







## RESEARCH

# Evaluating short-term forecasting of COVID-19 cases among different epidemiological models under a Bayesian framework

Qiwei Li <sup>1,\*</sup>, Tejasv Bedi <sup>1,†</sup>, Christoph U. Lehmann <sup>2,3,4</sup>,  
Guanghua Xiao <sup>3,4</sup> and Yang Xie <sup>3,4</sup>

<sup>1</sup>Department of Mathematical Sciences, The University of Texas at Dallas, 800 W Campbell Rd, Richardson, TX 75080, USA; <sup>2</sup>Department of Pediatrics, The University of Texas Southwestern Medical Center, Dallas, TX 75390, USA; <sup>3</sup>Lyda Hill Department of Bioinformatics, The University of Texas Southwestern Medical Center, Dallas, TX 75390, USA and <sup>4</sup>Department of Population and Data Sciences, The University of Texas Southwestern Medical Center, Dallas, TX 75390, USA

\*Correspondence address. Qiwei Li, Department of Mathematical Sciences, The University of Texas at Dallas, Richardson, TX 75080, USA. E-mail: [qiwei.li@utdallas.edu](mailto:qiwei.li@utdallas.edu)  <http://orcid.org/0000-0002-1020-3050>

<sup>†</sup>Contributed equally.

## Abstract

**Background:** Forecasting of COVID-19 cases daily and weekly has been one of the challenges posed to governments and the health sector globally. To facilitate informed public health decisions, the concerned parties rely on short-term daily projections generated via predictive modeling. We calibrate stochastic variants of growth models and the standard susceptible-infectious-removed model into 1 Bayesian framework to evaluate and compare their short-term forecasts.

**Results:** We implement rolling-origin cross-validation to compare the short-term forecasting performance of the stochastic epidemiological models and an autoregressive moving average model across 20 countries that had the most confirmed COVID-19 cases as of August 22, 2020. **Conclusion:** None of the models proved to be a gold standard across all regions, while all outperformed the autoregressive moving average model in terms of the accuracy of forecast and interpretability.

**Keywords:** COVID-19; SARS-CoV-2; stochastic growth model; stochastic SIR model; time-series cross-validation

## Background

COVID-19, a respiratory disease caused by the coronavirus SARS-CoV-2, rapidly caused an ongoing global pandemic. By October 2020, COVID-19 had become the third leading cause of death in the USA for individuals aged 45–84 years, and it continues to spread quickly in most countries. Given the extent of health and economic distress caused by the pandemic, there is an urgent public health need to improve prediction of the spread of COVID-19 locally, nationally, and globally.

Since its emergence, a myriad of predictive modeling approaches have been proposed to forecast trends of COVID-19 disease to aid public health officials in developing effective policies and measures to suppress spread and minimize casualties. Five general approaches to forecast new cases or expected combined mortality linked to COVID-19 exist: (i) time-series forecasting such as autoregressive integrated moving average (ARIMA) [1,2], (ii) growth curve fitting based on the generalized Richards curve (GRC) or its special cases [3–7], (iii) compartmental modeling based on the susceptible-infectious-removed

Received: 9 November 2020; Revised: 11 January 2021

© The Author(s) 2021. Published by Oxford University Press GigaScience. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

(SIR) models or their derivations [8–18], (iv) agent-based modeling [19], and (v) artificial intelligence (AI)-inspired modeling [20–23].

Each approach, whether deterministic or stochastic, has its own strengths. For example, the ARIMA model combines the regressive process and the moving average, allowing prediction of a given time series by considering its own lags and lagged forecast error. Curve-fitting approaches (also known as phenomenological modeling) fit a curve to the observed number of cumulative confirmed cases or deaths with a certain error structure (e.g., Gaussian or Poisson), enabling meaningful interpretation through curve parameters while accounting for measurement errors. Compartmental modeling (also known as mechanistic modeling) assigns partitions of the population to compartments corresponding to different stages of the disease and characterizes the disease transmission dynamics by the flow of individuals through compartments. Agent-based modeling approaches use computer simulations to study the dynamic interactions among the agents (e.g., people in epidemiology) and between an agent and the environment. AI-based modeling approaches usually combine time series, clustering, and forecasting, resulting in an exemplary predictive performance. Debate among researchers has grown over model performance evaluation and selection of the best model for a certain feature (e.g., cases, deaths), a particular regional level (e.g., county, state, country), and other parameters. Fair evaluation and comparison of the output of different forecasting methods have remained elusive [24] because models vary in their complexity and the variables and parameters that characterize the dynamic states of a system.

Although the literature has compared predictive models for infectious diseases, to our knowledge, existing work does not systematically compare their performance, specifically with the same amount of *a priori* available information. We calibrate stochastic variants of 6 different growth models (i.e., logistic, generalized logistic, Richards, generalized Richards, von Bertalanffy, and Gompertz) and the standard SIR model. All models can be included using an ordinary differential equation (ODE) into 1 flexible Bayesian modeling framework. We limited the analysis to these 2 modeling approaches because both not only produce good short- and long-term forecasts but also provide useful insights into the disease dynamics of COVID-19. The growth models provide an empirical approach without a specific theory on the mechanisms giving rise to the observed patterns in the cumulative infection data, while the compartmental models incorporate key mechanisms involved in the disease transmission dynamics that explain patterns in the observed data.

In our Bayesian modeling framework, the bottom level is represented by a negative binomial model that directly models infection count data and accounts for the over-dispersed observational errors. The top level is derived from a choice of growth or compartmental models that characterize certain disease transmission dynamics through ODE(s). The Markov chain Monte Carlo (MCMC) algorithm samples from the posterior distribution. The short-term forecasts derive from the resulting MCMC samples. We perform the rolling- origin cross-validation (ROCV) procedure to compare the prediction error of different stochastic models. We used the 20 countries with the most confirmed case numbers for a country-level analysis. Observations included that (i) as the models learned more, the predictive performance improved in general for all regions; (ii) none of the models proved to be a gold standard across all the regions; and (iii) the ARIMA model underperformed all stochastic models pro-

posed in the article. We designed a graphical interface that allows users to interact with future COVID-19 trends at different geographic locations in the USA based on real-time COVID-19 data. This web portal is updated daily and used to inform local policy-makers and the general public (<https://qiwei.shinyapp.sio/PredictCOVID19/> with Biotools ID: bayesepimodels.webapp) (BayesEpiModels Web App, [RRID:SCR\\_019292](https://pubmed.ncbi.nlm.nih.gov/38111111/)).

## Data Description

Let  $\mathbf{C} = (C_1, \dots, C_T)$  be a sequence of cumulative confirmed case numbers observed at  $T$  successive equally spaced points in time (e.g., day) in a specific region, where each entry  $C_t \in \mathbb{N}$  for  $t = 1, \dots, T$ . Further let  $C_0$  be the initial value and  $\hat{\mathbf{C}} = (\hat{C}_1, \dots, \hat{C}_T)$  be the lag 1 difference of  $\mathbf{C}$ , where  $\hat{C}_1 = C_1 - C_0$  and each following entry  $\hat{C}_t = C_t - C_{t-1}$ ,  $t = 2, \dots, T$ , i.e., the difference between 2 adjacent observations. In the analysis and modeling of a series of reported infectious disease daily data, the time-series data could also be the cumulative death numbers, recovery case numbers, or their sums, denoted by  $D$  (Death),  $E$  (Recovery), and  $R$  (Removed), and their corresponding new case numbers, denoted by  $\hat{D}$ ,  $\hat{E}$ , and  $\hat{R}$ . Assuming a closed population with size  $N$ , the time-series data could also be the number of susceptible people, denoted by  $S$ , with each entry  $S_t = N - C_t$ . In reality, only confirmed cases and deaths are reported in most regions. Recovery data are not available or are hindered by under-reporting issues if available. Thus, our main goal was to make predictions of the future trend of an infectious disease only based on the daily confirmed cases  $\hat{\mathbf{C}}$ .

## Analysis

In this section, we discuss the findings of our COVID-19 data analysis. We first implemented each of the growth models listed in Table 1 and the standard SIR model under the proposed Bayesian framework for the 20 countries with the most confirmed COVID-19 case numbers as of August 22, 2020. Input data were the sequence of daily confirmed cases  $\hat{\mathbf{C}}$  only, which were accessible from the Johns Hopkins University Center for Systems Science and Engineering COVID-19 Data Repository (<https://github.com/CSSEGISandData/COVID-19/>). Several recent COVID-19 studies also based their analyses on this resource (see e.g., [25–27]). For our MCMC algorithms, we set 100,000 iterations with the first half as burn-in and chose weakly informative priors. We present numerical and graphical summaries for posterior inference and short-term forecasting. Our final goal was to compare the predictive performance of all models using ARIMA as a benchmark model.

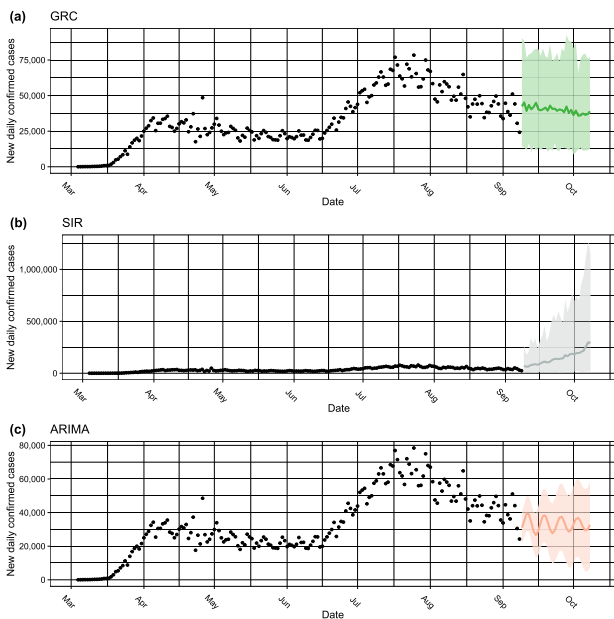
### Forecasting of daily confirmed cases in the USA

We first present the forecasting of U.S. daily confirmed cases made by the ARIMA model and our Bayesian framework with the choices of a GRC or SIR model. As seen in Figure 1, the GRC model demonstrates a downwards trend and the SIR model displays an upward trend, while the ARIMA model predicts a flat trajectory of daily predicted cases. A natural epidemiological interest is the estimated final size and end date of an epidemic. Growth models include a model parameter  $K$  that estimates the final epidemic size. For the SIR model, there is no available parameter that estimates the final size. Hence, the final case count is approximated as the predictive mean that converges to a specific value from the related MCMC samples. We applied a similar strategy to obtain the predicted mean of the final case counts us-

**Table 1:** List of  $g(\cdot)$ 's functions based on growth curves

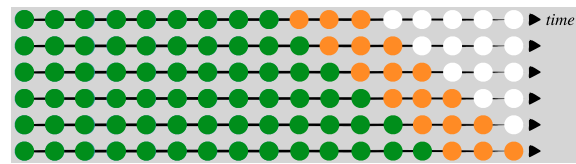
Model	$g(C_{t-1}, \Theta)$	Parameters $\Theta$	Continuous curve $y(u)$	Value at the turning point	Examples
GRC	$\lambda C_{t-1}^p \left[ 1 - \left( \frac{C_{t-1}}{K} \right)^\alpha \right]$	$K \in \mathbb{N}, \lambda \in \mathbb{R}^+, p \in (0, 1), \alpha \in \mathbb{R}^+$	NA	$\left( \frac{p}{p+\alpha} \right)^{1/\alpha} K$	[7,28,29]
Richards	$\lambda C_{t-1} \left[ 1 - \left( \frac{C_{t-1}}{K} \right)^\alpha \right]$	$K \in \mathbb{N}, \lambda \in \mathbb{R}^+, \alpha \in \mathbb{R}^+$	$K[1 + A \exp(-\lambda \alpha u)]^{-1/\alpha}$ , where $A = -1 + \left[ \frac{K}{y(0)} \right]^\alpha$	$\left( \frac{1}{1+\alpha} \right)^{1/\alpha} K$	[30-34]
GLC	$\lambda C_{t-1}^p \left( 1 - \frac{C_{t-1}}{K} \right)$	$K \in \mathbb{N}, \lambda \in \mathbb{R}^+, p \in (0, 1)$	N/A	$\frac{p}{p+1} K$	[7,35,36,37]
Logistic	$\lambda C_{t-1} \left( 1 - \frac{C_{t-1}}{K} \right)$	$K \in \mathbb{N}, \lambda \in (0, 1)$	$K[1 + A \exp(-\lambda u)]^{-1}$ , where $A = -1 + \frac{K}{y(0)}$	$\frac{1}{2} K$	[3,6,7,28,38]
von Bertalanffy	$\lambda C_{t-1}^{2/3} \left[ 1 - \left( \frac{C_{t-1}}{K} \right)^{1/3} \right]$	$K \in \mathbb{N}, \lambda \in \mathbb{R}^+$	$K \left[ 1 + A \exp(-\frac{1}{3} \gamma K^{-1/3} u) \right]^3$ , where $A = 1 - \left[ \frac{y(0)}{K} \right]^{1/3}$	$\frac{8}{27} K$	[6]
Gompertz	$\lambda C_{t-1} \log \frac{K}{C_{t-1}}$	$K \in \mathbb{N}, \lambda \in (0, 1)$	$K \exp[A \exp(-\lambda u)]$ , where $A = \log \frac{y(0)}{K}$	$\frac{1}{e} K$	[6]
GGC	$\lambda C_{t-1}^p$	$\lambda \in \mathbb{R}^+, p \in (0, 1)$	$\frac{[A + \lambda u(1-p)]^{1/(1-p)}}{y(0)^{1-p}}$ , where $A =$	NA	[7,29, 35, 36,39]

GGC: generalized growth curve; GLC: generalized logistic curve; GRC: generalized Richards curve; NA: not applicable.



**Figure 1:** The 1-month forecasting of new daily confirmed COVID-19 cases in the USA made by the (a) GRC and (b) SIR model under the proposed Bayesian framework, as well as the benchmark (c) ARIMA model. The black circles represent the observed COVID-19 case numbers since early March 2020, while the colored circles and ribbons represent the predicted means and 95% prediction intervals, respectively.

ing the ARIMA model [2]. The estimated cumulative confirmed cases by the end of 2020 were projected at 13.1, 106.1, and 10.0 (in millions), fitting the GRC, SIR, and ARIMA models, respectively. Assuming that the epidemic continues until the end of 2021, the final epidemic sizes are predicted to be 13.4, 187.3, and 22.0 (in millions) by the 3 models, respectively. To account for the discrepancies in forecasts and validate the forecast with actual reported figures, there is a need for an appropriate strategy to evaluate and compare the predictive performance of the concerned models.



**Figure 2:** A visual guide to rolling-origin cross-validation (ROCV), where the total sample size  $T = 17$ , the initial training sample size is 9, and the testing sample size is 3. The green, orange, and white circles indicate training, testing, and unused samples in 1 CV iteration.

### Model comparison through rolling-origin cross-validation

Cross-validation (CV) is a resampling procedure used to evaluate regression and classification models when only a limited data sample is available. The procedure randomly splits all data samples into 2 parts: training and testing sets, where the former is used to fit a model and the latter is used to evaluate the model's prediction performance in terms of certain error measures. The key assumption of CV is that all data points should be independent and identically distributed (i.i.d.). Unfortunately, time-series data are serially auto-correlated, meaning that the observations are dependent on the time they were recorded. To circumvent this issue, the ROCV technique was proposed [40]. It splits the data into training and testing sets without affecting the i.i.d. assumption. We used an adaption of this method to evaluate the short-term forecasting performance among different top-level choices under the proposed Bayesian framework and ARIMA. Figure 2 shows the ROCV representation for an example of time-series data ( $T = 17$ ). Algorithm 1 summarizes this evaluation procedure. The choice of initial training sample size (denoted by  $k$ ) was set to 7 days to evaluate how well the models are able to generate forecasts during the initial phase of the pandemic, while the testing sample size (denoted by  $\omega$ ) was chosen to be 3 days to meet with our objective of comparing short-term forecasting performance. We defined the first day  $t = 1$  as the date when cumulative confirmed case load per country reached 100, resulting in different days for different countries.

A CV algorithm requires a predictive error metric that can quantify model performance in terms of forecasting accuracy. Root mean square error (RMSE) and mean absolute deviations (MAD) are candidates for error measures for out-of-bag predictions but are dependent on scale. Thus, large values may influence the errors to be larger. Mean absolute percentage error (MAPE) has been a widely used predictive measure owing to its interpretability and its independence from scale, although the distribution of such percentage errors can be skewed if the data consist of values close to zero. Moreover, there is a possibility of this measure being undefined due to a zero in the denominator. In addition, MAPE can be subjected to unbounded extreme values if the actual data points are close to zero or if the absolute forecasting error  $(C_t - \hat{C}_t)$  is large. An improved percentage error metric, namely, symmetric mean absolute percentage error (sMAPE), was proposed to address these issues [40]. This measure bounded the error between 0% and 200% by incorporating the mean of actual and predicted cases  $(C_t + \hat{C}_t)/2$  in the denominator. Values close to 0% result from accurate predictions, while errors close to 200% signify inaccurate forecasting. This metric was considered in our analysis as it addressed the problem of having an unbounded measure and provided better symmetry and interpretability compared to MAPE.

---

**Algorithm 1** Rolling-origin cross-validation (ROCV)
 

---

- 1: Store the data starting day 1 to day  $T$
- 2: Initialize the number of initial training observations  $k$  ( $k = 7$ )
- 3: Set the size of the testing set  $\omega$  ( $\omega = 3$ )
- 4: **while**  $k + \omega \leq T$  **do**
- 5:   Learn the first  $k$  observations (green circles) as training data
- 6:   Hold out the next  $k + 1, \dots, k + \omega$  observations (orange circles) as the testing data
- 7:   Discard the remaining  $T - (k + \omega)$  observations (white circles)
- 8:   Compute an out-of-bag prediction error measure on the testing set (orange circles):

$$\text{sMAPE} = \frac{100\%}{\omega} \sum_{t=k+1}^{k+\omega} \left| \frac{C_t - \hat{C}_t}{(C_t + \hat{C}_t)/2} \right|, \quad 0.1$$

where  $\hat{C}_t$  denotes the predicted cumulative cases at time point  $t$ .

- 9:    $k = k + 1$
  - 10: **end while**
- 

Figure 3 displays the smoothed sMAPE curves generated by the ROCV across time for the 20 countries with the most confirmed case numbers as of August 22, 2020. All models performed poorly in the early stage, but as more data became available to be learned, the predictive performance gradually improved as the sMAPE decreased. The ARIMA and SIR models were performing significantly worse than the growth models in the early phase, which may be attributable to ARIMA (not having the growth-specific parameters) being unable to detect the early growth. However, due to assumptions of a fixed transmission rate  $\gamma$  and under-reporting of data, SIR performed poorly. The stochastic growth curves were able to learn the epidemiological data trend in the initial phase with the help of the growth and scaling parameters. In the latter half of the epidemic, all the models were performing equally well. Hence, we were unable to

conclude that any one particularly dominated the entire duration of the epidemic.

To answer the question whether we could pick 1 model with best predictive performance on average for any particular country, we constructed a Cleveland dot plot as shown in Figure 4 that allowed us to rank the model performance averaged over the entire pandemic by country. We arranged countries in ascending order of predictive performance.

## Discussion

We observed that all models performed best for Italy and worst for South Africa. The Richards model had the minimum averaged sMAPE for forecasting cumulative case counts in the USA, while the GRC model had the lowest averaged sMAPE across 7 countries followed by the GLC model with 4. The SIR model was the best performer for South Africa. The Richards, von Bertalanffy, and Gompertz models had a fair share of predictive dominance in the remaining countries. The ARIMA model performed below average across all countries.

The GRC and GLC models were consistent performers across all countries owing to their ability to detect subexponential growth rates at an early stage of an epidemic. The inclusion of the scale parameter  $\alpha$  that could account for any asymmetry in the data allowed the GRC and Richards models to generally perform best in countries that did not have symmetric “S”-shaped growth patterns and displayed randomness as well as multiple peaks. Countries including the USA, Peru, Saudi Arabia, Iran, Turkey, and France displayed multiple peaks in the daily confirmed case counts. As a result, the Richards model performed the best in the USA, UK, and Peru, while the GRC model dominated in the remaining countries with multiple peaks.

We observed a random structure in countries like Brazil, Chile, Bangladesh, and Mexico. The GRC was the most complex model and performed the best in these countries. However, the GLC model usually performed better in countries that had a single peak and an approximate “S”-shaped curvature. The GLC model was able to generalize better than the GRC model when data were well structured and less random. Argentina, Pakistan, Germany, Colombia, and India had a single peak without much randomness, and the GLC model performed better in these countries. In South Africa, the usual growth models performed the worst owing to a staggering growth rate in the initial and the middle phase of the epidemic. The SIR model performed the best out of the worst while the logistic model performed well due to its simplicity. The Gompertz model was the best performer in Russia, Spain, and Italy because it generalizes better than the other models.

## Conclusion

We developed a number of stochastic variants of growth and compartmental models in a unified Bayesian framework. The literature has discussed a theoretical comparison of growth models in great detail [5–7, 28, 29, 35, 36, 38]. However, to our knowledge, no work systematically compares the performances among all as well as against a compartmental model such as the SIR model and a time-series forecasting model such as the ARIMA model.

We conclude that the proposed Bayesian framework not only allows room for interpretation but also offers an exemplary predictive performance when it comes to COVID-19 daily report data. Moreover, ARIMA (being a pure learning algorithm) is not



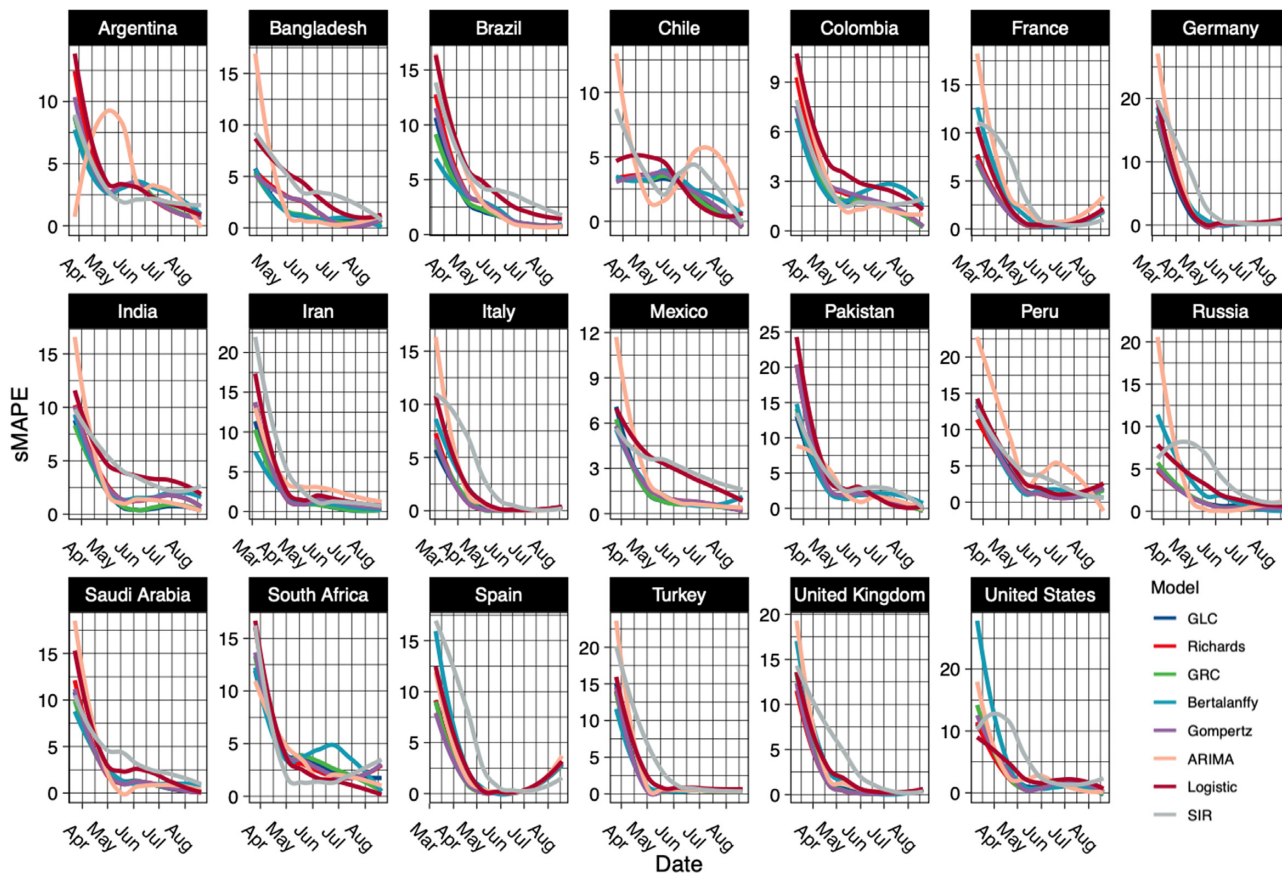


Figure 3: The smoothed sMAPE curves generated by the rolling-origin cross-validation (ROCV) over time for the 20 countries with the most confirmed COVID-19 case numbers as of August 22, 2020.

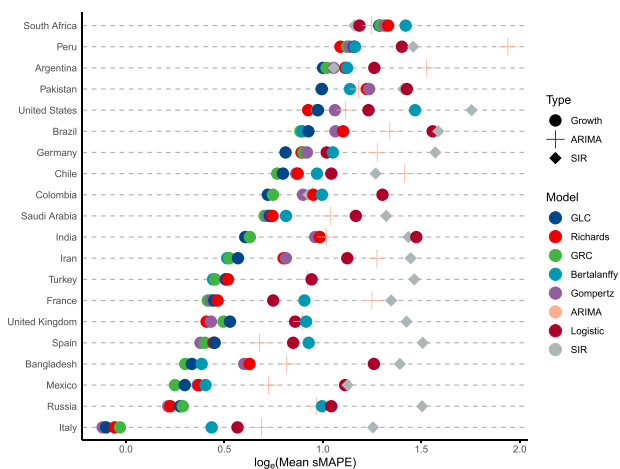


Figure 4: The Cleveland dot plot of the averaged sMAPE generated by the rolling-origin cross-validation (ROCV) for the 20 countries with the most confirmed COVID-19 case numbers as of August 22, 2020.

able to match the forecasting accuracy of stochastic models. Furthermore, the model parameters of ARIMA do not provide any information of epidemiological interest.

In the future, we aim to develop an ensemble model that can aggregate the prediction of each base model, resulting in 1 final prediction for the unseen data. Note that a group of re-

searchers have recently introduced a GGM-GLM ensemble model [35] and compared forecasting performance with the individual models for the Ebola Forecasting Challenge [41]. The ensemble model outperformed the others under some circumstances. We also plan to perform long-term forecasting evaluation using epidemic features described in [24]. A subepidemic wave model that could detect multiple peaks in the data has been recently developed [37] and has the potential to improve forecasting performance.

An “S”-shaped curvature on  $C$  attributes a simple growth model as it assumes that an epidemic would last only a short duration and that only a single peak would be observable on  $\dot{C}$ . This oversimplified assumption could be problematic because COVID-19 is more likely to be an endemic. Moreover, the changing government policies and public health guidelines as well as population behaviors (holiday) led to variable disease transmission rates, resulting in multiple peaks. Thus, developing stochastic growth models with the addition of a change-point detection mechanism to account for multiple peaks is worth investigating. We have demonstrated that an approach that combines a change-point detection model and a stochastic SIR model could significantly improve the short-term forecasting of the new daily confirmed cases [42]. To handle the sophisticated extensions of the present work, we need to utilize advanced versions of the Metropolis-Hastings (MH) algorithm in the MCMC algorithms. For example, the MH with delayed rejection [43], the combination of delayed rejection and adaptive Metropolis samplers [44], the multiple-try

Metropolis [45, 46], and the methods discussed in Liang et al. [47].

## Potential Implications

The proposed Bayesian epidemiological models in a unified framework lay the foundation for an integrative approach to model and predict epidemiological data with tremendous accuracy and interpretability. Growth and compartmental models obtained as solutions to ODEs are implemented to model epidemiological data under a deterministic setting as they provide a natural framework representative of such data types. However, the estimated model parameters crucial for providing insights into the nature of the epidemic are unreliable under the deterministic setting due to identifiability issues. The stochastic models mimic the structure of epidemiological models and incorporate parameter-specific priors and measurement errors to solve the issues. Researchers can follow a similar set-up to predict cases and deaths caused by an epidemic at any geographical level given the availability of data. Furthermore, the stochastic SIR model can be augmented by incorporating mobility, hospitalization, and recovery data, resulting in better forecasts. This work also promotes an algorithmic strategy to measure forecasting performances of time-series models in general.

On a much broader scale, this work encourages researchers to explore probabilistic approaches to model epidemiological data to develop computationally efficient algorithms that meet time and cost constraints.

## Methods

In this section, we present a bilevel Bayesian framework for predicting new confirmed cases during a pandemic in a closed society. The bottom level directly models the observed counts while accounting for measurement errors. Two alternatives for the top level are then introduced and characterize the epidemic dynamics through growth curve or compartmental trajectories, respectively. Before introducing the main components, we summarize the possibly observable data.

### Bottom level: time-series count-generating process

We consider that the new case numbers observed at time  $t$ , i.e.,  $\dot{C}_t$ , are sampled from a negative binomial (NB) model,

$$\dot{C}_t \sim \text{NB}(g(C_{t-1}, \Theta), \phi), \quad t = 2, \dots, T$$

as it automatically accounts for measurement errors and uncertainties associated with the counts. Here, we use  $\text{NB}(\mu, \phi)$ ,  $\mu, \phi > 0$  to denote an NB distribution with expectation  $\mu$  and dispersion  $1/\phi$ . We assume that this stochastic process is a Markov process, where the present state (at time  $t$ ) depends only upon its previous state (at time  $t - 1$ ). Therefore, the NB mean is a function, denoted by  $g(\cdot)$ , of the case number observed at time  $t - 1$ , characterized by a set of interpretable/uninterpretable model parameters  $\Theta$ . With this parameterization, the NB variance is  $\mu + \mu^2/\phi$ , indicating that  $\phi$  controls the variance of measurement error. A small value leads to a large variance to mean ratio, while a large value approaching infinity reduces the NB model to a Poisson model with the same mean and variance. We can write the full

data likelihood as

$$f(\dot{C} | \Theta, \phi) = \prod_{t=2}^T \frac{\Gamma(\dot{C}_t + \phi)}{\dot{C}_t! \Gamma(\phi)} \left[ \frac{\phi}{g(C_{t-1}, \Theta) + \phi} \right]^\phi \left[ \frac{g(C_{t-1}, \Theta)}{g(C_{t-1}, \Theta) + \phi} \right]^{\dot{C}_t}. \quad (1)$$

For the prior distribution of the dispersion parameter  $\phi$ , we choose a gamma distribution,  $\phi \sim \text{Ga}(a_\phi, b_\phi)$ . We recommend small values, such as  $a_\phi = b_\phi = 0.001$ , for a non-informative setting [48]. Note that the proposed framework can be viewed as a stochastic discrete-time state-space model with a negative binomial error structure. The proposed Bayesian models, on average, mimic the epidemic dynamics and are more flexible than those deterministic epidemiological models because they account for measurement error and have the potential to incorporate existing information into the prior structure of  $\Theta$ .

### Top level I: Growth model

We first discuss the choices of  $g(\cdot)$  when implementing growth models. The development of a variety of growth curves originates from population dynamics [49] and growth of biological systems [50–53] modeling. A number of growth curves have been adapted in epidemiology for trend characterization and forecasting of an epidemic, such as the severe acute respiratory syndrome (SARS) [30,31], dengue fever [32,33], pandemic influenza A (H1N1) [34], Ebola virus disease [28,38], Zika fever [29], and COVID-19 [3,6,7,54].

The underlying assumption is that the rate of growth of a population, organism, or infectious individuals eventually declines with size. The logistic curve (also known as sigmoid curve) is the simplest growth curve of continuous time  $u \in \mathbb{R}$ . It is a non-negative symmetric “S”-shaped curve with equation  $y(u) = K / (1 + \exp[-\lambda(u - u_0)])$ , where  $u_0$  is the midpoint,  $K$  is the maximum value, and  $\lambda$  reflects the steepness of the curve. Clearly,  $y(u)$  approaches  $K$  when  $u \rightarrow \infty$ , while it converges to zero when  $u \rightarrow -\infty$ . In fact, the continuous curve  $y(u)$  is the solution of a first-order non-linear ODE,

$$\frac{dy(u)}{du} = \lambda y(u) \left[ 1 - \frac{y(u)}{K} \right]$$

with condition  $y(u_0) = K/2$ , where  $dy(u)/du$  can be interpreted as time-variant growth rate of the curve  $y$ . The above ODE reveals: (i) a non-negative growth rate,  $dy(u)/du > 0$  as  $y(u) \in [0, K]$ ; (ii) an approximately exponential growth at the initial stage,  $y(u) \approx \exp(\lambda u)$  as  $dy(u)/du \approx \lambda y(u)$  when  $y(u) \rightarrow 0$ ; (iii) no growth at the final stage,  $y(u) dy(u)/du = 0$  when  $y(u) \rightarrow K$ ; (iv) a maximum growth rate of  $\lambda K/4$  occurred when  $y(u) = K/2$ , indicated by  $d^2y(u)/du^2 = \lambda dy(u)/du(1 - 2y(u)/K)$ . Based on those curve characteristics, we can use the growth curve to characterize the trend of cumulative confirmed cases  $C$ .

We mainly considered a family of growth curves that are derived from the GRC, which is the solution to the following ODE,

$$\frac{dy(u)}{du} = \lambda y(u)^p \left\{ 1 - \left[ \frac{y(u)}{K} \right]^\alpha \right\} \quad (2)$$

in continuous time  $u$ , while its discretized form at time point  $t$  is written as  $y_t - y_{t-1} = \lambda y_{t-1}^p \left[ 1 - (y_{t-1}/K)^\alpha \right]$ . For those model-specific parameters in the context of epidemiology,  $K$  is the final epidemic size and should be an integer in the range of  $(0, N]$ ,

where  $N$  is the total population,  $\lambda \in \mathbb{R}^+$  is the infectious rate at early epidemic stage,  $p \in (0, 1)$  is known as scaling of growth, and  $\alpha \in \mathbb{R}^+$  controls the curve symmetry. As our observed infectious disease data are usually counts collected at successive equally spaced discrete time points, we formulate the NB mean function  $g(\cdot)$  based on the discretized form of (2),

$$g(C_{t-1}, \Theta = \{K, \lambda, p, \alpha\}) = \lambda C_{t-1}^p \left[ 1 - \left( \frac{C_{t-1}}{K} \right)^\alpha \right]. \quad (3)$$

Table 1 provides a list of  $g(\cdot)$ 's for growth curves with their characteristics. All the listed growth curves have been utilized and discussed in previous epidemiological studies. We include all of those choices in our framework excluding the last one, which is based on the generalized growth curve (GGC), because it lacks the final epidemic size  $K$  specification.

Without any existing information, we assume that  $K$  is from a discrete uniform distribution in its range and  $\gamma$  is from a  $\gamma$ - or a  $\beta$ -distribution, depending on the choice of growth curves. For example, for both logistic and Gompertz curves, we assume  $\gamma \sim \beta(a_\gamma, b_\gamma)$ , a natural modeling choice for parameter value restricted to the (0,1) interval, and suggest choosing  $a_\gamma = b_\gamma = 1$  for a uniform setting; otherwise, we place a  $\gamma$  prior, i.e.,  $\gamma \sim \text{Ga}(a_\gamma = 0.001, b_\gamma = 0.001)$ . For the choice of GRC and generalized logistic curve (GLC), the prior of  $p$  is chosen to be  $\beta(a_p = 1, b_p = 1)$ . Last, we set  $\alpha \sim \text{Ga}(a_\alpha = 0.001, b_\alpha = 0.001)$  for fitting a GRC or Richards curve.

## Top level II: Compartmental model

The SIR model was developed to simplify the mathematical modeling of human-to-human infectious diseases by Kermack and McKendrick [55]. It is a fundamental compartmental model in epidemiology. At any given time  $u$ , each individual in a closed population with size  $N$  is assigned to 3 distinctive compartments with labels: susceptible ( $S$ ), infectious ( $I$ ), or removed ( $R$ , being either recovered or dead). The standard SIR model describes the flow of people from  $S$  to  $I$  and then from  $I$  to  $R$  by the following set of nonlinear ODEs:

$$\begin{cases} \frac{dS(u)}{du} = -\beta S(u) \frac{I(u)}{N} \\ \frac{dI(u)}{du} = \beta S(u) \frac{I(u)}{N} - \gamma I(u) \\ \frac{dR(u)}{du} = \gamma I(u) \end{cases},$$

where  $S(u)$ ,  $I(u)$ , and  $R(u)$  are the population numbers of susceptible, infectious, and removed compartments measured in time  $u$ , subject to  $S(u) + I(u) + R(u) = N$ ,  $\forall u$ . Another nature constraint is  $dS(u)/du + dI(u)/du + dR(u)/du = 0$ . Here,  $\beta \in \mathbb{R}^+$  is the disease transmission rate,  $\gamma \in \mathbb{R}^+$  is the removal rate, and their ratio  $\mathcal{R}_0 = \beta/\gamma$  is defined as the ‘‘basic reproduction number.’’ The rationale behind the first equation is that the number of new infections during an infinitesimal amount of time,  $-dS(u)/du$ , is equal to the number of susceptible people,  $S(u)$ , times the product of the contact rate,  $I(t)/N$ , and the disease transmission rate  $\beta$ . The third equation reflects that the infectious individuals leave the infectious population per unit time,  $dI(u)/du$ , at a rate of  $\gamma I(u)$ . The second equation follows immediately from the first and third equations as a result of  $dS(u)/du + dI(u)/du + dR(u)/du = 0$ . Assuming that only a small fraction of the population is infected or removed in the onset phase of an epidemic, we have  $S(u)/N \approx 1$  and thus the second equation reduces to  $dI(u)/du = (\beta - \gamma)I(u)$ , revealing that the infectious population is growing if and only if  $\beta > \gamma$ . As the expected lifetime of an infected case is given by

$\gamma^{-1}$ , the ratio  $\mathcal{R}_0 = \beta/\gamma$  is the average number of new infectious cases directly produced by an infected case in a completely susceptible population. The so-called basic reproduction number is a good indicator of an infectious disease’s transmissibility.

We only consider the standard SIR model, although it is still feasible to design  $g(\cdot)$ 's from its variations (see a comprehensive summary [56]), such as the susceptible-infectious (SIS) model, the susceptible-infectious-recovered-deceased (SIRD) model, the susceptible-exposed-infectious-removed (SEIR) model, the susceptible-exposed-infectious-susceptible (SEIS) model, and their versions with a maternally derived immunity compartment [57], as well as the recently developed extended-SIR (eSIR) model [14]. For modeling discrete time-series data, we use the discrete-time version of the standard SIR model,

$$\begin{cases} \dot{S}_t = -\beta S_{t-1} \frac{I_{t-1}}{N} \\ \dot{I}_t = \beta S_{t-1} \frac{I_{t-1}}{N} - \gamma I_{t-1} \\ \dot{R}_t = \gamma I_{t-1} \end{cases}, \quad (4)$$

where  $\dot{S}_t = S_t - S_{t-1}$ ,  $\dot{I}_t = I_t - I_{t-1}$ , and  $\dot{R}_t = R_t - R_{t-1}$  are the differences between 2 adjacent observations in the sequence of susceptible, infectious, and removed case numbers, respectively. The model has 3 trajectories, 1 for each compartment. The compositional nature of the 3 trajectories implies that we only need 2 of the 3 sequence data, e.g.,  $S_t = N - C_t$  and  $R_t$  for  $t = 1, \dots, T$ . However, recovery data only exist in a few regions and are hindered by under-reporting even if they exist, which makes both model inference and predictions infeasible. Alternatively, we consider both the removed and actively infectious cases as missing data and mimic their relationship in spirit to some compartmental models in epidemiology. Specifically, we assume that the number of new removed cases at time  $t$ , i.e.,  $\dot{R}_t$ , is sampled from a Poisson distribution with mean  $\gamma I_{t-1}$ , i.e.,  $\dot{R}_t \sim \text{Poi}(\gamma I_{t-1}) = \text{Poi}(\gamma(N - C_{t-1} - R_{t-1}))$ , where  $\gamma$  should be specified. Such a strategy but with different error structure was also considered in some other compartmental models in epidemiology [16,58,59]. We can estimate the value of  $\gamma$  from publicly available high-quality data where confirmed cases, deaths, and recovered cases are all well documented, or from prior epidemic studies due to the same under-reporting issue in actual data. In this article, we choose the removal rate  $\gamma = 0.1$  as suggested by Weitz et al. [60]. Based on this simplification, we rewrite the first equation in (4) as,

$$(N - C_t) - (N - C_{t-1}) = -\beta(N - C_{t-1}) \frac{N - C_{t-1} - R_{t-1}}{N},$$

resulting in

$$\dot{C}_t = \beta(N - C_{t-1}) \frac{N - C_{t-1} - R_{t-1}}{N}.$$

Thus, we formulate the NB mean function  $g(\cdot)$  for the standard SIR model as,

$$g(C_{t-1}, \Theta = \{\beta\} | R) = \beta(N - C_{t-1}) \frac{N - C_{t-1} - R_{t-1}}{N}, \quad (5)$$

where  $R$  can be sequentially inferred from  $C$ .

Without any existing information, in our Bayesian framework we assume  $\beta$  from a  $\gamma$ -distribution with hyperparameters that makes both the mean and variance of the transformed variable  $\mathcal{R}_0 = \beta/\gamma$  equal to 1, i.e.,  $\beta \sim \text{Ga}(1, 1/\gamma)$ .

## Model Fitting

In this section, we briefly describe the MCMC algorithm for posterior inference and forecasting. Our Bayesian inferential strategy allows us to simultaneously infer all model-specific parameters and quantify their uncertainties.

### MCMC algorithms

We first describe how to update the dispersion parameter  $\phi$  in the proposed Bayesian framework because it does not depend on the choice of models. At each MCMC iteration, we perform the following step:

Update of dispersion parameter  $\phi$ . We update  $\phi$  by using a random walk Metropolis-Hastings (RWMH) algorithm. We first propose a new  $\phi^*$ , of which logarithmic value is generated from  $N(\log \phi, \tau_\phi^2)$  and then accept the proposed value  $\phi^*$  with probability  $\min(1, m_{\text{MH}})$ , where the Hastings ratio is

$$m_{\text{MH}} = \frac{f(\hat{C}|\phi^*, \Theta) \pi(\phi^*) J(\phi \leftarrow \phi^*)}{f(\hat{C}|\phi, \Theta) \pi(\phi) J(\phi^* \leftarrow \phi)}.$$

Here we use  $J(\cdot \leftarrow \cdot)$  to denote the proposal probability distribution for the selected move. Note that the last term, which is the proposal density ratio, cancels out for this RWMH update.

#### Top level as a growth model

We only present the updates of each parameter in the GRC model because all other derivative models are its special cases. Our primary interest lies in the estimation of the final pandemic size  $K$  and the infectious rate at early epidemic stage  $\lambda$ .

Update of final epidemic size parameter  $K$ . We update  $K$  by using an RWMH algorithm. We first propose a new  $K^*$ , of which logarithmic value is generated from a truncated Poisson distribution  $\text{Poi}(\log K)$  within  $[\log C_T, \log N]$ , and then accept the proposed value  $K^*$  with probability  $\min(1, m_{\text{MH}})$ , where the Hastings ratio is

$$m_{\text{MH}} = \frac{f(\hat{C}|\phi, K^*, \lambda, p, \alpha) \pi(K^*) J(K \leftarrow K^*)}{f(\hat{C}|\phi, K, \lambda, p, \alpha) \pi(K) J(K^* \leftarrow K)}.$$

Note that with a discrete uniform prior on  $K$ , the last 2 terms cancel out for this RWMH update.

Update of infectious rate parameter  $\lambda$ . We update  $\lambda$  by using an RWMH algorithm. We first propose a new  $\lambda^*$ , of which logarithmic value is generated from  $N(\log \lambda, \tau_\lambda^2)$ , and then accept the proposed value  $\lambda^*$  with probability  $\min(1, m_{\text{MH}})$ , where the Hastings ratio is

$$m_{\text{MH}} = \frac{f(\hat{C}|\phi, K, \lambda^*, p, \alpha) \pi(\lambda^*) J(\lambda \leftarrow \lambda^*)}{f(\hat{C}|\phi, K, \lambda, p, \alpha) \pi(\lambda) J(\lambda^* \leftarrow \lambda)}.$$

Note that the last term, which is the proposal density ratio, cancels out for this RWMH update.

Update of growth scaling parameter  $p$ . We update  $p$  by using an RWMH algorithm. We first propose a new  $p^*$ , of which logarithmic value is generated from a truncated normal distribution  $N(\log p, \tau_p^2)$  within  $[-\infty, 0]$ , and then accept the proposed value  $p^*$  with probability  $\min(1, m_{\text{MH}})$ , where the Hastings ratio is

$$m_{\text{MH}} = \frac{f(\hat{C}|\phi, K, \lambda, p^*, \alpha) \pi(p^*) J(p \leftarrow p^*)}{f(\hat{C}|\phi, K, \lambda, p, \alpha) \pi(p) J(p^* \leftarrow p)}.$$

Note that with a uniform prior on  $p$  over its range  $[0, 1]$ , the last 2 terms cancel out for this RWMH update.

Update of symmetry parameter  $\alpha$ . We update  $\alpha$  by using an RWMH algorithm. We first propose a new  $\alpha^*$ , of which logarithmic value is generated from  $N(\log \alpha, \tau_\alpha^2)$ , and then accept the proposed value  $\alpha^*$  with probability  $\min(1, m_{\text{MH}})$ , where the Hastings ratio is

$$m_{\text{MH}} = \frac{f(\hat{C}|\phi, K, \lambda, p, \alpha^*) \pi(\alpha^*) J(\alpha \leftarrow \alpha^*)}{f(\hat{C}|\phi, K, \lambda, p, \alpha) \pi(\alpha) J(\alpha^* \leftarrow \alpha)}.$$

Note that the last term, which is the proposal density ratio, cancels out for this RWMH update.

#### Top level as a compartmental model

Our primary interest lies in the estimation of the reproduction number  $\mathcal{R}_0 = \beta/\gamma$ , which reflects the transmissibility of the disease. With our assumption that  $\gamma$  is given, we propose the following updates in each MCMC iteration.

Generate  $\mathbf{R}$  based on  $\mathbf{C}$ . We assume  $I_1 = C_1$ , i.e., all the confirmed cases are capable of passing the disease to all susceptible individuals in a closed population at the very beginning. In other words,  $R_1 = 0$ . Then we sample  $\hat{R}_2 \sim \text{Poi}(\gamma I_1)$ , where  $\gamma$  is a prespecified tuning parameter and  $\hat{R}_2 = R_2 - R_1$  ( $R_2 = R_2$  here in that  $R_1 = 0$ ) is the new removed case numbers at time  $t = 2$ . Owing to the compositional nature, we can compute  $I_2 = I_1 + \hat{C}_2 - \hat{R}_2$ , where  $\hat{C}_2 = C_2 - C_1$  is the new confirmed case numbers at time  $t = 2$ . Next, we repeat this process of sampling  $\hat{R}_t \sim \text{Poi}(\gamma I_{t-1})$  and computing  $I_t = I_{t-1} + \hat{C}_t - \hat{R}_t$ ,  $t = 3, \dots, T$ , to generate the sequence  $\mathbf{R}$ .

Update of reproduction number parameter  $\beta$ . We update  $\beta$  by using an RWMH algorithm. We first propose a new  $\beta^*$ , of which logarithmic value is generated from a truncated normal distribution  $N(\log \beta, \tau_\beta^2)$ , and then accept the proposed value  $\beta^*$  with probability  $\min(1, m_{\text{MH}})$ , where the Hastings ratio is

$$m_{\text{MH}} = \frac{f(\hat{C}|\beta^*, \mathbf{R}) \pi(\beta^*) J(\beta \leftarrow \beta^*)}{f(\hat{C}|\beta, \mathbf{R}) \pi(\beta) J(\beta^* \leftarrow \beta)}.$$

Note that the last term, which is the proposal density ratio, cancels out for this RWMH update.

### Posterior inference

We obtain posterior inference by post-processing the MCMC samples. Update of reproduction number parameter after burn-in. Suppose that a sequence of MCMC samples on  $\theta$ ,  $\theta \in \{\phi, K, \lambda, p, \alpha, \beta\}$ ,

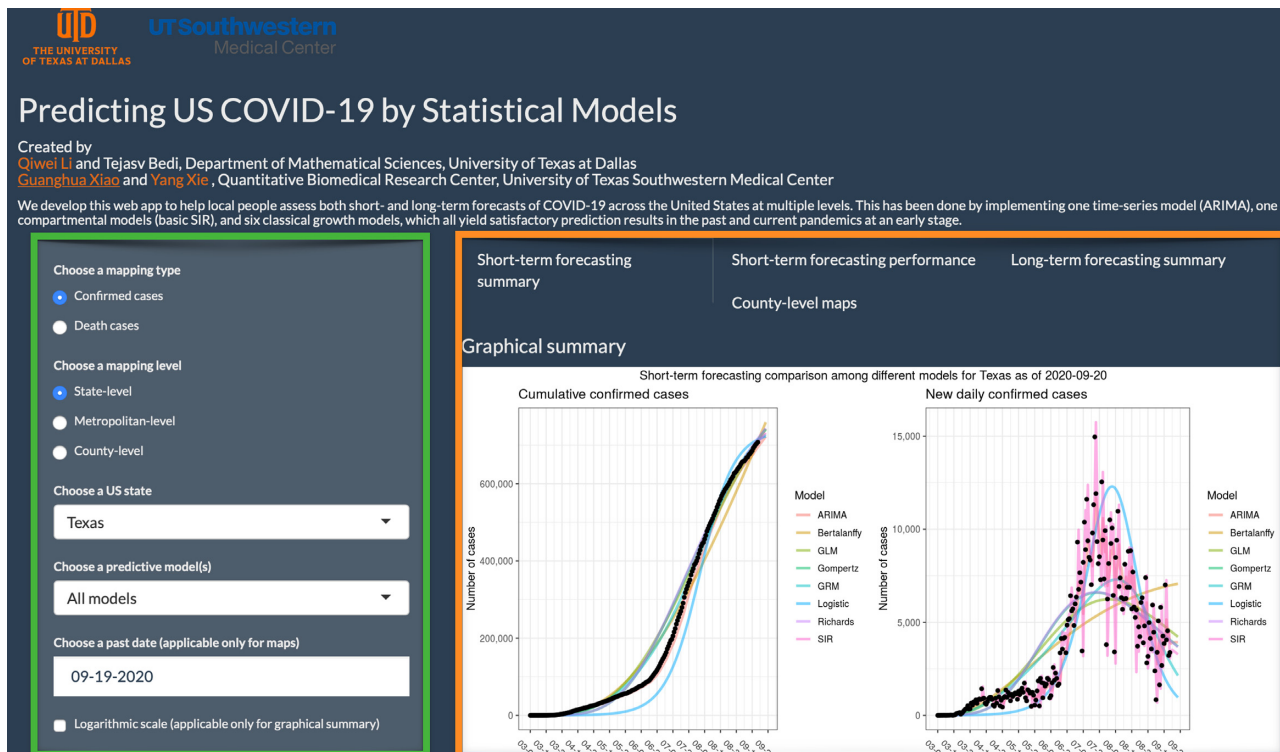
$$\theta^{(1)}, \dots, \theta^{(U)}$$

has been collected, where  $u$ ,  $u = 1, \dots, U$  indexes the iteration after burn-in. An approximate Bayesian estimator of each parameter can be obtained simply by averaging over the samples,  $\hat{\theta} = \sum_{u=1}^U \theta^{(u)} / U$ . In addition to that, a quantile estimation or credible interval for each parameter of interest can be obtained from this sequence as well.

### Forecasting

On the basis of the sequences of MCMC samples on  $K, \lambda, p$ , and  $\alpha$  in the growth model or  $\beta$  in the compartmental model, we





**Figure 5:** The web interface of the COVID-19 trend analysis page. The green box highlights the input panel that allows users to choose different mapping types and levels for a region. The orange box highlights the visualizations for short-term forecasting as per the instructions of users. Other tabs offer different graphs for summarizing the model performance, long-term forecasting, and county-level spatial maps.

can predict the cumulative or new confirmed cases at any future time  $T_f$  by Monte Carlo simulation. Particularly, from time  $T + 1$  to  $T_f$ , we sequentially generate

$$\hat{C}_t^{(u)} \sim \text{NB}(g(C_{t-1}, \Theta^{(u)}), \phi^{(u)}), \quad t = T + 1, \dots, T_f. \quad (6)$$

Then, both short- and long-term forecast can be made by summarizing the  $(T_f - T)$ -by- $U$  matrix of MCMC samples. For example, the mean predictive number of cumulative and new confirmed cases at time  $T + 1$  are  $\sum_{u=1}^U C_{T+1}^{(u)}/U$  and  $\sum_{u=1}^U \hat{C}_{T+1}^{(u)}/U$ , respectively.

## Software

This article introduces a user-friendly interactive web application (<https://qiwei.shinyapps.io/PredictCOVID19/> with Biotoools ID: bayesepimodels.webapp) (BayesEpiModels Web App, RRID: SCR.019292) integrated with the R Shiny package. Shiny is a web platform that allows users to interact with real-time data and use a myriad of visualization tools to analyze them. Figure 5 shows a screenshot of the web application. The web application has been developed to help the general public assess both short- and long-term forecasts of COVID-19 across the USA at both state and metropolitan level. The numbers of cumulative or new daily confirmed cases as well as deaths are projected by different growth models and the SIR model under the proposed Bayesian framework. Alongside the numerical summaries, users can view and interpret the trends that cover the same information. To validate the short-term forecasting, numerical and graphical summaries of MAE and MAPE of the predictions are

provided for the more advanced users. Moving on to the long-term forecasting, the models estimate the peak number of cases and deaths, as well as their respective dates. Moreover, predictive estimates for the final size and date are also offered. Finally, for the users keen on visualizing the currently observed cases at a geographical level, the website offers county-level spatial maps.

## Availability of Source Code and Requirements

- Project name: BayesEpiModels
- Project home page: <https://github.com/liqiwei2000/BayesEpiModels>
- Operating systems: Windows and Linux
- Programming language: R (version 3.6.0)
- Other requirements: None
- License: GNU General Public License v3.0
- Biotoools ID: bayesepimodels
- RRID:SCR.019291

## Data Availability

The related R/C++ codes for model preparation and execution are available on GitHub at <https://github.com/liqiwei2000/BayesEpiModels>, with snapshots in the GigaScience GigaDB repository [61]. The R Shiny web application is available for users at <https://qiwei.shinyapps.io/PredictCOVID19/>. The COVID-19 data repository is operated by the Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE) and is freely available on GitHub at <https://github.com/CSSEGISandData/COVID-19/>.

## Abbreviations

AI: artificial intelligence; ARIMA: autoregressive integrated moving average; COVID-19: coronavirus disease–2019; CV: cross-validation; eSIR: extended susceptible-infectious-removed; GGC: generalized growth curve; GGM-GLM: generalized growth model-generalized logistic model; GLC: generalized logistic curve; GRC: generalized Richards curve; i.i.d.: independent and identically distributed; JHU CSSE: Johns Hopkins University Center for Systems Science and Engineering; MAD: mean absolute deviations; MAPE: mean absolute percentage error; MCMC: Markov chain Monte Carlo; NB: negative binomial; ODE: ordinary differential equation; RMSE: root mean square error; ROCV: rolling-origin cross-validation; RWMH: random walk Metropolis-Hastings; SARS-Cov-2: severe acute respiratory syndrome coronavirus 2; SEIR: susceptible-exposed-infectious-removed; SEIS: susceptible-exposed-infectious-susceptible; SIR: susceptible-infectious-removed; SIRD: susceptible-infectious-recovered-deceased; SIS: susceptible-infectious-susceptible; SMAPE: symmetric mean absolute percentage error.

## Competing Interests

The authors declare that they have no competing interests.

## Funding

This work was supported by the University of Texas at Dallas (UT Dallas) Office of Research (UT Dallas Center for Disease Dynamics and Statistics) and partially supported by the National Institutes of Health (1R01GM115473, 1R01GM140012, 5R01CA152301, P30CA142543, P50CA70907, R35GM136375); and the Cancer Prevention and Research Institute of Texas (RP180805, RP190107).

## Authors' Contributions

Q.L. developed the Bayesian framework, designed the MCMC algorithms, and constructed the R Shiny web application. Q.L. and T.B. contributed to the review of different methods and collaborated in the real data analysis. Q.L., C.U.L., G.X., and Y.X. conceived the study and supervised the web application development and the statistical analyses. All authors contributed to the writing of the manuscript. All authors have read and approved the final manuscript.

## Acknowledgments

The authors thank Drs. Swati Biswas, Min Chen, Pankaj Choudhary, and Vladimir Dragovic in the Department of Mathematical Sciences at the University of Texas at Dallas for their support and valuable suggestions on this work.

## References

1. Elmousalami HH, Hassanien AE. Day level forecasting for coronavirus disease (COVID-19) spread: analysis, modeling and recommendations. *arXiv 2020*: 2003.07778.
2. Perone G. An ARIMA model to forecast the spread and the final size of COVID-2019 epidemic in Italy. *medRxiv 2020*, doi:10.1101/2020.04.27.20081539.
3. Batista M. Estimation of the final size of the COVID-19 epidemic. *medRxiv 2020*, doi:10.1101/2020.02.16.20023606.
4. IHME COVID-19 health service utilization forecasting team, Murray CJ. Forecasting COVID-19 impact on hospital bed-days, ICU-days, ventilator-days and deaths by US state in the next 4 months. *medRxiv 2020*, doi:10.1101/2020.03.27.20043752.
5. Roosa K, Lee Y, Luo R, et al. Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24th, 2020. *Infect Dis Model 2020*;5:256–63.
6. Jia L, Li K, Jiang Y, et al. Prediction and analysis of coronavirus disease 2019. *arXiv 2020*:2003.05447.
7. Wu K, Darcet D, Wang Q, et al. Generalized logistic growth modeling of the COVID-19 outbreak in 29 provinces in China and in the rest of the world. *Nonlinear Dyn 2020*;101:1561–81.
8. Fanelli D, Piazza F. Analysis and forecast of COVID-19 spreading in China, Italy and France. *Chaos Solitons Fractals 2020*;134:109761.
9. Kucharski AJ, Russell TW, Diamond C, et al. Early dynamics of transmission and control of COVID-19: A mathematical modelling study. *Lancet Infect Dis 2020*;20(5): 553–8.
10. Li R, Pei S, Chen B, et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science 2020*;368(6490):489–93.
11. Liu Z, Magal P, Seydi O, et al. Predicting the cumulative number of cases for the COVID-19 epidemic in China from early data. *Math Biosci Eng 2020*;17(4):3040–51.
12. Pan A, Liu L, Wang C, et al. Association of public health interventions with the epidemiology of the COVID-19 outbreak in Wuhan, China. *JAMA 2020*;323(19):1915–23.
13. Pei S, Shaman J. Initial simulation of SARS-CoV2 spread and intervention effects in the continental US. *medRxiv 2020*, doi:10.1101/2020.03.21.20040303.
14. Song PX, Wang L, Zhou Y, et al. An epidemiological forecast model and software assessing interventions on COVID-19 epidemic in China. *J Data Sci 2020*;18(3): 409–32.
15. Sun H, Qiu Y, Yan H, et al. Tracking reproductivity of COVID-19 epidemic in China with varying coefficient SIR model. *J Data Sci 2020*;18(3):455–72.
16. Wang L, Wang G, Gao L, et al. Spatiotemporal dynamics, nowcasting and forecasting of COVID-19 in the United States. *arXiv 2020*:2004.14103.
17. Yamana T, Pei S, Kandula S, et al. Projection of COVID-19 cases and deaths in the US as individual states re-open May 4, 2020. *medRxiv 2020*, doi:10.1101/2020.05.04.20090670.
18. Yang Z, Zeng Z, Wang K, et al. Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions. *J Thorac Dis 2020*;12(3): 165.
19. Gomez J, Prieto J, Leon E, et al. INFEKTA: A general agent-based model for transmission of infectious diseases: Studying the COVID-19 propagation in Bogotá-Colombia. *medRxiv 2020*, doi:10.1101/2020.04.06.20056119.
20. Bullock J, Pham KH, Lam CSN, et al. Mapping the landscape of artificial intelligence applications against COVID-19. *J Artif Intell Res 2020*;69, doi:10.1613/jair.1.12162.
21. Hu Z, Ge Q, Jin L, et al. Artificial intelligence forecasting of COVID-19 in China. *arXiv 2020*:2002.07112.
22. Distante C, Pereira IG, Gonçalves LMG, et al. Forecasting COVID-19 outbreak progression in Italian regions: A model based on neural network training from Chinese data. *medRxiv 2020*, doi:10.1101/2020.04.09.20059055.

23. Vaishya R, Javaid M, Khan IH, et al. Artificial Intelligence (AI) applications for COVID-19 pandemic. *Diabetes Metab Syndr* 2020;**14**(4):337–9.
24. Tabataba FS, Chakraborty P, Ramakrishnan N, et al. A framework for evaluating epidemic forecasts. *BMC Infect Dis* 2017;**17**(1):345.
25. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect Dis* 2020;**20**(5):533–4.
26. Zhou T, Ji Y. Semiparametric Bayesian inference for the transmission dynamics of COVID-19 with a state-space model. *Contemp Clin Trials* 2020;**97**:106146.
27. Toda AA. Susceptible-infected-recovered (SIR) dynamics of COVID-19 and economic impact. arXiv 2020:2003.11221.
28. Chowell G, Simonsen L, Viboud C, et al. Is West Africa approaching a catastrophic phase or is the 2014 Ebola epidemic slowing down? Different models yield different answers for Liberia. *PLoS Curr* 2014;**6**, doi:10.1371/currents.outbreaks.b4690859d91684da963dc40e00f3da81.
29. Chowell G, Hincapie-Palacio D, Ospina J, et al. Using phenomenological models to characterize transmissibility and forecast patterns and final burden of Zika epidemics. *PLoS Curr* 2016;**8**, doi:10.1371/currents.outbreaks.f14b2217c902f453d9320a43a35b9583.
30. Hsieh YH, Lee JY, Chang HL. SARS epidemiology modeling. *Emerg Infect Dis* 2004;**10**(6):1165.
31. Hsieh YH. Richards model: A simple procedure for real-time prediction of outbreak severity. In: *Modeling and Dynamics of Infectious Diseases*. World Scientific; 2009:216–36.
32. Hsieh YH, Ma S. Intervention measures, turning point, and reproduction number for dengue, Singapore, 2005. *Am J Trop Med Hyg* 2009;**80**(1):66–71.
33. Hsieh YH, Chen C. Turning points, reproduction number, and impact of climatological events for multi-wave dengue outbreaks. *Trop Med Int Health* 2009;**14**(6):628–38.
34. Hsieh YH. Pandemic influenza A (H1N1) during winter influenza season in the southern hemisphere. *Influenza Other Respir Viruses* 2010;**4**(4):187–97.
35. Chowell G, Luo R, Sun K, et al. Real-time forecasting of epidemic trajectories using computational dynamic ensembles. *Epidemics* 2020;**30**:100379.
36. Chowell G. Fitting dynamic models to epidemic outbreaks with quantified uncertainty: A primer for parameter uncertainty, identifiability, and forecasts. *Infect Dis Modell* 2017;**2**(3):379–98.
37. Chowell G, Tariq A, Hyman JM. A novel sub-epidemic modeling framework for short-term forecasting epidemic waves. *BMC Med* 2019;**17**(1):164.
38. Pell B, Kuang Y, Viboud C, et al. Using phenomenological models for forecasting the 2015 Ebola challenge. *Epidemics* 2018;**22**:62–70.
39. Viboud C, Simonsen L, Chowell G. A generalized-growth model to characterize the early ascending phase of infectious disease outbreaks. *Epidemics* 2016;**15**:27–37.
40. Tashman L. Out-of sample tests of forecasting accuracy: A tutorial and review. *Int J Forecast* 2000;**16**(4):437–50.
41. Viboud C, Sun K, Gaffey R, et al. The RAPIDD ebola forecasting challenge: Synthesis and lessons learnt. *Epidemics* 2018;**22**:13–21.
42. Jiang S, Zhou Q, Zhan X, et al. BayesSMILES: Bayesian Segmentation Modeling for Longitudinal Epidemiological Studies. medRxiv 2020, doi:10.1101/2020.10.06.20208132.
43. Mira A. On Metropolis-Hastings algorithms with delayed rejection. *Metron* 2001;**59**(3-4):231–41.
44. Haario H, Laine M, Mira A, et al. DRAM: Efficient adaptive MCMC. *Stat Comput* 2006;**16**(4):339–54.
45. Liu JS, Liang F, Wong WH. The multiple-try method and local optimization in Metropolis sampling. *J Am Stat Assoc* 2000;**95**(449):121–34.
46. Martino L. A review of multiple try MCMC algorithms for signal processing. *Digit Signal Process* 2018;**75**:134–52.
47. Liang F, Liu C, Carroll R. *Advanced Markov Chain Monte Carlo Methods: Learning from past samples*. Wiley; 2011.
48. Gelman A. Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Anal* 2006;**1**(3):515–34.
49. Haberman R. *Mathematical models: Mechanical vibrations, population dynamics, and traffic flow*. Society for Industrial and Applied Mathematics; 1998, doi:10.1137/1.9781611971156.
50. Werker A, Jaggard K. Modelling asymmetrical growth curves that rise and then fall: Applications to foliage dynamics of sugar beet (*Beta vulgaris* L.). *Ann Bot* 1997;**79**(6):657–65.
51. Desta F, Mac Siurtain M, Colbert J. Parameter estimation of nonlinear growth models in forestry. *Silv Fenn* 1999;**33**(4):327–36.
52. Kaps M, Herring W, Lamberson W. Genetic and environmental parameters for traits derived from the Brody growth curve and their relationships with weaning weight in Angus cattle. *J Anim Sci* 2000;**78**(6):1436–42.
53. Topal M, Bolukbasi Ş. Comparison of nonlinear growth curve models in broiler chickens. *J Appl Anim Res* 2008;**34**(2):149–52.
54. Lee SY, Lei B, Mallick B. Estimation of COVID-19 spread curves integrating global data and borrowing information. *PLoS One* 2020;**15**(7), doi:10.1371/journal.pone.0236860.
55. Kermack WO, McKendrick AG. A contribution to the mathematical theory of epidemics. *Proc R Soc Lond A Math Phys Sci* 1927;**115**(772):700–21.
56. Bailey NT. *The Mathematical Theory of Infectious Diseases and Its Applications*. Charles Griffin; 1975.
57. Hethcote HW. *The mathematics of infectious diseases*. *SIAM Rev Soc Ind Appl Math* 2000;**42**(4):599–653.
58. Siettos CI, Russo L. Mathematical modeling of infectious disease dynamics. *Virulence* 2013;**4**(4):295–306.
59. Anastassopoulou C, Russo L, Tsakris A, et al. Data-based analysis, modelling and forecasting of the COVID-19 outbreak. *PLoS One* 2020;**15**(3):e0230405.
60. Weitz JS, Beckett SJ, Coenen AR, et al. Modeling shield immunity to reduce COVID-19 epidemic spread. *Nat Med* 2020;**26**(6):849–54.
61. Li Q, Bedi T, Lehmann CU, et al. Supporting data for “Evaluating short-term forecasting of COVID-19 cases among different epidemiological models under a Bayesian framework.” GigaScience Database 2021. <http://dx.doi.org/10.5524/100863>.