This observation provides us with an instrument for predicting the time of the processes of decline in an epidemic wave.
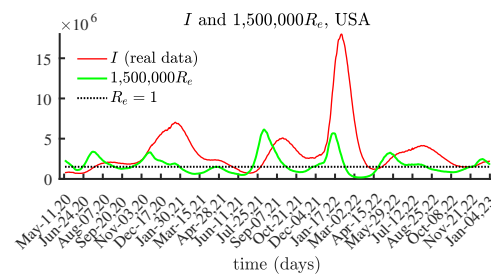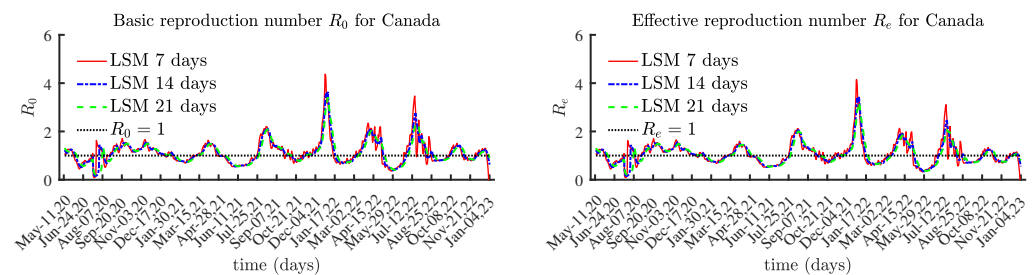
Figure 5. The number of currently infectious people $I(t)$ and the effective reproduction number $R_e(t)$ scaled by 1,500,000, United States.

The shapes of the graphs for the currently infectious people and the effective reproduction number R_e are similar. To demonstrate the difference, we present a graph with two curves, one for the number of infectious individuals, and another for the effective reproduction number R_e, scaled by 1,500,000.

4.2. Canada

Figure 6 presents the obtained results for $R_0(t)$ and $R_e(t)$. As expected, the shapes of the graphs for the currently infectious population and the effective reproduction number $R_e(t)$ are similar. Figure 7 compares the number of currently infectious individuals $I(t)$ and the effective reproduction number $R_e(t)$, scaled by 50,000. Figure 7 confirms the connection between the local maximum values of the number of currently infectious people and the corresponding local maxima of the effective reproduction number.
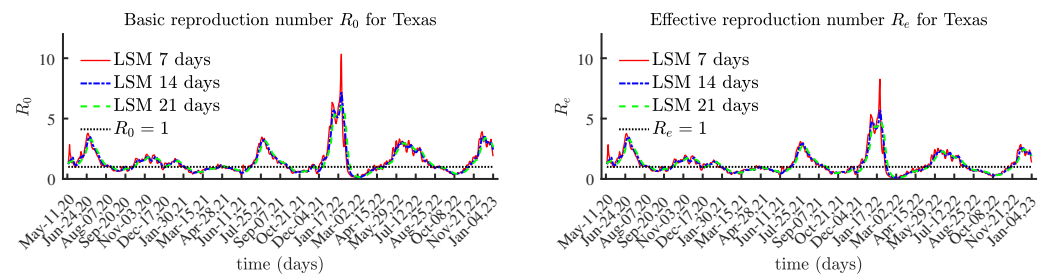

Figure 6. The basic and effective reproduction numbers for Canada.
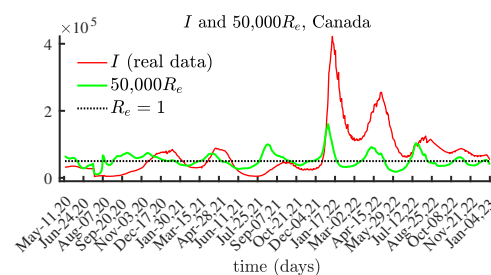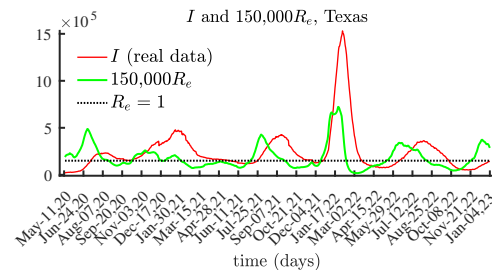

Figure 7. The number of currently infectious people $I(t)$ and the effective reproduction number $R_e(t)$ scaled by 50,000, Canada.

4.3. Texas

The results for Texas, shown in Figures 8 and 9, show a similar tendency. Figure 9 clearly illustrates the connection between local maximum values of the number of currently infectious individuals and the corresponding local maxima of the effective reproduction number $R_e(t)$, multiplied by 150,000.
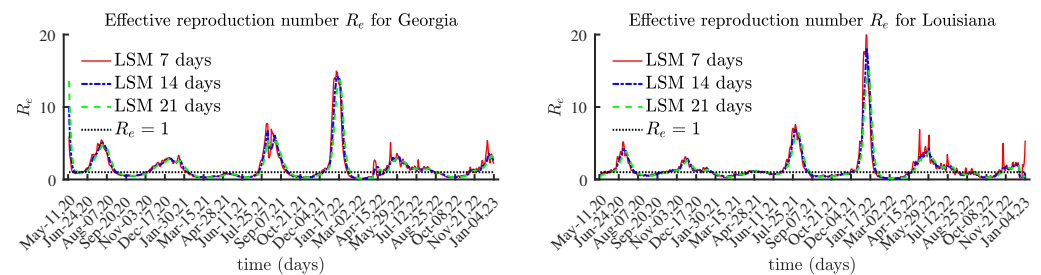
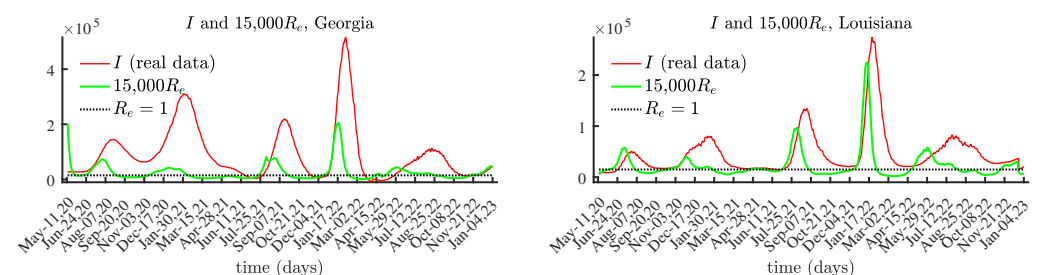**Figure 8.** The basic and effective reproduction numbers for Texas.

Figure 9. The number of currently infectious people $I(t)$ and the effective reproduction number $R_e(t)$ scaled by 150,000, Texas.

4.4. Georgia and Louisiana

Georgia and Louisiana have relatively small populations compared to Texas, Canada, and the United States. Figure 10 presents the results for $R_e(t)$, while Figure 11 displays $I(t)$ and 15,000 $R_e(t)$ for the states of Georgia and Louisiana. These results clearly confirm the claim that the local maximum values of the effective reproduction number $R_e(t)$ occur a few weeks before the local maximum values of the number of currently infectious individuals $I(t)$.

Figure 10. The effective reproduction number for Georgia and Louisiana.

Figure 11. The number of currently infectious people $I(t)$ and the effective reproduction number $R_e(t)$ scaled by 15,000, Georgia and Louisiana.

5. Discussion

This work presents an instrument for the study and analysis of the spread of infectious diseases within the scope of compartmental SIR-type models. We have developed efficient methods for determining the time-dependent basic and effective reproduction numbers,

$R_0(t)$ and $R_e(t)$, through an inverse problem approach. We have applied the developed method to publicly reported COVID-19 data in [44] for Canada and the United States, as well as for the states of Texas, Georgia, and Louisiana. The method uses direct integration of equations resulting from the SIR system and allows for time-dependent infection and recovery rates, thus enabling the results to capture the changes in the reproduction numbers.

We applied the proposed methods to obtain explicit expressions for the time-dependent basic and effective reproduction numbers, $R_0(t)$ and $R_e(t)$, respectively. The inverse problem for the time-dependent reproduction numbers uses available data for the number of susceptible and infectious individuals. We developed two approaches for the identification of the reproduction numbers, namely, the Boundary Value Problem (BVP) and Least Squares Method (LSM), for the derived equations.

Next, we presented and analyzed the numerical results for the selected counties and states. The obtained results demonstrate a relationship between the values of the number of infected individuals $I(t)$ and the scaled effective reproduction number $R_e(t)$. The curves follow a distinct pattern, where the local maxima and minima differ by a few weeks, with the effective reproduction number extrema preceding those of the reported data for $I(t)$. In the future, we aim to provide a mathematical proof of the observed correlation between the maximum (and minimum) values of the number of currently infectious individuals and those of the effective reproduction number.

Mathematical modeling of infectious diseases has many known limitations. This study involves assumptions, such as that the reported data include all the cases; that the population under study is well mixed; and that infectious individuals leave the $I(t)$ class and move directly into the $R(t)$ class, although it is now known that some who recovered from COVID-19 individuals can be infected again. The proposed model does not consider asymptomatic and presymptomatic transmissions, which play an important role in spreading an epidemic [45]. Furthermore, numerous symptomatic cases of COVID-19 are unreported. Neglecting to account for these cases leads to artificially inflated transmission rates and, consequently, higher reproduction numbers.

We tested the developed algorithms against reported data for the active cases from [44]. However, we acknowledge that the assumptions for the reported data pose a limitation, since they are not exactly true for COVID-19. Negative tests do not guarantee that individuals are not currently infectious, and there are diagnostic delays (especially with the omicron variant) which may not be detected on the widely used lateral flow test kits until five or six days later, according to [46]. The SIR-type models assume mixing of the population, lack of reinfection, a constant population, and accuracy and completeness of the reported data. Therefore, the models may not capture all the complexities of the pandemic, and the results should be interpreted with caution.

6. Conclusions

This work presents a novel method for modeling epidemics by using an inverse problem approach to efficiently estimate time-dependent reproduction numbers. We solve the inverse problem by defining the boundary value problem (BVP) and the least squares method (LSM) problems, using datasets of the examined population obtained from publicly available COVID-19 data. The estimated values of the reproduction numbers provide a tool for predicting the spread of infectious diseases. Despite the limitations of the model and the data quality, this study presents a framework for analyzing the spread of infectious diseases. In particular, the observed relationship between the effective reproduction number and the infection data can provide insight into the behavior of the infection curve.

Author Contributions: T.T.M., R.S.M., R.T.M. and N.S. contributed to the design and implementation of the research, to the analysis of the results, and to the writing and reviewing of the manuscript. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All the data used in this work are available and published. The COVID-19 data used in this research are available at https://www.worldometers.info/coronavirus/.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

SIR   Susceptible–Infectious–Removed
BVP   Boundary-Value Problem
LSM   Least-Squares Method

References

1. Cooper, D.L.; Smith, G.E.; Regan, M.; Large, S.; Groenewegen, P.P. Tracking the spatial diffusion of influenza and norovirus using telehealth data: A spatiotemporal analysis of syndromic data. *BMC Med.* **2008**, *6*, 16. https://doi.org/10.1186/1741-7015-6-16.
2. Pelosi, P.; Tonelli, R.; Torregiani, C.; Baratella, E.; Confalonieri, M.; Battaglini, D.; Marchioni, A.; Confalonieri, P.; Clini, E.; Salton, F.; et al. Different Methods to Improve the Monitoring of Noninvasive Respiratory Support of Patients with Severe Pneumonia/ARDS Due to COVID-19: An Update. *J. Clin. Med.* **2022**, *11*, 1704. https://doi.org/10.3390/jcm11061704.
3. Sousa Rêgo, L.O.; Alves Braga, L.L.; Vilas-Boas, G.S.; Oliveira Cardoso, M.S.; Duraes, A.R. Cardiovascular and Neurological Complications of COVID-19: A Narrative Review. *J. Clin. Med.* **2023**, *12*, 2819. https://doi.org/10.3390/jcm12082819.
4. Silva, M.J.A.; Ribeiro, L.R.; Gouveia, M.I.M.; Marcelino, B.d.R.; Santos, C.S.d.; Lima, K.V.B.; Lima, L.N.G.C. Hyperinflammatory Response in COVID-19: A Systematic Review. *Viruses* **2023**, *15*, 553. https://doi.org/10.3390/v15020553.
5. van den Driessche, P. Reproduction numbers of infectious disease models. *Infect. Dis. Model.* **2017**, *2*, 288–303. https://doi.org/10.1016/j.idm.2017.06.002.
6. Dharmaratne, S.; Sudaraka, S.; Abeyagunawardena, I.; Manchanayake, K.; Kothalawala, M.; Gunathunga, W. Estimation of the basic reproduction number (R0) for the novel coronavirus disease in Sri Lanka. *Virol. J.* **2020**, *17*, 144. https://doi.org/10.3390/app10228296.
7. Locatelli, I.; Trächsel, B.; Rousson, V. Estimating the basic reproduction number for COVID-19 in Western Europe. *PLoS ONE* **2021**, *16*, e0248731. https://doi.org/10.1371/journal.
8. Demongeot, J.; Griette, Q.; Magal, P. SI epidemic model applied to COVID-19 data in mainland China. *R. Soc. Open Sci.* **2020**, *7*, 144. https://doi.org/10.1186/s12985-020-01411-0.
9. Korolev, I. Identification and estimation of the SEIRD epidemic model for COVID-19. *J. Econom.* **2021**, *220*, 63–85. https://doi.org/10.1016/j.jeconom.2020.07.038.
10. Yaagoub, Z.; Allali, K. Global Stability of Multi-Strain SEIR Epidemic Model with Vaccination Strategy. *Math. Comput. Appl.* **2023**, *28*, 9. https://doi.org/10.3390/mca28010009.
11. Sy, K.T.L.; White, L.; Nichols, B. Population density and basic reproductive number of COVID-19 across United States counties. *PLoS ONE* **2021**, *16*, e0249271. https://doi.org/10.1371/journal.pone.0249271.
12. Barratt, H.; Epidemiology. In *Public Health Textbook*; The UK Faculty of Public Health: London, UK, 2009.
13. Jorge, D.C.P.; Oliveira, J.F.; Miranda, J.G.V.; Andrade, R.F.S.; Pinho, S.T.R. Estimating the effective reproduction number for heterogeneous models using incidence data. *R. Soc. Open Sci.* **2022**, *9*, 220005. https://doi.org/10.1098/rsos.220005.
14. Alvarez, L.; Colom, M.; Morel, J.D.; Morel, J.M. Computing the daily reproduction number of COVID-19 by inverting the renewal equation using a variational technique. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2105112118. https://doi.org/10.1073/pnas.2105112118.
15. Nishiura, H.; Chowell, G. The effective reproduction number as a prelude to statistical estimation of time-dependent epidemic trends. In *Mathematical and Statistical Estimation Approaches in Epidemiology*; Chowell, G., Hyman, J.M., Bettencourt, L.M.A., Castillo-Chavez, C., Eds.; Springer: Dordrecht, The Netherlands, 2009; Volume 118, pp. 103–121.
16. Fraser, C. Estimating individual and household reproduction numbers in an emerging epidemic. *PLoS ONE* **2007**, *2*, e758.
17. Kermack, W.; McKendrick, A. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. Ser. A* **1927**, *115*, 700–721. https://doi.org/10.1098/rspa.1927.0118.
18. 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.* **2021**, *9*, 634251. https://doi.org/10.3389/fphy.2021.634251.

19. Cohen, K.; Linderman, S.; Moodie, Z.; Czartoski, J.; Lai, L.; Mantus, G.; Norwood, C.; Edara, V.; Floyd, K.; Rosa, S.D.; 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.* **2021**, 2, 100354. https://doi.org/10.1016/j.xcrm.2021.100354.

20. Griette, Q.; Magal, P. Clarifying predictions for COVID-19 from testing data: The example of New York State. *Infect. Dis. Model.* **2021**, *6*, 273–283. https://doi.org/10.1016/j.idm.2020.12.011.

21. Griette, Q.; Magal, P.; Seydi, O. Unreported cases for age dependent COVID-19 outbreak in Japan. *Biology* **2020**, *9*, 132. https://doi.org/10.1016/j.jtbi.2020.110501.

22. Kucharski, A.; Russell, T.; Diamond, C.; Liu, Y. Early dynamics of transmission and control of COVID-19: A mathematical modelling study. *Lancet Infect. Dis.* **2020**, *20*, 512–5013. https://doi.org/10.1016/S1473-3099(20)30144-4.

23. Kurkina, E.S.; Koltsova, E.M. Mathematical Modeling of the Propagation of Covid-19 Pandemic Waves in the World. *Comput. Math. Model.* **2021**, *32*, 147–170. https://doi.org/10.1007/s10598-021-09523-0.

24. Li, Y.; Wang, B.; Peng, R.; Zhou, C.; Zhan, Y.; Liu, Z.; Jiang, X.; Zhao, B. Mathematical Modeling and Epidemic Prediction of COVID-19 and Its Significance to Epidemic Prevention and Control Measures. *Ann. Infect. Dis. Epidemiol.* **2020**, *5*, 1052.

25. Lin, Y.; Duan, Q.; Zhou, Y.; Yuan, T.; Li, P.; Fitzpatrick, T.; Fu, L.; Feng, A.; Luo, G.; Zhan, Y.; et al. Spread and Impact of COVID-19 in China: A Systematic Review and Synthesis of Predictions From Transmission-Dynamic Models. *Front. Med.* **2020**, *7*, 321. https://doi.org/10.3389/fmed.2020.00321.

26. Lobo, A.; dos Santos, A.C.; Rocha, M.; Pinheiro, R.; Bremm, J.; Macário, E.; de Oliveira, W.; de França, G. COVID-19 epidemic in Brazil: Where are we at? *Int. J. Infect. Dis.* **2020**, *97*, 382–385. https://doi.org/10.1016/j.ijid.2020.06.044.

27. Nishiura, H.; Linton, N.; Akhmetzhanov, A. Serial interval of novel coronavirus (COVID-19) infections. *Int. J. Infect. Dis.* **2020**, *93*, 284–286. https://doi.org/10.1016/j.ijid.2020.02.060.

28. 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 Programs Biomed.* **2021**, *205*, 106078. https://doi.org/10.1016/j.cmpb.2021.106078.

29. Roosa, K.; Lee, Y.; Luo, R.; Kirpich, A.; Rothenberg, R.; Hyman, J.; Yan, P.; Chowell, G. Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24th, 2020. *Infect. Dis. Model.* **2020**, *5*, 256–263. https://doi.org/10.1016/j.idm.2020.02.002.

30. Shereen, M.; Khan, S.; Kazmi, A.; Bashir, N.; Siddique, R. COVID-19 infection: Origin, transmission, and characteristics of human coronaviruses. *J. Adv. Res.* **2020**, *24*, 91–98. https://doi.org/10.1016/j.jare.2020.03.005.

31. Wacker, B.; Schlüter, J. Time-continuous and time-discrete SIR models revisited: Theory and applications. *Adv. Differ. Equ.* **2020**, *2020*, 556. https://doi.org/10.1186/s13662-020-02995-1.

32. Georgiev, S.; Vulkov, L. Numerical Coefficient Reconstruction of Time-Depending Integer- and Fractional-Order SIR Models for Economic Analysis of COVID-19. *Mathematics* **2022**, *10*, 4247. https://doi.org/10.3390/math10224247.

33. Margenov, S.; Popivanov, N.; Ugrinova, I.; Hristov, T. Mathematical Modeling and Short-Term Forecasting of the COVID-19 Epidemic in Bulgaria: SEIRS Model with Vaccination. *Mathematics* **2022**, *10*, 2570. https://doi.org/10.3390/math10152570.

34. Leonov, A.; Nagornov, O.; Tyuflin, S. Modeling of Mechanisms of Wave Formation for COVID-19 Epidemic. *Mathematics* **2023**, *11*, 167. https://doi.org/10.3390/math11010167.

35. Chen, X. How Does Social Distancing Change COVID-19 Diseasescape: Case Modeling Across Connecticut Towns in the Early Outbreak. *Northeast. Geogr.* **2022**, *13*.

36. Sen, D.; Sen, D. Use of a Modified SIRD Model to Analyze COVID-19 Data. *Ind. Eng. Chem. Res.* **2021**, *60*, 4251–4260. https://doi.org/10.1021/acs.iecr.0c04754.

37. 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.* **2021**, *11*, 15729. https://doi.org/10.1038/s41598-021-95025-3.

38. Zhao, Z.; Li, X.; Liu, F.; Jin, R.; Ma, C.; Huang, B.; Wu, A.; Nie, X. 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* **2021**, *9*, 1292. https://doi.org/10.3390/healthcare9101292.

39. Marinov, T.; Marinova, R. Adaptive SIR model with vaccination: Simultaneous identification of rates and functions illustrated with COVID-19. *Sci. Rep.* **2022**, *12*, 15688. https://doi.org/10.1038/s41598-022-20276-7.

40. Marinov, T.; Marinova, R. Dynamics of COVID-19 Using Inverse Problem for Coefficient Identification in SIR Epidemic Models. *Chaos, Solitons Fractals X* **2020**, *5*, 100041.

41. Marinov, T.; Marinova, R. Inverse problem for adaptive SIR model: Application to COVID-19 in Latin America. *Infect. Dis. Model.* **2022**, *7*, 134–148.

42. Murray, J. *Mathematical Biology. I. An Introduction*, 3rd ed.; Springer: New York, NY, USA, 2002.

43. Marinov, T.; Marinova, R.; Omojola, J.; Jackson, M. Inverse problem for coefficient identification in SIR epidemic models. *Comput. Math. Appl.* **2014**, *67*, 2218–2227. https://doi.org/10.1016/j.camwa.2014.02.002.

44. Worldometer. COVID-19 Coronavirus Pandemic. Available online: https://www.worldometers.info/coronavirus/ (accessed on 5 January 2023).

45. Wu, P.; Liu, F.; Chang, Z.; Lin, Y.; Ren, M.; Zheng, C.; Li, Y.; Peng, Z.; Qin, Y.; Yu, J.; et al. Assessing Asymptomatic, Presymptomatic, and Symptomatic Transmission Risk of Severe Acute Respiratory Syndrome Coronavirus 2. *Clin. Infect. Dis.* **2021**, *73*, e1314–e1320. https://doi.org/10.1093/cid/ciab271.

46. Ng, Q.X.; Lim, Y.L.; Han, M.X.; Teoh, S.E.; Thumboo, J.; Tan, B.H. Performance of Lateral Flow Tests in the Age of the Omicron: A Rapid Systematic Review. *Life* **2022**, *12*, 1941. https://doi.org/10.3390/life12111941.