Mathematical Biosciences and Engineering

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

MBE, 19(3): 2800–2818.
DOI: 10.3934/mbe.2022128
Received: 29 October 2021
Revised: 20 December 2021
Accepted: 22 December 2021
Published: 13 January 2022

Research article

Calibration and prediction for the inexact SIR model

Yan Wang[1], **Guichen Lu** [2] **and Jiang Du**[1,*]

[1] School of Statistics and Data Science, Beijing University of Technology, Beijing 100124, China
[2] School of Science, Chongqing University of Technology, Chongqing 400054, China

* **Correspondence:** Email: dujiang@bjut.edu.cn.

Abstract: A Susceptible Infective Recovered (SIR) model is usually unable to mimic the actual epidemiological system exactly. The reasons for this inaccuracy include observation errors and model discrepancies due to assumptions and simplifications made by the SIR model. Hence, this work proposes calibration and prediction methods for the SIR model with a one-time reported number of infected cases. Given that the observation errors of the reported data are assumed to be heteroscedastic, we propose two predictors to predict the actual epidemiological system by modeling the model discrepancy through a Gaussian Process model. One is the calibrated SIR model, and the other one is the discrepancy-corrected predictor, which integrates the calibrated SIR model with the Gaussian Process predictor to solve the model discrepancy. A wild bootstrap method quantifies the two predictors' uncertainty, while two numerical studies assess the performance of the proposed method. The numerical results show that, the proposed predictors outperform the existing ones and the prediction accuracy of the discrepancy-corrected predictor is improved by at least 49.95%.

Keywords: Inexact SIR model; Gaussian Process model; Wild bootstrap; uncertainty quantification; Heteroscedastic noise; calibration

1. Introduction

The Susceptible Infective Recovered (SIR) model is a commonly used mathematical model for understanding the dynamics of infectious diseases [1–3]. It assumes that the total population can be divided into three distinct classes of sub-populations: susceptible, infectious, and recovered, whose numbers are denoted by S, I, and R, respectively. The susceptible class of individuals includes members of the population that have the potential to contract a disease. The infected class of individuals is assumed to have contracted the disease, while the recovered class comprises those who recovered and cannot contract the disease again. Moreover, the SIR model assumes that the number of individuals per class changes with time, i.e., $S(t)$, $I(t)$ and $R(t)$ are functions of time t and the total population size

$N = S(t) + I(t) + R(t)$.

The SIR model can be expressed by the following set of ordinary differential equations:

$$\begin{aligned} \frac{dS}{dt} &= -\frac{\beta IS}{N}, \\ \frac{dI}{dt} &= \frac{\beta IS}{N} - \gamma I, \\ \frac{dR}{dt} &= \gamma I, \end{aligned} \tag{1.1}$$

with initial conditions

$$S(0) = S_0 > 0, I(0) = I_0 > 0, R(0) = 0.$$

Here, β is the transmission parameter. The transmission rate can dramatically decrease during the epidemic due to governmental lockdown policies, with an exponential function describing the transmission rate decrease [4]:

$$\beta(t) = \beta_0 + \beta_1 exp(-\mu t). \tag{1.2}$$

where $\beta_0 + \beta_1$ is the initial infection kinetic, with β_0 denoting the infection kinetic at infinite time, which can be near zero due to a long enough period of people isolation [5] and μ characterizes the time of decrease: the larger the μ, the faster $\beta(t)$ decreases to β_0. The model parameter γ is the recovery rate, which is affected by the medical resources and medical level. Improving knowledge on infectious diseases allows healthcare systems to manage the epidemic risks effectively, i.e., the recovery rate can be treated as a constant.

The accuracy of the SIR model is affected by the four model parameters β_0, β_1, μ, and γ. Estimating these parameters by a given set of observations is called *calibration* for the SIR model, and these parameters are the *calibration parameters*. To efficiently estimate the calibration parameters, two uncertainty types between the actual observations and the SIR model must be fully considered.

The first uncertainty is related to the observation errors between the actual observations and the actual epidemiological system. A most commonly used assumption for the observation errors is that these are independent and identically distributed random variables obeying a normal distribution [6]. However, assuming the observation errors are homoscedastic is unrealistic because the observation errors for S (I or R) tend to be more prominent when the number of susceptible (infectious or recovered) people is high [7, 8].

Another uncertainty is the *model uncertainty* between the true epidemiological system and the SIR model [9]. Compared with the true epidemiological system, the SIR model is built under many assumptions, and simplifications that are not true in reality [10]. For example,

- Once a person is recovered, he is no longer susceptible and is immune
- Age, sex, race, and social status do not affect the probability of a person being affected
- The infection and recovery rates are much faster than the time scale of births and deaths

Even with an optimal set of calibration parameters, i.e., optimal parameter, these assumptions, and simplifications impose a SIR model output that does not perfectly fit the actual observations [9].

Therefore, the actual epidemiological system and the SIR model discrepancy cannot be ignored, and the SIR model is an *inexact* model.

In summary, both the heteroscedastic observation errors and the model uncertainties must be considered in the SIR model's calibration and prediction processes.To date, a variety of approaches have been used to estimate the calibration parameters. Several approaches assume that the deterministic SIR model can precisely describe the epidemiological system and the optimal parameters [7, 8, 11, 12]. Among them, several approaches assume that the epidemiological system can be precisely described by the deterministic SIR model and the optimal parameters. Hence, these methods ignore the model discrepancy between the actual epidemiological system and the SIR model. Moreover, most of the existing calibration methods involving an inexact model assume the observation errors are homoscedastic [13–16]. Nevertheless, Sung *et al.* [17] utilized replicated actual observations to calibrate an inexact model with heteroscedastic observation errors. However, in many cases, replicates are not available. In this work, we proposed a novel calibration and prediction method for the SIR model with one-time reported infected data. Our contributions can be summarized as follows:

- We proposed a Weighted Least Squares estimator for the calibration parameters. The proposed estimator makes full use of the monotonically increasing relationship between the observation errors and the infected cases.
- We proposed two different predictors to predict the actual epidemiological system. The proposed predictors take full consideration of the model's uncertainty and heteroscedastic observation errors.
- The proposed method improves the prediction accuracy of the SIR model without complicating the model.
- A wild bootstrap method is adopted to quantify the uncertainties of the proposed predictors.

The remainder of this paper is organized as follows: Section 2 introduces a new calibration and prediction method for the SIR model. Section 3 conducts two numerical studies to compare the prediction performance of the proposed method against existing techniques. Finally, Section 4 concludes this work and suggests some future search topics.

2. Methodology

We first introduce a full calibration model that considers the model uncertainty and observation error. Then, Subsection 2.1 proposes a Weighted Least Squares (WLS) estimation method to calculate the calibration parameters in this full model along with two predictors for the actual epidemiological system. Subsection 2.2 demonstrates some computational difficulties in calibration and prediction, while Subsection 2.3 presents the uncertainty quantification of these two predictors by using a wild bootstrap method.

2.1. WLS calibration

Let the reported infected data $y_1, y_2, \ldots, y_n$ refer to times $t_1, t_2, \ldots, t_n$, where $t_i \in \mathbb{R}, i = 1, \ldots, n$ and $\zeta(\cdot)$ denotes the true epidemiological system. We assume the real observation generated by:

$$y_i = \zeta(t_i) + \epsilon_i, i = 1, \ldots, n. \tag{2.1}$$

where ϵ_i are the observation errors assumed to be independent and have zero mean. Let $I(\cdot, \boldsymbol{\theta})$ be the estimated number of infected population by the SIR model, where $\boldsymbol{\theta} = (\beta_0, \beta_1, \mu, \gamma)^T$ is a set of calibration parameters. Kennedy and O'Hagan [9] employ the following model to consider model uncertainty:

$$\zeta(\cdot) = I(\cdot, \boldsymbol{\theta}^*) + \delta^*(\cdot). \tag{2.2}$$

where $\boldsymbol{\theta}^*$ is the optimal but unknown set of calibration parameters, δ^* is the *model discrepancy* between the true epidemiological system and the SIR model with $\boldsymbol{\theta} = \boldsymbol{\theta}^*$. The combination of (2.1) and (2.2) is named the *KO's model*, which has been widely used in calibrating inexact models.

The calibration goal for the SIR model is to determine $\boldsymbol{\theta}^*$ utilizing the KO's model. Following [9], we assume that the model's discrepancy function is a Gaussian Process $GP(0, \tau^2 K)$, where $\tau^2 K$ is a covariance function with variance τ^2 and K is the correlation function. A common choice for K is the power exponential correlation function given by:

$$K_\phi(h) = \exp\left\{-\frac{h^\alpha}{\phi}\right\}, \tag{2.3}$$

where $h = |t_i - t_j|$ is the distance between t_i and t_j, $\phi > 0$ is an unknown range parameter to be estimated, and $\alpha \in (0, 2]$ is a fixed positive parameter. In this work we adopt [18] and set $\alpha = 1.9$.

As the infected cases increase, the observation errors increase and thus we adopt [7, 8] and assume that the observation errors are in the form:

$$\epsilon_i = \zeta_i^\eta e_i, \tag{2.4}$$

where $(e_i)_{i=1}^n$ are independent $N(0, \sigma^2)$ random variables with $\sigma^2 < \infty$. If $\eta = 0$, $(\epsilon_i)_{i=1}^n$ are independent and identically distributed random variables following the normal distribution $N(0, \sigma^2)$ and since $\zeta(\cdot) \geq 0$, a positive value of β indicates that the variance of ϵ increases as ζ increases. Especially, if $\eta = 1$, the standard deviation of ϵ is assumed to scale linearly with ζ. This situation is often referred to as relative noise. If $\eta = 1/2$, the variance of ϵ scales linearly with ζ and in this case the observation error is called Poisson noise.

Let $\boldsymbol{I}_{\boldsymbol{\theta}} = (I(t_1, \boldsymbol{\theta}), \ldots, I(t_n, \boldsymbol{\theta}))^T$ be a vector of the number of infected people estimated by the SIR model with fixed $\boldsymbol{\theta}$. We denote $\boldsymbol{Y} = (y_1, \ldots, y_n)^T$,; $\boldsymbol{\delta}^{\boldsymbol{\theta}} = (\delta^{\boldsymbol{\theta}}(t_1), \ldots, \delta^{\boldsymbol{\theta}}(t_n))^T$, where $\delta^{\boldsymbol{\theta}}(t_i) = \zeta(t_i) - I(t_i, \boldsymbol{\theta})$ and $\boldsymbol{E} = (\epsilon_1, \ldots, \epsilon_n)^T$. The full model is presented as

$$\begin{aligned} \boldsymbol{Y} &= \boldsymbol{I}_{\boldsymbol{\theta}} + \boldsymbol{\delta} + \boldsymbol{E}, \\ \boldsymbol{\delta}^{\boldsymbol{\theta}} &\sim MN(0, \tau^2 \mathbf{K}_\phi), \\ \boldsymbol{E} &\sim MN(0, \sigma^2 \mathbf{V}_\eta), \end{aligned} \tag{2.5}$$

where $\mathbf{K}_\phi = [K_\phi(t_i, t_j)]_{ij}, i, j = 1, \ldots, n$ is the correlation matrix of $\delta^{\boldsymbol{\theta}}(t_1), \ldots, \delta^{\boldsymbol{\theta}}(t_n)$ and $\tau^2 < \infty$ is the variance of $\delta^{\boldsymbol{\theta}}(t_i)$ and $\mathbf{V}_\eta$ is a $n \times n$ diagonal matrix with (i, i) element $\zeta_i^{2\eta}, i = 1, \ldots, n$. We can easily obtain that $\boldsymbol{Y}$ follows a multivariate normal distribution. The expectation of Y is:

$$\mathbb{E}(\boldsymbol{Y}) = \boldsymbol{I}_{\boldsymbol{\theta}}, \tag{2.6}$$

and the variance of $\boldsymbol{Y}$ is

$$\mathbb{V}(\boldsymbol{Y}) = \tau^2 \mathbf{K}_\phi + \sigma^2 \mathbf{V}_\eta. \tag{2.7}$$

Although a maximum likelihood estimator based on the distribution of $\boldsymbol{Y}$ is a natural choice to estimate $\boldsymbol{\theta}$, Tuo *et al.* [19] prove the unsatisfactory convergence properties of this maximum likelihood estimator. Hence, we suggest a WLS estimator for $\boldsymbol{\theta}$, which is defined as the solution to the minimization problem:

$$\hat{\boldsymbol{\theta}}_{WLS} = \text{argmin}_{\boldsymbol{\theta}} \frac{1}{n} \sum_{i=1}^{n} \frac{(y_i - I(t_i, \boldsymbol{\theta}))^2}{\zeta_i^{2\eta}}. \tag{2.8}$$

Given $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}_{WLS}$, the posterior distribution of $\delta^{\hat{\boldsymbol{\theta}}_{WLS}}(t)$ at an unobserved input t has the closed form expression [20]:

$$\delta^{\hat{\boldsymbol{\theta}}_{WLS}}(t)|\boldsymbol{Y}, \boldsymbol{I}_{\hat{\boldsymbol{\theta}}_{WLS}} \sim N\left(\hat{\delta}_n(t), \tau^2 s_n^2(t)\right). \tag{2.9}$$

The posterior mean and variance of $\delta^{\hat{\boldsymbol{\theta}}_{WLS}}(t)$ are:

$$\begin{aligned} \hat{\delta}_n(t) &= \boldsymbol{K}_\phi^T(t)\boldsymbol{\Sigma}^{-1}\left[\boldsymbol{Y} - \boldsymbol{I}_{\hat{\boldsymbol{\theta}}_{WLS}}\right], \\ \tau^2 s_n^2(t) &= \tau^2\left\{1 - \boldsymbol{K}_\phi^T(t)\boldsymbol{\Sigma}^{-1}\boldsymbol{K}_\phi(t) + \boldsymbol{U}^T(t)(\mathbf{1}^T\boldsymbol{\Sigma}^{-1}\mathbf{1})^{-1}\boldsymbol{U}(t)\right\}, \end{aligned} \tag{2.10}$$

where $\boldsymbol{K}_\phi(t)$ is the correlation vector between $\delta(t)$ and $\boldsymbol{\delta}$, $\mathbf{1}$ is a $n \times 1$ vector with entries of 1, $\boldsymbol{\Sigma} = \mathbf{K}_\phi + \sigma^2/\tau^2\mathbf{V}_\eta$ and $\boldsymbol{U}(t) = 1 - \boldsymbol{K}_\phi^T(t)\boldsymbol{\Sigma}^{-1}\mathbf{1}$. The posterior mean $\hat{\delta}_n(t)$ is considered as the best linear unbiased estimator of $\delta^{\hat{\boldsymbol{\theta}}_{WLS}}(t)$ and $\tau^2 s_n^2(t)$ the corresponding variance [20]. It should be noted that $\delta^{\hat{\boldsymbol{\theta}}_{WLS}}(t)$ is different from δ^*, which is defined in Eq. (2.2). δ^* was defined relative to the optimal set of calibration parameters, rather than the estimated value $\hat{\boldsymbol{\theta}}_{WLS}$. Since $\boldsymbol{\theta}^*$ is unknown, $\hat{\delta}^*$ can be approximated by $\hat{\delta}_n$.

When we predict the underlying number of infected people at time t, two predictors can be used [21].

- Predicting the true epidemiological system by utilizing the calibrated SIR model

$$\hat{\zeta}_I(t) = I(t, \hat{\boldsymbol{\theta}}_{WLS}). \tag{2.11}$$

- A discrepancy-reduced predictor for the true epidemiological system

$$\hat{\zeta}_R(t) = I(t, \hat{\boldsymbol{\theta}}_{WLS}) + \hat{\delta}_n(t). \tag{2.12}$$

The predictor in Eq. (2.11) inputs the estimation $\hat{\boldsymbol{\theta}}_{WLS}$ into the SIR model and runs the model at time t to obtain a prediction. This predictor is called *model predictor*, and its accuracy is determined by the variance of $\hat{\boldsymbol{\theta}}_{WLS}$. The second predictor in Eq. (2.12) is corrected by an estimation of the model discrepancy between the true epidemiological system and the calibrated SIR. This predictor is called *discrepancy-corrected predictor*. Bayarri *et al.* [21] highlights that since an estimator of the model discrepancy is available, the prediction accuracy of the discrepancy-corrected predictor $\hat{\zeta}_R(t)$ is much higher than the pure model predictor $\hat{\zeta}_I(t)$.

2.2. Computational difficulties in calibration and prediction

Evaluating $\hat{\boldsymbol{\theta}}_{WLS}$ and $\hat{\delta}_n$ utilizing Eqs. (2.8) and (2.10) has two major practical difficulties. The first one is that the true value ζ_i is unknown, and thus Eq. (2.8) cannot be calculated. To address this problem, assuming that the observation error ϵ_i is much smaller than ζ_i, we replace ζ_i with y_i. The second problem is that the hyper-parameters η, σ^2, ϕ, and τ^2 are also unknown. Hence, we estimate them based on maximum likelihood estimations.

By replacing ζ_i with y_i, we define $\tilde{\mathbf{V}}_\eta$ as an $n\times n$ diagonal matrix with (i,i) element $y_i^{2\beta}$. Furthermore, we denote $\tilde{\boldsymbol{\Sigma}} = \mathbf{K}_\phi + g\tilde{\mathbf{V}}_\eta$ with $g = \sigma^2/\tau^2$. Given $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$, the parameters β, ϕ, g andτ^2 can be estimated by maximizing the likelihood function:

$$L(\eta,\phi,g,\tau^2) = \left(\frac{1}{2\pi\tau^2}\right)^{-\frac{n}{2}} |\tilde{\boldsymbol{\Sigma}}|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2\tau^2}(\boldsymbol{Y}-\boldsymbol{I}_{\hat{\theta}})^T\tilde{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{Y}-\boldsymbol{I}_{\hat{\theta}})\right\},$$

where $|A|$ is the determinant of the matrix A. With some direct calculations [18], the MLE of (η,ϕ,g) is the maximizer of:

$$L(\eta,\phi,g) \propto -\frac{n}{2}\log\left\{(\boldsymbol{Y}-\boldsymbol{I}_{\hat{\theta}})^T\tilde{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{Y}-\boldsymbol{I}_{\hat{\theta}})\right\} - \frac{1}{2}\log|\tilde{\boldsymbol{\Sigma}}|, \tag{2.13}$$

and the MLE of τ^2 given η,ϕ and g is:

$$\hat{\tau}^2 = \frac{(\boldsymbol{Y}-\boldsymbol{I}_{\hat{\theta}})^T\tilde{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{Y}-\boldsymbol{I}_{\hat{\theta}})}{n}. \tag{2.14}$$

We propose estimating the calibration parameters $\boldsymbol{\theta}$ and the hyper-parameters (η,ϕ,g,τ^2) in two stages. In the first stage, let $\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}_{OLS}$ which is defined as:

$$\hat{\boldsymbol{\theta}}_{OLS} = \text{argmin}_\theta \frac{1}{n}\sum_{i=1}^{n}(y_i - I(t_i,\boldsymbol{\theta}))^2,$$

and we estimate the hyper-parameters $\hat{\eta}_0$, $\hat{\phi}_0$, $\hat{g}_0$ and $\hat{\tau}_0^2$. In the second stage, we evaluate $\hat{\boldsymbol{\theta}}_{WLS}$ by using Eq. (2.8), where η is replaced by its estimation $\hat{\eta}_0$. Let $\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}_{WLS}$, we re-estimate the MLE of β, σ^2, ϕ and τ^2, written as $\hat{\eta}^*$, $\hat{\phi}^*$, $\hat{g}^*$ and $\hat{\tau}^{*2}$.

Algorithm 1 provides the details of the calibration and prediction for the SIR model.

2.3. Uncertainty quantification by using wild bootstrap

This section considers the uncertainty quantification for $\hat{\zeta}_I$ and $\hat{\zeta}_R$. Since the SIR model is deterministic the uncertainties of $\hat{\zeta}_I$ and $\hat{\zeta}_R$ depend on the randomness of $\hat{\boldsymbol{\theta}}_{WLS}$ and $\hat{\delta}_n$. Bootstrap methods have been applied successfully to quantify the model uncertainty when the sample size is finite [22, 23]. Among various bootstrap methods, the wild bootstrap [24] effectively deals with the weighted least squares inference. Thus, wild bootstrap samples for $\hat{\zeta}_I$ and $\hat{\zeta}_R$ can be generated utilizing Algorithm 2.

It should be noted that our estimation procedure involves non-parametric regression of the model discrepancy function. To keep the approach simple, we adopt Wong *et al.* [23] stating that the confidence regions for $\hat{\delta}_n$ can also be evaluated from the posterior distribution of Eq. (2.10) and thus the confidence regions for $\hat{\zeta}_R$ can be evaluated by the sum of the confidence regions for $\hat{\zeta}_R$ and the confidence regions for $\hat{\delta}_n$.

Algorithm 1 Calibration and prediction for the SIR model

Step 1. Evaluate $\hat{\boldsymbol{\theta}}_{OLS} = \text{argmin}_{\boldsymbol{\theta}} \frac{1}{n} \sum_{i=1}^{n} (y_i - I(t_i, \boldsymbol{\theta}))^2$ and let $\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}_{OLS}$;
Step 2. Evaluate $\hat{\eta}_0$, $\hat{\phi}_0$, $\hat{g}_0$ by maximizing (2.13); Plug $\hat{\eta}_0$, $\hat{\phi}_0$, $\hat{g}_0$ into (2.14) to get $\hat{\tau}_0^2$;
Step 3. Minimize the loss function $l(\boldsymbol{\theta})$ (2.15) to get $\hat{\boldsymbol{\theta}}_{WLS}$:

$$l(\boldsymbol{\theta}) = \frac{1}{n} \sum_{i=1}^{n} \frac{(y_i - I(t_i, \boldsymbol{\theta}))^2}{y_i^{2\hat{\eta}_0}}; \tag{2.15}$$

Step 4. Let $\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}_{WLS}$, re-estimate the hyper-parameters by using the method in step 2. Write the estimations of hyper-parameters as $\hat{\eta}^*$, $\hat{\phi}^*$, $\hat{g}^*$ and $\hat{\tau}^{*2}$;
Step 5. Plug $\hat{\boldsymbol{\theta}}_{WLS}$, $\hat{\eta}^*$, $\hat{\phi}^*$, $\hat{g}^*$ and $\hat{\tau}^{*2}$ into (2.10) to obtain $\hat{\delta}_n(t)$;
Step 6. Compute $\hat{\zeta}_I$ and $\hat{\zeta}_R$ respectively.

Algorithm 2 Uncertainty quantification

Evaluate $\hat{\boldsymbol{\theta}}_{WLS}$ and $\hat{\delta}_n(t)$ by using Algorithm 1;
for iteration $k = 1, \cdots, N$ **do**
Evaluate the errors $\{\hat{\epsilon}_1, \ldots, \hat{\epsilon}_n\}$:

$$\hat{\epsilon}_i = y_i - I(t_i, \hat{\boldsymbol{\theta}}_{WLS}) - \hat{\delta}_n(t_i),$$

Next, we generate the bootstrap sample data set:

$$y_i' = I(t_i, \hat{\boldsymbol{\theta}}_{WLS}) + \hat{\delta}_n(t_i) + \hat{\epsilon}_i v_i, i = 1, \ldots, n \tag{2.16}$$

where

$$v_i = \begin{cases} -(\sqrt{5} - 1)/2 & \text{with probability } (\sqrt{5} + 1)/(2\sqrt{5}), \\ (\sqrt{5} + 1)/2 & \text{with probability } (\sqrt{5} - 1)/(2\sqrt{5}). \end{cases} \tag{2.17}$$

Evaluate $\hat{\boldsymbol{\theta}}_{WLS}^{(k)}$, $\hat{\delta}_n^{(k)}(t)$, $\hat{\zeta}_I^{(k)}$ and $\hat{\zeta}_R^{(k)}$ by using Algorithm 1, where the vector of the real observations $\boldsymbol{Y}$ is replaced by $\boldsymbol{Y'} = (y_1', \ldots, y_n')^T$;
end for
Print the confidence regions for $\hat{\zeta}_I$ and $\hat{\zeta}_R$.

3. Results

This section challenges the performance of $\hat{\zeta}_I$ and $\hat{\zeta}_R$ against the model predictors where $\boldsymbol{\theta}$ is estimated by the ordinary least squares (OLS) method [6], and the weighted least squares (WLS) method [8]. These two model predictors ignore the model discrepancy and assume that the actual epidemiological system can be precisely described by the SIR model together with the optimal parameters.

The ordinary least squares estimation of $\boldsymbol{\theta}$ is defined as:

$$\hat{\boldsymbol{\theta}}^E_{OLS} = \text{argmin}_{\boldsymbol{\theta}} \sum_{i=1}^{n} (y_i - I(t_i, \boldsymbol{\theta}))^2. \tag{3.1}$$

The superscript E means that the SIR model is exact with $\boldsymbol{\theta} = \boldsymbol{\theta}^*$. The weighted least squares estimator of $\boldsymbol{\theta}$ is defined as:

$$\hat{\boldsymbol{\theta}}^E_{WLS} = \text{argmin}_{\boldsymbol{\theta}} \sum_{i=1}^{n} w_i (y_i - I(t_i, \boldsymbol{\theta}))^2, \tag{3.2}$$

where $\{w_i = 1/var(\epsilon_i)\}_{i=1,\dots,n}$ is a set of weights and ϵ_i is the observation error for y_i. Current works including [7, 8] assume that $w_i \propto 1/I(t_i, \boldsymbol{\theta})^{2\xi}$ and ξ is chosen from $\{1/2, 1\}$. Given that the observation errors is independent from the calibration parameters, assuming the variance of ϵ_i to be a function of $\boldsymbol{\theta}$ is unrealistic. Thus, this work assumes $w_i \propto 1/y_i^{2\xi}$, and ξ is $1/2$ or 1. We write the WLS estimator of $\boldsymbol{\theta}$ as $\hat{\boldsymbol{\theta}}^E_{WLS\,1}$ when $\xi = 1/2$, and as $\hat{\boldsymbol{\theta}}^E_{WLS\,2}$ when $\xi = 1$.

To fulfill the calibration and prediction for the SIR model, Section 3.1 provides two reported infected data. A sensitivity analysis for the SIR model is conducted in Section 3.2, while Section 3.3 shows the comparison results of four different predictors.

3.1. Data

Influenza epidemic data. Table 1 presents data for an influenza outbreak at a boys' boarding school in England [25]. The total population is $N = 763$ boys.

Table 1. Influenza epidemic data from a boys' boarding school.

Day	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Confirmed	3	6	25	73	222	294	258	237	191	125	69	27	11	4

The data given above is the number of infected students confined to bed throughout the epidemic, which lasted a fortnight. Beyond being confined to bed, only ten of the ill boys received antibiotics, i.e., the treatment the boys received was minimal.

COVID-19 data. Since Jan. 2, 2021, China's Hebei province has suffered from a new COVID-19 wave, and the related health authorities announced a new lockdown for 11 million people in the Hebei province. The epidemiologic data illustrated in Figure 1 are from the official open-source repository operated by the Health Commission of Hebei Province. It should be noted that since the lockdown measures, the Hebei outbreak is considered a closed system.

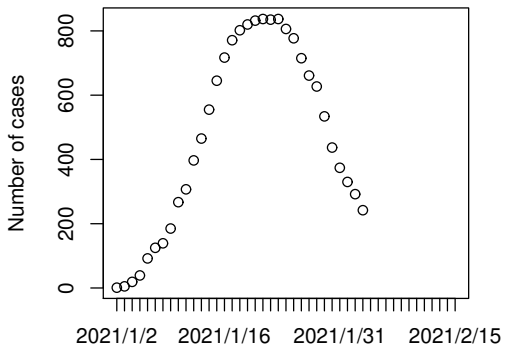

Figure 1. Number of confirmed cases in Hebei, China.

3.2. Sensitivity analysis for the SIR model

With the help of the influenza epidemic data, we explore the relationship between the number of infected people estimated from the SIR model and the calibration parameter values. One of the simplest and most common approaches is the changing one-factor-at-a-time (OAT) [26] scheme, in which we change the value of only one calibration parameter and the remaining ones are preserved 3.1.

Let $S_0 = 760$, $I_0 = 3$ and $R_0 = 0$. Set $\hat{\boldsymbol{\theta}}_B = (0.538, 0.065, 1.392, 0.429)$ be the baseline values for $\boldsymbol{\theta}$, which is randomly generated by the maximin latin hypercube sampling method [20]. Four SIR models are executed with different calibration parameter setups, as listed in Table 2.

Table 2. Values of the calibration parameters ($k = 1, 2, \ldots, 6$).

Model	β_0	β_1	μ	γ
model (a)	$\frac{2(k-1)}{5}$	0.065	1.392	0.429
model (b)	0.538	$\frac{2(k-1)}{5}$	1.392	0.429
model (c)	0.538	0.065	$\frac{k-1}{5}$	0.429
model (d)	0.538	0.065	1.392	$\frac{2(k-1)}{5}$

The estimated number of infected boys utilizing each SIR model is illustrated in Figure 2 highlighting that the calibration parameters have a significant impact on the number of infected people. Specifically, (a) For $\beta_0 = 0$, the number of infected people increases slowly, reaches the peak after about 9 days, and then decreases slowly. For a higher value of β_0, the time to reach the peak of infections is shorter, and the number of infected people decreases to a smaller value more rapidly. (b) The relationship between β_1 and $I(t)$ is similar to the relationship of β_0 and $I(t)$. (c) A smaller value of μ leads to a shorter time to reach the peak of infections, a faster decrease from the peak, and a smaller number of infected people at the end of the simulation period. For $\mu = 0.5$, the number of infected people on the 12th day reaches its peak, which is less than 100. (d) For $\gamma = 0$, infected people increase rapidly and reach 763 after about 6 days. Hence, if the recovery rate is 0, then all boys will ultimately

be infected. For a higher value of γ, the peak of infections is smaller, but the time to reach the peak is almost the same, while the number of infected people at the end of the simulation period is smaller.

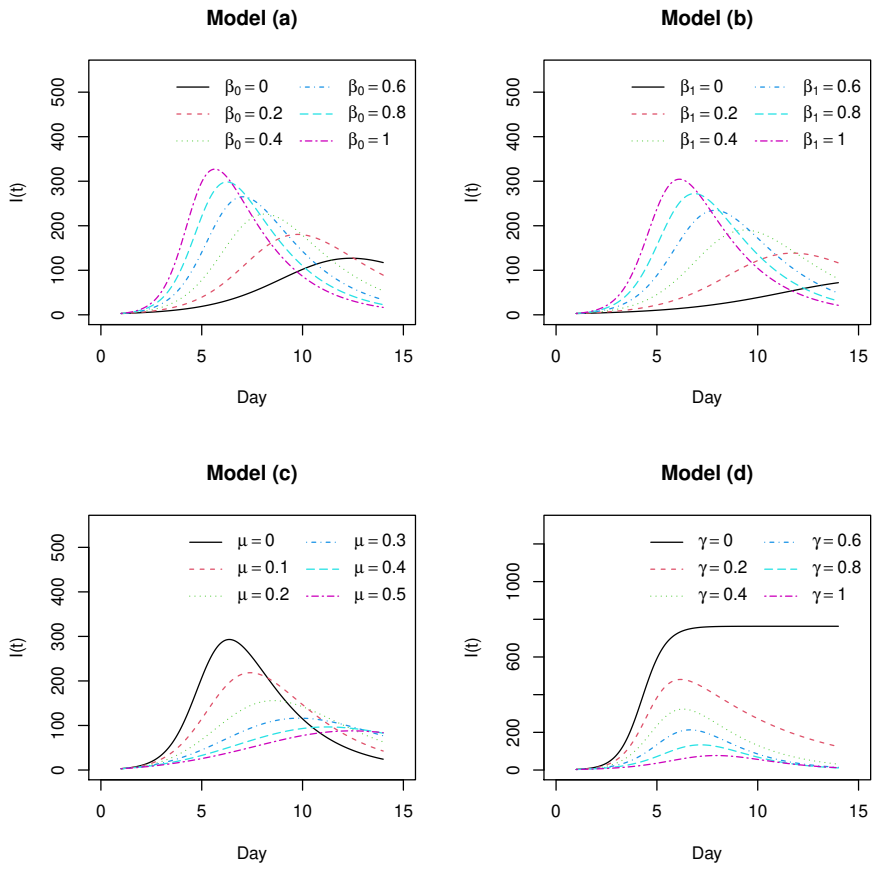

Figure 2. Infected data estimated by the SIR model with different calibration parameter setups.

It is worth noticing that, except for the OAT method, there are many other commonly used global sensitivity analysis methods, such as the Sobol' method [27]. This method decomposes the variance of the model output into fractions that can be attributed to each input parameter. Then it measures the global sensitivity of each calibration parameter by the percentage of the corresponding fraction. With the help of *sobolshap_knn* function in the R package *sensitivity* [28], we obtain the first-order Sobol' indices of the SIR model, which is shown in Figure 3.

From Figure 3, we can see that at the beginning of the epidemic, the first-order effects of all these four calibration parameters are not negligible. As the epidemic spreads, the first-order effect of γ becomes higher. If the aim is to improve the prediction accuracy of the SIR model during the whole epidemic, these results suggest adjusting the values of all the calibration parameters. And if the goal is to improve the model accuracy at the end of the epidemic, then it is more important to focus the efforts on better quantifying the recovery rate γ. In this work, we focus on the prediction accuracy of the SIR model throughout the epidemic. As a result, all the calibration parameters have significant impacts on the number of infected people. It coincides with the result obtained by the OAT method.

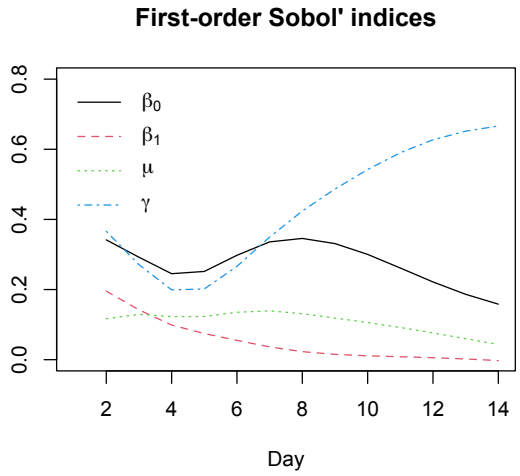

Figure 3. First-order Sobol' indices of the SIR model .

3.3. Calibration and prediction results

This section compares the prediction performance of five different predictors: three model predictors $I(\cdot, \hat{\boldsymbol{\theta}})$, with $\hat{\boldsymbol{\theta}} = \{\hat{\boldsymbol{\theta}}^E_{OLS}, \hat{\boldsymbol{\theta}}^E_{WLS1}, \hat{\boldsymbol{\theta}}^E_{WLS2}\}$, the model predictors $\hat{\zeta}_I = I(\cdot, \hat{\boldsymbol{\theta}}_{WLS})$ and the discrepancy-corrected predictor $\hat{\zeta}_R = \hat{\zeta}_I + \hat{\delta}_n$. *Absolute errors* and *Relative errors* are used to measure the discrepancy between the reported number of infected case and the predictions given by the different predictors. The absolute error (AE) and relative error (RE) are defined as:

$$\begin{aligned} \mathrm{AE}_i &= |y_i - \hat{y}_i|, \\ \mathrm{RE}_i &= \left|\frac{y_i - \hat{y}_i}{y_i}\right|, i = 1, \ldots, n, \end{aligned} \tag{3.3}$$

where $\hat{y}_i$ is the predicted number of infected cases at time t_i by each method. It is evident that the absolute and the relative error values describe the accuracy of a given prediction method. We display the comparison results in Table 4 and Table 6 using the Influenza epidemic data and COVID-19 data, respectively.

Influenza epidemic data. We first apply the NEWUOA algorithm [29] to evaluate $\hat{\boldsymbol{\theta}}^E_{OLS}$ and $\hat{\boldsymbol{\theta}}^E_{WLS}$. Specifically, we randomly generate 50 sets of pseudo-random numbers as the starting points utilizing the maximin latin hypercube sampling method, utilize the *NLOPT* package in R to perform local searches on all starting points to obtain the ordinary least squares estimation of $\boldsymbol{\theta}$, and use the *NLOPT* package to estimate $\hat{\boldsymbol{\theta}}^E_{WLS}$ where the initial value is $\hat{\boldsymbol{\theta}}^E_{OLS}$. Finally Algorithm 1 is used to obtain $\hat{\boldsymbol{\theta}}_{WLS}$, $\hat{\zeta}_I$ and $\hat{\zeta}_R$. Table 3 displays the estimated calibration parameters.

The prediction results of the five different predictors are illustrated in Figure 4 highlight that the proposed predictor $\hat{\zeta}_R$ outperforms the competitor predictors at most observation time points. Moreover, the performance of $I(\cdot, \hat{\boldsymbol{\theta}}^E_{OLS})$ and $\hat{\zeta}_R$ are similar at $(t_1, \ldots, t_{10})$, but $\hat{\zeta}_R$ attains more accurate predictions at $(t_{11}, \ldots, t_{14})$. The predictions at $(t_1, \ldots, t_4)$ and (t_{13}, t_{14}) given by $I(\cdot, \hat{\boldsymbol{\theta}}^E_{WLS2})$ are closest to the reported number of infected boys. However, at $(t_5, \ldots, t_{11})$ the performance of $I(\cdot, \hat{\boldsymbol{\theta}}^E_{WLS2})$ is the poorest, underestimating the infected population. The performance of $I(\cdot, \hat{\boldsymbol{\theta}}^E_{WLS1})$ and $\hat{\zeta}_I$ is similar throughout the period.

Table 3. Estimated calibration parameters.

Estimator	β_0	β_1	μ	γ
$\hat{\boldsymbol{\theta}}^E_{OLS}$	0.066	1.626	0	0.448
$\hat{\boldsymbol{\theta}}^E_{WLS1}$	0.115	1.653	0	0.528
$\hat{\boldsymbol{\theta}}^E_{WLS2}$	0.294	1.587	0	0.720
$\hat{\boldsymbol{\theta}}_{WLS}$	0.105	1.649	0	0.512

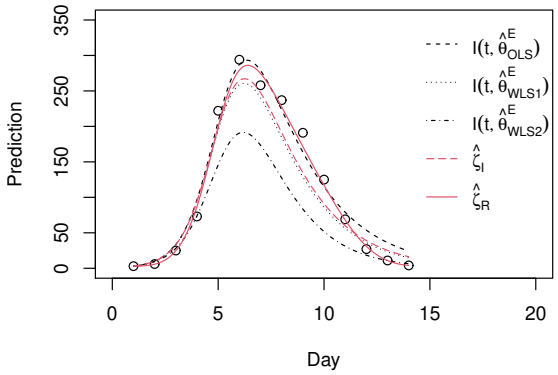

Figure 4. Prediction results of various method against the reported number of infected boys (circles). The predictors are the OLS predictor $I(\cdot, \hat{\boldsymbol{\theta}}^E_{OLS})$ (dashed line), WLS predictor $I(\cdot, \hat{\boldsymbol{\theta}}^E_{WLS1})$ with $\xi = 1/2$ (dotted line), WLS predictor $I(\cdot, \hat{\boldsymbol{\theta}}^E_{WLS2})$ with $\xi = 1$ (dotdash line), and the two proposed predictors $\hat{\zeta}_I$ (longdash line) and $\hat{\zeta}_R$ (solid line) .

To further investigate the effectiveness of the proposed method, Table 4 presents the median absolute error (MAE) and median relative error (MRE) of the evaluated predictors.

Table 4. Median absolute error (MAE) and median relative error (MRE).

Error	$I(\cdot, \hat{\boldsymbol{\theta}}^E_{OLS})$	$I(\cdot, \hat{\boldsymbol{\theta}}^E_{WLS1})$	$I(\cdot, \hat{\boldsymbol{\theta}}^E_{WLS2})$	$\hat{\zeta}_I$	$\hat{\zeta}_R$
MAE	14.249	15.196	23.232	13.271	6.002
MRE	0.144	0.299	0.348	0.283	0.046

From Table 4, we observe that the proposed predictor $\hat{\zeta}_R$ outperforms the competitor methods: the MAE of $\hat{\zeta}_R$ is reduced by at least 57.88% and the MRE of $\hat{\zeta}_R$ is reduced by at least 68.06%. Compared with the OLS predictor, the suggested predictor $\hat{\zeta}_I$ has a smaller MAE but a larger MRE. The performance of $I(\cdot, \hat{\boldsymbol{\theta}}^E_{WLS1})$ is slightly inferior to $\hat{\zeta}_I$, with MAE and MRE of $I(\cdot, \hat{\boldsymbol{\theta}}^E_{WLS2})$ being the largest indicating the poor prediction performance of this method. Due to the poor performance of $I(\cdot, \hat{\boldsymbol{\theta}}^E_{WLS2})$, we neglect this method in the following uncertainty quantification trials.

Next, we use the wild bootstrap method (Algorithm 2) to obtain the 95% prediction intervals of the remaining four different predictions, as depicted in Figure 5.

From the upper- left subfigure we observe that only the reported numbers at $(t_1, t_5, t_6, t_8, t_{10})$ lie in the

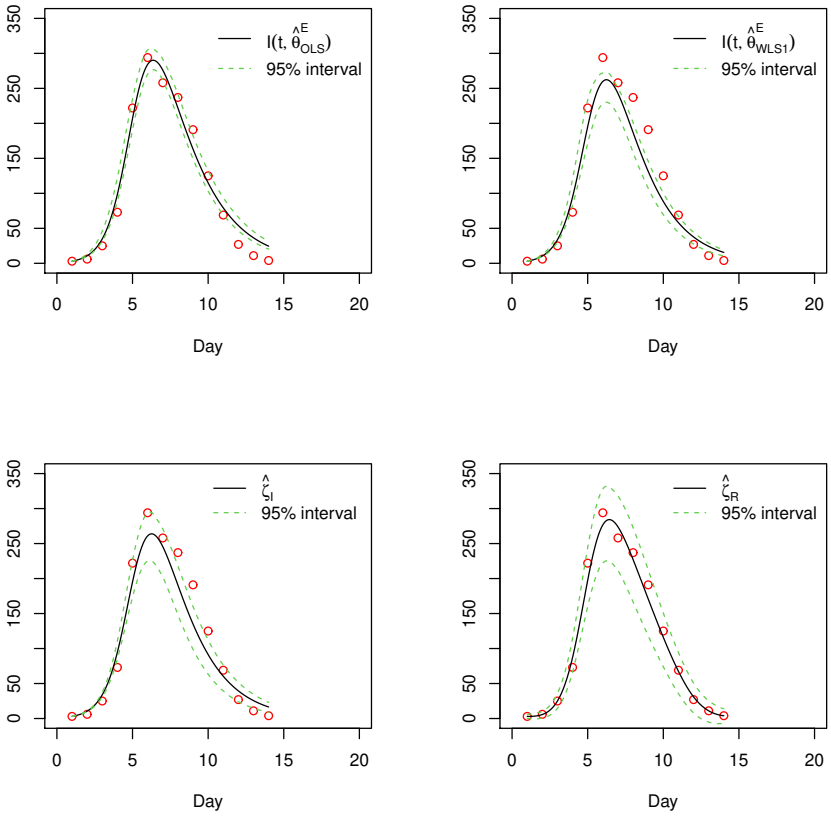

Figure 5. 95% prediction intervals of four different predictors, where circles in the subfigures represent the reported number of infected boys.

95% prediction intervals of $I(\cdot, \hat{\boldsymbol{\theta}}^E_{OLS})$. The lower endpoints of the 95% prediction interval at (t_{12}, t_{13}, t_{14}) are larger than the corresponding reported numbers. The predictions given by $I(\cdot, \hat{\boldsymbol{\theta}}^E_{WLS\,1})$ are closer to the reported numbers at (t_{12}, t_{13}, t_{14}), but present an inferior accuracy at $(t_6, \ldots, t_{10})$. By comparing the the 95% prediction interval of $\hat{\zeta}_I$ and $I(\cdot, \hat{\boldsymbol{\theta}}^E_{WLS\,1})$, we observe that although the prediction accuracy of $\hat{\zeta}_I$ and $I(\cdot, \hat{\boldsymbol{\theta}}^E_{WLS\,1})$ are similar, the performance of $\hat{\zeta}_I$ is indeed better than the performance of $I(\cdot, \hat{\boldsymbol{\theta}}^E_{WLS\,1})$. Finally, the bottom-right subfigure shows that all the reported data (nearly) lie in the 95% prediction interval of $\hat{\zeta}_R$.

As our methodology in Section 2.1 suggests, we also display the plots (Figure 6) estimating the model's discrepancy and residuals for the predictors $\hat{\zeta}_R$. This is important to verify that the SIR model is inexact and that the constant variance assumption about the observation errors is valid.

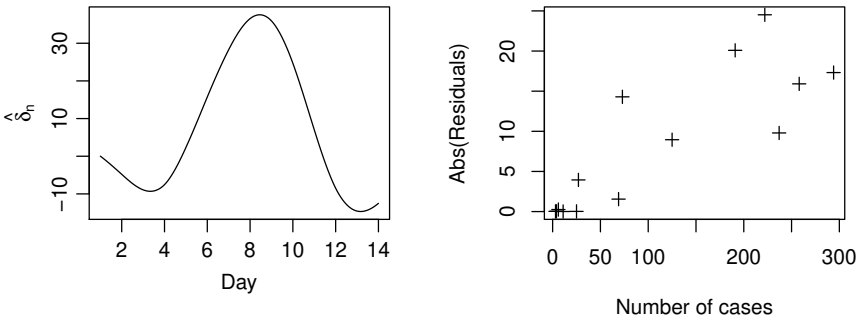

Figure 6. Left: model discrepancy function estimation. Right: absolute values of the residuals for $\hat{\zeta}_R$ vs. the reported number of infected boys.

The left subfigure depicts the estimation of the model discrepancy function, which is a smooth function of time. Through some calculations, we obtain $\|\hat{\delta}_n\|_{L_2} = \sqrt{\int_0^{14} \hat{\delta}_n^2(t)dt} = 19.282 > 0$ and thus model discrepancy between the SIR model and the true epidemiological system cannot be ignored. The right subfigure illustrates the relationship between the absolute values of the residuals for $\hat{\zeta}_R$ and the reported number of infected boys. This subfigure indicates that as the infected cases increase, the residuals for $\hat{\zeta}_R$ increase. The correlation coefficient between them is 0.835, indicating a strong relationship between the observation errors and the number of cases.

COVID-19 data. The same method is used to $\hat{\boldsymbol{\theta}}^E_{OLS}$, $\hat{\boldsymbol{\theta}}^E_{WLS}$ and $\hat{\boldsymbol{\theta}}_{WLS}$, with Table 5 displaying the corresponding calibration parameter estimations.

Table 5. Calibration parameter Estimations.

Estimator	β_0	β_1	μ	γ
$\hat{\boldsymbol{\theta}}^E_{OLS}$	0.006	1.278	0.103	0.157
$\hat{\boldsymbol{\theta}}^E_{WLS\,1}$	0.0001	1.391	0.102	0.196
$\hat{\boldsymbol{\theta}}^E_{WLS\,2}$	0.023	1.415	0.111	0.195
$\hat{\boldsymbol{\theta}}_{WLS}$	0.0009	1.319	0.099	0.184

Since $\beta_1 > 0$ and $\mu > 0$, Figure 7 describes the decreasing of $\beta(t)$ with t, where the calibration parameter is set to be $\hat{\boldsymbol{\theta}}_{WLS}$. We observe that the initial transmission rate is approximately 1.320, and due to the lockdown policies the transmission rate shall gradually decrease to zero. Table 6 presents the MAE and MRE values of different predictors, proving that the developed predictor $\hat{\zeta}_R$ outperforms the competitor methods. Compared to $I(\cdot, \hat{\boldsymbol{\theta}}^E_{WLS\,1})$, $I(\cdot, \hat{\boldsymbol{\theta}}^E_{WLS\,2})$ affords a smaller MAE but a larger MRE. Due to the poor performance of $I(\cdot, \hat{\boldsymbol{\theta}}^E_{OLS})$, we neglect this method in the following uncertainty quantification trials.

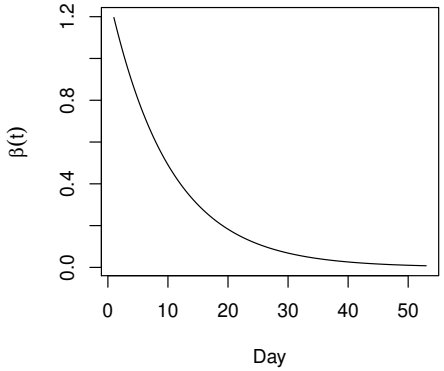

Figure 7. $\beta(t)$ with $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}_{WLS}$.

Table 6. Median absolute error and median relative error.

Error	$I(\cdot, \hat{\boldsymbol{\theta}}^E_{OLS})$	$I(\cdot, \hat{\boldsymbol{\theta}}^E_{WLS\,1})$	$I(\cdot, \hat{\boldsymbol{\theta}}^E_{WLS\,2})$	$\hat{\zeta}_I$	$\hat{\zeta}_R$
MAE	121.049	73.809	97.087	96.568	30.571
MRE	69.173	31.120	17.394	44.631	8.705

The prediction results of the competitor predictors are depicted in Figure 8, where the proposed predictor $\hat{\zeta}_R$ outperforms the other predictors at most of the observation time points. The MAE of $\hat{\zeta}_R$ is reduced by at least 58.58% and the MRE of $\hat{\zeta}_R$ is reduced by at least 49.95%. The predictor $I(\cdot, \hat{\boldsymbol{\theta}}^E_{WLS\,1})$ underestimates the number of cases between $15 < t < 22$ and overestimates the number of cases for $3 < t < 8$. The predictor $I(\cdot, \hat{\boldsymbol{\theta}}^E_{WLS\,2})$ overestimate the number of cases between $21 < t < 28$, while the two WLS predictors $I(\cdot, \hat{\boldsymbol{\theta}}^E_{WLS\,1})$ and $I(\cdot, \hat{\boldsymbol{\theta}}^E_{WLS\,2})$ perform similar at $t \leq 13$ and $t \geq 29$. The performance of $\hat{\zeta}_I$ is much better than $I(\cdot, \hat{\boldsymbol{\theta}}^E_{WLS\,1})$ and $I(\cdot, \hat{\boldsymbol{\theta}}^E_{WLS\,2})$. By comparing the the 95% prediction interval of $\hat{\zeta}_I$ and $\hat{\zeta}_R$ we conclude that although the prediction accuracy of $\hat{\zeta}_R$ is higher, the variance of $\hat{\zeta}_R$ is larger than the variance of $\hat{\zeta}_I$.

Figure 9 describes the model discrepancy estimation and the residuals for predictor $\hat{\zeta}_R$. Through calculations, we obtain $\|\hat{\delta}_n\|_{L_2} = \sqrt{\int_0^{14} \hat{\delta}_n^2(t)dt} = 35.924 > 0$, verifying that the model discrepancy cannot be ignored. The correlation coefficient between the absolute values of the residuals for $\hat{\zeta}_R$ and the reported number of infected cases is 0.519, i.e., the constant variance assumption about the observation errors is valid.

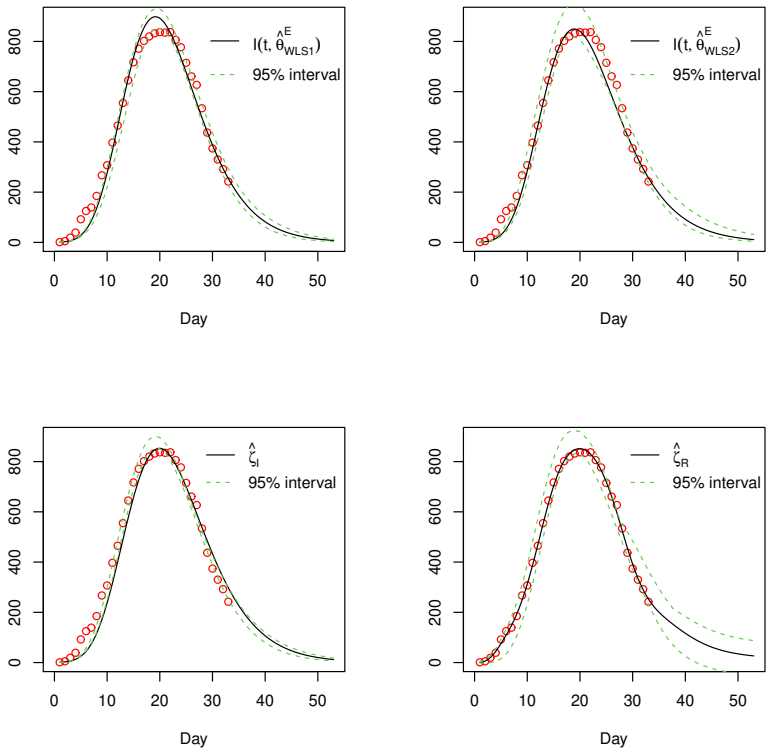

Figure 8. 95% prediction intervals of four different predictors: WLS predictor $I(\cdot, \hat{\boldsymbol{\theta}}^E_{WLS1})$, where $\xi = 1/2$; WLS predictor $I(\cdot, \hat{\boldsymbol{\theta}}^E_{WLS2})$ where $\xi = 1$; two proposed predictors $\hat{\zeta}_I$ and $\hat{\zeta}_R$.

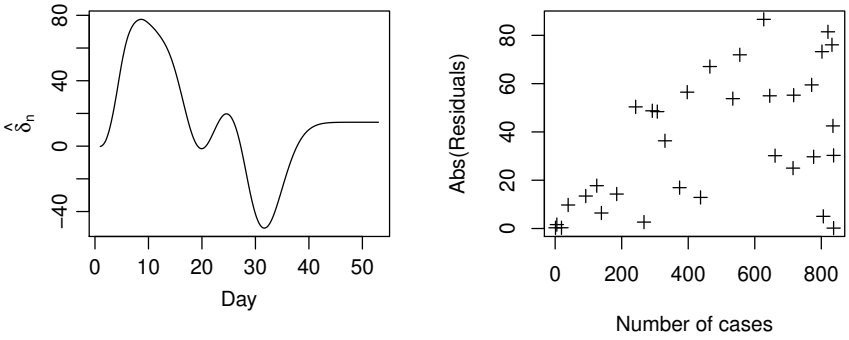

Figure 9. Left: estimation of the model discrepancy function. Right: absolute values of the residuals for $\hat{\zeta}_R$ vs. the reported number of infected cases.

4. Conclusion and future work

In this work, we proposed a new WLS estimator for the calibration parameters by modeling the discrepancy between the SIR model and the actual epidemiological system through a Gaussian Process model. The proposed estimator takes full consideration of the heteroscedastic observation errors with one-time reported infected data. We proposed two different predictors for the epidemiological system. One is the calibrated SIR model, and the other one is the discrepancy-corrected predictor. A wild bootstrap method is adopted to quantify the predictors' uncertainties. We used two numerical studies to assess the performance of the proposed predictors. The numerical results show that the prediction accuracy of the discrepancy-corrected predictor is improved by at least 49.95%. The 95% prediction intervals show that the calibrated SIR model enjoys a smaller variance than the discrepancy-corrected predictor. These results demonstrate that the proposed method improves the prediction accuracy of the SIR model without complicating the model. We also displayed the estimation of the SIR model's discrepancy and residuals for the proposed predictors. These results show that the SIR model is inexact and that there is a strong relationship between the observation errors and the number of cases.

Future research directions are multifold. First, we assumed that finite unknown parameters can determine the variance of the observation errors. Nevertheless, there is no information about the observation errors. Hence, it is desirable to develop a method to test whether the observation error variance is constant. For heteroscedastic errors specifically, a non-parametric method can be employed to estimate the variance of the observation errors. Second, the theoretical properties of the proposed predictors should be further explored. Third, developing a Bayesian version of the proposed method is appealing, as it is more convenient to perform uncertainty quantification trials.

Acknowledgments

Dr. Wang's work was supported by the National Natural Science Foundation of China (12101024), the Natural Science Foundation of Beijing Municipality (1214019). Dr. Lu's work was supported by Scientific and Technological Research Program of Chongqing Municipal Education Commission(Grant No. KJQN201801136). Dr. Du's work was supported by the National Natural Science Foundation of China (11971045,11771032), the Natural Science Foundation of Beijing Municipality (1202001) .

Conflict of interest

The authors declare no conflict of interest.

References

1. F. Brauer, C. Castillo-Chavez, C. Castillo-Chavez, *Mathematical models in population biology and epidemiology*, vol. 2, Springer, 2012.
2. W. O. Kermack, A. G. McKendrick, A contribution to the mathematical theory of epidemics, *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, **115** (1927), 700–721.

3. M. Martcheva, *An introduction to mathematical epidemiology*, vol. 61, Springer, 2015.
4. D. Caccavo, Chinese and italian covid-19 outbreaks can be correctly described by a modified sird model, *medRxiv*. Doi: 10.1101/2020.03.19.20039388.
5. M. Shen, Z. Peng, Y. Xiao, L. Zhang, Modeling the epidemic trend of the 2019 novel coronavirus outbreak in china, *The Innovation*, **1** (2020), 100048. Doi:10.1016/j.xinn.2020.100048.
6. K. Hadeler, Parameter identification in epidemic models, *Math. Biol.*, **229** (2011), 185–189. Doi: 10.1016/j.mbs.2010.12.004.
7. A. Capaldi, S. Behrend, B. Berman, J. Smith, J. Wright, A. L. Lloyd, Parameter estimation and uncertainty quantication for an epidemic model, *Math. Biosci. Eng.*, **9** (2012), 553–576. Doi: 10.3934/mbe.2012.9.553.
8. S. Venkatramanan, B. Lewis, J. Chen, D. Higdon, A. Vullikanti, M. Marathe, Using data-driven agent-based models for forecasting emerging infectious diseases, *Epidemic*, **22** (2018), 43–49. Doi: 10.1016/j.epidem.2017.02.010.
9. M. C. Kennedy, A. O'Hagan, Bayesian calibration of computer models, *J. R. Statist. Soc. B* , **63** (2001), 425–464. Doi: 10.1111/1467-9868.00294.
10. R. Beckley, C. Weatherspoon, M. Alexander, M. Chandler, A. Johnson, G. S. Bhatt, Modeling epidemics with differential equations, *Tennessee State University Internal Report*.
11. C. Tönsing, J. Timmer, C. Kreutz, Profile likelihood-based analyses of infectious disease models, *Stat. Methods Med. Res.*, **27** (2018), 1979–1998. Doi: 10.1177/0962280217746444.
12. N. C. Roberty, L. S. de Araujo, Sir model parameters estimation with covid-19 data, *J. Adv. Math.*, **36** (2021), 97–117. Doi: 10.9734/jamcs/2021/v36i330349.
13. D. Higdon, M. Kennedy, J. C. Cavendish, J. A. Cafeo, R. D. Ryne, Combining field data and computer simulations for calibration and prediction, *SIAM J. Sci. Comput.*, **26** (2004). 448–466. Doi: 10.1137/S1064827503426693.
14. C. J. Chang, V. R. Joseph, Model calibration through minimal adjustments, *Technometrics*, **56** (2014), 474–482. Doi: 10.1080/00401706.2013.850113.
15. Y. Wang, X. Yue, R. Tuo, J. H. Hunt, J. Shi, Effective model calibration via sensible variable identification and adjustment, with application to composite fuselage simulation, *Ann. Appl. Stat.*, **14** (2020), 1759–1776, Doi: 10.1214/20-AOAS1353.
16. W. Sun, M. Plumlee, J. Hu, J. Jin, Robust system design with limited experimental data and an inexact simulation model, *SIAM-ASA J. Uncertain.*, **9** (2021), 483–506, Doi: 10.1137/20M1316287.
17. C. L. Sung, B. D. Barber, B. J. Walker, Calibration of computer models with heteroscedastic errors and application to plant relative growth rates, *arXiv preprint arXiv: 1910.11518*.
18. R. B. Gramacy, *Surrogates: Gaussian Process Modeling, Design, and Optimization for the Applied Sciences*, CRC Press, 2020.
19. R. Tuo, Y. Wang, C. F. J. Wu, On the improved rates of convergence for mat\'ern-type kernel ridge regression, with application to calibration of computer models, *SIAM-ASA J. Uncertain.*, **8** (2020), 1522–1547. Doi: 10.1137/19m1304222.

20. T. J. Santner, B. J. Williams, W. I. Notz, *The Design and Analysis of Computer Experiments*, Springer Science & Business Media, 2013.

21. M. J. Bayarri, J. O. Berger, R. Paulo, J. Sacks, J. A. Cafeo, J. Cavendish, et al., A framework for validation of computer models, *Technometrics*, **49** (2007), 138–154. Doi: 10.1198/004017007000000092.

22. M. R. Chernick, *Bootstrap methods: A guide for practitioners and researchers*, vol. 619, John Wiley & Sons, 2011.

23. R. K. Wong, C. B. Storlie, T. C. Lee, A frequentist approach to computer model calibration, *J. R. Stat. Soc. Series B. Stat. Methodol.*, **79** (2017), 635–648. Doi: 10.1111/rssb.12182.

24. C. F. J. Wu, Jackknife, bootstrap and other resampling methods in regression analysis, *Ann. Stat.*, **14** (1986), 1261–1295.

25. Anonymous, Influenza in a boarding school, *British Med. J.*, **1** (1978), 586–590. https://www.mendeley.com/catalogue/c6c2c239-c377-3ef9-945c-1afb8f91200f/

26. J. M. Murphy, D. M. Sexton, D. N. Barnett, G. S. Jones, M. J. Webb, M. Collins, et al., Quantification of modelling uncertainties in a large ensemble of climate change simulations, *Nature*, **430** (2004), 768–772. Doi: 10.1038/nature02771.

27. A. Olivares, E. Staffetti, Uncertainty quantification of a mathematical model of covid-19 transmission dynamics with mass vaccination strategy, *Chaos. Solit.*, **146** (2021), 110895. Doi: 10.1016/j.chaos.2021.110895.

28. G. Pujol, B. Iooss, M. B. Iooss, S. DiceDesign, Package 'sensitivity': Sensitivity analysis, 2015.

29. M. J. Powell, The newuoa software for unconstrained optimization without derivatives, in *Large-scale nonlinear optimization*, Springer, 2006, 255–297.

AIMS Press

© 2022 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0)