COVID-19 Predictions for India
An Initial Attempt

Abhijit Das
Professor
Department of Computer Science and Engineering
Indian Institute of Technology Kharagpur
Kharagpur 721302, India
Email: abhij@cse.iitkgp.ac.in
Home URL: http://cse.iitkgp.ac.in/~abhij/
Date: 13-June-2020

(Technical Report)

***Abstract*—In this report, some attempts for making predictions about the COVID-19 pandemic in India are documented. The logistic model is used to fit the available daily counts of infection cases. The best estimates are obtained using the concept of minimizing the root-mean-square (RMS) errors. Moving averages are used to smooth out daily noises. The tool is applied on national data and also on state-level data.**

The model shows remarkably good fitting with the past data. However, the predictions for the future change quite rapidly with time. It appears that the spread of the infection in India has not yet reached a steady pattern. Consequently, predictive systems (like the one reported here) are unable to make reliable and definitive predictions about the future of the disease in India. Nevertheless, the trend revealed by the model may be a good indicator, and can be used for future planning.

***Index Terms*—SARS-COV-2, COVID-19, pandemic, prediction, logistic model, least-square curve fitting, nonlinear regression.**

I. Introduction

The global pandemic caused by the coronavirus (SARS-COV-2) is one of the most studied topics in 2020. Because of the pervasive influence of the disease (COVID-19, an acronym for COrona VIrus Disease that originated in 2019) in health-care, industry, economy, and even academics, predictive models—even if only approximate and/or inaccurate—are used in many parts of the globe. These models help predict the future spread of the disease, and can be used as suggestions to make educated plannings for immediate future. In this report, a prediction system based on the logistic model and developed in IIT Kharagpur is described.

Devising a robust predictive model for the spread of a disease has several challenges. First, this involves access to sensitive medical data which are not available in the public domain. Running a large-scale simulation of human movements is technically possible, but calls for private information like contact-tracing data for a large fraction of the population. These data are important because treating humans as random objects (like gas molecules) cannot be a valid assumption.

The only public knowledge that can be reasonably reliably used is the listing of daily and cumulative infection counts reported in several sites [1], [2], [3]. In particular, the COVID-19 India site [3] has detailed data both for the entire country and for all individual states that are affected by the disease. The daily data for India are available from 30-Jan-2020, whereas the state-level data are available from 14-Mar-2020. Since data of more than two months are at our disposal, we can proceed to use this database with reasonable confidence.

We assume that the daily counts $d_0, d_1, d_2, \ldots, d_n$ are available for n days. The corresponding cumulative counts are defined as $D_t = d_0 + d_1 + d_2 + \cdots + d_t$ for $t = 0, 1, 2, \ldots, n$. The idea is to find a curve that fits these cumulative counts in the best possible manner. A logistic function (also called a sigmoid function) [4], [5] is frequently used to model the spread of infectious diseases. This report describes the results from an implementation of the logistic model, carried out by the author. Other popular models for the purpose of prediction are the susceptible-infectious-removed (SIR) model and its several variants [6], [5]. The SIR model is not implemented so far. As reported in [5], the logistic model and the SIR model provide fairly close predictions. However, the predicted numbers of cases available at the site [7], which are based on the SEIR (susceptible-exposed-infected-removed) model, are much higher than what is revealed by the logistic model reported in this document.

This work is motivated by the predictive-monitoring site [8]. This site uses the implementation of the SIR model by Batista [9], but its outputs are internalized from 11-May-2020. Subsequently, the author of the current report finds it expedient to design a code himself for this purpose. As per the design choice of the author, the code should be a standalone utility without the use of any opaque toolbox. The outputs from this system are updated daily in the author's website [10].

The rest of this article describes efforts to make the model more and more robust. The improvement ideas are accompanied by the results obtained from the predictor. This article reports the predictions for the entire country, and for the two states Maharashtra (the most affected in the country) and West Bengal (the local state). The website [10] displays predictions for eight most affected states (Maharashtra, Tamil Nadu, Delhi, Gujarat, Uttar Pradesh, Rajasthan, Madhya Pradesh, and West Bengal). Since the source code can be downloaded from the site, curves for other states (or regions or countries) can be generated in a straightforward manner (only the input data files need to be prepared).

Despite several attempts to fine tune the predictor, stable predictions are not achieved. This is considered normal in [11], but it is prudent to conclude that India has not yet reached a stage where meaningful and reliable predictions are possible.

II. The Logistic Model

A. The Algorithm Used

The logistic model [4], [5] starts with the data of cumulative infection counts $C_0, C_1, C_2, \ldots, C_N$, where C_0 is the initial count (not necessarily corresponding to the start date of the infection or of the available data), and where $C_1, C_2, \ldots, C_N$ are the data available for the next N consecutive days. The cumulative count satisfies the differential equation

$$\frac{1}{C(t)} \frac{dC(t)}{dt} = r \left(1 - \frac{C(t)}{K}\right)$$

for some constant parameters K and r. The solution of the differential equation is

$$C(t) = \frac{K}{1 + A \exp(-rt)},$$

where

$$A = \frac{K - C(0)}{C(0)}.$$

The observed values C_t of $C(t)$ are available. From these, the best estimates of the two parameters are to be computed:

$$\begin{aligned} K &= \text{The total number of cases} = C(\infty), \\ r &= \text{The infection rate.} \end{aligned}$$

The method of nonlinear regression [12] is used for that. Let us write $C(t)$ as $C(t, K, r)$. Then, we have the following partial derivatives:

$$\begin{aligned} \frac{\partial C}{\partial K} &= \frac{1 - \exp(-2t)}{\left[1 + \left(\frac{K - C_0}{C_0}\right) \exp(-rt)\right]^2}, \\ \frac{\partial C}{\partial r} &= \frac{K \left(\frac{K - C_0}{C_0}\right) t \exp(-rt)}{\left[1 + \left(\frac{K - C_0}{C_0}\right) \exp(-rt)\right]^2}. \end{aligned}$$

One needs to start with some initial guesses of K and r. One possibility is to choose

$$\begin{aligned} K^{(0)} &= 2C_N, \\ r^{(0)} &= \frac{\ln\left(\frac{2C_N - C_0}{C_0}\right)}{N}. \end{aligned}$$

Next, the guesses are iteratively improved as follows. Assume that the i-th approximations $K^{(i)}, r^{(i)}$ are available for some $i \geqslant 0$. For $t = 0, 1, 2, \ldots, N$, compute

$$J_{t,K} = \frac{\partial C}{\partial K}(t, K^{(i)}, r^{(i)}),$$
$$J_{t,r} = \frac{\partial C}{\partial r}(t, K^{(i)}, r^{(i)}).$$

Consider the $(N+1)\times 2$ matrix J in which the first column consists of the $J_{t,K}$ values and the second column consists of the $J_{t,r}$ values. Also compute the current errors

$$\Delta C_t^{(i)} = C_t - C(t, K^{(i)}, r^{(i)})$$

for $t = 0, 1, 2, \ldots, N$. Let $\Delta C^{(i)}$ be the column vector consisting of these $N+1$ error values. The least-square regression method gives (see [12] for a proof)

$$J^{\,t} J \begin{pmatrix} \Delta K^{(i)} \\ \Delta r^{(i)} \end{pmatrix} = J^{\,t} \Delta C^{(i)}.$$

If the matrix $J^{\,t}J$ is invertible, the refinements $\Delta K^{(i)}, \Delta r^{(i)}$ are available, and the next guesses for K and r are computed:

$$K^{(i+1)} = K^{(i)} + \Delta K^{(i)},$$
$$r^{(i+1)} = r^{(i)} + \Delta r^{(i)}.$$

This refinement process may stop when $|\Delta K^{(i)}| < 0.5$ and $|\Delta r^{(i)}| < 0.0001$. The procedure may, however, fail because of numerical instability and/or the singularity of the matrix $J^{\,t}J$. If K or r ever becomes NaN (not a number), or r moves out of the interval $(0, 1)$, the loop is forcibly broken.

B. Experimental Observations

In order to understand the current trend in the progress of the disease, a good idea is to use the data for last N days. We have the daily counts $d_0, d_1, d_2, \ldots, d_n$ and the corresponding cumulative counts $D_t = \sum_{i=0}^{t} d_i$ for n days. In order to use these as the input data for the logistic model, we set

$$C_t = D_{n-N+t} = D_{n-N} + \sum_{i=1}^{t} d_i$$

for $t = 0, 1, 2, \ldots, N$.

Table I lists the predictor outputs based on the last N days of data, for several values of N. The tables in this report are prepared on 13-06-2020. On this day, data are available from the beginning to 12-06-2020. Only three regions are reported as samples: Entire India (IN), Maharashtra (MH), and West Bengal (WB). At this point, the actual cumulative counts for the three regions are respectively 623349, 140228, and 10244. The peak corresponds to the date on which the maximum number of new cases are added according to the estimated values of K and r. The expected end date is calculated as the time (day) t on which $C(t) = 0.999K$. The predictor fails for $N = 30$ for West Bengal.

It is evident that the predicted values vary widely depending upon the choice of N. This table does not prescribe any realistic optimal choice for N. On the contrary, it tends to indicate that the real-life patterns in India changed considerably over time, rendering these predictions quite unreliable.

TABLE I
VARIATION OF K AND r FOR DIFFERENT CHOICES OF N

	N	K	r	Peak on	End on
IN	15	725761	0.061	17-06-2020	07-10-2020
	30	780099	0.059	20-06-2020	14-09-2020
	45	641937	0.064	14-06-2020	29-09-2020
	60	537559	0.070	09-06-2020	15-09-2020
	75	362463	0.095	29-05-2020	09-08-2020
MH	15	205960	0.057	13-06-2020	13-10-2020
	30	142510	0.078	02-06-2020	29-08-2020
	45	142220	0.077	02-06-2020	30-08-2020
	60	127540	0.086	30-05-2020	18-08-2020
	75	105528	0.111	25-05-2020	26-07-2020
WB	15	20651	0.083	13-06-2020	03-09-2020
	45	27559	0.069	20-06-2020	27-09-2020
	60	16819	0.084	08-06-2020	29-08-2020
	75	13320	0.099	13-06-2020	11-08-2020

TABLE II
MINIMIZATION OF RMS ERROR BY VARYING N

	N	K	r	E_{rms}
IN	15	725761	0.061	1975.9
	30	780099	0.059	2438.7
	45	641937	0.064	1724.3
	60	537559	0.070	2104.9
	75	362463	0.095	7019.9
Best	52	623349	0.065	1703.3
MH	15	205960	0.057	3378.3
	30	142510	0.078	635.0
	45	142220	0.077	636.4
	60	127540	0.086	898.1
	75	105528	0.111	2595.2
Best	49	140228	0.078	622.4
WB	15	20651	0.083	296.1
	45	27559	0.069	156.0
	60	16819	0.084	239.1
	75	13320	0.099	330.6
Best	41	29561	0.068	153.5

TABLE III
CHANGE OF PREDICTION PARAMETERS WITH TIME

	Date	N	K	r	E_{rms}	End on
IN	03-06-2020	46	467353	0.069	1408	13-09-2020
	04-06-2020	47	482624	0.069	1456	14-09-2020
	05-06-2020	48	496703	0.069	1498	15-09-2020
	06-06-2020	36	534399	0.067	1546	20-09-2020
	07-06-2020	49	555637	0.067	1597	21-09-2020
	08-06-2020	50	564909	0.066	1612	22-09-2020
	09-06-2020	51	572899	0.066	1625	23-09-2020
	10-06-2020	51	587702	0.066	1646	24-09-2020
	11-06-2020	52	597997	0.066	1674	25-09-2020
	12-06-2020	52	623349	0.065	1703	27-09-2020
MH	03-06-2020	44	123813	0.083	469	20-08-2020
	04-06-2020	44	128012	0.082	468	22-08-2020
	05-06-2020	45	128113	0.082	465	22-08-2020
	06-06-2020	45	128113	0.082	465	22-08-2020
	07-06-2020	47	129579	0.081	471	23-08-2020
	08-06-2020	48	130493	0.081	477	23-08-2020
	09-06-2020	49	131091	0.081	479	23-08-2020
	10-06-2020	49	132546	0.081	498	24-08-2020
	11-06-2020	42	136580	0.080	550	26-08-2020
	12-06-2020	49	140228	0.078	622	28-08-2020
WB	03-06-2020	39	12882	0.077	135	02-09-2020
	04-06-2020	40	14682	0.075	141	07-09-2020
	05-06-2020	41	16826	0.073	146	12-09-2020
	06-06-2020	42	19247	0.072	151	16-09-2020
	07-06-2020	43	21872	0.071	155	20-09-2020
	08-06-2020	37	27874	0.069	156	29-09-2020
	09-06-2020	38	29406	0.068	156	30-09-2020
	10-06-2020	39	29382	0.068	155	30-09-2020
	11-06-2020	40	29365	0.068	154	30-09-2020
	12-06-2020	41	29561	0.068	153	01-10-2020

III. THE BEST ESTIMATE

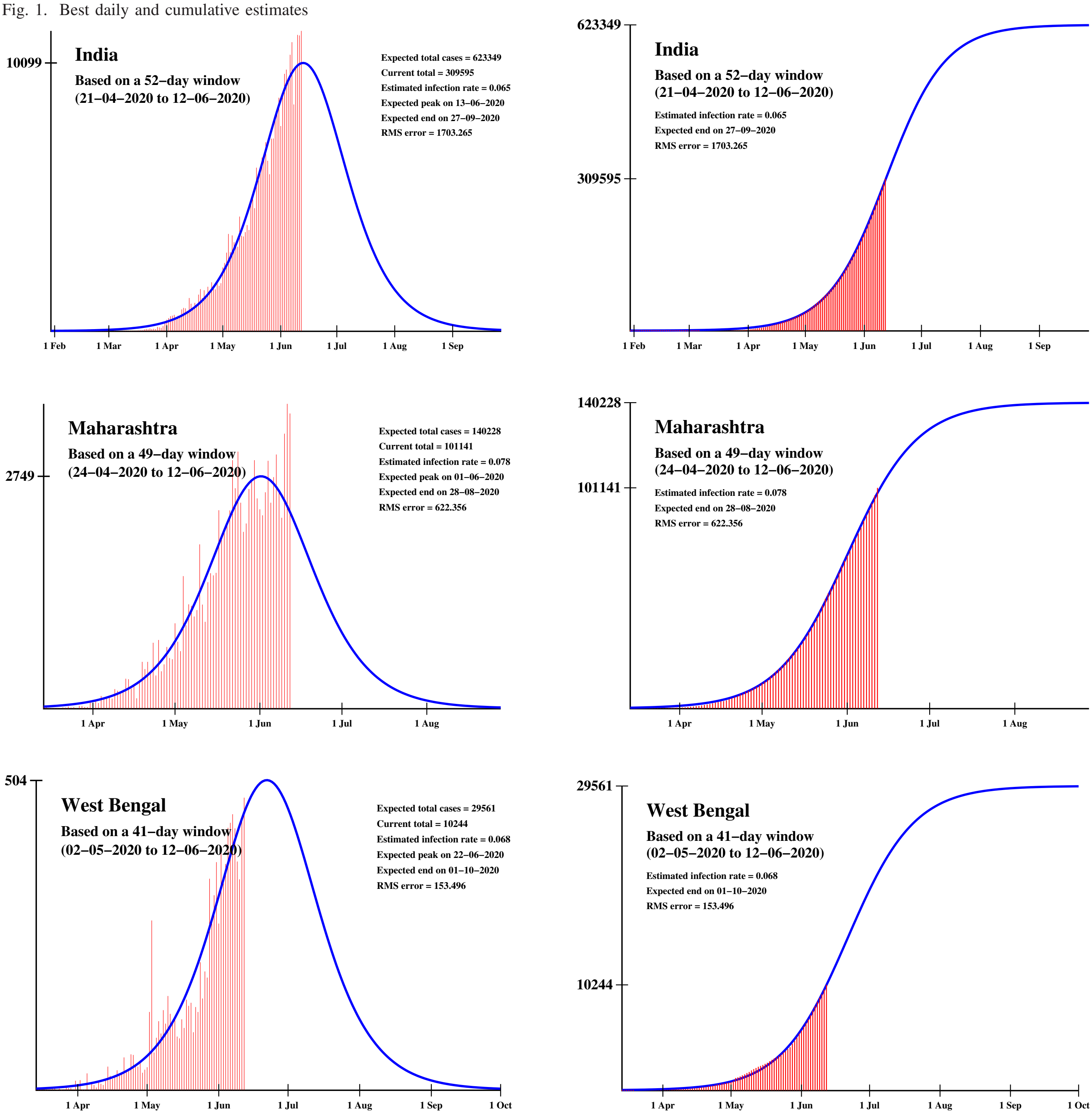
A. Minimizing RMS Errors

In order to avoid the difficulty arising out of the strong dependence of predictions on the number N of days considered for curve fitting, the concept of best estimators is introduced. Given N, one estimates K and r as explained above. The predicted cumulative sums are

$$P_t = C(t + N - n, K, r)$$

for $t = 0, 1, 2, \ldots, n$.

Fig. 1. Best daily and cumulative estimates



The root-mean-square (RMS) error is then calculated as

$$E_{rms} = \sqrt{\frac{1}{n+1}\sum_{t=0}^{n}(P_t - D_t)^2}.$$

Although the predictions are based on N days, the RMS error is calculated over the entire set of n days of data. N is varied in the range $[30, n]$. The value of N which gives the minimum RMS error is taken as the best choice for N. One should take $N \geqslant 30$. Looking at the data for less than one month is arguably not a good idea.

A better approach is to look at N-day windows not only at the end, but anywhere on the n-day data.

B. Experimental Observations

The effects of the minimization of E_{rms} are illustrated in Table II. These figures correspond to the data of infections up to 12-06-2020. For India, the best value of N happens to be 52, whereas for Maharashtra and West Bengal, this is 49 and 41, respectively. Although the N-day window can appear anywhere in the available data period, the data for the last N days usually lead to the best estimates. On a very few occasions, a window not flushed to the last date gives the best estimator. Large values of N (two months or more) lead to large RMS errors, again because over time, the parameters in the country have changed quite a lot.

Figure 1 provides a visual representation of the best estimates (blue curves) alongside the actual data (red lines). Although the daily data are too noisy to closely approximate the predicted curves, the fit of the cumulative predictions with the actual data is encouragingly good.

This apparent optimism quickly fades as one looks at the prospects of long-term and short-term predictions by the estimators. Table III shows the pattern in the change of the parameters K and r during a period of ten days. On all these days, the best estimator is used. Over this period, the estimated total count K increases quite significantly.

For the entire country, the increase is by a factor of 1.33, and for West Bengal, the factor is 2.29. Only Maharashtra shows some stability (increase by a factor of 1.13). The expected end date extends by 14 days for India, by 8 days for Maharashtra, and by about a month for West Bengal. The RMS error increases almost monotonically. Short-term predictions based on these rapidly changing parameters do not make sufficient sense. The paper [11] considers these changes as normal. But then, the basic motivation behind achieving a reliable and stable long-term prediction system continues to remain unfulfilled.

IV. Moving Averages

A. Removing Daily Noises

Moving averages are capable of smoothing out short-term fluctuations. A $(2w+1)$-day moving window is considered. The average daily count for each position of the window is assigned to the central day of the window. Using these averages in place of D_t is expected to reduce RMS errors. On the other hand, the predictions made are w days old (although partial contributions are taken from the last w days). One-day moving-average predictions are the same as daily-data predictions. The window size should not be too large. It appears reasonable to take $2w+1 \in \{3, 5, 7\}$.

B. Experimental Observations

Table IV shows the data from the moving-average predictors MA$(2w+1)$ alongside the data from the daily-data predictor DD. In each case, the best estimator which minimizes the RMS error is used. Data from the MA$(2w+1)$ predictors are shifted backward by w days for meaningful comparisons.

The table indicates that the RMS error decreases slightly if we consider moving averages. However, the predictions in terms of the total number K of cases and the expected end dates do not change much, and exhibit the same pattern of evolution with time as revealed by the daily-data predictor. In short, the use of moving averages fails to stabilize the predictions in any useful sense.

V. Initialization of the Logistic Model

A. The Initialization Parameter

The estimation loop for K and r starts with the initial value $K^{(0)} = 2C_n$ and the value $r^{(0)}$ derived from this initial estimate of K. In practice, one can take any real number $\xi > 1$, and start the loop with the initial values

$$\begin{aligned} K^{(0)} &= \xi C_N, \\ r^{(0)} &= \frac{\left(\dfrac{\xi C_N - C_0}{(\xi - 1)C_0}\right)}{N}. \end{aligned}$$

So far, the results reported pertain to $\xi = 2$. A natural question to ask in this context is whether the initialization could be a critical factor leading the finally estimated values of K and r to end up in different local minima. In order to address this concern, ξ is varied from 2 to 10 in steps of 1, and the best estimates are obtained for these different initializing values.

B. Experimental Observations

Table V lists the outcomes for the daily-data predictor for input data ending on 12-06-2020. The predictor fails on the data from Maharashtra for $\xi \geqslant 7$. The table reveals that there do exist local minima, but the RMS error is the smallest around $\xi = 2$, indicating that 2 is in general a good choice for ξ.

VI. Conclusion

In this report, the details of an implementation of the logistic model for COVID-19 predictions are discussed. Despite several attempts to improve the performance of the model, the results produced by the predictor are unreliable. It appears that this is because of considerable changes in the spread pattern of the disease in India over time. This may be attributed to various causes, like different mobility patterns of Indian people in different phases of lockdown, large-scale migration of laborers, change in diagnostic facilities, evolution of the coronavirus, and so on. These causes are well beyond the control of the logistic model (or any other prediction model for that matter).

Although the implementation fails to generate stable and reliable predictions, the trend clearly reveals that the disease is going to stay in the country for many more months. It is unlikely to get rid of it before the end of September 2020. This does not leave us in a region of comfort, but the reality has to be accepted, and appropriate plans must be chalked out to address all the issues associated with the outbreak of the pandemic.

References

[1] Worldometer, "COVID-19 coronavirus pandemic." https://www.worldometers.info/coronavirus/.

[2] Worldometer, "COVID-19 coronavirus pandemic: India." https://www.worldometers.info/coronavirus/country/india/.

[3] A. Shukla, "Covid19India: A crowdsourced initiative." https://www.covid19india.org/.

[4] Wikipedia, "Logistic function." https://en.wikipedia.org/wiki/Logistic_function.

[5] M. Batista, "Estimation of the final size of the COVID-19 epidemic," *Preprint from MedRxiv*, 2020.

[6] Wikipedia, "Compartmental models in epidemiology." https://en.wikipedia.org/wiki/Compartmental_models_in_epidemiology.

[7] COVID Analytics, "DELPHI epidemiological case predictions." https://www.covidanalytics.io/projections.

[8] SUTD Data-Driven Innovation Lab, "Predictive monitoring of COVID-19." https://ddi.sutd.edu.sg/.

[9] M. Batista, "Estimation of coronavirus COVID-19 epidemic evaluation by the SIR model." https://in.mathworks.com/matlabcentral/fileexchange/74658-fitviruscovid19.

[10] A. Das, "Covid-19 predictions: An initial attempt." http://cse.iitkgp.ac.in/~abhij/COVID19/.

[11] J. Luo, "Predictive monitoring of COVID-19," May 2020. https://people.sutd.edu.sg/jianxi_luo/public_html/COVID19PredictionPaper.pdf.

[12] Wikipedia, "Non-linear least squares." https://en.wikipedia.org/wiki/Non-linear_least_squares.

TABLE IV
COMPARISON OF DAILY-DATA AND MOVING-AVERAGE PREDICTORS

		K				E_{rms}				End on			
	Date	DD	MA(3)	MA(5)	MA(7)	DD	MA(3)	MA(5)	MA(7)	DD	MA(3)	MA(5)	MA(7)
IN	01-06-2020	423785	429730	423202	426888	1341	1328	1314	1298	08-09-2020	09-09-2020	08-09-2020	09-09-2020
	02-06-2020	435360	442007	448048	451234	1370	1359	1350	1337	09-09-2020	10-09-2020	11-09-2020	12-09-2020
	03-06-2020	467353	470501	461234	464887	1408	1398	1387	1376	13-09-2020	14-09-2020	12-09-2020	13-09-2020
	04-06-2020	482624	484887	487201	490438	1456	1440	1427	1418	14-09-2020	15-09-2020	15-09-2020	15-09-2020
	05-06-2020	496703	499714	502210	504115	1498	1486	1474	1456	15-09-2020	16-09-2020	16-09-2020	16-09-2020
	06-06-2020	534399	532309	529899	528259	1546	1533	1515	1493	20-09-2020	19-09-2020	19-09-2020	19-09-2020
	07-06-2020	555637	546007	542452	540719	1597	1570	1547	1526	21-09-2020	20-09-2020	20-09-2020	20-09-2020
	08-06-2020	564909	557343	567598	564036	1612	1597	1577	1557	22-09-2020	21-09-2020	22-09-2020	22-09-2020
	09-06-2020	572899	582271	578202	575265	1625	1614	1601	1586	23-09-2020	24-09-2020	23-09-2020	23-09-2020
	10-06-2020	587702	591442	587833		1646	1633	1623		24-09-2020	24-09-2020	24-09-2020	
	11-06-2020	597997	601522			1674	1659			25-09-2020	25-09-2020		
	12-06-2020	623349				1703				27-09-2020			
MH	01-06-2020	129549	127550	127914	126338	474	453	440	424	23-08-2020	22-08-2020	22-08-2020	22-08-2020
	02-06-2020	124038	126915	127586	126337	472	451	437	422	20-08-2020	22-08-2020	22-08-2020	22-08-2020
	03-06-2020	123813	126618	127334	126382	469	448	434	419	20-08-2020	21-08-2020	22-08-2020	22-08-2020
	04-06-2020	128012	126658	127315	126621	468	445	431	417	22-08-2020	21-08-2020	22-08-2020	22-08-2020
	05-06-2020	128113	127030	127685	127057	465	444	430	416	22-08-2020	22-08-2020	22-08-2020	22-08-2020
	06-06-2020	128113	127670	128327	127643	465	445	431	417	22-08-2020	22-08-2020	22-08-2020	22-08-2020
	07-06-2020	129579	128549	129014	128462	471	450	434	422	23-08-2020	22-08-2020	23-08-2020	22-08-2020
	08-06-2020	130493	129440	129925	129550	477	456	441	433	23-08-2020	22-08-2020	23-08-2020	23-08-2020
	09-06-2020	131091	130437	131162	133275	479	467	458	455	23-08-2020	23-08-2020	23-08-2020	25-08-2020
	10-06-2020	132546	134309	134261		498	488	485		24-08-2020	25-08-2020	25-08-2020	
	11-06-2020	136580	136220			550	533			26-08-2020	26-08-2020		
	12-06-2020	140228				622				28-08-2020			
WB	01-06-2020	9719	11057	11539	12052	117	116	111	107	22-08-2020	29-08-2020	01-09-2020	03-09-2020
	02-06-2020	11231	13276	13951	14587	128	124	120	115	28-08-2020	06-09-2020	09-09-2020	11-09-2020
	03-06-2020	12882	15991	16852	17774	135	130	126	122	02-09-2020	13-09-2020	16-09-2020	19-09-2020
	04-06-2020	14682	19148	20428	21596	141	135	131	127	07-09-2020	20-09-2020	23-09-2020	26-09-2020
	05-06-2020	16826	22886	24599	26101	146	139	135	131	12-09-2020	26-09-2020	29-09-2020	02-10-2020
	06-06-2020	19247	27321	29250	30705	151	142	138	133	16-09-2020	02-10-2020	05-10-2020	07-10-2020
	07-06-2020	21872	31978	33766	34651	155	144	139	134	20-09-2020	07-10-2020	09-10-2020	11-10-2020
	08-06-2020	27874	35659	36959	37539	156	144	140	135	29-09-2020	10-10-2020	12-10-2020	13-10-2020
	09-06-2020	29406	37054	38484	39196	156	144	139	134	30-09-2020	11-10-2020	13-10-2020	14-10-2020
	10-06-2020	29382	36725	38641		155	143	138		30-09-2020	11-10-2020	13-10-2020	
	11-06-2020	29365	35927			154	142			30-09-2020	10-10-2020		
	12-06-2020	29561				153				01-10-2020			

TABLE V
BEST PREDICTOR FOR DIFFERENT INITIALIZATION PARAMETERS ξ

	ξ	K	r	E_{rms}
IN	2	623349	0.065	1703
	3	623349	0.065	1703
	4	580838	0.067	1786
	5	656082	0.062	2107
	6	323226	0.120	51718
	7	323226	0.120	51718
	8	323226	0.120	51718
	9	323226	0.120	51718
	10	323226	0.120	51718
MH	2	140228	0.078	622
	3	140228	0.078	622
	4	143230	0.078	642
	5	143230	0.078	642
	6	187897	0.074	3182
WB	2	29561	0.068	153
	3	29561	0.068	153
	4	29561	0.068	153
	5	29561	0.068	153
	6	49427	0.057	268
	7	49427	0.057	268
	8	49427	0.057	268
	9	49427	0.057	268
	10	49427	0.057	268

APPENDIX

In what follows, the prediction trends are reported for India and for the eight most affected states. The old data for the country posted in the site [3] change slightly. Sometimes, daily counts as old as belonging to the last tenth day increase or decrease. The tables in the appendix are based on the data as on 13-June-2020. If the input data (daily counts) change only slightly (as it has been the case so far), the predictions will not be significantly affected.

Each row of the table starts with a prediction date. Data up to that date are used to obtain the results from the best daily-data estimator. The quantities reported against each prediction date are the total number of reported cases up to that day, the best window of N days that minimizes the RMS error, the logistic-model parameters K (the total number of cases) and r (the infection rate), the RMS error, and the predicted end date (when 99.9% of the total cases are expected to be detected). The tables illustrate the evolution of the predictor outputs over a period of 21 consecutive days.

The website [10] maintained by the author updates the predictor outputs daily. On each day, the best prediction results based on the data up to the previous day are posted. Moreover, the prediction trends are updated based on the latest data.

TABLE VI
PREDICTION TREND FOR INDIA (IN)

Date	Total cases	Best window	K	r	E_{rms}	End on
23-05-2020	131425	40 (13-04-2020 – 23-05-2020)	288566	0.077	1069	24-08-2020
24-05-2020	138536	41 (13-04-2020 – 24-05-2020)	305572	0.076	1123	26-08-2020
25-05-2020	144950	42 (13-04-2020 – 25-05-2020)	319619	0.075	1159	27-08-2020
26-05-2020	150857	43 (13-04-2020 – 26-05-2020)	328517	0.075	1174	28-08-2020
27-05-2020	158103	44 (13-04-2020 – 27-05-2020)	338122	0.075	1192	29-08-2020
28-05-2020	165357	43 (14-04-2020 – 27-05-2020)	352244	0.074	1208	01-09-2020
29-05-2020	173495	45 (14-04-2020 – 29-05-2020)	372555	0.073	1236	03-09-2020
30-05-2020	181859	45 (15-04-2020 – 30-05-2020)	388525	0.072	1272	05-09-2020
31-05-2020	190648	45 (16-04-2020 – 31-05-2020)	412601	0.071	1315	07-09-2020
01-06-2020	198371	46 (16-04-2020 – 01-06-2020)	423785	0.071	1341	08-09-2020
02-06-2020	207183	47 (16-04-2020 – 02-06-2020)	435360	0.071	1370	09-09-2020
03-06-2020	216872	46 (18-04-2020 – 03-06-2020)	467353	0.069	1408	13-09-2020
04-06-2020	226719	47 (18-04-2020 – 04-06-2020)	482624	0.069	1456	14-09-2020
05-06-2020	236191	48 (18-04-2020 – 05-06-2020)	496703	0.069	1498	15-09-2020
06-06-2020	246599	36 (30-04-2020 – 05-06-2020)	534399	0.067	1546	20-09-2020
07-06-2020	257481	49 (19-04-2020 – 07-06-2020)	555637	0.067	1597	21-09-2020
08-06-2020	266017	50 (19-04-2020 – 08-06-2020)	564909	0.066	1612	22-09-2020
09-06-2020	275998	51 (19-04-2020 – 09-06-2020)	572899	0.066	1625	23-09-2020
10-06-2020	287154	51 (20-04-2020 – 10-06-2020)	587702	0.066	1646	24-09-2020
11-06-2020	298289	52 (20-04-2020 – 11-06-2020)	597997	0.066	1674	25-09-2020
12-06-2020	309595	52 (21-04-2020 – 12-06-2020)	623349	0.065	1703	27-09-2020

TABLE VII
PREDICTION TREND FOR MAHARASHTRA (MH)

Date	Total cases	Best window	K	r	E_{rms}	End on
23-05-2020	47190	33 (20-04-2020 – 23-05-2020)	106558	0.085	443	15-08-2020
24-05-2020	50231	34 (20-04-2020 – 24-05-2020)	115748	0.084	478	18-08-2020
25-05-2020	52667	35 (20-04-2020 – 25-05-2020)	121836	0.084	491	19-08-2020
26-05-2020	54758	36 (20-04-2020 – 26-05-2020)	123481	0.083	490	20-08-2020
27-05-2020	56948	37 (20-04-2020 – 27-05-2020)	122926	0.083	487	19-08-2020
28-05-2020	59546	38 (20-04-2020 – 28-05-2020)	122752	0.083	483	19-08-2020
29-05-2020	62228	39 (20-04-2020 – 29-05-2020)	123016	0.083	480	20-08-2020
30-05-2020	65168	38 (21-04-2020 – 29-05-2020)	129701	0.081	480	23-08-2020
31-05-2020	67655	40 (21-04-2020 – 31-05-2020)	130257	0.081	477	23-08-2020
01-06-2020	70013	41 (21-04-2020 – 01-06-2020)	129549	0.081	474	23-08-2020
02-06-2020	72300	43 (20-04-2020 – 02-06-2020)	124038	0.083	472	20-08-2020
03-06-2020	74860	44 (20-04-2020 – 03-06-2020)	123813	0.083	469	20-08-2020
04-06-2020	77793	44 (21-04-2020 – 04-06-2020)	128012	0.082	468	22-08-2020
05-06-2020	80229	45 (21-04-2020 – 05-06-2020)	128113	0.082	465	22-08-2020
06-06-2020	82968	45 (21-04-2020 – 05-06-2020)	128113	0.082	465	22-08-2020
07-06-2020	85975	47 (21-04-2020 – 07-06-2020)	129579	0.081	471	23-08-2020
08-06-2020	88529	48 (21-04-2020 – 08-06-2020)	130493	0.081	477	23-08-2020
09-06-2020	90787	49 (21-04-2020 – 09-06-2020)	131091	0.081	479	23-08-2020
10-06-2020	94041	49 (22-04-2020 – 10-06-2020)	132546	0.081	498	24-08-2020
11-06-2020	97648	42 (30-04-2020 – 11-06-2020)	136580	0.080	550	26-08-2020
12-06-2020	101141	49 (24-04-2020 – 12-06-2020)	140228	0.078	622	28-08-2020

TABLE VIII
PREDICTION TREND FOR TAMIL NADU (TN)

Date	Total cases	Best window	K	r	E_{rms}	End on
23-05-2020	15512	53 (31-03-2020 – 23-05-2020)	25450	0.108	362	22-07-2020
24-05-2020	16277	30 (24-04-2020 – 24-05-2020)	29135	0.099	356	30-07-2020
25-05-2020	17082	31 (24-04-2020 – 25-05-2020)	29452	0.099	353	31-07-2020
26-05-2020	17728	32 (24-04-2020 – 26-05-2020)	29511	0.099	351	31-07-2020
27-05-2020	18545	33 (24-04-2020 – 27-05-2020)	29773	0.099	348	31-07-2020
28-05-2020	19372	34 (24-04-2020 – 28-05-2020)	30166	0.098	346	01-08-2020
29-05-2020	20246	35 (24-04-2020 – 29-05-2020)	30705	0.098	345	01-08-2020
30-05-2020	21184	36 (24-04-2020 – 30-05-2020)	31415	0.097	346	02-08-2020
31-05-2020	22333	37 (24-04-2020 – 31-05-2020)	32501	0.096	352	03-08-2020
01-06-2020	23495	39 (23-04-2020 – 01-06-2020)	35244	0.092	363	08-08-2020
02-06-2020	24586	62 (01-04-2020 – 02-06-2020)	36722	0.091	373	10-08-2020
03-06-2020	25872	41 (23-04-2020 – 03-06-2020)	38332	0.090	390	12-08-2020
04-06-2020	27256	43 (22-04-2020 – 04-06-2020)	42158	0.086	409	17-08-2020
05-06-2020	28694	64 (02-04-2020 – 05-06-2020)	44738	0.085	432	20-08-2020
06-06-2020	30152	46 (21-04-2020 – 06-06-2020)	49734	0.081	455	26-08-2020
07-06-2020	31667	48 (20-04-2020 – 07-06-2020)	54107	0.079	477	30-08-2020
08-06-2020	33229	50 (19-04-2020 – 08-06-2020)	59988	0.076	497	05-09-2020
09-06-2020	34914	66 (04-04-2020 – 09-06-2020)	63628	0.076	517	07-09-2020
10-06-2020	36841	54 (17-04-2020 – 10-06-2020)	70077	0.074	541	11-09-2020
11-06-2020	38716	56 (16-04-2020 – 11-06-2020)	76820	0.072	563	16-09-2020
12-06-2020	40698	58 (15-04-2020 – 12-06-2020)	86567	0.070	584	21-09-2020

TABLE IX
PREDICTION TREND FOR DELHI (DL)

Date	Total cases	Best window	K	r	E_{rms}	End on
23-05-2020	12910	43 (10-04-2020 – 23-05-2020)	24567	0.077	176	20-08-2020
24-05-2020	13418	44 (10-04-2020 – 24-05-2020)	25323	0.077	178	21-08-2020
25-05-2020	14053	45 (10-04-2020 – 25-05-2020)	26187	0.076	181	22-08-2020
26-05-2020	14465	46 (10-04-2020 – 26-05-2020)	26581	0.076	181	23-08-2020
27-05-2020	15257	47 (10-04-2020 – 27-05-2020)	27439	0.076	185	24-08-2020
28-05-2020	16281	34 (24-04-2020 – 28-05-2020)	32840	0.071	200	03-09-2020
29-05-2020	17386	48 (11-04-2020 – 29-05-2020)	36356	0.069	224	08-09-2020
30-05-2020	18549	48 (12-04-2020 – 30-05-2020)	41622	0.068	255	13-09-2020
31-05-2020	19844	49 (12-04-2020 – 31-05-2020)	48301	0.067	291	19-09-2020
01-06-2020	20834	40 (22-04-2020 – 01-06-2020)	67268	0.063	312	03-10-2020
02-06-2020	22132	46 (17-04-2020 – 02-06-2020)	90116	0.060	330	14-10-2020
03-06-2020	23645	47 (17-04-2020 – 03-06-2020)	117916	0.059	351	21-10-2020
04-06-2020	25004	41 (23-04-2020 – 03-06-2020)	109299	0.060	375	18-10-2020
05-06-2020	26334	54 (12-04-2020 – 05-06-2020)	108562	0.062	396	13-10-2020
06-06-2020	27654	55 (12-04-2020 – 06-06-2020)	122919	0.062	402	16-10-2020
07-06-2020	28936	56 (12-04-2020 – 07-06-2020)	132702	0.061	403	18-10-2020
08-06-2020	29943	57 (12-04-2020 – 08-06-2020)	132263	0.061	401	18-10-2020
09-06-2020	31309	41 (24-04-2020 – 04-06-2020)	147178	0.059	392	25-10-2020
10-06-2020	32810	42 (23-04-2020 – 04-06-2020)	146448	0.059	386	25-10-2020
11-06-2020	34687	41 (25-04-2020 – 05-06-2020)	167362	0.059	374	28-10-2020
12-06-2020	36824	42 (30-04-2020 – 11-06-2020)	179259	0.059	375	30-10-2020

TABLE X
PREDICTION TREND FOR GUJARAT (GJ)

Date	Total cases	Best window	K	r	E_{rms}	End on
23-05-2020	13669	34 (19-04-2020 – 23-05-2020)	18260	0.097	228	22-07-2020
24-05-2020	14063	34 (19-04-2020 – 23-05-2020)	18260	0.097	237	22-07-2020
25-05-2020	14468	35 (20-04-2020 – 25-05-2020)	18853	0.095	237	25-07-2020
26-05-2020	14829	35 (21-04-2020 – 26-05-2020)	19455	0.092	257	28-07-2020
27-05-2020	15205	39 (18-04-2020 – 27-05-2020)	17637	0.106	232	15-07-2020
28-05-2020	15572	30 (20-04-2020 – 20-05-2020)	18207	0.096	254	23-07-2020
29-05-2020	15944	31 (20-04-2020 – 21-05-2020)	18371	0.096	257	23-07-2020
30-05-2020	16356	32 (21-04-2020 – 23-05-2020)	19176	0.092	264	27-07-2020
31-05-2020	16794	43 (18-04-2020 – 31-05-2020)	18591	0.103	259	18-07-2020
01-06-2020	17217	34 (21-04-2020 – 25-05-2020)	19357	0.092	282	28-07-2020
02-06-2020	17632	43 (20-04-2020 – 02-06-2020)	20349	0.091	259	29-07-2020
03-06-2020	18117	45 (19-04-2020 – 03-06-2020)	20443	0.092	267	29-07-2020
04-06-2020	18609	46 (19-04-2020 – 04-06-2020)	20778	0.091	282	30-07-2020
05-06-2020	19119	45 (21-04-2020 – 05-06-2020)	21707	0.087	307	04-08-2020
06-06-2020	19617	47 (20-04-2020 – 06-06-2020)	21756	0.088	320	03-08-2020
07-06-2020	20097	48 (20-04-2020 – 07-06-2020)	22164	0.087	341	04-08-2020
08-06-2020	20574	50 (19-04-2020 – 08-06-2020)	22347	0.088	366	04-08-2020
09-06-2020	21044	51 (19-04-2020 – 09-06-2020)	22762	0.087	388	05-08-2020
10-06-2020	21554	50 (21-04-2020 – 10-06-2020)	23815	0.082	399	11-08-2020
11-06-2020	22067	51 (21-04-2020 – 11-06-2020)	24274	0.081	421	12-08-2020
12-06-2020	22562	52 (21-04-2020 – 12-06-2020)	24738	0.080	442	14-08-2020

TABLE XI
PREDICTION TREND FOR UTTAR PRADESH (UP)

Date	Total cases	Best window	K	r	E_{rms}	End on
23-05-2020	6017	35 (18-04-2020 – 23-05-2020)	9044	0.075	155	16-08-2020
24-05-2020	6268	36 (18-04-2020 – 24-05-2020)	9988	0.073	164	22-08-2020
25-05-2020	6497	37 (18-04-2020 – 25-05-2020)	10857	0.071	169	27-08-2020
26-05-2020	6724	38 (18-04-2020 – 26-05-2020)	11606	0.069	172	30-08-2020
27-05-2020	6991	39 (18-04-2020 – 27-05-2020)	12324	0.068	175	02-09-2020
28-05-2020	7170	40 (18-04-2020 – 28-05-2020)	12762	0.068	176	04-09-2020
29-05-2020	7445	41 (18-04-2020 – 29-05-2020)	13211	0.067	176	06-09-2020
30-05-2020	7701	42 (18-04-2020 – 30-05-2020)	13609	0.066	177	08-09-2020
31-05-2020	8075	43 (18-04-2020 – 31-05-2020)	14226	0.066	179	10-09-2020
01-06-2020	8361	44 (18-04-2020 – 01-06-2020)	14796	0.065	180	12-09-2020
02-06-2020	8729	45 (18-04-2020 – 02-06-2020)	15495	0.064	183	14-09-2020
03-06-2020	8870	46 (18-04-2020 – 03-06-2020)	15778	0.064	183	15-09-2020
04-06-2020	9237	47 (18-04-2020 – 04-06-2020)	16210	0.064	184	17-09-2020
05-06-2020	9733	47 (19-04-2020 – 05-06-2020)	19038	0.059	187	30-09-2020
06-06-2020	10103	48 (19-04-2020 – 06-06-2020)	19947	0.059	190	02-10-2020
07-06-2020	10536	49 (19-04-2020 – 07-06-2020)	21023	0.058	194	05-10-2020
08-06-2020	10947	50 (19-04-2020 – 08-06-2020)	22144	0.057	197	08-10-2020
09-06-2020	11335	51 (19-04-2020 – 09-06-2020)	23196	0.057	200	10-10-2020
10-06-2020	11610	52 (19-04-2020 – 10-06-2020)	23820	0.057	201	11-10-2020
11-06-2020	12088	53 (19-04-2020 – 11-06-2020)	24655	0.056	203	13-10-2020
12-06-2020	12616	53 (20-04-2020 – 12-06-2020)	26944	0.055	206	19-10-2020

TABLE XII
PREDICTION TREND FOR RAJASTHAN (RJ)

Date	Total cases	Best window	K	r	E_{rms}	End on
23-05-2020	6742	41 (12-04-2020 – 23-05-2020)	11237	0.070	192	25-08-2020
24-05-2020	7028	42 (12-04-2020 – 24-05-2020)	11904	0.069	195	28-08-2020
25-05-2020	7300	43 (12-04-2020 – 25-05-2020)	12543	0.068	197	30-08-2020
26-05-2020	7536	44 (12-04-2020 – 26-05-2020)	13060	0.068	198	01-09-2020
27-05-2020	7816	45 (12-04-2020 – 27-05-2020)	13558	0.067	199	03-09-2020
28-05-2020	8067	46 (12-04-2020 – 28-05-2020)	13968	0.067	200	05-09-2020
29-05-2020	8365	47 (12-04-2020 – 29-05-2020)	14391	0.066	200	06-09-2020
30-05-2020	8617	47 (13-04-2020 – 30-05-2020)	15875	0.063	200	14-09-2020
31-05-2020	8831	48 (13-04-2020 – 31-05-2020)	16014	0.063	199	15-09-2020
01-06-2020	9100	49 (13-04-2020 – 01-06-2020)	16148	0.063	198	15-09-2020
02-06-2020	9373	50 (13-04-2020 – 02-06-2020)	16284	0.063	197	16-09-2020
03-06-2020	9652	51 (13-04-2020 – 03-06-2020)	16428	0.063	197	16-09-2020
04-06-2020	9862	52 (13-04-2020 – 04-06-2020)	16473	0.062	195	16-09-2020
05-06-2020	10084	53 (13-04-2020 – 05-06-2020)	16474	0.062	194	16-09-2020
06-06-2020	10337	54 (13-04-2020 – 06-06-2020)	16487	0.062	193	16-09-2020
07-06-2020	10599	55 (13-04-2020 – 07-06-2020)	16521	0.062	192	16-09-2020
08-06-2020	10876	56 (13-04-2020 – 08-06-2020)	16588	0.062	191	17-09-2020
09-06-2020	11245	57 (13-04-2020 – 09-06-2020)	16764	0.062	191	17-09-2020
10-06-2020	11600	58 (13-04-2020 – 10-06-2020)	17007	0.062	192	18-09-2020
11-06-2020	11838	59 (13-04-2020 – 11-06-2020)	17200	0.062	193	19-09-2020
12-06-2020	12068	60 (13-04-2020 – 12-06-2020)	17347	0.062	193	19-09-2020

TABLE XIII
PREDICTION TREND FOR MADHYA PRADESH (MP)

Date	Total cases	Best window	K	r	E_{rms}	End on
23-05-2020	6371	38 (15-04-2020 – 23-05-2020)	10449	0.071	201	23-08-2020
24-05-2020	6665	39 (15-04-2020 – 24-05-2020)	11070	0.070	203	26-08-2020
25-05-2020	6859	40 (15-04-2020 – 25-05-2020)	11533	0.069	205	29-08-2020
26-05-2020	7024	41 (15-04-2020 – 26-05-2020)	11800	0.068	204	30-08-2020
27-05-2020	7261	42 (15-04-2020 – 27-05-2020)	12063	0.068	204	31-08-2020
28-05-2020	7453	43 (15-04-2020 – 28-05-2020)	12236	0.068	204	01-09-2020
29-05-2020	7645	44 (15-04-2020 – 29-05-2020)	12348	0.067	203	02-09-2020
30-05-2020	7891	45 (15-04-2020 – 30-05-2020)	12498	0.067	202	02-09-2020
31-05-2020	8089	46 (15-04-2020 – 31-05-2020)	12607	0.067	201	03-09-2020
01-06-2020	8283	47 (15-04-2020 – 01-06-2020)	12684	0.067	200	03-09-2020
02-06-2020	8420	48 (15-04-2020 – 02-06-2020)	12678	0.067	199	03-09-2020
03-06-2020	8588	49 (15-04-2020 – 03-06-2020)	12655	0.067	198	03-09-2020
04-06-2020	8762	50 (15-04-2020 – 04-06-2020)	12632	0.067	196	03-09-2020
05-06-2020	8996	51 (15-04-2020 – 05-06-2020)	12661	0.067	195	03-09-2020
06-06-2020	9228	52 (15-04-2020 – 06-06-2020)	12727	0.067	195	03-09-2020
07-06-2020	9401	53 (15-04-2020 – 07-06-2020)	12779	0.067	194	04-09-2020
08-06-2020	9638	54 (15-04-2020 – 08-06-2020)	12865	0.067	193	04-09-2020
09-06-2020	9849	55 (15-04-2020 – 09-06-2020)	12960	0.066	193	05-09-2020
10-06-2020	10049	56 (15-04-2020 – 10-06-2020)	13056	0.066	193	05-09-2020
11-06-2020	10241	57 (15-04-2020 – 11-06-2020)	13150	0.066	193	06-09-2020
12-06-2020	10443	36 (07-05-2020 – 12-06-2020)	13849	0.063	197	12-09-2020

TABLE XIV
PREDICTION TREND FOR WEST BENGAL (WB)

Date	Total cases	Best window	K	r	E_{rms}	End on
23-05-2020	3459	44 (09-04-2020 – 23-05-2020)	4559	0.110	36	15-07-2020
24-05-2020	3667	48 (06-04-2020 – 24-05-2020)	4790	0.107	40	18-07-2020
25-05-2020	3816	48 (07-04-2020 – 25-05-2020)	4949	0.105	43	19-07-2020
26-05-2020	4009	41 (15-04-2020 – 26-05-2020)	5203	0.102	48	22-07-2020
27-05-2020	4192	35 (22-04-2020 – 27-05-2020)	5479	0.100	54	25-07-2020
28-05-2020	4536	34 (24-04-2020 – 28-05-2020)	5990	0.096	67	30-07-2020
29-05-2020	4813	34 (25-04-2020 – 29-05-2020)	6707	0.091	80	05-08-2020
30-05-2020	5130	35 (25-04-2020 – 30-05-2020)	7490	0.088	93	10-08-2020
31-05-2020	5501	36 (25-04-2020 – 31-05-2020)	8577	0.084	108	17-08-2020
01-06-2020	5772	37 (25-04-2020 – 01-06-2020)	9719	0.081	117	22-08-2020
02-06-2020	6168	38 (25-04-2020 – 02-06-2020)	11231	0.079	128	28-08-2020
03-06-2020	6508	39 (25-04-2020 – 03-06-2020)	12882	0.077	135	02-09-2020
04-06-2020	6876	40 (25-04-2020 – 04-06-2020)	14682	0.075	141	07-09-2020
05-06-2020	7303	41 (25-04-2020 – 05-06-2020)	16826	0.073	146	12-09-2020
06-06-2020	7738	42 (25-04-2020 – 06-06-2020)	19247	0.072	151	16-09-2020
07-06-2020	8187	43 (25-04-2020 – 07-06-2020)	21872	0.071	155	20-09-2020
08-06-2020	8613	37 (02-05-2020 – 08-06-2020)	27874	0.069	156	29-09-2020
09-06-2020	8985	38 (02-05-2020 – 09-06-2020)	29406	0.068	156	30-09-2020
10-06-2020	9328	39 (02-05-2020 – 10-06-2020)	29382	0.068	155	30-09-2020
11-06-2020	9768	40 (02-05-2020 – 11-06-2020)	29365	0.068	154	30-09-2020
12-06-2020	10244	41 (02-05-2020 – 12-06-2020)	29561	0.068	153	01-10-2020