Less Wrong COVID-19 Projections with Interactive Assumptions

Aditya Nagori[1,2], Raghav Awasthi[2], Vineet Joshi[2], Suryatej Reddy Vyalla[2], Akhil Jarodia[2], Chandan Gupta[2], Amogh Gulati[2], Harsh Bandhey[2], Ponnurangam Kumaraguru[2], Tavpritesh Sethi[2,3*]

1. CSIR-Institute of genomics and integrative biology, New Delhi
2. Indraprastha Institute of Information technology Delhi, New Delhi
3. All India Institute of Medical Sciences, New Delhi

*Correspondence: tavpriteshsethi@iiitd.ac.in

Abstract:

COVID-19 pandemic is an enigma with uncertainty caused by biological and health systems factors. Although many models have been developed all around the world, transparent models that allow interacting with the assumptions will become more important as we test various strategies for lockdown, testing and social interventions and enable effective policy decisions. In this paper we developed a suite of models to guide development of policies under different scenarios when the lockdown opens. These had been deployed to create an interactive dashboard called *COVision* which includes the Agent based Models (ABM) and classical compartmental models i.e Susceptible-Infected-Recovered (SIR) and Susceptible-Exposed-Infected-Recovered (SEIR) approaches. Our tool allows simulation of scenarios by changing strength of lockdown, basic reproduction number(R0), asymptomatic spread, testing rate, contact rate (Beta), recovery rate (Gamma), incubation period and starting number of cases. We optimized ABMs and classical compartmental models to fit the actual data, both of which performed well in terms of R-squared, root mean squared error (RMSE) and mean absolute percentage error (MAPE). Out of the three models in our suite, ABM was able to capture the data better than SIR and SEIR and achieved an RSQ of 92.3% for India and 89% for Maharashtra for the next 30 days. We also computed R0 using SIR and SEIR models which were found to be decreasing over the different periods of lockdown indicating the effectiveness of policies and interventions. Finally, we formulated ICU bed requirements using our best models. Our evaluation suggests that ABM models were able to capture the dynamic nature of the epidemic for a longer duration of time while classical SIR and SEIR models performed inefficiently for longer terms. The visual interactivity and ability to simulate outcomes under different parameters will allow the policymakers to make informed decisions for estimating the strength of lockdown to be implemented and testing rates. Further, our models were able to highlight the differences at state level for the parameters such as R0 and contact rates and hence can be applied for state specific decision making. An interactive dashboard http://covision.tavlab.iiitd.edu.in have been hosted as a web-server for the war level monitoring of the covid19 pandemic in India in public domain

Keywords: Covid19, projections, Agent Based models, Epidemic Models, Compartmental models.

Introduction

The recent novel coronavirus disease, COVID-19, has shook the world by storm. A communicable disease with no existing medical treatment, the World Health Organisation (WHO) has already declared it a pandemic. Many countries have adopted the social distancing norms to limit the spread of this virus and are trying their best to contain the virus by nation wide lockdowns, testing, marking disease hotspots and contact tracing. Similarly, India had also implemented a 21 day lockdown upto 15th April 2020 to assess the situation which has now been extended to 3rd May 2020. Since prolonged lockdowns, including closing of organizations, schools and educational institutions and other important places can cause depression in economy and may cause mayhem among people in shortage of necessary living supplies, including food, shelter and hygienic environments, it is both necessary and imperative to model the flow of the infection so that the policy makers can assess the situation and consequently take constructive steps to save lives as well as the economy.

Since the start of the epidemic, many new studies and papers have come to the surface, aiming to model the spread of the disease in many countries, including India. While many have tried to give projections using machine and deep learning techniques, these tend to overfit the data [1], and are unable to accomodate the overtime changes in the parameters such as contact rate due to government interventions, such as lockdowns, increased testing, etc. However, most of the papers have used classical population level epidemiological models, mainly Susceptible - Infected - Recovered (SIR) model, and some other models which are just some minor modifications of the same [1-2]. Even though these models are reliable for modern modelling of an epidemic or pandemic, they cannot contain the changes in the parameters, as is the case with the machine learning techniques. Even though the researchers have tried to take the effects of social distancing and lockdown while setting some parameters like contact rate or reproduction number, the numbers predicted using such dynamical models have not been reliable, as in the case of [3], which projects a sharp decrease after 25th March, which was not the case.

In this paper, we found that the classical epidemiological models cannot be used to model the disease effectively and reliably, and promote the applications of agent based models (ABMs). Many studies had used agent based models instead of equation based models in the past to study epidemics. Venkataraman et al. [4] argued against using compartmental models as they fail to capture individual level dynamics as well as social heterogeneity in real world epidemics. Instead, they used an agent based modelling mechanism for forecasting the 2014-2015 Ebola epidemic which generated close to real world scenarios for longer term forecasts. Tuomisto et al. [5] used ABMs to identify potentially destructive

policies that may have been undertaken by the government due to the shortcomings as explained accompanying various model strategies and their underlying assumptions.

METHODS

1. Dataset

Datasets were obtained from the ministry of health and family welfare website[6]. The statewise data were downloaded on the dailybasis and collated in a time-series format. Data were in the cumulative form; we extracted the infected counts by subtracting the cured/recovered and migrated counts from the total to get the infected counts. Exposure counts were obtained by taking the infected time series at a lag of incubation period. Population data were taken from the official Indian government website[7]. The population data is used to represent susceptible populations in the models.

We had done our projection modeling in every state of India but here we will only present the results of the top 6 states which were most affected till 20th May (see Figure 1).

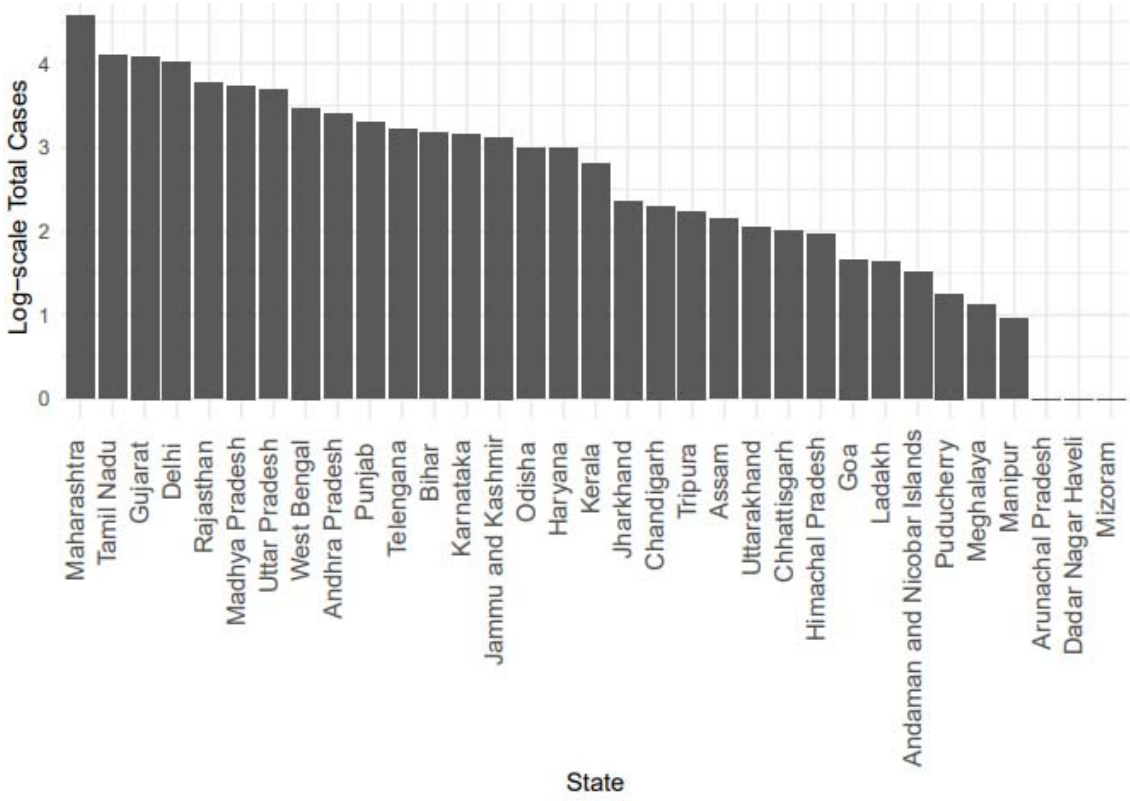

Figure 1. Barplot $\log_{10}$ of total number of covid-19 cases until 20th May.

2. Agent Based Model: The agent-based model was used to predict the statewise daily and total new cases by simulating the spread on a population of agents. The agent-based model is built on top of an SEIR model with the entire population divided among four primary states of susceptible, exposed, infected & recovered. In order to incorporate the

effect of interventions like lockdown and testing rate, we also introduced two additional states, ‘Under Lockdown’ and ‘Tested & Quarantined’. On initialization, a specified number of agents were randomly spread over an area with few agents exposed to the virus and the rest in susceptible state. After an incubation period, the earlier exposed agents move to the infected state and start infecting more based on the provided contact rate value. The infected agents were further symptomatic & asymptomatic in nature, with the symptomatic agent having higher chances of being tested and isolated. Figure 2 shows the model’s state chart that guides an agent's behaviour based on the state they were currently in.

The ABM was developed in Anylogic, a multi-agent simulation software [8], because of its ability to allow highly customizable and scalable models.

S.No.	Parameter	Description	Default Values
1.	Population	Number of susceptible agents.	State-wise (Supplementary Material)
2.	Area	Dimension of the grid environment on which agents are spread.	Square root of population.
3.	Incubation Period	Number of days an agent stays in the exposed state before moving to the infected one.	7-14 days.
4.	Percentage Symptomatic	Percentage of infected population showing symptoms.	20%
5.	Testing Rate	Percentage of agents tested.	40% - 60% of susceptible agents
6.	Contact Rate	Average number of agents an infected agent meets each day.	2-7
7.	Initial Infected	Number of infected agents before the model starts.	2
8.	Lockdown Rate	Percentage of susceptible population that is under total lockdown.	50-70%
9.	Recovery Rate	Percentage of infected population that gets recovered.	30%
10.	Mortality Rate	Percentage of infected population that dies.	0.5%
11.	Average Illness Duration	Number of days after which an agent recovers or dies.	21-42 days.

Table1: Scenarios/Parameters in Agent Based Models.

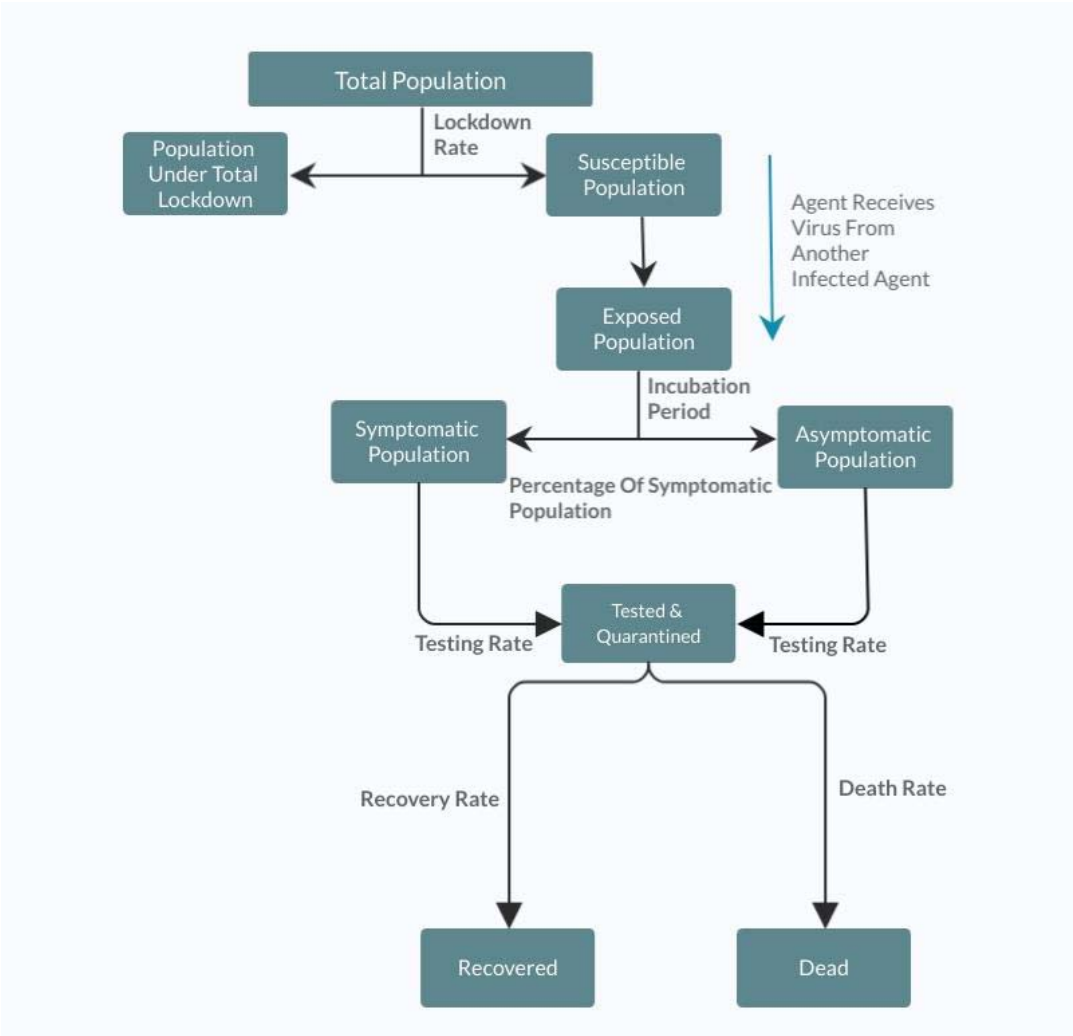

Figure 2. Flowchart showing the possible states of an agent and effect of interventions on them.

Different combinations of parameters were evaluated using R^2 Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE) of Predicted new cases vs actual new cases. The optimal model was selected from a grid search of combinations of parameters for future projections (Supplementary Table1).

3. Classical Compartmental models & their optimisation: We used the standard compartmental ecological models called susceptible-infection-recovery (SIR) and susceptible-exposure-infection-recovery (SEIR). The model parameters involved the

Contact rate (beta), Recovery rate (gamma) and Delta were optimised using ODE Solver and squared error minimisation. We numerically solved the ordinary differential equations in the SIR & SEIR (Supplementary text2) using Runge-Kutta method as a first order ordinary differential equation [9]. Assumptions for model parameters for recovery rate were taken from covid literature; we found that the average time taken to recover by an individual with non-chronic illness is 2-3 weeks and for a chronic illness individual 3-6 weeks as mentioned in the WHO report [10]. We therefore used the range to optimise the recovery rate. The incubation period was taken in the range of 1-14. The exposure time-series were taken at a lag of the incubation period.

4. Dashboard :

Scenarios: A user friendly, interactive dashboard is developed in order to allow policy makers, researchers and the general public to get a better sense of the situation on a daily basis. We have currently incorporated optimised ABM and SIR, SEIR models along with various parameters. These models were optimised and updated with new data and user inputs in fractions of seconds.

Interactivity: The dashboard allows users to set R0, Recovery rate and incubation period in a literature backed range for classical compartmental models. For ABM, the dashboard provides a variety of simulated scenarios including the best fitted and the unlock models. These scenarios provide a sense on the spread of the disease and help in studying the effects of interventions. “Under the hood” option is added to read more about the methods used to produce these scenarios. The interface is shown in supplementary Figure1.

5. Hospital beds reporting : For ICU beds requirement forecasting we assumed an average of 9% [11] infected people needed critical care which varies in the range of 6%-12% [12,13,14] using projection outcome of infected count from different models we forecasted the Requirement.

Results :

1. Quantitative evaluation of Models

To identify the best suite of forecasting models we quantitatively evaluated the compartmental and Agent based models in different scenarios using Rsq, Rmse and Mape. We observed that in the longer duration ABM has better performance than compartmental models (Figure 4), but for the shorter duration compartmental (SIR and SEIR models) models had comparable results.

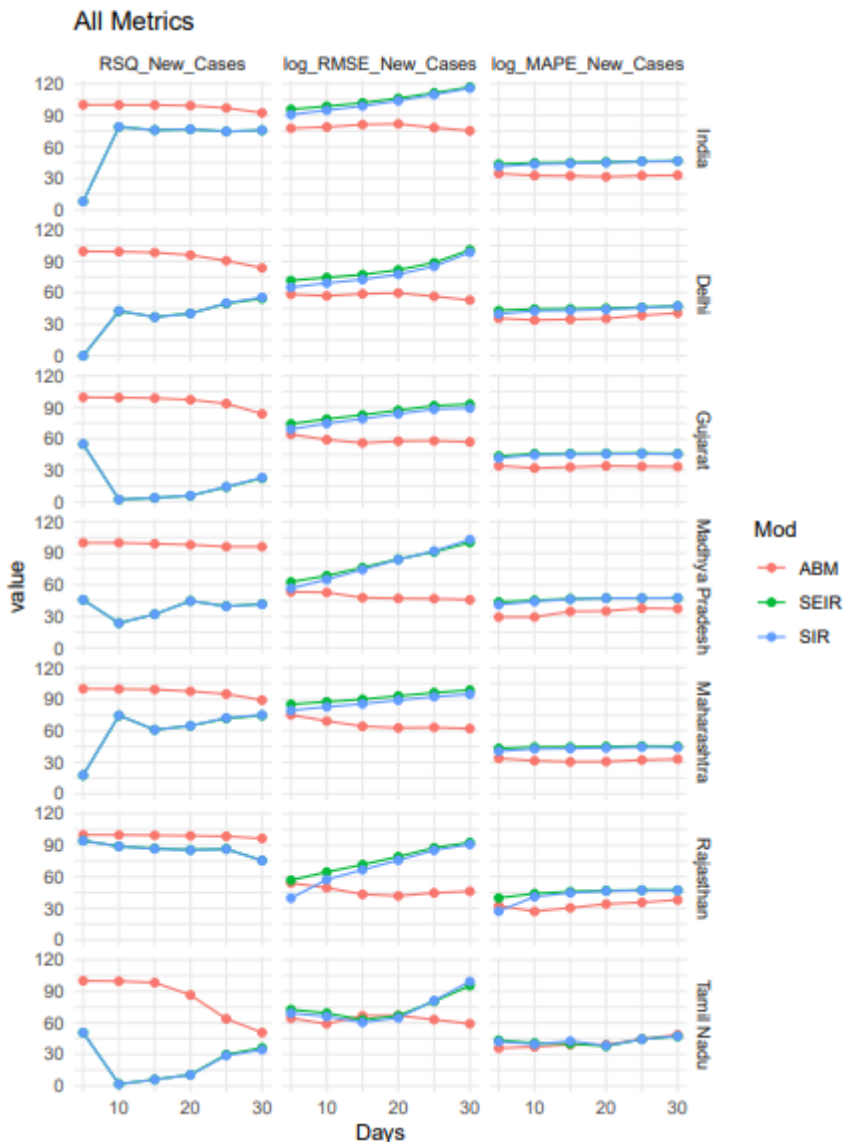

Figure 3 : *Comparison of State wise R Squared from different models, state wise log Root Mean Square Error of daily new cases from different models, log Mean Absolute Percentage Error of daily new cases from different models. (Log base 1.1 is taken to match the scales for plotting) (Plot shows evaluation for India, Delhi, Gujrat, Madhya Pradesh, Maharashtra, Rajasthan, Tamil Nadu)*

2. State-wise scenario based (ABM) projections :

Every state is different and holds geo-political heterogeneity. The agent based models (ABM) provide a variety of parameters that can all be tuned independently in order to capture the dynamic nature of the epidemic and heterogeneity. Table 1 shows the list of all the parameters that are customizable in our model along with the values of all parameters that were kept the same for all the states. Parameters such as incubation rate, average illness duration, recovery rate, death rate and percent symptomatic[15] were taken from

literature whereas the model was optimized around the fixed values of area and initial infection. The remaining 4 parameters, susceptible population, lockdown rate, contact rate and testing rate were empirically changed to simulate the scenarios that provide the best fit.

Optimal fitting scenarios for different states are given below:

- **Delhi:** 10 Lakh susceptible population with 65% lockdown rate, 50% testing rate and contact rate of 4. Lockdown starts from 23rd March 2020.
- **Gujarat:** 10 Lakh susceptible population with 65% lockdown rate, 43% testing rate and contact rate of 4. Lockdown starts from 23rd March 2020.
- **Maharashtra:** 30 Lakh susceptible population with 58% lockdown rate, 50% testing rate and contact rate of 2 which changes post 30th March to 5.
- **Madhya Pradesh:**10 Lakh susceptible population with 40% lockdown rate, 60% testing rate and contact rate of 3. Lockdown starts from 23rd March 2020.
- **Rajasthan:** 10 Lakh susceptible population with 60% lockdown rate, 40% testing rate and contact rate of 3. Lockdown starts from 23rd March 2020.
- **Tamil Nadu:** 10 Lakh susceptible population with 60% lockdown rate, 40% testing rate and contact rate of 4. Lockdown starts from 23rd March 2020.
- **India:** 1 Crore susceptible population with 67% lockdown rate, 47 % testing rate and contact rate of 2 which changes post 30th March to 7.

Figure4: *ABM based Projections of daily new cases in India and states i.e Delhi, Gujrat, Madhya Pradesh, Maharashtra, Rajasthan, Tamil Nadu.*

3. Unlock1: Lifting of the lockdown

Indian government announced that from June 8th, 2020 [16], more services will be resumed in a planned manner. This is being done in leu to subvert the economic crisis in the country. But this might increase the contact rate and hence the exposure of the larger percentage of the population to infection. We simulated these scenarios in ABM models for future predictions of new cases for the next 30 days. We observed a hockey stick point in the projections around the 5th to 8th June as shown in figure 5. These results were obtained by putting parameters same as the best fit till 31th May, later than that, lockdown was set to 33%, testing 50% and contact rate were tuned according to state-level government guidelines on restriction and relaxations. These simulations are available on our dashboard and can be tested for different parameters as well.

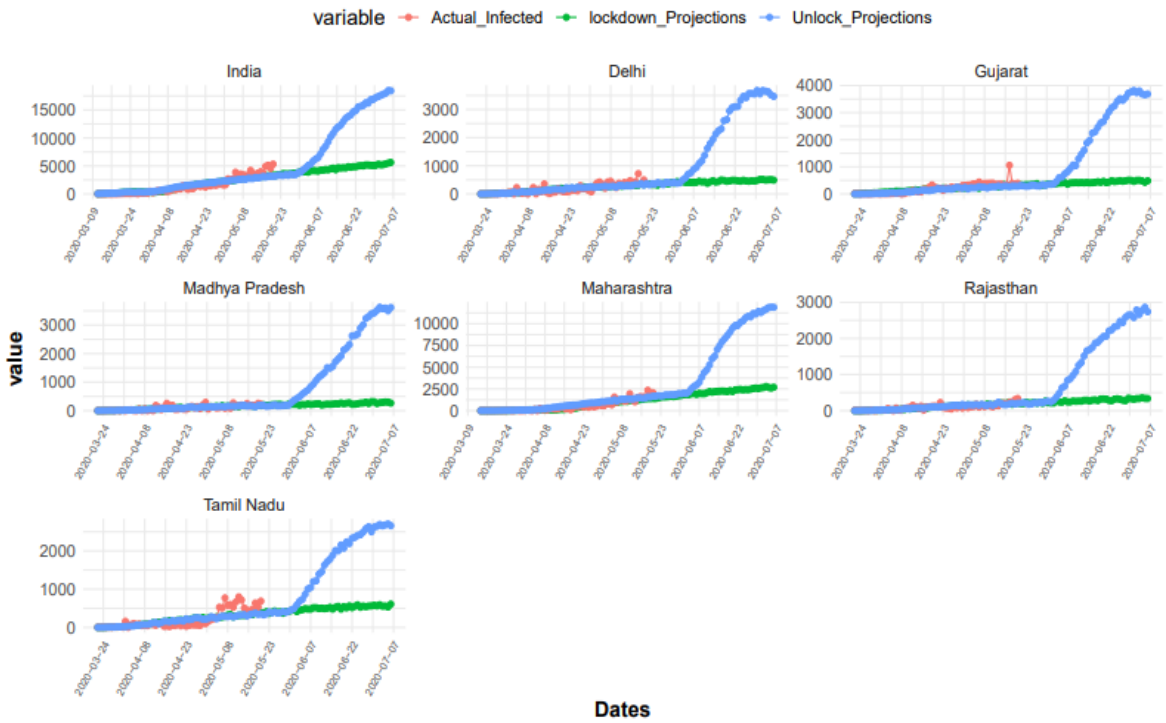

Figure 5: *Unlock 1 and Lockdown differences of new cases using ABM shown for India, Delhi, Gujrat, Madhya Pradesh, Maharashtra, Rajasthan, Tamil Nadu.*

4. Decreasing trend of R0 revealed the effectiveness of interventions :

Since R0 is a good indicator to assess the policies and interventions, we trained our models on different stages of pandemic and calcuted R0, we found that the trend of R0 was decreasing (See Figure 6) for every model, which strongly suggests the effectiveness of the interventions. On the basis of R0 .

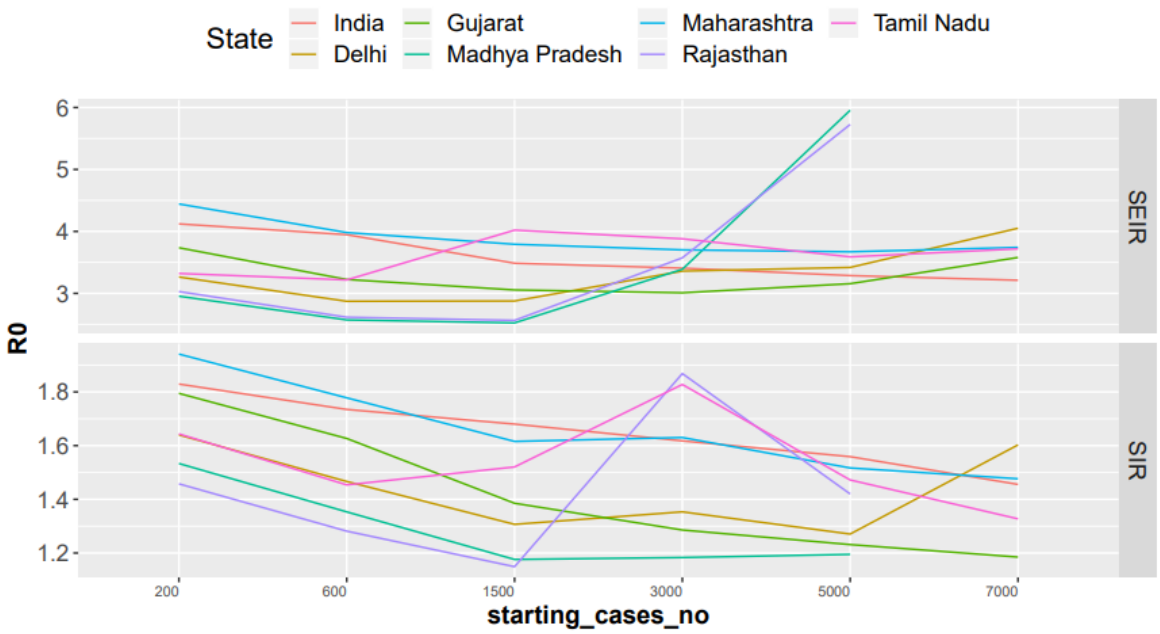

Figure 6: *State Wise Variation of R0 obtained from SIR and SEIR model in different scenarios of pandemic*

5. Statewise ICU Beds Forecasting:

Assuming that 9% [6%,12%] of infected people require critical care, from the Agent based model, we estimated the ICU beds requirement for every state from 21st may 2020 to 6 july, 2020 (Figure7). In Maharashtra, we found that the daily requirement of ICU Beds for Covid critical patients reached upto 1,000 on 28 june 2020 with a maximum value of 1,069 [713,1426] on 6th july. Daily ICU Beds requirement for Delhi, Gujrat, Madhya Pradesh was found almost similar with maximum value 311, 331, 325 respectively on 6th july 2020.

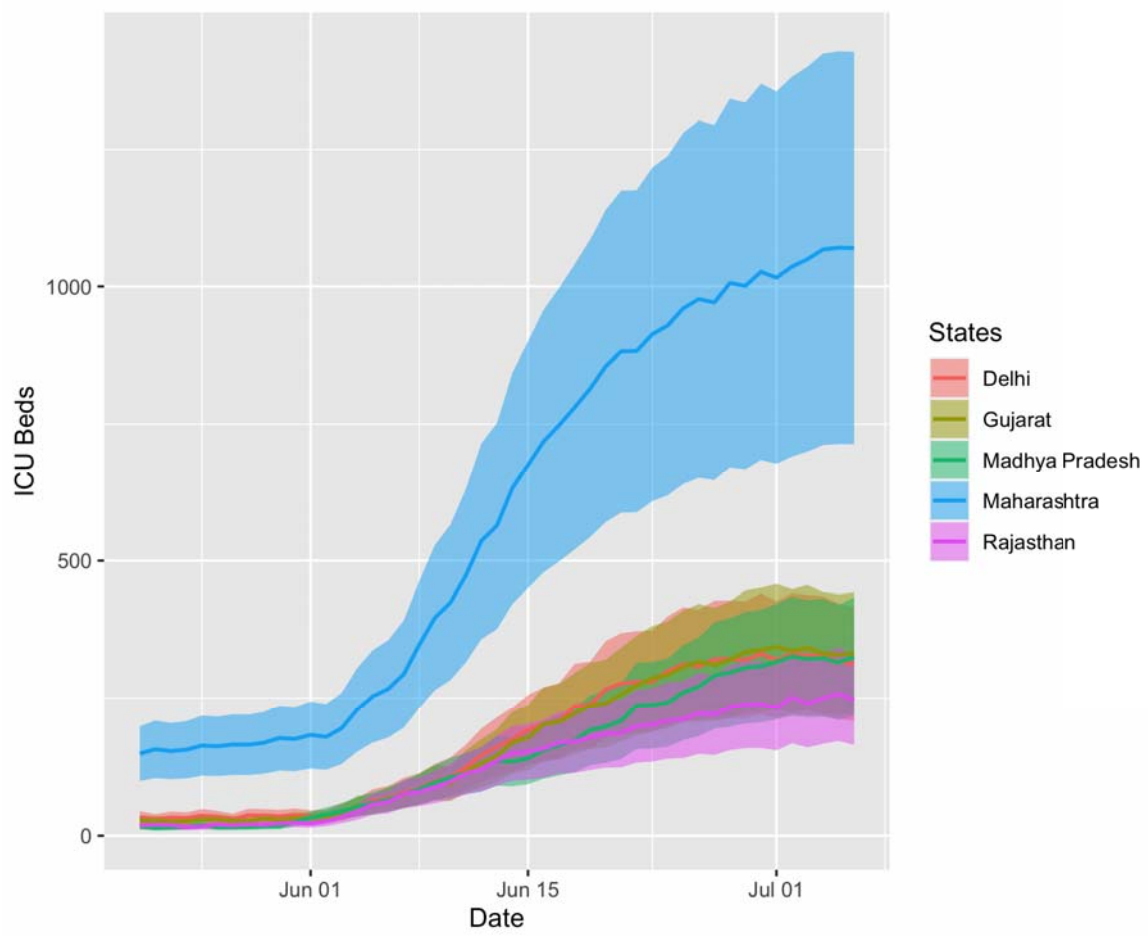

Figure 7: *Cumulative ICU beds requirement forecasting from 21 may to 6 july using Agent Based Model*

Discussion:

Exponential spread of covid-19 is becoming a threat for human survival and until no treatment is discovered social distancing is the only vaccine. India is the second most populated country in the world [17] and in high population density areas social distancing becomes difficult. One possible solution to tackle this is lockdown. Indian government implemented lockdown very timely and effectively and that controlled the spread of disease significantly. Since the spread of covid is impacting multiple sectors like industry, education, agriculture , health and many more, effective interventions are required which need the projection of spread , in this study we made available different models with different scenarios to project disease with an interactive dashboard. The ecological models were trained on the data downloaded from the ministry of health and welfare website. Building these models involves parameter optimisation, a slight error in a given parameter can lead to error propagation in the final projections. We had therefore used a minimum parameter optimisation. We did an extensive literature review to get a reasonable range of parameters. The WHO reports the recovery interval for healthy and

chronic ill subjects as 2-3 weeks and 4-6 weeks respectively. [10]. The recovery rate in models were therefore optimised in this range. The incubation period has also been reported that it might take 1-14 [10] days to show the symptoms, thus, incubation period in 1-14 days is a good guess for the optimisation. We have less clue about contact rates so we allowed our model to optimise in a wider range for the contact rates. We looked at the literature and found out that our contact rates were in close proximity to what has been reported recently [2]. But these are classical models making them robust by involving the dynamicity of human interventions and behaviors is little trickier, thus, we involved the Agent based Models (ABM) which allowed us to incorporate the wider interventional measures. Lockdown in India had a gradual implementation as we saw migrant worker movement and religious gatherings increase the spread during the first lockdown starting from 24th March, 2020. The ability of ABMs to customize interventions and incorporate such scenarios allows them to capture this behaviour in a better way and study how lockdown and testing might affect the upcoming trend. Of course we kept a number of assumptions, ABMs assumes that the population under lockdown is not susceptible towards the disease. Moreover, testing rate affects the asymptomatic population more as symptomatic agents are more likely to be tested at the first place. It should be noted that there are many models being currently studied and deployed related to Covid 19 in India and other countries. As noted by Holmdahl and Bucklee [18], the selectivity for a particular model depends upon various factors, like the time period for the application, the underlying assumptions, and the parameters used and their estimations. It goes on to further argue to check for the specificity of the model, and identify the type of data used to fit the model. The most important parameter, the contact rate [18], needs to be correctly estimated and must take into account various lockdown and reopening scenarios, if the model is to be used in the long run. Hence, it is extremely important to specify the assumptions that were made while constructing a model, which are able to accommodate for rapidly changing crises and select the model which complements our task.

Acknowledgements:
This work was partially supported by the Wellcome Trust/DBT India Alliance Fellowship IA/CPHE/14/1/501504 awarded to Tavpritesh Sethi and the Center for Artificial Intelligence at IIIT-Delhi. We thank AnyLogic North America LLC for providing an extended free trial for conducting AnyLogic simulations. We also thank CSIR India for supporting Aditya Nagori and UGC India for supporting Raghav Awasthi. We acknowledge Rakesh Lodha from All India Institute of Medical Sciences, New Delhi, for his valuable inputs and Bhavika Rana for her design input for the dashboard frontend.

Authors Contributions:
Concept TP, PK, AN, RA Compartmental models: AN, RA, CG, TP ABM Models: VJ, AJ, TP Dashboard Frontend: SR, AG, BR, PK Dashboard Backend AN, VJ, AJ, HB

Statistical Analysis: AN, RA, VJ Data Extraction & Curation: AN, HB Manuscript: AN, RA,VJ,CG,TP

References:

1. Jay Naresh Dhanwant and V. Ramanathan. “Forecasting COVID 19 growth in India using the Susceptible-Infected-Recovered (S.I.R) model.”

2. Rajesh Ranjan. “Predictions For COVID-19 Outbreak In India Using Epidemiological Models.”
3. Singh and Adhikari. “Age-structured impact of social distancing on the COVID-19 epidemic in India.”

4. Venkatramanan et al “Using data-driven agent-based models for forecasting emerging infectious diseases.” ScienceDirect

5. Tuotino et al “An agent-based epidemic model REINA for COVID-19 to identify destructive policies” preprint MedaRxiv.

6. https://www.mohfw.gov.in/

7. https://data.gov.in/

8. Emrich, Stefan, Sergej Suslov, and Florian Judex. "Fully agent based modelling of epidemic spread using AnyLogic." *Proc. eurosim*. 2007.

9. Soetaert, Karline ER, Thomas Petzoldt, and R. Woodrow Setzer. "Solving differential equations in R: package deSolve." *Journal of Statistical Software* 33 (2010).

10. WHO report https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf#:~:text=Using%20available%20preliminary%20data%2C,severe%20or%20critical%20disease.

11. JM, Rodriguez Llanes, M. G. Pedersen, and M. Meneghini. "Confronting COVID-19: Surging critical care capacity in Italy." (2020).

12. Xie, Tong, Guan, Du, Haibu, Arthur S,. “Critical care crisis and some recommendations during Covid - 19 epidemic in China.” *Intensive Care Medicine, SpringerLink* (2020)

13. Phua, Weng, Ling, Egi, Ling, Divatia, “Intensive care management of coronavirus diseases 2019 (COVID 19) : challenges and recommendations”

14. Vincent, Jean-Louis, and Fabio S. Taccone. "Understanding pathways to death in patients with COVID-19." *The Lancet Respiratory Medicine* (2020).

15. https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200306-sitrep-46-covid-19.pdf?sfvrsn=96b04adf_4

16. Unlock 1.0 Guidelines: https://mha.gov.in/sites/default/files/MHADOLrDt_3052020.pdf

17. Population Data https://www.worldometers.info/world-population/population-by-country/

18. Holmdahl et al “Wrong but Useful — What Covid-19 Epidemiologic Models Can and Cannot Tell Us” The New England Journal of Medicine (2020)

19. Beretta, Edoardo, and Yasuhiro Takeuchi. "Global stability of an SIR epidemic model with time delays." *Journal of mathematical biology* 33.3 (1995): 250-260.

20. Li, Michael Y., et al. "Global dynamics of a SEIR model with varying total population size." *Mathematical biosciences* 160.2 (1999): 191-213.

21. Zhu, Caillou, et al. "Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization." *ACM Transactions on Mathematical Software (TOMS)* 23.4 (1997): 550-560.

22. Kermack and McKendrick. “A contribution to the mathematical theory of epidemics.”