An epidemic model SIPHERD and its application for prediction of the spread of COVID-19 infection in India

Ashutosh Mahajan [a,*], Namitha A Sivadas [a], Ravi Solanki [b]

[a] *Vellore Institute of Technology, Vellore 632 014, India*
[b] *Visvesvaraya National Institute of Technology, Nagpur 440 010, India*

ABSTRACT

Originating from Wuhan, China, in late 2019, and with a gradual spread in the last few months, COVID-19 has become a pandemic crossing 9 million confirmed positive cases and 450 thousand deaths. India is not only an overpopulated country but has a high population density as well, and at present, a high-risk nation where COVID-19 infection can go out of control. In this paper, we employ a compartmental epidemic model SIPHERD for COVID-19 and predict the total number of confirmed, active and death cases, and daily new cases. We analyze the impact of lockdown and the number of tests conducted per day on the prediction and bring out the scenarios in which the infection can be controlled faster. Our findings indicate that increasing the tests per day at a rapid pace (10k per day increase), stringent measures on social-distancing for the coming months and strict lockdown in the month of July all have a significant impact on the disease spread.

© 2020 Elsevier Ltd. All rights reserved.

1. Introduction

The outbreak of novel Corona Virus Disease (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), originated from a wet market in Wuhan, China, is now widespread in the world and has severely affected many counties including India. The first case of COVID-19 in India was reported on January, 30 2020 at Thrissur, Kerala, in a student who had returned from China. India has reached the fourth position in the world in the number of confirmed COVID-19 cases and presently has 449,613 confirmed cases and 14,162 deaths as of June 23, 2020, which is really an alarming situation. Social distancing is the best method for mitigating this pandemic until an effective medicine or vaccine is invented [1]. The first nationwide lockdown was ordered by the Prime Minister of India on March 24, for 21 days and it was further extended till May 3, by relaxing certain essential fields . Later, though the lockdown period is extended to June 30, the freedom is given to the states to impose restrictions assessing the situations in the respective states.

Mathematical models for the epidemic have a major role to make predictions of the transmission dynamics of the disease and thus assist the authority to take necessary actions for the containment.

Several epidemic models are already reported in the literature, however, the COVID-19 is different type of infection showing certain special characteristics. The transmission of the disease from the persons who are infected without any symptoms (asymptomatic cases) is one of the special characteristics of this disease [3] and to be considered in the modeling. Also, the disease can be spread from the infected who is in the incubation period [4]. In this study, we present a new mathematical model named SIPHERD, incorporating the aforementioned characteristics of the COVID-19.

Many mathematical models for the infectious disease spread are reported in the literature, and the classical and widely used is SIR model described in [5]. An approximate spatial epidemiological model of the COVID-19, is initially proposed in [6] in which the spread of the disease within and between the countries is analysed. The infected and undetected cases and the spread of COVID-19 from those persons are incorporated in [7] but this study is associated with only China.

A modified compartmental SIR model is discussed in [8]. In this study, the total population of the country is divided into eight compartments, and this work is carried out only for Italy, and also, the model does not undertake the purely Asymptomatic cases of infected. A different compartmental model SEIR [9] predicts the dynamics of the transmission of the COVID-19 for certain countries, and the impact of quarantine of the infected persons are also studied in it. Another improved SIR model is depicted in [10], and the time dependency of the parameters of the SIR model is also

examined. Prediction of transmission of COVID-19 using curve fitting algorithms are reported in [11,12] and [13].

A stochastic mathematical model is proposed in [14] to analyse the impact of COVID-19 on the healthcare system in India. Assessment of the preventive measures of COVID-19 such as lockdown and prediction of its spread in India is studied in [2,15]. The SEIR epidemic and regression model is extended for predicting and evaluating the transmission of COVID-19 in India [16]. Progression of the epidemic in India is also determined using the mathematical modelling in [17] and [18].

In India, complete lockdown is limited to the containment zones and hotspots from June 1, on-wards. The forbidden activities in the places outside the containment zones are re-opened in a phased manner with the conditions to follow the standard operating guidelines given by the Health Ministry of India. Interstate and intra-state travel excluding the containment zones is allowed in the present situation without any pass/permission. Religious places, restaurants, shopping malls and hospitality facilities are allowed to open whereas the educational institutions, entertainment zones, international air travel and railway services remain prohibited as of June 23. Despite all these restrictions, the infection is growing exponentially, and policymakers need to consider different ways for the containment of the disease. The most hit cities in India by COVID-19 pandemic are Mumbai, Delhi and Chennai with a collective population of around 40 million.

In this paper, we bring out different possible ways for better control of the infection spread. We employ an improved mathematical model SIPHERD [19] for the COVID-19 pandemic embedding the purely asymptomatic infected cases and the transmission of the disease from them. The model simulations bring out the efficacy of different ways for the containment, by predicting the total number of active and confirmed cases, total deaths, and daily new positive cases considering various social distancing/lockdown conditions and the number of tests done per day.

2. Methods

We model the evolution of the COVID-19 disease by dividing the population into different categories as listed below which is described in detail in [19,20]. The SIPHERD model equations for the defined entities (S,I,P,H,E,R,D) are a set of coupled ordinary differential Eqs. (1) to (7). The population is divided into different categories, as Susceptible (S), Exposed (E), Symptomatic (I), Purely Asymptomatic (P), Hospitalized or Quarantined (H), Recovered (R) and Deceased (D).

$$\frac{dS}{dt} = -S(\alpha E + \beta I + \gamma P + \delta H) \tag{1}$$

$$\frac{dE}{dt} = S(\alpha E + \beta I + \gamma P + \delta H) - (\mu + \xi_I + \xi_P)E \tag{2}$$

$$\frac{dH}{dt} = \mu(E + P) + \nu I - \sigma H(t - t_R) - \tau H(t - t_D) \tag{3}$$

$$\frac{dI}{dt} = \xi_I E - (\nu + \omega)I \tag{4}$$

$$\frac{dP}{dt} = \xi_P E - (\mu + \eta)P \tag{5}$$

$$\frac{dR}{dt} = \omega I + \eta P + \sigma H(t - t_R) \tag{6}$$

$$\frac{dD}{dt} = \tau H(t - t_D) \tag{7}$$

where, t_R and t_D are the delay associated with the recovery and death respectively with respect to active cases H. The various parameters are shown in Fig. 1 and their optimized values for India
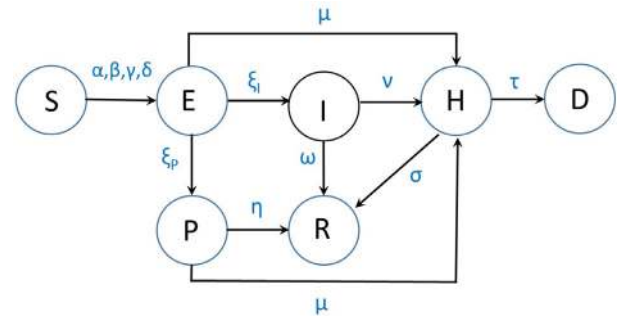

Fig. 1. The SIPHERD Model.

Table 1
Parameters values for India.

Parameter	India	Value
Population	N	1380M
Starting Date of Simulation	T_0	Mar 4
Number of days till Lockdown	T_{LD}	20,56
Trans. rate, E to S before LD	α_1	0.34
Trans. rate, E to S after LD	α_2, α_3	0.15,0.11
Trans. rate, I to S before LD	β_1	0.19
Trans. rate, I to S after LD	β_2, β_3	0.09,0.12
Trans. rate, P to S before LD	γ_1	0.34
Trans. rate, P to S after LD	$\gamma_2 \, \gamma_3$	0.15,0.11
Trans. rate, H to S	δ	6E-3
Conversion rate, E to I	ξ_I	0.2
Conversion rate, E to P	ξ_P	0.072
Detection Probability of E,P	μ_0	5.7e-3
Detection Prob Coefficient	μ_1	1.9E-6
Detection Probability of I	ν_0	0.055
Detection Prob Coefficient	ν_1	1.9E-6
Home Recovery rate of Sym.	ω	0.07
Home Recovery rate of Asy.	η	0.1
Recovery Rate	σ	0.065
Recovery delay	t_R	10
Mortality delay	t_D	1
Total Initially Infected	$E_0 + P_0 + I_0$	100

COVID-19 data are listed in Table 1. All fractions add up to unity that can also be seen from summing the above equations.

$$\frac{d}{dt}(S + I + P + H + E + R + D) = 0 \tag{8}$$

The detection of the Asymptomatic and Symptomatic cases can be taken dependent on the number of tests done per day (T_{PD}).

$$\nu = \nu_0 + \nu_1 T_{PD} \tag{9}$$

$$\mu = \mu_0 + \mu_1 T_{PD} \tag{10}$$

where, μ_0, μ_1, ν_0, and ν_1 are positive constants. The effectiveness of the tests increases if contact tracing is performed. So far in India contact tracing is performed well, we assume that the increased tests are also performed on the suspects more carefully and the detection probability increases with increasing tests linearly. Since the severe cases are going to approach for the tests, one component of the detection probability is not taken dependent on the number of tests.

Recovery of Asymptomatic cases is taken faster than the Symptomatic cases. The total confirmed cases are the addition of the active cases, extinct cases, and a part of the recovered that were detected. This can be written as

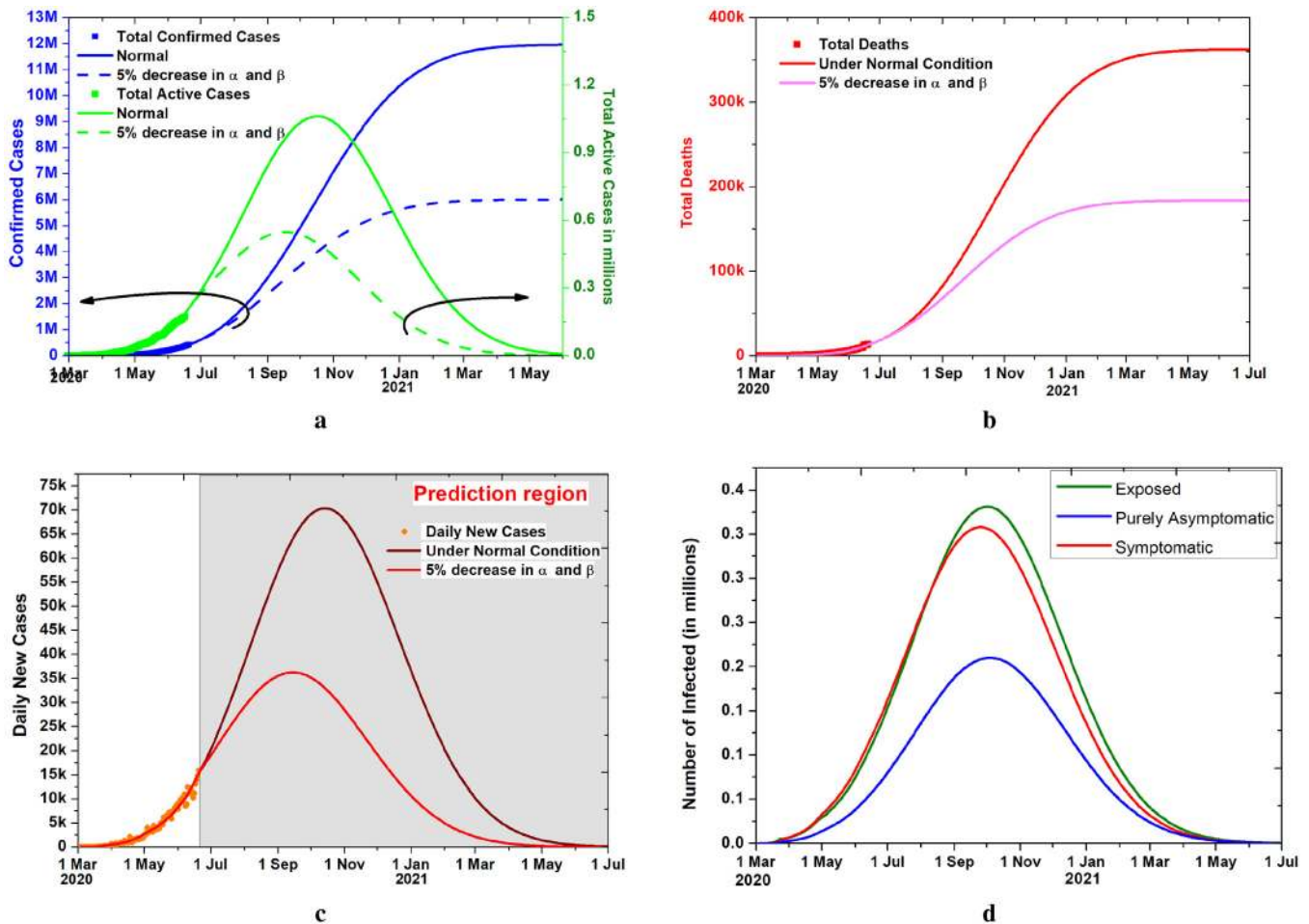$$T_c(t) = H(t) + D(t) + \int^t \sigma H(\tau')d\tau' \tag{11}$$

Fig. 2. Model predictions for Stringent Measures on Social Distancing. The present social distancing conditions are taken improved further with additional measures such that transmission rates comes down by 5%. **a,b.** Model prediction for the confirmed, active and total deaths with and without additional control measures. **c** Comparison of daily new cases in "stringent measures conditions" with the current trend. **d** The evolution of the undetected Exposed, Symptomatic, and purely Asymptomatic cases for 2.5k increase in TPD with social distancing conditions kept the same.

3. Numerical implementation and simulation

The set of coupled ordinary differential equations for the model can be solved numerically for a given set of parameters and initial values. It is however important that the parameters are determined accurately so that the model demonstrates the real situation of the infection spread. We take into account the data sets of the total number of confirmed cases, active cases, cumulative deaths and tests done per day, and find the model parameters that generates the best possible match between the actual data and model. For this purpose, a cost function is written in terms of errors between the actual and solver data sets. The minimizer of the cost gives the optimized set of parameters. The model and the optimization codes are implemented in MATLAB.

4. Results and discussion

The number of total positive or confirmed cases, present active cases and deaths are collected from [21,22], and the number of tests per day from [23], which is plotted in Supplementary Material Fig. S1 a. The parameters determined by our model for COVID-19 spread in India are listed in Table 1 and the simulation data from the model is compared with the actual data in Supplementary Material Fig. S6. The parameter values related to the characteristics of the disease are discussed in more detail in [19].

4.1. Predictions for India

For the available data till June 20, 2020, we run the model for 109 days i.e. from March 4 till June 20, to extract parameters listed in Table 1, and then with the extracted parameters, the model is run for 500 days starting from March 4. The current increase in tests per day is around 2.5k, we assume the same trend for the test per day and take the current value of transmission rate to generate the simulated data for the prediction.

4.2. Effect of lockdown

Two scenarios are considered for lockdown and social distancing conditions. One possible scenario is that the conditions are kept the same, and the second one is that they are made stricter by taking into account some measures after June 20, such that the transmission rates α and β decrease by 5%. These measures can include restrictions on travel, large gathering of people for social events, distribution of low-cost masks and hand sanitizers in hot spots.

Test per day assumed to be increased by 2.5k, which is close to the current trend and taken saturated at 1 million for both the scenarios. The mortality rate is calculated from the active cases and daily new deaths data and is fitted in steps with values 3.17E-3, 5.37E-3. 3.14E-3, 2.29E-3, 2.79E-3 on March 4, to June 20, as seen
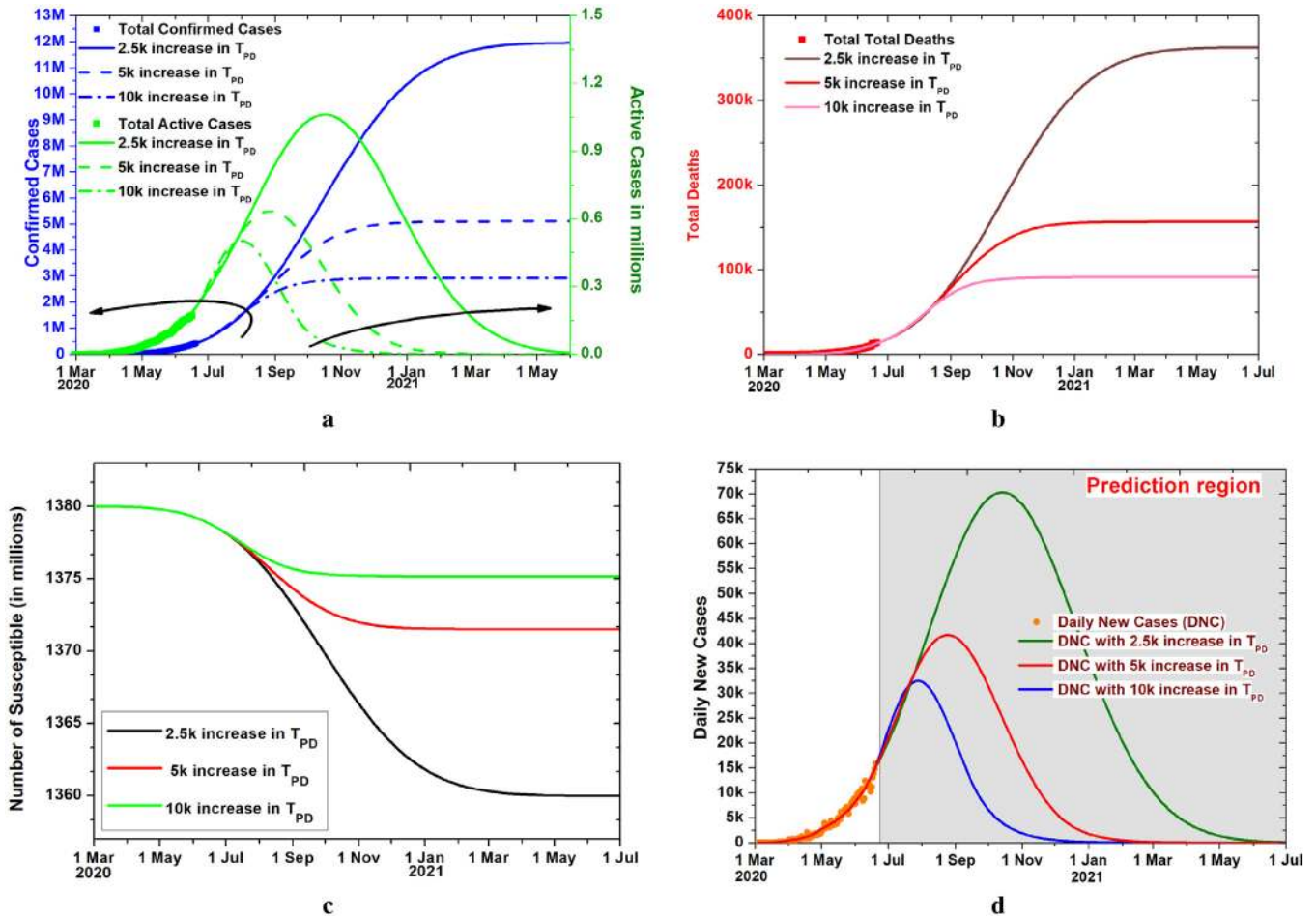
Fig. 3. **Effect of Increase in Tests. a,b.** Model prediction for the confirmed, active and total deaths cases for different testing conditions. **c.** Evolution of the Susceptible. **d.** The daily new cases for 2.5k, 5k and 10k increase per day in TPD.
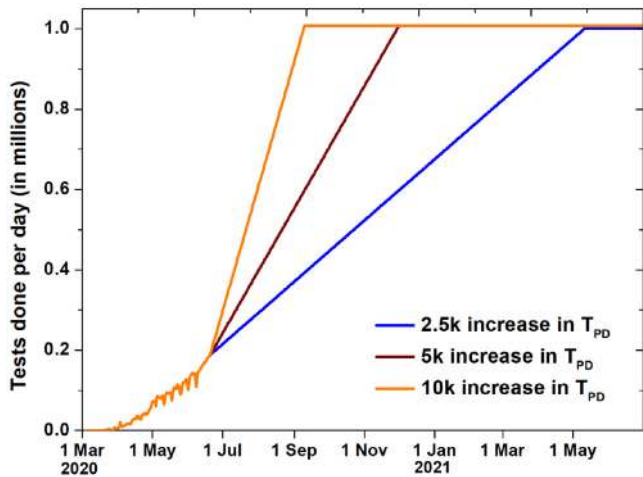

Fig. 4. Comparison of 2.5k, 5k and 10k increase in Tests per day (T_{PD}) in future saturated at 1 million.

in Supplementary Material Fig. S1.b. For the future, mortality rate is taken improved to a fixed value 2e-3.

A comparison of the predictions for the two scenarios i.e. with and without the stringent measures is plotted in Fig. 2. Reproduction number variation with time and evolution of the undetected infected cases can be seen in Supplementary Material Fig. S2. The

total number of reported cases is predicted to be around 12 million. This number appears very high. However, compared to the USA reported cases 2.4 million, the number is reasonable, given the fact that India's population is roughly four times higher.

We also study one more possible scenario in which a total lockdown is imposed for July 2020. The rate of transmission of infection is going to decrease in the imposed lockdown, and we take the α and β values to decrease by 10% from the current value. The prediction with 2.5k increase in tests per day and saturation at 1 million tests, is compared in Fig. 5a,b and in Fig. 5d, we plot the prediction for the daily new cases. Detailed plots and evolution of infected can be seen in Supplementary Material Fig. S5 for this condition of stricter lockdown for July and relaxed after that.

4.3. Effect of increase in testing

We compare the effect of testing on the prediction of total, active and death cases in Fig. 3 a,b. Total, Active and extinct cases are plotted for the coming months if tests per day are increased by 2.5k, 5k and 10k per day after June 21, and saturated at 1 Million as seen in Fig. 4. The factor by which an increase in testing can contain the infection can be estimated. The daily new positive cases data and the prediction for the 2.5k and 10k increase in tests per day are plotted in Fig. 3d.

The initial reproduction number 3.2 is seen go down to 1.5 after the imposition of lockdown on 20th day i.e. March 24. After the second lockdown on 56th day i.e. April 27, it has reached 1.32. It shows a downward trend further and goes below 1 on October
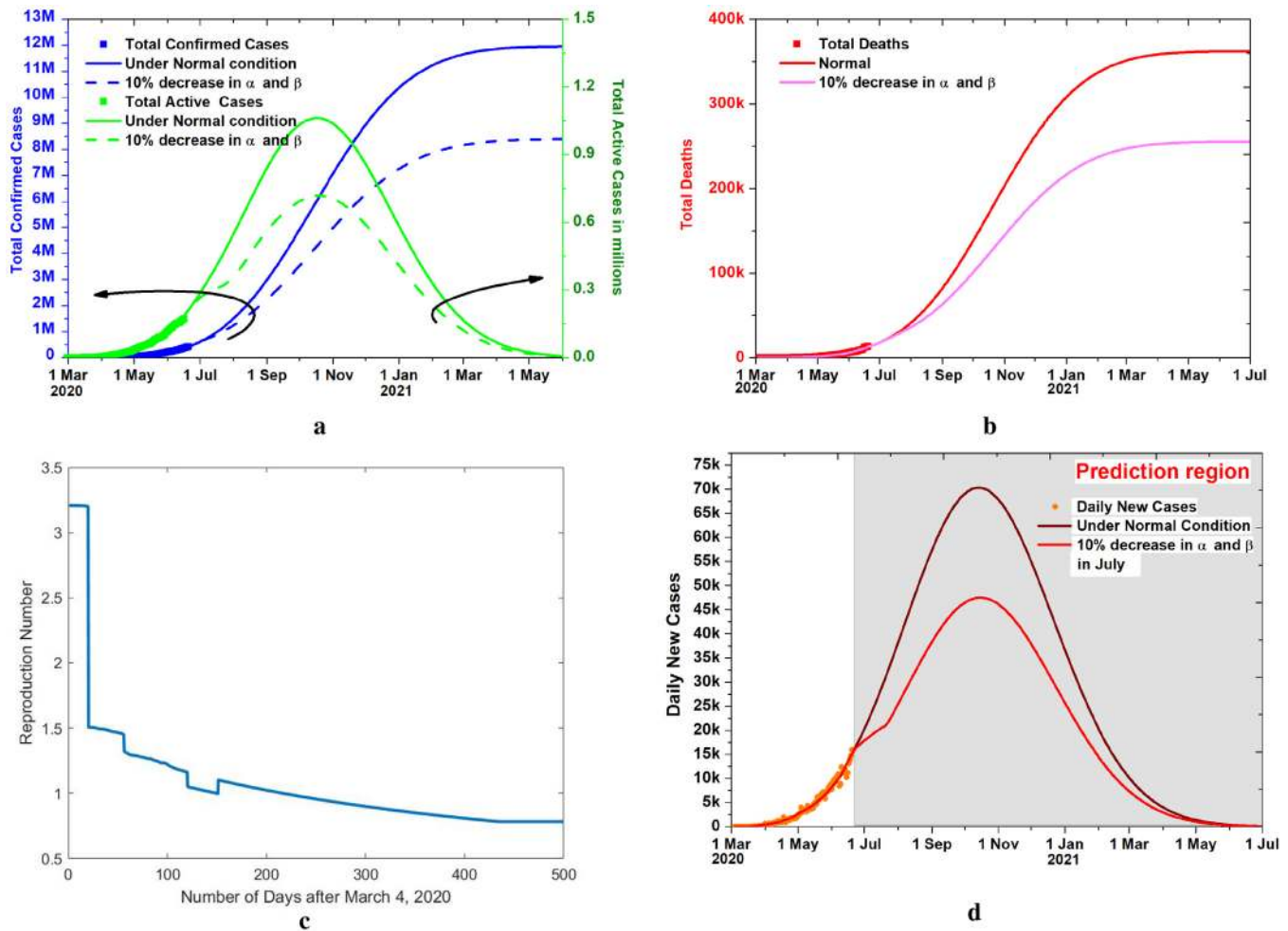
Fig. 5. **Effect of lockdown in the month of July. a,b.** Effect of one month Lockdown on the total, active and extinct cases, with 2.5k increase in Tests per day (T_{PD}) saturated at 1 million.**c.** Reproduction number variation with number of days. Effect of lockdown can be seen as a dip in reproduction number. **d.** The daily new cases with and without lockdown in July.

5,2020. If tests per day are increased 10k per day, then the reproduction number comes below one on July, 16, which can also be seen in Fig. 5c.

The reproduction number is seen to go below one on July 30, if the transmission rate decreases by 10% due to lockdown in July. The initial basic reproduction number 3.2 is in the range of the mean reported value [24,25].

A sensitivity study is carried out for the different parameters, as seen in Supplementary Material Fig. S4 and S5. The parameters are increased and decreased by 5% from the optimized values to see the changes in the outcomes. It can be seen in Supplementary Material Fig. S5 that the model prediction are most sensitive to the transmission rates α_3, β_3 and γ_3. Just a 5% change in these estimated parameters gives a huge change in total, active and extinct cases. Supplementary Material Fig. S4 show that ν_0 is a sensitive parameter. It indicates that if infected people spontaneously approach the test and quarantine themselves, it will give better control. The fear of testing due to the subsequent isolated quarantine if tested positive should be alleviated.

5. Conclusion

SIPHERD model is employed for COVID-19 spread that considers purely Asymptomatic category of infected cases in addition to the Symptomatic, and the disease spread by the Exposed. The effect of lockdown on the rates of transmission of infection and the

influence of tests per day on detection rates has been incorporated in the model. With the current trend, total infections would be 12 million when disease ends and can lead to 362k total deaths. Our findings suggest that increasing the number of tests at 10k per day with highly efficient contact tracing, rather than the current 2.5k per day rise leads to a reduction of 9 million reported cases and reduction of 270k in total extinct cases. In the absence of a vaccine, the infection can last long till the end of this year and number of deaths could be around 350k if social distancing conditions and increase in tests remain at the current trend.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.chaos.2020.110156.

References

[1] Singh R, Adhikari R. Age-structured impact of social distancing on the covid-19 epidemic in india. arXiv preprint arXiv:200312055 2020.

[2] Sardar T, Nadim SS, Chattopadhyay J. Assessment of 21 days lockdown effect in some states and overall India: a predictive mathematical study on covid-19 outbreak. arXiv preprint arXiv:200403487 2020.

[3] Mizumoto K, Kagaya K, Zarebski A, Chowell G. Estimating the asymptomatic proportion of coronavirus disease 2019 (covid-19) cases on board the diamond princess cruise ship, yokohama, japan, 2020. Eurosurveillance 2020;25(10):2000180.

[4] Rothe C, Schunk M, Sothmann P, Bretzel G, Froeschl G, Wallrauch C, et al. Transmission of 2019-ncov infection from an asymptomatic contact in germany. N top N Engl J Med 2020;382(10):970–1.

[5] Kermack WO, McKendrick AG. A contribution to the mathematical theory of epidemics. Proc R Soc London Ser A, Contain Paper Math Phys Character 1927;115(772):700–21.

[6] Ivorra B, Ramos AM. Application of the be-codis mathematical model to forecast the international spread of the 2019–20 wuhan coronavirus outbreak. ResearchGate 2020;31460. https://doi.org/10.13140/RG.2.2

[7] Ivorra B, Ferrández M, Vela-Pérez M, Ramos A. Mathematical modeling of the spread of the coronavirus disease 2019 (covid-19) considering its particular characteristics. the case of china. Tech. Rep.. Technical report, MOMAT, 03 2020; 2020.

[8] Giordano G, Blanchini F, Bruno R, Colaneri P, Di Filippo A, Di Matteo A, et al. A sidarthe model of covid-19 epidemic in italy. arXiv preprint arXiv:200309861 2020.

[9] Efimov D, Ushirobira R. On an interval prediction of covid-19 development based on a seir epidemic model; 2020.

[10] Chen Y-C, Lu P-E, Chang C-S, Liu T. A time-dependent sir model for covid-19 with undetectable infected persons. arXiv preprint arXiv:200300122 2020.

[11] Zhao S, Lin Q, Ran J, Musa SS, Yang G, Wang W, et al. Preliminary estimation of the basic reproduction number of novel coronavirus (2019-ncov) in china, from 2019 to 2020: adata-driven analysis in the early phase of the outbreak. Int J Infect Diseas 2020;92:214–17.

[12] Zeng T, Zhang Y, Li Z, Liu X, Qiu B. Predictions of 2019-ncov transmission ending via comprehensive methods. arXiv preprint arXiv:200204945 2020.

[13] Hu Z, Ge Q, Jin L, Xiong M. Artificial intelligence forecasting of covid-19 in china. arXiv preprint arXiv:200207112 2020.

[14] Chatterjee K, Chatterjee K, Kumar A, Shankar S. Healthcare impact of covid-19 epidemic in India: astochastic mathematical model. Med J Armed Force India 2020.

[15] Tomar A, Gupta N. Prediction for the spread of covid-19 in India and effectiveness of preventive measures. Sci Total Environ 2020:138762.

[16] Pandey G, Chaudhary P, Gupta R, Pal S. Seir and regression model based covid-19 outbreak predictions in India. arXiv preprint arXiv:200400958 2020.

[17] Ranjan R. Predictions for covid-19 outbreak in India using epidemiological models. medRxiv 2020.

[18] Das S. Prediction of covid-19 disease progression in India: under the effect of national lockdown. arXiv preprint arXiv:200403147 2020.

[19] Mahajan A, Solanki R, Sivadas N. Estimation of undetected symptomatic and asymptomatic cases of covid-19 infection and prediction of its spread in USA. medRxiv 2020. https://www.medrxiv.org/content/early/2020/06/23/2020.06.21.20136580.full.pdf

[20] Mahajan A, Solanki R, Namitha AS. An epidemic model sipherd and its application for prediction of the spread of covid-19 infection for India and USA. arXiv preprint:200500921 https://arxiv.org/abs/2005.00921.

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

[22] https://www.covid19india.org.

[23] https://ourworldindata.org/grapher/full-list-covid-19-tests-per-day.

[24] Liu Y, Gayle AA, Wilder-Smith A, Rocklöv J. The reproductive number of covid-19 is higher compared to sars coronavirus. J Travel Med 2020.

[25] Read JM, Bridgen JR, Cummings DA, Ho A, Jewell CP. Novel coronavirus 2019-ncov: early estimation of epidemiological parameters and epidemic predictions. MedRxiv 2020.