medRxiv preprint doi: https://doi.org/10.1101/2020.07.05.20146324; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Study of the Dependence of Effective Reproduction Number of COVID-19 on the Temperature and Humidity: A Case Study with the Indian States

Manoj Mandal[1], Subhradeep Patra[1], Sabyasachi Pal*[1], Suman Acharya[1], and Mangal Hazra[1]

[1]Midnapore City College, Kuturia, Paschim Medinipur, India, 721129

Abstract

Corona Virus Disease 2019 (COVID-19) started in Wuhan province of China in November 2019 and within a short time, it was declared as a worldwide pandemic by World Health Organisation due to very fast worldwide spread of the virus. In the absence of any vaccine, various mitigation measures were used. In the past, the effect of temperature and humidity on the spread of the virus was studied for a very early phase of the data with mixed results. We studied the impact of COVID-19 on the maximum temperature and relative humidity of a place using Indian states as test cases for SIR, SIRD, and SEIR models. We used a linear regression method to look for any dependency between the effective reproduction number with the maximum temperature and relative humidity. Most of the states showed a correlation with the negative slope between the effective reproduction number with the maximum temperature and the relative humidity. It indicates that the effective reproduction number goes down as maximum temperature or relative humidity rise. But, the regression coefficient R^2 was low for these correlations which means that the correlation is not strong.

Keywords— Coronavirus, COVID-19, Environmental factors

1 Introduction

Corona Virus Disease 2019 (COVID-19) started in Wuhan province of China in November 2019 and within a short time, it infected millions of people worldwide [2, 3, 4]. The disease is caused by severe acute respiratory syndrome coronavirus named SARS-CoV-2 [32]. World Health Organisation recognised the disease as pandemic [35, 32].

In India, the first case was reported on 30 January 2020 [33]. Till 31st May 2020, India had 190609 total confirmed cases which were highest in Asia [30] and the fourth-highest

*corresponding author: sabya.pal@gmail.com

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.05.20146324; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

number of confirmed cases in the world [30]. Study with India is interesting as it is the world's second largest population with about one sixth of the world's total population [34]. During the above mentioned period, India had 91852 recoveries, 1 migration, and 5408 death [19]. From 24th March, India imposed a few lockdowns in a staged manner. We have limited our study till 31st May as from first June India started phased unlocking [5] and various other parameters contributed after that in the spread of the virus.

The spread of an epidemic depends on a large number of factors. Earlier, it was found that humidity is one of the driving factors to explain the seasonality of influenza [27]. Shaman & Kohn [27] studied the effect of relative and absolute humidity on the transmission of the influenza virus. They have shown 50% of the influenza virus transmission variability can be explained by absolute humidity but they restrict the observation with limited data.

At the beginning of the spread of the disease, it was widely believed that summer will cause the virus to reduce [8]. Experimental results of other coronaviruses also indicated that the virus may be less effective with rising temperature and humidity [9]. There were a few studies to investigate the dependency of transmission of COVID-19 on temperature and humidity [28, 1, 9, 17, 18, 16, 21] using early data of COVID-19 pandemic which show contradictory results. Using very early data of 100 Chinese cities, Wang et al. [28] found that high temperature and high humidity significantly reduce the transmission of COVID-19 while Bhattacharjee [1] concluded the opposite using initial data from four of the most affected places from China and five of the most affected places in Italy. A rigorous study was necessary with a large amount of data to effectively study the effect of temperature and humidity on the spread of coronavirus.

In the present paper, we investigate the influence of maximum air temperature and relative humidity on effective contact rate and reproduction number [26] of COVID-19. We studied the effective reproduction number $\mathbb{R}(t)$ for different states of India till 31st May 2020.

In Section 2, we summarised different mathematical models of the pandemic. We highlighted the SIR model in Subsection 2.1. Subsection 2.2 carried a brief introduction of the SIRD model. A short introduction of the SEIR model was given in Subsection 2.3. In Section 3, we presented our findings. The temperature dependence of the effective reproduction number was discussed in the Subsection 3.1 and the humidity dependence of effective reproduction number $\mathbb{R}(t)$ was discussed in Subsection 3.2. In Section 4, we discussed the implication of our findings and Section 5 we summarised our result.

2 Mathematical models

Different mathematical models were designed to simulate the effect of the disease from different perspectives [22, 23, 25, 20, 21, 12]. These helped us to get an idea about the spread rate of the disease over the population. Here we study the most common SIR, SIRD, and SEIR models. The spread rate depends on a large number of parameters such as mean incubation time [1, 6], mean infectious period, social distancing, and risk of international spread [7, 36]. The most important part of these models is to calculate the basic reproduction number (R_0) which is the contagiousness of the diseases. R_0 determines the average number of people can be affected by a single infected person over a course of time. $R_0 > 1$ indicates that

medRxiv preprint doi: https://doi.org/10.1101/2020.07.05.20146324; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

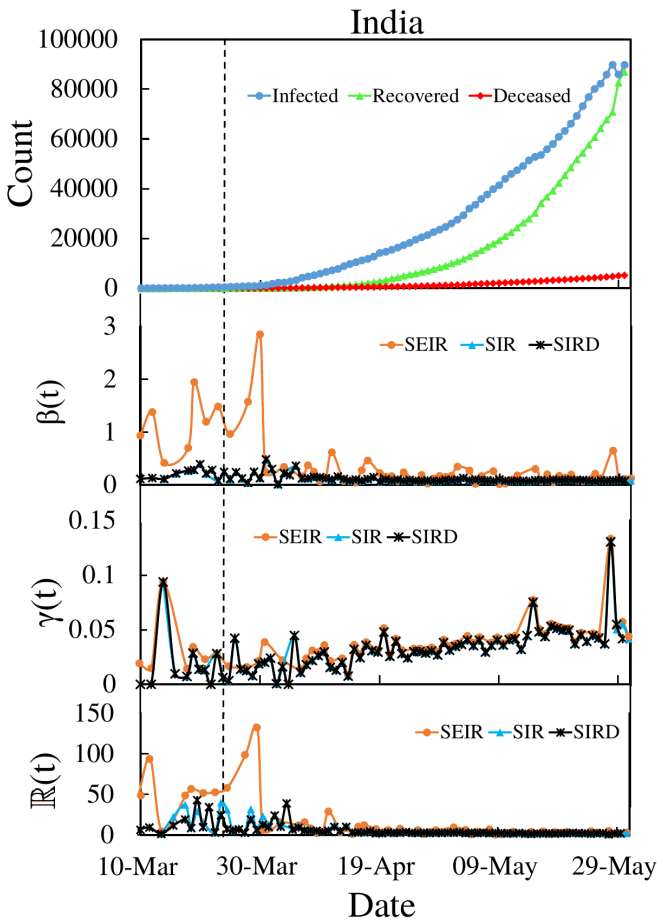

Figure 1: Variation of different parameters with time for India. First row: Variation of the number of infected, recovered, and deceased individuals with time. Second row: Variation of effective contact rate ($\beta(t)$) with time. Third row: Variation of recovery rate ($\gamma(t)$) with time. Fourth row: Variation of effective reproduction number ($\mathbb{R}(t)$) with time. The time of the imposition of the first lockdown is shown with a dotted vertical line.

the spread is increasing, $R_0 = 1$, indicates that the spread is stable and $R_0 < 1$, indicates that the spread is expected to stop. We studied the effective reproduction number $\mathbb{R}(t)$ for India using different models. We also studied the evolution of $\mathbb{R}(t)$ using the same models for different individual states of India. In Figure 1, we have shown the variation of $\mathbb{R}(t)$ for India, using three different models.

2.1 SIR model

There exists a lot of models for analyzing the spread of an epidemic in a large population. The Susceptible-Infected-Recovered (SIR) model [14] is the simplest compartmental model that can describe the evolution of the epidemic at the population level. We used this model to study the different time dependent parameters like recovery rate (γ), contact rate (β), and effective reproduction number ($\mathbb{R}$).

medRxiv preprint doi: https://doi.org/10.1101/2020.07.05.20146324; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

At any time t, $S(t)$ be the total number of susceptible individuals, $I(t)$ be the total number of infected individuals, and $R(t)$ be the total number of recovered (till a given time) individuals from the epidemic.

We shall mostly deal with the fractions:
$s(t) = \frac{S(t)}{N}$, $i(t) = \frac{I(t)}{N}$, $r(t) = \frac{R(t)}{N}$
If the total population is N, then the SIR model can be expressed as

$$\frac{ds(t)}{dt} = -\beta_n i_n s_n \tag{1}$$

$$\frac{di(t)}{dt} = \beta_n i_n s_n - \gamma_n i_n \tag{2}$$

$$\frac{dr(t)}{dt} = \gamma_n i_n \tag{3}$$

Finally, in discrete form, the equation (1), (2) and (3) can be written as

$$s_{n+1} - s_n = -\beta_n i_n s_n \tag{4}$$

$$i_{n+1} - i_n = \beta_n i_n s_n - \gamma_n i_n \tag{5}$$

$$r_{n+1} - r_n = \gamma_n i_n \tag{6}$$

From Eq (5) we can write

$$\beta_n = \frac{i_{n+1} - (1-\gamma_n) i_n}{i_n s_n} \tag{7}$$

The basic reproduction number R_0 can be given by

$$R_0 = \frac{\beta}{\gamma} \tag{8}$$

The effective reproduction number $\mathbb{R}(t)$ can be given by

$$\mathbb{R}(t) = \frac{\beta_n}{\gamma_n} \tag{9}$$

In Figure 1, we have shown various parameters for the SIR model. In the second and third row, the variation of effective contact rate ($\beta_{\text{SIR}}(t)$) and the variation of recovery rate ($\gamma_{\text{SIR}}(t)$) were shown. The fourth row represented the variation of effective reproduction number ($\mathbb{R}_{\text{SIR}}(t)$) for the SIR model.

In Figure 2, we have shown the variation of different parameters for the SIR model for few selected Indian states. In the second and third row, the variation of $\beta_{\text{SIR}}(t)$ and the variation of $\gamma_{\text{SIR}}(t)$ were shown. The fourth row showed the variation of effective reproduction number ($\mathbb{R}_{\text{SIR}}(t)$) for the SIR model.

For the study of SIR, SIRD, and SEIR models, we assumed that the total population N was the same throughout the study. It was also assumed that those individuals, who have recovered from the disease can not become infected again. For a few special cases, β_n and corresponding $\mathbb{R}_n$ may be negative for a certain value of n [12] which are not physical. For a few particular days, the value of $\mathbb{R}_n$ was very high which may bias the correlation studies in later Sections. We have not used these values for our study.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.05.20146324; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

2.2 SIRD model

The SIRD model [20] is the modified version of the SIR model. We used this model to study the different time dependent parameters like recovery rate (γ), contact rate (β), and effective reproduction number ($\mathbb{R}$). At any time t, $S(t)$ be the total number of susceptible individuals, $I(t)$ be the total number of infected individuals, $R(t)$ be the total number of recovered, $D(t)$ be the total number of deceased individuals from the epidemic.

Here, β is the contact rate and γ is the recovery rate and α is the mortality rate. We shall mostly deal with fractions: $s(t) = \frac{S(t)}{N}$, $i(t) = \frac{I(t)}{N}$, $r(t) = \frac{R(t)}{N}$, $d(t) = \frac{D(t)}{N}$. From conservation law s(t)+i(t)+r(t)+d(t)=1
The SIRD model can be expressed as

$$\frac{ds(t)}{dt} = -\beta_n i_n s_n \tag{10}$$

$$\frac{di(t)}{dt} = \beta_n i_n s_n - (\gamma_n + \alpha_n) i_n \tag{11}$$

$$\frac{dr(t)}{dt} = \gamma_n i_n \tag{12}$$

$$\frac{dd(t)}{dt} = \alpha_n i_n \tag{13}$$

We use the discrete form of the equations (10) to (13).

From Eq (11) we can write

$$\beta_n = \frac{i_{n+1} + (\gamma_n + \alpha_n - 1) i_n}{i_n s_n} \tag{14}$$

The effective reproduction number $\mathbb{R}(t)$ can be given by

$$\mathbb{R}(t) = \frac{\beta_n}{\gamma_n + \alpha_n} \tag{15}$$

In Figure 1, we have shown various parameters for the SIRD model. In the second and third rows, the variation of effective contact rate ($\beta_{\text{SIRD}}(t)$) and the variation of recovery rate ($\gamma_{\text{SIRD}}(t)$) were shown. The fourth row represented the variation of effective reproduction number ($\mathbb{R}_{\text{SIRD}}(t)$) for the SIRD model.

In Figure 2, we have shown a variation of different parameters for the SIRD model for few selected Indian states. In the second and third row, the variation of effective contact rate ($\beta_{\text{SIRD}}(t)$) and the variation of recovery rate ($\gamma_{\text{SIRD}}(t)$) were shown. The fourth row represents the variation of effective reproduction number ($\mathbb{R}_{\text{SIRD}}(t)$) for the SIRD model.

2.3 SEIR model

The Susceptible-Exposed-Infectious-Removed (SEIR) model [15] is the most studied one. The SEIR model is a variant of the SIR model (Section 2.1). At any time t, $S(t)$ be the total number of susceptible individuals, $E(t)$ be the total number of exposed individuals, $I(t)$ be the total number of infected individuals, and $R(t)$ be the total number of removed

medRxiv preprint doi: https://doi.org/10.1101/2020.07.05.20146324; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

(recovered or deceased till a given time) individuals from the epidemic. In the SEIR model, we considered the average incubation period $1/\sigma$ as a constant.

The parameter β is the product of the average number of contacts per person and per unit time by the probability of disease transmission in contact between susceptible and infectious individuals. γ is the recovery rate. The compartment E of the exposed individuals in the SEIR model makes the model slightly more delicate. The SEIR model is a compartmental model used to understand the mathematical modeling of infectious diseases in a large population. If N is the total population size then in fraction form $s(t) = \frac{S(t)}{N}$, $e(t) = \frac{E(t)}{N}$, $i(t) = \frac{I(t)}{N}$, $r(t) = \frac{R(t)}{N}$
Conservation of the total number of individuals, we have the relation
$s(t) + e(t) + i(t) + r(t) = 1$
For any time $t \geq 0$, the SEIR model can be expressed by the differential equations

$$\frac{ds(t)}{dt} = -\beta_n s_n i_n \tag{16}$$

$$\frac{de(t)}{dt} = \beta_n s_n i_n - \sigma e_n \tag{17}$$

$$\frac{di(t)}{dt} = \sigma e_n - \gamma_n i_n \tag{18}$$

$$\frac{dr(t)}{dt} = \gamma_n i_n \tag{19}$$

We use the discrete form of equations (16) to (19).

Now we use equation (17) to find β_n

$$\beta_n = \frac{e_{n+1} - e_n + \sigma e_n}{s_n i_n} \tag{20}$$

Using Eq(17), Eq(18) and Eq(20) we can show

$$\beta_n = \frac{1}{s_n i_n \sigma}[i_{n+2} + (\gamma_n + \sigma - 2)i_{n+1} + (\sigma - 1)(\gamma_n - 1)i_n] \tag{21}$$

The effective reproduction number $\mathbb{R}$(t) can be given by

$$\mathbb{R}(t) = \frac{\beta_n}{\gamma_n} \tag{22}$$

In Figure 1, we have shown various parameters for the SEIR model. In the second and third rows, the variation of effective contact rate ($\beta_{\text{SEIR}}(t)$) and the variation of recovery rate ($\gamma_{\text{SEIR}}(t)$) were shown. The fourth row represented the variation of effective reproduction number ($\mathbb{R}_{\text{SEIR}}(t)$) for the SEIR model.

In Figure 2, we have shown the variation of different parameters for the SEIR model for few selected Indian states. In the second and third row, the variation of effective contact rate ($\beta_{\text{SEIR}}(t)$) and the variation of recovery rate ($\gamma_{\text{SEIR}}(t)$) were shown. The fourth row showed the variation of effective reproduction number ($\mathbb{R}_{\text{SEIR}}(t)$) for the SEIR model.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.05.20146324; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

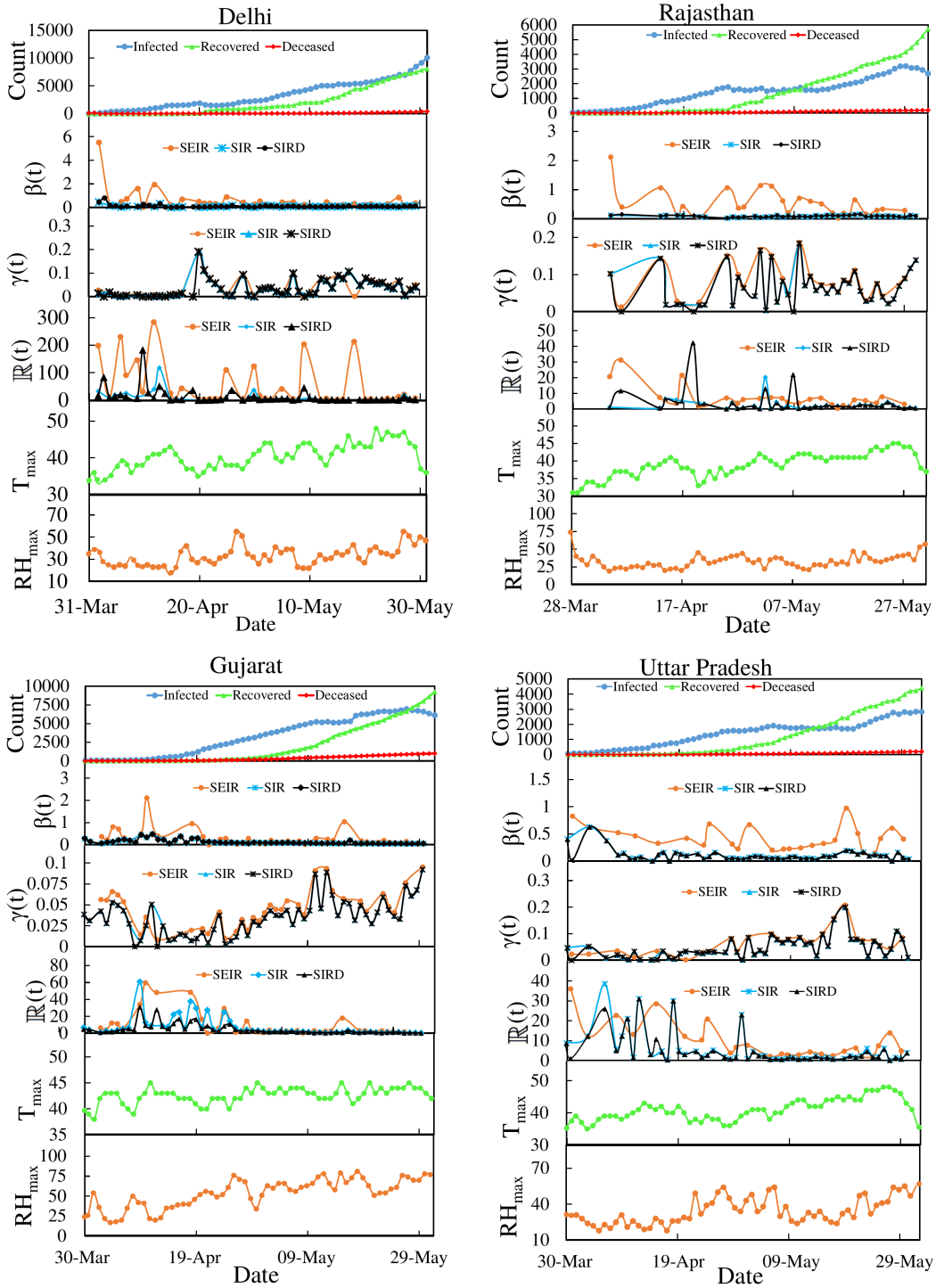

Figure 2: Variation of different parameters for four Indian states Delhi, Rajasthan, Gujarat, and Uttar Pradesh. First row: Variation of the total number of infected, recovered, and deceased individuals with time. Second row: Variation of effective contact rate $\beta(t)$ with time using three different models. Third row: Variation of recovery rate $\gamma(t)$ with time using three different models. Fourth row: Variation of the effective reproduction number $\mathbb{R}(t)$ with time using three different models. Fifth row: Variation of maximum temperature T_{max} with time. Sixth row: Variation of maximum relative humidity RH_{max} with time.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.05.20146324; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

3 Result

We studied the evolution of effective reproduction number $\mathbb{R}(t)$ for India using three different models. For our study, we used the time series data for infected, recovered, and deceased individuals of COVID-19 [11]. To study the dependency of $\mathbb{R}(t)$, $\beta(t)$, $\gamma(t)$ on the temperature and humidity variation [31] of a state, we used temperature and humidity of the capital of the corresponding state as representative. We restricted our study for states for which the number of infected individuals was greater than 1000 till 31st May. In the first row of Figure 1, we have shown the daily variation of the number of infected, recovered, and deceased individuals. In the second row of Figure 1, we have shown the variation of the effective contact rate $\beta(t)$ for three different models. The SEIR model fluctuates at the beginning of the pandemic but the value of $\beta(t)$ is close for SIR and SIRD models. In the third row of the figure, we have shown the variation of recovery rate $\gamma(t)$ with time which is close for the three models and varies between 0 to 0.15. In the fourth row of Figure, we have shown the variation of $\mathbb{R}(t)$ for three different models. The SEIR model showed more fluctuation at the beginning of the study. But, since we restricted our study of temperature and humidity dependency (in Section 3.1 and 3.2) with effective reproduction number for individual states for the time when the number of infected was more than 50 individuals, the initial time where β_{SEIR} and $\mathbb{R}_{\text{SEIR}}$ fluctuate was not be used for our study.

We have studied the evolution of different parameters for different Indian states using three different models. In Figure 2, we have shown an example of the variation of different parameters for four states Delhi, Rajasthan, Gujarat, Uttar Pradesh. In Figure 2, the first row showed the variation of the number of infected, recovered and deceased individuals with time, the second row showed the variation of effective contact rate $\beta(t)$ with time, the third row showed the variation of recovery rate $\gamma(t)$ with time, the fourth row showed the variation of effective reproduction number $\mathbb{R}(t)$ with time and the fifth and the sixth rows showed the variation of maximum air temperature T_{max} and maximum relative humidity RH_{max} respectively.

3.1 Temperature dependence of $\mathbb{R}(t)$

We studied the variation of effective reproduction number $\mathbb{R}(t)$ on temperature for different states of India. The linear regression method [10] is used to find the regression coefficient R^2 which is the goodness of fit for linear regression models. Table 1 showed the result of the linear regression method for three different models. The positive sign indicated the correlation between effective reproduction number ($\mathbb{R}(t)$) and maximum temperature (T_{max}) with a positive slope and the negative sign indicated a correlation between $\mathbb{R}(t)$ and temperature with a negative slope. It should be noted that the correlation with negative slope indicated that the effective reproduction number ($\mathbb{R}(t)$) will decrease with the increase of the temperature.

Table 1 indicated the correlation between effective reproduction number with maximum air temperature was not strong and for different states of India, correlations with positive and negative slopes were found. For nine states out of fourteen states, we have found a correlation with negative slopes for all three models, and for only two states (Karnataka and West Bengal) we have found a correlation with positive slopes for all three models. The

medRxiv preprint doi: https://doi.org/10.1101/2020.07.05.20146324; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

State	R^2_{SIR} (%)	R^2_{SIRD} (%)	R^2_{SEIR} (%)
Andhra Pradesh	13.9(−)	13.5(−)	2.9(−)
Bihar	5.6(−)	6.3(−)	0.7(−)
Delhi	2.3(−)	6.4(−)	2.3(−)
Gujarat	10.4(−)	6.7(−)	1.0(−)
Jammu and Kashmir	0.3(−)	0.0(+)	10.9(+)
Karnataka	0.4(+)	0.7(+)	0.9(+)
Madhya Pradesh	4.1(−)	1.9(−)	7.5(−)
Maharashtra	0.1(+)	0.0(+)	0.4(−)
Punjab	11.1(−)	9.2(−)	12.4(−)
Rajasthan	3.9(−)	0.4(−)	15.9(−)
Tamil Nadu	7.1(−)	9.2(−)	12.4(−)
Telangana	1.8(+)	1.2(-)	2.2(+)
Uttar Pradesh	6.7(−)	6.3(−)	11.6(−)
West Bengal	2.0(+)	0.1(+)	0.9(+)

Table 1: Value of regression coefficient R^2 using different models for different states of India from $\mathbb{R}$(t) vs T_{max} plot.

states Andhra Pradesh, Bihar, Delhi, Gujarat, Madhya Pradesh, Punjab, Rajasthan, Tamil Nadu, and Uttar Pradesh showed a correlation between $\mathbb{R}(t)$ and maximum air temperature T_{max} with a negative slope for all the three models. For the remaining states in the table, we saw mixed results between different models. Figure 3 represented the variation of effective reproduction number $\mathbb{R}(t)$ with maximum temperature T_{max} for four representative states Delhi, Rajasthan, Gujarat, and Uttar Pradesh. All four states showed a correlation with the negative slope between effective reproduction number $\mathbb{R}(t)$ and maximum air temperature T_{max} for all the three models. For the state Assam, the number of infected individuals crossed 50 after the second week of May resulting less number of points in correlation study. So, we did not include it in the table.

3.2 Humidity dependence of $\mathbb{R}$(t)

Humidity also plays a crucial role as a driving environmental parameter on the spread of a virus. We also investigated whether there is any dependence of relative humidity on the spread rate of COVID-19. We studied the variation of effective reproduction number $\mathbb{R}(t)$ with maximum relative humidity (RH_{max}). The linear regression method was used to find the parameter R^2. We summarised the result of variation of the effective reproduction number ($\mathbb{R}(t)$) with the maximum relative humidity (RH_{max}) for different states of India in the Table 2. The table indicated that there was a correlation with negative a slope for seven states Delhi, Gujarat, Jammu and Kashmir, Maharashtra, Rajasthan, Tamil Nadu, and Uttar Pradesh for all the three models and we have found a correlation with a positive slope for three states Andhra Pradesh, Punjab, West Bengal for all three models. The rest of the states in the Table 2 showed a mixed correlation for different models. Figure 4 represents the variation of the effective reproduction number $\mathbb{R}(t)$ with RH_{max} for four representative states Delhi, Rajasthan, Gujarat, and Uttar Pradesh. All four states showed a correlation

medRxiv preprint doi: https://doi.org/10.1101/2020.07.05.20146324; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

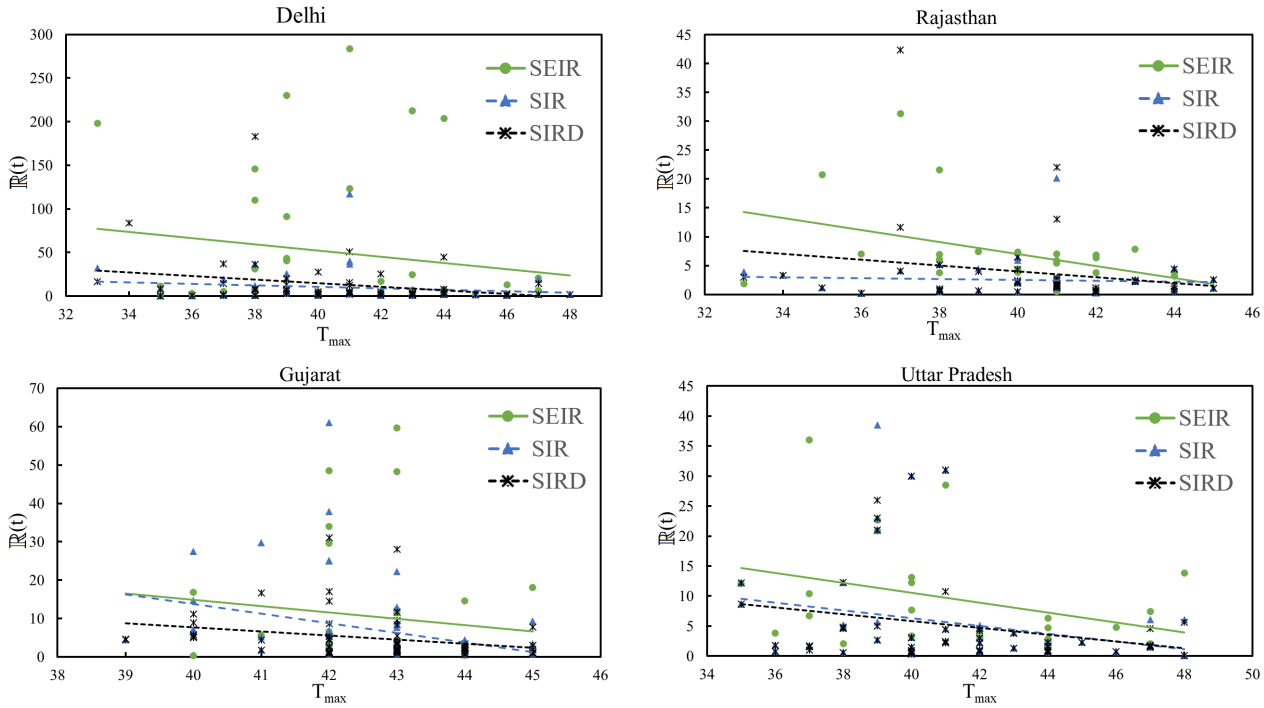

Figure 3: Variation of effective reproduction number $\mathbb{R}(t)$ with maximum temperature T_{max} for four different Indian states Delhi, Rajasthan, Gujarat, and Uttar Pradesh.

with a negative slope between $\mathbb{R}(t)$ and RH_{max} for all the three models.

4 Discussion

In the absence of any vaccine, different mitigation measures were used to control the spread of the disease like lockdown, using masks, personal hygiene, restriction of movement, etc. So, it was difficult to identify the effect of a single parameter on the spread of the disease.

It should be noted that except for Kerala and Maharashtra, for all other states we have used data from dates when strict lockdown measures were in place since we have used only dates when the infected number for an individual state is more than 50 and we restricted our study till 31st May as India started stepped unlocking from 1st June [5]. So, during the almost whole span of the study, the lockdown was imposed over India.

In India, the lockdown was started on 24th March. Figure 1 showed that the average value of $\beta(t)$ and $\mathbb{R}(t)$ in India for all models went down after the imposition of lockdown. The average value of $\beta(t)$ for SIR and SIRD before the lockdown was 0.23 and 0.22 respectively, while the average value of $\beta(t)$ after lockdown went down to 0.11 and 0.12 respectively. For the SEIR model, the value of $\beta(t)$ went down from 1.15 to 0.28. The value of $\mathbb{R}(t)$ also went down for all the models. For the SIR model, it changed from 19.7 to 5.2, and for SIRD model $\mathbb{R}(t)$ went down from 15.4 to 4.6. So, it is clear that lockdown helped to control the spread of the disease.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.05.20146324; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

State	R^2_{SIR} (%)	R^2_{SIRD} (%)	R^2_{SEIR} (%)
Andhra Pradesh	12.8(+)	12.1(+)	10.9(+)
Bihar	1.2(+)	1.7 (+)	1.0(−)
Delhi	3.5(−)	7.1(−)	2.8(−)
Gujarat	8.0(−)	17.7(−)	16.2(−)
Jammu and Kashmir	1.1(−)	1.3(−)	0.6(−)
Karnataka	0.3(+)	0.2(+)	5.7(−)
Madhya Pradesh	1.1(+)	2.5(−)	0.3(+)
Maharashtra	4.1(−)	6.0(−)	5.9(−)
Punjab	19.2(+)	22.8(+)	36.3(+)
Rajasthan	8.6(−)	2.3(−)	21.3(−)
Tamil Nadu	13.4(−)	7.0(−)	19.1(−)
Telangana	0.3(−)	0.4(−)	6.4(+)
Uttar Pradesh	11.2(−)	11.3(−)	9.4(−)
West Bengal	2.7(+)	5.0(+)	12.1(+)

Table 2: Value of regression coefficient R^2 using different models for different states of India for $\mathbb{R}$(t) vs RH_{max} plot.

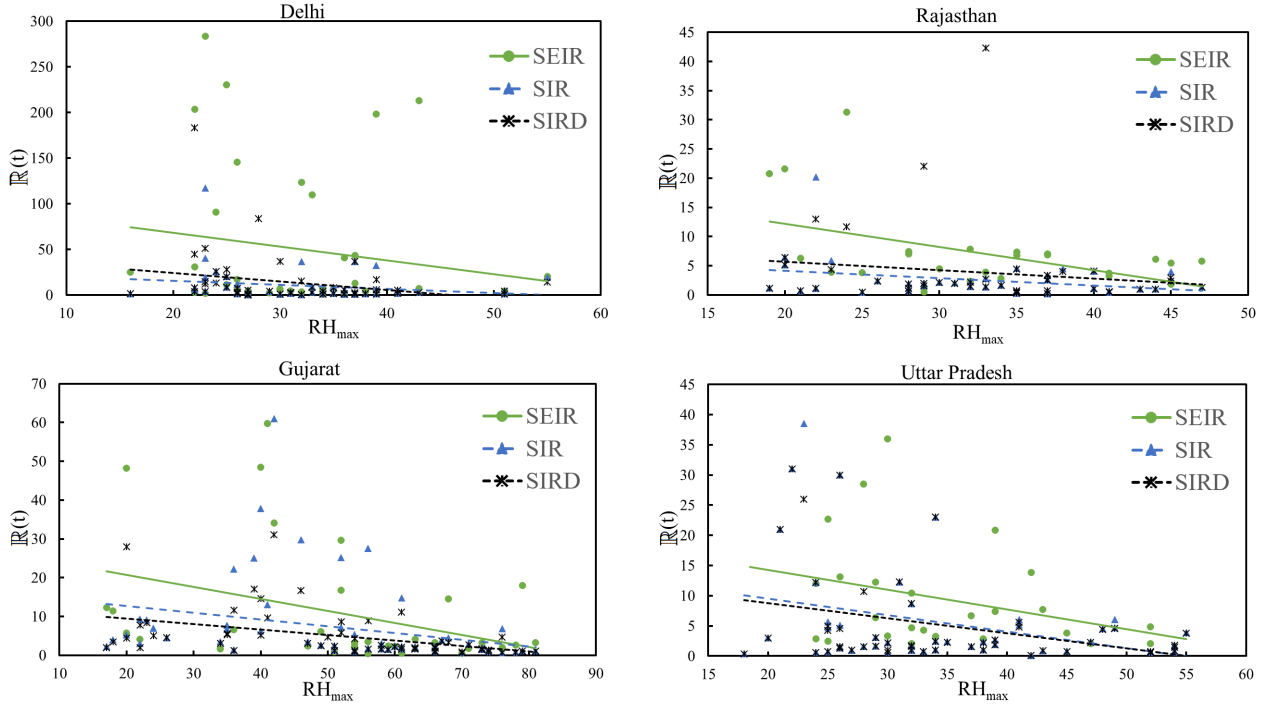

Figure 4: Variation of effective reproduction number $\mathbb{R}(t)$ with maximum relative humidity RH_{max} for four different Indian states Delhi, Rajasthan, Gujarat, and Uttar Pradesh.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.05.20146324; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

For the study of the correlation between the maximum temperature and effective reproduction number, the SIR model showed a maximum correlation (13.9%) with a negative slope for the state Andhra Pradesh, whereas the SIRD model showed a maximum correlation (13.5%) for the same state. The SEIR model showed a maximum correlation (15.9%) with a negative slope for the state Rajasthan for $\mathbb{R}(t)$ vs T_{max} plot. The SEIR model was showing a higher correlation coefficient compared to SIR and SIRD. For all of SIR, SIRD, and SEIR models, 10 states out of 14 states were showing correlation with a negative slope. Among the states showing correlation with negative slopes for all three models, the median values of R^2 for SIR, SIRD, and SEIR models were 6.7, 6.4, and 7.5 respectively. The corresponding mean values were 7.2%, 6.7%, and 7.4%.

For the study of the correlation between the relative humidity and effective reproduction number, the SEIR model showed a maximum correlation (21.3%) with a negative slope for the states Rajasthan, whereas the SIR and SIRD models showed maximum correlations with negative slope 13.4% and 17.7% for the states Tamilnadu and Gujarat respectively. For the SIRD and SEIR models, 9 out of 14 states showed a correlation with negative slope whereas, for the SIR model, 8 out of 14 sources showed a correlation with a negative slope for $\mathbb{R}(t)$ vs RH_{max} plot. The SEIR model was showing a higher correlation coefficient compared to SIR and SIRD models. Among the 7 states showing correlation with negative potential for all states, the median values of R^2 for SEIR, SIR, and SIRD models were 9.4, 8.0, and 7.0 respectively. The corresponding mean values were 10.8%, 7.1%, and 7.5%. The correlation coefficient was close for SIR and SIRD models.

5 Conclusion

We studied the relation between effective reproduction number with maximum temperature and relative humidity of a place using Indian states as a test case. We used the time when most of the states were in lockdown to minimise the effect of other parameters. We used SIR, SIRD, and SEIR models to calculate the effective reproduction number of a place. The linear regression method was used to look for any dependency between effective reproduction number with maximum temperature and relative humidity. Most of the states showed correlation with a negative slope between effective reproduction number with maximum temperature and relative humidity for all three models which indicate that the effective reproduction number tends to go down when maximum temperature and relative humidity increase. But, the regression coefficients R^2 is low for these correlations which indicates that the correlation is not strong. Lockdown also played a crucial role in controlling the spread of the virus.

References

[1] S. Bhattacharjee, *Statistical investigation of relationship between spread off corona virus disease(COVID-19) and environmental factors based on study of four mostly affected places of China and five mostly affected places of Italy*, https://arxiv.org/abs/2003.11277, 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.05.20146324; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

[2] *Pneumonia of unknown cause - China*, https://www.who.int/csr/don/05-january-2020-pneumonia-of-unkown-cause-china/en/, 2020.

[3] J. Ma, *Coronavirus: China's first confirmed Covid-19 case traced back to November 17*, https://www.scmp.com/news/china/society/article/3074991/coronavirus-chinas-first-confirmed-covid-19-case-traced-back, 2020.

[4] M. L. Holshue et al., *First Case of 2019 Novel Coronavirus in the United States*, N Engl J Med, 382, 929, https://www.nejm.org/doi/full/10.1056/NEJMoa2001191, 2020.

[5] *New Guidelines to fight COVID-19 to be effective from 1st June 2020*, https://mha.gov.in/sites/default/files/PR_UNLOCK1Guidelines_30052020.pdf, 2020.

[6] S. G. Baum, *COVID-19 Incubation Period: An Update*, NEJM J. Watch, 2020.

[7] I. I. Bogoch et al., *Pneumonia of Unknown Etiology in Wuhan, China: Potential for International Spread Via Commercial Air Travel. J. Travel Med*, https://doi.org/10.1093/jtm/taaa008, 2020.

[8] Q. Bukhari, Y. Jameel, *Will Coronavirus Pandemic Diminish by Summer?*, https://ssrn.com/abstract=3556998, 2020.

[9] K. H. Chan et al. *The Effects of Temperature and Relative Humidity on the Viability of the SARS Coronavirus*, Advances in Virology., 2011, 2011.

[10] D. A. Freedman, *Statistical Models: Theory and Practice*, Cambridge University Press. p 26, 2009.

[11] https://github.com/datasets/covid-19

[12] G. Goswami, J. Prasad, M. Dhuria, *Extracting the the effective contact rate of COVID-19 pandemic*, https://arxiv.org/abs/2004.07750, 2020.

[13] *Coronavirus in summer-Explore-Google Trends*, https://trends.google.com/trends/explore?q=coronavirusinsummer.

[14] W. O. Kermack and A. G. McKendrick, *Contribution to the mathematical theory of epidemics*, Proceedings of the Royal Society of London. Series A, Containing papers of a mathematical and physical character, 115, 700, 1927.

[15] M. Y. Li, *An Introduction to Mathematical Modeling of Infectious Diseases*, Springer International Publishing, Mathematics of Planet Earth, Vol 2, 2018.

[16] W. Luo et al., *The role of absolute humidity on transmission rates of the COVID-19 outbreak*, https://www.medrxiv.org/content/early/2020/02/17/2020.02.12.20022467, 2020.

[17] M. B. Araujo, B. Naimi. *Spread of SARS-CoV-2 Coronavirus likely to be constrained by climate*, https://www.medrxiv.org/content/early/2020/03/16/2020.03.12.20034728, 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.05.20146324; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

[18] M. M. Sajadi et al., *Temperature, Humidity and Latitude Analysis to Predict Poten-tial Spread and Seasonality for COVID-19*, https://ssrn.com/abstract=3550308, 2020.

[19] Ministry of Health and Family Welfare Government of India, https://www.mohfw.gov.in, 2020.

[20] B. M. Ndiaye, L. Tengana, D. Seck, *Comparative prediction of confirmed cases of COVID-19 pandemic by machine learning, deterministic and stochastic SIR models*, https://arxiv.org/abs/2004.13489, 2020.

[21] A. Notari, *Temperature dependence of COVID-19 transmission*, https://arxiv.org/abs/2003.12417[q-bio.PE], 2020.

[22] G. Pandey et. al., *SEIR and Regression Model based COVID-19 outbreak predictions in India*, https://arxiv.org/pdf/2004.00958[q-bio.PE], 2020.

[23] E. L. Piccolomini, F. Zama, *Preliminary analysis of COVID-19 spread in Italy with an adaptive SEIRD model*, https://arxiv.org/pdf/2003.09909, 2020.

[24] J. Prasad, *A data first approach to modelling Covid-19*, https://www.medrxiv.org/content/10.1101/2020.05.22.20110171v2, 2020.

[25] L. Pribylova and Veronika Hajnova, *SEIAR model with asymptomatic cohort and consequences to efficiency of quarantine government measures in COVID-19 epidemic*, https://arxiv.org/pdf/2004.02601[q-bio.PE], 2020.

[26] Department of Health, *2.2 The reproduction number*, https://www.health.gov.au/internet/publications/publishing.nsf/Content/mathematical-models~mathematical-models-models.htm~mathematical-models-2.2.htm.

[27] J. Shaman, M. Kohn, *Absolute humidity modulates influenza survival, transmission, and seasonality*, Proc. Natl. Acad. Sci. USA, 106, 3243, 2009.

[28] J. Wang et al. *High temperature and high humidity reduces the transmission of COVID-19*, https://arxiv.org/abs/2003.05003, 2020.

[29] *Coronavirus disease (COVID-2019) situation reports*, https://www.who.int/

[30] https://www.worldometers.info/coronavirus/

[31] worldweatheronline.com/

[32] *Naming the coronavirus disease (COVID-19) and the virus that causes it*, https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-2019)-and-the-virus-that-causes-it.

[33] *Update on Novel Coronavirus: one positive case reported in Kerala*, https://pib.gov.in/PressReleaseIframePage.aspx?PRID=1601095, 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.05.20146324; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

[34] https://www.census.gov/popclock/world, 2020 (data fetched on 3rd July).

[35] *Archived: WHO Timeline - COVID-19*, https://www.who.int/news-room/detail/27-04-2020-who-timeline---covid-19, 2020.

[36] S. Zhao et al. *Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak*, Int. J. Infect. Dis., 92, 214, 2020.