J. Inst. Eng. India Ser. B (December 2021) 102(6):1167–1176
https://doi.org/10.1007/s40031-021-00608-3

ORIGINAL CONTRIBUTION

CARD Predictive Modeling and SEI Formulation: COVID-19 Statistics in India

Debjit Majumder[1] · Sougata Mazumder[1] · Prasun Ghosal[1]

Received: 3 July 2020 / Accepted: 23 April 2021 / Published online: 20 May 2021

Abstract The current scenario of the pandemic COVID-19 has been a source of anchorage for researchers, healthcare professionals, and statisticians. Based on the immense data, it has been observed that the role of statistics has been crucial in researching and at the same for predicting the COVID-19 scenario of the entire globe. This paper deals with extensive data collection and predictive modeling to derive a CARD model using statistical tools like regression curve fitting. The exponential growth model has been prevalent in live updates via COVID-19 dashboards maintained by different organizations like WHO, Johns Hopkins University, Indian Council of Medical Research. In a similar tone, the paper discusses a time-varying exponential growth model specific to the Indian condition. However, a generic model has been derived by different researchers of other countries. The model accuracy has been considered satisfactory. Moreover, a State-wise Evaluation Indexing has been performed considering parameters like sanitation, population below the poverty line, literacy rate, and population density. Results have been shown for better data visualization through cartograms. The conclusions are noteworthy, and the CARD model can be trained and developed with better accuracy using the concept of machine and deep learning, keeping in context the huge amount of instantaneous data being generated every day all over the world.

Keywords CARD · Predictive modeling · Statistical analysis · Time-varying exponential growth model · Regression · SEI · Cartogram

Introduction

With the recent outbreak of the pandemic of COVID-19, every country has made significant steps to combat this widely spreading threat to the human community. Mathematical and statistical models have been produced to forecast the future of COVID-19 [15]. For India, relative to other countries in the EU, the spread has been less, which can be attributed to early lockdown implementation throughout the country. However, it has been seen that the cases of confirmed and death rates have been steadily increasing in the country, as shown in Figs. 1 and 2. Both of these show increasing trends and have been considered as a major parameter in world COVID-19 dashboards, maintained by many organizations like Center for Systems Science and Engineering (CSSE) at Johns Hopkins University [4], World Health Organization (WHO) [5]. This situation emphasizes the need for a reliable and accurate statistical model that can not only fit well to the existing data, but also can predict with considerable accuracy. There have been predictive models published by World Health Organization [2] that discussed a model of susceptible, exposed, infected, and removed (SEIR) cases, and it has been reformed by alternating various parameters by many leading research organizations.

✉ Sougata Mazumder
sougatamazumder97@gmail.com

Debjit Majumder
debjit.ankur@gmail.com

Prasun Ghosal
p_ghosal@it.iiests.ac.in

[1] Indian Institute of Engineering Science and Technology, Shibpur, Howrah, West Bengal 711103, India

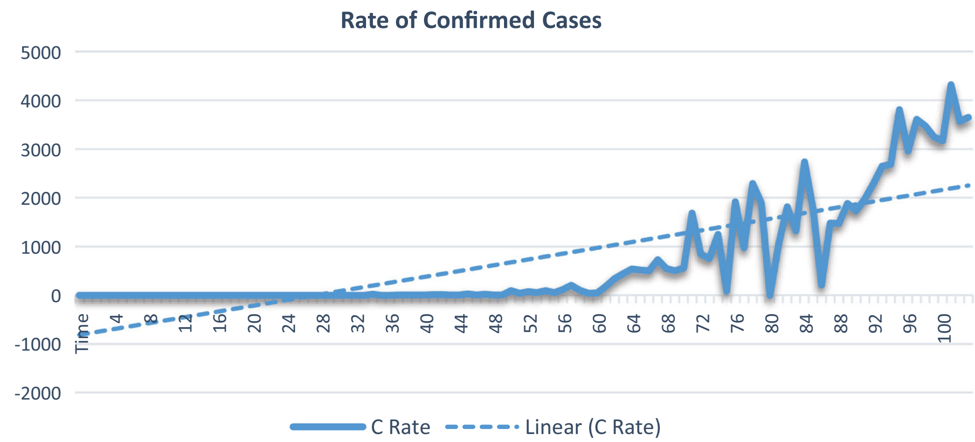

Fig. 1 Variation of rate of confirmed cases

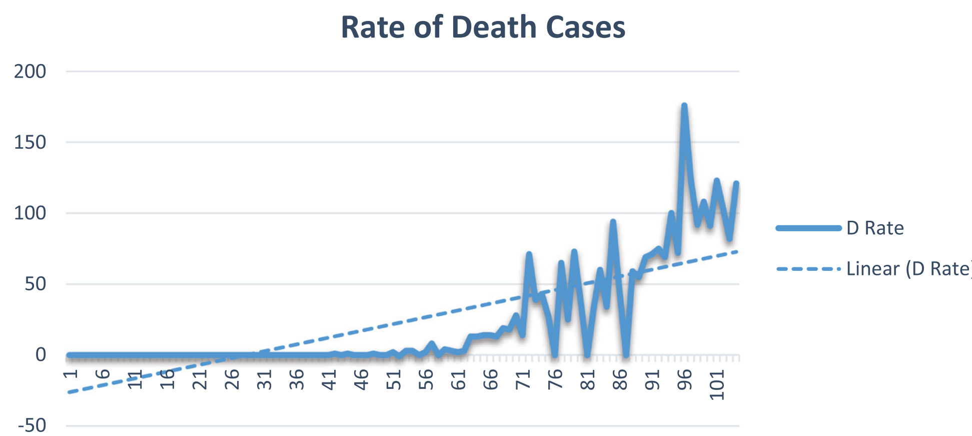

Fig. 2 Variation of rate of death cases

Motivation

We decided to produce a predictive growth model during our research using statistical tools and generalize it for various countries. This study comprised of obtaining the specific growth model of the Indian COVID-19 scenario, updated from 30th January 2020 (when the first confirmed case was detected in India, taken as $t = 1$) till 13th May 2020 ($t = 104$). However, this predictive empirical model can be used country-wise, using the data available. The growth model had a very high R^2 value, and the deviations of the forecasted data from the actual data were considerably small. The model can predict the future of Indian COVID-19 and be of great help for policy-makers, researchers, scientists, and medical professionals to combat this deadly virus and control its spread. The forecasted values can also be termed reliable, and suitable hospitalization can be improved in India. Based on this, future research can be carried out on how to control the reproduction number and doubling time and improve test per million, keeping in mind the socio-demographic diversity of a country like India. The data collection methodology, statistical analysis, growth models, and data validation have been shown.

Related Works

On 11th February, 2020, the World Health Organisation announced that severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) as the virus responsible for the pandemic of COVID-19 [21]. To support a sound ecosystem of research in this pandemic, WHO agreed to work with other countries. It was directed that countries affected by this pandemic shall implement proper measures to detect, isolate, and treat potential cases, trace contact along with incorporating social distancing measures and to take crucial steps during the early spread of this disease [22, 25–27, 31] . On 11th March, 2020, there were reportedly more than 118,000 cases in 114 countries and 4291 people have lost their lives. Keeping this in context and probable risk of foreseeing, alarming levels of this spread WHO declared COVID-19 as a global pandemic [23]. Other situation reports of every month released by WHO have been very essential to emphasize the need of this model at this stage of ever increasing cases of COVID-19 in India and abroad [24]. A joint mission was launched between WHO and China for better knowledge sharing and combating this pandemic. Updates regarding this mission helped countries to monitor how it dealt with the problems

[28–30]. As on 13th May, 2020, globally 4,170,424 cumulative cases and 287,399 deaths [32] have been recorded highlighting the fact that models such as SEIR, SIQR and CARD shall be a great boon for predicting the future of COVID-19. The spread of this disease shall be tackled by following the latest updated guidelines by WHO.

In a similar note as the SEIR model released in the WHO bulletin, other models such as susceptible-infectious-quarantined-recovered (SIQR) have been derived [17]. Indicators were compared with other countries such as Italy and Brazil. However, a purely statistical model using the specific case of India has been missing. Thus, there is an ardent need for the CARD model, which can be used for improved governance and counteracting the threat of SARS-Cov-2, a novel virus responsible for the rapid spread of COVID-19 around the world [8, 9].

Several statisticians and mathematicians are publishing models online, and the race to establish a particular model is enormous. However, loopholes are there due to the maximum of such models, not taking into consideration the dynamic nature of COVID-19. The forecasts and nowcasts are being highlighted, and the peer review process does not encounter most of the models published online, and thus have erroneous predictions [14]. The CARD model and the SEI takes into consideration the actual statistical modeling parameters which genuinely contributes to the spread of this deadly virus, along with a comparative study. The need for such an established model is the need of the hour. It can be a massive boon for the researchers (specific not only to India) to strategic combat plans and increasing data. It can also train this particular predictive model to give more accurate results shortly [3]. The challenges addressed through this model include the following.

- To represent the current scenario of COVID-19 in India.
- Absence of an accurate model to predict future status.
- Apart from the SEIR model, the need for a purely statistical model generic to all countries.
- To understand exactly how the different Indian states and union territories are performing using four basic parameters and rating them.

Similar rating styles can be adopted by other countries to have an idea of what parameters shall be improved in individual states to help them perform better. Funds allocated for fighting this pandemic through data-driven decision making shall be a huge help for convenient governance. Most of the models have not been supplemented with the State-wise Evaluation Index (SEI), and this CARD model being time-varying has considerable accuracy. At this stage of the pandemic in India, the models

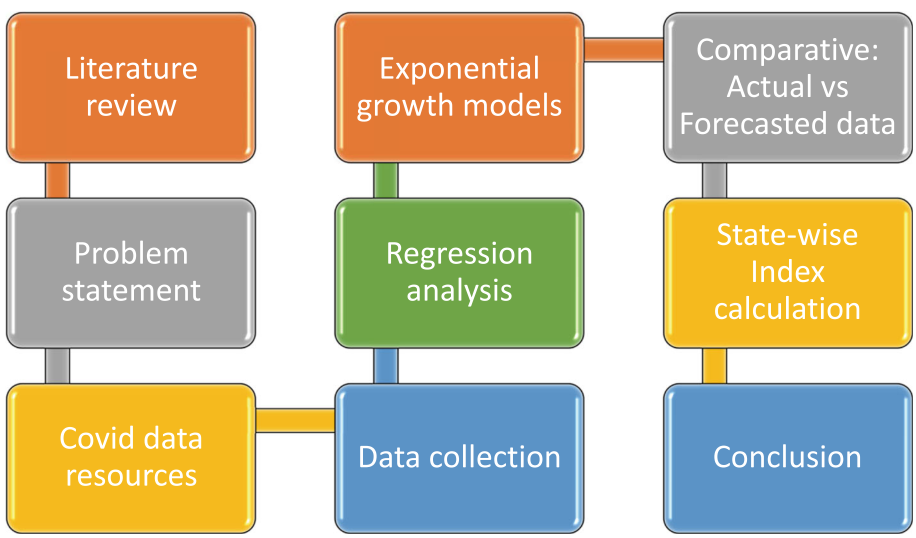

Fig. 3 Methodology flowchart

Table 1 Regressional growth model

Serial no.	Parameter	Growth model	Residual sum of squares (RSS)	R^2
1	Confirmed cases (C)	$C(t) = 10668833.22e^{-522.54/t}$	35,807,181.64	0.997864646
2	Recovered cases (R)	$R(t) = (1.103855539)^t$	9,631,542.75	0.9918079304
3	Death cases (D)	$D(t) = 446315.56e^{-544.51/t}$	60,915.24	0.9966717303

Table 2 CARD system data validation: actual versus derived with deviations

Date	Time	Confirmed			Recovered			Death			Active		
		Actual	Derived	Deviation	Actual	Derived	Deviation	Actual	Derived	Deviation	Actual	Derived	Deviation
18-Mar	49	142	249	− 107	14	126.67	− 112.67	3	6.66	− 3.66	125	115.98	9.01
19-Mar	50	149	309	− 160	15	139.83	− 124.83	3	8.31	− 5.31	131	160.44	− 29.44
20-Mar	51	244	379	− 135	20	154.35	− 134.35	5	10.29	− 5.29	219	214.12	4.87
21-Mar	52	283	461	− 178	23	170.38	− 147.38	4	12.64	− 8.64	256	278.23	− 22.23
22-Mar	53	360	558	− 198	27	188.08	− 161.08	7	15.40	− 8.40	326	354.07	− 28.07
23-Mar	54	415	669	− 254	27	207.61	− 180.61	10	18.63	− 8.63	378	443.00	− 65.00
24-Mar	55	511	798	− 287	40	229.17	− 189.17	10	22.38	− 12.38	461	546.43	− 85.43
25-Mar	56	562	946	− 384	43	252.97	− 209.97	12	26.71	− 14.71	507	665.85	− 158.85
26-Mar	57	680	1114	− 434	45	279.25	− 234.25	20	31.68	− 11.68	615	802.78	− 187.78
27-Mar	58	887	1304	− 417	73	308.25	− 235.25	20	37.36	− 17.36	794	958.81	− 164.81
28-Mar	59	987	1520	− 533	84	340.26	− 256.26	24	43.80	− 19.80	879	1135.56	− 256.56
29-Mar	60	1,024	1761	− 737	95	375.60	− 280.60	27	51.09	− 24.09	902	1334.65	− 432.65
30-Mar	61	1,071	2032	− 961	100	414.61	− 314.61	29	59.28	− 30.28	942	1557.75	− 615.75
31-Mar	62	1,251	2333	− 1082	102	457.67	− 355.67	32	68.46	− 36.46	1117	1806.53	− 689.53
1-Apr	63	1,590	2667	− 1077	148	505.20	− 357.20	45	78.70	− 33.70	1397	2082.66	− 685.66
2-Apr	64	2,032	3036	− 1004	148	557.67	− 409.67	58	90.08	− 32.08	1826	2387.78	− 561.78
3-Apr	65	2,567	3442	− 875	192	615.59	− 423.59	72	102.68	− 30.68	2303	2723.55	− 420.55
4-Apr	66	3,082	3888	− 806	229	679.52	− 450.52	86	116.58	− 30.58	2767	3091.55	− 324.55
5-Apr	67	3,588	4375	− 787	229	750.10	− 521.10	99	131.86	− 32.86	3260	3493.34	− 233.34
6-Apr	68	4,314	4907	− 593	328	828.00	− 500.00	118	148.60	− 30.60	3868	3930.43	− 62.43
7-Apr	69	4,858	5485	− 627	382	913.99	− 531.99	136	166.88	− 30.88	4340	4404.23	− 64.23
8-Apr	70	5,360	6112	− 752	468	1008.92	− 540.92	164	186.80	− 22.80	4728	4916.09	− 188.09
9-Apr	71	5,916	6789	− 873	506	1113.70	− 607.70	178	208.43	− 30.43	5232	5467.27	− 235.27
10-Apr	72	7,600	7520	80	645	1229.36	− 584.36	249	231.85	17.14	6706	6058.89	647.10
11-Apr	73	8,446	8306	140	840	1357.04	− 517.04	288	257.16	30.83	7318	6691.96	626.03
12-Apr	74	9,205	9150	55	951	1497.97	− 546.97	331	284.44	46.55	7923	7367.36	555.63
13-Apr	75	10,453	10053	400	1052	1653.55	− 601.55	358	313.76	44.23	9043	8085.79	957.20
14-Apr	76	10,541	11018	− 477	1205	1825.28	− 620.28	358	345.21	12.78	8978	8847.78	130.21
16-Apr	77	12,456	12047	409	1513	2014.84	− 501.84	423	378.87	44.12	10520	9653.66	866.33
17-Apr	78	13,430	13142	288	1768	2224.10	− 456.10	448	414.83	33.16	11214	10503.56	710.43
18-Apr	79	15,722	14306	1416	2463	2455.08	7.91	521	453.15	67.84	12738	11397.36	1340.63
19-Apr	80	17,615	15539	2076	2854	2710.06	143.93	559	493.93	65.06	14202	12334.68	1867.31
20-Apr	81	17,615	16844	771	3273	2991.51	281.48	559	537.22	21.77	13783	13314.85	468.14
21-Apr	82	18,658	18222	436	3273	3302.20	− 29.20	592	583.12	8.87	14793	14336.91	456.08
22-Apr	83	20,471	19676	795	3976	3645.15	330.84	652	631.69	20.30	15843	15399.53	443.46
23-Apr	84	21,797	21208	589	4376	4023.72	352.27	686	683.00	2.99	16735	16501.03	233.96
24-Apr	85	24,530	22818	1712	5498	4441.61	1056.38	780	737.13	42.86	18252	17639.32	612.67
25-Apr	86	26,283	24509	1774	5939	4902.89	1036.10	825	794.13	30.86	19519	18811.84	707.15
26-Apr	87	26,496	26282	214	5939	5412.08	526.91	825	854.08	− 29.08	19732	20015.57	− 283.57
27-Apr	88	27,977	28138	− 161	6523	5974.16	548.83	884	917.04	− 33.04	20570	21246.97	− 676.97
28-Apr	89	29,451	30080	− 629	7137	6594.61	542.38	939	983.07	− 44.07	21375	22501.87	− 1126.87
29-Apr	90	31,332	32107	− 775	7747	7279.50	467.49	1008	1052.22	− 44.22	22577	23775.52	− 1198.52
30-Apr	91	33,062	34223	− 1161	8437	8035.51	401.48	1079	1124.55	− 45.55	23546	25062.45	− 1516.45
1-May	92	35,043	36427	− 1384	9068	8870.05	197.94	1154	1200.12	− 46.12	24821	26356.42	− 1535.42
2-May	93	37,336	38721	− 1385	10007	9791.25	215.74	1223	1278.98	− 55.98	26106	27650.37	− 1544.37
3-May	94	39,980	41106	− 1126	10819	10808.13	10.86	1323	1361.18	− 38.18	27838	28936.34	− 1098.34
4-May	95	42,670	43583	− 913	11782	11930.61	− 148.61	1395	1446.76	− 51.76	29493	30205.34	− 712.34

Table 2 continued

Date	Time	Confirmed			Recovered			Death			Active		
		Actual	Derived	Deviation	Actual	Derived	Deviation	Actual	Derived	Deviation	Actual	Derived	Deviation
5-May	96	46,476	46153	323	12849	13169.67	− 320.67	1571	1535.77	35.22	32056	31447.31	608.68
6-May	97	49,436	48817	619	14183	14537.42	− 354.42	1695	1628.25	66.74	33558	32650.98	907.01
7-May	98	53,045	51575	1470	15331	16047.21	− 716.21	1787	1724.24	62.75	35927	33803.74	2123.25
8-May	99	56,516	54429	2087	16867	17713.80	− 846.80	1895	1823.78	71.21	37754	34891.57	2862.42
9-May	100	59,765	57379	2386	17897	19553.48	− 1656.48	1986	1926.90	59.09	39882	35898.81	3983.18
10-May	101	62,939	60426	2513	19358	21584.22	− 2226.22	2109	2033.64	75.35	41472	36808.08	4663.91
11-May	102	67259	63570	3689	20969	23825.86	− 2856.86	2212	2144.02	67.97	44078	37600.06	6477.93
12-May	103	70827	66812	4015	22549	26300.30	− 3751.30	2294	2258.07	35.92	45984	38253.31	7730.68
13-May	104	74480	70152	4328	24453	29031.74	− 4578.74	2415	2375.82	39.17	47612	38744.05	8867.94

Table 3 Mean standard error for last 10 days data set

Serial no.	Parameter	Mean standard error (%)
1	Confirmed	2.59
2	Recovered	7.78
3	Death	1.79
4	Active	7.56

produced earlier have a limited data set for operation, and this CARD model incorporates the data in its modeling until 13th May 2020, latest. Moreover, it can control the new outbreak of this pandemic which can result in further waves. [13, 18]. The only limitation of this model is the effects due to partial lifting of lockdown in India on community spreading, which has not been considered. However, further research is already being done by us to estimate the effect of the partial lifting of lockdown on COVID-19 statistics in the Indian context.

Summary of Major Contributions

Novelties of this work lie in the manifold and may be summarized as follows.

- A well-established growth model (CARD) is developed, which can be used to show the relationship between time and COVID-19 cases in India.
- A generic relation is established between time and COVID-19 cases, which can be applied to any country in the world.
- The CARD model can act as a founding stone for statisticians and scientists to develop a more well-trained and advanced model.
- In contrary to other mathematical models, the CARD model is purely statistical in nature and is based on reliable data, which is updated daily by credible sources like WHO, ICMR, MHA.
- 37 different states and union territories of India are given index points based on their performances on certain parameters like population BPL, sanitation, literacy rate, and population density. This variation has been shown using a cartogram, which can help the decision-makers to improve on these parameters and act judiciously.

Organization of the Paper

The overall organization of the rest of the paper is as follows. Section 1.2 discusses the background of the present work and details regarding the related state-of-the-art research works. It highlighted the recent works done in modeling. It emphasizes the need to undertake this research study and its limitations to enable future research to improve this CARD model by feeding more data into it. Section 1.3 enlists the novel contributions of this work and its impact on future researches, decision-making processes of policymakers, socio-economic parameters, and the ardent need for this to reach the researchers. Section 2 in the paper summarizes the time-varying growth models for India and the deduction of a generic version for the world. Section 3 deals with the comparative study between the actual and derived data, showcasing the accuracy of this CARD model with the deviations mentioned, which has a very small mean standard error for the last ten day's data set. The accuracy can be further increased as we go on with data for the next 1-2 months and a study can surely be done to analyze the statistical differences in COVID-19, after partial lock-down relaxation in India starting from 17th May 2020. Section 4 is a section that supplements the entire CARD model in its approach towards helping stakeholders in their decision-making process. The State-wise Evaluation Index (SEI) has been formed tactically to

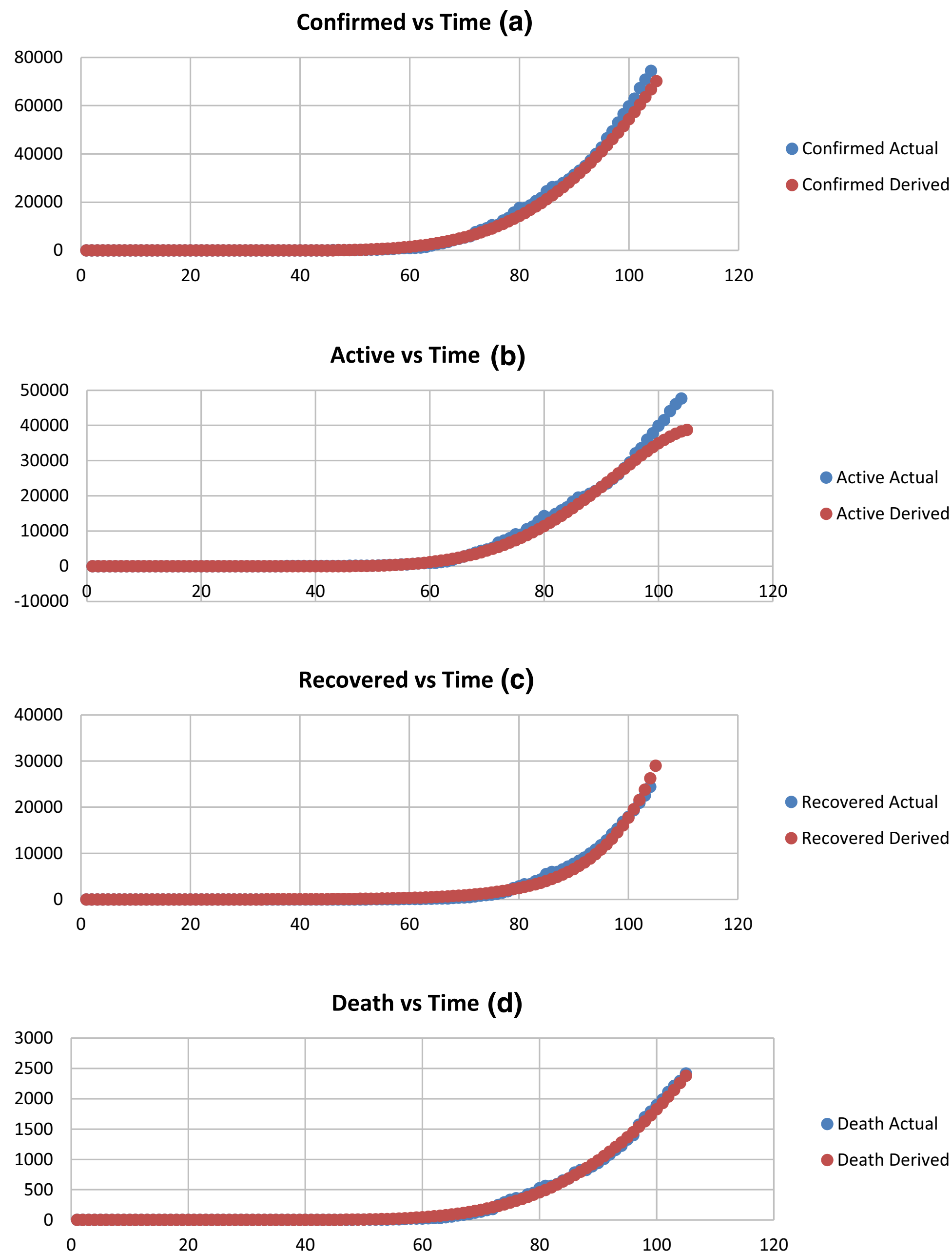

Fig. 4 Curve fitting actual versus derived data

rate the states based on some fixed parameters such as sanitation, population below poverty line, literacy rate, and population density. Most of the researches done [11, 12, 19] highlighted the concept of reproduction number, doubling time, temperature, humidity, etc. This aims to bridge the gap, and those researches can now be combined to more accurately determine the states' performance in lieu of its response to COVID-19 and what factors shall be improved. Visualization at a glance has been given using a cartogram [4, 5]. The paper concludes in Sect. 5 with some noteworthy suggestions which can surely prove to be crucial, followed by references in the end. An overview of the workflow adopted across the entire work is depicted in the flowchart (see Fig. 3).

Modeling

Raw data are one of the most significant parts of this study. The data that support the findings of this study are available in Statista at www.statista.com [7]. On this data, a series of regressions was performed, and ultimately a particular solution was reached. A lot of trial and error was carried out choosing various linear and nonlinear regression models, but ultimately, the following solutions have been proven to be the most optimized empirically. As per CSSE by Johns Hopkins University [4] and WHO COVID-19 status dashboard [5], it was observed that most of the models were exponential growth models, which were back validated with our trial and error method. The models obtained have been tabulated in Table 1.

where $C(t)$, $R(t)$, and $D(t)$ represent time-varying functions of confirmed, recovered, and death cases (specifically for India), respectively, and RSS denotes the residual sum of squares which should be minimized in order to get an optimized equation, and R^2 is an indicator of the extent of curve fitting which shall ideally be 1. It is quite evident that the total number of confirmed cases till date is the summation of the total number of recovered, death and active cases, respectively, till date. Thus, the total number of active cases till date can be found out by subtracting the summation of the total number of recovered and death cases, respectively, from the total number of confirmed cases.

$$C(t) = R(t) + D(t) + A(t) \tag{1}$$

or,

$$A(t) = 10668833.22e^{-522.54/t} - (1.103855539^t + 446315.56e^{-544.51/t}) \tag{2}$$

where $A(t)$ denotes the time-varying function of active cases.

As a matter of fact, a more generic version of this model can be found out which will be applicable to every country in the world, where in, the time variant equation of confirmed, recovered and death cases will change in the following way:

$$C(t) = \kappa e^{-\alpha/t} \tag{3}$$

$$R(t) = \mu^t \tag{4}$$

$$D(t) = \lambda e^{-\beta/t} \tag{5}$$

Therefore, using equation (2) above,

$$A(t) = \kappa e^{-\alpha/t} - \mu^t - \lambda e^{-\beta/t} \tag{6}$$

The values of κ, α, μ, λ and β are country specific and can be found out by carrying out a series of regressions on the COVID-19 data of that particular country.

Comparative Study

Using the equations mentioned above in Table 1, obtained from the regression analysis, a comparative study between the actual data and the derived data has been carried out. Due to inconsistency in data from 30th January 2020 ($t = 1$) to 17th March 2020 ($t = 48$), there was an anomaly in the forecasting. This anomaly is attributed to the poor quality of data and thus has been eliminated. Although the regression has been done for the entire set of data, only the result from 18th March 2020 ($t = 49$) to 13th May 2020 ($t = 104$) has been shown below in Table 2 for confirmed, recovered, death and active cases, respectively. Deviations have also been shown from the actual data to emphasize the accuracy of this model. Mean standard error calculation of each parameter has been shown for a sample space of the last 10 days, i.e., from 3rd May 2020 to 13th May 2020 in Table 3. The accuracy of this confirmed-active-recovered-death (CARD) model is demonstrated by a parity between the actual data and derived data through curve fitting (see Fig. 4).

State-Wise Evaluation Index (SEI) Calculation

A state-wise rating has been formulated to calculate the index of every state based on some of the fixed basic parameters, which can presumably play an important role

Table 4 Corresponding scales for parametric rating

Parameters-rating	Sanitation	Population below poverty line	Literacy rate	Population density
1	16.7–24.944	32.6–29.34	61.8–65.02	$\geq$ 2598
2	24.944–33.188	29.34–26.08	65.02–68.24	2598–2309.64
3	33.188–41.432	26.08–22.82	68.24–71.46	2309.64–2021.29
4	41.432–49.676	22.82–19.56	71.46–74.68	2021.29–1732.93
5	49.676–57.92	19.56–16.3	74.68–77.9	1732.93–1444.58
6	57.92–66.164	16.3–13.04	77.9–81.12	1444.58–1156.22
7	66.164–74.408	13.04–9.78	81.12–84.34	1156.22–867.87
8	74.408–82.652	9.78–6.52	84.34–87.56	867.87–579.51
9	82.652–90.896	6.52–3.26	87.56–90.78	579.51–291.16
10	90.896–99.14	3.26–0.00	90.78–94	291.16–2.8

in the fight against COVID-19. The parameters are taken for index calculation, and the corresponding scales have been mentioned in Table 4. All the states and union territories have been considered the latest updated in 2020. The sources of the data have also been mentioned on why that particular parameter has been chosen for indexing.

The data (for sanitation and population below poverty line of every state and union territories) that support this study's findings are openly available in the Handbook of Urban Statistics 2019, Ministry of Housing and Urban Affairs, Government of India [16].

Sanitation denotes the percentage of the population having necessary hand-washing facilities. As per the Interim Guidance Report released by WHO on 23rd April 2020 [20], the water, sanitation, and hygiene (WASH) play a considerable part in this context. Therefore, it is a very crucial point to evaluate states based on sanitation levels. Population below poverty line (expressed in percentage of population) is essential to highlight the effect of COVID-19 and how it shall affect the vital topic of cash fluidity in the bottom portion of the pyramid, ensuring financial stability. Literacy rate (denoted by the percentage of the total

Table 5 State-wise Evaluation Indexing (SEI)

Sl. no.	State	Sanitation	Population BPL	Literacy rate	Population density	Total rating
1	Andaman and Nicobar Islands	99.14	0	86.63	46	38
2	Andhra Pradesh	94.45	5.8	67.02	303	30
3	Arunachal Pradesh	84.82	20.3	65.39	17	25
4	Assam	93.07	20.5	72.19	397	27
5	Bihar	95.27	31.2	61.8	1102	19
6	Chandigarh	98.84	22.3	89.05	9252	24
7	Chhattisgarh	92.58	24.8	70.28	189	25
8	Dadra and Nagar Haveli	86.1	15.4	76.24	970	27
9	Daman and Diu	95.75	12.6	87.1	970	32
10	Delhi	91.08	9.8	86.21	11297	26
11	Goa	96.44	4.1	88.7	394	37
12	Gujarat	95.52	10.1	78.03	308	32
13	Haryana	95.26	10.3	75.55	573	31
14	Himachal Pradesh	94.77	4.3	82.8	123	36
15	Jammu and Kashmir	98.72	7.2	67.16	98	30
16	Jharkhand	84.81	24.8	66.41	414	23
17	Karnataka	94.52	15.3	75.37	319	30
18	Kerala	95.57	5	94	859	37
19	Ladakh	16.7	0	79.98	2.8	27
20	Lakshadweep	95.1	3.4	91.85	2013	33
21	Madhya Pradesh	94.19	21	69.32	236	27
22	Maharashtra	96.9	9.1	82.34	365	34
23	Manipur	90	32.6	76.9	122	25
24	Meghalaya	96.47	9.3	74.43	132	32
25	Mizoram	98.77	6.4	91.33	52	39
26	Nagaland	91.24	16.5	79.6	119	31
27	Odisha	76.06	17.3	72.89	269	27
28	Puducherry	94.64	6.3	85.85	2598	29
29	Punjab	98.97	9.2	75.84	550	32
30	Rajasthan	96.13	10.7	66.11	201	29
31	Sikkim	97.77	3.7	81.42	86	36
32	Tamil Nadu	90.74	6.5	80.09	555	33
33	Telangana	94.4	22.89	66.86	312	24
34	Tripura	90.18	7.4	87.22	350	34
35	Uttar Pradesh	97.19	26.1	67.68	828	22
36	Uttarakhand	97.22	10.5	78.82	189	32
37	West Bengal	85.31	14.7	76.26	1029	27

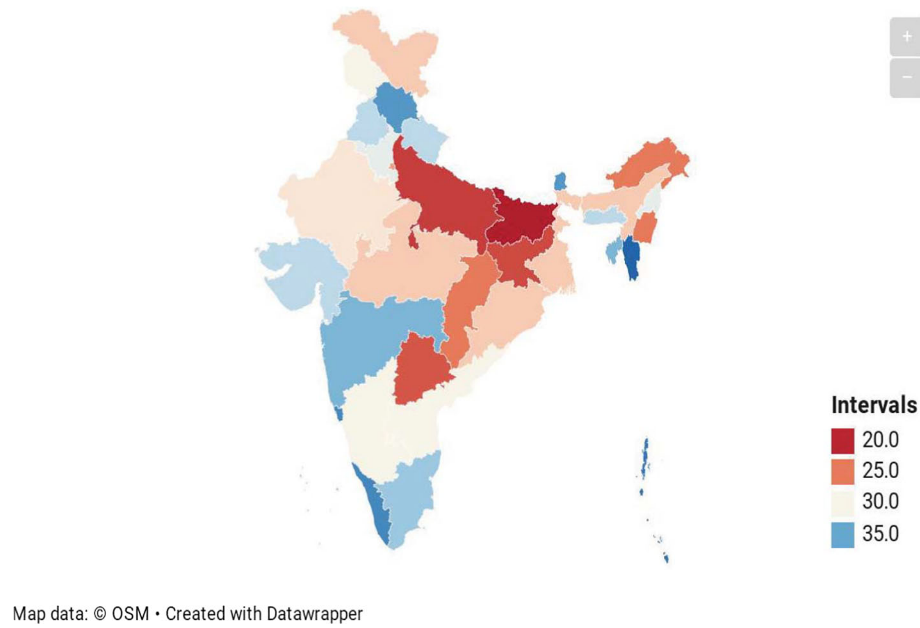

Fig. 5 SEI visualization

population) can be related to the understanding of the standard population and its sequential reaction to this pandemic, state-wise. In general, responding to the awareness campaigns and being socially aware can be considered an essential yardstick for fighting against COVID-19 and preventing community spreading. Population density (inferred about population/square kilometer) is an indicator of expectation of community spreading, rising, and correlation among all the above parameters. The data (for literacy rate and population density of every state and union territories) that support this study's findings are openly available in Wikipedia under "List of states and union territories of India by population" [10] to increase accuracy with updated data till 2020. Table 5 shows the Indian state-wise data along with the calculated index using the rating system discussed already.

For better visualization of the index calculated for every state and union territories, Datawrapper has been used to create a cartogram, which will help state-wise analysis in India based on the four parameters. It has been shown in Fig. 5. Credit for the map tile used for cartography goes to OSM and its contributors [1, 6] [©OpenStreetMap contributor, more details here: https://www.openstreetmap.org/copyright].

Conclusion

We have developed a CARD model using statistical tools and predictive modeling, which emphasizes on the need for a time-variant growth model to focus on the effect of COVID-19 in Indian conditions. However, a generic equation is generated for further scopes of study specific to other countries. The model is quite accurate, as indicated by the mean standard error of the sample data. Curve fitting further highlights the parity between the actual and derived data and can be a real asset for predicting the future of COVID in India, provided the current circumstances remain constant.

The model is applicable to the lockdown condition and does not consider the dynamic parameters which play a role during unlock down scenario. In that condition, the model does not perform well and gives a significant deviation in the actual and derived curves. Potential dynamic factors are mobility, migration, etc. With more data being generated every day, this model can be trained, and the accuracy can be increased using machine learning and deep learning techniques, which is one of the future scopes of the researchers.

From the cartogram (Fig. 5) and SEI:

The states, having a rating of more than 35, have set themselves a class apart in dealing with the pandemic of COVID-19 appropriately. Although the index scores are considerably better, there are shortcomings in other dynamic parameters for the states with ratings between 30-35, such as for Maharashtra. Its location and constant exposure to international travelers through airports are critical parameters attributing to its latest worsening condition. The rest of the states whose index numbers are below 30 lacks planning and its implementation. Development of factors such as test per million, reproduction number, doubling time, hospitalization, and availability of beds shall be a benevolence for the country, in this fight against COVID-19.

References

1. J. Bennett, *OpenStreetMap* (Packt Publishing Ltd, Birmingham, 2010)
2. F.A. Binti Hamzah, C. Lau, H. Nazri, D.V. Ligot, G. Lee, C.L. Tan, et al. *CoronaTracker: Worldwide COVID-19 Outbreak Data Analysis and Prediction.* [Preprint]. Bull World Health Organ. E-pub: 19 March 2020. https://doi.org/10.2471/BLT.20.255695
3. P. Brown, Studying COVID-19 in light of critical approaches to risk and uncertainty research pathways, conceptual tools, and some magic from Mary Douglas (2020)
4. coronavirus.jhu.edu/map.html. Accessed 17 May 2020
5. covid19.who.int/. Accessed 17 May 2020
6. M. Haklay, P. Weber, Openstreetmap: user-generated street maps. IEEE Pervasive Comput. **7**(4), 12–18 (2008)
7. India- Coronavirus Covid-19 Daily Confirmed Recovered Death Cases, https://www.statista.com/statistics/1104054/india-coronavirus-covid-19-daily-confirmed-recovered-death-cases/. Accessed 17 May 2020
8. A.J. Kucharski, T.W. Russell, C. Diamond, Y. Liu, J. Edmunds, S. Funk, R.M. Eggo, F. Sun, M. Jit, J.D. Munday, N. Davies, Early dynamics of transmission and control of COVID-19: a mathematical modelling study. Lancet Infect Dis **20**, 553–558 (2020)
9. R. Li, S. Pei, B. Chen, Y. Song, T. Zhang, W. Yang, J. Shaman, Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). Science **368**(6490), 489–493 (2020)
10. List of states and union territories of India by population, https://en.wikipedia.org/wiki/List of states and union territories of India by population (For data: Literacy Rate and Population Density). Accessed 17 May 2020
11. Y. Liu, A.A. Gayle, A. Wilder-Smith, J. Rocklöv, The reproductive number of COVID-19 is higher compared to SARS coronavirus. J. Travel Med. **1**, 4 (2020)
12. W. Luo, M. Majumder, D. Liu, C. Poirier, K. Mandl, M. Lipsitch, M. Santillana, The role of absolute humidity on transmission rates of the COVID-19 outbreak (2020)
13. J. Qiu, Covert coronavirus infections could be seeding new outbreaks. Nature (2020)
14. T. Rhodes, et al., A model society: maths, models and expertise in viral outbreaks (2020)
15. K. Roosa, Y. Lee, R. Luo, A. Kirpich, R. Rothenberg, J. Hyman, Real-time forecasts of the COVID-19 epidemic in china from February 5th to February 24th, 2020. Infect Dis Model **5**, 256–263 (2020). https://doi.org/10.1016/j.idm.2020.02.002
16. The Handbook of Urban Statistics 2019, Ministry of Housing and Urban Affairs, Government of India, http://mohua.gov.in/pdf/5c80e2225a124Handbook. Accessed 17 May 2020
17. A. Tiwari, Modelling and analysis of COVID-19 epidemic in India (2020), https://www.medrxiv.org/content/10.1101/2020.04.12.20062794v1.full.pdf
18. C. Wang, P.W. Horby, F. Hayden, G.F. Gao, A novel coronavirus outbreak of global health concern. Lancet **395**, 470–473 (2020). https://doi.org/10.1016/S0140-6736(20)30185-9
19. M. Wang, A. Jiang, L. Gong, L. Luo, W. Guo, C. Li, B. Yang, J. Zeng, Y. Chen, Temperature significant change COVID-19 Transmission in 429 cities. MedRxiv (2020)
20. Water Sanitation Hygiene and Waste Management for Covid-19, https://www.who.int/publications-detail/water-sanitation-hygiene-and- waste management-for-covid-19. Accessed 17 May 2020
21. World Health Organization: naming the coronavirus disease (COVID-19) and the virus that causes it (2020), https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-2019)-and-the-virus-that-causes-it
22. World Health Organization: statement on the second meeting of the international health regulations (2005). Emergency committee regarding the outbreak of novel coronavirus (2019-ncov) (2020), https://www.who.int/news-room/detail/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-(2019-ncov)
23. World Health Organization: Director-general's opening remarks at the media briefing on COVID-19—11 March 2020 (2020), https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19—11-march-2020
24. World Health Organization: Coronavirus disease (COVID-2019) Situation reports (2020), https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/
25. World Health Organization: Coronavirus disease 2019 (COVID-19) situation report-24 (2020), https://doi.org/10.1016/j.idm.2020.02.0020
26. World Health Organization: Coronavirus disease 2019 (COVID-19) situation report-28 (2020), https://doi.org/10.1016/j.idm.2020.02.0021
27. World Health Organization: Coronavirus disease 2019 (COVID-19) situation report-31 (2020), https://doi.org/10.1016/j.idm.2020.02.0022
28. World Health Organization: report of the WHO-China Joint Mission on Coronavirus Disease 2019 (2020), https://doi.org/10.1016/j.idm.2020.02.0023
29. World Health Organization: rolling updates in coronavirus disease COVID-19 (2020), https://doi.org/10.1016/j.idm.2020.02.0024
30. World Health Organization: report of the who-China joint mission on coronavirus disease 2019 (covid-19) (2020), https://doi.org/10.1016/j.idm.2020.02.0023
31. World Health Organization: Coronavirus disease 2019 (COVID-19) situation report-25 (2020), https://doi.org/10.1016/j.idm.2020.02.0026
32. World Health Organization: Coronavirus disease 2019 (COVID-19) situation report-114 (2020), https://doi.org/10.1016/j.idm.2020.02.0027