FORECASTING COVID-19 PANDEMIC: A DATA-DRIVEN ANALYSIS

Khondoker Nazmoon Nabi

Department of Mathematics, Bangladesh University of Engineering and Technology (BUET), Dhaka-1000, Bangladesh.

Abstract

In this paper, a new Susceptible-Exposed-Symptomatic Infectious-Asymptomatic Infectious-Quarantined-Hospitalized-Recovered-Dead (SEI_DI_UQHRD) deterministic compartmental model has been proposed and calibrated for describing the transmission dynamics of the novel coronavirus disease (COVID-19). A calibration process is executed through the solution of an inverse problem with the help of a Trust-Region-Reflective algorithm, used to determine the best parameter values that would fit the model response. The purpose of this study is to give a tentative prediction of the epidemic peak for Russia, Brazil, India and Bangladesh which could become the next COVID-19 hotspots in no time. Based on the publicly available epidemiological data from late January until 10 May, it has been estimated that the number of daily new symptomatic infectious cases for the above mentioned countries could reach the peak around the beginning of June with the peak size of $\sim 15,774$ symptomatic infectious cases in Russia, $\sim 26,449$ cases in Brazil, $\sim 9,504$ cases in India and $\sim 2,209$ cases in Bangladesh. Based on our analysis, the estimated value of the basic reproduction number (R_0) as of May 11, 2020 was found to be ~ 4.234 in Russia, ~ 5.347 in Brazil, ~ 5.218 in India, ~ 4.649 in the United Kingdom and ~ 3.5 in Bangladesh. Moreover, with an aim to quantify the uncertainty of our model parameters, Latin hypercube sampling-partial rank correlation coefficient (LHS-PRCC) which is a global sensitivity analysis (GSA) method is applied which elucidates that, for Russia, the recovery rate of undetected asymptomatic carriers, the rate of getting home-quarantined or self-quarantined and the transition rate from quarantined class to susceptible class are the most influential parameters, whereas the rate of getting home-quarantined or self-quarantined and the inverse of the COVID-19 incubation period are highly sensitive parameters in Brazil, India, Bangladesh and the United Kingdom which could significantly affect the transmission dynamics of the novel coronavirus. Our analysis also suggests that relaxing social distancing restrictions too quickly could exacerbate the epidemic outbreak in the above mentioned countries.

Keywords: Compartmental model, COVID-19, coronavirus, asymptomatic carrier, quarantined class, model calibration, sensitivity.

Email address: nabinil@yahoo.com (Khondoker Nazmoon Nabi)

1. Introduction

The coronavirus disease 2019 (COVID-19) has become a public health emergency of global concern affecting 212 countries and territories around the world as of May 12, 2020 Worldometer, (2020). It is a respiratory disease caused by the novel coronavirus (SARS-CoV-2) that was first identified in Wuhan, Hubei province, China, late December 2019 Li et al., (2020). This novel virus started transmitting around the world rapidly, and on 30 January, WHO declared the outbreak a Public Health Emergency of International Concern (PHEIC). On 11 March, WHO Director-General marked COVID-19 as a pandemic. As of May 11, 2020, an outbreak of COVID-19 has resulted in 4,271,689 confirmed cumulative cases with reported deaths of 287,613 worldwide Worldometer, (2020). In most of the countries, infected patients are struggling to get the proper treatment due to highly transmissible and virulent nature of the virus. Nevertheless, numerous COVID-19 mitigation strategies has been adapted so far such as quarantine, isolation, promoting the wearing of face masks, travel restrictions, and lockdowns with a view to reducing community transmission of the disease.

In the absence of either proven effective therapy or a vaccine, the rising numbers of infections can overburden the already fragile health care systems of different developed and developing countries in the coming months if the spread is not controlled. This pandemic is giving rise to significant social, economic, and health impacts and has highlighted the significance of detecting the development trend of the disease and prediction of disease future dynamics for designing infectious disease prevention and control strategies, effective public health policies and economical activity guidelines. Different mathematical paradigms have always played a notable role in providing deeper understanding of the transmission mechanisms of a disease outbreak, contributing considerable insights for controlling the disease outbreak. One of the familiar models that is reasonably predictive for human-to-human transmission is Susceptible-Infectious-Removed (SIR) epidemic model proposed by Kermack-Mckendrick in 1927 Calafiore et al., (2020); Nesteruk (2020). After that, the SIR epidemic model has been extended to Susceptible-Exposed-Infectious-Removed (SEIR) model and many of its variants to explore the risk factors of a disease or predict the dynamics of a disease outbreak Wu et al., (2020); Calafiore et al., (2020); kucharski et al., (2020); Simha et al., (2020); Anastassopoulou et al., (2020); Nesteruk (2020). It is really challenging in population based model to incorporate certain real-world complexities. In fact, analysis and prediction could go wrong in the absence of adequate historical real data. On the other hand, different stochastic agent-based models where individuals typically interact on a network structure and exchange infection stochastically, have been established as useful tools for tracing fine-grained effects of heterogeneous intervention policies in diverse epidemic and pandemic settings Ferguson et al., (2020); Chang et al., (2020); Wilder et al., (2020); Ruiz et al., (2020), including for policy advice currently in place in the USA and the UK. However, accuracy of this approach can be a vital issue due to the time-varying nature of network-structure.

Since the outbreak of the virus, Li et al., (2020) proposed a metapopulation structure onto an SEIR-model incorporating travel between major cities in China and considered both documented and undocumented infections. Their model depicts that as many as 86% of all cases went undetected in Wuhan before travel restrictions took effect on January 23, 2020. According to their estimation, on a per person basis,

asymptomatic individuals were only 55% as contagious, yet were responsible for 79% of new infections, given their increased prevalence. The importance of accounting for asymptomatic individuals has been solidified by other studies Ferguson et al., (2020); Verity et al., (2020). Later, Calafiore et al., (2020), using a modified SIR-model, estimated that, on average, cases in Italy went underreported by a factor of 63, as of March 30, 2020. Anastassopoulou et al., (2020) apply the SIRD model to Chinese official statistics, estimating parameters using linear regression and predict the COVID-19 pandemic in Hubei province by using these models and parameters. Nonetheless, long-term forecasting is debatable while using a simple mathematical model. Diego Caccavo, (2020) and Peter Turchin, (2020) independently apply modified SIRD models, in which parameters change overtime following specific function forms. Parameters govern these functions are estimated by minimizing the sum-of-square-error. However, using the sum-of-square method causes over-fitting and always favors a complex model, therefore it is not suitable to access policy effectiveness. Moreover, fitting the SIRD model in the early stage of infection is questionable as well.

In the light of above shortcomings of several prior mathematical models, a more refined deterministic system of nonlinear differential equations, considering all possible interactions, a Susceptible-Exposed-Symptomatic Infectious-Asymptomatic Infectious-Quarantined-Hospitalized-Recovered-Dead (SEI_DI_UQHRD) model has been proposed, which can give more accurate and robust short-term as well as long-term predictions of COVID-19 future dynamics. This model could be considered as a generalization of SEIR, which is based on the introduction of asymptomatic infectious state, quarantined state in order to understand the effect of preventive actions and hospitalized (isolated) state. Some assumptions are also based on similar studies performed over SEIR dynamical systems Wu et al., (2020); Calafiore et al., (2020); Anastassopoulou et al., (2020). As of May 11, it is evident that how the early stage mathematical model parameters have changed drastically as detection rate was really low until middle of February, however the outbreak situation has improved comprehensively in several countries due to massive scale testing Qianying et al. (2020). We have considered the nominal values of the model parameters understanding the characteristics of the coronavirus infection, quantitatively estimated in the literature or published by health organizations Read et al., (2020); Li et al., (2020); Wu et al., (2020); Verity et al., (2020). In sequence, a rigorous process of model calibration is performed through the formulation and solution of an inverse problem. Our analysis was based on the publicly available data of the new confirmed daily cases reported for the future probable COVID-19 hotspots which are Russia, Brazil, India and Bangladesh from late January until May 09, 2020 Center for Systems Science and Engineering at Johns Hopkins University, (2020). Further, we have validated our model for the United Kingdom using same time frame. Based on the released data, we attempted to estimate the mean values of the crucial epidemiological parameters for COVID-19, such as the infectiousness factor for asymptotic carriers, isolation period, the estimation of the inflection point with probable date, recovery period for both symptomatic and asymptomatic individuals, case-fatality ratio and mortality rate in those potential hotspots. Perfect data-driven and curve-fitting methods for the prediction of any disease outbreak have always been question of interest in epidemiological research. Importantly, for an epidemic model to be truly useful it must undergo a judicious process of validation, which

consists in comparing model predictions with real-time data in order to evaluate if they are realistic and useful. In general, the first predictions of a model do not agree with the observations, possibly due to inadequacies in the model hypotheses or because of a poor choice of the model nominal parameters. This can invalidate the model, however model calibration can bring tremendous accuracy, where a set of parameters that promote a fine agreement between predictions and observations is sought. This approach could be efficient in order to have considerable insights on disease outbreak dynamics and to reach worthwhile conclusions in curtailing the disease burden. Furthermore, by calibrating the parameters of the proposed SEI_DI_UQHRD model to fit the reported data, we have also provided probable forecasts until late August.

This paper is organized as follows. In section 2, the mathematical model is described and estimation of nominal values for the model parameters is discussed. In section 3, the proposed model has been analyzed. In section 4, the forward and inverse problems are formulated and solved to calibrate the model. In section 5, the main contributions of this work are emphasized by giving a subsequent comparison between model predictions and real data. In section 6, one of the robust global sensitivity analysis techniques is used to quantify the most influential mechanisms in our model. This paper ends with some qualitative and quantitative observations and discussions.

2. Formulation of the Mathematical Model

In this work, we use a compartmental differential equation model for the spread of COVID-19 in the world. The spread of the infection starts with the introduction of a small group of infected individuals to a large population. In this model we focus our study on eight components of the epidemic flow, i.e. $\{S(t), E(t), I_D(t), I_U(t), Q(t), H(t), R(t), D(t)\}$ which indicate the number of the susceptible individuals, exposed individuals (infected however not yet to be infectious, in an incubation period), symptomatic infectious individuals (confirmed with infectious capacity), asymptomatic infectious individuals (undetected but infectious), quarantined, hospitalized (under treatment), recovered cases (immuned) and death cases (or closed cases). Then the entire number of population in a certain region or country is $N = S + E + I_D + I_U + Q + H + R + D$. The following assumptions are considered in the formulation of the model:

- Emigration from the population and immigration into the population have not been taken into account in model formulation as there is negligible proportion of individuals move in and out of the population at a specific time frame.
- Births and natural deaths in the population are not considered.
- The susceptible population are exposed to a latent class .
- Hospitalized patients cannot spread disease while in isolation treatment after confirmed diagnosis.
- Recovered individuals do not return to susceptible class as they develop certain immunity against the disease. Hence they cannot become re-infected again and cannot infect susceptible either.

Figure 1 is the flow diagram of the proposed model, where susceptible individuals (S) get infected at a baseline infectious contact rate β, via contact with either symptomatic infectious individuals (I_D) or at a rate $\beta\lambda$ of asymptomatic infectious carriers and move to exposed (E) class. Here λ is the relative infectiousness factor of asymptomatic carriers (in comparison to symptomatic individuals). The rate of getting home-quarantined or self-quarantined of S is q, whereas κ is the rate of progression of symptoms of COVID-19 (hence $\frac{1}{\kappa}$ is the mean incubation period of COVID-19). Detected symptomatic infected individuals are generated at a proportion σ_1 and undetected asymptomatically-infected cases are generated at a rate σ_2 from the exposed class. γ is the rate of confirmation and isolation after symptom onset in detected infectious individuals and $1/\gamma$ is the time period when the infection can spread. The rate of transition from quarantined to symptomatic infectious, asymptomatic infectious and susceptible are $r_2\eta$, $r_1\eta$ and $(1-r_1-r_2)\eta$ respectively. δ_U is the disease-induced death rate for undetected asymptomatic patients and δ_U is the death rate for hospitalized patients. ϕ_D, ϕ_U and ϕ_H are recovery rates for the symptomatic patients, asymptomatic carriers and hospitalized patients respectively. In addition, $1/\phi_D$, $1/\phi_U$ and $1/\phi_H$ represent the mean period of isolation for the detected symptomatic infectious individuals, undetected asymptomatic carriers and hospitalized patients. Based on these

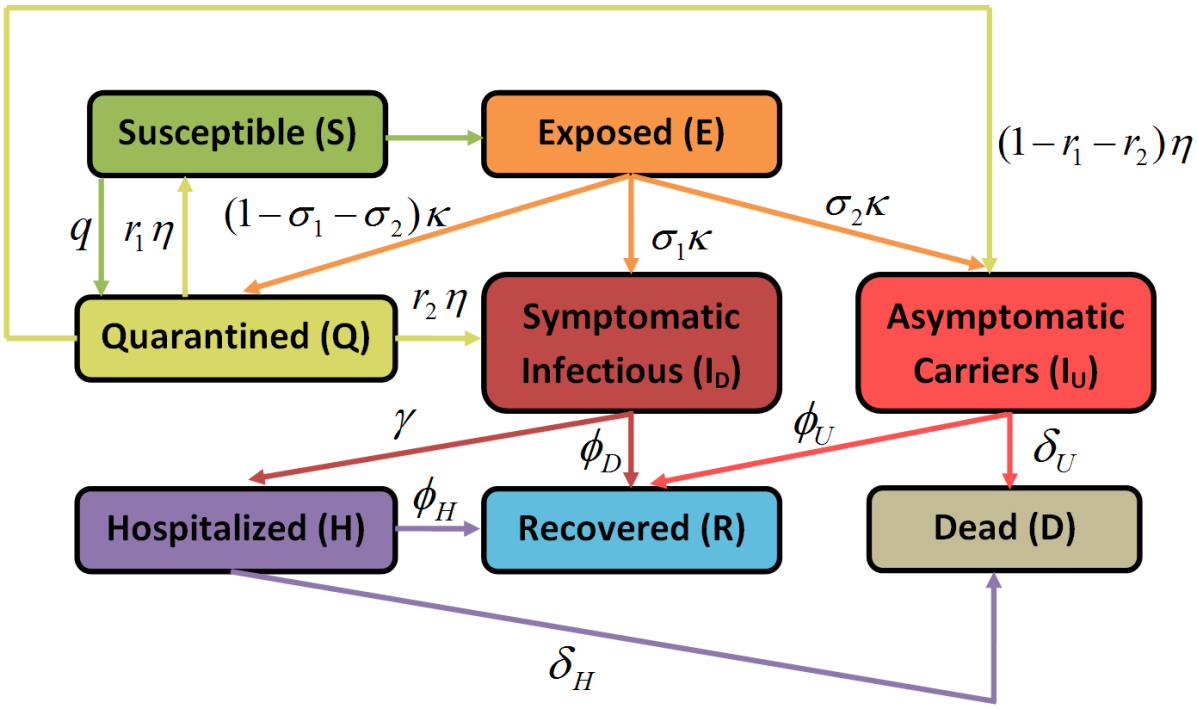


Figure 1: A schematic diagram that illustrates the proposed COVID-19 model

assumptions, the basic model structure for the transmission dynamics of COVID-19 is

illustrated by the following deterministic system of nonlinear differential equations (1).

$$
\begin{cases}
\dfrac{dS}{dt} &= -\beta S\dfrac{(I_D+\lambda I_U)}{N-D} + r_1\eta Q - qS,\\
\dfrac{dE}{dt} &= \beta S\dfrac{(I_D+\lambda I_U)}{N-D} - \kappa E,\\
\dfrac{dI_D}{dt} &= \sigma_1\kappa E - \gamma I_D - \phi_D I_D + r_2\eta Q,\\
\dfrac{dI_U}{dt} &= \sigma_2\kappa E - \phi_U I_U - \delta_U I_U + (1-r_1-r_2)\eta Q\\
\dfrac{dQ}{dt} &= (1-\sigma_1-\sigma_2)\kappa E - \eta Q + qS\\
\dfrac{dH}{dt} &= \gamma I_D - \phi_H H - \delta_H H\\
\dfrac{dR}{dt} &= \phi_D I_D + \phi_U I_U + \phi_H H\\
\dfrac{dD}{dt} &= \delta_U I_U + \delta_H H
\end{cases}
\tag{1}
$$

2.1. Baseline epidemiological parameters

The incubation period for COVID-19 is estimated to average 5.1 *days* Lauer et al., (2020), similar to other model-based estimates Li et al., (2020), giving $\kappa = 1/5.1\,day^{-1}$. Some previous model-based estimates of infectious duration are on the order of several days Li et al., (2020); Ferguson et al., (2020); Tang et al., (2020), with giving about 7 *days* for asymptomatic individuals to recover. However, the clinical course of the disease is typically much longer: In a study of hospitalized patients Verity et al., (2020), average total duration of illness until hospital discharge or death was 21 *days*, and moreover, the median duration of viral shedding was 20 *days* in survivors. The effective transmission rate (as a constant), β, ranges from around $0.2 - 1.5\,day^{-1}$ in prior modeling studies Shen et al. (2020); Read et al., (2020); Li et al., (2020), and typically trends down with time Li et al., (2020); Tang et al., (2020). We have left this as a free parameter in our fits. Baseline model parameters with brief description, likely ranges based on modeling and clinical studies, and default value chosen for the study is presented in the Table 1.

2.2. Dataset

The updated epidemic data of five different countries which are Russia, Brazil, India, Bangladesh and the United Kingdom is collected from authoritative and known sources. The data is gathered from Center of Disease Control and Prevention (CDC) and the COVID Tracking Project (testing and hospitalizations). The data repository is operated by the Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE) and Supported by ESRI Living Atlas Team and the Johns Hopkins University Applied Physics Lab (JHU APL). The repository is publicly available Center for Systems Science and Engineering at Johns Hopkins University, (2020).

Parameter	Definition	Units	Likely Range	Default Value	References
β	infectious contact rate	day^{-1}	$0.2 - 1.5$	0.5	Shen et al. (2020); Read et al., (2020); Li et al., (2020)
λ	infectiousness factor for undetected infected carrier	-	0.4-0.6	0.5	Li et al., (2020); Ferguson et al., (2020)
r_1	fraction of quarantined that become susceptible	-	0.7-0.99	0.9	Calibrated
r_2	fraction of quarantined that become detected symptomatic	-	0.005-0.1	0.02	Calibrated
η	transition from quarantined to either infectious or susceptible	day^{-1}	0.005-0.3	0.01	Calibrated
κ	transition exposed to infectious	day^{-1}	$\frac{1}{14} - \frac{1}{3}$	1/5.1	Li et al., (2020); Lauer et al., (2020)
q	transition susceptible to quarantined	day^{-1}	0.001-0.5	0.01	Calibrated
σ_1	fraction of infections that become detected symptomatic	-	0.0001-0.1	0.0032	Li et al., (2020); Ferguson et al., (2020); Verity et al., (2020)
σ_2	fraction of infections that become undetected symptomatic	-	0.01-0.95	0.5	Calibrated
γ	transition detected symptomatic to hospitalized	day^{-1}	0.2-0.9	0.5	Ferguson et al., (2020); Zhou et al., (2020)
ϕ_D	Recovery rate, detected symptomatic	day^{-1}	$\frac{1}{30} - \frac{1}{3}$	1/7	Tang et al., (2020); Zhou et al., (2020)
ϕ_U	Recovery rate, undetected symptomatic	day^{-1}	$\frac{1}{30} - \frac{1}{3}$	1/7	Tang et al., (2020); Zhou et al., (2020)
ϕ_H	Recovery rate, Hospitalized	day^{-1}	$\frac{1}{30} - \frac{1}{3}$	1/7	Tang et al., (2020); Zhou et al., (2020)
δ_U	Disease induced death rate, undetected symptomatic	day^{-1}	0.001-0.1	0.015	Ferguson et al., (2020)
δ_H	disease induced death rate, hospitalized	day^{-1}	0.001-0.1	0.015	Ferguson et al., (2020)

Table 1: Baseline model parameters with brief description, likely ranges based on model calibration and clinical studies, and default value chosen for the study

3. Analysis of the Model

3.1. Disease Free Equilibrium (DFE)

The disease-free equilibrium (DFE) of (1) can be obtained easily by setting $S = E = I_D = I_U = Q = H = R = D = 0$. Therefore the DFE of (1) is :

$$E_0 = (S_0^*, E_0^*, I_{D_0}^*, I_{U_0}^*, Q_0^*, H_0^*, R_0^*, D_0^*) = (0, 0, 0, 0, 0)$$

3.2. Basic Reproduction Number for Proposed Model

The local stability of the DFE is explored using the next generation operator method Van den Driessche et al. (2002). Using the notation in Diekmann et al., (1990), it follows that the associated non-negative matrix, $\mathcal{F}$, for the new infection terms, and the non-singular matrix, $\mathcal{V}$, for the remaining transfer terms, are given, respectively, by :

$$\mathcal{F} = \begin{pmatrix} 0 & \beta & \beta\lambda \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

and

$$\mathcal{V} = \begin{pmatrix} k & 0 & 0 \\ -\sigma_1 k & \gamma + \phi_D & 0 \\ -\sigma_2 k & 0 & \phi_U + \delta_U \end{pmatrix}$$

The associated basic reproduction number, denoted by $\mathcal{R}_0$ is then given by, $\mathcal{R}_0 = \rho\left(\mathcal{F}\mathcal{V}^{-1}\right)$, where ρ is the spectral radius of $\mathcal{F}\mathcal{V}^{-1}$. It follows that,

$$\mathcal{R}_0 = \frac{\beta\left[\delta_U\sigma_1 + (\gamma + \phi_D)\,\lambda\sigma_2 + \sigma_1\phi_U\right]}{(\gamma + \phi_D)\,(\delta_U + \phi_U)} \tag{2}$$

Thus, by Theorem 2 of LaSalle (1976), the following result is established.

Lemma 1. *The DFE, E_0 of the system (1), is locally-asymptotically stable (LAS) if $\mathcal{R}_0 < 1$, and unstable if $\mathcal{R}_0 \geq 1$*

The threshold quantity, $\mathcal{R}_0$, measures the average number of secondary cases generated by a single infected individual in a completely susceptible human population Hethcote (2000). The above result implies that a small influx of infected individuals would not generate large outbreaks if $\mathcal{R}_0 < 1$, and the disease will persist (be endemic) in the population if $\mathcal{R}_0 > 1$.

4. Calibration of the epidemic model

The proposed epidemic model (1), supplemented by an appropriate set of initial conditions, is a continuous-time dynamical system

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(x(t), \boldsymbol{p}), \quad \boldsymbol{x}(t_0) = \boldsymbol{x_0} \tag{3}$$

where $x(t) = (S(t), E(t), I_D(t), I_U(t), Q(t), H(t), R(t), D(t)) \in \mathbb{R}^8$ is the vector of states at time t, $x_0 = (S^i, E^i, I_D^i, I_U^i, Q^i, H^i, R^i, D^i) \in \mathbb{R}^8$ is a prescribed initial condition vector referring to the initial time t_0 of the analysis, the vector $\boldsymbol{p} =$

$(N, \beta, \lambda, r_1, r_2, \eta, k, \sigma_1, \sigma_2, \gamma, q, \phi_D, \phi_U, \phi_H, \delta_U, \delta_H) \in \mathbb{R}^{16}$ aggregates the model parameters and $\boldsymbol{f} : U \subset \mathbb{R}^8 \times \mathbb{R}^{16} \to \mathbb{R}^8$ is a nonlinear map which gives the evolution law of this dynamics, defined (for fixed t) on the open set

$$U = \left\{(\boldsymbol{x}(t), \boldsymbol{p}) \in \mathbb{R}^8 \times \mathbb{R}^{16} \mid x_m(t) > 0 \text{ and } p_n > 0, \text{ for } m = 1, 2, ..., 8 \text{and } n = 1, ..., 16\right\}$$

The forward problem consists in providing initial conditions (IC) and a set of parameters, represented by the pair $\alpha = (\boldsymbol{x_0}, \boldsymbol{p})$, and compute by means of numerical integration the model response $\boldsymbol{x}(t)$ from which a scalar observable $\boldsymbol{\psi}(\boldsymbol{\alpha}, t)$ is obtained. In the forward problem, $\boldsymbol{\alpha}$ represents all IC and system parameters from (1), while $\boldsymbol{\psi}(\boldsymbol{\alpha}, t)$ is the new cases $\mathcal{N}_w$ system response which can be calculated using daily reported cumulative cases real-time data.

Since the map $\boldsymbol{f}$ has a polynomial nature in $\boldsymbol{x}$, it is continuously differentiable function in $\boldsymbol{x}$. Thus, the existence and uniqueness theorem for ordinary differential equations guarantees that the initial value problem of (3) has a unique solution. In addition, it can also be showed that this solution is continuously dependent on $\boldsymbol{\alpha}$, as well as the *forward map* $\boldsymbol{\psi}$ Richard et al., (2012); Yaman et al., (2013).

In this work, the evaluation of the system response in the *forward problem* is solved numerically by using Runge-kutta (4,5) method and the scalar observable of interest $\mathcal{N}_w$ is used to assess the simulation when compared with real data of the COVID-19 outbreak made available by Center for Systems Science and Engineering at Johns Hopkins University, (2020). The referred data consists of new infected, recovered and death cases of COVID-19 per day.

4.1. Inverse Problem

The model calibration problem seeks to find a set of parameters that, to a certain degree, makes the model response as close as possible to the empirical observations (reference data), once, due to the errors on model conception and reference data acquisition, it is (practically) impossible for the *forward map* to reproduce outbreak observations.

The mathematical setting for this case considers the *parameters vector* $\boldsymbol{\alpha}$ defined in the *parameter space*, $E = \mathbb{R}^{16}$, since here $\boldsymbol{\alpha}$ comprises all IC and system parameters from (1), excepting N (entire population size), which is kept fixed in its nominal value. For the purpose of comparison between observations and predictions, a discrete set with M time-instants is considered, so that scalar observations and predictions are respectively aggregated into $\boldsymbol{y} = (y_1, y_2, ..., y_M)$ and $\boldsymbol{\psi}(\boldsymbol{\alpha}) = (\psi_1, \psi_2, ..., \psi_M)$, both defined in the *data space* $F = \mathbb{R}^M$. Note that the *forward map* $\boldsymbol{\psi} : E \to F$ associates to each parameters vector $\boldsymbol{\alpha}$ an observable vector $\boldsymbol{\psi}(\boldsymbol{\alpha})$ where the component represents the number of new cases in each day of the year, i.e., $\boldsymbol{\psi}_w = \boldsymbol{\mathcal{N}}_w$. In practice, the parameters vector is restricted to be on the convex set of *admissible values*, $C = \{\boldsymbol{\alpha} \in E \mid \mathbf{lb} \leq \boldsymbol{\alpha} \leq \mathbf{ub}\}$ in which $\mathbf{lb}$ and $\mathbf{ub}$ are lower and upper vector bounds for $\boldsymbol{\alpha}$, respectively.

In formal terms, given an *observation vector* $\boldsymbol{y} \in F$ and a *prediction vector* $\boldsymbol{\psi}(\boldsymbol{\alpha}) \in F$, the calibration aims at finding a vector of parameters $\boldsymbol{\alpha}^*$ such that

$$\boldsymbol{\alpha}^* = arg \min_{\boldsymbol{\alpha} \in C} J(\boldsymbol{\alpha})$$

for a misfit function

$$J(\boldsymbol{\alpha}) \;=\; \|\boldsymbol{y}-\boldsymbol{\psi}(\boldsymbol{\alpha})\|^2 = \;\left\{\sum_{m=1}^{M}\left|y_m-\psi_m(\boldsymbol{\alpha})\right|^2\right\}$$

This is the *inverse problem* associated to the epidemic model. In general this type of extremely nonlinear, with none or low regularity, multiple solutions (or even none), being much more complicated to crack in comparison with the forward problem Richard et al., (2012); Yaman et al., (2013).

The inverse problem attempts to estimate a finite number of parameters on a finite dimension space, being defined in terms of a typical nonlinear misfit function. Therefore, Theorem 4.5.1 of Chavent, (2010) can be invoked to guarantee a proper sense of well-posedness (existence, uniqueness, unimodality and local stability) for the inverse problem.

The Trust-Region-Reflective method (TRR) is employed here to numerically approximate a solution for the inverse problem. The core idea of the method is to minimize a simpler function that represents the behavior of $J(\boldsymbol{\alpha})$ in a neighborhood (trust-region) around $\boldsymbol{\alpha}$. The simpler function is defined as dependent on the trial step s, characterizing the Trust-Region subproblem, and its computation is optimized by restricting the subproblem to a two-dimensional subspace. The subspace is linear spanned by a multiple of the gradient $\boldsymbol{g}$ and (in the bounded case) a vector obtained in a scaled modified Newton step, used for the convergence condition $D(\boldsymbol{\alpha})^{-2}\boldsymbol{g}(\boldsymbol{\alpha})=0$, where D is a diagonal matrix that depends on $\boldsymbol{\alpha}$, $\boldsymbol{g}$, $\mathbf{lb}$, and $\mathbf{ub}$ Coleman, (1996). Eventually, the trial step is constructed through the subproblem as

$$s^* \;=\; arg\min_s\left\{\frac{1}{2}s^TQs+g^Ts \mid \|Ds\|_2\le\Delta\right\} \tag{4}$$

where Δ is a scalar associated with the trust region size; Q is a matrix involving D, a Jacobian matrix (also dependent on $\boldsymbol{\alpha}$, $\boldsymbol{g}$, $\mathbf{lb}$, and $\mathbf{ub}$) and an approximation of the Hessian matrix Coleman, (1996). The quadratic approximation in (4) has well-behaved solutions Andrew, (2000) and if $J(\boldsymbol{\alpha}+s)<J(\boldsymbol{\alpha})$ then $\boldsymbol{\alpha}$ is updated to $\boldsymbol{\alpha}+s$ and the process iterates, otherwise Δ is decreased. In addition, a reflection step also occurs if a given step intersects a bound: the reflected step is equivalent to the original step except in the intersecting dimension, where it assumes the opposite value after reflection.

The TRR algorithm also requires an initial guess for each parameter, identified in the next section as "TRR input". The stopping criteria are the norm of the step and the change in the value of the objective function, with a tolerance of 10^{-9}.

5. Numerical Experiments and Forecasting

In this section, we have used dynamically calibrated compartment model for real-time analysis and real-time prediction of COVID-19 outbreak for five different countries which are Russia, Brazil, India, Bangladesh and the United Kingdom. Figure 2, 4, 6, 8, 10 illustrates the best result for the $\mathcal{N}_w$ system response fitting problem using the nominal parameters from Table 1 as initial guesses for the TRR algorithm. The upper

Parameter	TRR input	**lb**	**ub**	TRR output	References
β	0.4	0.2	1.5	0.4852	Shen et al. (2020); Li et al., (2020)
λ	0.45	0.4	0.6	0.5249	Li et al., (2020); Ferguson et al., (2020)
r_1	0.76	0.1	0.999	0.9810	Estimated
r_2	0.03	0.01	0.9	0.0224	Estimated
η	0.05	0.01	0.9	0.0779	Estimated
κ	0.196	1/14	1/3	0.2443	Li et al., (2020); Lauer et al., (2020)
σ_1	0.01	0.0001	0.9	0.0067	Li et al., (2020); Ferguson et al., (2020)
σ_2	0.5	0.01	0.95	0.6894	Estimated
γ	0.3	0.1	0.9	0.7297	Ferguson et al., (2020); Zhou et al., (2020)
ϕ_D	0.1428	1/30	1/3	0.3332	Tang et al., (2020); Zhou et al., (2020)
ϕ_U	0.1428	1/30	1/3	0.0401	Tang et al., (2020); Zhou et al., (2020)
ϕ_H	0.1428	1/30	1/3	0.0433	Tang et al., (2020); Zhou et al., (2020)
δ_U	0.09	0.001	0.01	0.0013	Ferguson et al., (2020)
δ_H	0.09	0.001	0.01	0.0095	Ferguson et al., (2020)
q	0.015	0.01	0.5	0.0107	Estimated

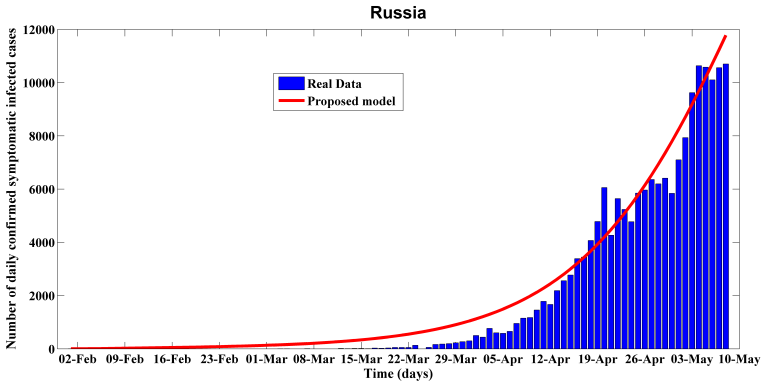
Table 2: TRR setup for the Figure 3 calibrated response

and lower bounds used for the parameters were set compatible with the literature suggested intervals and are also presented in 2, 3, 4, 5, 6 along with the parameters resulted from the calibration ("TRR output"). The minimum for S was set to $0.8\,N$ to establish a high number of susceptible individuals as expected for the beginning of an outbreak.

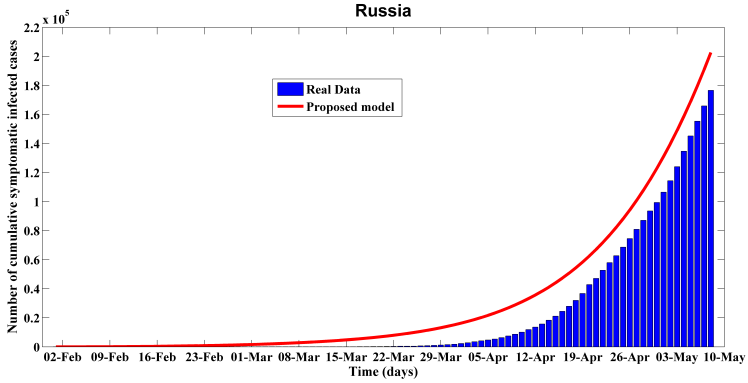
5.1. Analysis and prediction for Russia

With much of Europe now easing itself out of confinement, Russia could become the continent's new Covid-19 hotspot according to our analysis. The model fitting and projection results for Russia from early February to late August are shown in Figure 2 and 3. We took real data from February 01 to May 08, 2020 to calibrate the model parameters.

As we can see the results from the proposed model match the real data very well. Based on the proposed model, we want to project that from Figure 3, the number of daily detected symptomatic infectious cases in Russia will reach the peak at May 27 with about 15.774K cases. As time progresses, the projected error goes down and is limited about 10% for the cumulative cases and daily new cases according to our calculated daily projected mean error. The basic reproduction number is 4.234 as of May 08, which is close to the observed basic reproduction number for COVID-19, estimated about 2-7 for COVID-19 Liu et al., (2020). The number of cumulative infected cases is projected to reach 970K around August 30, and the estimated total death cases will reach to about 35K in the end. To date, Russia's official death toll of 1,827 is relatively low because not all deaths of people who have contracted the virus are being counted as Covid-19 deaths. The upper and lower bounds used for the parameters were set compatible with the literature suggested intervals and are presented in Table 2, along with the parameters and IC values resulted from the calibration ("TRR output").

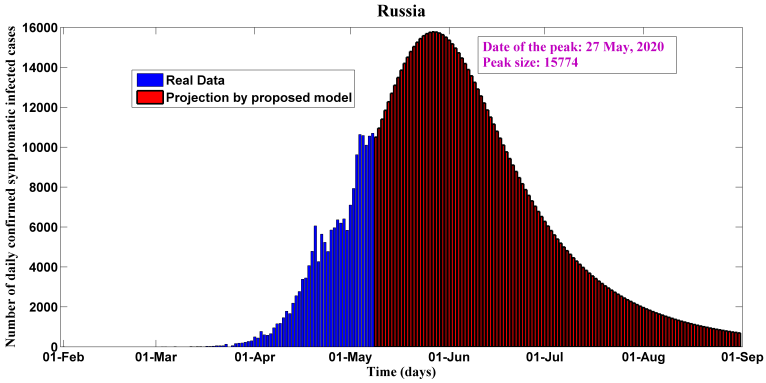


(a) The number of daily detected symptomatic infectious cases measured and fitted from early February to early May, 2020

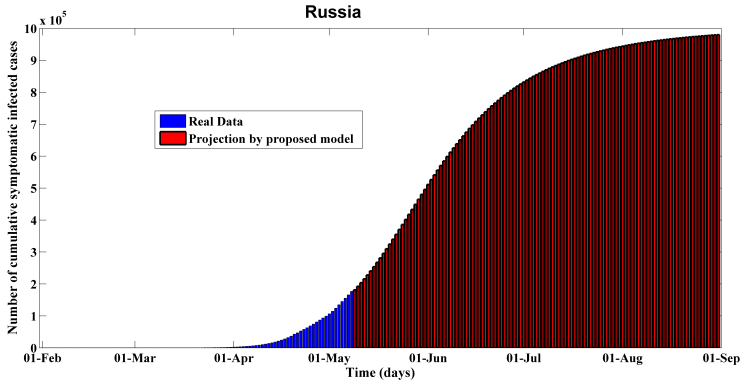


(b) The number of cumulative detected symptomatic infectious cases cases measured and fitted from early February to early May, 2020

Figure 2: Fitting performance of calibrated SEI_DI_UQHRD model for Russia from February 1 to May 08, 2020



(a) The number of daily detected symptomatic infectious cases measured and projected for Russia from early February to late August, 2020



(b) The number of cumulative detected symptomatic infectious cases measured and projected for Russia from early February to late August, 2020

Figure 3: Predictions of the proposed SEI_DI_UQHRD model for Russia from early March to late August, 2020

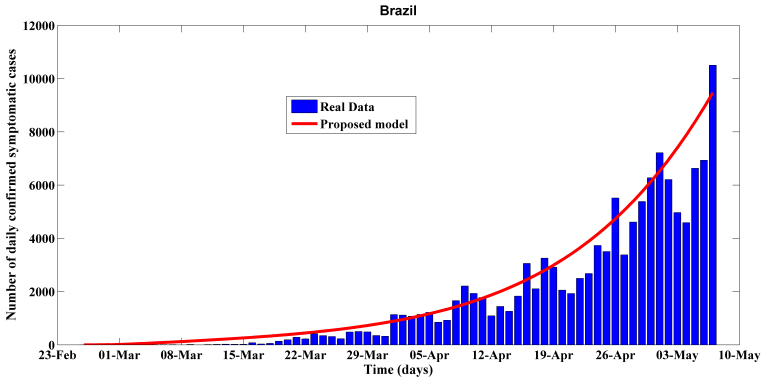
5.2. Analysis and prediction for Brazil

The coronavirus disease 2019 (COVID-19) pandemic reached Latin America later than other continents. The first case recorded in Brazil was on Feb 25, 2020. But now, Brazil has the most cases and deaths in Latin America (155939 cases and 10627 deaths as of May 9), and these are probably substantial underestimates. Our analysis projects that Brazil is developing as one of the world's next coronavirus hotspots. The model fitting and projection results for Brazil from late February to late August are shown in Figure 4 and 5. We took real data from February 25 to May 08 to calibrate the model parameters. As we can see the results from the proposed model fit the real data very well. Based on the proposed model, we project that from Figure 5, the number of daily detected symptomatic infectious cases in Brazil seem reaching the peak around June 11 with about 26.449K cases. The basic reproduction number is estimated about 5.3467 as of May 11, which is in between the observed basic reproduction number for COVID-19, estimated about 2-7 for COVID-19 Liu et al., (2020). This crucial epidemiological parameter could blow up because of scant diagnostics and impermanent non-pharmaceutical interventions. The case-fatality rate is hovering around 9.3% as of May 11, which is necessary to assess how much community transmission has occurred and its burden. According to our projection, this ratio could be doubled within two

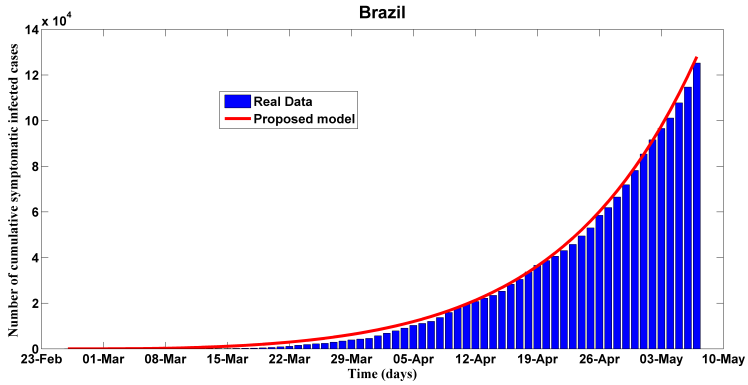
Parameter	TRR input	**lb**	**ub**	TRR output	References
β	0.3	0.2	1.5	0.5155	Shen et al. (2020); Li et al., (2020)
λ	0.45	0.4	0.6	0.5855	Li et al., (2020); Ferguson et al., (2020)
r_1	0.76	0.1	0.999	0.9855	Estimated
r_2	0.03	0.01	0.9	0.0159	Estimated
η	0.05	0.01	0.9	0.0897	Estimated
κ	0.196	1/14	1/3	0.1867	Li et al., (2020); Lauer et al., (2020)
σ_1	0.01	0.0001	0.9	0.0072	Moriarty et al., (2020); Verity et al., (2020)
σ_2	0.5	0.1	0.95	0.7308	Estimated
γ	0.3	0.1	0.9	0.7	Ferguson et al., (2020); Zhou et al., (2020)
ϕ_D	0.1428	1/30	1/3	0.3333	Tang et al., (2020); Zhou et al., (2020)
ϕ_U	0.1428	1/30	1/3	0.0356	Tang et al., (2020); Zhou et al., (2020)
ϕ_H	0.1428	1/30	1/3	0.0431	Tang et al., (2020); Zhou et al., (2020)
δ_U	0.09	0.001	0.01	0.001	Ferguson et al., (2020)
δ_H	0.09	0.001	0.01	0.009	Ferguson et al., (2020)
q	0.015	0.01	0.5	0.04	Estimated

Table 3: TRR setup for the Figure 5 calibrated response

months. The number of cumulative infected cases is projected to reach 1800K around August 30 if current trend is held, and the estimated total death cases will reach to about 108K in the end. The upper and lower bounds used for the parameters were set compatible with the literature suggested intervals and are presented in Table 3, along with the parameters and IC values resulted from the calibration ("TRR output").

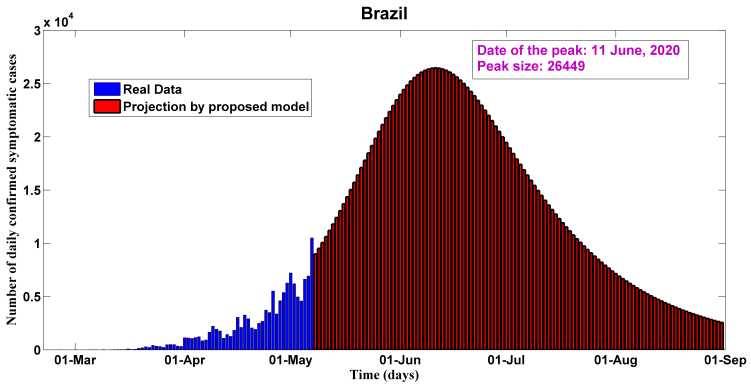


(a) The number of daily detected symptomatic infectious cases measured and fitted from late February to early May, 2020

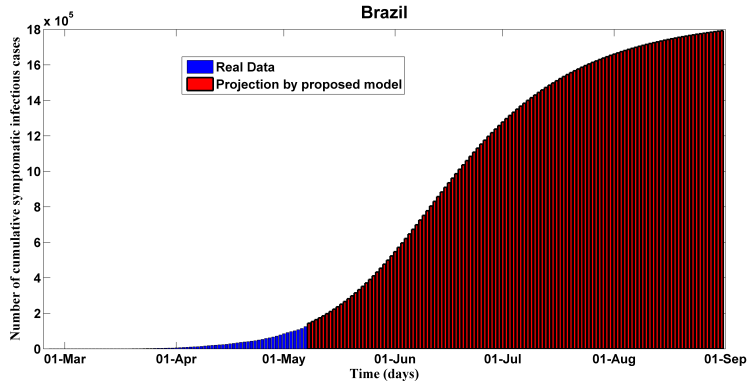


(b) The number of cumulative detected symptomatic infectious cases cases measured and fitted from late February to early May, 2020

Figure 4: Fitting performance of calibrated SEI_DI_UQHRD model for Brazil from 26 February to 8 May, 2020



(a) The number of daily detected symptomatic infectious cases measured and projected for Brazil from late February to late August, 2020



(b) The number of cumulative detected symptomatic infectious cases measured and projected for Brazil from late February to late August, 2020

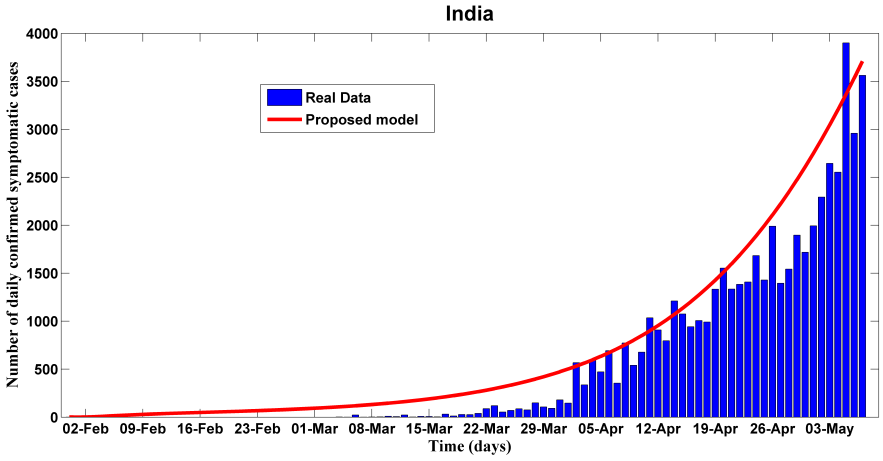
Figure 5: Predictions of the proposed SEI_DI_UQHRD model for Brazil from late February to late August, 2020

5.3. Analysis and prediction for India

Some people might think about the possibilities of the presence of a less virulent strain of the virus in India, along with the possibility that its hot weather could diminish the contagion. Nevertheless, our mathematical analysis suggests that India could become the new Covid-19 hotspot in South Asia. According to our projection, India could continue seeing spikes in the number of cases as it time progresses despite of following lockdown and other disease mitigation measures. The first case of the COVID-19 pandemic in India was reported on 30 January 2020, originating from China. But now, India has the most cases and deaths in South Asia (62,808 cases and 2,101 deaths as of May 9), and these are probably substantial underestimates. Figure 6 and 7 illustrate COVID-19 disease modeling and prediction for India from late January to late August. We took real-time data from January 30 to May 8 to calibrate the model parameters. As we can see, our proposed model match well for the historical real data. Based on our prediction from Figure 7, the number of daily detected symptomatic infectious cases in India seem reaching the peak around June 15 with about 9.504K cases. The basic reproduction number is 5.218 as of May 11, which lies between the studied observations Liu et al., (2020). This estimation may be considered as an overestimation of this critical parameter. Notwithstanding, this is owing to the infectiousness factor of the asymptomatic spreaders. The asymptomatic and minimally or moderately symptomatic should be quarantined to avoid the spread of the virus, with the severely symptomatic managed and isolated in health care settings. Contract tracing could be carried out to find those at risk of being in the incubation period by virtue of their exposure. The number of cumulative infected cases is projected to reach 730K around August 30 if current pattern is held, and the estimated total death cases will reach to about 43.8K in the end. The upper and lower bounds used for the parameters were set compatible with the literature suggested intervals and are presented in Table 4, along with the parameters values resulted from the calibration ("TRR output"). Interestingly, we have found in our analysis that India's case-fatality rate is at 3.5% and the country's recovery rate is at 33% which commensurate with real reported statistics precisely.

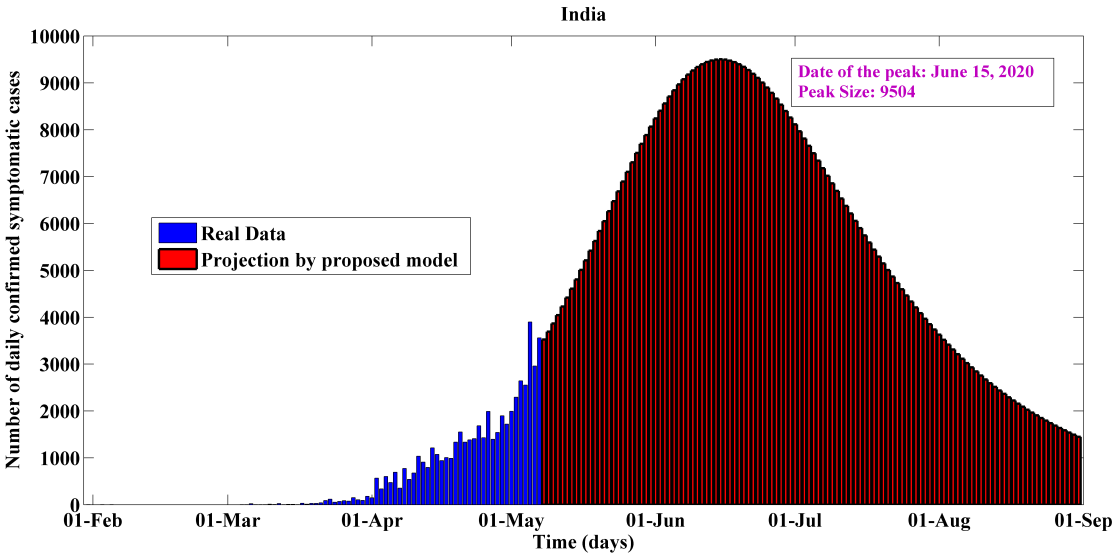
Parameter	TRR input	**lb**	**ub**	TRR output	References
β	0.3	0.2	1.5	0.3995	Shen et al. (2020); Li et al., (2020)
λ	0.4	0.4	0.6	0.5995	Li et al., (2020); Ferguson et al., (2020)
r_1	0.76	0.1	0.999	0.9911	Estimated
r_2	0.03	0.01	0.9	0.01	Estimated
η	0.05	0.01	0.9	0.0997	Estimated
κ	0.196	1/14	1/3	0.3296	Li et al., (2020); Lauer et al., (2020)
σ_1	0.01	0.0001	0.1	0.00046	Moriarty et al., (2020); Verity et al., (2020)
σ_2	0.5	0.1	0.95	0.7495	Estimated
γ	0.3	0.1	0.9	0.8	Ferguson et al., (2020); Zhou et al., (2020)
ϕ_D	0.1428	1/30	1/3	0.3333	Tang et al., (2020); Zhou et al., (2020)
ϕ_U	0.1428	1/30	1/3	0.0334	Tang et al., (2020); Zhou et al., (2020)
ϕ_H	0.1428	1/30	1/3	0.3335	Tang et al., (2020); Zhou et al., (2020)
δ_U	0.09	0.001	0.01	0.001	Ferguson et al., (2020)
δ_H	0.09	0.001	0.01	0.01	Ferguson et al., (2020)
q	0.015	0.01	0.5	0.0507	Estimated

Table 4: TRR setup for the Figure 7 calibrated response

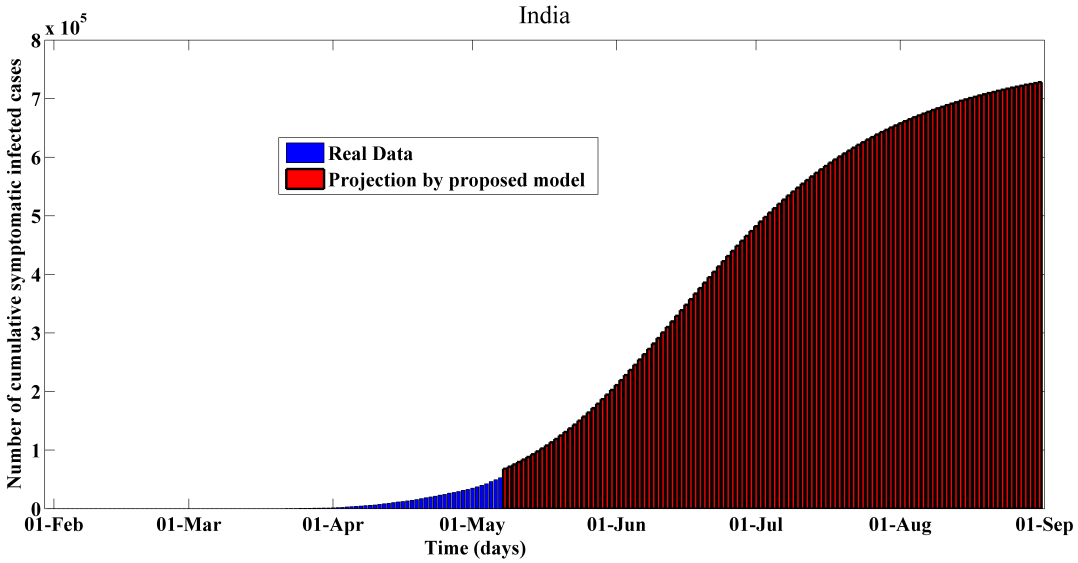


(a) The number of daily detected symptomatic infectious cases measured and fitted from early March to early May, 2020

Figure 6: Fitting performance of calibrated SEI_DI_UQHRD model for India from 8 March to early May, 2020



(a) The number of daily detected symptomatic infectious cases measured and projected for India from late January to late August, 2020



(b) The number of cumulative detected symptomatic infectious cases measured and projected for India from early March to late August, 2020

Figure 7: Predictions of the proposed SEI_DI_UQHRD model for India from late January to late August, 2020

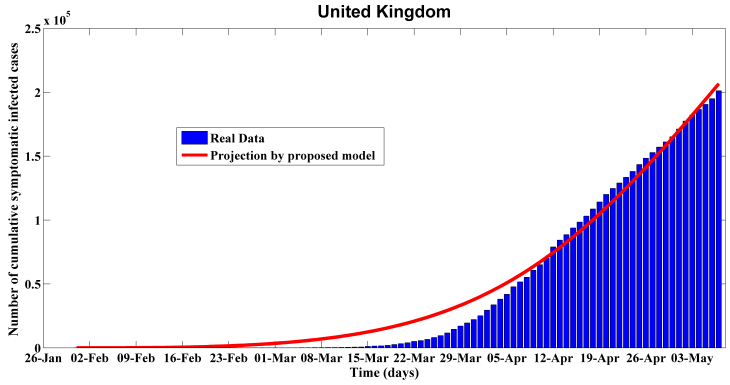
5.4. Analysis and prediction for the United Kingdom

Our analysis indicates that the United Kingdom could face risk of a second wave of coronavirus infections as the United Kingsom gradually eases a nationwide lockdown. According to Worldometer, (2020), by 11 May there had been 223,060 confirmed cases and 32,065 deaths overall a rate of 465 deaths per million population . The outbreak in London has the highest number and highest rate of infections, while England and Wales are the UK countries with the highest recorded death rate per capita. The prime minister of the UK is expected to unveil a coronavirus warning system when he outlines his plans to gradually ease the lockdown while dropping the "stay home"

Parameter	TRR input	**lb**	**ub**	TRR output	References
β	0.3	0.2	1.5	0.73	Shen et al. (2020); Li et al., (2020)
λ	0.4	0.4	0.6	0.59	Li et al., (2020); Ferguson et al., (2020)
r_1	0.76	0.6	0.999	0.97	Estimated
r_2	0.03	0.01	0.1	0.024	Estimated
η	0.05	0.04	0.1	0.09	Estimated
κ	0.196	1/14	1/3	0.3333	Li et al., (2020); Lauer et al., (2020)
σ_1	0.01	0.0001	0.01	0.0044	Moriarty et al., (2020); Verity et al., (2020)
σ_2	0.5	0.3	0.95	0.37	Estimated
γ	0.3	0.1	0.9	0.79	Ferguson et al., (2020); Zhou et al., (2020)
ϕ_D	0.1428	1/30	1/3	0.3333	Tang et al., (2020); Zhou et al., (2020)
ϕ_U	0.1428	1/30	1/3	0.0542	Tang et al., (2020); Zhou et al., (2020)
ϕ_H	0.1428	1/30	1/3	0.0333	Tang et al., (2020); Zhou et al., (2020)
δ_U	0.09	0.001	0.01	0.001	Ferguson et al., (2020)
δ_H	0.09	0.001	0.01	0.001	Ferguson et al., (2020)
q	0.015	0.01	0.5	0.015	Estimated

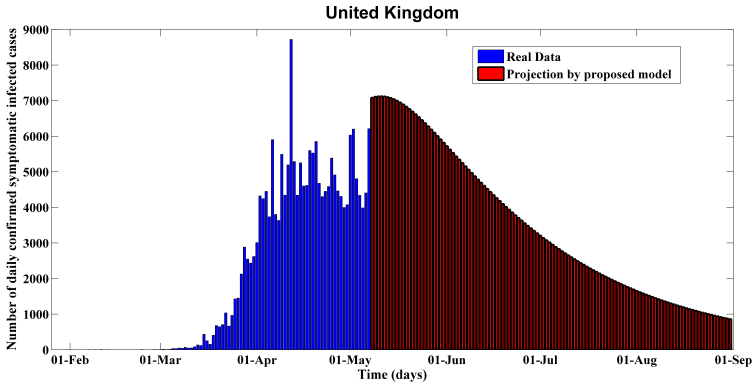
Table 5: TRR setup for the Figure 9 calibrated response

slogan. For the UK, the modeling and projection results from January 31 to August 31 are shown in Figure 8 and 9. As we can see from Figure 9, the number of daily detected symptomatic infectious cases reached the peak around April 10. Since then the curve has been maintaining a plateau which is a really unusual scenario. This phenomenon has closely been captured by our proposed model. This case study was really important for the validation of our model. In fact, this elucidates the fact that this model can give more precise and robust short-term predictions of COVID-19 dynamics. According to our calculation, the basic reproduction number is around 4.649 as of May 09, which is close to the observed basic reproduction number for COVID-19, estimated about 2-7 for COVID-19 Liu et al., (2020). The number cumulative infected cases is projected to reach 618K around August 30, and the estimated total death cases will reach to about 63.48K in the above mentioned period. Importantly, as of May 11, we have found in our analysis that the UK's case-fatality rate is at 17.2%, which is the worst among our five studied cases. This could exacerbate as time progresses in the absence of proven effective therapy or a vaccine. In addition, Our study suggests that relaxing social distancing too quickly could cause thousands of additional death in the UK. The upper and lower bounds used for the parameters were set compatible with the literature suggested intervals and are presented in Table 5, along with the parameters values resulted from the calibration ("TRR output").

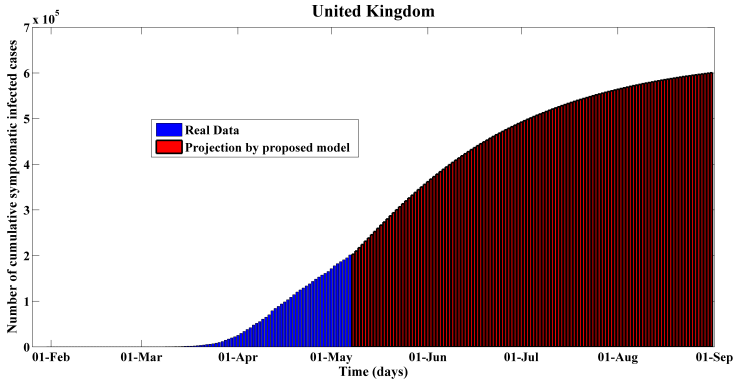


(a)

Figure 8: Fitting performance of calibrated SEI_DI_UQHRD model for the United Kingdom from 31 January to early May, 2020



(a) The number of daily detected symptomatic infectious cases measured and projected for the United Kingdom from late January to late August, 2020



(b) The number of cumulative detected symptomatic infectious cases measured and projected for India from late January to late August, 2020

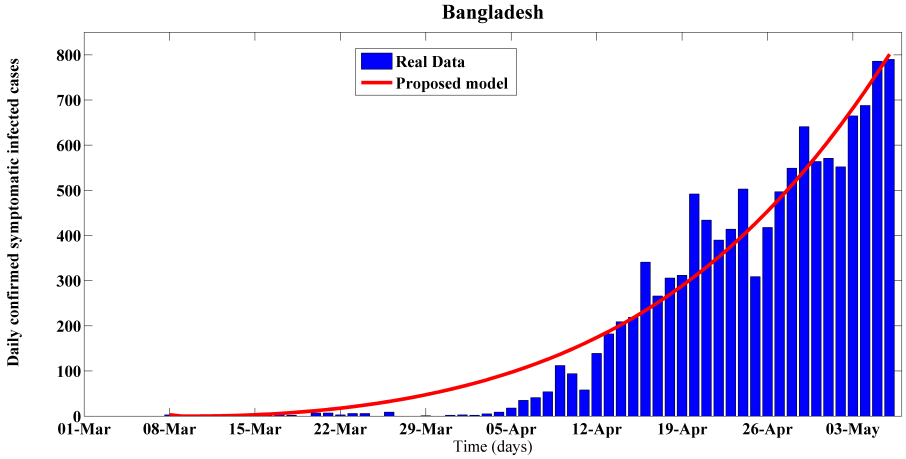
Figure 9: Predictions of the proposed SEI_DI_UQHRD model for the United Kingdom from late January to late August, 2020

Parameter	TRR input	**lb**	**ub**	TRR output	Reference
β	0.3	0.2	1.5	0.3972	Shen et al. (2020); Li et al., (2020)
λ	0.4	0.4	0.6	0.5498	Li et al., (2020); Ferguson et al., (2020)
r_1	0.76	0.1	0.999	0.9454	Estimated
r_2	0.03	0.01	0.9	0.02	Estimated
η	0.05	0.01	0.9	0.06	Estimated
κ	0.196	1/14	1/3	0.0714	Li et al., (2020); Lauer et al., (2020)
σ_1	0.01	0.0001	0.9	0.001	Moriarty et al., (2020); Verity et al., (2020)
σ_2	0.5	0.3	0.98	0.95	Estimated
γ	0.3	0.1	0.9	0.5	Ferguson et al., (2020); Zhou et al., (2020)
ϕ_D	0.1428	1/30	1/3	0.3333	Tang et al., (2020); Zhou et al., (2020)
ϕ_U	0.1428	1/30	1/3	0.0333	Tang et al., (2020); Zhou et al., (2020)
ϕ_H	0.1428	1/30	1/3	0.3275	Tang et al., (2020); Zhou et al., (2020)
δ_U	0.09	0.001	0.01	0.0011	Ferguson et al., (2020)
δ_H	0.09	0.001	0.01	0.01	Ferguson et al., (2020)
q	0.015	0.01	0.5	0.012	Estimated

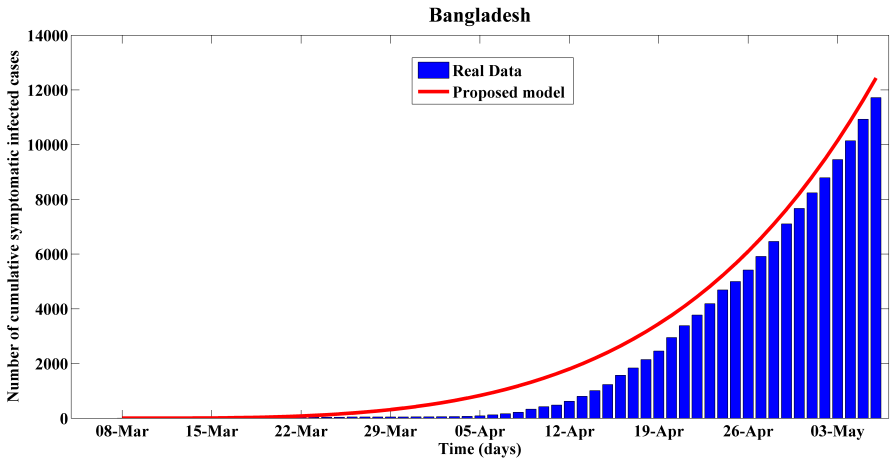
Table 6: TRR setup for the Figure 11 calibrated response

5.5. Analysis and prediction for Bangladesh

For Bangladesh, the modeling and projection results from early March to late August of are shown in Figure 10 and Figure 11. On 22 March, Bangladesh imposed a nationwide lockdown effective from 26 March to curb the spread of the novel coronavirus in the wake of four deaths and at least 39 infections. Notice that, there are some jumps in the number of confirmed daily new infected cases data from 16 april to 20 april due to the increase of limited testing system (from 2000 samples to 2700 samples per day) Institute of Epidemiology, Disease Control and Research (IEDCR). Bangladesh is still struggling (around 6700 samples per day) Institute of Epidemiology, Disease Control and Research (IEDCR) to incorporate massive scale testing as of May 11, 2020. Despite the fact, as we can see from Figure 10, our proposed model fits well for the historical data. As time progresses, the projected error goes down and is limited about 10% for the cumulative cases and daily new cases according to our calculated daily projected mean error. Though the estimated case-fatality rate in Bangladesh is at 2.5% which is really, without massive scale testing this rate could rise sharply incoming days. Moreover, the country's estimated recovery rate is at 24% which complements the real statistics precisely. The upper and lower bounds used for the parameters were set compatible with the literature suggested intervals and are presented in Table 6, along with the values of the model parameters resulted from the calibration ("TRR output").



(a) The number of daily detected symptomatic infectious cases measured and fitted from early March to early May, 2020

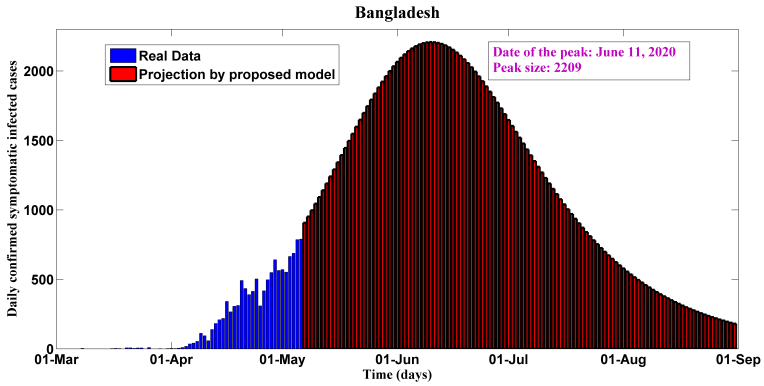


(b) The number of cumulative detected symptomatic infectious cases cases measured and fitted from early March to early May, 2020

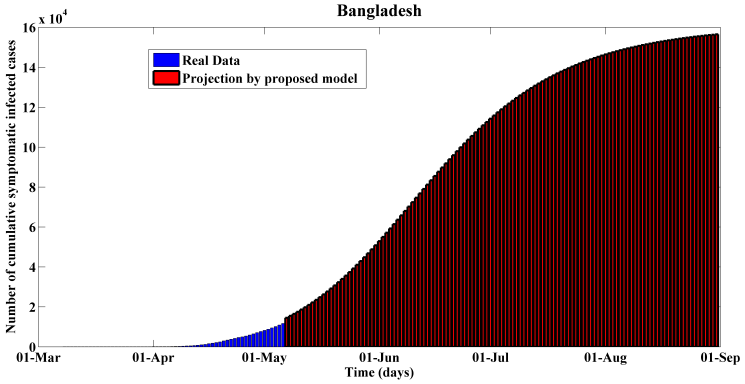
Figure 10: Fitting performance of calibrated SEI_DI_UQHRD model for Bangladesh from 8 March to early May, 2020

On 4 May, Bangladesh authorities intended to open up more factories, shopping malls and logistics operations, as they could diminish the economic impact of a coronavirus lockdown which they extended to May 16 Ministry of Health & Family Welfare of Bangladesh, (2020). This ease of lockdown could worsen the ongoing community transmission drastically. As we can see form Figure 11, the daily detected symptomatic infected cases (confirmed) will reach the peak (about 2209) around June 11, 2020 and then start to de-escalate. However, the probable peak time could occur in no time due to easing the coronavirus lockdown too quickly and this could bring a second wave of infections in the outbreak after post-peak period. The effective reproduction number is around 3.5 as of May 11, which is close to the observed basic reproduction number for COVID-19, estimated about 2-7 for COVID-19 Liu et al., (2020). Our estimation is relatively high which due to the fact that we have considered the infectiousness factor of the asymptomatic carriers. Massive level of testing is highly required to identify the asymptomatic spreaders. An aggressive test-and-isolate approach could curtail the burden of COVID-19, unlike other tactics like reporting contacts, wearing masks, and

maintaining social distance. Otherwise, this basic reproduction number could increase upto 5.7 within 20 days. Again, we want to express that such estimation is subject to differ due to various changing factors such as mitigation policies and people awareness in the near term.



(a) The number of daily detected symptomatic infectious cases measured and projected for Bangladesh from early March to late August, 2020



(b) The number of cumulative detected symptomatic infectious cases measured and projected for Bangladesh from early March to late August, 2020

Figure 11: Predictions of the proposed SEI_DI_UQHRD model for Bangladesh from early March to late August, 2020

6. Global Sensitivity Analysis

PRCC analysis which is a global sensitivity analysis method that calculates the partial rank correlation coefficient for the model inputs (sampled by Latin hypercube sampling method) and outputs Marino et al., (2008); Blower et al., (1994); Nabi et al., (2020). The PRCC method assumes a monotonic relationship between the model input parameters and the model outputs.

The calculated PRCC values are between -1 and 1 and they are comparable among different model inputs. The sign of the PRCC values shows the qualitative relationship between the model input and model output. A positive PRCC value implies that when the corresponding model input increases, the model output will also increase. A negative PRCC value indicates a negative correlation between the model input and output. The magnitude of the PRCC sensitivity measures the significance of the model input in contributing to the model output.

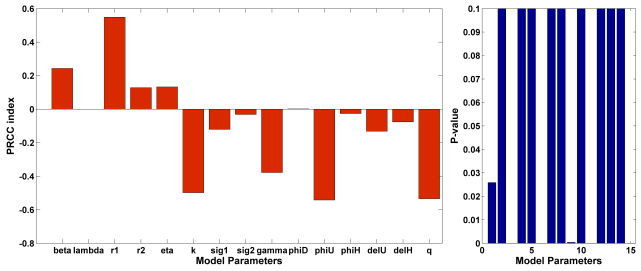
As our proposed epidemic model contains a moderate number of empirical parameters, uncertainty analysis can give considerable insights regarding the quantitative relationship between model responses and model input parameters. However, it is really challenging for complex models to determine the relationship with sufficient accuracy. Importantly, we have got startling yet realistic results from our sensitivity analysis. As we can see from Figure 12, we found nearly the same qualitative and significant quantitative relationship between the number of symptomatic infectious individuals (one of the crucial model responses) and three parameters which are rate of getting quarantined of susceptible individuals (q), transition rate from exposed to infectious or quarantined (κ) which can also be referred as the inverse of the average incubation period of COVID-19 and recovery rate of undetected asymptomatic (undetected) infectious carriers for our proposed model.

In case of Russia, from Figure 12a recovery rate of undetected asymptomatic carriers (ϕ_U) is the most negatively influential parameter when the number of detected infectious individuals is our selected model response. The PRCC index is found to be -0.585. In addition, rate of entering into home-quarantine or self-quarantine of susceptible individuals (q) and the fraction of wrongly quarantined people who become susceptible after certain latent period (r_1) are the other two influential empirical features with PRCC indexes are -0.5672 and 0.567 respectively.

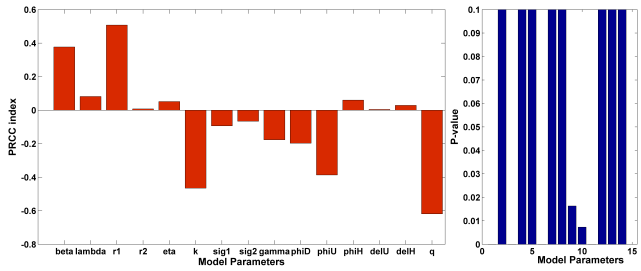
In case of Brazil, rate of getting home-quarantined or self-quarantined of susceptible individuals (q), the fraction of quarantined people who become susceptible due to avoiding home-quarantine (r_1), and the inverse of the COVID-19 mean latent period are the most influential parameter on the symptomatic infectious population size ($I_D(t)$). The corresponding PRCC indices are -0.62, 0.53 and -0.499. The figure 12b qualitatively elucidates that a high quarantine rate of the susceptible individuals can curtail the number of symptomatic infected individuals. In broader view, high efficacy of home or self-quarantine and a lower unlockdown rate could flatten the $I_D(t)$ curve in Brazil.

In case of the India, rate of getting home-quarantined or self-quarantined of susceptible individuals (q) and the inverse of the COVID-19 incubation period are the most influential parameter on the symptomatic infectious population size ($I_D(t)$) and the corresponding PRCC indexes are -0.64 and -0.62. Both the parameters are negatively sensitive to the size of detected infected individuals which is illustrated in the figure 12c. This elucidates that a higher quarantine rate of the susceptible individuals can reduce the number of symptomatic infected individuals. Therefore, it is obvious that early unlockdown in India could worsen the outbreak situation in the blink of an eye.

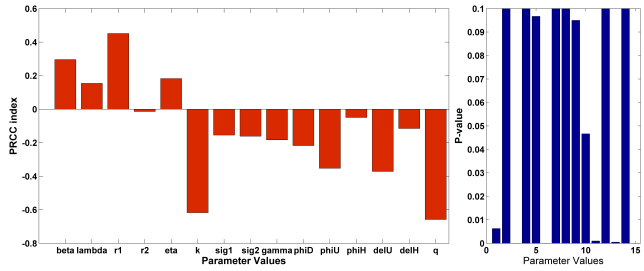
In case of the UK, rate of getting home-quarantined or self-quarantined of susceptible individuals (q) and the inverse of the COVID-19 incubation period are the most influential parameter on the symptomatic infectious population size ($I_D(t)$) and the corresponding PRCC indices are -0.58 and -0.499. Both the parameters are negatively sensitive to the size of detected infected individuals which is illustrated in the figure 12d. This elucidates that a higher quarantine rate of the susceptible individuals can curtail the number of symptomatic infected individuals. Nevertheless, early easing lockdown measures could bring a second wave of infection in the UK in no time.



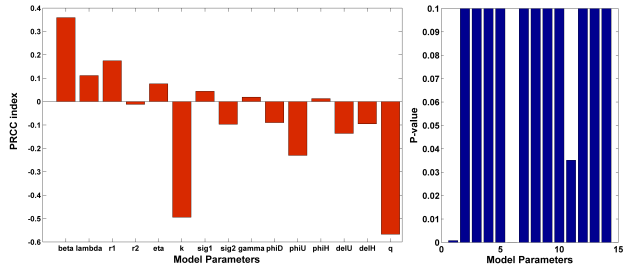
(a) Sensitivity analysis for Russia



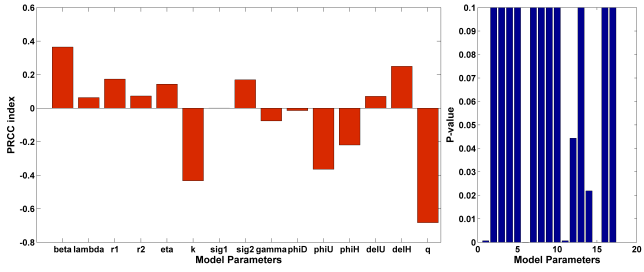
(b) Sensitivity analysis for Brazil



(c) Sensitivity analysis for India



(d) Sensitivity analysis for the UK



(e) Sensitivity analysis for Bangladesh

Figure 12: Sensitivity of the symptomatic infected cases while changing parameters in the proposed SEI_DI_UQRHD model as indicated by the PRCC index for five different countries

For Bangladesh, it has been found in our analysis, home-quarantine rate is the most negatively sensitive parameter on the size of symptomatic infectious individuals and the corresponding PRCC index is -0.68 which is the highest among our five studied cases. The result in Figure 12e elucidates that if the people in Bangladesh start breaking social distancing restrictions extensively, then it would be difficult to control the disease outbreak there. As of May 11, when Bangladesh is seeing continuous spikes in COVID-19 infections every day, the government has announced an easing of movement for all on a limited scale with a view to reviving the country's economy. Moreover, the length of the inverse of the latent period is the another negatively sensitive parameter with PRCC index lying at -0.43.

7. Concluding Remarks

We have proposed a methodology for the estimation of the key epidemiological parameters as well as forecasting of the spread of the COVID-19 epidemic in Russia, Brazil, India, Bangladesh and the UK while considering publicly available data from early January 2020 to early May 2020 with the introduction of an real-time differential SEI_DI_UQHRD epidemic model which can give more accurate and robust short-term predictions. Nominal quantities for the parameters are selected from the related literature concerning the COVID-19 infection. The calibration process is done through the solution of an inverse problem with the aid of a Trust-Region-Reflective method, used to pick the best parameter values that would fit the model response for the number of new infectious cases per week into the disease's empirical data. Results within realistic values for the parameters are presented, stating reasonable descriptions with the curve shape similar to the outbreak.

Numerical results on the recent COVID-19 data from Russia, Brazil, India, Bangladesh and the UK have been analyzed. Based on the projection as of May 11, 2020, Russia will reach the peak in terms of daily infected cases and death cases around end of May. Brazil and India will reach the peak in terms of daily infected cases and death cases around beginning of June. Based on the projection the United Kingdom may have a second wave of infection provided that lockdown is lifted. To end the pandemic, it is mandatory to actively seek out potential spreaders, particularly the asymptomatic spreaders which requires massive scale testing. Such a level of testing would effectively bring the R_0 below 1 in the above mentioned countries which will eventually cause the pandemic to fade out.

References

Shen, M., Peng, Z., Xiao, Y., & Zhang, L. (2020). Modelling the epidemic trend of the 2019 novel coronavirus outbreak in China. bioRxiv.

Read, J. M., Bridgen, J. R., Cummings, D. A., Ho, A., & Jewell, C. P. (2020). Novel coronavirus 2019-nCoV: early estimation of epidemiological parameters and epidemic predictions. MedRxiv.

Li, R., Pei, S., Chen, B., Song, Y., Zhang, T., Yang, W., & Shaman, J. (2020). Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2), *Science*, 368(6490), 489-493.

Ferguson, N., Laydon, D., Nedjati Gilani, G., Imai, N., Ainslie, K., Baguelin, M., & Dighe, A. (2020). Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand.

Lauer, S. A., Grantz, K. H., Bi, Q., Jones, F. K., Zheng, Q., Meredith, H. R., & Lessler, J. (2020). The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. *Annals of internal medicine*, 172(9), 577-582.

Tang, B., Bragazzi, N. L., Li, Q., Tang, S., Xiao, Y., & Wu, J. (2020). An updated estimation of the risk of transmission of the novel coronavirus (2019-nCov). *Infectious Disease Modelling*, 5, 248–255.

Zhou, F., Yu, T., Du, R., Fan, G., Liu, Y., Liu, Z., & Guan, L. (2020). Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *The Lancet*, 395 (10229), 1054-1062.

Moriarty, L. F., Plucinski, M., M., Marston, & B. J. et al. (2020). Public Health Responses to COVID-19 Outbreaks on Cruise Ships-Worldwide, February-March 2020. Morbidity and Mortality Weekly Report, 69, 347-352.

Wu, J. T., Leung, K., & Leung, G. M. (2020). Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *The Lancet*, 395(10225), 689-697.

Calafiore, G. C., Novara, C., & Possieri, C. (2020). A Modified SIR Model for the COVID-19 Contagion in Italy. arXiv preprint arXiv: 2003.14391.

Kucharski, A. J., Russell, T. W., Diamond, C., Liu, Y., Edmunds, J., Funk, S., & Davies, N. (2020). Early dynamics of transmission and control of COVID-19: a mathematical modelling study. *The Lancet Infectious Diseases*, 20(5), 553-558.

Simha, A., Prasad, R. V., & Narayana, S. (2020). A simple Stochastic SIR model for COVID 19 Infection Dynamics for Karnataka: Learning from Europe. arXiv preprint arXiv:2003.11920.

Anastassopoulou, C., Russo, L., Tsakris, A., & Siettos, C. (2020). Data-Based Analysis, Modelling and Forecasting of the novel Coronavirus (2019-nCoV) outbreak. *PLOS ONE*, 15(3), e0230405.

Nesteruk, I. (2020). Statistics based predictions of coronavirus 2019-nCoV spreading in mainland China. MedRxiv.

Chang, S. L., Harding, N., Zachreson, C., Cliff, O. M., & Prokopenko, M. (2020). Modelling transmission and control of the COVID-19 pandemic in Australia. arXiv preprint arXiv:2003.10218.

Wilder, B., Charpignon, M., Killian, J. A., Ou, H. C., Mate, A., Jabbari, S., & Majumder, M. S. (2020). The Role of Age Distribution and Family Structure on COVID-19 Dynamics: A Preliminary Modeling Assessment for Hubei and Lombardy. Available at SSRN 3564800.

Ruiz Estrada, M. A., & Koutronas, E. (2020). The Networks Infection Contagious Diseases Positioning System (NICDP-System): The Case of Wuhan-COVID-19. Available at SSRN 3548413.

Center for Systems Science and Engineering at Johns Hopkins University. (2020). COVID-19. Github Repository. https://github.com/CSSEGISandData/COVID-19. Last accessed May 08, 2020.

Verity, R., Okell, L. C., Dorigatti, I., Winskill, P., Whittaker, C., Imai, N., & Dighe, A. (2020). Estimates of the severity of coronavirus disease 2019: a model-based analysis. *The Lancet Infectious Diseases*.

Diego Caccavo, Chinese and Italian COVID-19 outbreaks can be correctly described by a modified SIRD model, MedRxiv.

Peter Turchin. Analyzing covid-19 data with sird models, (2020). https://github.com/pturchin/CSH-Covid-19-Project/blob/master/Turchin_2020_Covid19.pdf

Hethcote, H. W., (2000). The Mathematics of Infectious Diseases, *SIAM Review*, 42(4), 599–653.

LaSalle, JP., (1976). The Stability of Dynamical Systems,*SIAM*, Regional Conference Series in Applied Mathematics.

Van den Driessche, P., & Watmough, J., (2002). Reproduction numbers and subthreshold endemic equilibria for compartmental models of disease transmission, *Mathematical biosciences*, 180(1-2), 29–48.

Diekmann, O., Heesterbeek, J. A. P., & Metz, J. A., (1990). On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations, *Journal of mathematical biology*, 28(4), 365–382.

582 Liu, Y., Gayle, A. A., Wilder-Smith, A. , & Rocklov, J., (2020). The reproductive
583 number of COVID-19 is higher compared to SARS coronavirus. *Journal of Travel*
584 *Medicine*, 27(2).

585 Marino, S., Hogue, I. B., Ray, C. J., and Kirschner, D. E., (2008). A methodology for
586 performing global uncertainty and sensitivity analysis in systems biology, *Journal of*
587 *theoretical biology*, 254 (1), 178–196.

588 Blower, S. M., Dowlatabadi, H., (1994). Sensitivity and uncertainty analysis of
589 complex-models of disease transmission – an HIV model, as an example, *Int. Stat.*
590 *Rev.*, 62(2), 229–243.

591 Nabi, K. N. & Podder, C. N., (2020). Sensitivity analysis of chronic hepatitis C
592 virus infection with immune response and cell proliferation. *International Journal*
593 *of Biomathematics*, 13(3), 2050017.

594 Worldometer, https://www.worldometers.info/coronavirus, accessed: 12-05-2020.

595 Qianying, L., Shi, Z., Daozhou, G., Yijun, L., Shu, Y., Salihu, S. M., Maggie, H. W.,
596 Yongli, K., Weiming, W., Lin, Y., Daihai, H., (2020). A conceptual model for the
597 outbreak of Coronavirus disease 2019 (COVID-19) in Wuhan, China with individual
598 reaction and governmental action. *International Journal of Infectious Diseases*, 93,
599 211-216.

600 G. Chavent. Nonlinear Least Squares for Inverse Problems: Theoretical Foundations
601 and Step-by-Step Guide for Applications. Springer, 2010.

602 Richard, C. A., Brian, B., Clifford, H. T., Parameter Estimation and Inverse Problems,
603 Elsevier, second edition, 2012. ISBN 978-0-128-10092-9.

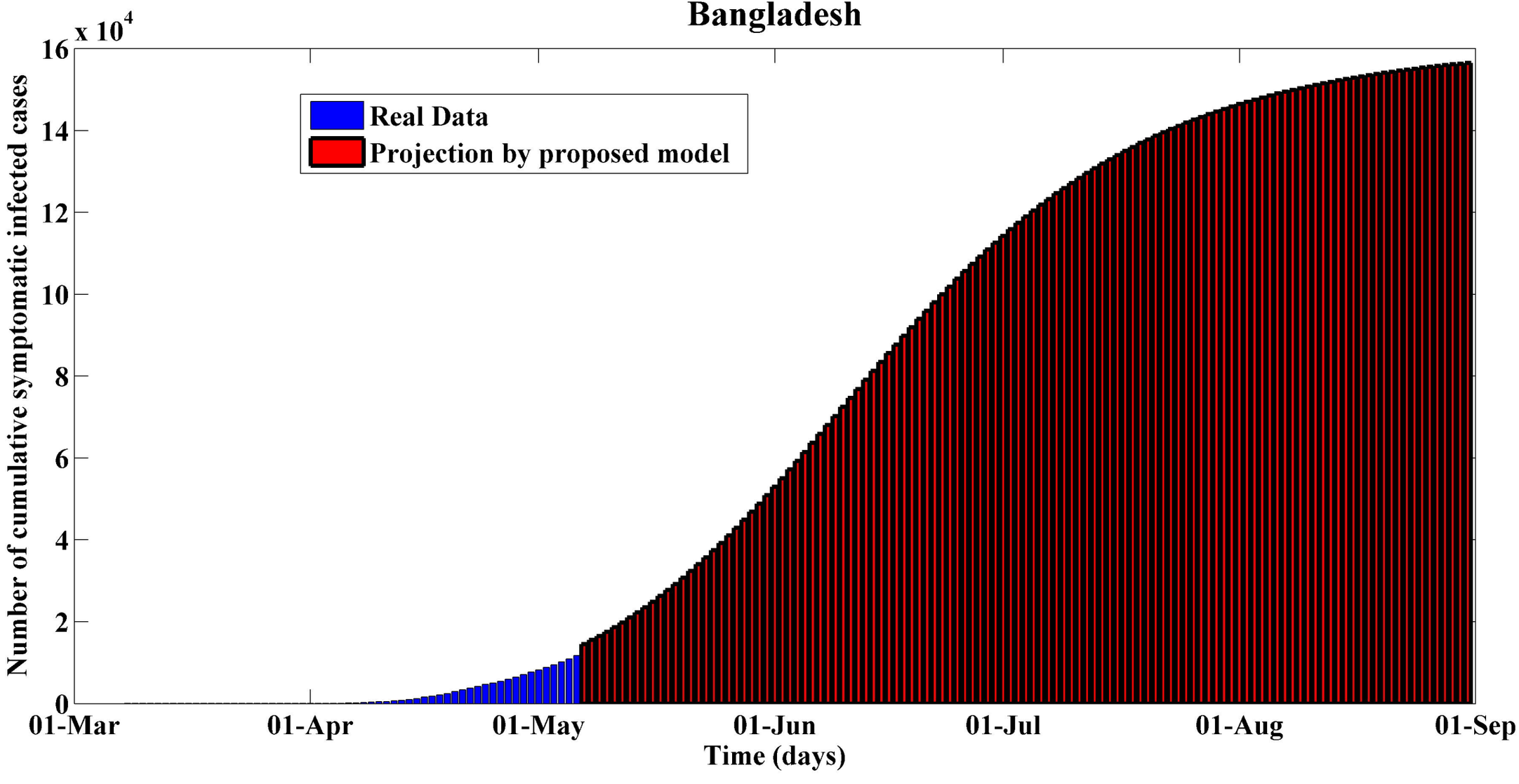
604 Yaman, F., Yakhno, V. G., & Potthast, R., 2013. A Survey on Inverse Problems for
605 Applied Sciences, *Mathematical Problems in Engineering*, Volume 2013, Article ID
606 976837.

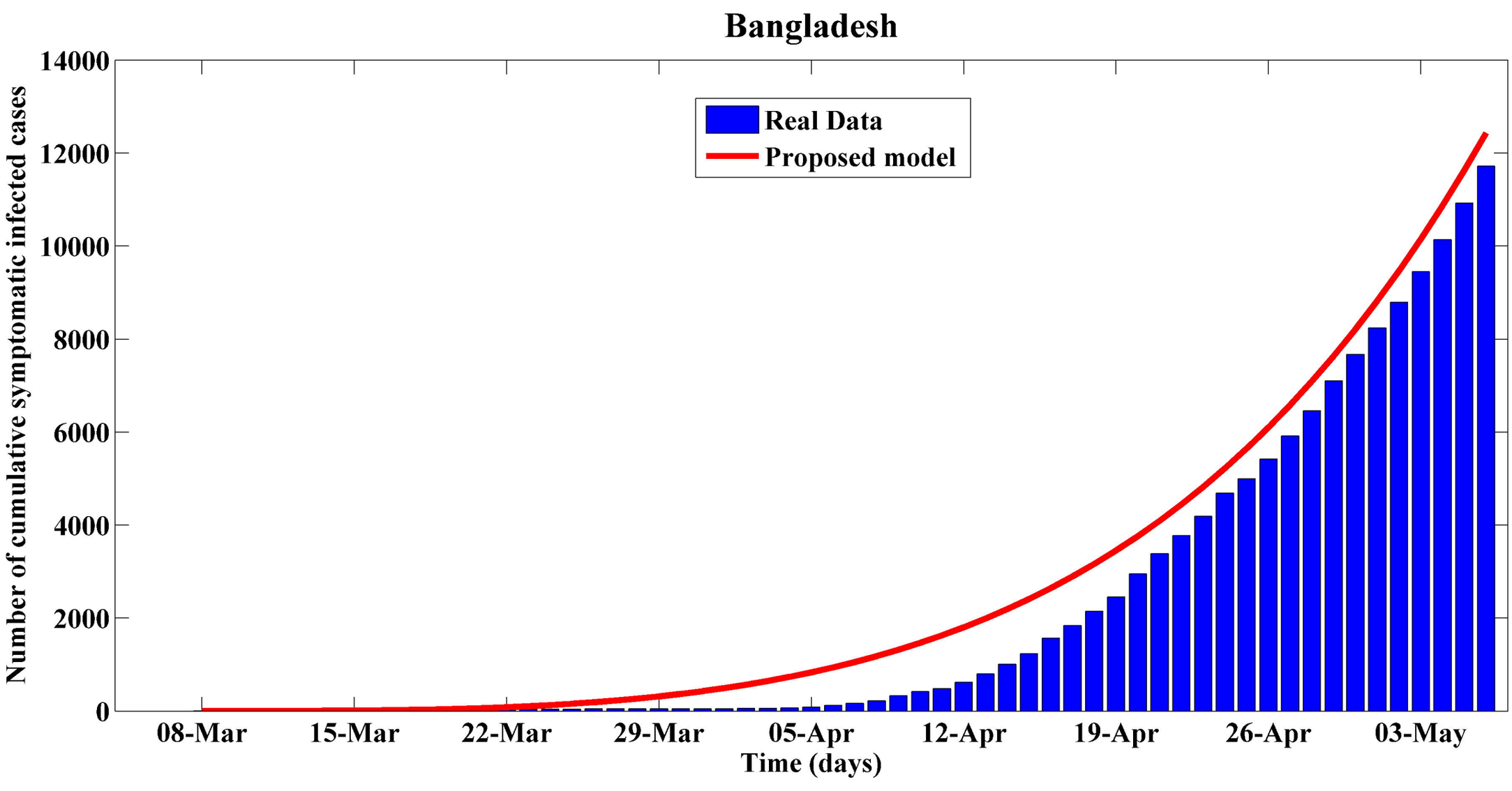
607 Coleman, T. F. & Y. Li. An interior, trust region approach for nonlinear minimization
608 subject to bounds. *SIAM J. Optimization*, 6(2), 418-445.

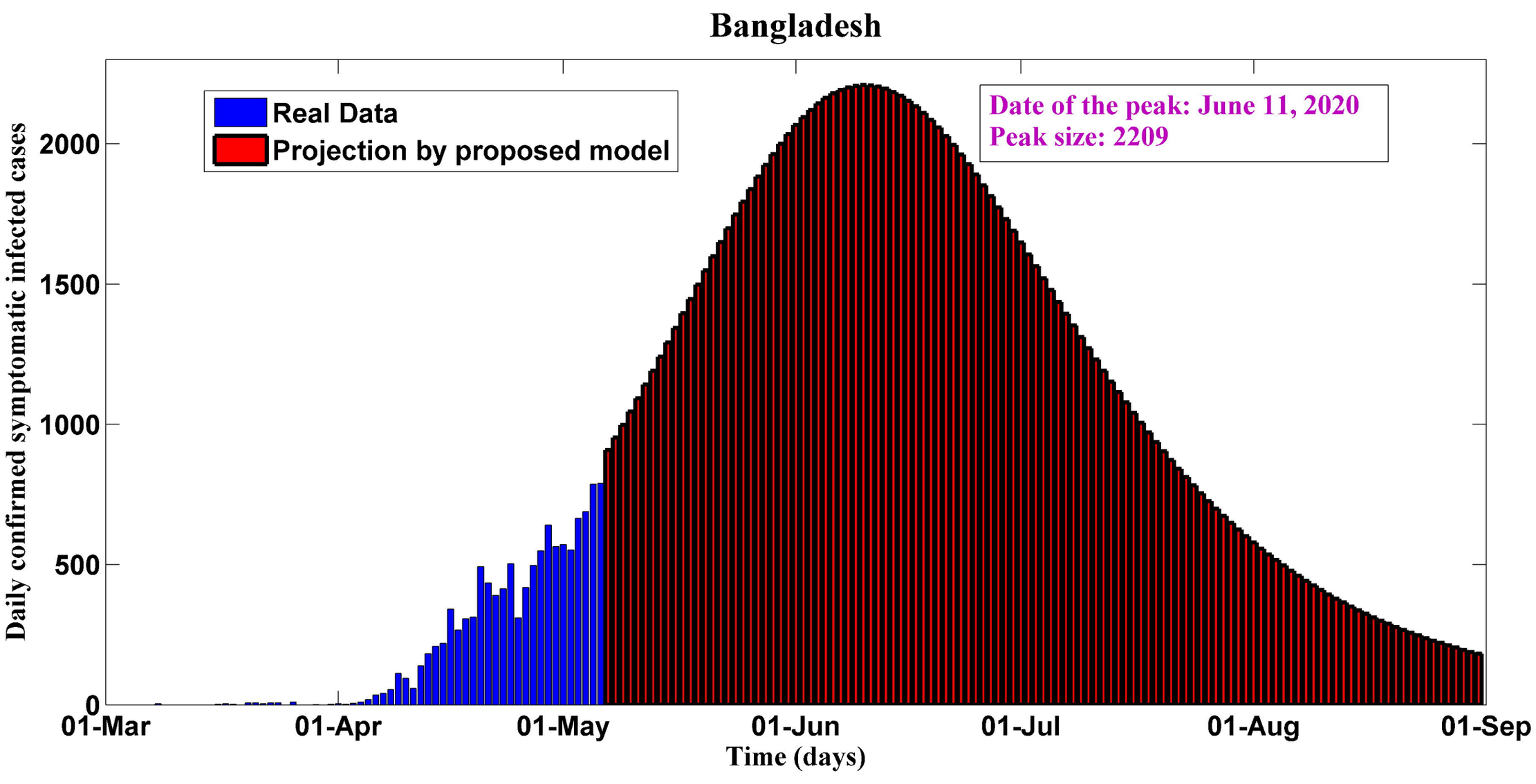
609 Andrew, R. C., Nicholas, I. M. G., and Philippe, L. T., Trust Region Methods. SIAM,
610 2000.

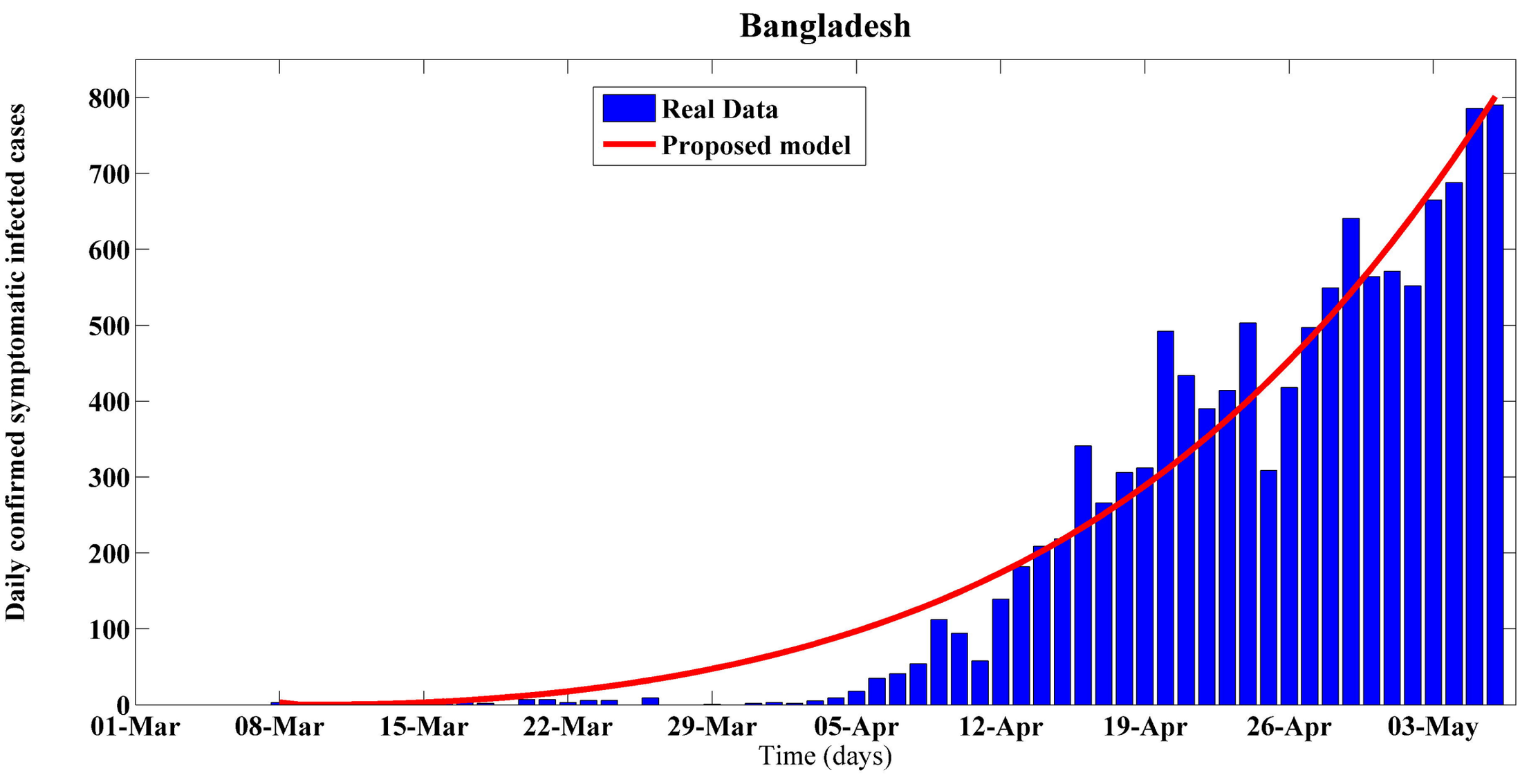
611 Institute of Epidemiology, Disease Control and Research (IEDCR), Bangladesh https:
612 //www.iedcr.gov.bd, accessed: 11-05-2020.

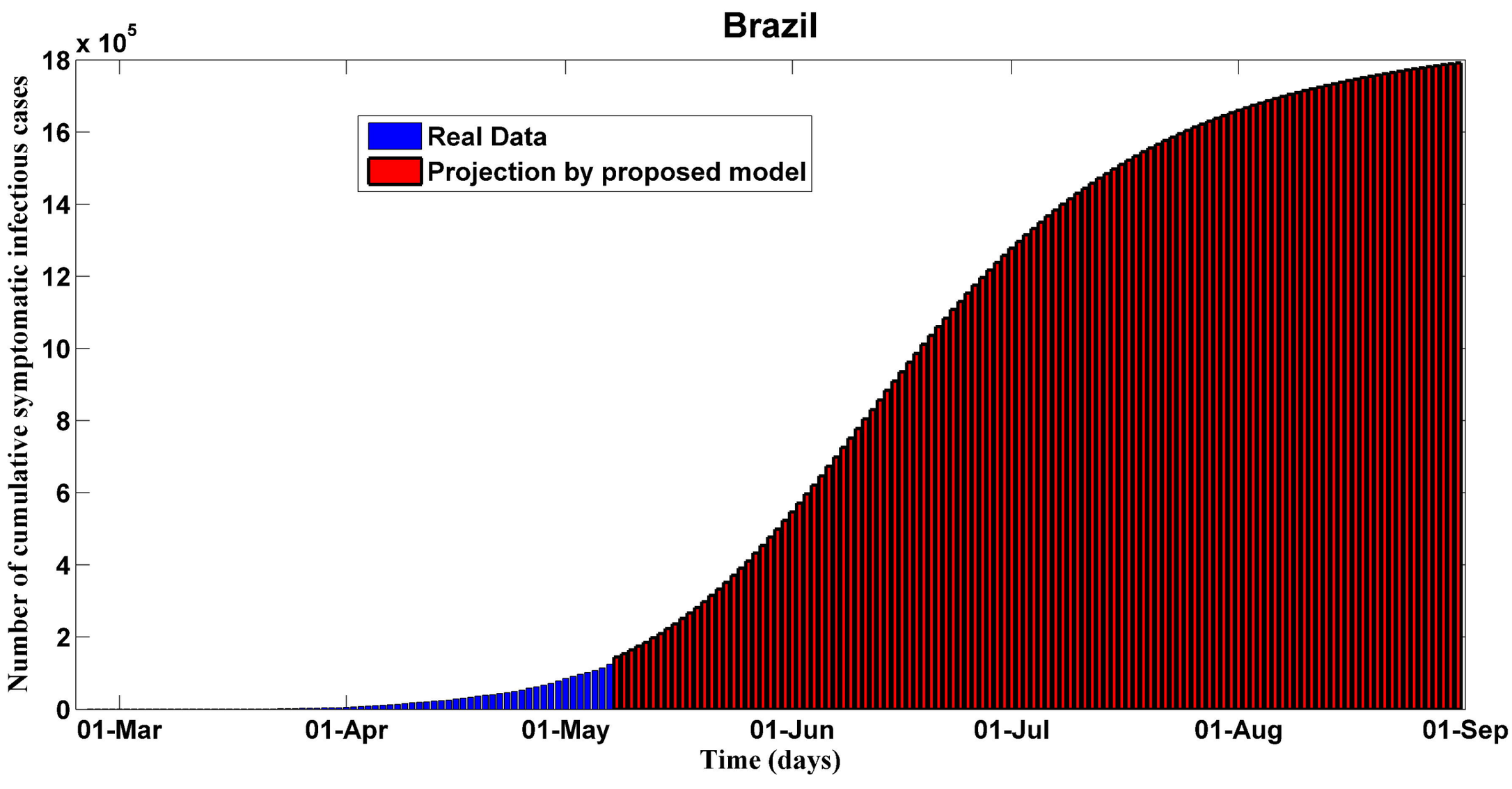
613 Ministry of Health & Family Welfare of Bangladesh, https://dghs.gov.bd/index.
614 php/en/home/5343-covid-19-update, accessed: 11-05-2020.

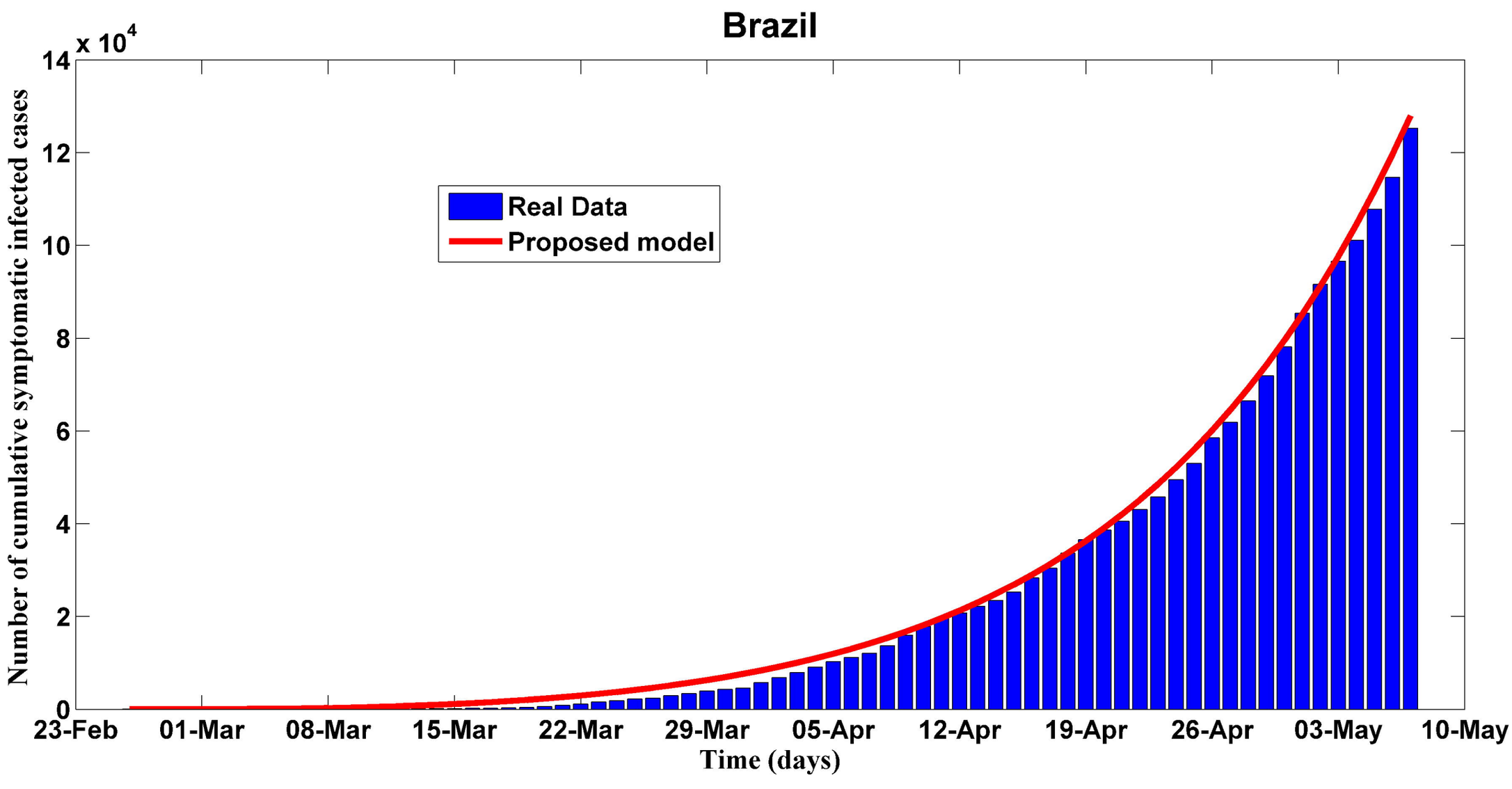

Bangladesh
x 10^4
16
14
12
10
8
6
4
2
0
Number of cumulative symptomatic infected cases
Real Data
Projection by proposed model
01-Mar
01-Apr
01-May
01-Jun
01-Jul
01-Aug
01-Sep
Time (days)

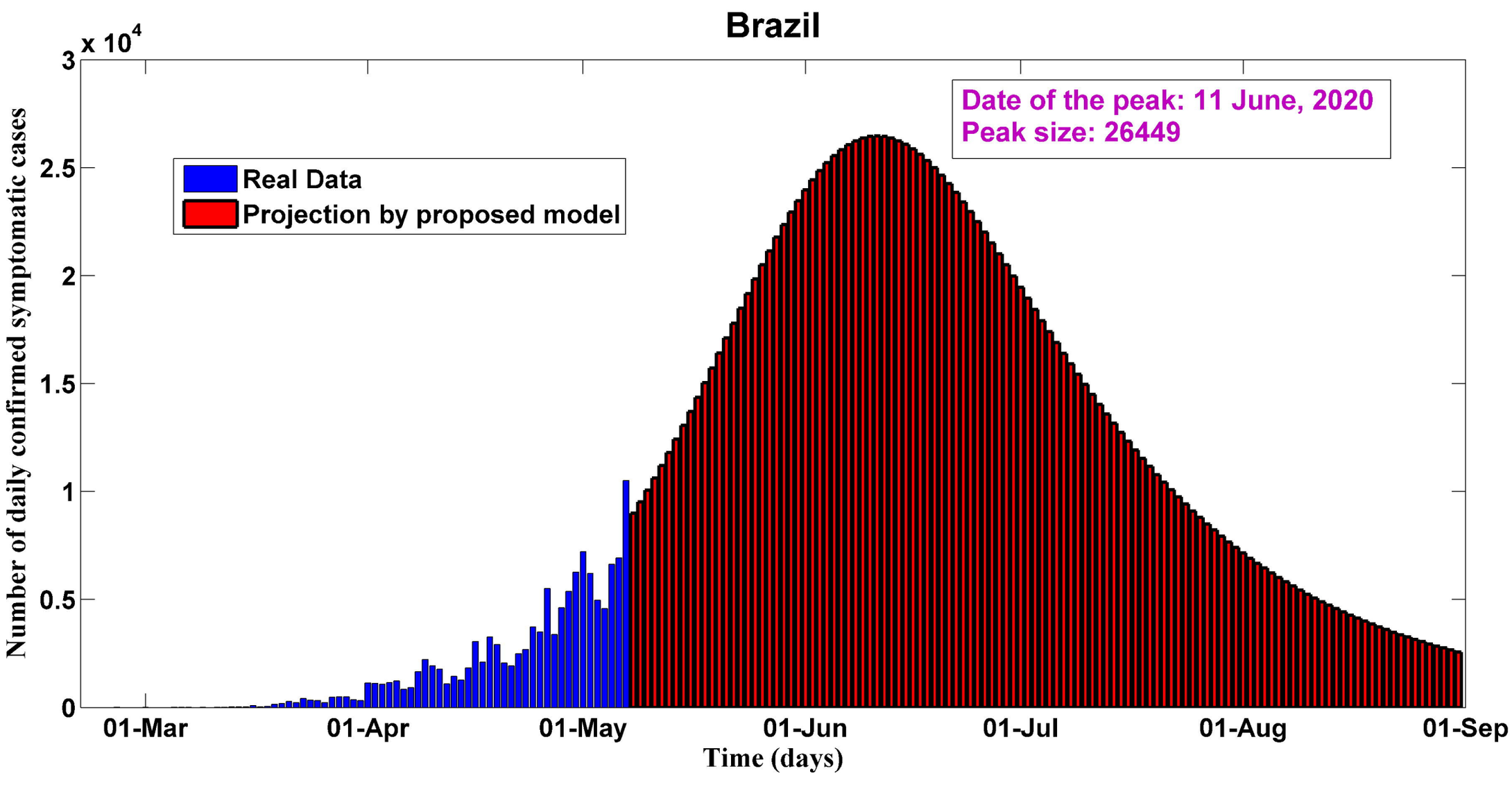

Bangladesh
Real Data
Proposed model
Number of cumulative symptomatic infected cases
14000
12000
10000
8000
6000
4000
2000
0
08-Mar
15-Mar
22-Mar
29-Mar
05-Apr
12-Apr
19-Apr
26-Apr
03-May
Time (days)

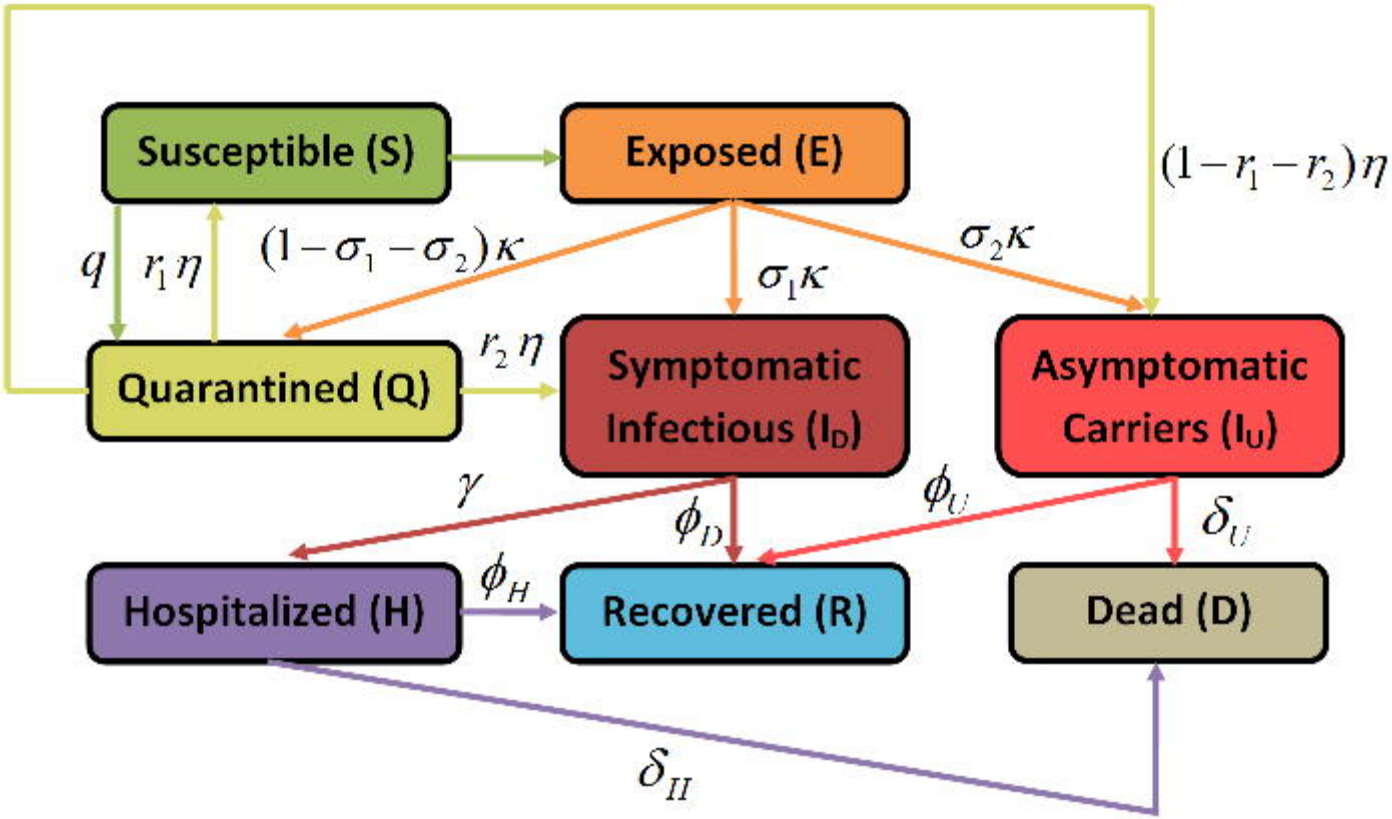

Bangladesh
Real Data
Projection by proposed model
Date of the peak: June 11, 2020
Peak size: 2209
Daily confirmed symptomatic infected cases
2000
1500
1000
500
0
01-Mar
01-Apr
01-May
01-Jun
01-Jul
01-Aug
01-Sep
Time (days)

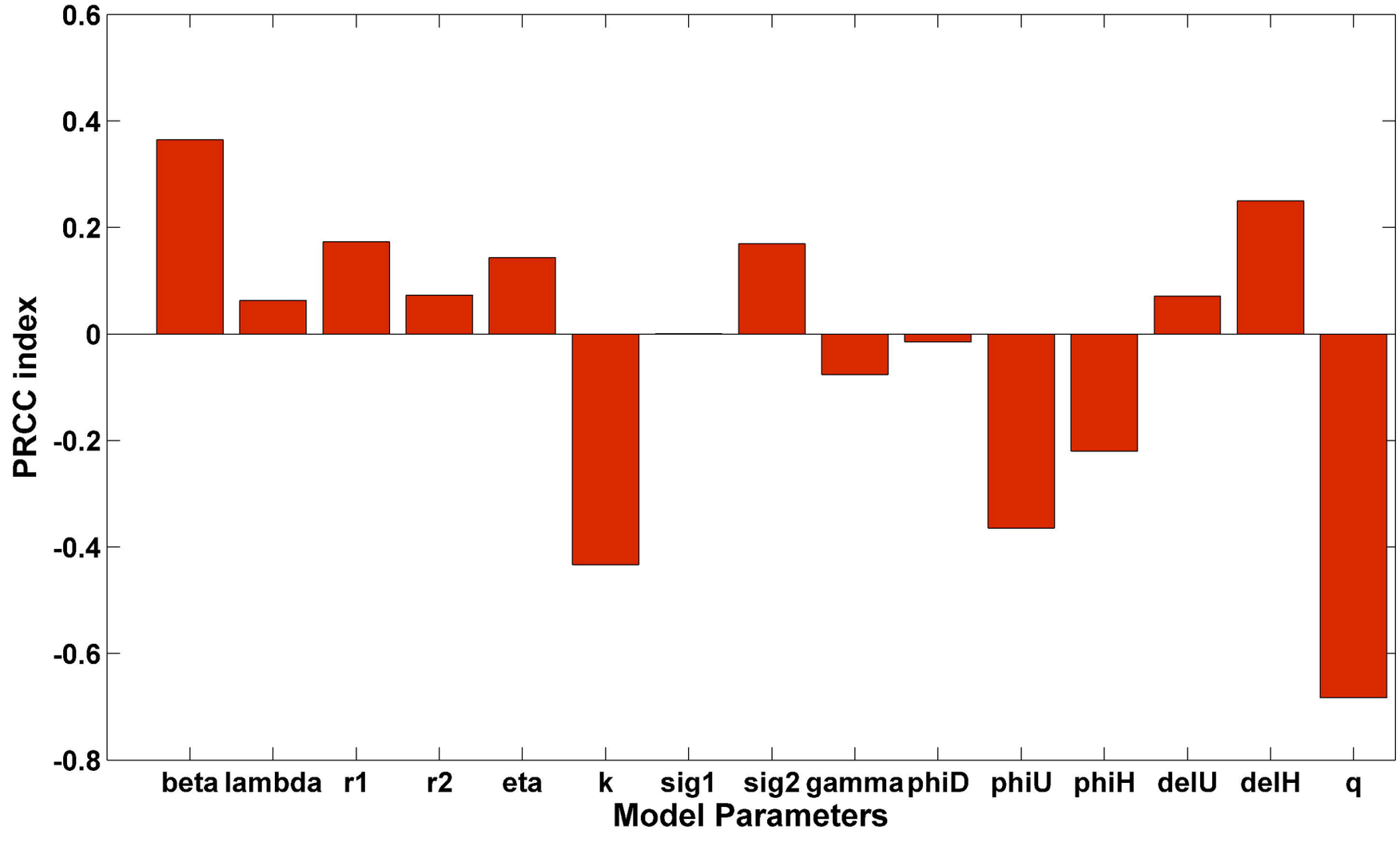

Bangladesh
Real Data
Proposed model
Daily confirmed symptomatic infected cases
0
100
200
300
400
500
600
700
800
01-Mar
08-Mar
15-Mar
22-Mar
29-Mar
05-Apr
12-Apr
19-Apr
26-Apr
03-May
Time (days)

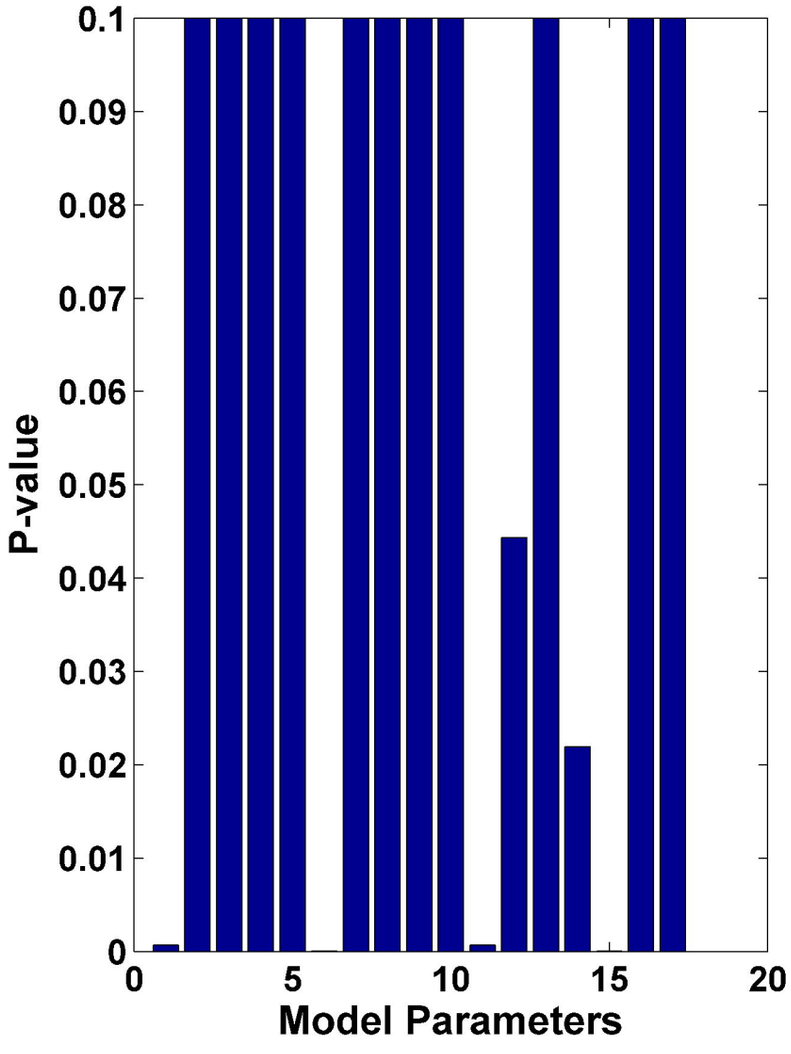

Brazil
x 10^5
18
16
14
12
10
8
6
4
2
0
Number of cumulative symptomatic infectious cases
Real Data
Projection by proposed model
01-Mar
01-Apr
01-May
01-Jun
01-Jul
01-Aug
01-Sep
Time (days)

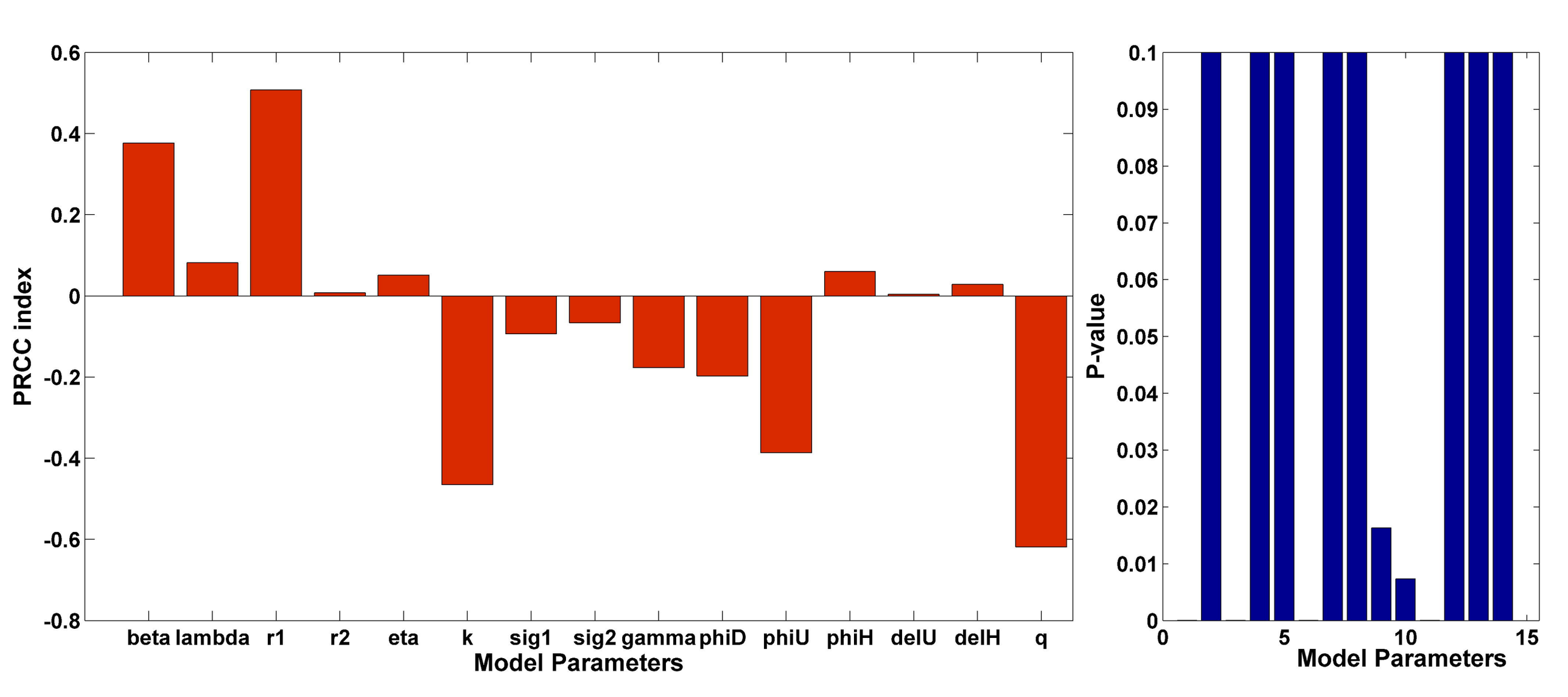

Brazil
x 10^4
14
12
10
8
6
4
2
0
Number of cumulative symptomatic infected cases
Real Data
Proposed model
23-Feb
01-Mar
08-Mar
15-Mar
22-Mar
29-Mar
05-Apr
12-Apr
19-Apr
26-Apr
03-May
10-May
Time (days)

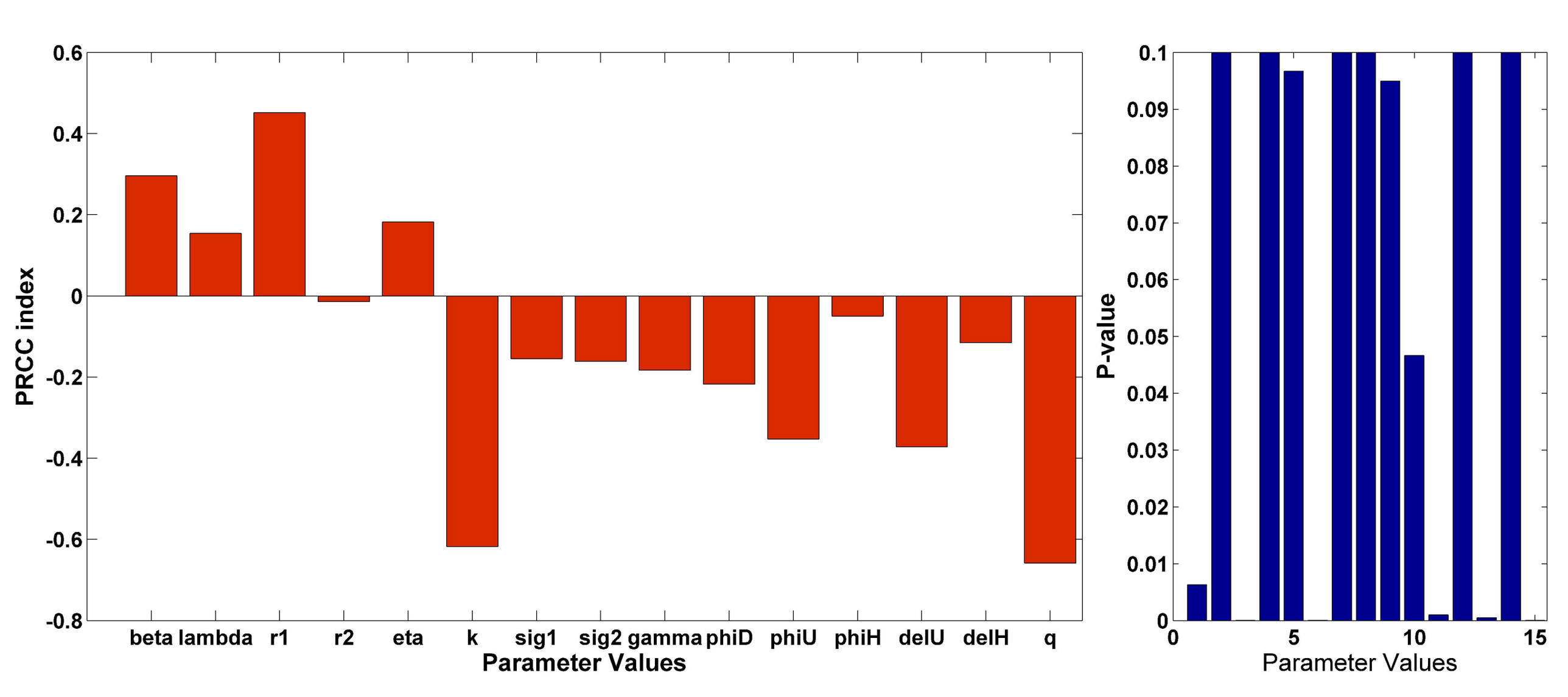

Brazil
x 10^4
3
2.5
2
1.5
1
0.5
0
Number of daily confirmed symptomatic cases
Real Data
Projection by proposed model
Date of the peak: 11 June, 2020
Peak size: 26449
01-Mar
01-Apr
01-May
01-Jun
01-Jul
01-Aug
01-Sep
Time (days)

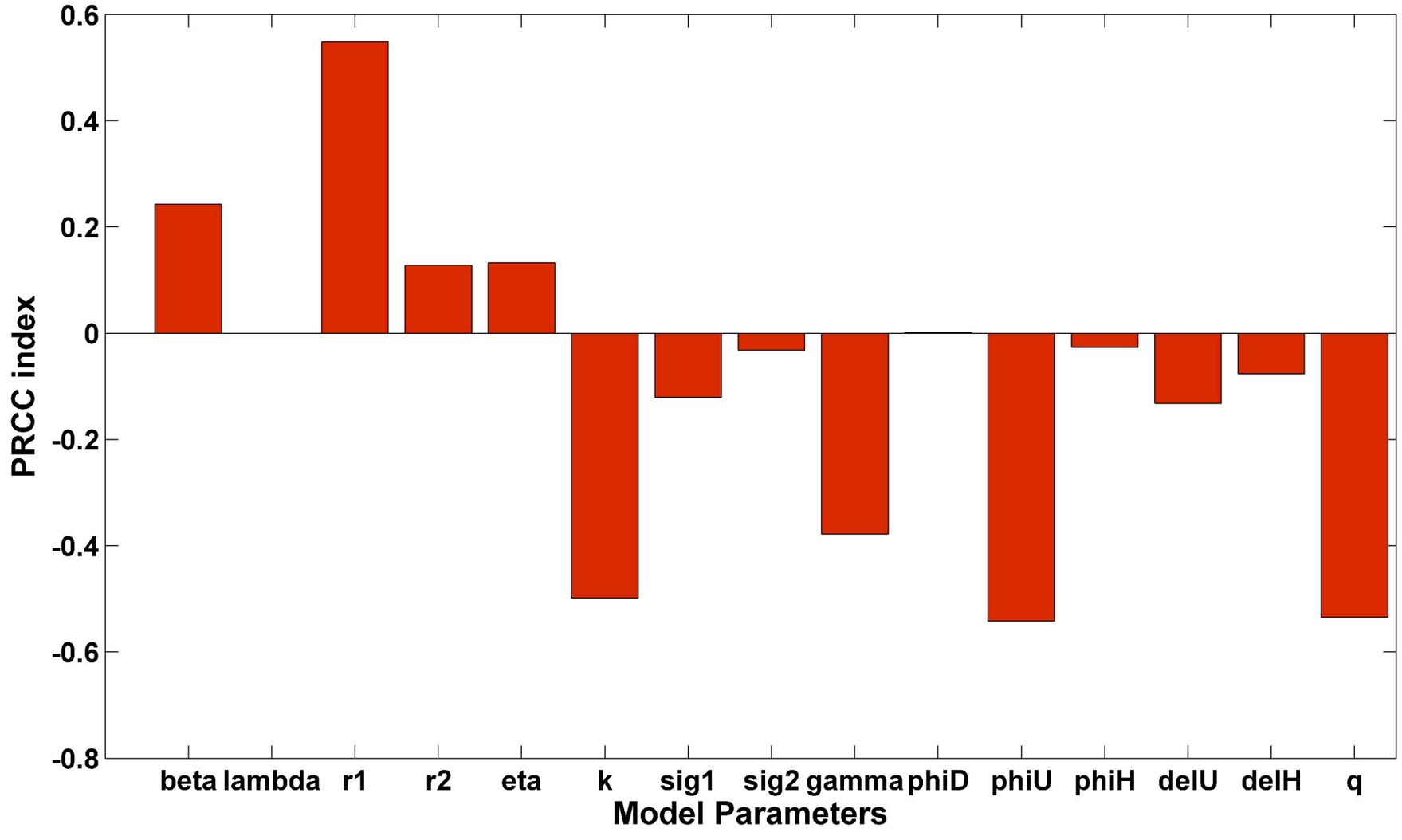

Susceptible (S)
Exposed (E)
Quarantined (Q)
Symptomatic Infectious (I_D)
Asymptomatic Carriers (I_U)
Hospitalized (H)
Recovered (R)
Dead (D)
$(1-r_1-r_2)\eta$
q
$r_1\eta$
$(1-\sigma_1-\sigma_2)\kappa$
$\sigma_1\kappa$
$\sigma_2\kappa$
$r_2\eta$
γ
ϕ_D
ϕ_U
δ_U
ϕ_H
δ_H

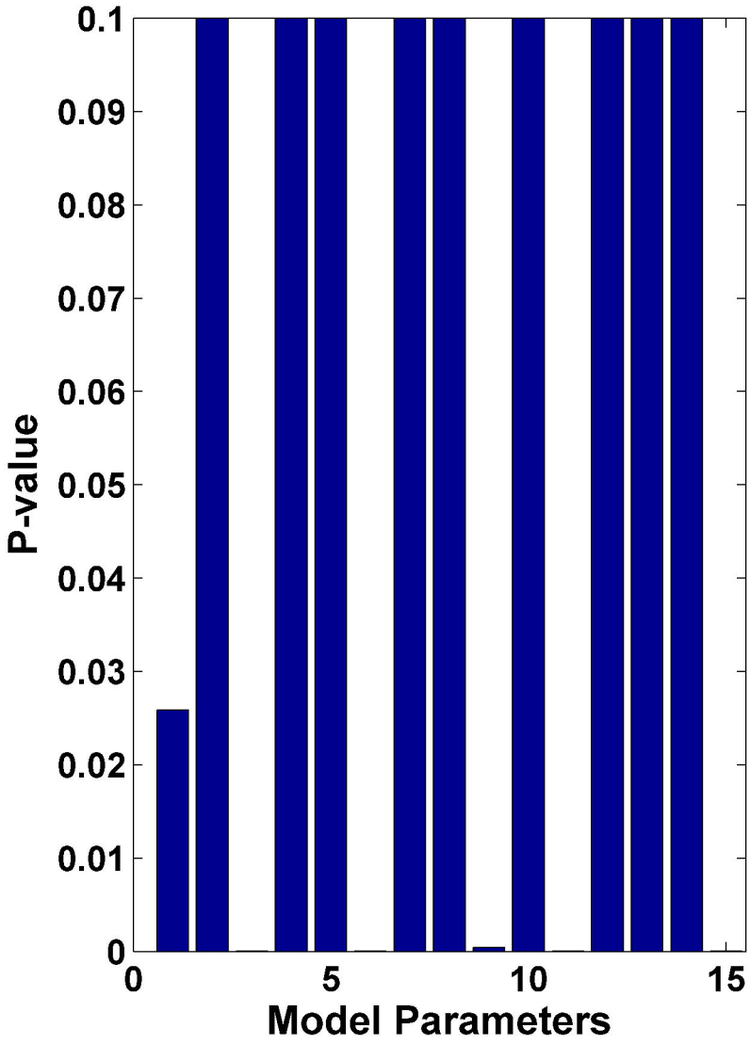

0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
PRCC index
beta
lambda
r1
r2
eta
k
sig1
sig2
gamma
phiD
phiU
phiH
delU
delH
q
Model Parameters

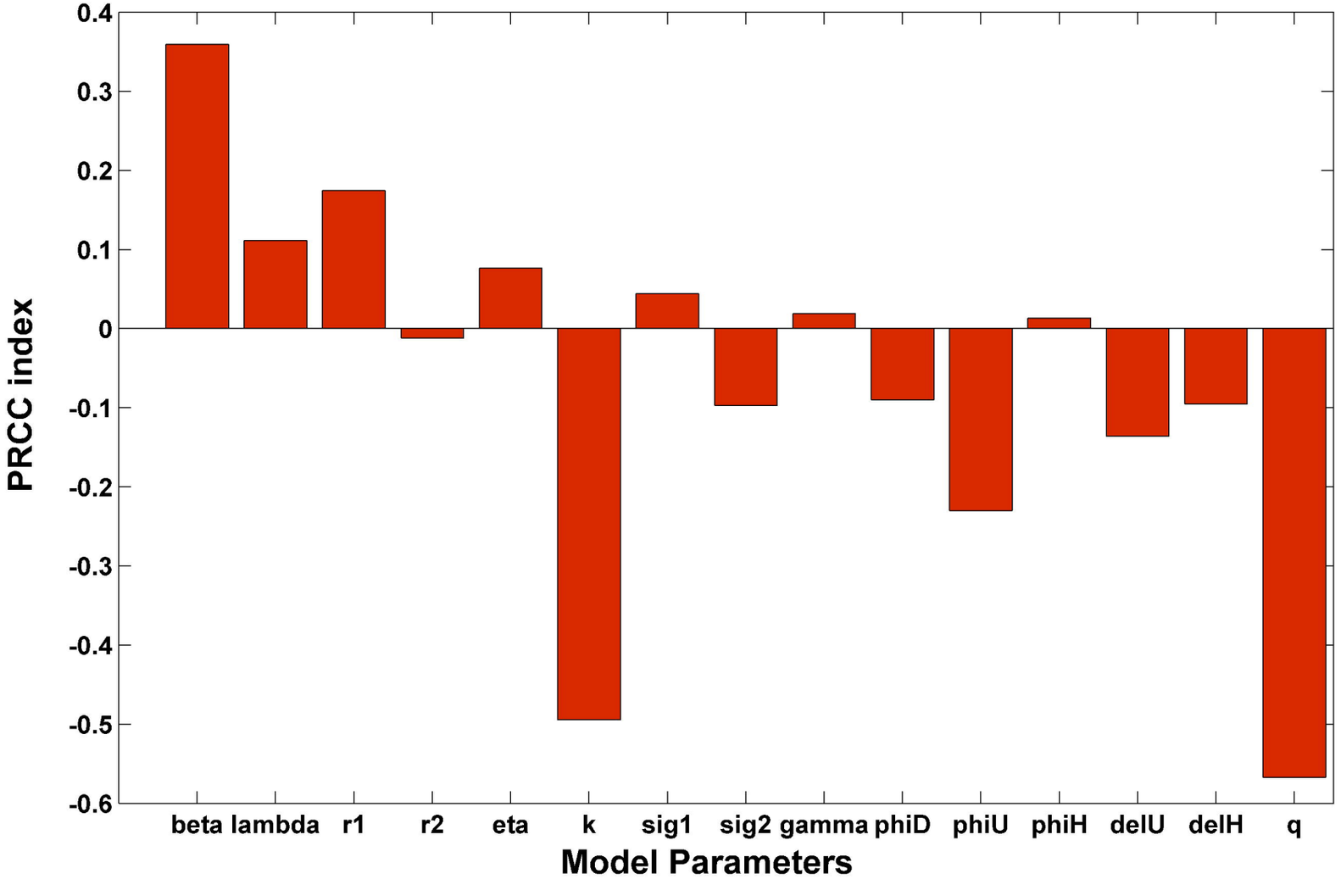

0.1
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
P-value
0
5
10
15
20
Model Parameters

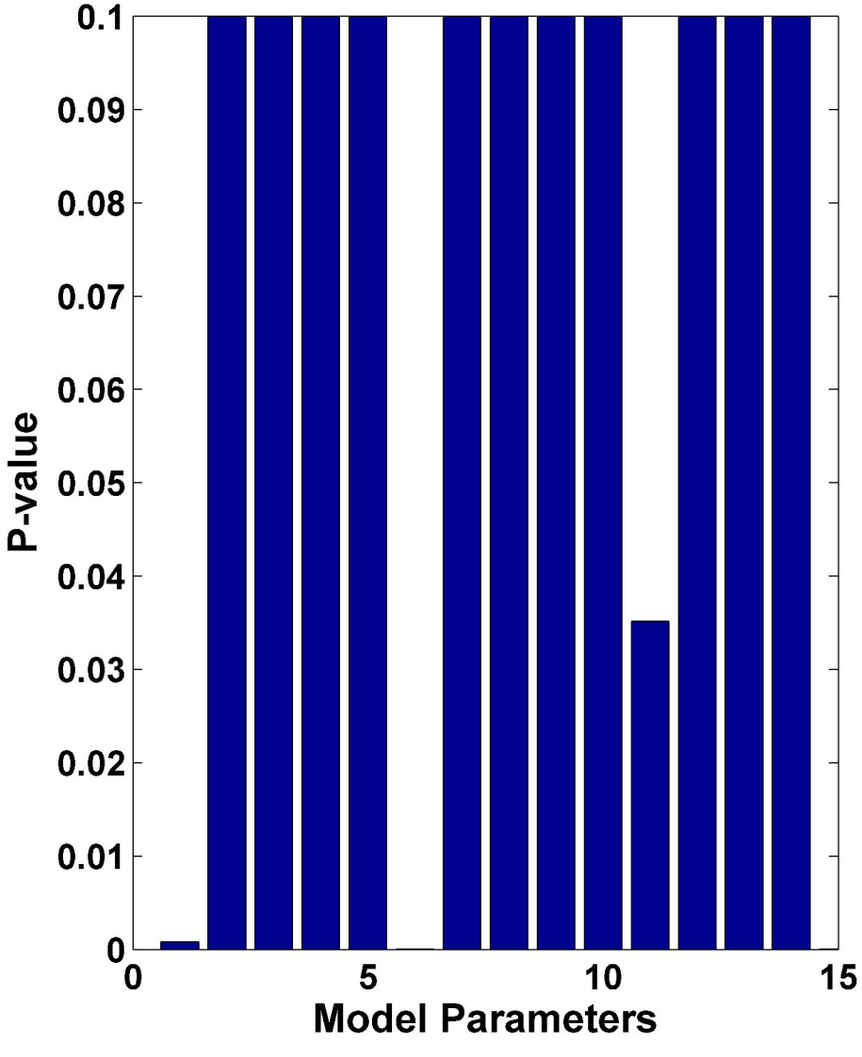

PRCC index
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
beta
lambda
r1
r2
eta
k
sig1
sig2
gamma
phiD
phiU
phiH
delU
delH
q
Model Parameters
P-value
0.1
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
0
5
10
15
Model Parameters

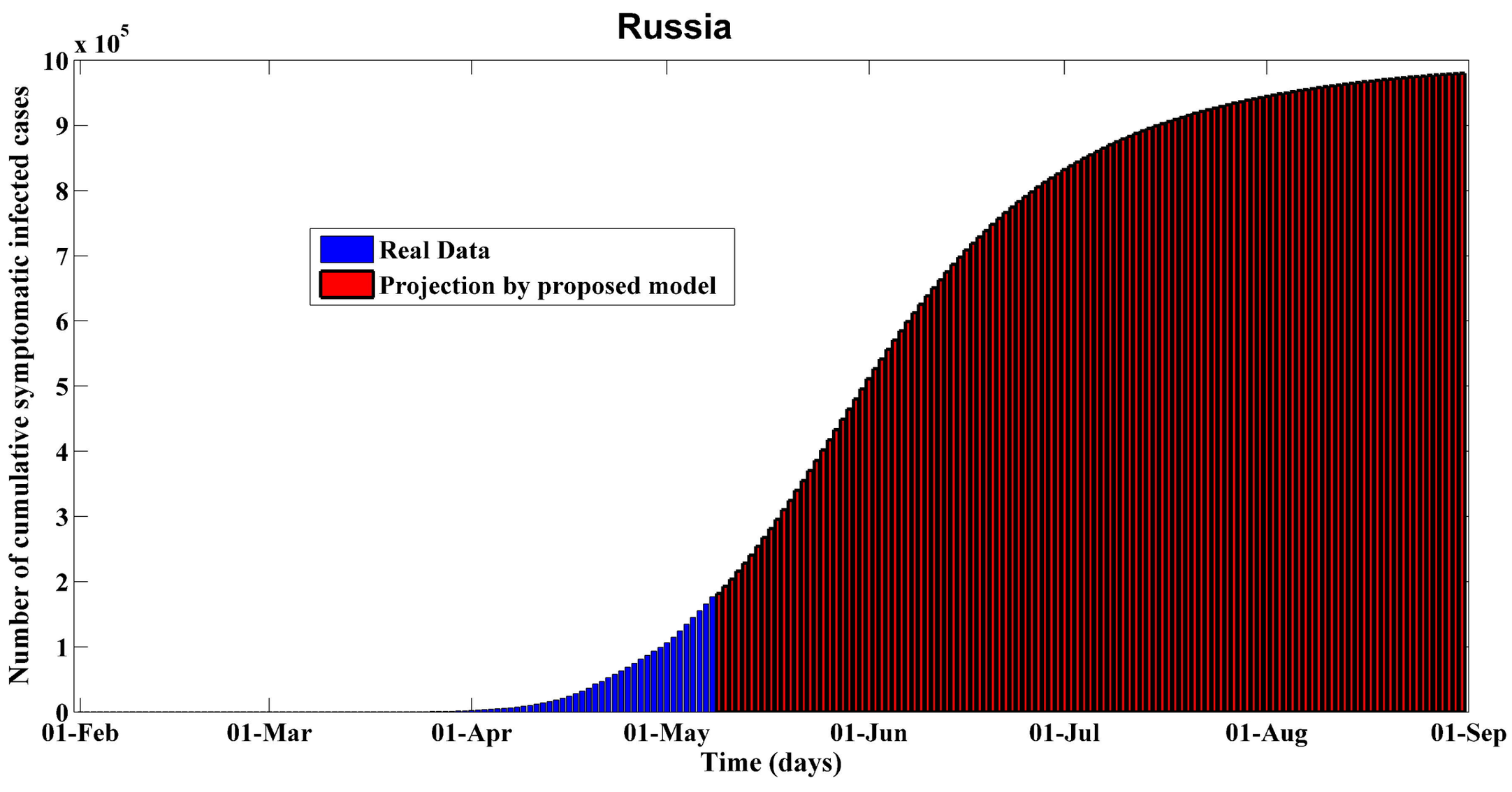

PRCC index
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
beta
lambda
r1
r2
eta
k
sig1
sig2
gamma
phiD
phiU
phiH
delU
delH
q
Parameter Values
P-value
0.1
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
0
5
10
15
Parameter Values

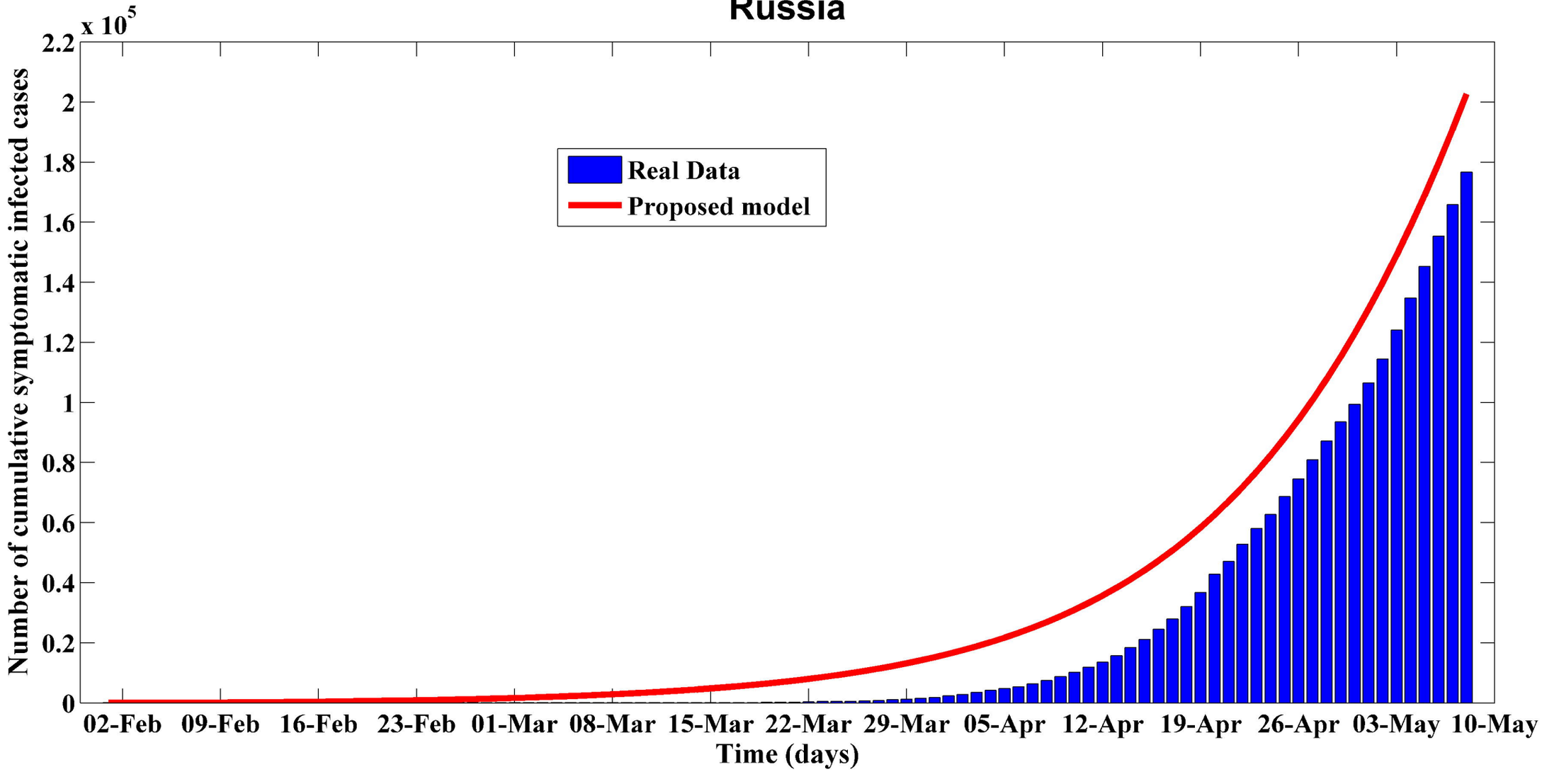

PRCC index
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
beta
lambda
r1
r2
eta
k
sig1
sig2
gamma
phiD
phiU
phiH
delU
delH
q
Model Parameters

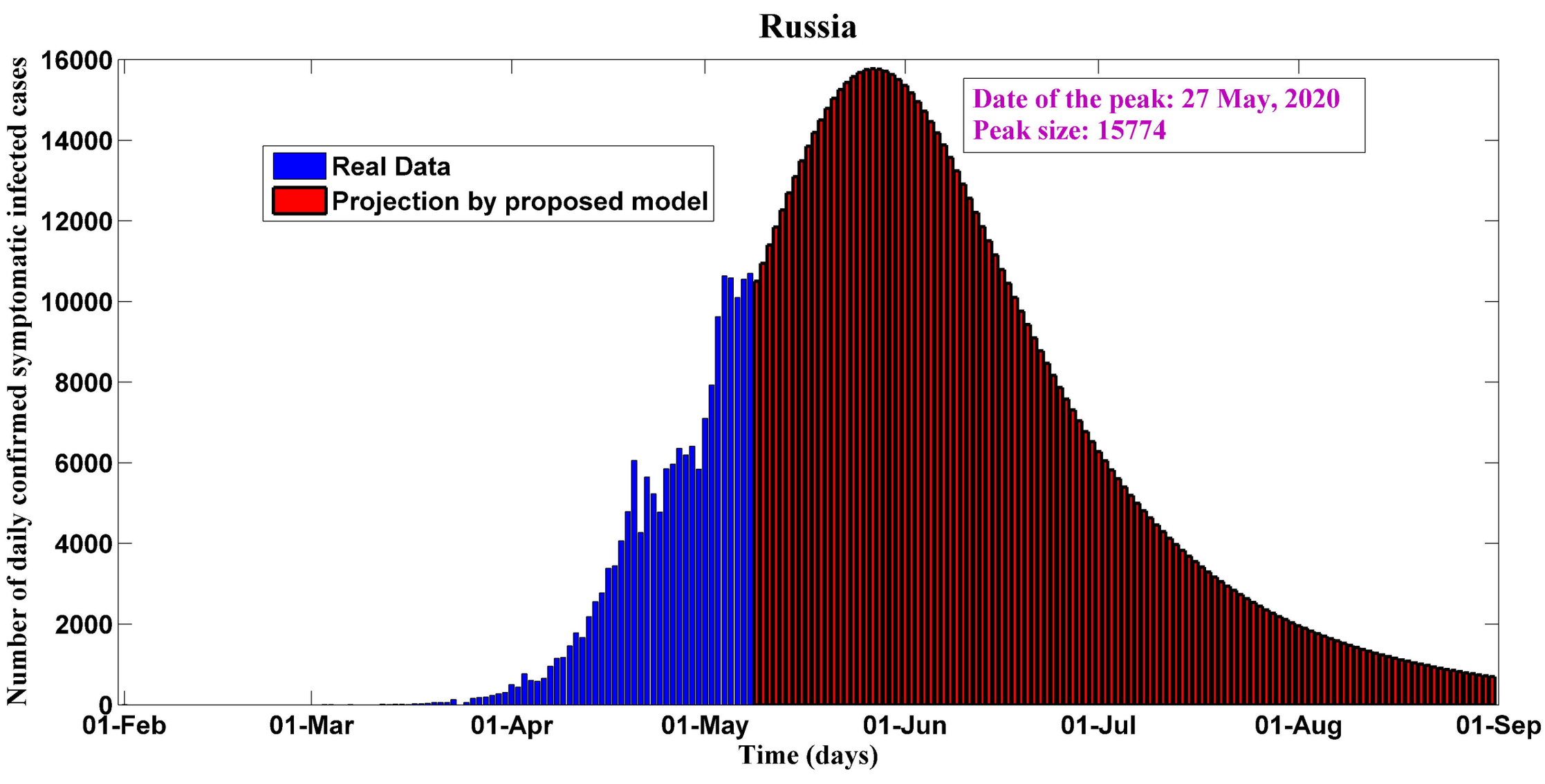

P-value
0.1
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
0
5
10
15
Model Parameters

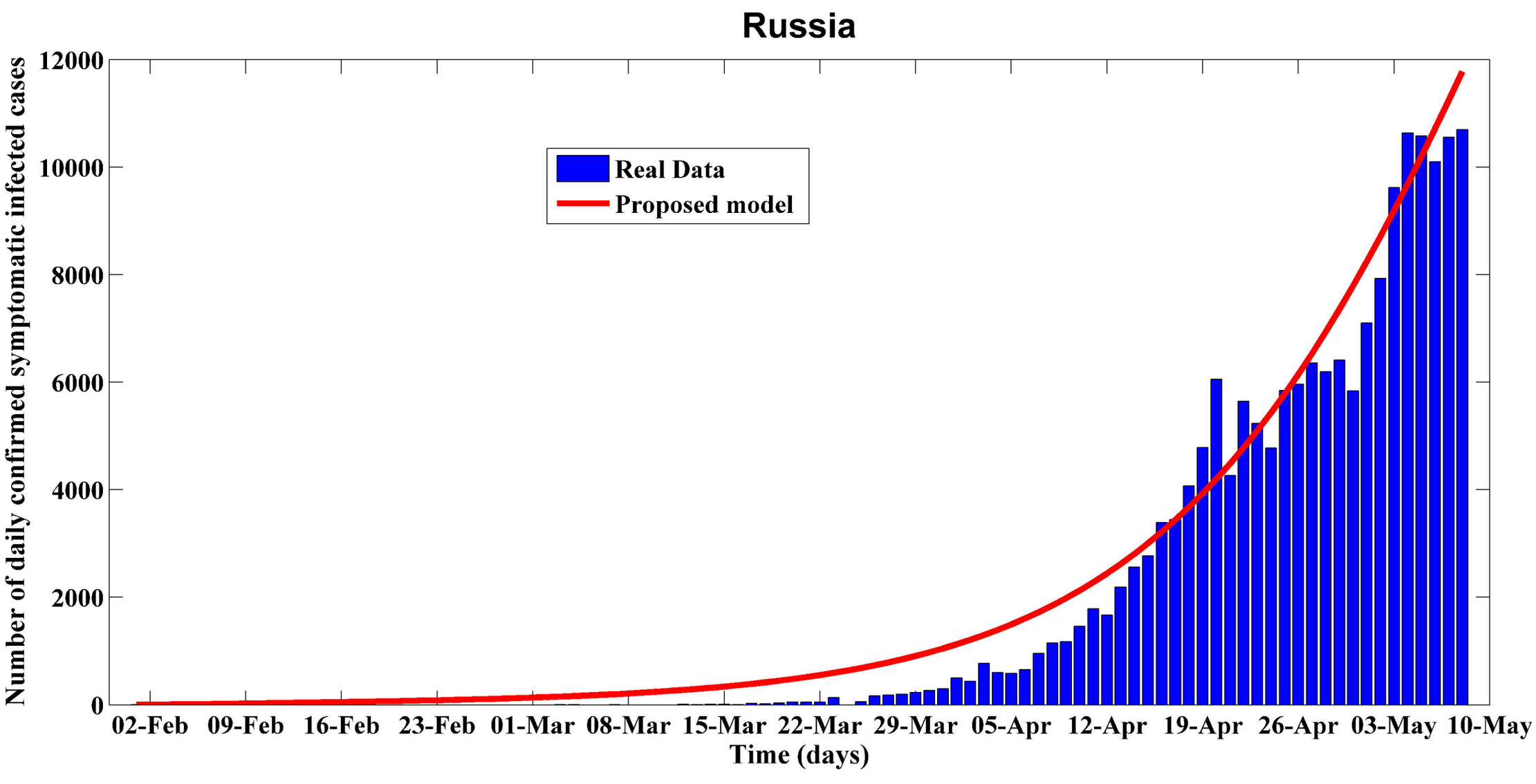

PRCC index
0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
beta
lambda
r1
r2
eta
k
sig1
sig2
gamma
phiD
phiU
phiH
delU
delH
q
Model Parameters

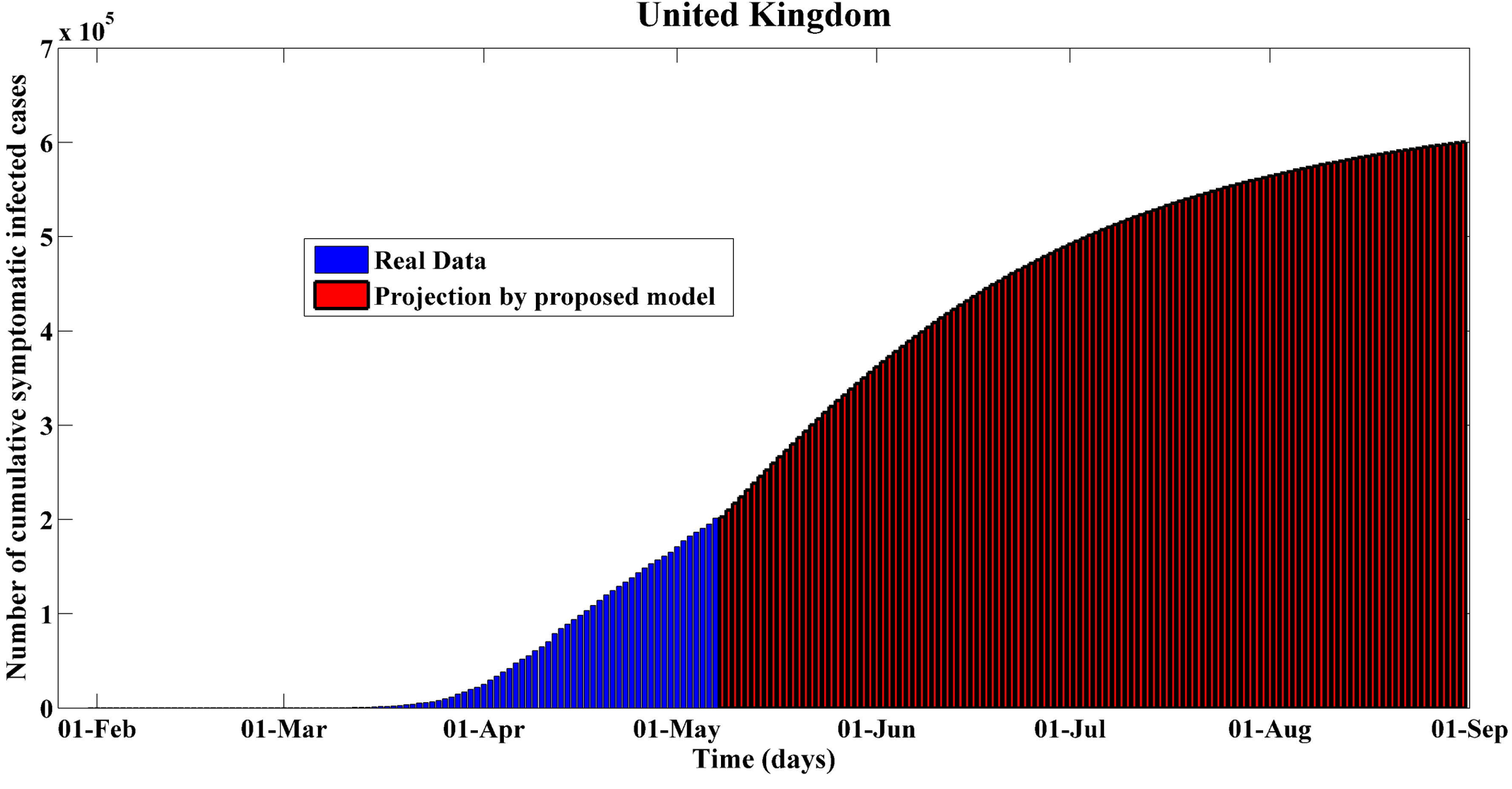

P-value
0.1
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
0
5
10
15
Model Parameters

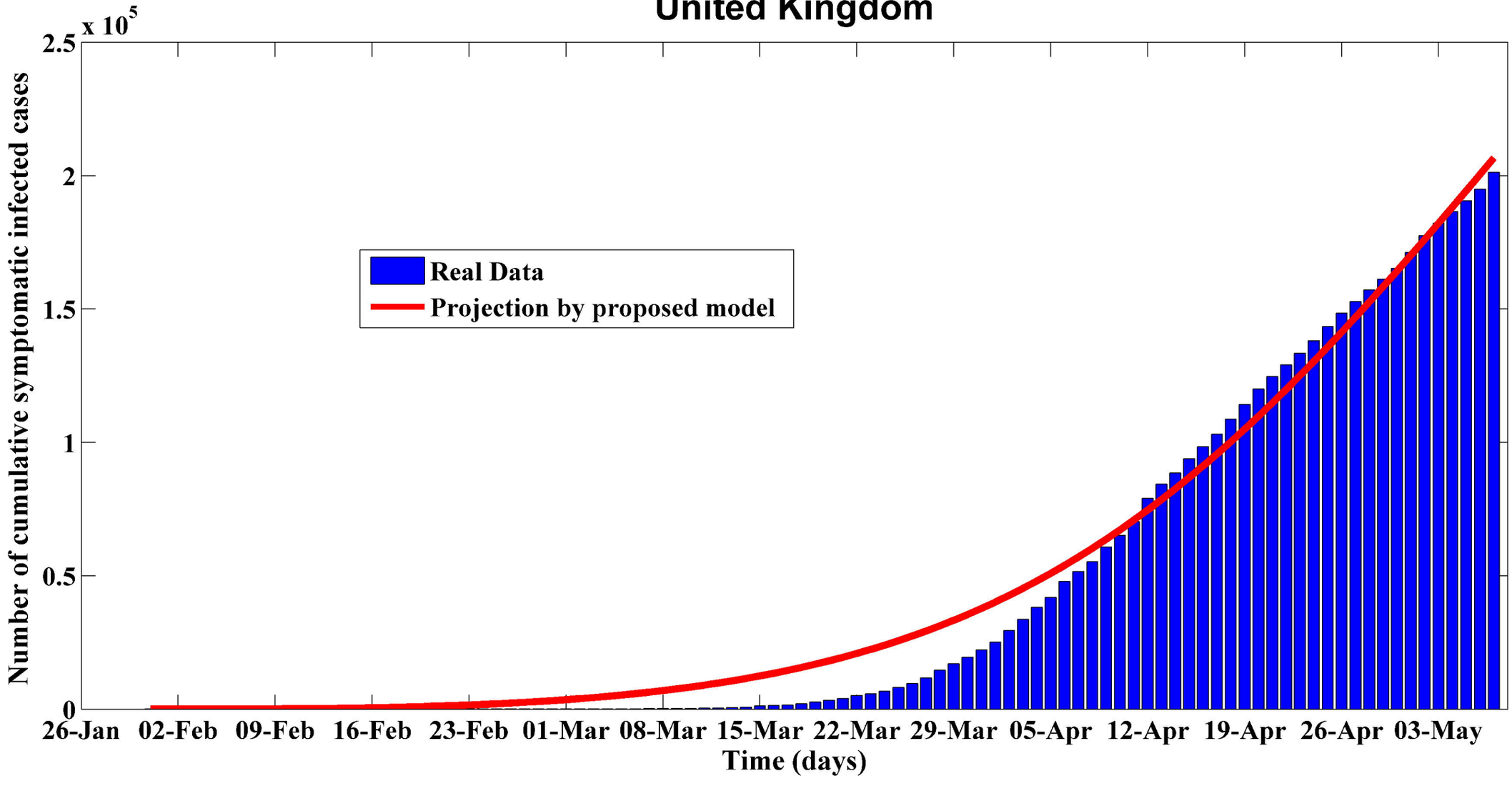

Russia
x 10^5
10
9
8
7
6
5
4
3
2
1
0
Number of cumulative symptomatic infected cases
Real Data
Projection by proposed model
01-Feb
01-Mar
01-Apr
01-May
01-Jun
01-Jul
01-Aug
01-Sep
Time (days)

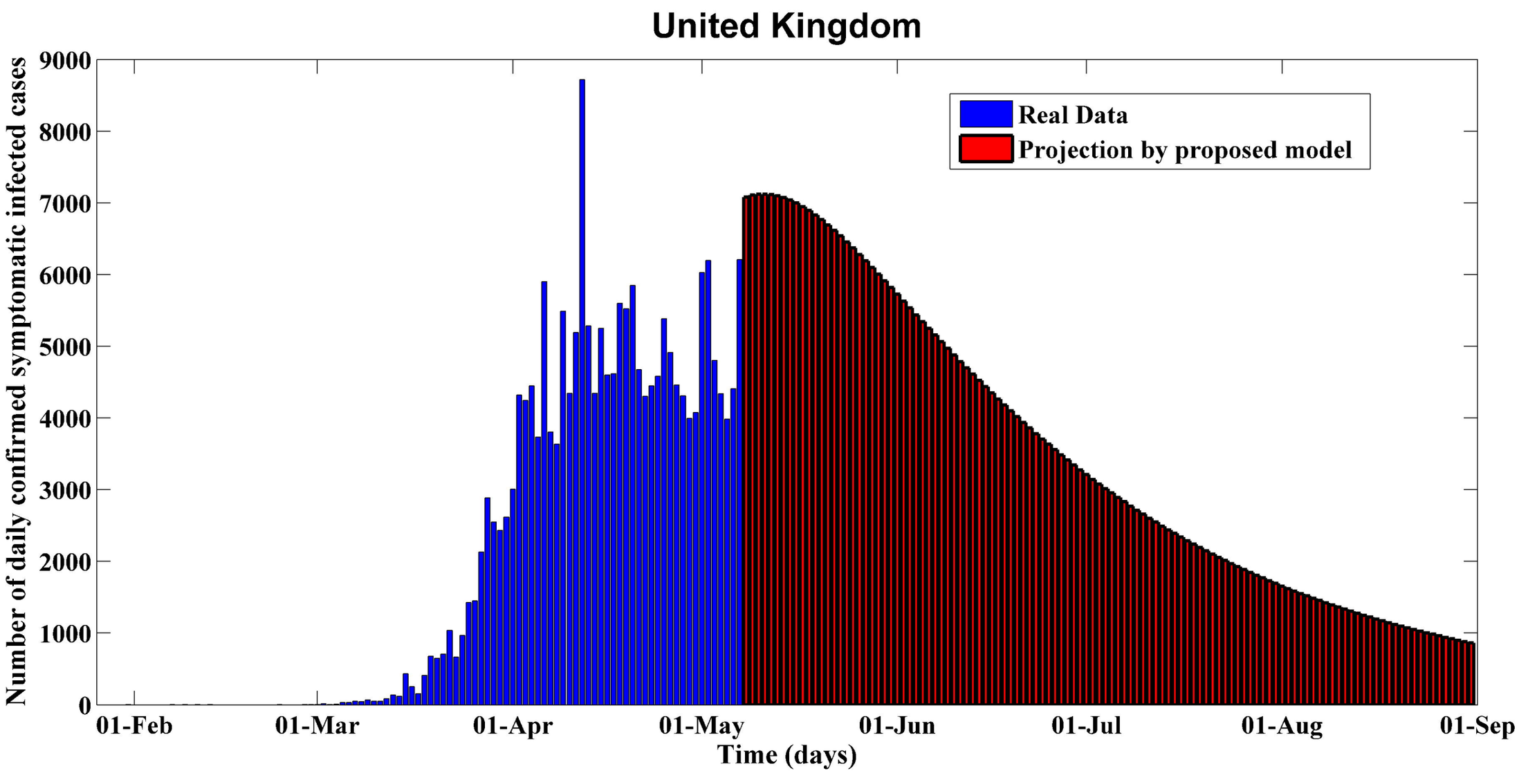

Russia
x 10^5
2.2
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Number of cumulative symptomatic infected cases
Real Data
Proposed model
02-Feb
09-Feb
16-Feb
23-Feb
01-Mar
08-Mar
15-Mar
22-Mar
29-Mar
05-Apr
12-Apr
19-Apr
26-Apr
03-May
10-May
Time (days)


Russia
Number of daily confirmed symptomatic infected cases
16000
14000
12000
10000
8000
6000
4000
2000
0
Real Data
Projection by proposed model
Date of the peak: 27 May, 2020
Peak size: 15774
01-Feb
01-Mar
01-Apr
01-May
01-Jun
01-Jul
01-Aug
01-Sep
Time (days)


Russia
Real Data
Proposed model
Number of daily confirmed symptomatic infected cases
0
2000
4000
6000
8000
10000
12000
02-Feb
09-Feb
16-Feb
23-Feb
01-Mar
08-Mar
15-Mar
22-Mar
29-Mar
05-Apr
12-Apr
19-Apr
26-Apr
03-May
10-May
Time (days)


United Kingdom
x 10^5
Number of cumulative symptomatic infected cases
0
1
2
3
4
5
6
7
Real Data
Projection by proposed model
01-Feb
01-Mar
01-Apr
01-May
01-Jun
01-Jul
01-Aug
01-Sep
Time (days)


United Kingdom
x 10^5
2.5
2
1.5
1
0.5
0
Number of cumulative symptomatic infected cases
Real Data
Projection by proposed model
26-Jan
02-Feb
09-Feb
16-Feb
23-Feb
01-Mar
08-Mar
15-Mar
22-Mar
29-Mar
05-Apr
12-Apr
19-Apr
26-Apr
03-May
Time (days)


United Kingdom
Real Data
Projection by proposed model
Number of daily confirmed symptomatic infected cases
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
01-Feb
01-Mar
01-Apr
01-May
01-Jun
01-Jul
01-Aug
01-Sep
Time (days)