npj | Climate and Atmospheric Science www.nature.com/npjclimatsci

ARTICLE OPEN

Exploring dependence of COVID-19 on environmental factors and spread prediction in India

Hemant Bherwani[1,2]✉, Ankit Gupta[1,2]✉, Saima Anjum[1], Avneesh Anshul[1,2] and Rakesh Kumar[1,2]

COVID-19 has taken the world by storm, with the majority of nations still being challenged by the novel coronavirus. The present work attempts to evaluate the spread of COVID-19 in India using the Susceptible-Exposed-Infectious-Removed (SEIR) model to establish the impact of socio-behavioural aspects, especially social distancing. The impact of environmental factors like temperature and relative humidity (RH) using statistical methods, including Response Surface Methodology (RSM) and Pearson's correlation, is also studied on numbers of COVID-19 cases per day. Here we report the resultant changes of lockdowns-unlocks initiated by the Government of India for COVID-19, as against the scenario of total lockdown. The phased unlocks and crowded gatherings result in an increase in the number of cases and stretch the mitigation timeline of COVID-19 spread, delaying the flattening of the curve. The SEIR model predictions have been fairly validated against the actual cases. The daily spread of COVID-19 cases is also fairly correlated with temperature in Indian cities, as supported by well-established causation of the role of higher temperatures in disrupting the lipid layer of coronavirus, but is greatly undermined by the key factor of social distancing and gets confounded with other multiple unknown co-varying environmental factors. However, the analysis couldn't clearly establish the role of RH in affecting daily COVID-19 cases. Hence, it becomes essential to include environmental parameters into epidemiological models like SEIR and to systematically plan controlled laboratory experiments and modeling studies to draw conclusive inferences, assisting policymakers and stakeholders in formulating comprehensive action plans to alleviate the COVID-19 spread.

npj Climate and Atmospheric Science (2020)3:38 ; https://doi.org/10.1038/s41612-020-00142-x

INTRODUCTION

During the end of December 2019, an outbreak of atypical pneumonia [now being called as coronavirus disease 2019 (COVID-19)] started in Wuhan, China[1–4]. The virus is being considered of zoonotic origin. It is being referred to as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) or in general as novel coronavirus, and the disease-associated is being called COVID-19[3–5]. With the onslaught of SARS-CoV-2 in India, major interventions in epidemic preparedness started. These interventions include, but not limited to, public awareness, deployment of widespread testing facilities, medical institutions preparedness, and surveillance and tracking of individual movement and quarantines of suspected cases[6]. Presently social distancing and regularly washing of hands are some of the best ways to keep this virus at bay[7].

Countries all over the world are challenged with this virus and have declared lockdowns in their various cities and states[8–11]. In India, nationwide lockdown is declared for 21 days starting 25th March 2020[12]. People are instructed to stay indoors unless for emergency services[13]. All tourist visas and e-visas for travellers are suspended till 15th April 2020. Travellers, who returned after 15th February 2020, are quarantined for a minimum of 14 days upon their arrival in India. The researchers estimate that the virus proliferates to more than two persons from every infected person, highlighting the possibility to infect millions. The effectiveness of lockdown is contingent on people avoiding social gatherings and limiting population movement. India reported its first COVID-19 case, originating from China, on 30th January 2020 in Kerala. Government of India (GoI) reported its first COVID-19 death on 10th March 2020 in Karnataka. Till 26th May 2020, India has reported 145,380 confirmed cases with 67,404 recoveries and 4,167 deaths by COVID-19.

Researchers are trying to understand the trend of the SARS-COV-2 movement, to predict the scenarios of new potential cases and to plan effective remediations for their country[14,15]. The role of environmental conditions in the survivability of this virus has also emerged as a potential factor impacting the spread[16,17]. Few studies have suggested that virus should behave similar to most influenza or flu, or also SARS, commonly suggesting its reduced activity in hot and humid conditions[18–20]. However, it is essentially necessary to investigate the impact of local environmental conditions, on the virus spread. The alarming community spread of SARS-CoV-2 majorly impacts the public health, economics and behavioural aspect of the society.

SARS-CoV-2 affects the human body in a similar fashion as other influenza viruses, attacking the respiratory system and is affected by temperature and humidity as well[21]. There have been researches that correlate the decrease in mortality due to COVID-19 to unit degree rise in temperature, which is similar to other respiratory ailment causing viruses[22,23]. The survival of this virus is evidently affected by environmental factors such as temperature and relative humidity (RH) as reinforced by recent works of literature[24,25]. Some of the recent evidences show link between air pollutants and SARS-CoV-2 susceptibility[26–28]. Past evidence can also be taken into consideration where laboratory studies on SARS-CoV and MERS CoV have indicated to be more stable at lower temperatures and dry conditions compounding the help in transitivity through various factors such as increased suspension time, improved conditions for attachment, reduced function of cilia and so on[29,30]. The ability of a MERS-CoV virus to be transmitted via aerosols remains feasible in an airborne state.

[1]CSIR-National Environmental Engineering Research Institute (CSIR-NEERI), Nagpur 440020, India. [2]Academy of Scientific and Innovative Research [AcSIR], Ghaziabad 201002, India. ✉email: h.bherwani@neeri.res.in; a1_gupta@neeri.res.in

It is observed the decrease in capability of virus is significantly influenced by the temperature effect and human to human transmission[31]. The impact of environmental factors is exceedingly related to confirming COVID-19 cases as flu virus spreads rapidly in cold and dry condition and becomes inactive above 30 °C[32]. Hence from the past and recent studies, the importance of factors such as ambient temperature and RH is fairly clear and concrete.

The susceptible-exposed-infectious-recovered (SEIR models) has been reported as a successful tool to understand the pandemic dynamics and to evaluate the impact of environmental and social conditions on the spread of COVID-19[33–36]. Hence, the present work focuses on highlighting the impact of behavioural aspects of the society, and local environmental condition on COVID-19 spread using the SEIR model and statistical tools. The present assessments have been carried out in two folds. First is to predict the number of confirmed cases of COVID-19 for India and the time period for its subsidence. Secondly, to evaluate the dependency of COVID-19 spread on meteorological (may also be referred to as environmental) factors, namely, temperature, and RH.

RESULTS

SEIR model output

SEIR Model is simulated for two cases. Case A, where the model has considered the input data till 3rd April 2020 and Case B was modelled with data up to 9th June 2020. The difference between the two datasets is that India had declared lockdown on 25th March 2020 after understanding the situation that possible community spread of the COVID-19 cases has started and it is to be curbed almost immediately to flatten the curve[37]. Later on, it emerged that the case is true, as a large number of gatherings were happening during the first and second week of March, the spillover of which is seen at the end of March and early April, given the incubation period of novel coronavirus varies from 2 to 14 days[38–40].

For Case A, within the limit of the available datasets, the rise of the number of COVID-19 cases was gradual. The results for Case A are shown below in Table 1 and Fig. 1. Referring to SEIR model equations as supplementary data in Annexure-I, S[t] data is taken from COVID-19 cases in India from supplementary data in Annexure-II and S[t + 1], i.e., for the day plus one, is computed through the model.

From Table 1 and Fig. 1, it can be inferred that if social distancing is strictly implemented to control the community spread, the total number of infected cases is controlled to 10,050 by mid-June and follows a decreasing trend. In Case A, the drop in curve of number of Susceptible-Exposed-Infected-Recovered cases is attributable to hypothesized stringent lockdown in India. The decrease in number of infections is noticed on 19th May. The interstate movement (inflow and outflow) is restricted up to 16th June for Case A which leads to the fall in number of cases. The simulations for Case B are indicated in Fig. 2 and Table 2.

Lapse in lockdown, social distancing and reopening has led to increased number of cases as indicated above. The increase in number of COVID-19 cases is influenced by the interstate movement of migrant workers, massive gathering of people and other major lapses as shown in Table 3. In Case B, the entry of possible community spread by the end of April and start of May has given the rise in the number of cases. The total maximum number of infected cases in Case B (Fig. 2) is 33.64 times of Case A (Fig. 1) for 16th June 2020. Figure 2 shows that the number of infected cases showed an increasing trend with the number of infections exceeding to 550,827 by the end of June.

Validation is performed for 4 weeks to estimate the predicted values with actual values. Long term validation may not be reliable given the ever-changing scenario of India as indicated in Table 3. Figure 3 and Table 4 shows the percentage error assessed within ~1–4% with an average of 2.0%. The number of susceptible, infected and recovered of COVID-19 cases in mid and end weeks of June and first few weeks of July are validated.

Table 1. Simulation result of SEIR model for Case A.

Date	S[t + 1]	E[t + 1]	I[t + 1]	R[t + 1]
31-Jan	7826	31,030	102	0
08-Feb	4781	15,836	206	0
16-Feb	12,807	31,694	510	0
24-Feb	13,125	40,999	765	0
27-Feb	18,014	65,276	612	0
03-Mar	10,079	32,226	514	0
10-Mar	13,885	42,761	703	4
17-Mar	133,124	478,921	1467	20
24-Mar	105,141	318,371	2497	71
31-Mar	112,752	158,692	2795	208
07-Apr	116,052	76,902	4723	774
14-Apr	114,883	59,122	10,513	1105
21-Apr	91,901	46,800	12,814	1182
28-Apr	88,217	39,422	16,818	1565
05-May	84,676	25,541	20,943	2003
12-May	75,347	20,454	27,447	2546
19-May	60,261	15,649	16,938	3552
26-May	48,192	10,894	12,941	4488
02-Jun	38,539	4895	11,591	5358
09-Jun	35,441	4510	10,791	6168
16-Jun	21,256	4313	10,050	6921

The SEIR model is run by considering full lockdown for Case A in which the peak in number of infected people {I[t]} appears to be on 12 May 2020, followed by drop in number of cases which reduces to 10050 by mid of June. Similarly, number of exposed persons {E[t]}, susceptible people {S[t]} also seem to follow a downward trend after a certain time. While number of Removed persons, which include removal by treatment or death, seems to be increasing given the lag in infection and treatment.

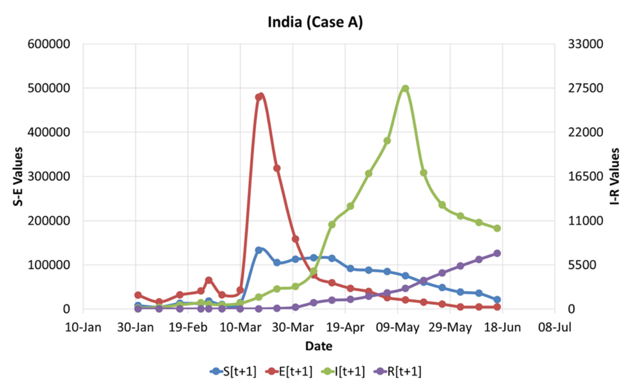

Fig. 1 Trend of rise and fall of cases in each category of the SEIR model (Case A).

Environmental parameters

The state's average data of temperature and RH are considered to study its relationship in the COVID-19 cases per day in two Indian states, namely; Maharashtra and Karnataka. The range of the parameters under study in these states is tabulated in Table 5.

These two states were chosen as they depict 2 different environmental conditions w.r.t. to the parameters under study, i.e., temperature and relative humidity, which is supported by two tailed t-test showing a significant difference in their meteorological condition with respect to temperature and RH as depicted in Table 6 with T = −3.67, $P < 0.001$ and T = −2.20, P-Value = 0.029 at 95% confidence interval, respectively. Table 6 also depicts the significant difference in the meteorological conditions of the cities under study, i.e. Mumbai, Srinagar, Kasaragod and New York, with $p < 0.05$ at CI of 95%.

Table 7 delineates the results of ANOVA for the COVID-19 cases per day in the states of Maharashtra and Karnataka, under the varied temperature and RH conditions. From Table 7, it can be inferred that the main effect of the temperature and RH was significant for Maharashtra with F = 18.46, $P < 0.001$ and F = 122.17, $P < 0.001$, respectively and for Karnataka F = 6.82, $P = 0.01$ and F = 12.83, $P < 0.001$, respectively. However, the squared effect of temperature was found to be negligible for both the states with $P > 0.05$ at CI-95%, the squared effect of the RH was also seen prominently in both the states, i.e. Maharashtra (F = 24.66, $P < 0.001$) and Karnataka (F = 11.5, $P < 0.001$). Also, a significant interaction between the input variables is observed affecting the number of COVID-19 cases per day in Maharashtra with F = 4.25, $P = 0.041$ (CI: 95%). However, Karnataka depicts no significant interactions between the temperature and RH, impacting the number of cases per day with F = 3.54, $P = 0.063$ (CI: 95%). The response surface mapped over the number of COVID-19 cases

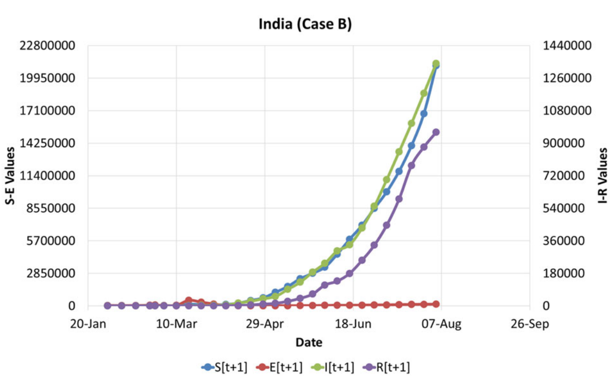

Fig. 2 Trend of rise and fall of cases in each category of the SEIR Model (Case B).

Table 2. Simulation result of SEIR model for Case B.

Date	S[t + 1]	E[t + 1]	I[t + 1]	R[t + 1]
31-Jan	7826	31,030	102	0
08-Feb	4781	15,836	206	0
16-Feb	12,807	31,694	510	0
24-Feb	13,125	40,999	765	0
27-Feb	18,014	65,276	612	0
03-Mar	10,079	32,226	514	0
10-Mar	13,885	42,760	703	3
17-Mar	133,124	478,914	1467	20
24-Mar	105,141	318,362	2497	71
31-Mar	72,752	158,690	2795	208
07-Apr	116,052	7690	4723	774
14-Apr	231,201	8275	15,757	2081
21-Apr	462,545	6067	25,764	5274
28-Apr	693,833	8245	39,104	9804
05-May	1,177,623	13,857	52,362	14,034
12-May	1,695,628	10,687	91,303	25,163
19-May	2,371,285	15,366	130,779	41,499
26-May	2,843,637	20,183	186,203	65,155
02-Jun	3,386,622	39,162	235,796	115,161
09-Jun	4,524,748	44,715	304,036	136,818
16-Jun	5,827,103	48,785	338,149	179,444
23-Jun	7,050,793	62,586	430,173	251,851
30-Jun	8,531,460	76,104	550,827	335,809
07-Jul	9,981,804	83,049	698,347	446,163
14-Jul	11,778,528	105,177	851,846	590,437
21-Jul	14,016,447	117,104	1,009,284	776,302
28-Jul	16,819,736	133,049	1,175,534	878,314
04-Aug	21,024,670	141,704	1,340,913	960,493

For Case B, changes in lockdown scenarios have been considered based on ground situation in the country. The number of infected people {I[t+1]} are increasing as due to relaxed or breach of social distancing norms. SEIR model is run upto 4th August shows the rate of susceptible, infected and exposed people are increasing. A drop in exposed people is seen in the first week of April due to strict lockdown in the end of March i.e., 25th. However, due to breach of social distancing norms, the cases relapse after a certain period of time. The model has been validated for a period of 4 weeks.

Table 3. List of events during lockdown.

Sr No.	Event Location	Date (Year: 2020)	Number of people breaching social distancing norms	References
1	Delhi (Tablighi Jamaat)	31 March	2300	56
2.	Mumbai (Migrants gather at Bandra station)	14 April	2000	41
3	Ahmedabad (Rations distributions)	31 March	Around 17,000 outlets distributing rations	40
4	Kerala (Aattukal Pongala festival)	8 March	1000	57
5	Gujrat (Rushing of Migrant labours)	11 April	1000	58
6	Karnataka (MLA birthday party)	11 April	500	59
7	Karnataka (Belagavi Wedding)	15 March	2000	60
8	Lockdown Phase - 1	25 March	21-day lockdown before 25th March 536 Confirmed cases	50
9	Lockdown Phase - 2	15 April	18-day lockdown slightly lower in cases after strictly lockdown implemented	50
10	Lockdown Phase - 3	4 May	13-day lockdown, zone wise distribution in country	50
11	Lockdown Phase - 4	18 May	14-day lockdown with some relaxation	50
12	Unlock - 1	1 June	30-day unlock	61
13	Unlock - 2	1 July	30-day unlock	61

There were some events occurred during lockdown in India which caused the increase in positive cases in India. In Delhi (Tablighi Jamat) over 2300 people found to be not following social distancing. Over 2000 migrants gathered at Bandra station in Mumbai found not following social distancing also migrant labours were rushing towards their home gathered in huge numbers in Gujarat. In Karnataka few events were organized which caused gathering of people not following social distancing. Lockdown occurred from 25 March, which was divided in 4 phases and from 1 June 2020, unlock phase started.

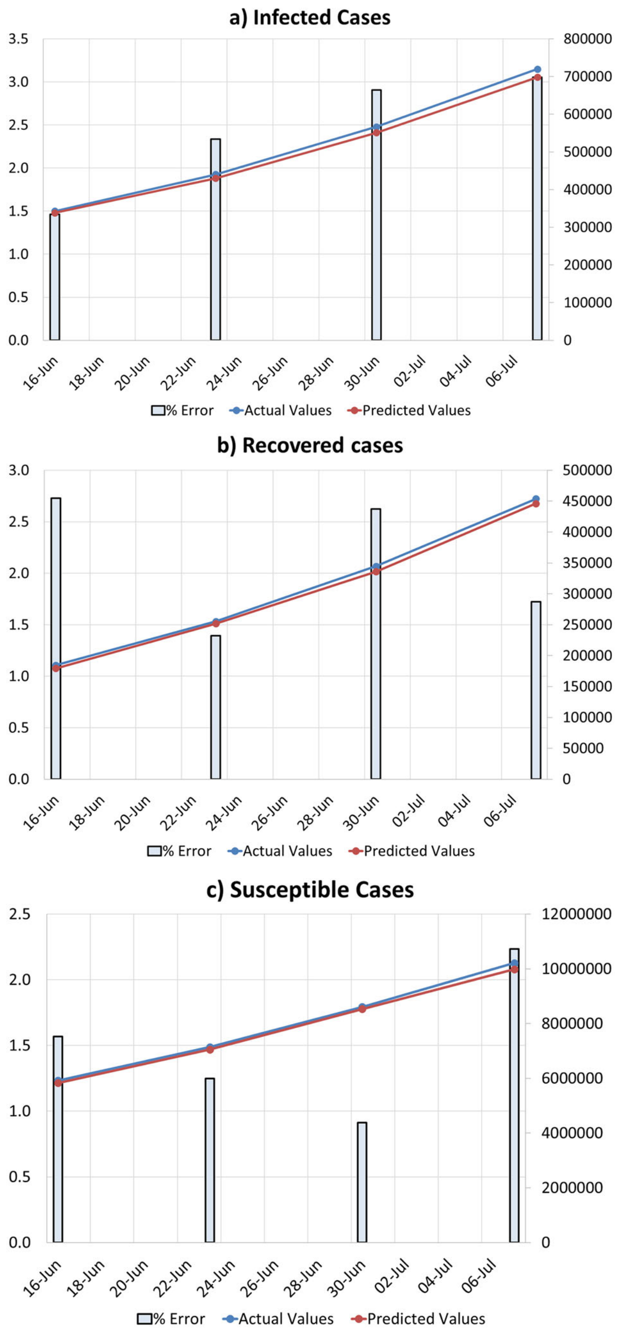

Fig. 3 Percentage error **a** infected cases **b** recovered cases **c** susceptible cases.

per day reflects a good model fit R-Sq of 83.63% and 57.22% for Maharashtra and Karnataka, respectively. The model fit is also confirmed with low standardized residuals, with good fits for both Maharashtra and Karnataka. Figure 4 depicts that modelled inputs data points are in close agreement with normal distribution curve line and residuals are closely distributed at different levels of the independent variable, and hence it can be said that equality of variance exists between varied levels of the response.

The results of dependence of meteorological parameters on COVID-19 cases per day, might depict a feeble causation of the parameters, as the meteorological parameters have been grossly averaged over the spatial-temporal variations across a large areas of these states. Hence, the detailed analysis was performed at city

Table 4. Percentage error between actual and predicted cases.

Infected cases				Recovered cases				Susceptible cases			
Date	Actual values	Predicted values	Error (%)	Date	Actual values	Predicted values	Error (%)	Date	Actual values	Predicted values	Error (%)
16-June	343,091	338,149	1.5	16-June	184,340	179,444	2.7	16-June	5,920,000	5,827,103	1.6
23-June	440,215	430,173	2.3	23-June	255,362	251,851	1.4	23-June	7,140,000	7,050,793	1.2
30-June	566,840	550,827	2.9	30-June	344,624	335,809	2.6	30-June	8,610,000	8,531,460	0.9
07-July	719,665	698,347	3.1	07-July	453,846	446,163	1.7	07-July	10,210,000	9,981,804	2.2

The number of cases for Case B are validated for four weeks i.e., mid of June to start of July 2020. The percentage error between the actual reported cases and modelled cases varies in between 0.9–2.9 for infected, recovered and susceptible cases. Weekly error percentages for all the parameters are shown and are acceptable in nature.

Table 5. Range of variables and coded units.

Environmental factors	Temperature (T in °C)		Relative Humidity (RH in %)	
	Low (Min.)	High (Max.)	Low (Min.)	High (Max.)
Values in coded unit (CU)	−1	+1	−1	+1
Actual values (6 days moving Avg.)				
Indian states				
Maharashtra(MH)	25.000	36.405	18.830	89.530
Karnataka (KR)	24.878	36.250	17.580	77.010
Cities				
Mumbai (MUM)	20.583	30.583	62.333	90.250
Kasaragod (KGD)	25.833	32.250	54.750	88.500
Srinagar (SRG)	9.833	27.667	38.250	75.083
New York (NYK)	5.167	27.500	40.250	76.000

The table depict the statistics of the prevalent environmental conditions in the states of Maharashtra (MH) and Karnataka (KR) along with 4 cities Mumbai (MUM), Kasaragod (KGD), Srinagar (SRG) and New York (NYK). The environmental factors, i.e. Temperature and Relative Humidity were considered and the ranges of these parameters are coded between −1 to +1, individually of each state and city. −1 represents minimum (Min.) value of the parameter and +1, representing maximum (Max.) value of the parameter, all the environmental factor values between them are linear interpolated.

Table 6. Two tailed t-test on the environmental parameters.

Parameters	Temperature		Relative Humidity	
Compared States/Cities	T-value	*P*-value	T-value	*P*-value
KR and MH	−3.67	<0.001	−2.20	0.029
MUM and KGD	−6.80	<0.001	10.33	<0.001
MUM and SNG	16.08	<0.001	31.07	<0.001
MUM and NYK	19.08	<0.001	29.09	<0.001
KGD and SNG	−19.70	<0.001	−14.51	<0.001
KGD and NYK	21.43	<0.001	12.95	<0.001
SNG and NYK	6.50	<0.001	−2.65	0.009

The two tailed t-test was performed to testify the distinct environmental conditions in the select states and city, namely Maharashtra (MH) and Karnataka (KR) along with 4 cities Mumbai (MUM), Kasaragod (KGD), Srinagar (SRG) and New York (NYK). The high T value > $T_{critical}$ and $p < 0.05$ at confidence interval of 95%, leads to rejection of the null hypothesis and depicts the significant difference between the means of the test variables i.e. environmental factor in states/cities.

level, to better understand the reliance of the parameters on COVID-19 cases per day. ANOVA, tabulated in Table 8a, b, has been performed for 3 cities in India and 1 city in USA as these cities are observed to be significantly different from each other in terms of meteorological parameters with $P < 0.05$ (CI: 95%) for both temperature and RH as shown in Table 6. The ANOVA results depict the significance of the dependence of the COVID-19 cases per day on temperature and RH, in all the four cities, with $P < 0.05$ at CI: 95%.

In case of New York, the main effect, and interaction effect of RH and temperature are found to be significantly dependent of the COVID-19 cases per day in the city with low p-values less than 0.05 at 95% CI, except for the squared effects with F = 1.16 and $P =$ 0.318. The response surface mapped shows a negative impact of the rising temperature and positive impact of relative humidity. The positive impact of RH is superseded (as indicated by the RSM model coefficients in coded units) by the negative combined effect of T and RH on the COVID-19 cases per day, as depicted by the RSM model coefficient, with a reasonable model fit with R-Sq = 53.70% and R-Sq(adj) = 51.86%.

Similarly, in case of Kasaragod, the RSM model has depicted a reasonable fit with R-Sq = 47.81% and R-Sq(adj) = 45.44, where the effect of temperature was found to have an significant and shows negative correlation with the COVID-19 cases per day in the city of Kasaragod with $p < 0.05$ (CI: 95%). The squared effect and interaction effect of the temperature and RH was the also found to be significant factors that impact COVID-19 cases with $P < 0.001$. For Kasaragod, the negative effect of T is superseded by the positive impact its squared effect and interaction effect.

For Srinagar, the main effect temperature was the only significant factor with F = 7.18 and $P = 0.009$ at CI of 95%, which is positively correlated with the response with R-Sq = 54.43% and R-Sq(adj) = 52.24%. In case of Mumbai, the main effect of Temperature, RH and squared effect of temperature were found be significant at $P < 0.001$ for all the factors with R-Sq = 24.91% R-Sq(adj) = 21.65%. However, the main effect temperature has shown a negative impact on the COVID-19 cases; its impact is subsided by the positively correlated squared effect of temperature, with higher model coefficient in coded units.

The overall impact of the temperature is observed to have a positive correlation on the COVID-19 cases for all the three Indian Cities i.e., Kasaragod, Srinagar and Mumbai, and is negatively correlated for New York This change is sign of correlation might be attributed to breaches in social distancing in the city of Mumbai[39,41], its extremely high population density[42], where the effect of environmental factors become secondary and less impactful, affecting the COVID-19 cases. Similarly, Srinagar has been witnessing regular lockdowns, owing to continued and ever-going political and public unrest[43] and the spread is contained in through a strict lockdown scenario already existing in Srinagar, mitigating the spread of the virus. However, Kasaragod was taken as a model and a much balanced city, which don't have extreme socio-political conditions as in case of Mumbai and Srinagar. But for Kasaragod, with India Unlocking from 1st June 2020 and considering a span of 42 unlocked days, the social/physical distance have breached at many levels across India (inclunding trans-state movements) and social/physical distance have been hard to maintain and monitor, despite of time-to time government advisories. However, considering the analysis for the city of Kasaragod before the unlock, i.e. upto 28th May, 2020 (Annexure IV), had been implemented, the temperarture was negatively correlated, strenthening and testifying the hypothesis that social distancing is the key and primary parameter impacting the spread and rise in COVID19 cases and environmental/ meteorological factors are undermined by social distancing factor and show a feeble impact. These impacts of meteorological parameters might also gets confounded with multiple environmental factors like bio-aerosols concentrations, air pollutions, particulates etc. and many other unknown co-varying factors. Figure 5 depicting the residual plot of RSM model confirms that the meteorological parameters are in agreement with the normal distribution line and residual being evenly distributed at all levels of input parameters and hence the assumption of the equality of variance holds true.

In addition to the RSM exploration of the above state and city cases dependency on temperature and RH, linear statistical correlation is also explored in order to understand if there is any significant difference between the results. Indian cities considered for analysis are Mumbai, Srinagar and Kasaragod. The Pearson's correlation matrix for each city is shown in Table 9. New York, which has seen unprecedented growth in COVID-19 cases, is also correlated for temperature and RH and validates similar results as Indian cities. The linear correlation is assessed for above-stated states as well.

Table 7. Summary of ANOVA results for COVID-19 cases per day in states of India: (a) Maharashtra (b) Karnataka and (c) T and P values for Maharashtra and Karnataka.

Source	DF	Seq SS	Adj SS	Adj MS	F	*P*
(a) Maharashtra						
Regression	5	28.2979	28.2979	5.65958	121.55	<0.001
Linear	2	26.6721	5.7102	2.85511	61.32	<0.001
X1	1	0.0002	0.8593	0.85933	18.46	<0.001
X2	1	26.6718	5.6886	5.68862	122.17	<0.001
Square	2	1.4279	1.2234	0.61169	13.14	<0.001
X1*X1	1	0.0405	0.0286	0.02863	0.61	0.435
X2*X2	1	1.3874	1.1481	1.14813	24.66	<0.001
Interaction	1	0.1979	0.1979	0.19794	4.25	0.041
X1*X2	1	0.1979	0.1979	0.19794	4.25	0.041
Residual error	119	5.541	5.541	0.04656		
Total	124	33.8389				
S = 0.215785 PRESS = 6.11034 R-Sq = 83.63% R-Sq(adj) = 82.94%						
(b) Karnataka						
Regression	5	12.6119	12.6119	2.52238	28.09	<0.001
Linear	2	9.4486	2.9548	1.47741	16.45	<0.001
X1	1	7.0128	0.612	0.61202	6.82	0.01
X2	1	2.4358	1.1516	1.15155	12.83	0.001
Square	2	2.8456	1.4256	0.71278	7.94	0.001
X1*X1	1	0.8792	0.1598	0.15981	1.78	0.185
X2*X2	1	1.9664	1.0323	1.03227	11.5	0.001
Interaction	1	0.3177	0.3177	0.31771	3.54	0.063
X1*X2	1	0.3177	0.3177	0.31771	3.54	0.063
Residual error	105	9.4279	9.4279	0.08979		
Total	110	22.0398				

S = 0.299649 PRESS = 10.3891 R-Sq = 57.22% R-Sq(adj) = 55.19%

(c) Terms	**Maharashtra**				**Karnataka**			
	Coef	SE Coef	T	*P*	Coef	SE Coef	T	*P*
Constant	−0.70846	0.05398	−13.125	<0.001	−0.816	0.08533	−9.563	<0.001
X1	0.43807	0.10197	4.296	<0.001	0.9819	0.37611	2.611	0.01
X2	0.69336	0.06273	11.053	<0.001	1.3233	0.3695	3.581	0.001
X1*X1	0.09904	0.12631	0.784	0.435	0.6901	0.51729	1.334	0.185
X2*X2	0.82839	0.16683	4.966	<0.001	1.3987	0.41251	3.391	0.001
X1*X2	0.5833	0.28291	2.062	0.041	1.6627	0.8839	1.881	0.063

Note: X1: Temperature, (°C); X2: Relative Humidity, %; both in coded units
The Analysis of Variance (ANOVA) helps understand the significance of the input parameters, i.e., Temperature and Relative Humidity on the response variable, i.e., COVID-19 cases per day. DF represents the degree of freedom from each source term in the model, Seq SS represents sequential sum of squares between the input variables (factor) and within the group (error). Adj MS represents Adjusted Mean Square, which is a ratio of SS/DF and F is calculated by dividing factor MS by MS of the residual error. The term Seq. SS, Adj. SS and Adj. MS, explain the model variability explained by each term of the response surface methodlogy (RSM) model.The significance of the each terms in the full quadratic RSM model is interpreted by F, T and P values, where high values of T and F (against a critical values T and F, computed from T and F tables for two-tailed test) and $p < 0.05$ at 95% confidence interval denotes the significant impact of the variable on the response. The R-sq (regression coefficients) and Adjusted R-sq (within 10% of the R-sq value), explains good model adequacy of the RSM model to map the responses (outputs) against the inputs variable. S represents the how far the actual data fall from the RSM fitted values.

From Table 9, Kasargod and New York show a negative correlation with temperature, while Mumbai and Srinagar show a positive correlation. It is interesting to see that RH for all states and cities show positive correlation except for Srinagar. The strength of these correlations varied from strong to weak relations, as indicated in Table 9. It has been well established that with the increase in temperature, the RH should drop given that the holding capacity of air will increase. It is also interesting to note that for Kasaragod, the pearson's correlation reflect a negative correlation for temperature in contract to a positive correlation depicted by RSM model. This change is correlations is attributed to the insufficiency of the linear correlation in pearson's models compared to quadratic RSM model, which is capable to capture the non-linearity of the dataset to a great extent. Hence, a cautious approach should be adopted using linear models while drawing conclusive inferences. Mumbai shows a weak positive correlation with both temperature and RH, while Temperature-RH relation is also positive. This might be due to the fact that as Mumbai is a coastal city and a densely packed one with respect to population and land use. Another reason could be that Mumbai being a metropolitan city and depends on public commutes, making the implementation of social distance practically very difficult. In such cases, the dominance of other factors, for disease spread, on environmental factors cannot be ignored and need

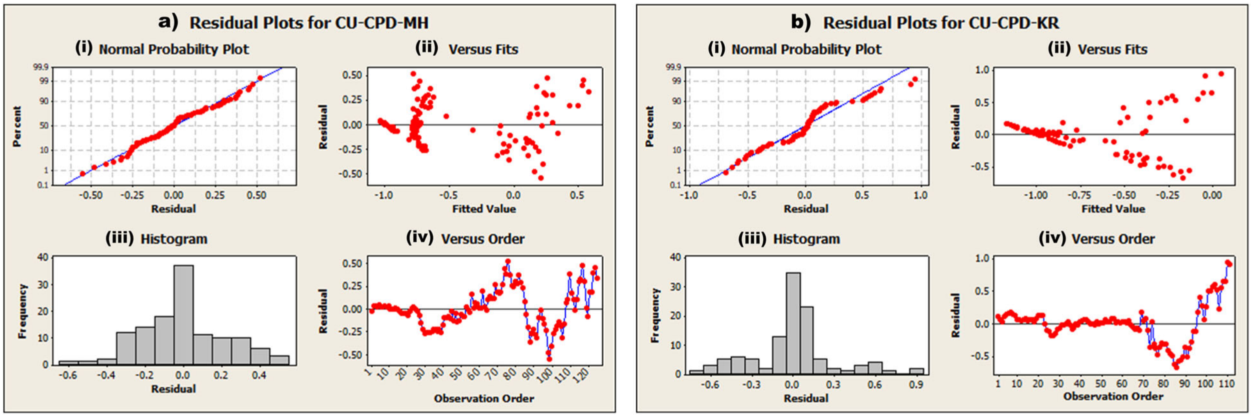

Fig. 4 Residual plots for states: **a** Maharashtra **b** Karnataka.

separate detailed research. Further, this also confirms that only environmental factors may not be able to mitigate this particular strain of novel coronavirus and other behavioural attributes are bound to pay a major role in its alleviation.

In the present work, single season short-term meteorological condition is studied given the availability of data and evolution of COVID-19 situation. RSM tools is used to understand the intricacies of multiple parameters on evolution of COVID-19 in the select cities. It is suggested that when considering the long term meteorological data series, it becomes essential to use the de-trended anomaly data, to establish an objective correlation between COVID-19 and meteorological parameters. Hence, future scope of work may include a de-trended anomaly perspective applied to all variables under consideration to remove longer term trends in the data and meteorological variables have to be calculated against the long-term average of the respective 6-day periods under consideration (taking thus into account the longer term background climatology), which will take care of the seasonal trend in temperature and humidity data.

DISCUSSION

In current time, when every COVID-19 affected nation is making efforts to mitigate and alleviate the spread of this virus, it becomes essential to study the correlations of the cases of COVID-19 reported per day with respect to behavioural and environmental attributes. India, similar to other countries, is making efforts to curb the spread by strictly monitoring the lockdowns and ensuring social distancing apart from the quest for targeted medical solutions. In a country like India, where the population exceeds 1 billion, social distancing remains a challenging but potential and practically implementable solution. In the current research, an effort is made to establish the impact of social distancing to subside the rising cases in India through simulating multiple cases using the SEIR model. The model concluded that breach of social distancing by people engagement in crowded gatherings, with the onset of possible community spread, might result in extending the time to control the spread with a total number of infected cases exceeding 1,350,000 and still increasing at a rapid rate. When compared to the normal case of strict compliance of lockdown, at present the cases are 33.64 times higher and increasing. The above scenario is considering number of gatherings which occurred in end of April and during May, June; however, if there are continued gatherings like this, the outcome can be disastrous in terms of number of people infected and respective deaths. Hence, the decision of GoI of nationwide lockdown was very timely to control the spread of the virus and continued efforts are needed towards that so that the pandemic curve keeps rising at a rate which does not become a burden on medical facilities. The finding of SEIR helps to conclude that social distancing is one of the best tools available with governments right now to control the spread of this disease. The SEIR model is validated for a period of four weeks and gives a better understanding of the dynamics of virus spread are changing.

Also, the impact of environmental conditions is found to be significantly affecting the spread of COVID-19, established through statistical models like RSM and correlation matrices. The COVID-19 mean incubation periods of 6 days have been considered for establishing the correlation with environmental factors. The effect of meteorological parameter on daily COVID-19 cases, although show a good RSM model fit with R-Sq 83.63% for Maharashtra and a reasonable fit for Karnataka at 57.22%, doesn't implicitly shows the relationships on COVID-19 cases per day as the environmental parameters gets grossly averaged over the states and doesn't essentially capture the essence of the relation between the meteorological parameters with COVID-19 cases. Hence, the analysis at the city level is desirable to draft more conclusive inferences.

As temperature rise has been reported to mitigate the spread of viruses, similar trends are observed in New York, USA under the study for SARS-CoV-2. However, the temperature is observed to have a positive correlation on the COVID-19 cases for all the Indian cities, i.e., Srinagar, Mumbai and Kasaragod, which is inverse of the relationship obtained for New York. The relation for Mumbai can be attributed to breaches in social distancing in the city of Mumbai and its extremely high population density, where the effect of environmental factors become secondary and less impactful, affecting the COVID-19 cases. Similarly, Srinagar has been witnessing regular lockdowns, owing to continued and ever-going political-public unrest and the spread is contained in through a strict lockdown scenario, unlike New York, mitigating the spread of the virus.

Also, in the case of Kasaragod with unlock implementation, the effect of temperature and humidity is undermined and is dominated by the social distance factor, thereby showing a positive correlation, which was negative till the lockdown periods. Learning for the analysis of Indian cities and comparison with a baseline city as New York, it can be inferred that although the meteorological parameters have found be correlate with the COVID-19 cases but are greatly undermined by the social distancing factor, either breached or very strictly imposed, as seen in the all the Indian cities. Hence, in such cities, social distancing is the key and primary parameter impacting the spread and rise in COVID-19 cases, undermining the environmental/meteorological factors, which might get confounded with other co-varying environmental factors.

Researches have suggested that hot climate should reduce the spread of COVID-19 owing the well-established causation of the role

Table 8. ANOVA for COVID-19 cases per day in New York, Kasaragod, Srinagar and Mumbai.

(i) Source	DF	Seq SS	Adj SS	Adj MS	F	*P*	DF	Seq SS	Adj SS	Adj MS	F	*P*
	a) New York						**c) Srinagar**					
Regression	5	21.0059	21.0059	4.2012	29.23	<0.001	5	6.3503	6.35034	1.27007	24.84	<0.001
Linear	2	18.2906	18.2801	9.14	63.59	<0.001	2	5.0432	3.2961	1.64805	32.24	<0.001
X1	1	11.9274	13.881	13.881	96.58	<0.001	1	5.0432	0.36702	0.36702	7.18	0.009
X2	1	6.3632	2.4437	2.4437	17	<0.001	1	0	0.00688	0.00688	0.13	0.714
Square	2	1.1471	0.3321	0.1661	1.16	0.318	2	1.3003	0.02696	0.01348	0.26	0.769
X1*X1	1	0.2846	0.0841	0.0841	0.58	0.446	1	1.147	0.02307	0.02307	0.45	0.503
X2*X2	1	0.8625	0.2592	0.2592	1.8	0.182	1	0.1533	0.00395	0.00395	0.08	0.782
Interaction	1	1.5683	1.5683	1.5683	10.91	0.001	1	0.0068	0.00678	0.00678	0.13	0.717
X1*X2	1	1.5683	1.5683	1.5683	10.91	0.001	1	0.0068	0.00678	0.00678	0.13	0.717
Residual Error	126	18.11	18.11	0.1437			104	5.3168	5.31679	0.05112		
Total	131	39.116					109	11.6671				
	S = 0.379118 PRESS = 19.8260 R-Sq = 53.70% R-Sq(adj) = 51.86%						**S = 0.226104 PRESS = 5.88259 R-Sq = 54.43% R-Sq(adj) = 52.24%**					
	b) Kasaragod						**d) Mumbai**					
Regression	5	1.52228	1.52228	0.304455	2.97	0.018	5	6.4983	6.4983	1.2997	7.63	<0.001
Linear	2	0.96293	1.12665	0.563323	5.5	0.006	2	2.0294	3.8953	1.9477	11.44	<0.001
X1	1	0.91382	0.8544	0.854401	8.35	0.005	1	0.0029	3.6813	3.6813	21.62	<0.001
X2	1	0.04911	0.00219	0.002187	0.02	0.884	1	2.0266	3.242	3.242	19.04	<0.001
Square	2	0.29411	0.55667	0.278335	2.72	0.073	2	3.8473	2.9026	1.4513	8.52	<0.001
X1*X1	1	0.28967	0.5446	0.544596	5.32	0.024	1	3.816	2.4865	2.4865	14.6	<0.001
X2*X2	1	0.00444	0.12586	0.125857	1.23	0.272	1	0.0312	0.2367	0.2367	1.39	0.241
Interaction	1	0.26524	0.26524	0.265236	2.59	0.112	1	0.6217	0.6217	0.6217	3.65	0.059
X1*X2	1	0.26524	0.26524	0.265236	2.59	0.112	1	0.6217	0.6217	0.6217	3.65	0.059
Residual error	65	6.65327	6.65327	0.102358			115	19.5837	19.5837	0.1703		
Total	70	8.17554					120	26.082				
	S = 0.197500 PRESS = 6.53204 R-Sq = 47.81% R-Sq(adj) = 45.44%						**S = 0.412665 PRESS = 21.6522 R-Sq = 24.91% R-Sq(adj) = 21.65%**					

	New York				Kasaragod			
(ii) Terms	Coef	SE Coef	T	*P*	Coef	SE Coef	T	*P*
Constant	−0.3152	0.07739	−4.073	<0.001	−0.8546	0.03959	−21.586	<0.001
X1	−0.5623	0.05722	−9.827	<0.001	−0.3645	0.18301	−1.992	0.049
X2	0.3747	0.09088	4.123	<0.001	−0.2049	0.16127	−1.27	0.207
X1*X1	−0.1044	0.13654	−0.765	0.446	1.5416	0.25872	5.959	<0.001
X2*X2	−0.2214	0.16486	−1.343	0.182	1.5716	0.22252	7.062	<0.001
X1*X2	−0.44	0.13321	−3.303	0.001	3.0761	0.3548	8.67	<0.001

	Srinagar				Mumbai			
	Coef	SE Coef	T	*P*	Coef	SE Coef	T	*P*
Constant	−0.84597	0.04004	−21.13	<0.001	−0.4926	0.08659	−5.689	<0.001
X1	0.3265	0.12186	2.679	0.009	−0.9608	0.20665	−4.649	<0.001
X2	−0.05388	0.14684	−0.367	0.714	0.7448	0.1707	4.363	<0.001
X1*X1	0.14775	0.21996	0.672	0.503	1.3213	0.34578	3.821	<0.001
X2*X2	0.10358	0.37282	0.278	0.782	0.3022	0.25631	1.179	0.241
X1*X2	−0.18598	0.51086	−0.364	0.717	−1.0198	0.53373	−1.911	0.059

The Analysis of Variance (ANOVA) depicts the significance of the linear, squared and interaction terms of Temperature and Relative Humidity on the response, i.e., COVID-19 cases per day. Here, DF represents the degree of freedom, Seq SS represents sequential sum of squares, Adj SS is Adjusted sum of squares and Adj. MS represents Adjusted Mean Square and all these terms, help determine the variability capture by each model terms in the RSM quadratic model. The RSM model fitness is captured by R-sq value and the Adj. R-sq values within 10% of the R-sq values confirm the model adequacy. High value of $T_{statistics} > T_{criticial}$ and $p < 0.05$ at confidence interval of 95% denotes significant impact of the model term on the response.

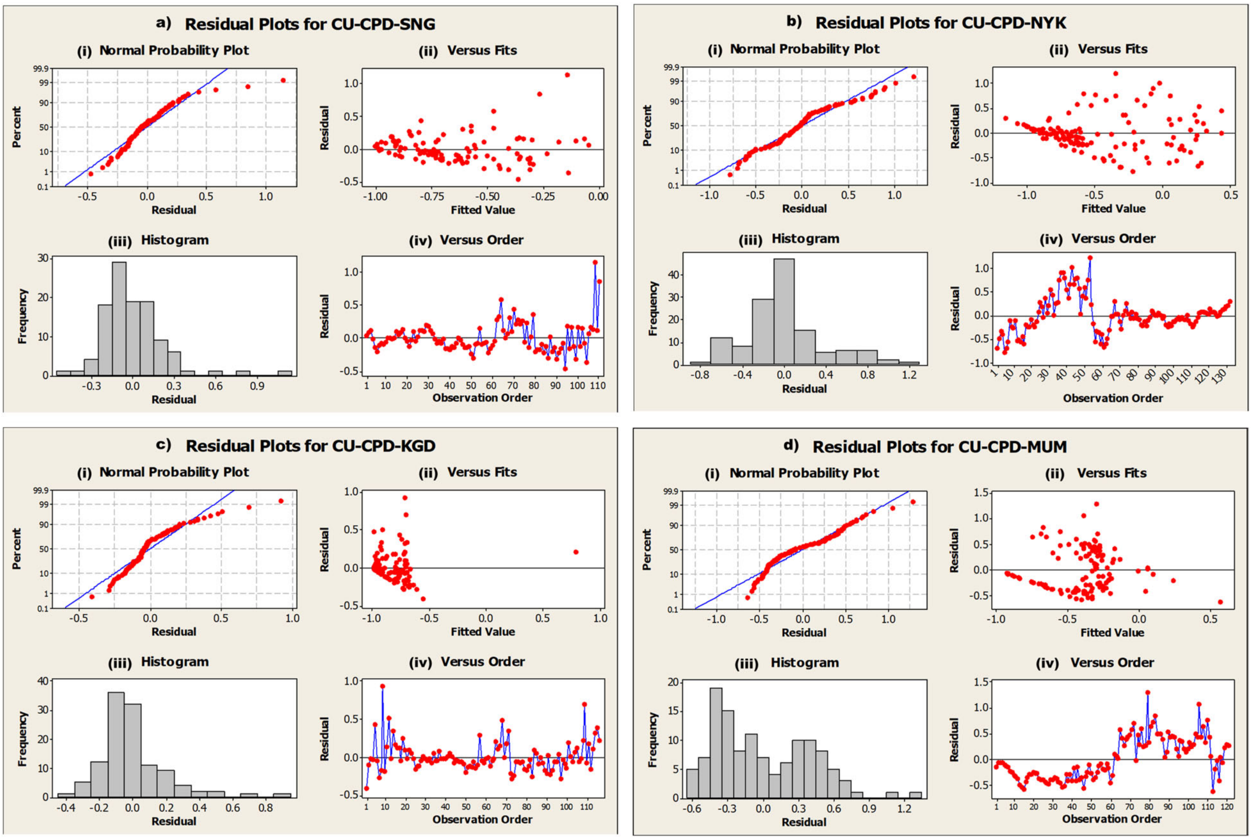

Fig. 5 Residual plot for cities **a** Srinagar **b** New York **c** Kasaragod and **d** Mumbai.

of higher temperatures within disrupting the lipid layer of coronavirus and viruses of similar nature, while the relation with RH has never been clear and established, till date. From the current research, it can be concluded that environmental factors impact on the COVID-19 can't be negated, however, the clear significance of these parameters would require more controlled laboratory experiments and modeling study, including environmental parameters into epidemiological models like SEIR, to objectively conclude its impact on the COVID-19 transmission and mitigation.

METHODS

SEIR model

The prediction of COVID-19 cases is based on the SEIR (Susceptible-Exposed-Infectious-Removed) model, which has been used to simulate two practical implementation conditions related to societal behavioural issue. Recently there have been advancements in terms of modified SEIR model and it has been upgraded to global model by incorporating potentially infected and infected but not isolated cases, however, the current analysis is restricted to globally accepted modified SEIR model, given the robust data available for this model. The first case is where the lockdown is followed diligently and the other being with community breaches, partial lockdowns, and failure of social distancing implementation protocols. Both of these scenarios actually happened in India and have been considered accordingly.

The SEIR model for the spread of Infectious disease is simulated, similar to severe acute respiratory syndrome[44]. Immunity, infected, exposure probability and recovery/removal are the compartments of the model[45]. The modified SEIR model with its codes and governing equations are stimulated for India having maximum number of reported COVID- 19 cases for understanding efficacy of measures in the current time and future. Governing equations along with the data for SEIR model are given as supplementary data in Annexure I. World Health Organization (WHO), India reported the numerals of COVID-19 pandemic spread in India on a day-to-day basis. The primary sources of the data are WHO situation reports for India. Daily migration index has been computed on the basis of mass movements of people from one city to another through various mode of transportation like air, road and rail. SEIR equations are modelled using python program v2.7.5 (open source) to predict the forthcoming Susceptible [S], Exposed [E], Infected [I], Recovery [R] cases for India[46]. The parameters considered with respect to the model are given as follows:

SEIR model parameters

Table 10 below gives the SEIR model parameters. Detailed equations and corresponding values are given as supplementary data in Annexure-I.

The model is run for two scenarios, one with complete lockdown throughout the duration (Case A) and the other with phased opening a breach in lockdown (Case B). The duration of model run is taken from 31st January until 4th August. The data for Case A is considered from 31st January to 3th April 2020 to ensure no breach in lockdown and it simulated further as a base scenario. Various breaches are observed 31st of March 2020 as given in Table 3 and are incorporated in Case B. The transmission rate is calculated on a weekly basis using methodologies adopted from the literature[47,48] for the given data sets for each case as stated above. The coding for SEIR model runs for each week by solving equations as given in supplementary data in Annexure-I. The simulation for both cases A and B are done on the basis of the single run as population and social distancing has proven to be one of the most important parameters in controlling the spread of this virus and is validated for a specific time period[48].

Breach in lockdown is reported by end of March 2020. Mass gathering of people at Mosque by Tablighi jamaat approximately 2300 people gathered from numerous neighbouring countries. Another vast event in Ahmedabad, thousands of people rushed to ration shops while distributing the ration approximately from 17,000 outlets. Another event like Unlock 1 and 2 has also gave rise to number of cases. Such events led to the massive spread of this infectious disease. Further, due to massive economic loss, phased relaxation is given in lockdown, which

Table 9. Pearson's Correlation Matrix for meteorological parameters on COVID-19 cases per day.

	T-KR	RH-KR	CPD-KR
Karnataka			
T-KR	1.00		
RH-KR	−0.92	1.00	
CPD-KR	−0.56	0.65	1.00
Maharashtra	**T-MH**	**RH-MH**	**CPD-MH**
T-MH	1.00		
RH-MH	−0.48	1.00	
CPD-MH	0.003	0.78	1.00
New York	**T-NYK**	**RH-NYK**	**CPD-NYK**
T-NYK	1.00		
RH-NYK	0.04	1.00	
CPD-NYK	−0.55	0.38	1.00
Srinagar	**T-SNG**	**RH-SNG**	**CPD-SNG**
T-SNG	1.00		
RH-SNG	−0.83	1.00	
CPD-SNG	0.66	−0.54	1.00
Kasaragod	**T-KGD**	**RH-KGD**	**CPD-KGD**
T-KGD	1.00		
RH-KGD	−0.94	1.00	
CPD-KGD	−0.23	0.30	1.00
Mumbai	**T-MUM**	**RH-MUM**	**CPD-MUM**
T-MUM	1.00		
RH-MUM	0.60	1.00	
CPD-MUM	0.01	0.23	1.00

Pearson's correlation i.e. linear correlation is evaluated in the form of lower triangular matrix in for Temperature (T), Relative and reported COVID-19 Cases per Day (CPD) for Indian States namely Karnataka (KR) and Maharashtra (MH) and Cities namely New York (NYK), Srinagar (SNG), Kasaragod (KGD) and Mumbai (MUM). Positive and negative signs indicate direct and inverse correlation of the parameters respectively while the number indicates strength of correlation, zero being no correlation and one showing complete correlation.

added to number of cases. Some of the relevant and important breaches related to lockdown and change in governing policies are highlighted in Table 3.

The predicted values is determined from the input data is shown in supplementary Table. 1. Dataset of incubation rate and recovery or death rate are used in the equations to estimate the upcoming cases. The data on confirmed cases for India which have been used in the above model is shown as supplementary data in supplementary file.

Statistical analysis

Two Indian States have been selected for studying the impact of the meteorological parameters on COVID-19 cases per day, namely, Maharashtra and Karnataka. The reasons for the selection of these states was, firstly the different environmental conditions w.r.t temperature and RH in these 2 states and secondly, the availability of consistent data on meteorological conditions and reporting COVID-19 cases. The select states have also witnessed breaches in social distancing norms[41,49,50]; hence it becomes important to study the impact of environmental conditions in such states to rationally weigh the impact of environmental factors and social distancing factor. The raw data on meteorological parameters for the states is calculated by averaging all the stations data, considering 24 h data including day and night values, available from CPCB (Central Pollution Control Board), India and the station-wise data is appended as supplementary information in Annexure-V.

Averaging meteorological conditions over the entire spatial domain of the state would be a gross representation of the environmental factors studied for dependence on confirmed COVID-19 cases per day. Hence, it becomes essentially important to explore the dependence of meteorological conditions at the cities level to overcome the assumptions made during the analysis at state level. Therefore, four additional cities were selected for in-depth analysis to establish the dependence of the environmental factors on COVID-19 cases per day. Three Indian Cities i.e., Mumbai, Kasaragod and Srinagar and 1 city abroad i.e., New York was selected for the analysis. The selection was primarily based on the different environmental conditions in the select Indian cities with the baseline city of New York, USA. These select cities had also witnessed one of the most rapid rises in COVID-19 cases per day i.e. Mumbai in India and New York in the USA or one of the first states to report the COVID*19 cases i.e. Kasaragod in India; Srinagar was chosen for a baseline comparison of an Indian city of similar environmental conditions to New York in the USA.

The hypotheses testing is performed using 2-sample t-test to establish the mean difference in the environmental conditions, w.r.t. temperature and RH of the selected cities is equal to 0 (vs. not = 0) at a 95% confidence interval. The analyses would help establish the distinct environmental conditions across the studied states and cities, aiding the selection and analyses of significant different states and cities, in terms of meteorological conditions, on the daily COVID-19 cases.

While studying the relation of environmental factors with COVID-19, RH and temperature are considered as the base parameters. The daily COVID-19 cases in various cities of India are correlated with above parameters to understand the relation. The COVID-19 mean incubation periods of 6 days have been considered for establishing the correlation with environmental factors[51,52]. Therefore, the average temperature and RH for six-day have been considered with one-day onset for the time taken for COVID-19 testing and reporting. Hence, six-day time series moving average of environmental conditions i.e. temperature and RH have been correlated with the results of the COVID-19 case reported on the seventh-day and the same have been consistently used in all the subsequent analysis to study the SEIR and impact of meteorological conditions on COVID-19 cases.

RSM is used as an effective statistical tool to analyse the linear and squared effect of meteorological parameters and understanding the interaction between these meteorological parameters affecting the daily COVID-19 cases. Hence, in addition to simple pearson's correlation (depicting linear relation), Response Surface Methodology (RSM) is used to establish an empirical association between the meteorological parameters, which included Temperature and RH, over the confirmed/reported COVID-19 cases per day in India. The raw data related to statistical modelling is given as supplementary data in Annexure III. RSM compared to the classical methods has been advantageous due to it reduce dataset requirement to map the responses over the complete domain of the inputs variable. The quadratic model in RSM also helps minimise the crude linear approximations of the inputs parameters on the response function, thus aids capturing the non-linearity of the meteorological processes and thus variables (being a consequence of those processes) through a full quadratic model as depicted by the RSM model Eq. (1).

A response surface is used to map the entire cases of the particular region using a full quadratic function. The association of the parameters can be depicted as a second-order polynomial equation.

$$R = \beta_0 + \sum_{i=1}^{i=n} \beta_i x_i + \sum_{i=1}^{i=n} \beta_{ii} x_i^2 + \sum\sum_{i<j} \beta_{ij} x_i x_j, \tag{1}$$

where R is the response variable, i.e. COVID-19 cases per day; β_0 a constant; β_i the linear coefficient; β_{ii} the squared coefficient; and β_{ij} the product-coefficient, x_i represent the linear or main effect of the input variable, x_i^2 represent the squared effect of the input variable, $x_i x_j$ represent the interaction term of the input variable and n is the number of factors[53,54]. The model coefficients (i.e., β_0,β_i,β_{ii} and β_{ij}) are computed using the least square method by minimizing the sum of squared residuals/error. The significance of the respective parameters is established by Analysis of Variance (ANOVA), which can interpret in terms of P values[55]. The lower the P-value, the higher is the significance of the corresponding parameters. For all the analysis, the p-value less than 0.05 at confidence interval of 95% is considered as a sufficient condition to reject the null hypothesis and represent the significance of the parameter. A good model fit can be affirmed with high R^2 (also referred as R-Sq) value[53,54], which help check the competency of the model. All the inputs and response variable have been modelled in coded units' i.e. highest parameter values as +1 and lowest parameter values as −1 and

Table 10. SEIR model parameters.

Parameter	Brief data definition	Source/Reference
Suspected people (S[t])	The total number of suspected people in India are taken from WHO highlights showing a number of suspected samples for testing out of which some have positive results (SARS-CoV-2). A suspected individual contract the disease when interacts with the infected agents. Number of suspected people (S[t]) is equal to the total number of screened people	Ministry of Health and Family Welfare (MoHFW)[62,63]
Population (N)	Total population data is as per the census report. The population of India is about 133.92 crores	Census of India[64]
Probability of transmission (β)	Probability of transmission changes over time with number of infections and is used to derive β_1 and β_2 [S]→[I] (β_1) susceptible to infected [S]→[E] (β_2) susceptible to exposed	[65]
Rate of transmission (β_1 *and* β_2)	The mortality rate of COVID-19 at about 3.434% in China and about 1.1% outside China. Rate of transmission is 1.5 - 4.5 for SARS-CoV-2, which is more than SARS rate of transmission. $\beta_1 = k_1 \times \beta$ $\beta_2 = k_2 \times \beta$ Where, k_1 = Probability of symptomatic infection (3 days) k_2 = Probability of asymptomatic infection (15 days)	[46,66–68]
Incubation rate (σ)	2 to 14 days has been publicly reported for SARS- CoV-2. The period between the exhibition to an infection and the appearance of first symptom due to COVID-19.	Publicly available report of confirmed cases[69–71]
Average rate of recovery or death (γ)	The average duration between the confirmation of detection to the time of removal i.e., death or recovered for India is taken as 14 days γ = 1/Average recovery time = 1/14 = 0.07	[70,72]
Number of contacts per person ($r[t]$)	The number of contacts per person can be considered as the number of exposed people during inflow and outflow of the population it is estimated depending upon population, age group, household size, etc	The daily average number of contacts per person[49,66,73]
Exposed People (E[t])	The number of exposed people can be calculated by the inflow rate of people coming from other places and the outflow rate of people flowing out of the country	The latent population of country[74,75]
Infected People (I[t])	The total number of infected people in the country from January 2020 data is collected from online sources till June 9th 2020	WHO pandemic report

Number of suspected people (S[t]) is equal to the total number of screened people. Total population (N) is as per the Celsius report. Probability of transmission (β) changes over time with number of infections and is used to derive β_1 and β_2, where β_1 is suspected to infected and β_2 is suspected to exposed. Rate of transmission is 1.5–4.5 for SARS-CoV-2, which is more than SARS rate of transmission. Incubation rate (σ2 to 14 days has been publicly reported for SARS- CoV-2. Average rate of recovery or death (γ), death or recovered for India is taken as 14 days. Number of contacts per person ($r[t]$) the number of contacts per person can be considered as the number of exposed people. Exposed People (E[t]), The number of exposed people can be calculated by the inflow rate of people coming from other places and the outflow rate of people flowing out of the country. Infected People (I[t]), the total number of infected people in the country from January 2020.

the remaining parameters have been linearly interpolated between −1 to +1. The analysis in coded unit helps understand and interpret the significance of the input parameters using the coefficient of the RSM model mapped within the limit of the dataset i.e. a particular city/state. Two- sample t-test and Response surface modelling (RSM) is carried using statistical software MINITAB 14[53].

DATA AVAILABILITY

The data used in the manuscript is publicly available:

• Office of the Registrar General & Census Commissioner, Ministry of Home Affairs, Government of India, Migration Index for India. Retrieved from https://censusindia.gov.in/Census_And_You/migrations.aspx (2nd April 2020).

• Novel Coronavirus Disease (COVID-19) India Situation Update Report 7, WHO Retrieved from https://www.who.int/docs/default-source/wrindia/situation-report/india-situation-report-7.pdf?sfvrsn=cf4a7312_2 (14 March 2020).

• Novel Coronavirus Disease (COVID-19) India Situation Update Report 8, WHO Retrieved from https://www.who.int/docs/default-source/wrindia/situation-report/india-situation-report-8bc9aca340f91408b9efbedb3917565fc.pdf?sfvrsn=5e0b8a43_2 (22 March 2020).

• Novel Coronavirus Disease (COVID-19) India Situation Update Report 9, WHO Retrieved from https://www.who.int/docs/default-source/wrindia/situation-report/india-situation-report-9.pdf?sfvrsn=c883d0c2_2 (28 March 2020).

• Novel Coronavirus Disease (COVID-19) India Situation Update Report 10, WHO Retrieved from https://www.who.int/docs/default-source/wrindia/situation-report/india-situation-report-10.pdf?sfvrsn=48298da5_2 (5 April 2020).

• Jason Oke and Carl Heneghan. Global Covid-19 Case Fatality Rates, CEBM Research. Retrieved from https://www.cebm.net/covid-19/global-covid-19-case-fatality-rates/ (March 17, 2020).

• COVID-19 CORONAVIRUS PANDEMIC, Confirmed Cases and Deaths by Country, Territory, or Conveyance. Worldometer. Retrieved from https://www.worldometers.info/coronavirus/#countries.

• 2020 coronavirus pandemic in India, Wikipedia. Retrieved from https://en.wikipedia.org/wiki/2020_coronavirus_pandemic_in_India (March 2020).

• Coronavirus (COVID-19) map. Google. Retrieved from https://www.google.com/covid19-map/ (3 April 2020).

• Office of the Registrar General & Census Commissioner, Ministry of Home Affairs, Government of India, Population data - Census of India: Retrieved from https://censusindia.gov.in/Census_And_You/migrations.aspx (31 March 2020).

• Thomas Stephens. Coronavirus: Head-shaking doctors and non-hand-shaking workers. Retrieved from: https://www.swissinfo.ch/eng/covid-19_coronavirus-head-shaking-doctors-and-non-hand-shaking-workers/45580772 (26 Feb 2020).
• COVID-19 Scientific Resources Retrieved from https://www.cerahgeneve.ch/resources/covid-19-free-online-scientific-resources/ (30 March 2020).
• Data from CPCB: https://app.cpcbccr.com/ccr/#/login (Accessed on 28 May).
• Novel Coronavirus Disease (COVID-19) India Situation Update Report 127, WHO Retrieved from https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200526-covid-19-sitrep-127.pdf?sfvrsn=7b6655ab_8 (27 May 2020).
• Migration figure during lockdown https://theconversation.com/mapping-the-lockdown-effects-in-india-how-geographers-can-contribute-to-tackle-covid-19-diffusion-136323 (Accessed on 26 May).
• COVID-19 Data Retrieved From https://ourworldindata.org/coronavirus-data?country=~IND (Accessed on 21 May 2020).

CODE AVAILABILITY

Any codes used in the paper available upon request from h.bherwani@neeri.res.in and a1_gupta@neeri.res.in.

Received: 8 April 2020; Accepted: 14 August 2020;
Published online: 22 September 2020

REFERENCES

1. Novel, C. P. E. R. E. & others The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) in China. *Zhonghua liu xing bing. xue za zhi= Zhonghua liuxingbingxue zazhi* **41**, 145–151 (2020).
2. Wu, F. et al. A new coronavirus associated with human respiratory disease in China. *Nature* **579**, 265–269 (2020).
3. Zi Y, et al. Coronavirus disease 2019 (COVID-19): a perspective from China. *Radiology*. https://doi.org/10.1148/radiol.2020200490 (2020).
4. Gorbalenya, A. E. Severe acute respiratory syndrome-related coronavirus—the species and its viruses, a statement of the Coronavirus Study Group. *BioRxiv*. https://doi.org/10.1101/2020.02.07.937862 (2020).
5. Lai, C.-C., Shih, T.-P., Ko, W.-C., Tang, H.-J. & Hsueh, P.-R. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and corona virus disease-2019 (COVID-19): the epidemic and the challenges. *Int. J. Antimicrob. Agents* **55**, 1–9 (2020).
6. Bhatia, R. Need for integrated surveillance at human-animal interface for rapid detection & response to emerging coronavirus infections using One Health approach. *Indian J. Med. Res.* **151**, 132–135 (2020).
7. Bedford, J. et al. COVID-19: towards controlling of a pandemic. *Lancet* **395**, 1015–1018 (2020).
8. Wang, C. et al. A novel coronavirus outbreak of global healthconcern. *Lancet* **395**, 470–473 (2020).
9. Sohrabi, C. et al. World Health Organization declares global emergency: a review of the 2019 novel coronavirus (COVID-19). *Int. J. Surg.* **76**, 71–76 (2020).
10. Fisher, D. & Wilder-Smith, A. The global community needs to swiftly ramp up the response to contain COVID-19. *Lancet* **395**, 1109–1110 (2020).
11. MacIntyre, C. R. Global spread of COVID-19 and pandemic potential. *Glob. Biosecur.* https://doi.org/10.31646/gbio.55 (2020).
12. Coronavirus: India enters 'total lockdown' after spike in cases, BBC News. https://www.bbc.com/news/world-asia-india-52024239 (2020).
13. Nolan Pinto. Lockdown is a curfew, only emergency services open. India Today. https://www.indiatoday.in/india/story/coronavirus-lockdown-curfew-bengaluru-police-commissioner-bhaskar-rao-karnataka-covid-19-1658912-2020-03-24 (2020).
14. Kamel Boulos, M.N., Geraghty, E.M. Geographical tracking and mapping of coronavirus disease COVID-19/severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) epidemic and associated events around the world: how 21st century GIS technologies are supporting the global fight against outbreaks and epidemics. Int J Health Geogr **19**, 8. https://doi.org/10.1186/s12942-020-00202-8 (2020).
15. Khan, N. & Naushad, M. Effects of Corona Virus on the World Community. https://doi.org/10.2139/ssrn.3532001 (2020).
16. Oke, J. & Heneghan, C. Global covid-19 case fatality rates. CEBM Res. https://www.cebm.net/covid-19/global-covid-19-case-fatality-rates/ (2020).
17. Poole, L. Seasonal Influences on the Spread Of SARS-CoV-2 (COVID19), Causality, and Forecastabililty (15 Mar 2020). https://doi.org/10.2139/ssrn.3554746 (2020).
18. Ma, Y. et al. Effects of temperature variation and humidity on the death of COVID-19 in Wuhan, China. *Sci. Total Environ.* **724**, 138226 (2020).
19. Oliveiros, B., Caramelo, L., Ferreira, N. C., & Caramelo, F. Role of temperature and humidity in the modulation of the doubling time of COVID-19 cases. *medRxiv*. https://doi.org/10.1101/2020.03.05.20031872 (2020).
20. Wang, M. et al. Temperature significant change COVID-19 transmission in 429 cities. *medRxiv*. https://doi.org/10.1101/2020.02.22.2002579 (2020).
21. Sajadi, et al. Temperature, humidity and latitude analysis to predict potential spread and seasonality for COVID-19. SSRN: https://ssrn.com/abstract=3550308 (2020).
22. Ma, Y. et al. Effects of temperature variation and humidity on the death of COVID-19 in Wuhan, China. *Sci. Total Environ.* **724**, 138226 (2020).
23. Park, J. et al. Effects of temperature, humidity, and diurnal temperature range on influenza incidence in a temperate region. *Influenza Respir. Viruses*. https://doi.org/10.1111/irv.12682 (2019).
24. Chan, K. H. et al. The effects of temperature and relative humidity on the viability of the SARS coronavirus. *Adv. Virol.* **2011**, 1–7 (2011).
25. Sun, Z., Thilakavathy, K., Kumar, S. S., He, G. & Liu, S. V. Potential factors influencing repeated SARS outbreaks in China. *Int. J. Environ. Res. Public Health* **17**, 1633 (2020).
26. Bherwani, H. et al. Valuation of air pollution externalities: comparative assessment of economic damage and emission reduction under COVID-19 lockdown. *Air Qual. Atmos. Health* **13**, 683–694 (2020).
27. Gupta, A. et al. Air pollution aggravating COVID-19 lethality? Exploration in Asian cities using statistical models. *Environ. Dev. Sustain.* https://doi.org/10.1007/s10668-020-00878-9 (2020).
28. Lian, X., Hang, J., Zhang, L. & Liu, C. Environmental indicator for effective control of COVID-19 spreading. *medRxiv*. https://doi.org/10.1101/2020.05.12.20099804 (2020).
29. Qi, H. et al. COVID-19 transmission in mainland China is associated with temperature and humidity: a time-series analysis. *Sci. Total Environ*. https://doi.org/10.1016/j.scitotenv.2020.138778 (2020).
30. Casanova, L. M., Jeon, S., Rutala, W. A., Weber, D. J. & Sobsey, M. D. Effects of air temperature and relative humidity on coronavirus survival on surfaces. *Appl. Environ. Microbiol.* **76**, 2712–2717 (2010).
31. Van Doremalen, N., Bushmaker, T. & Munster, V. J. Stability of Middle East respiratory syndrome coronavirus (MERS-CoV) under different environmental conditions. *Eur. Surveill.* **18**, 20590 (2013).
32. Pirouz, B., Haghshenas, S. S., Haghshenas, S. S. & Piro, P. Investigating a serious challenge in the Sustainable Development process: analysis of confirmed cases of COVID-19 (new type of coronavirus) through a binary classification using artificial intelligence and regression analysis. *Sustainability* **12**, 2427 (2020).
33. Chanprasopchai, P., Pongsumpun, P. & Tang, I. M. Effect of rainfall for the dynamical transmission model of the dengue disease in Thailand. *Comput. Math. Methods Med.* https://doi.org/10.1155/2017/2541862 (2017).
34. Liu, T. et al. Early rigorous control interventions can largely reduce dengue outbreak magnitude: experience from Chaozhou, China. *BMC Public Health* **18**, 90 (2017).
35. Niakan Kalhori, S. R., Ghazisaeedi, M., Azizi, R. & Naserpour, A. Studying the influence of mass media andenvironmental factors on influenza virus transmission in the US Midwest. *Public Health* **170**, 17–22 (2019).
36. Shi, P. et al. The impact of temperature and absolute humidity on the coronavirus disease 2019 (COVID-19) outbreak-evidence from China. *medRxiv*. https://doi.org/10.1101/2020.03.22.20038919 (2020).
37. PTI. Lockdown may help flatten India's COVID-19 curve. https://www.deccanherald.com/national/lockdown-may-help-flatten-covid-19-curve-in-india-says-study-820401.html (2020).
38. Coronavirus: search for hundreds of people after Delhi prayer meeting. BBC News. https://www.bbc.com/news/world-asia-india-52104753 (2020).
39. Defying social distancing norms by residents of Govandhi's Shivaji Nagar: https://www.mumbailive.com/en/politics/people-are-violating-social-distance-in-mumbai-govandi-bjp-leader-kirit-somaiya-criticizes-uddhav-thackeray-government-49754 (2020).
40. Crowd defying social distancing norms at Ahmadabad. https://www.business-standard.com/article/pti-stories/social-distancing-to-be-followed-during-ration-distribution-120033101535_1.html (2020). Accessed 18 May.
41. Aroosa Ahmed and Ashutosh Tripathi. 2,000 migrant workers gather at Mumbai's Bandra to catch train to Bihar. https://www.hindustantimes.com/india-news/2-000-migrant-workers-gather-at-mumbai-s-bandra-to-catch-train-to-bihar/story-H3p5B1z7erXL1697AAtxcl.html (2020). Accessed 19 May.
42. Population density of Mumbai as per census. https://www.census2011.co.in/census/district/357-mumbai-city.html (2020). Accessed on 25 May.
43. Muzamil, P. S. & Nabi, W. Schools in Kashmir reopen after seven months. *Wire*. https://thewire.in/rights/kashmir-school-students (2020).
44. Fukatani, Y. et al. Effect of ambient temperature and relative humidity on interfacial temperature during early stages of drop evaporation. *Phys. Rev*. https://doi.org/10.1103/PhysRevE.93.043103 (2016).

45. Shiflet, A. B. & Shiflet, G. W. *Introduction to Computational Science: Modeling and Simulation for the Sciences* (Princeton University Press, Princeton, 2014).
46. Hethcote, H. W. The mathematics of infectious diseases. *SIAM Rev.* **42**, 599–653 (2000).
47. Kucharski, A. J. et al. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. *Lancet Infect. Dis.* **20**, 1–7 (2020).
48. Asghar Ali, M., Rafiq, M. & Oazir Ahmad, M. Numerical analysis of a modified SIR epidemic model with the effect of time delay. *J. Math.* **51**, 79–90 (2018).
49. Adlhoch, C. et al. Estimation of the number of exposed people during highly pathogenic avian influenza virus outbreaks in EU/EEA countries, October 2016–September 2018. *Zoonoses Public Health* **66**, 874–878 (2019).
50. Srivastava, A. et al. Geographical variation in COVID-19 cases, prevalence, recovery and fatality rate by phase of national lockdown in India. *medRxiv.* https://doi.org/10.1101/2020.06.04.20122028 (2020).
51. Wang, J., Tang, K., Feng, K. & Lv, W. High temperature and high humidity reduce the transmission of COVID-19. Available at SSRN 3551767 (2020).
52. Myers, R. H., Montgomery, D. C., & Anderson-Cook, C. M. *Response Surface Methodology: Process and Product Optimization using Designed Experiments* 4th edn, 1–856 (Wiley, Hoboken, 2016).
53. Bower, K. M. Analysis of variance (ANOVA) using MINITAB. *Sci. Comput. Instrum.* **17**, 64–65 (2000).
54. Raj, I. et al. A high performance biological degradation of trimethylamine: experimental study and mathematical modeling. *J. Environ. Chem. Eng.* **7**, 103292 (2019).
55. Bi, Q. et al. Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study. *Lancet Infect. Dis.* **20**, 911–919 (2020).
56. India confronts its first coronavirus 'super-spreader' — a Muslim missionary group with more than 400 members infected. https://www.washingtonpost.com/world/asia_pacific/india-coronavirus-tablighi-jamaat-delhi/2020/04/02/abdc5af0-7386-11ea-ad9b254ec99993bc_story.html (2020). Accessed on 18 May.
57. Prashasti, A. Kerala celebrates Attukul Pongala festival amid coronavirus scare. https://www.thehindubusinessline.com/news/national/kerala-celebrates-attukul-pongala-festival-amid-coronavirus-scare/article31022882.ece (2020). Accessed on 18 May.
58. Mahesh, L. Migrant workers in Surat come out on road demanding salaries. https://www.thehindu.com/news/national/lockdown-migrant-workers-come-out-on-road-demanding-salaries/article31313403.ece (2020). Accessed on 18 May.
59. Social distancing, what? Karnataka MLA throws pompous birthday bash, invites 500 people. http://cms.newindianexpress.com/states/karnataka/2020/apr/11/social-distancing-what-karnataka-mlas-throws-pompous-birthday-bash-invites-500-people--2128534.html (2020). Accessed on 18 May.
60. Shreyas HS. Coronavirus: After restricting mass gatherings, Karnataka CM attends wedding which had 2,000 guests. https://timesofindia.indiatimes.com/city/bengaluru/coronavirus-after-restricting-mass-gatherings-karnataka-cm-attends-wedding-which-had-2000-guests/articleshow/74644166.cms (2020). Accessed on 18 May.
61. Kumar, S. Effect of meteorological parameters on spread of COVID-19 in India and air quality during lockdown. *Sci. Total Environ.* **745**, 141021 (2020).
62. Harko, T., Lobo, F. S. N. & Mak., M. K. Exact analytical solutions of the susceptible-infected-recovered (SIR) epidemic model and of the SIR model with equal death and birth rates. *Appl. Math. Comput.* **236**, 184–194 (2014).
63. Fanelli, D., & Piazza, F. *Analysis and Forecast of COVID-19 Apreading in China, Italy and France* 134 (Chaos, Solitons & Fractals, 2020).
64. Office of the Registrar General & Census Commissioner, Ministry of Home Affairs, Government of India, Population data—Census of India. https://censusindia.gov.in/Census_And_You/migrations.aspx (2020).
65. Rustan, R. & Handayani, L. (2020). THE outbreak's modeling of coronavirus (COVID-19) using the modified seir model in Indonesia. *Jurnal Fisika dan Aplikasinya* **5**, 61–68.
66. Adnan Shereen, M., Khan, S., Kazmi, A., Bashir, N. & Siddique, R. COVID-19 infection: origin, transmission, and characteristics of human coronaviruses. *J. Adv. Res.* **24**, 91–98 (2020).
67. Mahmud, A. & Lim, P. Y. Applying the SEIR model in forecasting the COVID-19 trend in Malaysia: a preliminary study. *medRxiv.* https://doi.org/10.1101/2020.04.14.20065607 (2020).
68. Lauer, S. A. et al. The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. *Ann. Intern. Med.* **172**, 577–582 (2020).
69. Ke, M. et al. Estimate the incubation period of coronavirus 2019 (COVID-19). *medRxiv.* https://doi.org/10.1101/2020.02.24.20027474 (2020).
70. Meng, L., Hua, F. & Bian, Z. Coronavirus disease 2019 (COVID-19): emerging and future challenges for dental and oral medicine. *J. Dent. Res.* **99**, 481–487 (2020).
71. Kraemer, M. U. G. et al. An Investigation of transmission control measures during the first 50 days of the COVID-19 pandemic in China. *Science* **368**, 638–642 (2020).
72. Mossong, J. et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med.* **5**, 381–391 (2008).
73. Singh, R. & Adhikari, R. Age-structured impact of social distancing on the COVID-epidemic in India. arXiv Preprint arXiv2003.12055 (2020).
74. Jiang, L. et al. Infectious disease transmission: survey of contacts between hospital-based healthcare workers and working adults from the general population. *J. Hosp. Infect.* **98**, 404–411 (2018).
75. Diaz, P., Constantine, P., Kalmbach, K., Jones, E. & Pankavich, S. A modified SEIR model for the spread of Ebola in Western Africa and metrics for resource allocation. *Appl. Math. Comput.* **324**, 141–155 (2018).

ACKNOWLEDGEMENTS

The authors thankfully acknowledge Council of Scientific and Industrial Research (CSIR), India and its constituent laboratory National Environmental Engineering Research Institute (NEERI) for support the research. The manuscript is checked for plagiarism using licensed iThenticate software wide CSIR-NEERI Knowledge Resource Centre [KRC], CSIR- CSIR-NEERI/KRC/2020/APRIL/CSUM-DRC-DIR/1 on 08 April 2020.

AUTHOR CONTRIBUTIONS

R.K. and H.B. conceived the idea and designed the study with inputs from A.A. and A. G. S.A. prepared the data. H.B. and A.G. conducted the analysis and worked on the models using datasets. A.A. and S.A. compiled the results and wrote the initial draft. H.B., A.G., A.A., and R.K. improved the discussion and results.

COMPETING INTERESTS

The authors declare that they have no conflict of interest associated with this publication and there has been no significant financial or non-financial support that could have influenced its outcome.

ADDITIONAL INFORMATION

Supplementary information is available for this paper at https://doi.org/10.1038/s41612-020-00142-x.

Correspondence and requests for materials should be addressed to H.B. or A.G.