medRxiv preprint doi: https://doi.org/10.1101/2020.05.29.20113571; this version posted June 2, 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 The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

A Time-Dependent SEIRD Model for Forecasting the COVID-19 Transmission Dynamics

Taarak Rapolu [1], Brahmani Nutakki[2], T. Sobha Rani[3], S. Durga Bhavani[4]

[1]*taarak.rapolu@gmail.com*
[2]*brahmani3110@gmail.com*
[3]*tsrcs@uohyd.ernet.in*
[4]*sdbcs@uohyd.ernet.in*

School of Computer and Information Sciences, University of Hyderabad
Hyderabad, India

29 May 2020

Abstract

The spread of a disease caused by a virus can happen through human to human contact or could be from the environment. A mathematical model could be used to capture the dynamics of the disease spread to estimate the infections, recoveries, and deaths that may result from the disease. An estimation is crucial to make policy decisions and for the alerts for the medical emergencies that may arise. Many epidemiological models are being used to make such an estimation. One major factor that is important in the forecasts using the models is the dynamic nature of the disease spread. Unless we can come up with a way of estimating the parameters that guide this dynamic spread, the models may not give accurate forecasts. In this work, using the SEIRD model, attempts are made to forecast Infected, Recovered and Death rates of COVID-19 up to a week using an *incremental* approach. A method of optimizing the parameters of the model is also discussed thoroughly in this work. The model is evaluated using the data taken from COVID-19 India tracker [2], a crowd-sourced platform for India. The model is tested with the whole country as well as all the states and districts. The results of all the states and districts obtained from our model can be seen in [12]. Forecasts for Infected and Deaths for the whole country and the state of Maharashtra are satisfactory with an average % error rate of 3.47 and 3.60 for infected and 3.88 and 1.61 for deaths respectively. It is supposed to be a reasonable estimate which can help the governments in planning for emergencies such as ICU requirements, PPEs, hospitalizations, and so on as the infection is going to be prevalent for some time to come.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.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/10.1101/2020.05.29.20113571; this version posted June 2, 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. It is made available under a CC-BY-NC 4.0 International license .

1 Introduction

Novel Coronavirus has become a pandemic within no time from the time of its detection in Wuhan, a province of China. This has been declared as a pandemic by WHO resulting in around 5,719,320 cases worldwide, by 27th of May[3]. Around 1,54,181 were affected in India alone. With 4,373 reported deaths, the cases are rapidly rising, where Maharashtra is leading the tally. Its rapid progress has necessitated the need to come up with models to model the spread of the virus under different conditions like lockdown, hotspots, and migration of people across the places, and so on. The outbreak of novel coronavirus Covid19 and the ensuing utter chaos and the utter uncertainty caused by the pandemic in the entire world is unprecedented. More than ever before, it emphasizes the need for robust mathematical models that can guide policies to control the spread of infection and help in planning the hospital requirements such as PPEs, ventilators, etc [10].

In the literature several epidemiological models such as Susceptible, Infected, Recovered(SIR), Susceptible, Exposed, Infected, Recovered(SEIR) and Susceptible, Exposed, Infected, Recovered and Death (SEIRD), etc have been proposed to model the virus spreads like H1N1, SARS, Ebola, and others. *EpiModel* is a very useful software package, developed in 'R' language, that allows simulation of compartmental models, stochastic individual contact models, and the more recent network models [6].

2 Existing Models

A few existing pandemic models, from which the current model is derived are discussed here. The first model used to model the pandemic virus spread is the SIR model.

2.1 SIR

One of the prediction models available is the SIR or the Susceptible, Infected, Recovered Model. This popular epidemic model considers a closed population. It initially considers a small part of the population as infected. This small percentage is considered to infect R_0 others, where R_0 is the Basic Reproduction Rate[1]. The SIR model can be described as

$$\frac{\partial S}{\partial t} = -\beta \frac{SI}{N}$$

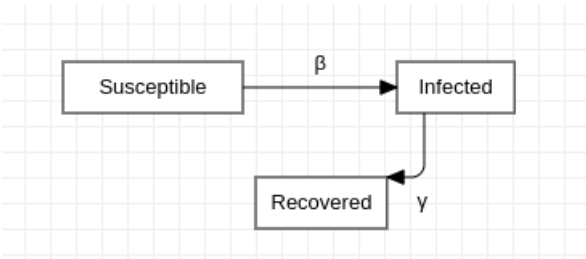
$$\frac{\partial I}{\partial t} = \beta \frac{SI}{N} - \gamma I$$

$$\frac{\partial R}{\partial t} = \gamma I$$

Here S, I, R stand for Susceptible, Infected and Recovered respectively. β is the Transmission rate and γ is the Recovery rate.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.29.20113571; this version posted June 2, 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. It is made available under a CC-BY-NC 4.0 International license .

Figure 1: SIR Model.



2.2 SEIR

The SIR model discussed here does not consider the percentage of the population who are exposed to the disease, but do not show any symptoms. When the incubation time i.e, the time elapsed before developing symptoms is significant, the SIR model will not be able to capture it. This leads to the SEIR model- Susceptible, Exposed, Infected, Recovered. The model is similar to SIR except that there is a transition from S to E instead of S to I. And the exposed percentage can also infect the Susceptibles. In a closed population, the SEIR model can be represented as

$$\frac{\partial S}{\partial t} = -\beta \frac{SI}{N}$$

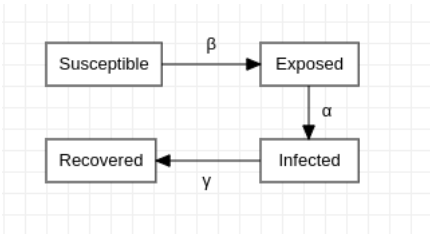
$$\frac{\partial E}{\partial t} = \beta \frac{SI}{N} - \alpha E$$

$$\frac{\partial I}{\partial t} = \alpha E - \gamma I$$

$$\frac{\partial R}{\partial t} = \gamma I$$

Here, β is the Transmission rate. α is the Incubation rate (Transition rate from E to I) while γ is the Recovery rate.

Figure 2: SEIR Model



Many extensions of the compartmental model have been proposed. These include extra compartments to denote, for example, the contaminated environmental reservoir [7], the eight compartment model of Tang et al.[9] to include

medRxiv preprint doi: https://doi.org/10.1101/2020.05.29.20113571; this version posted June 2, 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. It is made available under a CC-BY-NC 4.0 International license .

quarantined individuals and hospitalization, etc. Most of the papers in the literature consider the SEIR model with a deterministic approach by fixing the parameters to model the spread of infection.[8][11].

Yang and Wang consider the dynamic nature of the tuning parameters themselves. They consider the time-dependent parameters to model the spread of COVID 19 virus in Wuhan extending the SEIR model [7]. They have concluded that the disease is an endemic process and requires a long term plan to spread of the virus.

The model of B.Tang et al. [4] is one of the few which considers the parameters including the rate of transmission, contact rate, recovery rate as functions of time and simulate the model in order to predict the size of the infected population. They use the Markov Chain Monte Carlo (MCMC) procedure to fit the model to the data.

We observe that one of the main challenges in adopting the compartmental models lies in tuning the number of parameters involved in the model. The work in the literature fixes the parameters based on the indicators given by epidemiological experts in the scenario. The emphasis of the current work is to estimate the parameters in a dynamic manner.

3 Model formulation and Analysis

- Basic SEIR Model
  We initially ran our data against a basic SEIR model. It was observed that the results are not as accurate as expected. We were also not able to fit the Recovered rates as expected. So we extended our model to include parameter estimation- an optimized concept to estimate the parameters as per the data rather than assuming them.

- Models on Parameter Estimation

  - Grid Search
    In this model, a Grid Search is used to estimate parameters. A broad range is assigned to each of the parameters. The model then tunes the parameters to get possible values that fit the data.

    This model is computationally expensive. It takes about an hour and a half to run it on Google Colab. Once the range of parameters is narrowed down, it forecasts the rates which are more accurate than the previous model.

  - Walk forward with Grid Search
    After working on different models, it is evident that the parameters are non-stationary i.e, they change constantly. This model implements the Walk forward approach. Until the last model, the parameters are estimated for the training set as a whole. In this model, they are estimated incrementally one day at a time. The parameters obtained for the previous day are used to estimate the current day parameters.
    Though this model is relatively more accurate than the previous versions, it is extremely expensive in terms of computation. Efforts are made to extend this model using Parallel computation to no avail.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.29.20113571; this version posted June 2, 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. It is made available under a CC-BY-NC 4.0 International license .

It took about 3-4 hours on Google Colab to run this model with no results.

4 Time-Dependent SEIRD Model

In order to optimize the model, different approaches to the SEIR model are considered. This leads to the implementation of a time-dependent SEIRD model [4].
This approach resulted in accurate forecasts of Infected, Recovered, and Death rates for a week. Run time is exceptionally low, one minute at the most. This model can be represented as follows:

$$\Delta S = -\frac{\beta(t)S(t)I(t)}{N} \tag{1}$$

$$\Delta E = \frac{\beta(t)S(t)I(t)}{N} \; - \; \alpha(t)E(t) \tag{2}$$

$$\Delta I = \alpha(t)E(t) \; - \; \gamma(t)I(t) \; - \; \delta(t)I(t) \tag{3}$$

$$\Delta R = \gamma(t)I(t) \tag{4}$$

$$\Delta D = \delta(t)I(t) \tag{5}$$

From (4), we have

$$\gamma(t) = \frac{\Delta R}{I(t)} \tag{6}$$

From (5), we have

$$\delta(t) = \frac{\Delta D}{I(t)} \tag{7}$$

Using (6) and (7) in (3) yields

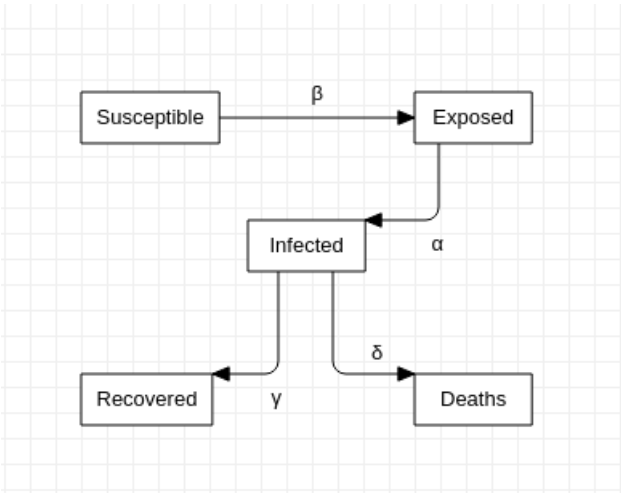
$$\alpha(t) = \frac{\Delta I + \Delta R + \Delta D}{E(t)} \tag{8}$$

Using (8) in (2) yields

$$\beta(t) = \frac{(\Delta E + \Delta I + \Delta R + \Delta D)N}{S(t)I(t)} \tag{9}$$

medRxiv preprint doi: https://doi.org/10.1101/2020.05.29.20113571; this version posted June 2, 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. It is made available under a CC-BY-NC 4.0 International license .

Figure 3: SEIRD Model



4.1 Data requirements and format to run the model

This model takes a *.csv* file with Cumulative Confirmed, Recovered, and Death values. It forecasts the results based on the parameters of the last five days of the test data. It takes the data file and 'N', the population of the area as input. Shown in Figure 4, is the format of data that is being used.

Figure 4: Data format

Day	Date	Confirmed	Recovered	Deaths
1	2020-03-24	571	40	10
2	2020-03-25	657	43	11
3	2020-03-26	730	50	16
4	2020-03-27	883	75	19
5	2020-03-28	1019	85	24
6	2020-03-29	1139	102	27

4.2 Model

In our model, the population is classified into 5 categories: the Susceptible, the Exposed, the Infected, the Recovered, and the Deaths. Parameters are estimated on a day to day basis. The data until the previous day and the current day is used to calculate the parameters of the previous day. Then the value for the Exposed population that is calculated is passed on to calculating the parameters for the next day. Unavailability of proper data on Exposed leads to improper results in the long run.

This model is run on the data [2] collected from 24th of March onward for states when the first lockdown is implemented in India. For districts, the data is available from 24th April in [2]. This model requires a minimum training of

medRxiv preprint doi: https://doi.org/10.1101/2020.05.29.20113571; this version posted June 2, 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. It is made available under a CC-BY-NC 4.0 International license .

18 days. In order to carry out forecasts, parameters for 4th day prior to the last day of training till the last day of the training are used. That is if T is the total number of training days, 1 to $(T-4)$ days data is used as training and the parameters α, β, γ, and δ are obtained for $(T-4)^{th}$ day. Then the data of the $(T-3)^{rd}$ day is added to the training and again α, β, γ and δ are computed for $(T-3)^{rd}$ day. Like this, all the parameters for $T-4, T-3, T-2, T-1, T$ days are obtained and used to forecast the values of the infections, deaths, and recoveries. Then out of these 5 forecasts, the median is taken as the final forecast value.

4.3 Algorithm

The code for the model is put up at the link given in [5]. Here, s[], e[], i[], r[], d[] are arrays to store calculated susceptible, exposed, infected, recovered and death values
alpha[], beta[], gamma[], delta[] –arrays to store calculated alpha, beta, gamma and delta values
preds are all the predictions stored in an array pred_values are stored in a stack that contains s, e, i, r, d array values in seird function.
start_date is starting date of the data taken from data.csv
T number of days in the training data taken from data.csv
incub_period is the Incubation Period

medRxiv preprint doi: https://doi.org/10.1101/2020.05.29.20113571; this version posted June 2, 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. It is made available under a CC-BY-NC 4.0 International license .

Algorithm 1 Incremental SEIRD Model

Input: N (population), data.csv
Output: Forecasts for Infected and Death cases for 5 days from the date training ends
$N \leftarrow population$
$Train \leftarrow data.csv$
$start_date \leftarrow Train.date$
$T \leftarrow Train.days$
$incub_period = 5$
for x = 1 to T **do**
  $i[x] = Train.c[x] - Train.r[x] - Train.d[x]$
  $r[x] = Train.r[x]$
  $d[x] = Train.d[x]$
end for
for t = 1 to T **do**
  parameter_estimation(t)
end for
prediction()
end

---

parameter_estimation(k)
if k = 1 **then**
  $alpha[k] = 1/incub_period$
  $gamma[k] = (r[k+1] - r[k])/i[k]$
  $gamma[k+1] = (r[k+2] - r[k+1])/i[k+1]$
  $delta[k] = (d[k+1] - d[k])/i[k]$
  $delta[k+1] = (d[k+2] - d[k+1])/i[k+1]$
  $e[k] = (i[k+1] - ((1 - gamma[k] - delta[k]) * i[k]))/alpha[k]$
  $e[k+1] = (i[k+2] - ((1 - gamma[k+1] - delta[k+1]) * i[k+1]))/alpha[k]$
  $s[k] = N - e[k] - i[k] - r[k] - d[k]$
else
  $alpha[k] = ((i[k+1] - i[k]) + (r[k+1] - r[k]) + (d[k+1] - d[k]))/e[k]$
end if
$beta[k] = (((e[k+1] - e[k]) + (i[k+1] - i[k]) + (r[k+1] - r[k]) + (d[k+1] - d[k])) * N)/(s[k] * i[k]))$
$gamma[k] = (r[k+1] - r[k])/i[k]$
$delta[k] = (d[k+1] - d[k])/i[k]$
seird(alpha[k], beta[k], gamma[k], delta[k], k, k+1)
end

---

seird(alpha, beta, gamma, delta, k, t)
$s[t] = s[k] - beta * s[k] * i[k]/N$
$e[t] = e[k] - beta * s[k] * i[k]/N - alpha * e[k]$
$e[t+1] = e[k+1] - beta * s[k+1] * i[k+1]/N - alpha * e[k+1]$
$i[t] = i[k] + alpha * e[k] - gamma * i[k] - delta * i[k]$
$r[t] = r[k] + gamma * i[k]$
$d[t] = d[k] + delta * i[k]$
pred_values = (s[t], e[t], i[t], r[t], d[t])
RETURN pred_values
end

medRxiv preprint doi: https://doi.org/10.1101/2020.05.29.20113571; this version posted June 2, 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. It is made available under a CC-BY-NC 4.0 International license .

```
prediction()
for x = T−4 to T do
  for y = x+1 to T+7 do
    pred[x][y] ← seird(alpha[x], beta[x], gamma[x], delta[x], x, y)
  end for
end for
for y = T+1 to T+7 do
  index = retrieve-first-index-of( median(pred.i[T-4][y],pred.i[T-3][y],
  pred.i[T-2][y],pred.i[T-1][y],pred.i[T][y] ) )
  Final_pred_i = pred.i[index][y]
  Final_pred_r = pred.r[index][y]
  Final_pred_d = pred.d[index][y]
end for
end
```

5 Results and Analysis

Forecasted infected and death values and plots for India, few states and districts are shown here. For these experiments, training data from 24th of March to 19th of May for states and 21st of April to 19th of May for districts is considered. Test data from 20th of May to 26th of May is used. Results are shown for the the states of Maharashtra, Gujarat, and the districts Mumbai, Thane and Pune in Maharashtra and Ahmedabad, Surat and Vadodara in Gujarat. The calculated error % for Maharashtra for Infected and Death cases over the week is 3.6 and 1.6 respectively.

Date	Actual	Forecast
2020-05-20	27590	27691
2020-05-21	28463	29529
2020-05-22	30483	31197
2020-05-23	32210	32969
2020-05-24	33997	34850
2020-05-25	35187	36847
2020-05-26	36013	38965

Date	Actual	Forecast
2020-05-20	1389	1353
2020-05-21	1453	1420
2020-05-22	1516	1491
2020-05-23	1576	1566
2020-05-24	1634	1645
2020-05-25	1694	1729
2020-05-26	1791	1817

Figure 5: Forecasted Infected and Death Cases for **Maharashtra** on test data

medRxiv preprint doi: https://doi.org/10.1101/2020.05.29.20113571; this version posted June 2, 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. It is made available under a CC-BY-NC 4.0 International license .

Date	Actual	Forecast
2020-05-20	6571	6523
2020-05-21	6649	6682
2020-05-22	6591	6854
2020-05-23	6671	7039
2020-05-24	6793	7237
2020-05-25	6944	7448
2020-05-26	6775	7673

Date	Actual	Forecast
2020-05-20	749	744
2020-05-21	773	770
2020-05-22	802	797
2020-05-23	829	824
2020-05-24	858	852
2020-05-25	888	881
2020-05-26	915	911

Figure 6: Forecasted Infected and Death Cases for **Gujarat** on test data

Date	Actual	Forecast
2020-05-20	18925	18502
2020-05-21	19916	19599
2020-05-22	20573	20774
2020-05-23	21772	22030
2020-05-24	22471	23371
2020-05-25	23863	24800
2020-05-26	23896	26323

Date	Actual	Forecast
2020-05-20	841	854
2020-05-21	882	901
2020-05-22	909	951
2020-05-23	949	1004
2020-05-24	988	1060
2020-05-25	1026	1119
2020-05-26	1065	1182

Figure 7: Forecasted Infected and Death Cases for **Mumbai** on test data

Date	Actual	Forecast
2020-05-20	3755	3977
2020-05-21	4084	4266
2020-05-22	4461	4580
2020-05-23	4816	4921
2020-05-24	4946	5292
2020-05-25	5373	5695
2020-05-26	5052	6131

Date	Actual	Forecast
2020-05-20	79	76
2020-05-21	81	87
2020-05-22	84	99
2020-05-23	85	112
2020-05-24	87	125
2020-05-25	93	140
2020-05-26	126	156

Figure 8: Forecasted Infected and Death Cases for **Thane** on test data

Date	Actual	Forecast
2020-05-20	2444	2380
2020-05-21	2557	2493
2020-05-22	2810	2617
2020-05-23	2916	2753
2020-05-24	2960	2900
2020-05-25	3063	3059
2020-05-26	3223	3230

Date	Actual	Forecast
2020-05-20	226	221
2020-05-21	234	230
2020-05-22	243	240
2020-05-23	257	250
2020-05-24	263	260
2020-05-25	274	271
2020-05-26	282	283

Figure 9: Forecasted Infected and Death Cases for **Pune** on test data

medRxiv preprint doi: https://doi.org/10.1101/2020.05.29.20113571; this version posted June 2, 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. It is made available under a CC-BY-NC 4.0 International license .

Date	Actual	Forecast
2020-05-20	5484	5442
2020-05-21	5500	5544
2020-05-22	5421	5663
2020-05-23	5468	5796
2020-05-24	5532	5945
2020-05-25	5681	6109
2020-05-26	5473	6287

Date	Actual	Forecast
2020-05-20	602	618
2020-05-21	619	650
2020-05-22	645	683
2020-05-23	669	716
2020-05-24	697	750
2020-05-25	722	786
2020-05-26	745	822

Figure 10: Forecasted Infected and Death Cases for **Ahmedabad** on test data

Date	Actual	Forecast
354 2020-05-20	354	323
2020-05-21	347	322
2020-05-22	349	321
2020-05-23	347	322
2020-05-24	362	323
2020-05-25	358	324
2020-05-26	368	326

Date	Actual	Forecast
2020-05-20	56	56
2020-05-21	57	58
2020-05-22	57	60
2020-05-23	60	62
2020-05-24	61	64
2020-05-25	62	66
2020-05-26	63	68

Figure 11: Forecasted Infected and Death Cases for **Surat** on test data

Date	Actual	Forecast
2020-05-20	231	248
2020-05-21	245	255
2020-05-22	261	262
2020-05-23	281	270
2020-05-24	304	277
2020-05-25	309	285
2020-05-26	331	293

Date	Actual	Forecast
2020-05-20	32	32
2020-05-21	35	32
2020-05-22	35	32
2020-05-23	35	32
2020-05-24	35	32
2020-05-25	35	32
2020-05-26	35	32

Figure 12: Forecasted Infected and Death Cases for **Vadodara** on test data

Date	Forecasted Infected	Forecasted Death
2020-05-28	90987	4662
2020-05-29	95033	4840
2020-05-30	99294	5025
2020-05-31	103777	5219
2020-06-01	108491	5422
2020-06-02	113445	5634
2020-06-03	118649	5855

Table 1: Forecasts of **India** for the upcoming week

medRxiv preprint doi: https://doi.org/10.1101/2020.05.29.20113571; this version posted June 2, 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. It is made available under a CC-BY-NC 4.0 International license .

5.1 Analysis

The Actual vs Forecasted plots for Infected and Death cases of Maharashtra, Gujarat, Mumbai, Pune, Thane, Ahmedabad, Surat and Vadodara from 13th May to 26th May are shown below:

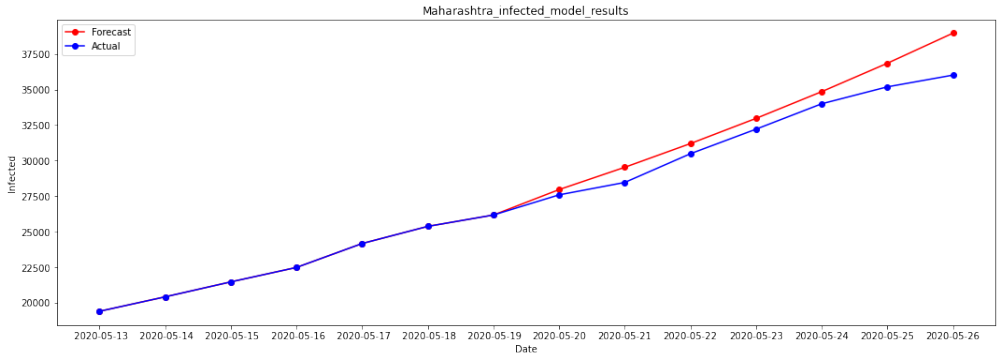


Figure 13: Plots for Actual vs Forecasted Infected Cases for **Maharashtra**; MAPE 3.60%

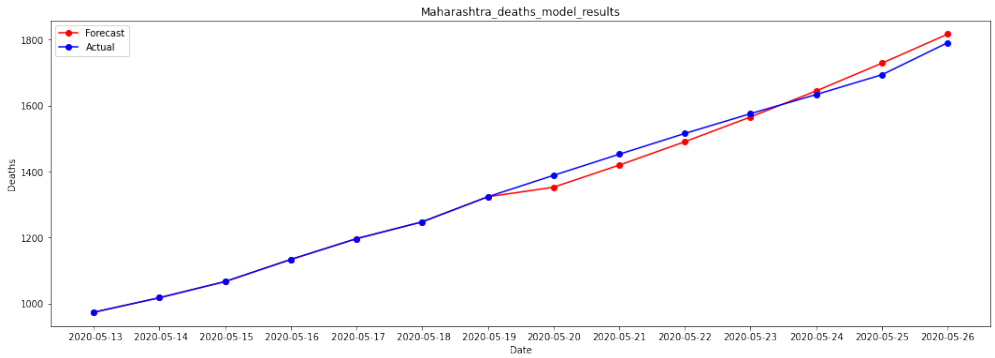


Figure 14: Plots for Actual vs Forecasted Death Cases for **Maharashtra**; MAPE 1.61%

medRxiv preprint doi: https://doi.org/10.1101/2020.05.29.20113571; this version posted June 2, 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. It is made available under a CC-BY-NC 4.0 International license .

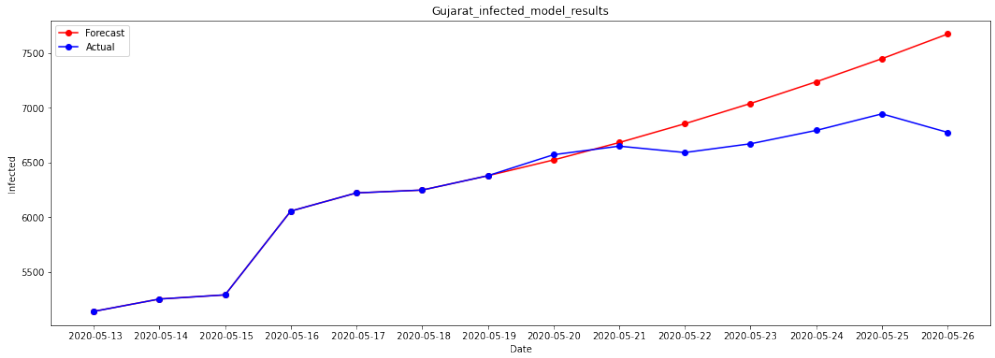


Figure 15: Plots for Actual vs Forecasted Infected Cases for **Gujarat**; MAPE 5.39%

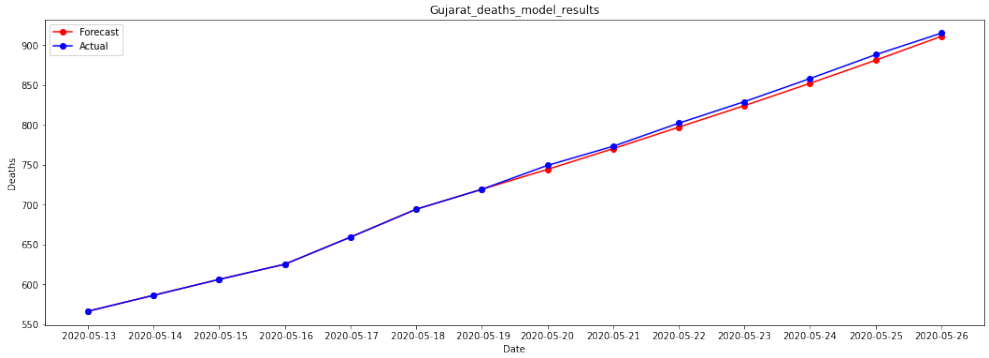


Figure 16: Plots for Actual vs Forecasted Death Cases for **Gujarat**; MAPE 0.6%

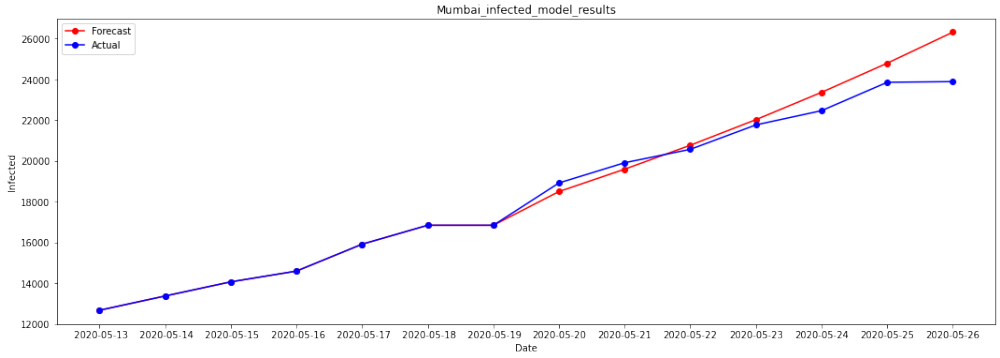


Figure 17: Plots for Actual vs Forecasted Infected Cases for **Mumbai**; MAPE 3.43%

medRxiv preprint doi: https://doi.org/10.1101/2020.05.29.20113571; this version posted June 2, 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. It is made available under a CC-BY-NC 4.0 International license .

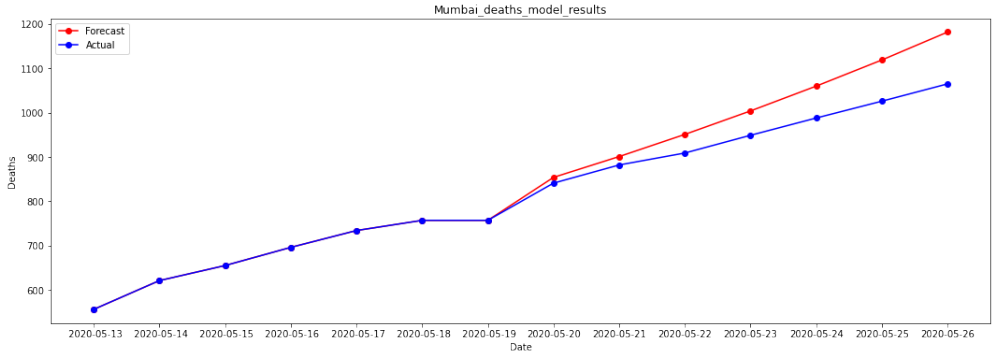


Figure 18: Plots for Actual vs Forecasted Death Cases for **Mumbai**; MAPE 5.92%

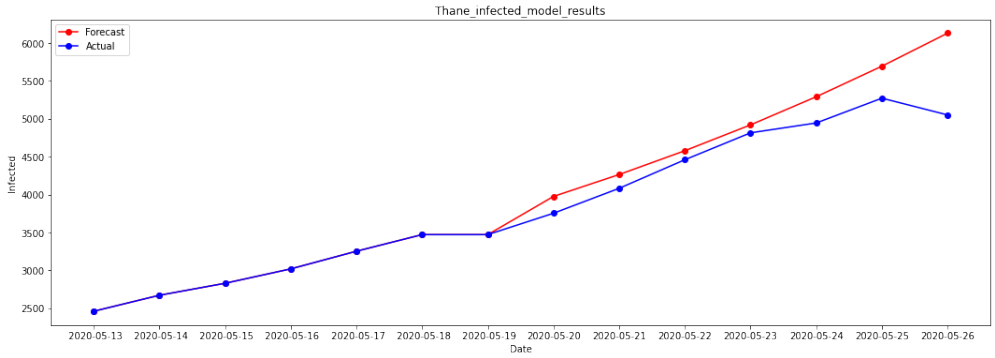


Figure 19: Plots for Actual vs Forecasted Infected Cases for **Thane** on Test Data; MAPE 7.36%

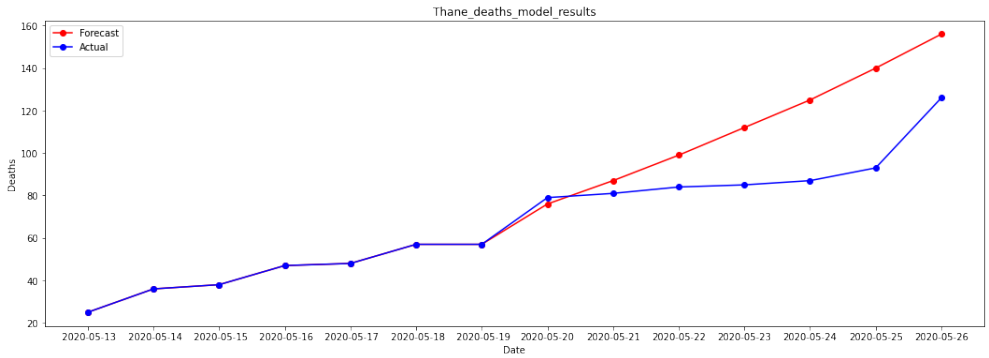


Figure 20: Plots for Actual vs Forecasted Death Cases for **Thane**; MAPE 25.55%

medRxiv preprint doi: https://doi.org/10.1101/2020.05.29.20113571; this version posted June 2, 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. It is made available under a CC-BY-NC 4.0 International license .

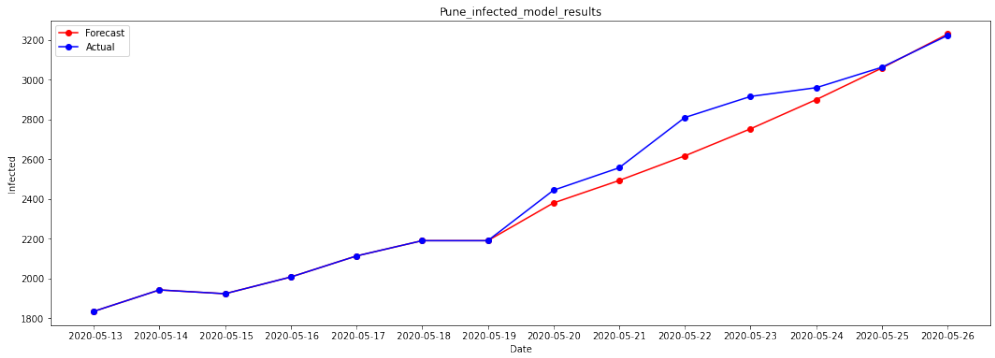


Figure 21: Plots for Actual vs Forecasted Infected Cases for **Pune**; MAPE 2.85%

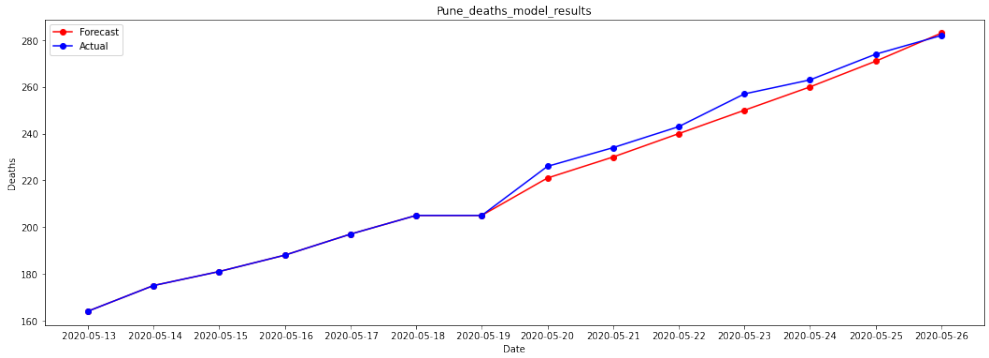


Figure 22: Plots for Actual vs Forecasted Death Cases for **Pune** on Test Data; MAPE 1.49%

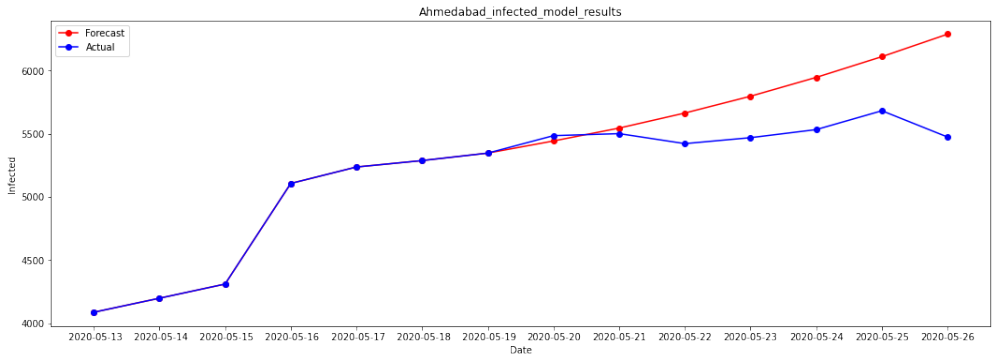


Figure 23: Plots for Actual vs Forecasted Infected Cases for **Ahmedabad**; MAPE 5.98%

medRxiv preprint doi: https://doi.org/10.1101/2020.05.29.20113571; this version posted June 2, 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. It is made available under a CC-BY-NC 4.0 International license .

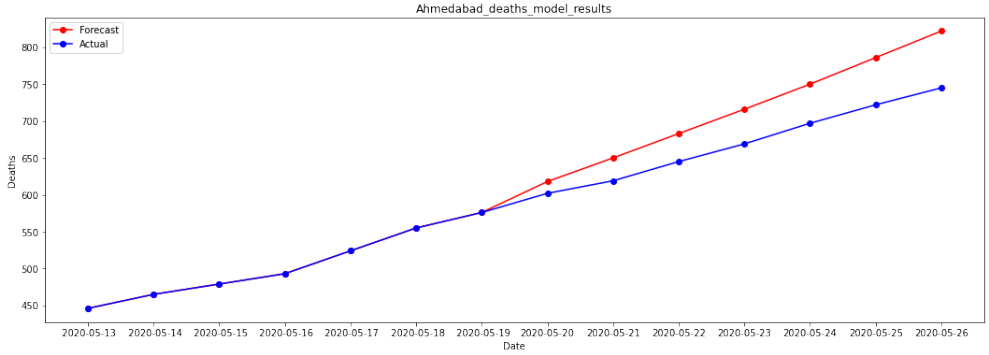


Figure 24: Plots for Actual vs Forecasted Death Cases for **Ahmedabad**; MAPE 6.76%

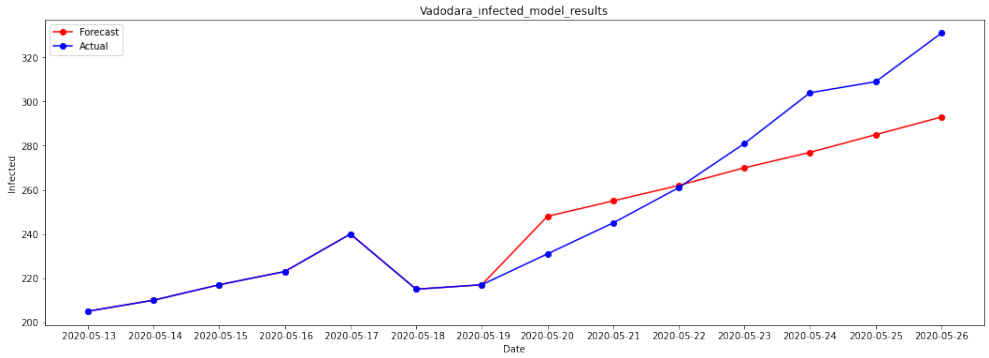


Figure 25: Plots for Actual vs Forecasted Infected Cases for **Vadodara**; MAPE 6.25%

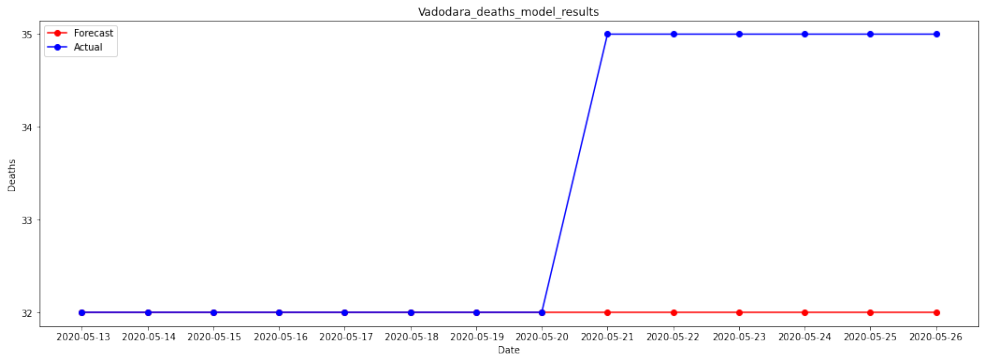


Figure 26: Plots for Actual vs Forecasted Death Cases for **Vadodara**; MAPE 7.34%

medRxiv preprint doi: https://doi.org/10.1101/2020.05.29.20113571; this version posted June 2, 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. It is made available under a CC-BY-NC 4.0 International license .

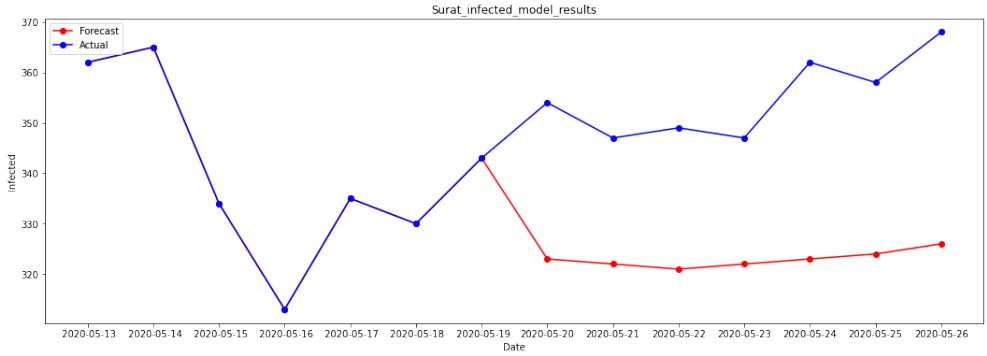


Figure 27: Plots for Actual vs Forecasted Infected Cases for **Surat**; MAPE 8.98%

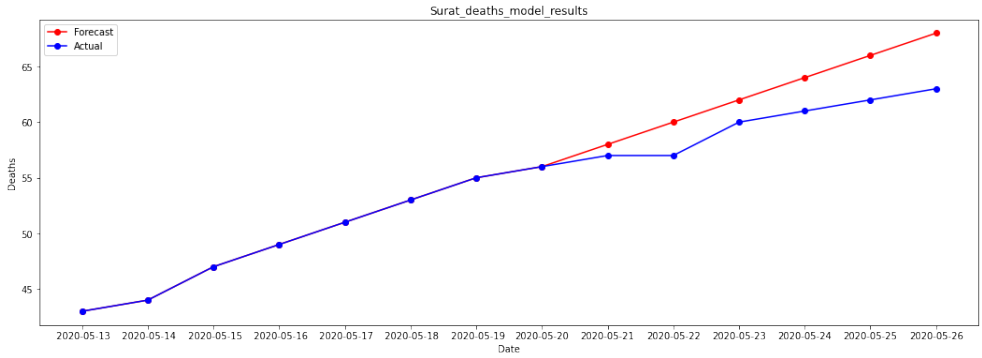


Figure 28: Plots for Actual vs Forecasted Death Cases for **Surat**; MAPE 4.23%

From the results obtained for the above mentioned states and districts, it can be summarized that the model can capture the current-trend properly. Its forecasts are based on the growth rate of the actual curve. If there is a sudden increase or decrease in the growth rate, the forecasts will be a bit inaccurate until the model stabilizes.
The average Infected and Death MAPE of 28 districts from various states in India from 20th of May to 26th of May is 6.69% and 1.76% respectively. Indivudual MAPE values can be seen in 2.

6 Limitations

- This model considers a closed population. Birth, Mortality rates, and others are not considered.
- This model is limited to short term forecasts as the parameters are dynamic and they can not be approximated to long term.
- The transmission rate for the exposed is not considered due to the uncertainty in calculating the exposed.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.29.20113571; this version posted June 2, 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. It is made available under a CC-BY-NC 4.0 International license .

7 Future Scope

- Considering population density instead of a homogeneous population to forecast accurate results.
- Transmission rate for exposed has to be tuned properly.
- Quarantine factor and others can be included to get a more detailed analysis of the situation.

8 Conclusions

Several papers are published which use SEIR to predict the results. The initial parameters are assumed to be constant in these papers. The parameters were assumed based on input from hospitals and other sources. In this model, we try to calculate these parameters from the data instead of assuming them to be constant. The goal was to forecast these results so that we can estimate and plan for Hospital equipment and Personal Protective Equipment in advance.

9 Acknowledgements

We would like to thank Dr. Bharat Sharma, Dakshas for initiating us into this study and Raakhal Rapolu, Bennett University for his inputs and useful discussions on this project.

References

[1] Sourish Das, Prediction of Covid-19 Disease Progression in India, arXiv:2004.03147 [q-bio.PE], 2020.

[2] *Covid19India* `https://api.covid19india.org/`

[3] Worldometer, `https://www.worldometers.info/coronavirus/`

[4] Yi-Cheng Chen, Ping-En Lu, Cheng-Shang Chang, Tzu-Husan Liu, A Time-dependent SIR Model for COVID-19 with Undetected Infected Persons, `http://gibbs1.ee.nthu.edu.tw/A_TIME_DEPENDENT_SIR_MODEL_FOR_COVID_19.PDF` , 2020.

[5] *Github Link* `https://github.com/Taarak9/COVID-19_Predictive_Model_SEIRD/`

[6] Samuel M. Jenness, Steven M. Goodreau, Martina Morris EpiModel: An R Package for Mathematical Modeling of Infectious Disease over Networks, J Stat Softw. doi:10.18637/jss.v084.i08, (2018).

[7] Chayu Yang and Jin Wang , A mathematical model for the novel coronavirus epidemic in Wuhan, China, Mathematical Biosciences and Engineering, Volume 17, Issue 3, 2708–2724, (2020).

medRxiv preprint doi: https://doi.org/10.1101/2020.05.29.20113571; this version posted June 2, 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. It is made available under a CC-BY-NC 4.0 International license .

[8] Zifeng Yang, Zhiqi Zeng, Ke Wang et al. Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions, J Thorac Dis, 12(3):165-174, (2020).

[9] Biao Tang, Nicola Luigi Bragazzi, Qian Li et al. An updated estimation of the risk of transmission of the novel coronavirus (2019-nCov), Infectious Disease Modelling 5, 248-255, (2020).

[10] Meher K. Prakash, Shaurya Kaushal, Soumyadeep Bhattacharya, Akshay Chandran, Aloke Kumar and Santosh Ansumali, A minimal and adaptive prediction strategy for critical resource planning in a pandemic, medRxiv preprint doi: https://doi.org/10.1101/2020.04.08.20057414, April, (2020).

[11] Gaurav Pandey, Poonam Chaudhary, Rajan Gupta, Saibal Pal, SEIR and Regression Model based COVID-19 outbreak predictions in India, arXiv:2004.00958 [q-bio.PE], (2020).

[12] `http://c19.dakshas.org:5000/`

A Infected and Death MAPE values

This section contains the Infected and Death MAPE values of 28 districts of various states from 20th of May to 26th of May.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.29.20113571; this version posted June 2, 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. It is made available under a CC-BY-NC 4.0 International license .

State	District	Infected MAPE (%)	Death MAPE (%)
Andhra Pradesh	Krishna	5.99	5.25
Andhra Pradesh	Visakhapatnam	5.28	0.00
Andhra Pradesh	Godavari	3.94	0.00
Bihar	Munger	8.53	0.00
Bihar	Supaul	9.14	0.00
Gujarat	Ahmedabad	5.98	6.77
Gujarat	Dahod	9.41	0.00
Gujarat	Surat	8.98	4.23
Gujarat	Vadodara	6.26	7.34
Haryana	Sirsa	7.14	0.00
Karnataka	Belagavi	9.74	0.00
Karnataka	Dakshina Kannada	9.39	7.14
Madhya Pradesh	Jabalpur	9.72	0.00
Maharashtra	Mumbai	3.44	5.92
Maharashtra	Pune	3.72	1.49
Rajasthan	Dausa	9.27	0.00
Rajasthan	Dholpur	9.36	0.00
Rajasthan	Karauli	2.85	0.00
Telangana	Medchal-Malkajgiri	8.16	0.00
Telangana	Nalgonda	4.76	0.00
Telangana	Suryapet	0.47	0.00
Telangana	Vikarabad	1.29	0.00
Uttar Pradesh	Auraiya	4.08	0.00
Uttar Pradesh	Ghaziabad	7.84	0.00
West Bengal	Howrah	6.11	11.29
West Bengal	Nadia	9.72	0.00
Jammu and Kashmir	Ramban	9.68	0.00
Ladakh	Leh	7.14	0.00

Table 2: MAPE values for 28 districts