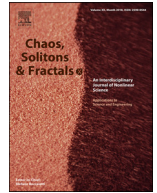




Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



# Modeling and forecasting the COVID-19 pandemic in India

Kankan Sarkar<sup>a</sup>, Subhas Khajanchi<sup>b,\*</sup>, Juan J. Nieto<sup>c</sup>

<sup>a</sup> Department of Mathematics, Malda College, Malda, West Bengal 732101, India

<sup>b</sup> Department of Mathematics, Presidency University, 86/1 College Street, Kolkata 700073, India

<sup>c</sup> Instituto de Matemáticas, Universidade de Santiago de Compostela, Santiago de Compostela, Spain

## ARTICLE INFO

### Article history:

Received 19 May 2020

Revised 12 June 2020

Accepted 22 June 2020

Available online 28 June 2020

### Keywords:

COVID-19

Mathematical model

Basic reproduction number

Sensitivity analysis

Isolation

Model prediction

## ABSTRACT

In India, 100,340 confirmed cases and 3155 confirmed deaths due to COVID-19 were reported as of May 18, 2020. Due to absence of specific vaccine or therapy, non-pharmacological interventions including social distancing, contact tracing are essential to end the worldwide COVID-19. We propose a mathematical model that predicts the dynamics of COVID-19 in 17 provinces of India and the overall India. A complete scenario is given to demonstrate the estimated pandemic life cycle along with the real data or history to date, which in turn divulges the predicted inflection point and ending phase of SARS-CoV-2. The proposed model monitors the dynamics of six compartments, namely susceptible (S), asymptomatic (A), recovered (R), infected (I), isolated infected ( $I_q$ ) and quarantined susceptible ( $S_q$ ), collectively expressed  $SARII_qS_q$ . A sensitivity analysis is conducted to determine the robustness of model predictions to parameter values and the sensitive parameters are estimated from the real data on the COVID-19 pandemic in India. Our results reveal that achieving a reduction in the contact rate between uninfected and infected individuals by quarantined the susceptible individuals, can effectively reduce the basic reproduction number. Our model simulations demonstrate that the elimination of ongoing SARS-CoV-2 pandemic is possible by combining the restrictive social distancing and contact tracing. Our predictions are based on real data with reasonable assumptions, whereas the accurate course of epidemic heavily depends on how and when quarantine, isolation and precautionary measures are enforced.

© 2020 Elsevier Ltd. All rights reserved.

## 1. Introduction

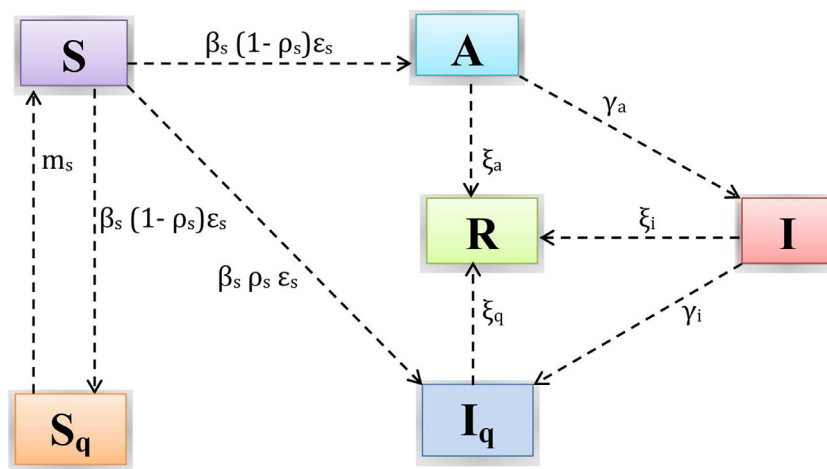
The ongoing coronavirus, SARS-CoV-2 epidemic has been announced a pandemic by the World Health Organization (WHO) on March 11, 2020 [1], and in the first phase the Govt. of India has announced 21 days nationwide lockdown from March 25, 2020 to April 14, 2020, and in the second phase the lockdown has been extended to May 03, 2020 to prevent stage-III spreading of the virus or human-to-human transmission [2]. According to the WHO report dated April 18, 2020 reported 4,619,477 total confirmed cases and 311,847 confirmed deaths worldwide [3]. COVID-19 or SARS-CoV-2 has already surpassed the earlier history of two coronavirus outbreaks, namely SARS-CoV (Severe Acute Respiratory Syndrome coronavirus) and MERS-CoV (Middle East respiratory syndrome coronavirus), posing the substantial endanger to the world-wide public health as well as global economy after the 2nd world war [4]. SARS-CoV-2 exhibits distinctive epidemiological traits collated with coronavirus epidemics of SARS-CoV and MERS-CoV.

The outbreak was first announced by the “Health Commission of Hubei province”, China, a cluster of unexplained cases of pneumonia of unknown etiology (unknown cases) [5], which is lethal, was first identified in Wuhan city of Hubei province, China, on December 31, 2019 [6,7]. After that an exceptionally large number of patients were diagnosed with SARS-CoV-2 in mainland of China, prodding Chinese Governments to initiate stringent measures to control the epidemic [5]. In spite of these precautions, SARS-CoV-2 pandemic evolved in the following months. The confirmed cases of COVID-19, the symptoms like sneeze, fever, or even a runny nose, dry cough, fatigue, breathing problem and bilateral lung infiltration to severely ill and dying. Due to the human mobility, this communicable disease has now spread throughout the world, making USA and Europe as new epicenters [4].

The first indigenous case of COVID-19 in India was reported on January 30, 2020 in Thrissur district of Kerala and the patient, a student of Wuhan University, China [8]. The authorities recommend the level of infection could be assload as the India's testing rates are very poor among the world [9]. The rate of infection of SARS-CoV-2 in India is chronicled to be 1.7 (that is, one coronavirus positive infects 1.7 in India), materially lower than in hot zones [10]. The estimated basic reproduction number  $R_0$  for

\* Corresponding author.

E-mail address: [subhaskhajanchi@gmail.com](mailto:subhaskhajanchi@gmail.com) (S. Khajanchi).



**Fig. 1.** Schematic representation of the model. The schematic flow diagram represents the interplays among different stages of infection in the model SARII<sub>q</sub>S<sub>q</sub>: susceptible or uninfected (S), asymptomatic or mildly symptomatic (A), recovered or healed (R), infected or symptomatic (I), isolated infected (I<sub>q</sub>) and quarantined susceptible (S<sub>q</sub>) individuals.

COVID-19 ranges from 2.0 to 3.5 [11,12], that seems similar, or perhaps higher than that of SARS-CoV and MERS-CoV. Higher level of viral loads for COVID-19 were observed in upper respiratory specimens of symptomatic patients resulting little or no symptoms, with a viral shedding pattern similar to that of influenza viruses [13]. Thus, uncertain viral transmission may take part a crucial and underestimated role in sustaining the epidemic.

India has suspended all the tourist visas as the majority of confirmed coronavirus cases were connected to other countries [14]. Governments will not be competent to reduce both fatalities from SARS-CoV-2 epidemic and the economic impact of viral outbreak. Maintaining the fatality rate as low as possible will be the utmost importance for the populations; therefore the Governments must put in place measures to mitigate the unavoidable economic downturn. Due to absence of any specific pharmaceutical interventions, government of various countries are imposing different strategies to prevent this outbreak and the lockdown is the most common one. As for examples, the measures adopted in this time incorporated social distancing, closing schools, universities, offices, churches, bars, avoid mass gatherings, other social places as well as contact of cases (quarantine, surveillance, contact tracing) [15]. On March 19, 2020 the Govt. of India suspended all the international flights till March 22, 2020 [16], and on March 23, 2020 the union Govt. also suspended all the domestic flights till March 25, 2020 [17] to maintain the social distancing among the people. The prime minister of India has announced a 14 hours voluntary public curfew ('Janata Curfew') on March 22, 2020 as a precautionary measure to combat against COVID-19. The Govt. of India followed it up with lockdowns on March 23, 2020 to prevent the emanating threat in 75 districts across the country including major cities where the COVID-19 infection was endemic [18]. Furthermore, on March 24, 2020 the Govt. of India has ordered a nationwide lockdown for 21 days, overwhelming the entire 1.3 billion public in India [19], and the lockdown has been extended to May 03, 2020 to prevent stage-III spreading of the virus or human-to-human transmission [20].

Predictive mathematical models play a key role to understand the course of the epidemic and for designing strategies to contain quickly spreading infectious diseases in lack of any specific antivirals or effective vaccine [21–25]. In the year 1927, Kermack and McKendrick [26] developed a fundamental epidemic model for human-to-human transmission to describe the dynamics of populations through three mutually exclusive phases of infection, namely susceptible (S), infected (I) and removed (R) classes. Mathematical modeling of infectious diseases are now ubiquitous and

many of them can precisely depict the dynamic spread of particular epidemics. Several mathematical models have been developed to study the transmission dynamics of COVID-19 pandemic. A Bats-Hosts-Reservoir-People network model has been developed by Chen et al. [27] to study the transmission dynamics of novel coronavirus. Lin et al. [28] extended the SEIR (susceptible-exposed-infectious-removed) compartment model to study the dynamics of COVID-19 incorporating public perception of risk and the number of cumulative cases. Khajanchi et al. [29] studied an extended SEIR model to study the transmission dynamics of COVID-19 and perform a short-term prediction based on the data from India. A fractional order extended SEIR (susceptible-exposed-infectious-removed) model studied by Khan and Atangana et al. [30] to portray the dynamics of COVID-19 outbreak in Wuhan, China, where the authors computed the basic reproduction  $R_0 = 2.4829$ . Wu et al. [6] studied a SEIR model to investigate the dynamics of 2019-nCoV human-human transmission dynamics based on the data from Wuhan, China from December 31, 2019 to January 28, 2020 and calculate the basic reproduction number was approximately 2.68. Wu et al. [31] used a SIR model to delineate the transmission dynamics of COVID-19 and also estimate the clinical severity for the coronavirus. To study the dynamics of COVID-19, a stochastic transmission model also developed by Kucharski et al. [32]. Analysis of viral dynamics using mathematical models have helped gain insights into the understanding of viral infections such as tuberculosis, dengue, and Zika virus [33–36]. Here, we developed a new epidemiological mathematical model for novel coronavirus or SARS-CoV-2 epidemic in India that extends the standard SEIR compartment model, alike to that studied by Tang et al. [37] for COVID-19. The transmission dynamics of our proposed model for COVID-19 is illustrated in the Fig. 1.

## 2. Dynamic model without effective control measures

We develop here a classical SEIR (susceptible-exposed-infectious-removed)-type epidemiological model by introducing contact tracing and other interventions such as quarantine, lockdown, social distancing and isolation that can represent the overall dynamics of novel coronavirus or COVID-19 (SARS-CoV-2). The model, named SARII<sub>q</sub>S<sub>q</sub>, monitors the dynamics of six compartments (classes), namely susceptible individuals (S) (uninfected), quarantined susceptible individuals (S<sub>q</sub>) (quarantined at home), infectious but not yet symptomatic or asymptomatic infectious individuals (A), infected or infectious with symptoms/clinically ill (I), isolated infected individuals (I<sub>q</sub>) (infected or life-threatening

or detected) and the recovered compartment ( $R$ ) (no more infectious). The total size of the individuals is  $N = S + S_q + A + I + I_q + R$ . Asymptomatic individuals have been exposed to the virus, but have not yet developed clinical symptoms of the COVID-19 or SARS-CoV-2 [38]. In our model, quarantine describes the separation of coronavirus infected populations from the susceptible individuals before progression of clinical symptoms, whereas the isolation refers to the dissociation of coronavirus infected populations with such clinical symptoms. The rate of change in each compartments at any time  $t$  is represented by the following system of nonlinear ordinary differential equations:

$$\begin{cases} \frac{dS}{dt} = \Lambda_s - (\beta_s + \rho_s(1 - \beta_s))\varepsilon_s S \frac{I}{N} - \delta S + m_s S_q, \\ \frac{dS_q}{dt} = (1 - \beta_s)\varepsilon_s \rho_s S \frac{I}{N} - (m_s + \delta)S_q, \\ \frac{dA}{dt} = \beta_s(1 - \rho_s)\varepsilon_s S \frac{I}{N} - (\gamma_a + \xi_a + \delta)A, \\ \frac{dI}{dt} = \gamma_a A - (\gamma_i + \xi_i + \delta)I, \\ \frac{dI_q}{dt} = \beta_s \varepsilon_s \rho_s S \frac{I}{N} + \gamma_i I - (\xi_q + \delta)I_q, \\ \frac{dR}{dt} = \xi_a A + \xi_i I + \xi_q I_q - \delta R, \end{cases} \quad (1)$$

the model is supplemented by the following non-negative initial values:

$$\begin{aligned} S(t_0) &= S_0, \quad S_q(t_0) = S_{q0}, \quad A(t_0) = A_0, \quad I(t_0) = I_0, \\ I_q(t_0) &= I_{q0} \quad \text{and} \quad R(t_0) = R_0. \end{aligned} \quad (2)$$

Herein,  $t \geq t_0$  represents time in days and  $t_0$  indicates the starting date for the system of the coronavirus epidemic.

In our model construction,  $\beta_s$  represents the probability of transmission per contact between an infective and a susceptible class, and  $\varepsilon_s$  is denoted by the daily contact rate per unit of time. Here the parameter  $\beta = \beta_s \varepsilon_s$  is explicitly associated with the measures like lock-down, social distancing, shaking hand, coughing and sneezing etc., which exactly decrease the number of social contacts. By enforcing contact tracing, a proportion  $\rho_s$ , of individuals exposed to the coronavirus is quarantined. The quarantined classes can either move to the compartment  $S_q$  or  $I_q$ , depending on whether they are effectively infected individuals or not, whereas the another proportion  $1 - \rho_s$ , consists of populations exposed to the coronavirus who are missed from contact tracing and move to the infectious class  $I$  (once infected) or remaining in susceptible class  $S$  (if uninfected). Then the quarantined classes, if uninfected (or infected), move to the class  $S_q$  (or  $I_q$ ) at a rate of  $(1 - \beta_s)\rho_s \varepsilon_s$  (or  $\beta_s \rho_s \varepsilon_s$ ). Those who are not quarantined individuals, but asymptomatic infectious individuals, will move to the asymptomatic compartment  $A$  at the rate of  $\beta_s(1 - \rho_s)\varepsilon_s$ . The proportion  $\rho_s$  ( $0 < \rho_s < 1$ ) encapsulates the effectiveness of different individual preventive measures, such as cough etiquettes, frequently wash hand by soap and use sanitizer etc. We utilize a constant  $m_s$  to indicate the transition rate from the susceptible individuals to the (COVID-19) quarantined susceptible class due to fever and/or illness-like clinical symptoms. We symbolize  $\xi_a$ ,  $\xi_i$  and  $\xi_q$  are the rates of recovery individuals of asymptomatic class, symptomatic or clinically ill patients and isolated individuals, respectively. Our model introduces some demographic effects by considering a proportional natural decay rate  $\delta$  in each of the six individuals, and  $\Lambda_s$  represents the constant inflow of susceptible individuals. Asymptomatic population develop to infected population at the rate  $\gamma_a$ , so the average time spent in the asymptomatic class is  $\frac{1}{\gamma_a}$  per unit time. In similar fashion,  $\frac{1}{\gamma_i}$  represents the mean duration for infected individuals. We ignore the rate of prob-

ability of transforming susceptible again after having cured (recovered) from the disease infection. It is to be noted that our  $SARII_q S_q$  model did not take into account many important ingredients that take part a key role in the transmission dynamics of COVID-19 such as the influence of the latency period, the inhomogeneous disease transmission network, the influence of the measures already considered to fight the coronavirus diseases, the features of the individuals (for example, the influence of the stage-structure, individuals who are already medically unfit). Some recent mathematical models incorporate asymptomatic such as in Ndairou et al. [39] but others do not include them [40].

## 2.1. Basic reproduction number

The basic reproduction number, symbolized by  $R_0$ , is 'the expected number of secondary cases produced, in a completely susceptible population, by a typical infective individual' [41,42]. The dimensionless basic reproduction number provides a threshold, which play a crucial role in determining the disease persists or dies out from the individual. In a more general way the basic reproduction number  $R_0$  can be stated as the number of new infections created by a typical infective population at a disease free equilibrium.  $R_0 < 1$  determines on average an infected population creates less than one new infected population during the course of its infective period, and the infection can die out. In reverse way,  $R_0 > 1$  determines each infected population creates, on average, more than one new infection, and the disease can spread over the population. The basic reproduction number  $R_0$  can be computed by using the concept of next generation matrix [41,42]. In order to do this, we consider the nonnegative matrix  $\mathcal{F}$  and the non-singular  $M$ -matrix  $\mathcal{V}$ , expressing as the production of new-infection and transition part respectively, for the system (1), are described by

$$\mathcal{F} = \begin{bmatrix} \beta_s(1 - \rho_s)\varepsilon_s S \frac{I}{N} \\ 0 \\ \beta_s \varepsilon_s \rho_s S \frac{I}{N} \end{bmatrix}, \quad \mathcal{V} = \begin{bmatrix} (\gamma_a + \xi_a)A \\ -\gamma_a A + (\gamma_i + \xi_i)I \\ -\gamma_i I + \xi_q I_q \end{bmatrix}.$$

The variational matrix of the model (1) computed at the infection free state ( $S_q = A = I = I_q = R = 0$ ) gives

$$F = \begin{bmatrix} 0 & \beta_s(1 - \rho_s)\varepsilon_s \frac{S}{N} & 0 \\ 0 & 0 & 0 \\ 0 & \beta_s \varepsilon_s \rho_s \frac{S}{N} & 0 \end{bmatrix}, \quad V = \begin{bmatrix} \gamma_a + \xi_a & 0 & 0 \\ -\gamma_a & \gamma_i + \xi_i & 0 \\ 0 & -\gamma_i & \xi_q \end{bmatrix}.$$

The basic reproduction number  $\mathcal{R}_0 = \rho(FV^{-1})$ , where  $\rho(FV^{-1})$  represents the spectral radius for a next generation matrix  $FV^{-1}$ . Thus, the basic reproduction number of the system (1) is

$$\rho(FV^{-1}) = \mathcal{R}_0 = \frac{\beta_s(1 - \rho_s)\varepsilon_s \gamma_a}{(\gamma_a + \xi_a)(\gamma_i + \xi_i)}. \quad (3)$$

## 3. Model calibration and data source

We calibrated our  $SARII_q S_q$  model for COVID-19 to the daily new infected cases and cumulative confirmed cases of SARS-CoV-2 viruses for 17 provinces of India, namely Andhra Pradesh, Delhi, Gujarat, Hariyana, Jammu & Kashmir, Karnataka, Kerala, Madhya Pradesh, Maharashtra, Punjab, Rajasthan, Tamil Nadu, Telangana, Uttar Pradesh, West Bengal, Bihar, Odisha and the overall India with real data till April 30, 2020. The source of real data are listed in the Table 7. The description of the  $SARII_q S_q$  model are given in Table 1, list of key estimated parameter values are specified in Table 2 and estimated initial population size are given in the Table 3. By calibrating the  $SARII_q S_q$  model parameters with real data up to 30 April 2020, we make an attempt to forecast the evolution of the epidemic in India and 17 provinces of India. In the model exploration, we did not consider the demographic effects

**Table 1**Parameters of the  $SARII_qS_q$  model system (1) and their biological interpretations.

Symbol	Description	Values & source
$\Lambda_s$	net inflow rate of susceptible individuals	0
$\beta_s$	probability of disease transmission	Estimated
$\rho_s$	quarantined rate of susceptible individuals	Estimated
$\varepsilon_s$	contact rate of entire individuals	Estimated
$\delta$	natural mortality rate	0
$m_s$	rate at which quarantined susceptible individuals are released into uninfected class	$\frac{1}{24}$ [46]
$\gamma_a$	probability rate at which asymptomatic individuals develops clinically symptoms	Estimated
$\xi_a$	recovery rate of asymptomatic infected individuals	Estimated
$\gamma_i$	probability rate at which infected individuals become isolated	$7.151 \times 10^{-2}$ [37]
$\xi_i$	rate of recovery for infected individuals	Estimated
$\xi_q$	recovery rate of isolated infected individuals	$\frac{1}{748}$ [46]

**Table 2**

Parameter values estimated from the observed daily new cases of COVID-19 in India and for 17 different provinces of India. Six sensitive parameters  $\beta_s$ ,  $\rho_s$ ,  $\varepsilon_s$ ,  $\gamma_a$ ,  $\xi_a$  and  $\xi_i$  are estimated among nine system parameters. The parameters are estimated by fitting time series solution of the  $SARII_qS_q$  model with the observed daily new cases of COVID-19 using the least square technique. Also, we compute the basic reproduction number  $R_0$  for 17 different provinces of India and the overall India.

India & Provinces	$\beta_s$	$\rho_s$	$\varepsilon_s$	$\gamma_a$	$\xi_a$	$\xi_i$	$R_0$
India	0.8799	0.3199	14.8300	0.0168	0.7100	0.0286	2.0490
Andhra Pradesh	1.0228	0.1000	5.6000	0.1000	0.4679	0.6840	1.2014
Delhi	0.3100	0.3360	9.6600	0.2840	0.6100	0.4034	1.3301
Gujarat	0.5748	0.0009	9.1140	0.1220	0.3240	0.9276	1.4342
Haryana	0.6699	0.2828	13.1220	0.1879	0.9000	0.8948	1.1269
Jammu & Kashmir	0.5569	0.4600	12.7564	0.1786	0.3968	0.8520	1.2894
Karnataka	0.6520	0.7360	13.3399	0.4499	0.6740	0.7399	1.1328
Kerala	1.3020	0.0090	12.7999	0.0090	0.4060	0.2080	1.2814
Madhya Pradesh	1.0620	0.8659	8.7600	0.2740	0.5600	0.1780	1.6415
Maharashtra	1.1400	0.1599	10.5999	0.0099	0.1660	0.1599	2.4690
Punjab	1.6099	0.1000	5.3200	0.0208	0.6399	0.1000	1.4149
Rajasthan	1.5299	0.3100	6.7999	0.0860	0.5020	0.7500	1.2779
Tamil Nadu	0.5499	0.8619	11.7960	0.8660	0.6959	0.3340	1.2248
Telangana	0.9160	0.7479	15.6400	0.2000	0.5860	0.5340	1.5177
Uttar Pradesh	1.3860	0.0208	9.0200	0.0882	0.8700	0.7940	1.3019
West Bengal	1.2180	0.4060	14.5139	0.0202	0.2800	0.3188	1.8120
Bihar	1.4100	0.5500	4.5169	0.0900	0.0990	0.8061	1.5552
Odisha	0.9700	0.0660	11.0700	0.0380	0.3000	0.8000	1.2938

**Table 3**

Initial population size used in numerical simulations for India and 17 different provinces of India.

India & Provinces	$S(0)$	$S_q(0)$	$A(0)$	$I(0)$	$I_q(0)$	$R(0)$
India	902,654	10,785	114	1	0	0
Andhra Pradesh	94,558	950	95	1	0	0
Delhi	58,210	505	16	1	0	0
Gujarat	68,774	715	76	2	0	0
Haryana	30,191	351	15	1	0	0
Jammu & Kashmir	21,695	750	12	2	0	0
Karnataka	60,156	900	12	2	0	0
Kerala	32,217	2151	37	1	0	0
Madhya Pradesh	41,712	978	21	4	0	0
Maharashtra	135,694	1000	140	18	0	0
Punjab	21,205	200	110	1	0	0
Rajasthan	103,704	1100	115	3	0	0
Tamil Nadu	119,748	1197	28	1	0	0
Telangana	19,278	201	10	1	0	0
Uttar Pradesh	78,013	708	39	1	0	0
West Bengal	16,525	575	18	1	0	0
Bihar	22,672	275	22	2	0	0
Odisha	34,133	301	5	1	0	0

because of the short epidemic time scale in compare to the demographic time scale, that is,  $\Lambda_s = \delta = 0$ .

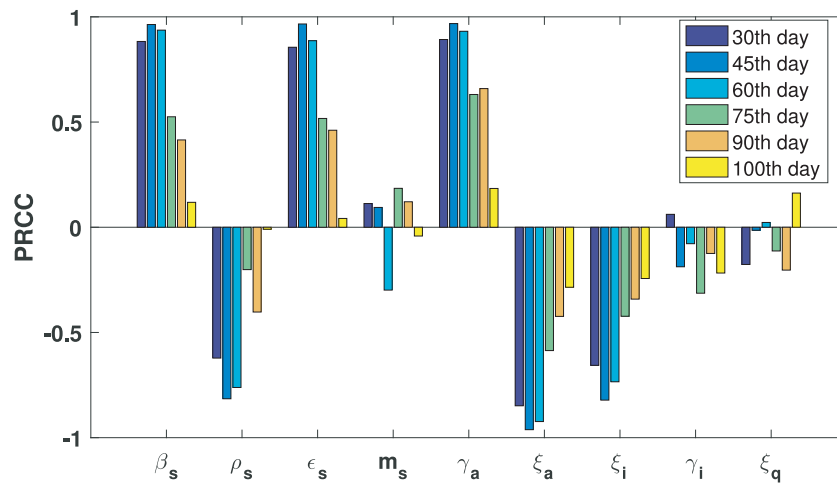
### 3.1. PRCC sensitivity analysis

To recognize the most influential parameters with respect to clinically ill infected population, we conduct a sensitivity analysis

by evaluating Partial Rank Correlation Coefficient (PRCC) technique for all the input parameters against the variable ( $I$ ). The set of most sensitive parameters, which have the ability to control the outbreak of COVID-19, can be estimated from 17 different provinces of India and the Republic of India. Our  $SARII_qS_q$  model has 11 parameters (among which  $\Lambda_s = \delta = 0$ ) for which we varied all the system parameters simultaneously. PRCC measures the relationship among the state variables and each of the model parameters. As a result PRCC aid to find the significant parameters, which contribute most to the model variability. For the model simulation, we have chosen the PRCC values between  $-1.0$  and  $+1.0$ .

Due to the method developed by Marino et al. [43], we conducted Latin hypercube sampling and generated 3200 samples to identify which parameters consistently affect each model individuals in time. The indices are assessed at six time points 30, 45, 60, 75, 90 and 100 days preceding to steady state, and the  $SARII_qS_q$  model variable selected for sensitivity analysis is infected or symptomatic individuals. The PRCC results has been shown in the Fig. 2 for six time points that represents the highest positively correlated parameters are the disease transmission rate  $\beta_s$ , contact rate  $\varepsilon_s$  of all the individuals, the probability rate  $\gamma_a$  at which the asymptomatic individuals develops clinically symptoms and highly negatively correlated parameters are the quarantined rate  $\rho_s$  of uninfected individuals, recovery rate  $\xi_a$  of asymptomatic infected individuals and the recovery rate  $\xi_i$  of infected individuals, accounts the most uncertainty with respect to the infected individuals. Thus, the PRCC analysis yields these six parameters  $\beta_s$ ,  $\rho_s$ ,  $\varepsilon_s$ ,  $\gamma_a$ ,  $\xi_a$ , and  $\xi_i$  are the most influential parameters out of 9 parameters.





**Fig. 2.** Parameter sensitivity. Partial rank correlation coefficients illustrating the sensitivity indices for the  $SARII_qS_q$  system for COVID-19 or SARS-CoV-2 model parameters with symptomatic or infected individuals ( $I$ ) at different days with  $p < 0.01$ .

Therefore, we estimated these six parameters by using least square method.

### 3.2. Parameter estimation

The most important challenge in any mathematical model based study is to estimate the model parameters and the initial population size. The solution of the  $SARII_qS_q$  model system (1) depends on both the parameter values and initial population size. The model parameters have been estimated assuming the initial population size and fitting the model simulation with the observed COVID-19 cases. The assumed initial population sizes are presented in the Table 3. We have estimated six parameters, probability of disease transmission ( $\beta_s$ ), quarantined rate of susceptible individuals ( $\rho_s$ ), contact rate of entire individuals ( $\epsilon_s$ ), probability rate at which asymptomatic individuals develop clinical symptoms ( $\gamma_a$ ), recovery rate of asymptomatic infected individuals ( $\xi_a$ ) and rate of recovery for infected individuals ( $\xi_i$ ) as these parameters are more sensitive in PRCC analysis. The parameters are estimated from the observed daily new COVID-19 or SARS-CoV-2 viruses. Although, we have shown the plot validating model simulation for cumulative cases, the parameters are not estimated from confirmed cumulative COVID-19 cases to optimize the error in parameter estimation [44] and errors are listed in the Table 5. First we have applied a five days moving average filter, which is a low pass filter, to smooth the random variation in the observed daily new COVID-19 cases. The observed daily COVID-19 cases are fitted with the model simulation by using least square method, which locally minimizes the sum of the square of errors. The square of sum of the error computed as  $\sum_{i=1}^n (C(i) - S(i))^2$ , where  $C(i)$  represents the observed daily new COVID-19 cases on  $i$ th day,  $S(i)$  is the  $SARII_qS_q$  model simulation on  $i$ th day and  $n$  is the sample size of the observed data. It has been observed that different set of parameter values can minimize the sum of the square of errors between the observed daily new COVID-19 cases and the  $SARII_qS_q$  model simulation but we have considered the set of parameter values, which produce realistic  $R_0$ . Varying the random values of six parameters and finding the minimum error requires computation power. We have varied the value of  $\beta_s$  from 0 to 1.75,  $\epsilon_s$  from 0 to 16.0 and the other parameters from 0 to 1.0. Initially, we have considered five intervals in each parameter value range, that is, 15,625 computation steps and around 60 seconds is required to run the FORTRAN program in a AMD Ryzen 5 3400G CPU. When we consider 10 intervals for each parameter value range requires 1,000,000

computation steps and which requires around one hour to run. Hence, to reduce the computation time, we have run the FORTRAN program in subsequent subsets of value range by considering 5 intervals for each parameter value range. The estimated values of the parameters are presented in the Table 2. These estimated values are used to predict the end of the COVID-19 pandemic in India and to study the effect of lockdown.

### 4. Numerical simulation

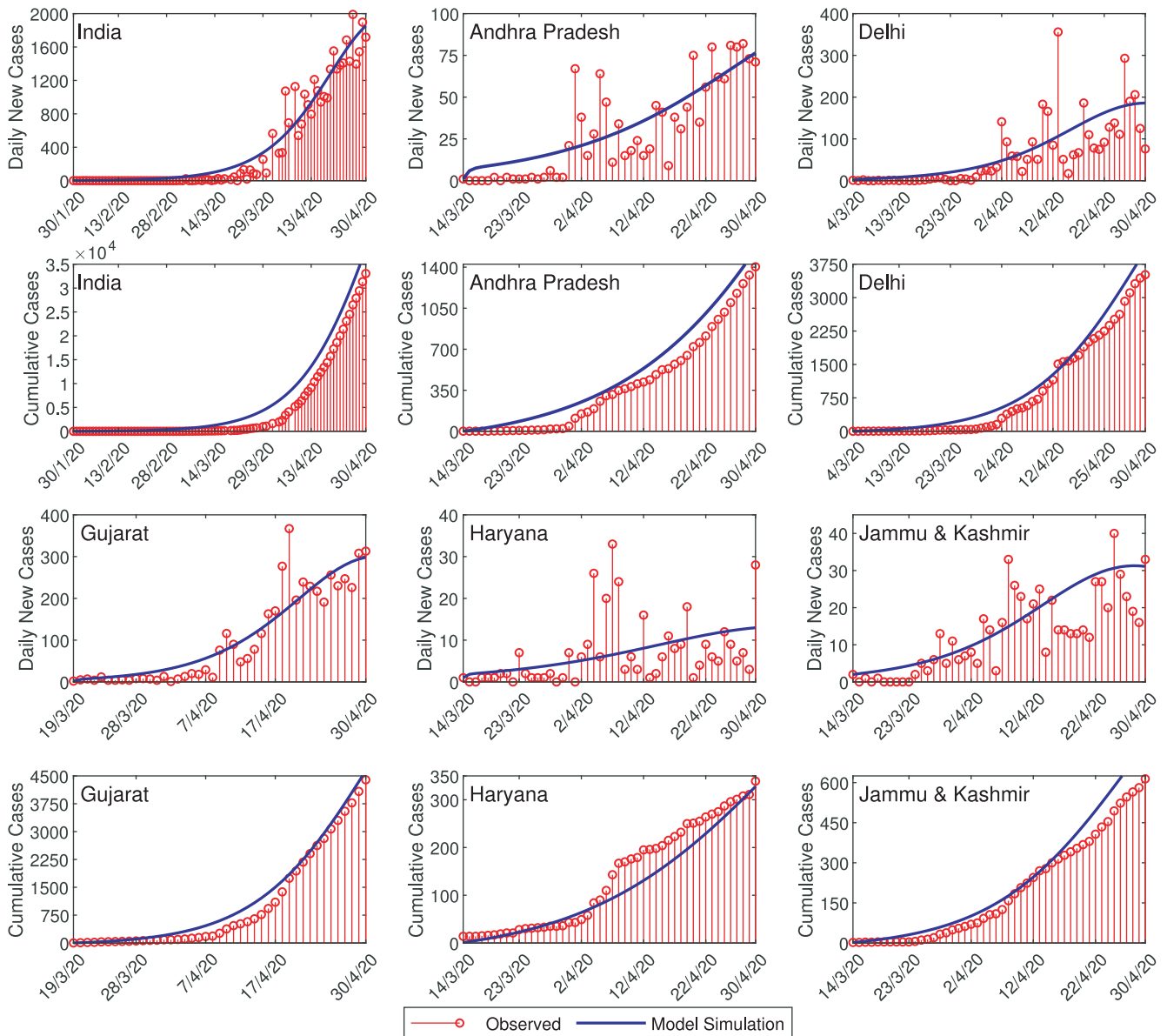
Initially we have validated the model simulation with the observed COVID-19 cases. The sources and duration of the observed data has been presented in Table 7. Model simulated from the first date of coronavirus infection and up to 30 April, 2020 for whole India and for seventeen states of India. The model simulation fitted with the observed daily new COVID-19 cases and cumulative COVID-19 cases. The parameter values are taken from Table 1 and the Table 2 and the initial population size from Table 3. Figs. 3–5 shows the observed COVID-19 cases and model simulations. Observed data points are displayed in the red dot histogram and the blue curve represents the best fitting curve for the model simulation. The first and third rows in each of these figures represents the daily new cases of coronavirus diseases, whereas the second and fourth rows represents the cumulative confirmed cases of COVID-19. In all the states of India except Telangana, the number of daily new COVID-19 cases is increasing and the model capture this increasing trend very well. However, in Telanga, the number of daily new COVID-19 cases starts decreasing from 12 April, 2020 onwards and our model captures this trend successfully (see the Fig. 5).

To measure the accuracy of fitting, we have computed Mean Absolute Error ( $E_{MAE}$ ) and Root Mean Square Error ( $E_{RMSE}$ ). The  $E_{MAE}$  and  $E_{RMSE}$  are defined as follows:

$$E_{MAE} = \frac{\sum_{i=1}^n |C(i) - S(i)|}{n},$$

$$E_{RMSE} = \sqrt{\frac{\sum_{i=1}^n (C(i) - S(i))^2}{n}}$$

where  $C(i)$  represents the observed value,  $S(i)$  is the model simulation and  $n$  is the sample size of the observed data. The values of  $E_{MAE}$  and  $E_{RMSE}$  for the seventeen states and for whole India are listed in the Table 5. Our model performs excellently well in all the provinces. The values of  $E_{MAE}$  and  $E_{RMSE}$  for whole India are 114.0705 and 187.4646 respectively. Also for all the states of India,



**Fig. 3.** Model estimation based on real data. The  $SARII_qS_q$  model fitted with the real data on daily new cases of COVID-19 in India and 5 provinces of India, namely Andhra Pradesh, Delhi, Gujarat, Haryana and Jammu & Kashmir. Observed data points are displayed in the red dot histogram and the blue curve represents the best fitting curve for the  $SARII_qS_q$  model. The first and third rows represents the daily new cases of coronavirus diseases, whereas the second and fourth rows represents the cumulative confirmed cases of COVID-19. The estimated parameter values are listed in the Table 2. The initial values used for this parameter values are presented in the Table 3. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the value of  $E_{MAE}$  varies from 2.3084 to 53.2686 and the value of  $E_{RMSE}$  varies from 3.7892 to 54.9932.

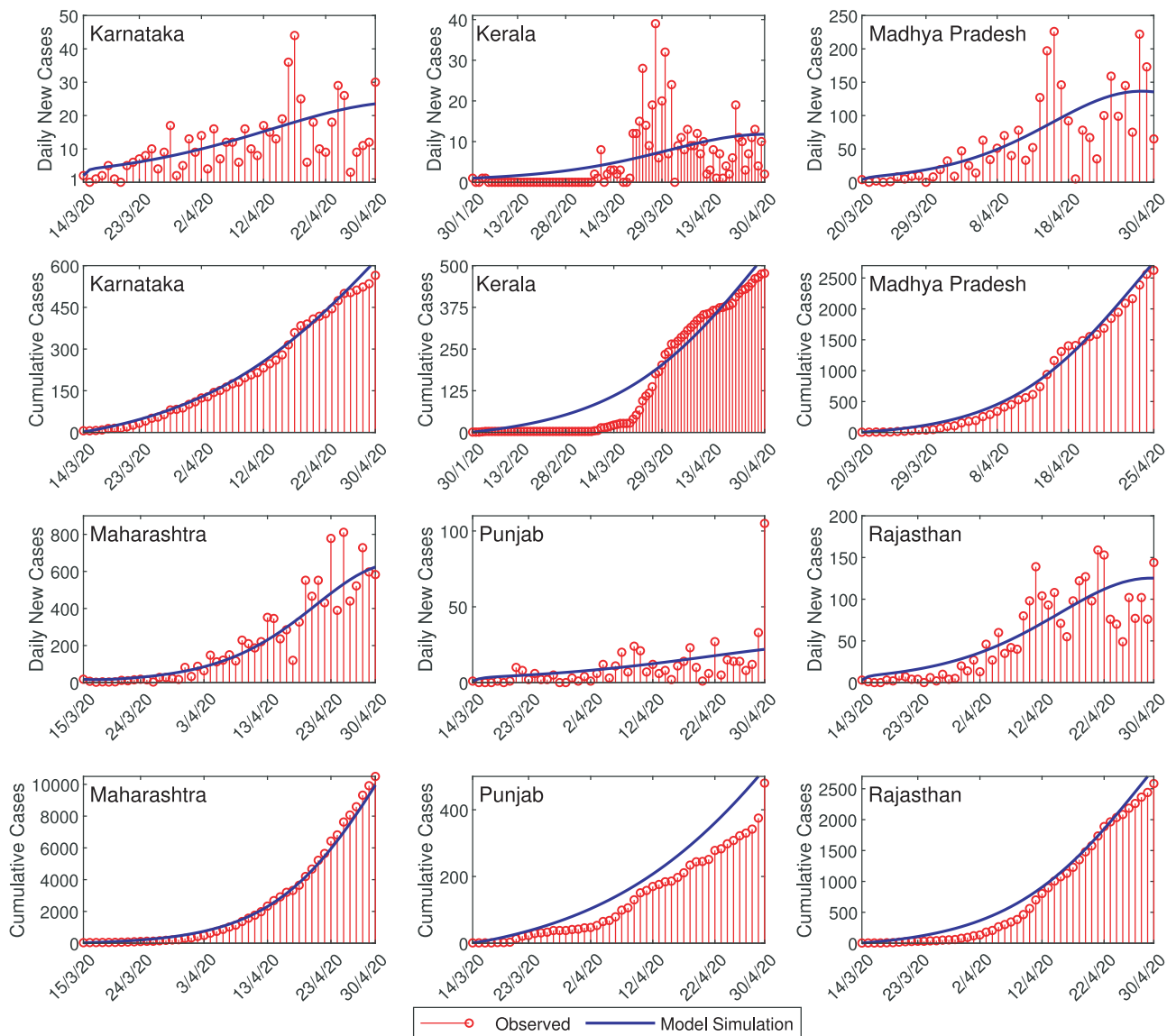
#### 4.1. Sensitivity analysis for $R_0$

To describe how best to minimize individuals impermanence and morbidity due to SARS-CoV-2, it is important to see the relative significance of various ingredients responsible for disease transmission. Transmission of SARS-CoV-2 is directly related to the basic reproduction number  $R_0$ . We compute the sensitivity indices for  $R_0$  for the parameters of the  $SARII_qS_q$  model. This indices apprise us how important each parameter is to disease transmission. Sensitivity analysis is mainly used to describe the robustness of the model predictions to the parameters, as there are generally errors in collection of data and assumed parameter values. Sensitivity indices quantify the relative change in a state variable when a parameter alters. The normalized forward sensitivity index for  $R_0$ ,

with respect to the disease transmission coefficient  $\beta_s$  can be defined as follows:

$$\Upsilon_{\beta_s}^{R_0} = \frac{\partial R_0}{\partial \beta_s} \times \frac{\beta_s}{R_0} = 1,$$

which demonstrates that  $R_0$  is a increasing function of  $\beta_s$ . This implies that probability of disease transmission has a high influence on COVID-19 control and management. The sensitivity indices of other parameters are given in the Table 4. In the Table 4, some of the indices are positive (and some are negative) which means if the parameter increases then increase the value of  $R_0$  (and if the parameter increases then decrease the value of  $R_0$ ). To control the outbreak of SARS-CoV-2, we must select the most sensitive parameters who have most influence to reduce the diseases. As for example, the transmission rate  $\beta_s$  has an impact in reducing the COVID-19 diseases, which can easily be observed from the Table 4. Therefore, we draw the contour plots for  $R_0$  in the Figs. 8 and 9 depen-



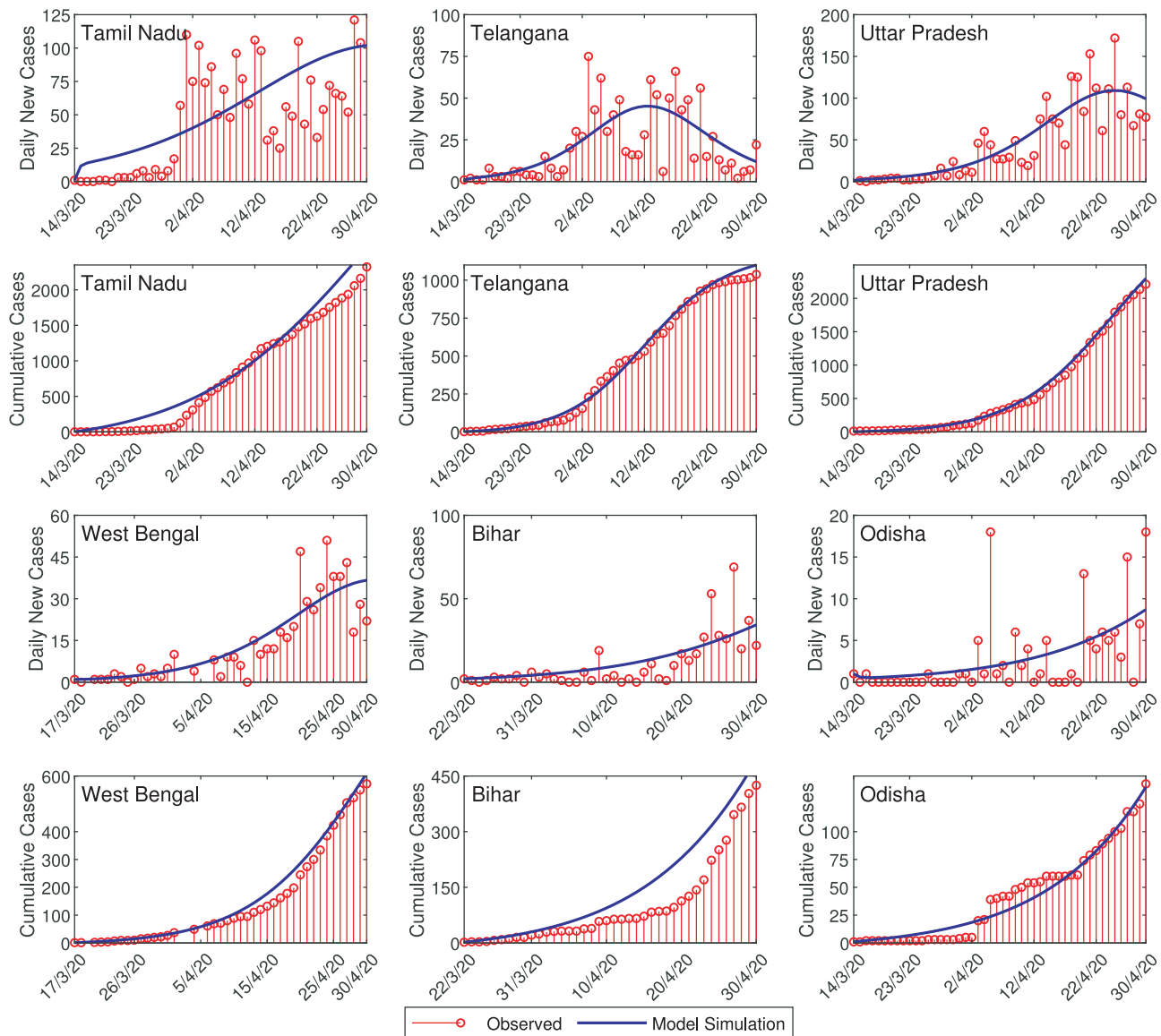
**Fig. 4.** Model estimation based on real data. The  $SARIL_qS_q$  model fitted with the real data on daily new cases of COVID-19 for 6 provinces of India, namely Karnataka, Kerala, Madhya Pradesh, Maharashtra, Punjab, and Rajasthan. Observed data points are shown in the red dot histogram and the blue curve represents the best fitting curve for the  $SARIL_qS_q$  model. The first and third rows represent the daily new cases of coronavirus diseases, whereas the second and fourth rows represent the cumulative confirmed cases of COVID-19. The estimated parameter values are listed in Table 2. The initial values used for this parameter values are presented in Table 3. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

**Table 4**

Sensitivity indices for the basic reproduction number  $R_0$  to the parameters for the COVID-19  $SARIL_qS_q$  model for 17 different provinces of India and overall India, assessed at the baseline parameter values listed in the Tables 1 and 2.

India & Provinces	$\beta_s$	$\rho_s$	Sensitivity $\varepsilon_s$	$\gamma_a$	Index $\gamma_i$	$\xi_a$	$\xi_i$
India	1.000	0.470	1.000	0.977	-0.714	-0.977	-0.286
Andhra Pradesh	1.000	-0.111	1.000	0.824	-0.095	-0.824	-0.905
Delhi	1.000	-0.506	1.000	0.682	-0.151	-0.682	-0.849
Gujarat	1.000	-0.001	1.000	0.726	-0.072	-0.726	-0.928
Haryana	1.000	-0.394	1.000	0.827	-0.074	-0.827	-0.926
Jammu & Kashmir	1.000	-0.852	1.000	0.689	-0.077	-0.689	-0.923
Karnataka	1.000	-2.788	1.000	0.599	-0.088	-0.599	-0.912
Kerala	1.000	-0.001	1.000	0.978	-0.256	-0.978	-0.744
Madhya Pradesh	1.000	-6.462	1.000	0.672	-0.287	-0.671	-0.713
Maharashtra	1.000	-0.191	1.000	0.944	-0.309	-0.944	-0.691
Punjab	1.000	-0.111	1.000	0.969	-0.417	-0.969	-0.583
Rajasthan	1.000	-0.449	1.000	0.854	-0.087	-0.854	-0.913
Tamil Nadu	1.000	-6.241	1.000	0.446	-0.176	-0.446	-0.824
Telangana	1.000	-2.967	1.000	0.746	-0.118	-0.746	-0.882
Uttar Pradesh	1.000	-0.021	1.000	0.908	-0.083	-0.908	-0.917
West Bengal	1.000	-0.684	1.000	0.933	-0.183	-0.933	-0.817
Bihar	1.000	-1.222	1.000	0.524	-0.081	-0.524	-0.919
Odisha	1.000	-0.071	1.000	0.888	-0.082	-0.888	-0.918



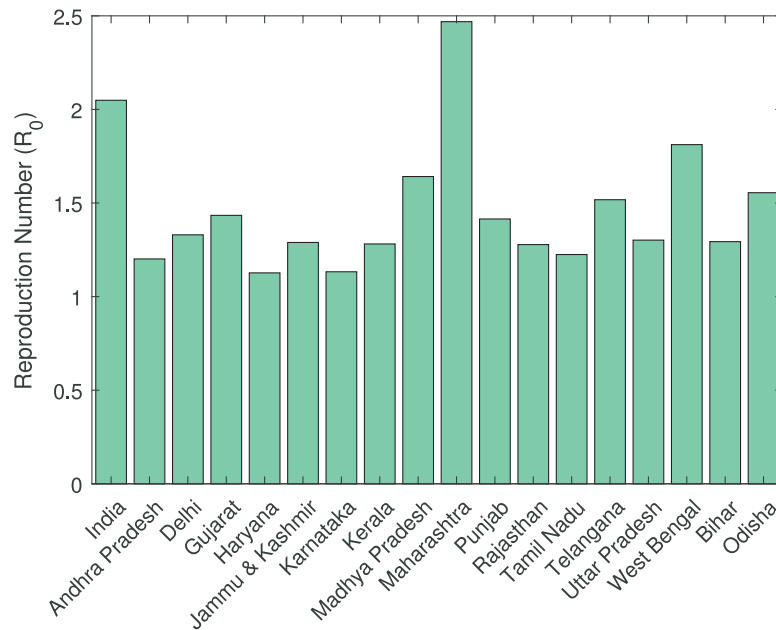


**Fig. 5.** Model estimation based on real data. The  $SARII_qS_q$  model fitted with the real data on daily new cases of COVID-19 for 6 provinces of India, namely Tamil Nadu, Telangana, Uttar Pradesh, West Bengal, Bihar, and Odisha. Observed data points are shown in the red dot histogram and the blue curve represents the best fitting curve for the  $SARII_qS_q$  model. The first and third rows represents the daily new cases of coronavirus diseases, whereas the second and fourth rows represents the cumulative confirmed cases of COVID-19. The estimated parameter values are listed in the Table 2. The initial values used for this parameter values are presented in the Table 3. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

**Table 5**

The values of the errors  $E_{MAE}$  and  $E_{RMSE}$  to estimate the accuracy of the  $SARII_qS_q$  model (1) simulation.

Provinces	India	Andhra Pradesh	Delhi	Gujarat	Haryana	Jammu & Kashmir
$E_{MAE}$	114.071	12.227	35.653	27.063	5.076	5.638
$E_{RMSE}$	187.465	15.446	54.993	42.333	7.442	7.439
Provinces	Karnataka	Kerala	Madhya Pradesh	Maharashtra	Punjab	Rajasthan
$E_{MAE}$	5.561	4.302	30.835	53.269	7.347	21.501
$E_{RMSE}$	7.796	6.677	44.477	86.456	13.777	27.058
Provinces	Tamil Nadu	Telangana	Uttar Pradesh	West Bengal	Bihar	Odisha
$E_{MAE}$	26.529	10.823	15.555	5.263	6.472	2.308
$E_{RMSE}$	31.018	15.129	23.011	8.609	10.049	3.789



**Fig. 6.** Basic reproduction number  $R_0$ . The bar-diagram represents the computation of basic reproduction number  $R_0$  from the estimated parameter values in the Table 2 for the Republic of India and 17 different provinces of India.

**Table 6**

Skewness and kurtosis of the numerically simulated daily new COVID-19 cases for 17 different provinces of India and the overall India. The number of active days computed starting from the first day of COVID-19 reported up to the day on which the daily new COVID-19 cases is less than one.

India & Provinces	Skewness	Kurtosis	Active days
India	1.328	0.341	254
Andhra Pradesh	0.559	-1.161	162
Delhi	1.068	-0.298	172
Gujarat	0.856	-0.748	107
Haryana	0.206	-1.406	117
Jammu & Kashmir	0.506	-1.222	108
Karnataka	0.293	-1.352	138
Kerala	0.285	-1.394	181
Madhya Pradesh	0.938	-0.601	131
Maharashtra	0.645	-1.066	103
Punjab	0.295	-1.364	155
Rajasthan	0.578	-1.147	114
Tamil Nadu	0.696	-0.971	175
Telangana	0.572	-1.167	66
Uttar Pradesh	0.741	-0.941	92
West Bengal	0.579	-1.161	92
Bihar	0.626	-1.101	138
Odisha	0.598	-1.131	157

dence on the rate of disease transmission probability  $\beta_s$  and the quarantine rate  $\rho_s$ . Contour plot shows that for the higher values of  $\beta_s$  the reproduction number  $R_0$  increases significantly, which means that the SARS-CoV-2 disease will persist among the human and spread throughout the community if the public not take the preventive measures. Thus, to control  $R_0$  must reduce the disease transmission coefficient  $\beta_s$  and increase the period of quarantine rate  $\rho_s$ . Thus, we may conclude that to end the COVID-19 outbreak enhance the quarantine and reduce the probability of disease transmission following contact tracing, social distancing, limit or stop theaters and cultural programme etc. For set of parameter values in the Table 1 and estimated parameter values in the Table 2, we plot a bar diagram for the basic reproduction number  $R_0$  in the Fig. 6. From the bar-diagram in the Fig. 6, it can be observed that the basic reproduction number  $R_0$  for the state Ma-

harashtra is too high, which indicates that the substantial outbreak of the COVID-19 in the state Maharashtra.

#### 4.2. Effect of lockdown

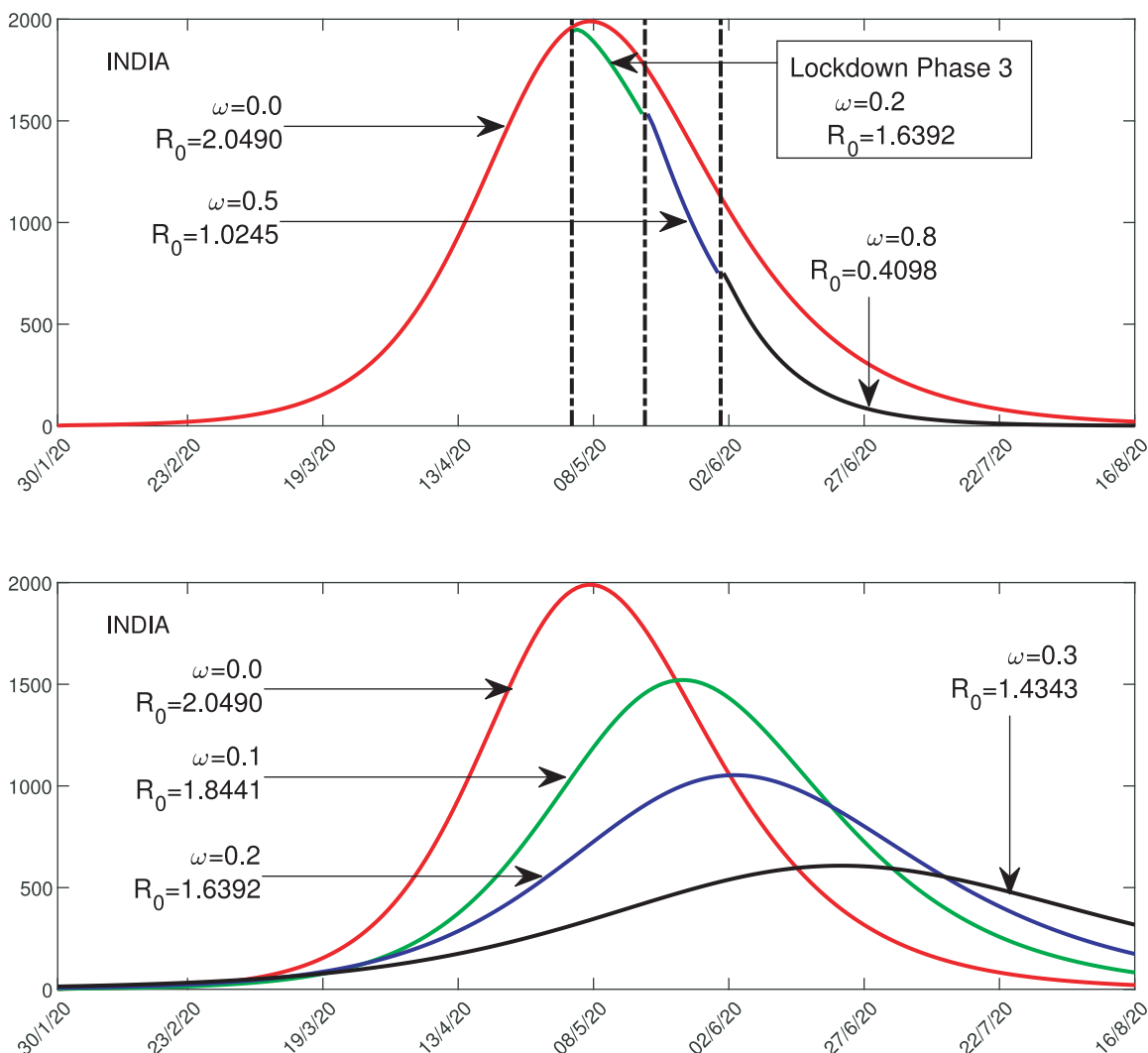
To investigate the influence of intervention policies in reducing the SARS-CoV-2 viruses, we modified the disease transmission rate  $\beta_s$  by  $(1 - \omega)\beta_s$  with  $0 < \omega < 1$ , where  $\omega$  encapsulates the efficacy of different individuals precautionary measures. Intervention strategies including the most recent lockdown progressively implemented since 25 March 2020, have influenced the spread of the outbreak. We also predict our model for 17 provinces of India and overall India, involving social distancing to control COVID-19.

For the estimated model parameters, before 3rd phase of lockdown ( $\omega = 0$ ) the basic reproduction number was  $R_0 = 2.0490$  that followed in a significant outbreak of COVID-19. The 3rd phase lockdown in India, has been implemented from 04 May 2020 and we divide into three time periods, first time period from 04 May 2020 to 17 May 2020, second time period from 18 May 2020 to 31 May 2020 and third time period from 01 June 2020 to 16 August 2020. In the first time window,  $R_0 = 1.6392$  as an outcome of the inclusion of lockdown, social distancing, hygiene precautions, and early measures by Indian Govt. (e.g. closing school, shopping mall etc.). In the second time window, the country-wide lockdown was implemented but not yet enforced, as result  $R_0 = 1.0245$ . The effectiveness of COVID-19 has been reduced due to lockdown but need to impose more restriction. The nation-wide lockdown was implemented in the third time period with strictly enforced, as a result  $R_0 = 0.4098$  reached below 1. The top panel of the Fig. 7 shows the effect of lockdown in three different time periods. Fig. 7 predict that on 17 May 2020, the cumulative number of COVID-19 infected population in India without Phase-3 lockdown is 70,948 but after strictly enforcing nation-wide Phase-3 lockdown in India the cumulative number of infected population reduced and it will be 64,838. Another interesting outcome can be noticed as an effect of nation-wide lockdown is the appearance of epidemic peak. The effect of lockdown not only reduce the infected population but also causes delay in an appearance of the peak, which has shown the bottom panel of the Fig. 7. The bot-

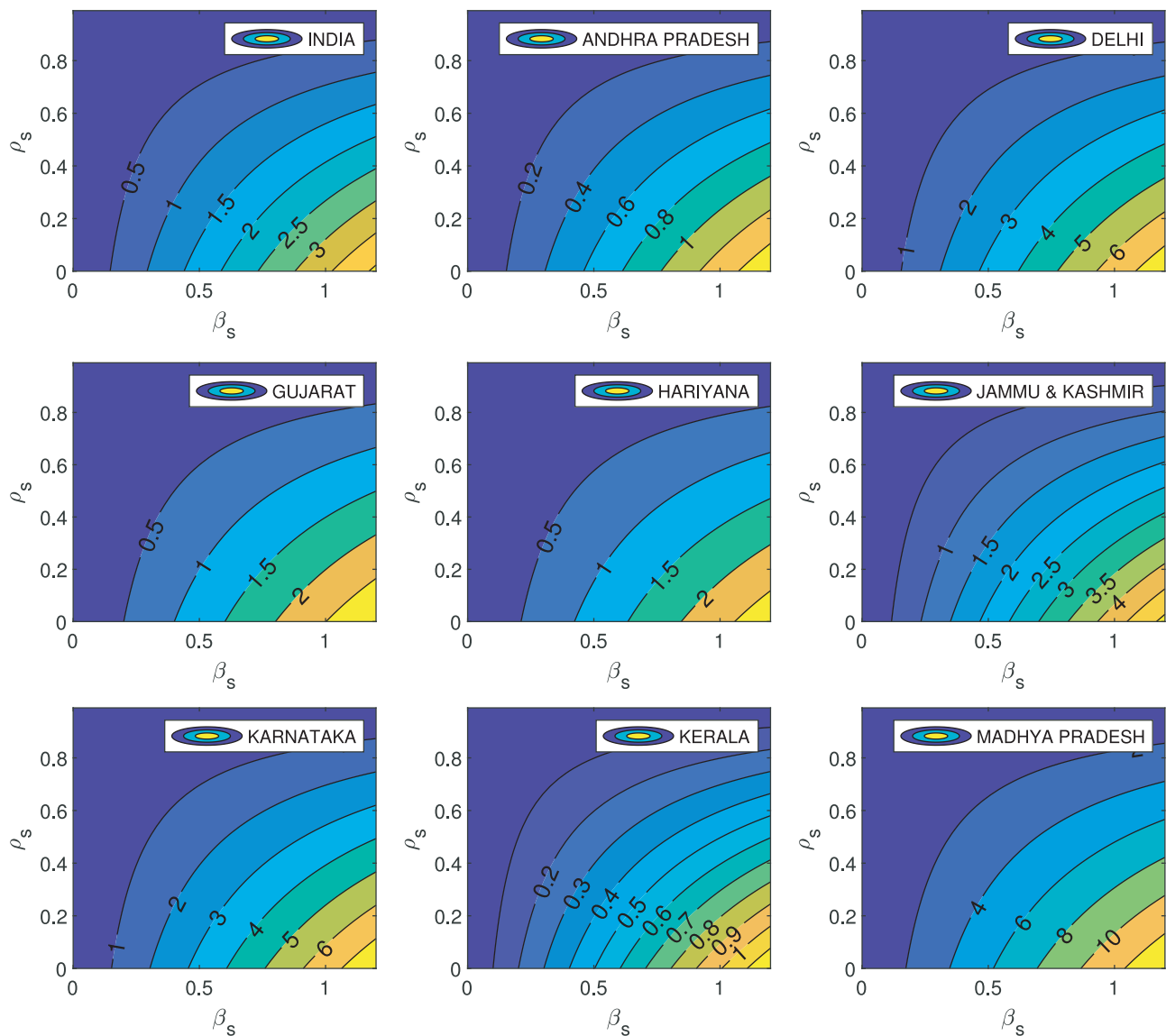
**Table 7**

Data duration and their sources. The first column list the name of India and its provinces, the second column list the source of data, the third column list the duration of data and the fourth column list the web address of the data sources.

India & Provinces	Source	Duration	Sample size	URL
India	Situation Report of World Health Organization (WHO)	30 January–30 April, 2020	87	<a href="http://www.who.int">www.who.int</a>
Delhi	Health & Family Welfare, Govt. of the National Capital Territory of Delhi	04 March–30 April, 2020	58	<a href="http://health.delhigovt.nic.in">health.delhigovt.nic.in</a>
Kerala	Directorate of Health Services, Govt. of Kerala	30 January–30 April, 2020	92	<a href="http://dhs.kerala.gov.in">dhs.kerala.gov.in</a>
West Bengal	Health & Family Welfare Dept., Govt. of West Bengal	17 March–30 April, 2020	42	<a href="http://www.wbhealth.gov.in">www.wbhealth.gov.in</a>
Bihar	COVID19 INDIA	22 March–30 April, 2020	40	<a href="http://www.covid19india.org">www.covid19india.org</a>
Gujarat	COVID19 INDIA	19 March–30 April, 2020	43	<a href="http://www.covid19india.org">www.covid19india.org</a>
Madhya Pradesh	COVID19 INDIA	20 March–30 April, 2020	42	<a href="http://www.covid19india.org">www.covid19india.org</a>
Maharashtra	COVID19 INDIA	15 March–30 April, 2020	47	<a href="http://www.covid19india.org">www.covid19india.org</a>
Andhra Pradesh	COVID19 INDIA	14 March–30 April, 2020	48	<a href="http://www.covid19india.org">www.covid19india.org</a>
Haryana	COVID19 INDIA	14 March–30 April, 2020	48	<a href="http://www.covid19india.org">www.covid19india.org</a>
Jammu & Kashmir	COVID19 INDIA	14 March–30 April, 2020	48	<a href="http://www.covid19india.org">www.covid19india.org</a>
Karnataka	COVID19 INDIA	14 March–30 April, 2020	48	<a href="http://www.covid19india.org">www.covid19india.org</a>
Odisha	COVID19 INDIA	14 March–30 April, 2020	48	<a href="http://www.covid19india.org">www.covid19india.org</a>
Punjab	COVID19 INDIA	14 March–30 April, 2020	48	<a href="http://www.covid19india.org">www.covid19india.org</a>
Rajasthan	COVID19 INDIA	14 March–30 April, 2020	48	<a href="http://www.covid19india.org">www.covid19india.org</a>
Tamil Nadu	COVID19 INDIA	14 March–30 April, 2020	48	<a href="http://www.covid19india.org">www.covid19india.org</a>
Telangana	COVID19 INDIA	14 March–30 April, 2020	48	<a href="http://www.covid19india.org">www.covid19india.org</a>
Uttar Pradesh	COVID19 INDIA	14 March–30 April, 2020	48	<a href="http://www.covid19india.org">www.covid19india.org</a>



**Fig. 7.** Effect of nation-wide lockdown. Epidemic evolution predicted by the  $SARIL_Sq$  model for SARS-CoV-2 epidemic in India for the time period 30 January 2020 to 16 August 2020. Before implemented 3rd Phase lockdown, the social distancing are not enforced, resulting a larger  $R_0 = 2.0490$ . After implementation of 3rd phase lockdown, the social distancing are enforced but not strong enough, as a result larger  $R_0 = 1.6302$  and  $R_0 = 1.0245$ , which shows a substantial outbreak of COVID-19. But after implementation of fully operational and strict lockdown, after 31 May 2020, resulting a smaller  $R_0 = 0.4098$ , yields below 1. Bottom panel shows the trajectories of COVID-19 diseases for different values of the strengthen of the intervention  $\omega$ .



**Fig. 8.** Contour plots of basic reproduction number  $R_0$ . Contour plots of  $R_0$  for India and eight different provinces of India, with respect to the probability of disease transmission rate  $\beta_s$  and the quarantined rate  $\rho_s$  of susceptible individuals. All parameter values other than  $\beta_s$  and  $\rho_s$  are listed in the Table 2. The contour plots demonstrates that higher disease transmission probability of the COVID-19 virus will remarkably increase the reproduction number  $R_0$ .

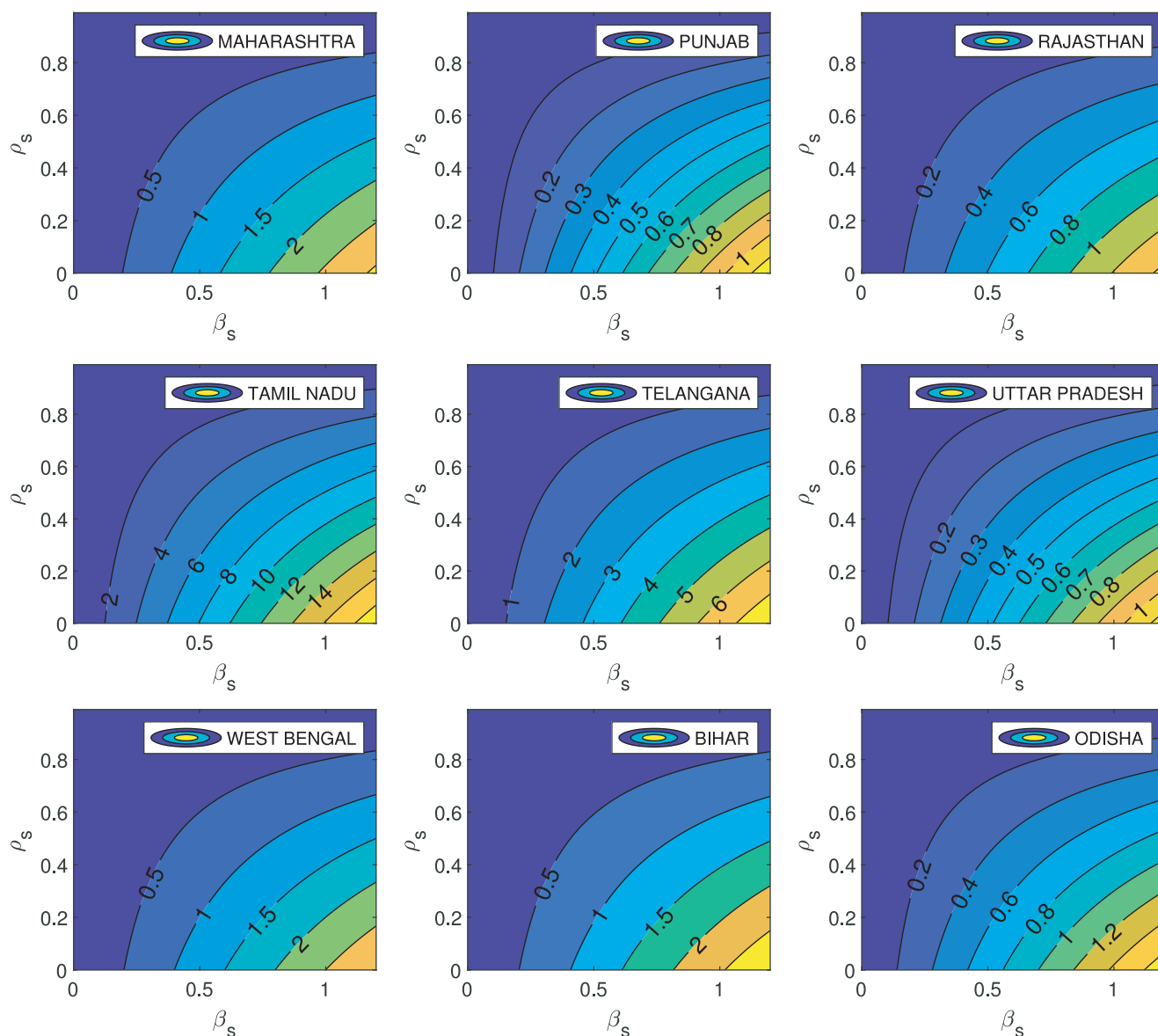
tom panel of Fig. 7 exhibits that the infected population reaches its peak at a different time.

#### 4.3. Model prediction

The first positive COVID-19 cases was observed in India on 30 January, 2020 and 02 March, 2020 onward and multiple new cases have been reported from all over the country. Forecast of the pandemic is most essential to take necessary administrative actions and health care planning and precautions. To estimate the end date of the pandemic, we have numerically simulated  $SARII_qS_q$  model and estimated the end days of COVID-19 infection. The numerical simulation of the model depends on the value of the parameters estimated and the parameters have been estimated from the observed data up to 30 April, 2020. So, if more observed data are available, the parameters values will alter and the long time prediction will be more accurate. However, this prediction gives us an overview of the pandemic, which will lead to decide future planning. In this study, we fitted  $SARII_qS_q$  model to forecast the pandemic trend over the period after 30 April, 2020 by using the ob-

served data from the first day of infection to 30 April, 2020 at national and state level.

Starting from the date of first COVID-19 case reported, we have simulated the  $SARII_qS_q$  model for 260 days for each states and for whole India to study the dynamics of the SARS-CoV-2 diseases. The simulated daily new COVID-19 cases are plotted for 175 days and for 200 days only for Kerala and for whole India in Figs. 10–12. The end date of the pandemic is the last date of COVID-19 case report. However, to take administrative decisions it is more essential to know the proportions of total infected populations rather estimating the end date of the pandemic. We have estimated the dates to reach 95% and 99% of the expected total infected COVID-19 cases in India and seventeen states of India. The dates are estimated by computing the total area under the simulated curve and finding the days upto which 95% or 99% area are enclosed. The model simulated forecasting with 95% and 99% end dates are shown in the Figs. 10–12. The blue curve represents the model forecast for the daily new confirmed cases of coronavirus diseases while the red dot histogram represents the observed cases. The dotted line rep-



**Fig. 9.** Contour plots of basic reproduction number  $R_0$ . Contour plots of  $R_0$  for nine different provinces of India, with respect to the probability of disease transmission rate  $\beta_s$  and the quarantined rate  $\rho_s$  of susceptible individuals. All parameter values other than  $\beta_s$  and  $\rho_s$  are listed in the Table 2. The contour plots demonstrates that higher disease transmission probability of the COVID-19 virus will remarkably increase the reproduction number  $R_0$ .

resents the 95% end date and dash-dot line represents the 99% end date.

The 95% end date of India is 26 June, 2020 and 99% end date of India is 27 July, 2020. The 95% and 99% end dates are different for different states of India, although the shapes of the curves are unimodal but have different skewness and kurtosis. We have observed that the 95% end date of states are on or before 26 June 2020 except for the states Andhra Pradesh, Kerala, Punjab, Tamil Nadu, Bihar and Odisha. Also, the 99% end of thirteen states are on or before 26 July, 2020 but the 99% end date of Andhra Pradesh, Punjab, Tamil Nadu, and Odisha are 30 July, 1 August, 3 August and 5 August, respectively. That is the 99% end days deviates 10 days from the 99% end day of India. This happens because within the 1% area (that is, from 99% to 100%) of the simulated curve for India contains some of the states which have 99% area after 26 July, 2020. The positively skewed and flatten (low kurtosis value) curves are more probable to have end dates after 26 July, 2020.

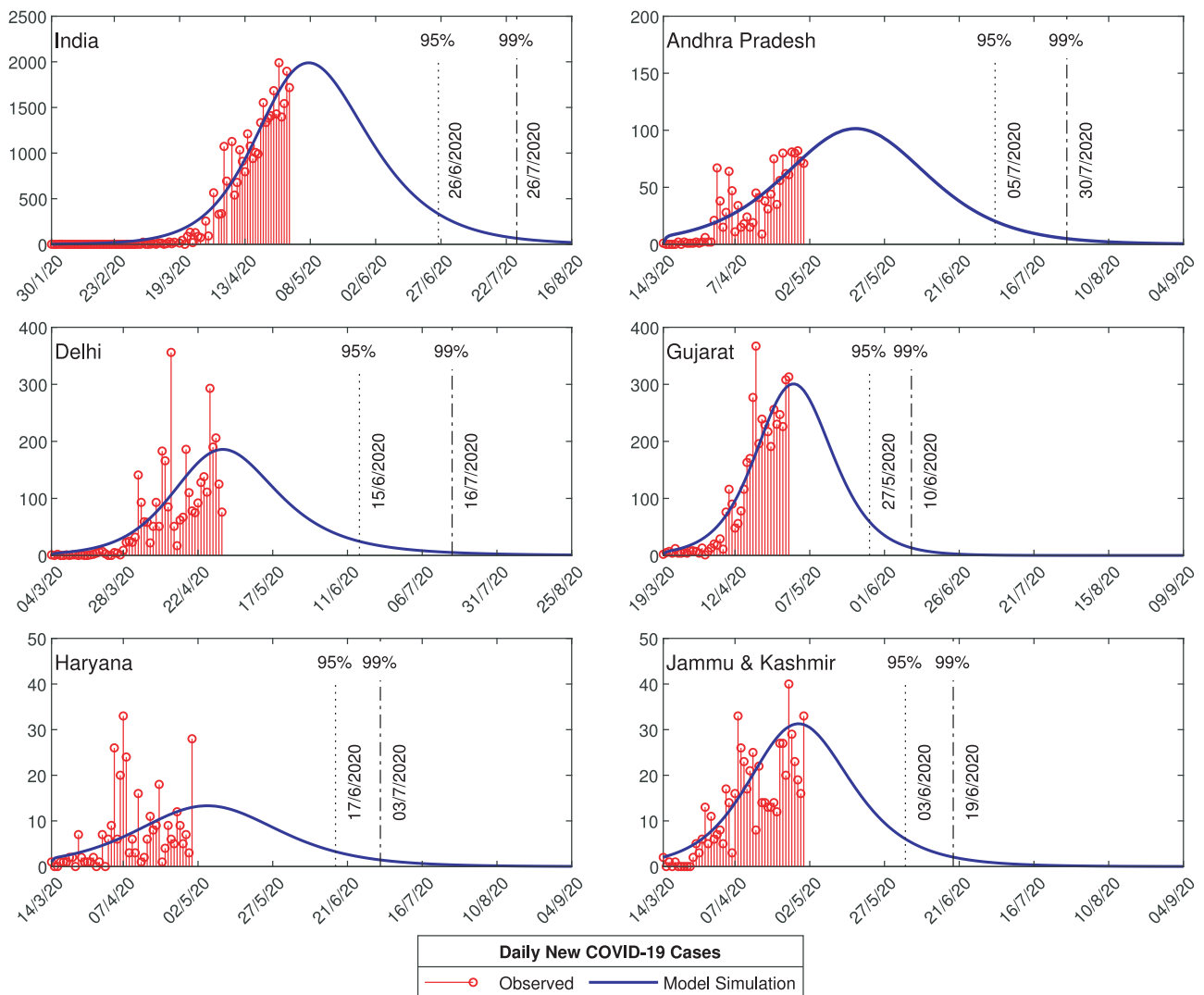
We have computed skewness, kurtosis and COVID-19 active days for each states and for whole India and listed in Table 6. All

the daily new COVID-19 distribution curves are positively skewed and platykurtic (kurtosis  $< 3$ ). The kurtosis of Andhra Pradesh, Punjab, Tamil Nadu and Odisha are  $-1.1613$ ,  $-1.3635$ ,  $-0.9711$  and  $-1.1314$ , respectively. Also, the active COVID-19 days in Andhra Pradesh, Punjab, Tamil Nadu and Odisha are 162, 155, 175 and 157 respectively. Hence, the curves, having maximum number of COVID-19 active days with kurtosis  $< -0.9$ , have the 99% end days after 26 July, 2020. Although, Kerala has kurtosis  $-1.3938$  and 181 active COVID-19 days, the 99% end date of Kerala is on 19 July, 2020. This might be due to the first 38 days, from 30 January, 2020 to 7 March, 2020, the number of COVID-19 case reported is very low.

## 5. Discussion

We have investigated a methodology in estimating the crucial epidemiological model parameters as well as the modeling and prediction of the spread of SARS-CoV-2 outbreak in 17 different



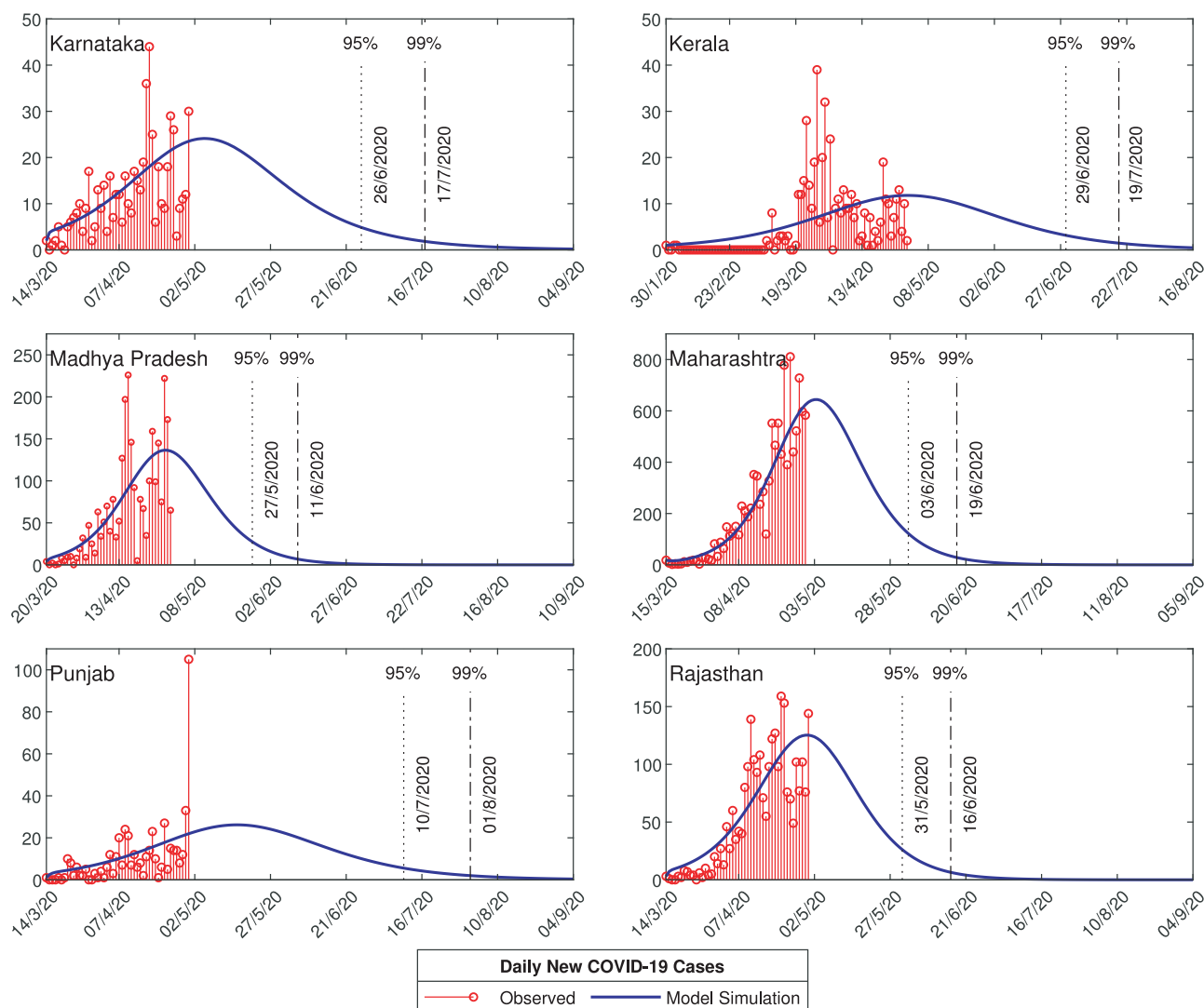


**Fig. 10.** Model-Based Data-Driven prediction of SARS-CoV-2 life cycle. Model predictions based on the estimated data for the Republic of India and five provinces of India, namely Andhra Pradesh, Delhi, Gujarat, Haryana and Jammu & Kashmir. The blue curve represents the model prediction for the daily new confirmed cases of coronavirus diseases while the red dot histogram represents the actual cases. The turning and ending dates of COVID-19 in India and five provinces are displayed in the figure. The estimated parameter values are specified in the Table 2. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

provinces of India and the overall India by evaluating publicly accessible real data till April 30, 2020.

Our model simulations provide an important tool to evaluate the consequences of possible policies, incorporating social distancing and lockdown. Our model findings, obtained by combining the  $SARII_qS_q$  model with the accessible data regarding the coronavirus outbreak in India, indicate that imposing social distancing is crucial, essential and effectual, in line with other details in the literature [6]. If the strong lockdown can be imposed earlier, the effective results can be achieved. We accept these indications can be handy to control the outbreak in India, together with the countries that are still in the early phases of epidemic.

The future is always unpredictable and we must keep this in mind when we read any kind of predictions. No one forecasted the outbreak of COVID-19 in India and worldwide, however Bill Gates warned about the potential harm of a worldwide infectious disease in a speech in 2015. Based on the real data on COVID-19 pandemic in India till April 30, 2020 we predicted our  $SARII_qS_q$  model. The total forecasted infected individual size is the total area under the curve. Our model forecasts give the following two estimates of end dates: (i) the date to reach 95% of expected total number of cases and (ii) the date to reach 99% of expected total number of cases. At the end, we want to conclude that the number of infected individuals may increase in India and the 17 provinces of

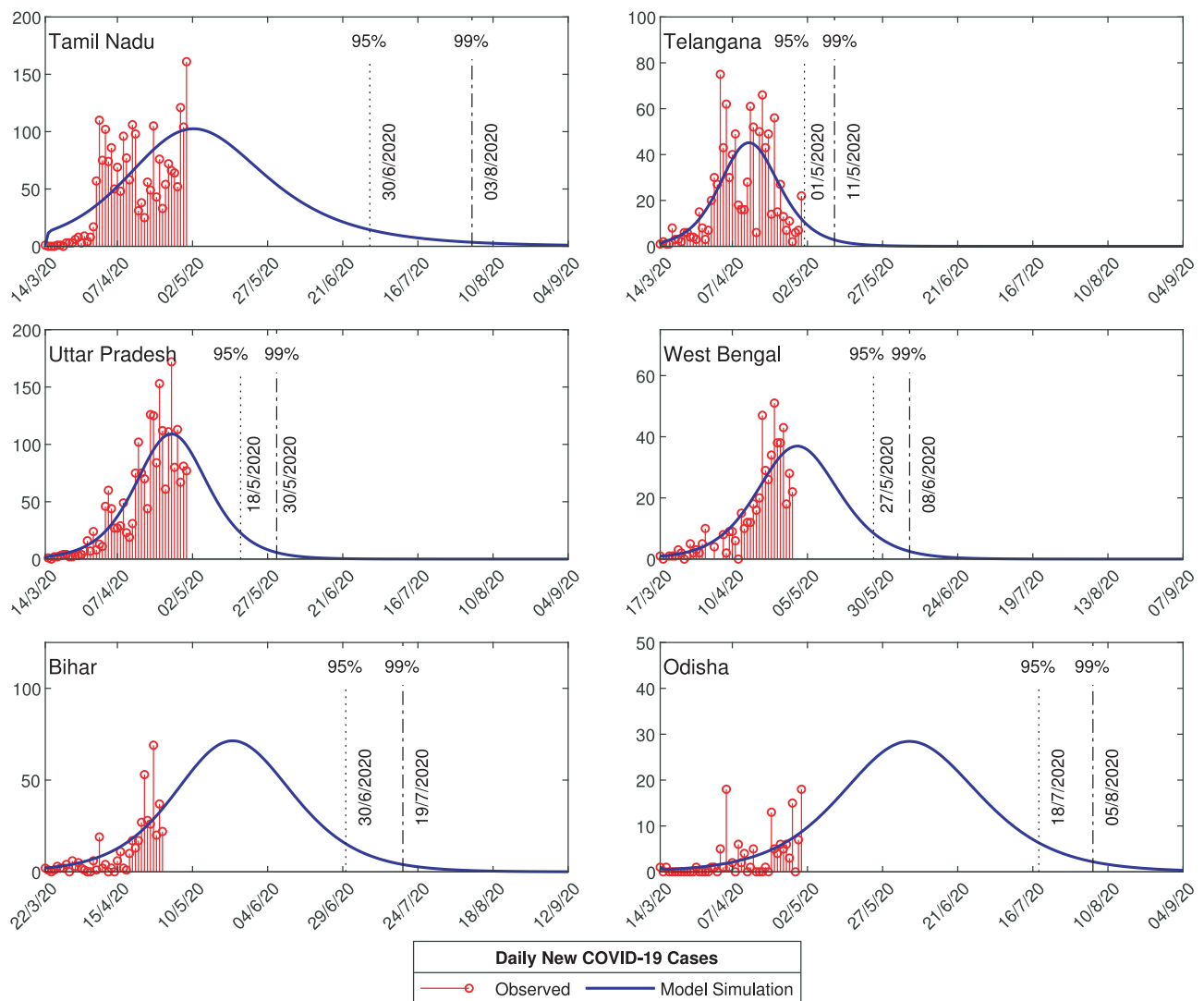


**Fig. 11.** Model-Based Data-Driven prediction of SARS-CoV-2 life cycle. Model predictions based on the estimated data for the five provinces of India, namely Karnataka, Kerala, Madhya Pradesh, Maharashtra, Punjab, and Rajasthan. The blue curve represents the model prediction for the daily new confirmed cases of coronavirus diseases while the red dot histogram represents the actual cases. The turning and ending dates of COVID-19 in India and five provinces are displayed in the figure. The estimated parameter values are specified in the Table 2. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

India for the next few months if human-to-human transmission and personal preventive measure resume with the existing rates as the testing status are very poor in India. It is worthy to mention that the researchers around the world is working for effective vaccine and/or therapeutics against coronavirus or SARS-CoV-2 diseases and the existence of such pharmaceutical interventions will remarkably change the results.

In the modern and developed world, data and detailed information regarding the COVID-19 or SARS-CoV-2 and the evolution of epidemic become accessible at an unparalleled pace. Howbeit, important questions still remain undetermined and precise answers

for forecasting the transmission dynamics of the epidemic simply cannot be acquired at this stage. We emphasize the uncertainty of accessible authentic data, specially concerning to the accurate baseline number of infected individuals, which may guide to the equivocal outcomes and inappropriate predictions by orders of size, as also identified by the other researches [45]. We hope that our predictions will be handy for Govt. and different companies as well as the people towards making resolutions and considering the suitable actions that contain the spreading of the coronavirus to the possible stage.



**Fig. 12.** Model-Based Data-Driven prediction of SARS-CoV-2 life cycle. Model predictions based on the estimated data for the five provinces of India, namely Tamil Nadu, Telangana, Uttar Pradesh, West Bengal, Bihar, and Odisha. The blue curve represents the model prediction for the daily new confirmed cases of coronavirus diseases while the red dot histogram represents the actual cases. The turning and ending dates of COVID-19 in India and five provinces are displayed in the figure. The estimated parameter values are specified in the Table 2. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

### Data availability

All the data used in this work has been obtained from official sources. All data supporting the findings of this study are in the paper and available from the corresponding author on request.

### Author contributions statement

Subhas Khajanchi, Kankan Sarkar, and Juan J. Nieto designed and performed the research as well as wrote the paper. The authors contributed equally to this work. All authors read and approved the final manuscript.

### Declaration of Competing Interest

The authors declare that they have no conflict of interest.

### Acknowledgments

The authors are thankful to Dr. Sandip Banerjee, Department of Mathematics, IIT Roorkee for valuable suggestions, which helped

in the better exposition of the manuscript. The work of Subhas Khajanchi is supported by Science and Engineering Research Board (SERB) (File No. ECR/ 2017/000234), Department of Science & Technology, Government of India. The work of J.J. Nieto has been partially supported by the Agencia Estatal de Investigacion (AEI) of Spain, co-financed by the European Fund for Regional Development (FEDER) corresponding to the 2014–2020 multi year financial framework, project MTM2016-75140-P; Instituto de Salud Carlos III, grant COV20/00617.

### References

- [1] World Health Organization (WHO), Coronavirus disease (COVID-19) outbreak situation. 2020. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>.
- [2] India COVID-19 tracker. 2020a. <https://www.covid19india.org/>.
- [3] WHO Situation Report, Coronavirus disease 2019 (COVID-19). 2020. <https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200306-sitrep-46-covid-19.pdf?sfvrsn=96b04adf-2>.
- [4] BBC News. <https://www.bbc.com/news/world-52114829>.
- [5] Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A novel coronavirus from patients with pneumonia in China, 2019. *N Engl J Med* 2020;382:727–33.
- [6] Wu Z, Googan JMM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of

- 72,314 cases from the Chinese center for disease control and prevention. *JAMA* 2020;323:1239–42.
- [7] Cohen J, Normile D. New SARS-like virus in China triggers alarm. *Science* 2020;367:234–5.
  - [8] Government of India. <https://www.mygov.in/covid-19>.
  - [9] BBC news. <https://www.bbc.com/news/world-asia-india-52077395>.
  - [10] The Indian Express. 2020b. <https://indianexpress.com/article/coronavirus/coronavirus-india-infection-rate-china-6321154/>.
  - [11] Anderson RM, Heesterbeek H, Klinkenberg D, Hollingsworth TD. How will country-based mitigation measures influence the course of the COVID-19 epidemic? *Lancet Comment* 2020;395(10228):931–4.
  - [12] Zhao S, Lin Q, Ran J, Musa SS, Yang G, Wang W, et al. Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: a data-driven analysis in the early phase of the outbreak. *Int J Inf Dis* 2020;92:214–17.
  - [13] Zou L, Ruan F, Huang M, et al. SARS-CoV-2 viral load in upper respiratory specimens of infected patients. *N Engl J Med* 2020;382:1177–9.
  - [14] NDTV News. <https://www.ndtv.com/india-news/coronavirus-impact-visas-to-india-suspended-till-april-15-2193382>.
  - [15] Ferguson NM, Laydon D, Nedjati-Gilani G, Imai N, Ainslie K, Baguelin M, Bhatia S, Boonyasiri A, et al. Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. In: Imperial college COVID-19 response team; 2020. p. 1–20.
  - [16] INDIA TV News, Suspension of international flight. 2020c. <https://www.indiatvnews.com/news/india/no-international-commercial-flight-will-be-allowed-to-land-india-from-march-22-599644>.
  - [17] Livemint, India suspends domestic flights. 2020a. <https://www.livemint.com/news/india/coronavirus-india-suspends-domestic-flights-from-25-march-11584962827845.html>.
  - [18] Regan H, Mitra E, Gupta S. Millions in India under coronavirus lockdown as major cities restrict daily life, CNN. 2020. <https://edition.cnn.com/2020/03/23/asia/coronavirus-covid-19-update-india-intl-hnk/index.html>.
  - [19] NDTV NEWS, India's Coronavirus Lockdown: What it Looks Like when India's 1.3 Billion People Stay Home. <https://www.ndtv.com/india-news/indias-coronavirus-lockdown-clear-roads-clean-air-80s-tv-2209373>.
  - [20] Livemint, Extension of lockdown till 3 may. Livemint. 2020b. <https://www.livemint.com/news/india/pm-modi-announces-extension-of-lockdown-till-3-may-11586839412073.html>.
  - [21] Anderson RM, May RM. Infectious diseases of humans. London: Oxford University Press; 1991.
  - [22] Diekmann O, Heesterbeek JAP. Mathematical epidemiology of infectious diseases: model building, analysis and interpretation. New York: Wiley; 2000.
  - [23] Hethcote HW. The mathematics of infectious diseases. *SIAM Rev* 2000;42:599–653.
  - [24] Banerjee S, Khajanchi S, Chaudhuri S. A mathematical model to elucidate brain tumor abrogation by immunotherapy with t11 target structure. *PLoS One* 2015;10(5):E0123611.
  - [25] Khajanchi Subhas, Sarkar Kankan. Forecasting the daily and cumulative number of cases for the COVID-19 pandemic in India. arXiv 2020 Submitted for publication. arxiv: 2006.14575.
  - [26] Kermack WO, Kendrick AGM. A contribution to the mathematical theory of epidemics. *Proc R Soc Lond* 1927;115:700–21.
  - [27] Chen T, Rui J, Wang Q, Zhao Z, Cui J, Yin L. A mathematical model for simulating the phase-based transmissibility of a novel coronavirus. *Infect Dis Rev* 2020;9(24).
  - [28] Lina Q, Zhao S, Gao D, Lou Y, Yang S, Musa SS, et al. A conceptual model for the coronavirus disease 2019 (COVID-19) outbreak in Wuhan, China with individual reaction and governmental action. *Int J Inf Dis* 2020;93:211–16.
  - [29] Khajanchi S, Sarkar K, Mondal J, Perc M. Dynamics of the COVID-19 pandemic in India2020; ArXiv preprint arXiv:2005.06286.
  - [30] Khan MA, Atangana A. Modeling the dynamics of novel coronavirus (2019-nCoV) with fractional derivative. *Alex Eng J* 2020. doi:10.1016/j.aej.2020.02.033.
  - [31] Wu JT, Leung K, Bushman M, Kishore N, Niehus R, de Salazar PM, et al. Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China. *Nat Med* 2020;26:506–10.
  - [32] Kucharski AJ, Russell TW, Diamond C, Liu Y, Edmunds J, Funk S, Eggo RM. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. *Lancet Infect Dis* 2020. doi:10.1016/S1473-3099(20)30144-4.
  - [33] Das DK, Khajanchi S, Kar TK. The impact of the media awareness and optimal strategy on the prevalence of tuberculosis. *Appl Math Comput* 2020;366:124732.
  - [34] Windarto Khan MA. Fatmawati, parameter estimation and fractional derivatives of dengue transmission model. *AIMS Math* 2020;5(3):2758–79.
  - [35] Khan MA, Shah SW, Ullah S, Gomez-Aguilar JF. A dynamical model of asymptomatic carrier zika virus with optimal control strategies. *Nonlinear Anal Real World Appl* 2019;50:144–70.
  - [36] Khajanchi S, Das DK, Kar TK. Dynamics of tuberculosis transmission with exogenous reinfections and endogenous reactivations. *Phys A* 2018;497:52–71.
  - [37] Tang B, Wang X, Li Q, Bragazzi NL, Tang S, Xiao Y, et al. Estimation of the transmission risk of the 2019-nCoV and its implication for public health interventions. *J Clin Med* 2020;9(2):462.
  - [38] Gumel AB, Ruan S, Day T, Watmough J, Brauer F, Van den Driessche P, et al. Modelling strategies for controlling SARS outbreaks. *Proc R Soc Lond B* 2004;271:2223–32.
  - [39] Ndairou F, Area I, Nieto JJ, Torres DFM. Mathematical modeling of COVID-19 transmission dynamics with a case study of Wuhan. *Chaos Soliton Fractals* 2020;135:109846.
  - [40] Fanelli D, Piazza F. Analysis and forecast of COVID-19 spreading in China, Italy and France. *Chaos Soliton Fractals* 2020;135:109761.
  - [41] Diekmann O, Heesterbeek JSP, Metz JA. On the definition and the computation of the basic reproduction ratio  $r_0$  in models for infectious diseases in heterogeneous populations. *J Math Biol* 1990;28:365–82.
  - [42] van den Driessche P, Watmough J. Reproduction numbers and subthreshold endemic equilibria for compartmental models of disease transmission. *Math Biosci* 2002;180:29–48.
  - [43] Marino S, Hogue IB, Ray CJ, Kirschner DE. A methodology for performing global uncertainty and sensitivity analysis in systems biology. *J Theor Biol* 2008;254(1):178–96.
  - [44] King AA, de Celles MN, Magpantay FMC, Rohani P. Avoidable errors in the modelling of outbreaks of emerging pathogens, with special reference to Ebola. *Proc Biol Sci* 2015;282:20150347.
  - [45] Battegay M, Kuehl R, Tschudin-Sutter S, Hirsch HH, Widmer AF, Neher RA. 2019-novel coronavirus (2019-nCoV): estimating the case fatality rate—a word of caution. *Swiss Medical Weekly*; 2020.
  - [46] Tang SY, Xiao YN, Yuan L, Cheke RA, Wu J. Campus quarantine (Fengxiao) for curbing emergent infectious diseases: lessons from mitigating A/H1N1 in Xi'an, China. *J Theor Biol* 2012;295:47–58.