Anticipating the novel coronavirus disease (COVID-19) pandemic

Taranjot Kaur[1], Sukanta Sarkar[1], Sourangsu Chowdhury[2], Sudipta Kumar Sinha[3], Mohit Kumar Jolly[4], and Partha Sharathi Dutta[1]*

[1] *Department of Mathematics, Indian Institute of Technology Ropar, Punjab, India 140 001.*
[2] *Department of Atmospheric Chemistry, Max Planck Institute for Chemistry, 55128 Mainz, Germany.*
[3] *Department of Chemistry, Indian Institute of Technology Ropar, Punjab, India 140 001.*
[4] *Centre for BioSystems Science & Engineering, Indian Institute of Science, Bengaluru, India 560 012*

(Dated: April 8, 2020; Received :to be included by reviewer)

The infectious novel coronavirus disease COVID-19 outbreak has been declared as a public health emergency of international concern, and later as an epidemic. To date, this outbreak has infected more than one million people and killed over fifty thousand people across the world. In most countries, the COVID-19 incidence curve rises sharply in a short span of time, suggesting a transition from a disease free (or low-burden disease) equilibrium state to a sustained infected (or high-burden disease) state. Such a transition from one stable state to another state in a relatively short span of time is often termed as a critical transition. Critical transitions can be, in general, successfully forecasted using many statistical measures such as return rate, variance and lag-1 autocorrelation. Here, we report an empirical test of this forecasting on the COVID-19 data sets for nine countries including India, China and the United States. For most of the data sets, an increase in autocorrelation and a decrease in return rate predict the onset of a critical transition. Our analysis suggests two key features in predicting the COVID-19 incidence curve for a specific country: a) the timing of strict social distancing and/or lockdown interventions implemented, and b) the fraction of a nation's population being affected by COVID-19 at the time of implementation of these interventions. Further, using satellite data of nitrogen dioxide which is emitted predominantly as a result of anthropogenic activities, as an indicator of lockdown policy, we find that in countries where the lockdown was implemented early and strictly have been successful in reducing the extent of transmission of the virus. These results hold important implications for designing effective strategies to control the spread of infectious pandemics.

I. INTRODUCTION

The outbreak of the COVID-19 disease caused by a novel pathogenic coronavirus (SARS-CoV-2) - which initiated in Wuhan, China in December 2019 - is a global challenge for the healthcare, economy and the society [1]. The World Health Organization (WHO) assessed the massive epidemics of the disease (COVID-19) and declared it as a Public Health Emergency of International Concern (PHEIC) [2]. Since the Wuhan outbreak, nearly all the United Nations member countries have experienced a rapid spread of the virus and are taking preventive measures to overcome the threats posed by the pandemic [3].

The COVID-19 disease can spread in a population through infected symptomatic/asymptomatic individuals who come in contact directly/indirectly [4]. Thus, concerned with the public health and well-being affected due to COVID-19, various countries have adopted comprehensive strategies converging to social distancing such as the closure of schools, ban of large gatherings, isolation of symptomatic individuals, and monitoring travelers, particularly to those from COVID-19 hotspots. Large scale lockdowns of populations have been implemented to prevent social contacts and reduce reproduction of the infected cases [5–7]. Evidence also highlights the importance of mitigation interventions in controlling the transmission of the virus [5, 8, 9]. Nonetheless, the timing of the implementation of strategies vary between countries and can significantly influence the growth curve of the epidemic [10].

The growth curve of total confirmed cases for many countries initially demonstrates a gradual increase near the start of the epidemic and is often followed by a sudden shoot or a critical transition [11, 12], as the disease spreads. This shoot places a considerable burden on the limited availability of the public health resources required to treat the disease and inhibit its further spread. Thus, it is crucial to anticipate this inevitable transition to take effective controlling measures for the outbreak. There exists a rich history of investigations that can predict processes that could lead to ecological outbreaks [13–16]. Theory suggests practical applicability of a variety of leading generic indicators, widely known as Early Warning Signals (EWSs) (e.g. return rate, variance, autocorrelation, and conditional heteroskedasticity) to identify the proximity of a system to such a critical transition [13, 14, 17]. For instance, in time-series data following ancient abrupt climate shifts, EWSs can be identified before the critical transition took place [18]. Similarly, EWSs were seen in the resurgence of malaria in Kericho, Kenya [19].

EWSs are hallmarks of critical slowing down (CSD) of a system as it approaches a critical transition. The phenomenon of CSD owes to the loss of resilience in the system such that even small disturbances can invoke an

* Corresponding author; parthasharathi@iitrpr.ac.in

often irreversible transition to an alternative stable state [11, 20]. In particular, the phenomenon of CSD can be captured as a large time taken by a system to return to its previous states due to which the rate of return of a system decreases prior to a transition. Moreover, it leads to an increase in the short term memory of a system, this feature can be identified by the changes in the correlation structure of a time-series preceding a critical transition [13, 14, 21–23]. Thus, it is crucial to understand whether these indicators capture the characteristics of slowing down in the epidemic growth curves.

Acquiring prior information about an upcoming transition that may occur in many countries with confirmed cases of COVID-19 is urgent to curb the impact of the rising pandemic and take appropriate containment measures. While the effectiveness of transmission of the disease and its severity in an individual depends on various factors [24], the timing of implementation of various social distancing measures can play a crucial role to influence the spread of the epidemic, assuming that the effectiveness of social communication/interaction in transmitting the disease can be of similar or equal probability across countries.

To mitigate the epidemic, China strictly restricted public movement and followed measures of quarantine and symptomatic isolation 24 days later (i.e., 23 January) to the arrival of the first reported case. The total reported cases (confirmed) at the time of the lockdown were nearly 623 (accounting for approximately 4.4732×10^{-7} of the total population). The daily increase in the number of confirmed cases in China, thus, saturated nearly in mid of March, hence flattening the growth curve of the total confirmed cases. European countries adopted different non-pharmaceutical measures to intervene in the disease transmission. The spread began later in Italy as compared to China; however, the strict interventions were initiated on 9 March, which marks a gap of nearly 40 days from the first reported case. Spain, which is continued to suffer severely by the virus, reported its first infected case on 1st Feb and took nearly 45 days to put the country on lockdown(see Table I). India confirmed its first case on 30 January and prompted "Janata curfew" and lockdown measures for complete cessation of public contacts nearly around 22 March, with approximately 2.36×10^{-7} of its population as COVID-19 infected cases, while this proportion was more than $1.7\times10^{-4}\%$ in the US. Therefore, it is essential to unfold how prolonged gaps between the arrival of the epidemic and non-pharmaceutical interventions such as quarantining/social distancing can influence public health and the environment at a national as well as a global scale. More interesting is to understand if the statistical tools can be useful to formulate laws to stifle the spread of an epidemic.

In this work, we analyse how the timing of strict controlling strategies influence the COVID-19 growth curve of the total confirmed cases in different countries. We use statistical tools to calculate the return rate and lag-1 autocorrelation function of the time-series data of the cumulative confirmed cases each in nine different countries. We investigate the EWSs for the alarming situations observed in the growth curves in each of the countries and record the timing of implementation of containment strategies to slow down the outbreak. Our work suggests that the dynamics of growth curve in the initial 40 days since the first reported case can signal an upcoming sudden rise in the cumulative number of infected cases. Thus, preliminary actions of at least 20 days before the timing of observed EWSs is crucial for an effective and timely containment of the disease. Delay in the strict surveillance and control measures can increase the time to contain the spread, which in turn will affect a larger proportion of the population. Furthermore, the proportion of the affected cases on the commencement of public health measures plays a significant role in containing the epidemic in each country. The timeline of implementation of strict intervention strategies coincided with that of emergence of EWSs for many countries such as India, Italy and Germany. However, the relatively low proportion of the affected cases in the case of India compared to Italy or Germany can be a significant factor explaining the slow rise for India but a relatively disruptive situation in the other countries. Thus, a combination of these two factors for India may restrict the extent of COVID-19 spread in the country, as compared to many other countries across the world. We conclude that model-independent forecasting systems can be applied to clinical data sets for predictability of the disease re-occurrence and formulate control policies.

II. RESULTS

We obtain the data sets of the cumulative number of the COVID-19 cases from the date of reporting of the first affected person up to 25 March, 2020 each for India, China, South Korea, the United States (US), Singapore, Germany, Italy, United Kingdom (UK) and Spain (for the data source see Materials and Methods section). Figure 1 depicts the growth curve of the fraction of the affected population in each of these countries. Interestingly, it is noted that the growth curve of the confirmed cases follows a slow increment for around initial 20-25 days, which can be interpreted as a time window to control the epidemic promptly and effectively. Since human to human contact is a leading transmitter of the disease, therefore, by-passing a certain threshold of infected cases, the growth curve shows an increasing slope and finally depicts a sudden large shift/transition in the fraction of affected populations (Fig. 1) [25]. It is important to note that the growth in number of cases for China and South Korea, the countries which initiated public monitoring/social distancing actions relatively earlier as compared to the other countries, saturates after nearly 3-4 weeks from the initiation of the lockdown. The shift of the COVID-19 from a low-burden to a high-burden state can be associated with the phenomenon of critical

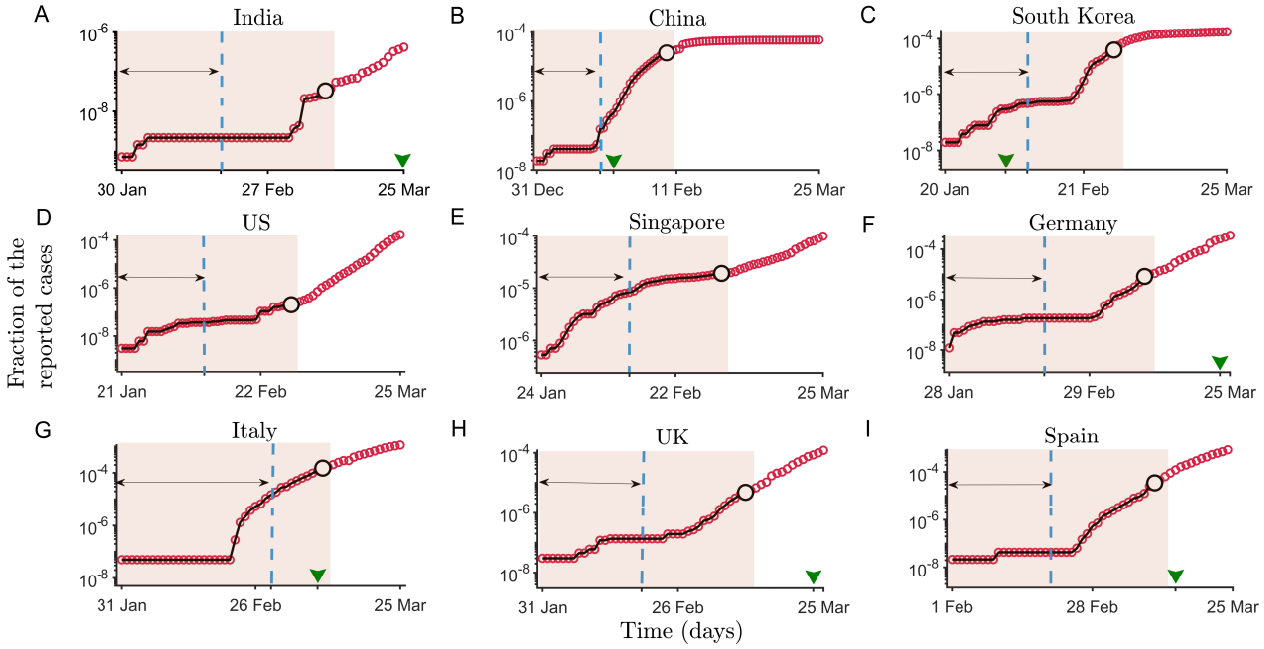

FIG. 1. Time-series constructed as a fraction of total infected cases in nine different countries across the world. The increasing curves depict growth in the number of cases after the onset of the epidemic in the respective countries up to 25 March. Shaded regions identify the transition phase nearly after around 40 days (35 in the US) and marks the data used to compute the indicators of slowing down. The arrows mark the size of the moving window. In the subfigures, A, B, C, F, G, and H, the arrow-heads on the x-axis mark the beginning of the officially recorded social-distancing and/or lockdown dates, and for other countries, the recorded social-distancing and/or lockdown dates are after 25 March.

transition. Thus we employ statistical methods that can monitor the onset of the transition phase and provide insights into the growth curve so as to suggest establishing worldwide disease elimination campaigns.

Signals of critical slowing down

To estimate statistical indicators anticipating the upcoming shifts in the growth curve, we extract the cumulative daily number of COVID-19 cases up to 35-40 days from the beginning of the epidemic (shaded regions in Fig. 1), for each country (note that the EWSs analyses on the fraction of reported case datasets result in qualitatively similar outcomes). To examine whether the system slows down to recover from perturbation while approaching the transition, we estimate the changes in the return rate and autocorrelation at first lag (ACF(1)) of each extracted data for all the nine countries (see Methods). Critical slowing down is reflected in systems near a critical transition through an increase in the autocorrelation. We observe that after nearly 40 days of the onset of the epidemic, the short term memory of the time-series data exhibits an increasing trend in most of the countries (Fig. 2). However, there are no significant signals of CSD exhibited by ACF(1) for the datasets of India as well as Italy (Figs. 2J and 2P). The return rates for all, except India (Fig. 2A), decreases, thus signaling to expect a sudden rise in the number of the COVID-19 cases for these countries. Furthermore, the strength of signals varies amongst countries depending upon the data sets determining the fraction of affected populations in individual countries. For instance, the trends in the UK are observed to be very strong, with ACF(1) approaching close to 1 (see Fig. 2Q) [21]. Since the time lag of up to almost two weeks is expected for the detection of symptomatic cases, the observed signals of CSD around 40 days indicate that the total cases gathered till then must be infected with the disease around two weeks ago. Thus, suitable preventive and surveillance strategies adopted in the initial 20-25 days are capable of suppressing the COVID-19 outbreak [26].

EWSs and enforcement of interventions

The timing of intervention measures varies among the countries. China was the first country to take the containment measures, nearly after 24 days at the beginning of the epidemic, while Italy took around 40 days, and other countries followed afterward. As a consequence, the COVID-19 growth curve in China flattened after 20-25 days of implementing the intervention measures. Similar to China, South Korea adopted different combinations of controlling measures around mid February (in the time window of 20-25 days since the epidemic began there).

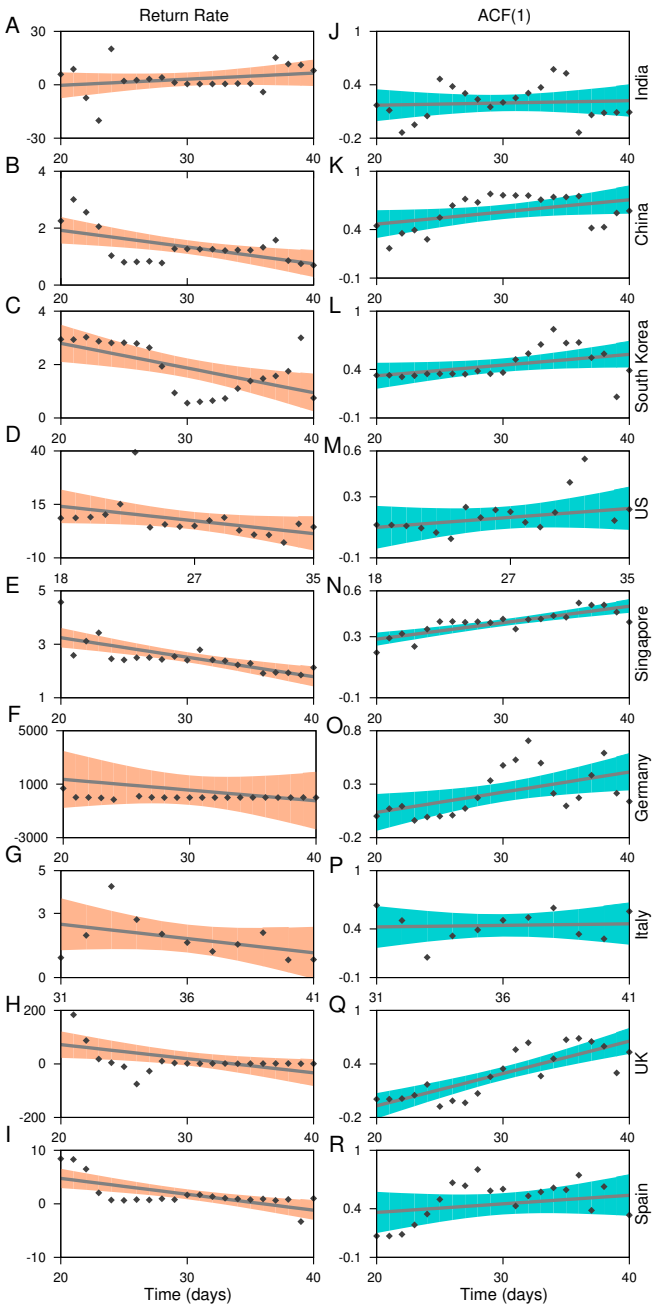

FIG. 2. Statistical estimates used to analyse the signals of a forthcoming transition in the COVID-19 growth curve. Figures on the left panel depicts the return rate measures each for India, China, South Korea, USA, Singapore, Germany, Italy, UK, and Spain. The right panel shows the lag-1 autocorrelation of the time-series data analysed in the corresponding countries. Scattered points are the estimated values of the respective slowing down indicators. Solid lines reflect the increasing/decreasing trend in the indicators and are obtained by fitting linear regression models. The shaded regions are the confidence bounds for the fitted models.

This measure was accompanied by a drop in the number of cases, and the curve followed the pattern as observed

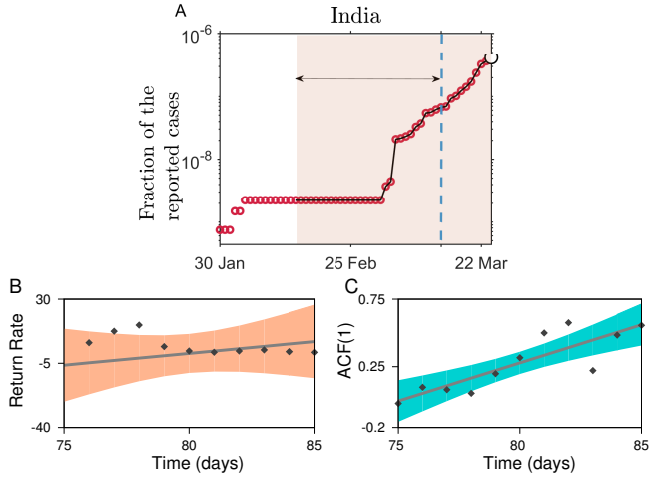

FIG. 3. Statistical analyses to measure the indicators of CSD using a backward approach while data extraction, in the case of India. A: The growth curve depicting the fraction of people infected from 30 January up to 25 March. The shaded region is the data used to calculate the return rate and lag-1 autocorrelation function. It demonstrates the epidemic situation during the period of 40 days prior to the commencement of the nationwide lockdown. The arrow marks the size of the rolling window (75%) used to calculate the statistical signals. B-C: The estimated values of the slowing down indicators. Solid lines are the fitted linear regression models to analyse the trend in the indicators along with the confidence bounds (shaded regions).

for China (Fig. 1 C). The rising indicators of CSD around 40 days also suggest that the time gap in implementing the protocols such as the closure of public gatherings, controlled public movement, lockdown beyond 15-25 days from the onset of the crisis can significantly influence the growth curve and result in the more extended time required to flatten it. However, the interventions during the initial 40 days in each of these countries can slowly hamper the daily increase in the number of cases. In the US, the hints of an approaching transition are visible relatively early (around 35 days after the first case was reported) by both the return rate as well as the ACF(1) (Figs. 2D and 2M).

The scenario is quite different in the case of India. The COVID-19 growth curve does not signal the behavior of CSD within the initial 40 days of the outbreak (Figs. 2A and 2J). Due to a continuous rise in the number of cases and the exponential curve observed for many countries, we also analyse the EWSs in the growth curve for India, following a backward approach. We consider the cumulative cases of 40 days prior to the beginning of the country-wide lockdown (25 March, Fig. 3A). The autocorrelation captures the signals of CSD (Fig. 3C), whereas the return rate does not show a decreasing trend (Fig. 3B).

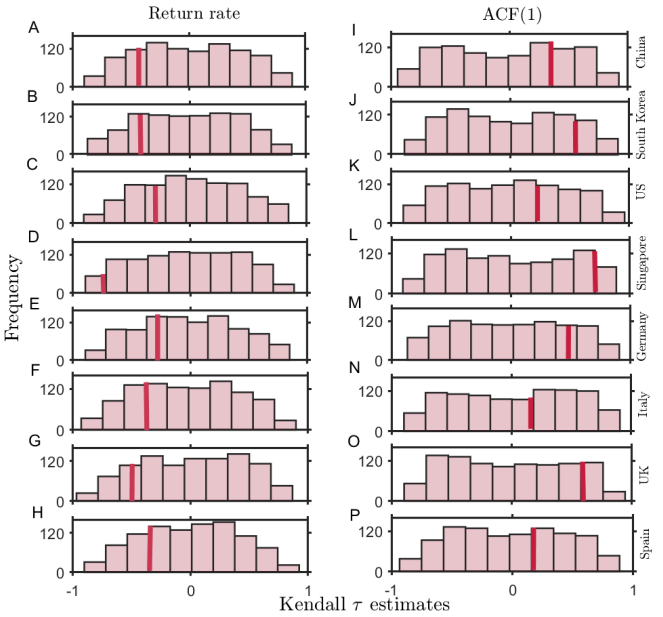

FIG. 4. The probability distribution of Kendall-τ test statistic on a set of 1000 surrogate time-series generated by bootstrapping residual time-series of the original data. Histograms depict the distribution of the test statistic for the surrogate time-series return rate (left panels) and autocorrelation function at lag-1 (right panels). Solid lines indicate the limit beyond which the Kendall-τ of the surrogate data is higher (lower) than the statistic observed in the ACF(1) (return rate) of the original time-series.

Onset of social distancing practices and the affected population density

Another important aspect is to consider the reported proportion of a population affected at the time of the implementation of intervention measures. So far, Germany and Italy, which had the largest outbreak in Europe around mid of the March, had visible signals of the forthcoming transition (Figs. 2F and 2O). It is noted that each of the countries, namely India, Germany, and Italy, adopted concerned public health measures around the time when the EWSs were visible in their respective data sets (see the arrow-heads on the x-axis in Fig. 1). However, the fraction of the population affected by the time in Germany and Italy was much higher (approximately 2.9×10^{-4} and 1.2×10^{-4}, respectively) as compared to that for India $- 2.36 \times 10^{-7}$. Thus, the growth curve projected a significant rise in these two countries, whereas the rise in the number of cases in India is still gradual and is expected to follow a similar response owing to the effectiveness of these interventions. Overall, our analyses suggest that delayed interventions (depending upon the signals of CSD) along with the fraction of the affected population can influence the country-wise variation in the daily number of rising cases.

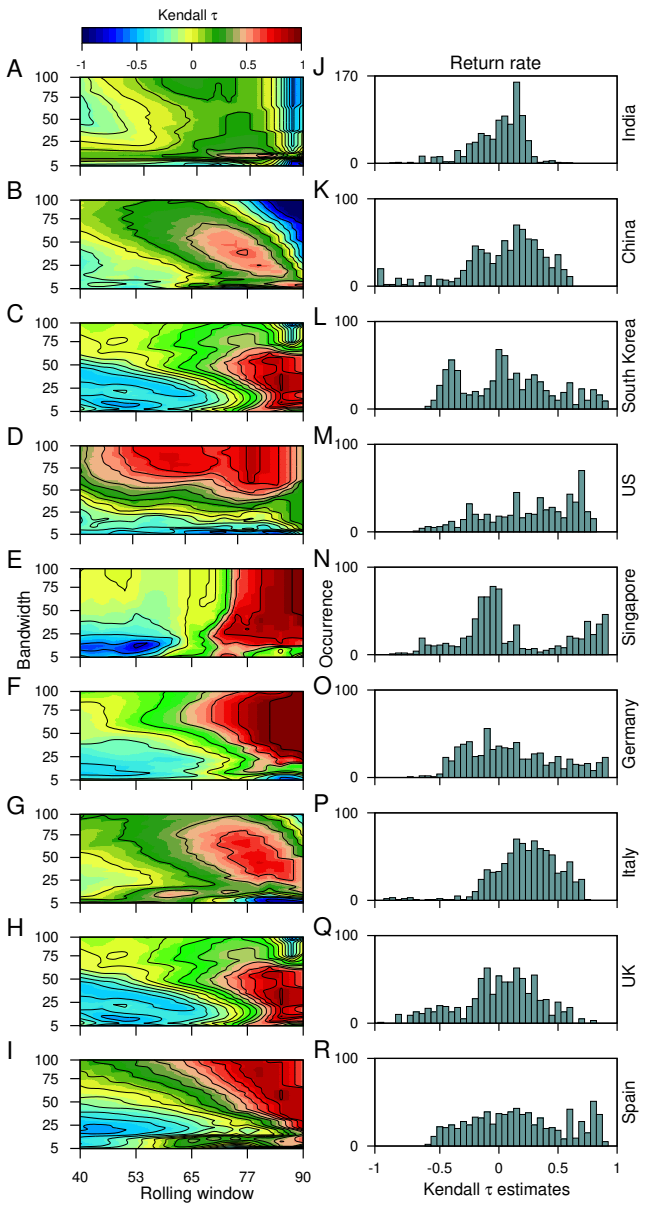

FIG. 5. The sensitivity of the choice of the rolling window size and the filtering bandwidth to estimate the EWSs. Contour plots demonstrate the effect of moving window size and the filtering bandwidth on the trends observed while calculating the return rate, using the Kendall-τ test statistic. Panels on right show the frequency distribution of the trend statistic.

Surrogate analyses

The lower number of data points available for the analyses can lead to feeble trends and influence the probability of occurrence of the increased signals of CSD by chance. Further, due to undocumented patients, there is always a chance of stochasticity in the number of reported cases. Thus, we studied the likelihood of coincidence in the occurrence of trends in the ACF(1) and the return

rate observed in our original data sets by investigating the indicators in the surrogate time-series (see Materials and Methods). The surrogate time-series is generated to follow similar distribution (mean and variance) of the data time-series before the episode of a sudden rise in the number, denoted by shaded regions in Fig. 1 (see Materials and Methods). Figure 4 depicts the distribution of the test statistic of the surrogate time-series. Solid lines show the trend estimate obtained for the original time-series. We calculate the probability of randomness of our observed estimates as the fraction of 1000 surrogate time-series having trend statistic of same or higher absolute values than the original trend, i.e., $P(\tau^* \leq \tau)$. The probability to, by chance, obtain similar trend statistic varies from country to country, depicting most significant estimates for Singapore (Figs. 4D and 4L) and UK (Figs. 4G and 4O). The probability estimates P obtained by bootstrapping the data sets for each of the countries are given in the Table II. While in the case of the US, the probability of randomness in our observed estimates is relatively high (Figs. 4C and 4K), rapid spreading in the epidemic makes it keystone to consider applicability of EWSs to warn-off such events. Overall, we find a low probability of randomness in both the ACF(1) and the return rate estimates for all the cases. However, the observations are more significant for the return rate. This analysis suggests the robustness of the return rate as an EWS in predicting the signals of CSD.

Sensitivity analyses of the generic indicators

The choices made in data transformations such as filtering/ Gaussian detrending can also influence the trends observed. Thus, it is necessary to test the robustness of the estimated trends towards the choice of rolling window size and the filtering bandwidth. Here, we employ sensitivity analyses for the return rate using the CSD dataset. (see Fig. 5). We use Kendall-τ estimates of the return rate for all the combinations of these two parameters (for details, see Materials and Methods).

We find that the observed trends in the return rate are sensitive to the choice of parameters and can significantly vary between the datasets. High bandwidths reveal the opposite outcome of the return rates for most of the data sets analysed (Figs. 5C–5I). Since we use Silverman's thumb rule to select the bandwidth, which gives the best fit to the data, therefore, the choice of window size can influence the observed trends (Figs. 5A–5I). In our work, we find a large size of the rolling window can alter the EWSs analyses and produce misleading estimates for the return rate. For instance, sensitivity analyses show that it is challenging to disentangle accurate signals of an impending transition from the false ones (Figs. 5G and 5P) for a wide range of window sizes and bandwidths. False signals of an alarming situation can deviate from understanding the gravity of any situation and intensity of surveillance needed. Thus, the choice of these parameters is crucial in anticipating the signals of a forthcoming transition and implementing convincing public health measures.

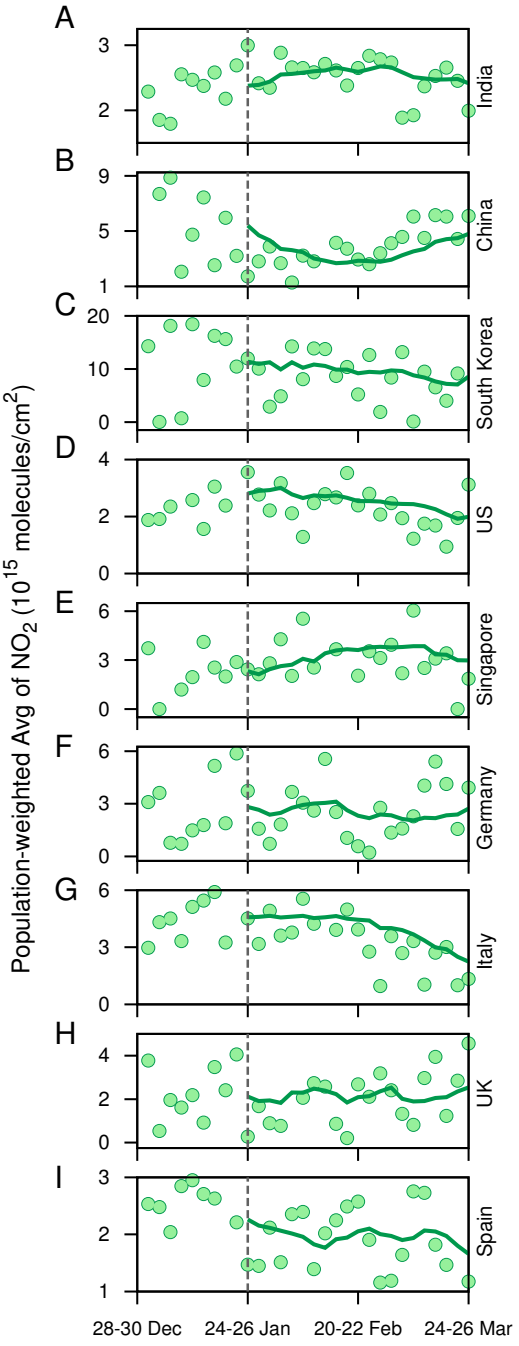

FIG. 6. Time-series of triads of the population-weighted total-column NO_2 (molecules/cm^2) density over the length of the study period for the nine countries considered in this work (depicted by the circular points). The solid curve in each subfigure represents a 10-triad moving window average of the time-series.

Impact of COVID-19 spread on the atmospheric total-column NO_2 density

The rigor of social distancing/intervention strategies can be measured by atmospheric data, as the lockdown periods have witnessed better air quality across the globe. Thus, we first obtain time-series of triads of the population-weighted total-column NO_2 (molecules/cm^2) density over the length of the study period for the nine countries considered in this work (the circular points in Fig. 6). The solid curve in each subfigure of Fig. 6 represents a 10-triad moving window average. In the majority of the countries, the timing of NO_2 decline concurs with the spread of the virus and the onset of pragmatic lockdown in a country may be hypothesized by the reversal (or break) in the trend of NO_2. In China (Fig 6B), the decreasing trend in NO_2 is evident from January end till February; after that, it starts increasing which is coincident with the dynamics of the spread of COVID-19 disease. In India (Fig 6A), South Korea (Fig 6C), US (Fig 6D), Italy (Fig 6G) and Spain (Fig 6F), the decreasing trend in NO_2 coincides with time of the rapid spread in the virus (Fig. 1). We estimate that after the date of official enforcement of lockdown, time-averaged NO_2 decreased by 35.9% in China and 50.4% in Italy compared to the pre-lockdown period. India has also seen a significant drop by 5% of NO_2 in the last 2 triads around the day of "Janata curfew". The decreased pattern in NO_2 is expected to continue as the lockdown prevails up to 14 April. It should be noted that we did not control for meteorological variations which may have a significant impact on total-column NO_2 over the period of our study [27]. Overall, amidst the fears of the novel coronavirus, the countries where the lockdown intervened are expecting a rejuvenated environment. However, at the same time, possibilities to lower down air pollutants when the world is not facing such harsh conditions is also important to understand.

A minimal stochastic model

We propose a minimal kinetic model for the short term prediction of the spreading of COVID-19 disease. Suppose that the only processes are infection and recovery. The processes can be described as:

$$I + H \xrightarrow{k_i} I + I; \quad \text{and} \quad I \xrightarrow{k_r} H, \tag{1}$$

where I and H are infected and healthy people, respectively, and k_i and k_r are rate constants for infection and recovery. The first equation shows that if I is the infected people, then H becomes I at a rate k_i; and the second equation indicates that I recovers at a rate k_r. A minimal kinetic model can be formulated as ordinary differential equations for the population of I, as:

$$\frac{dI}{dt} = k_i I(1 - \frac{I}{K}) - k_r I, \tag{2}$$

where K is the size of the population.

We develop a master equation for the infected population by considering the two elementary processes (Eq. 1). The transition probability at which the number of infected population increases from i to $(i+1)$ is $w(i+1|i) = k_i i(1-i/K)$, and the rate at which the number of infected population reduces from i to $(i$-1) is $w(i-1|i) = k_r i$. From these, the probability of finding i infectives in the system at time t, $P(i,t)$ can be obtained from the following equation:

$$\frac{dP(i,t)}{dt} = w(i|i-1)P(i-1,t) + w(i|i+1)P(i+1,t) - (w(i+1|i) + w(i-1|i))P(i,t). \tag{3}$$

The above probabilistic model is solved by the kinetic Monte Carlo simulations by means of the Gillespie algorithm, which incorporates the intrinsic noise [28]. The algorithm considers each of the events as individual realisations of Markov process. The time and species numbers are updated stochastically by choosing the random processes.

To simulate the system (Eq. 3), we first obtain the parameters from the cumulative time series data of confirmed cases for India, China, and South Korea. In the data sets, we fitted the below logistic function (which is a solution of Eq. 2):

$$f(t) = \frac{a}{1 + b\exp(-ct)}, \tag{4}$$

where a, b, and c are parameters. Once we obtain these parameters for an individual country, we map them to our model and find the system parameters k_i, k_r and K, and i_0 is the initial infected poplation. We list those parameters below:

Country	k_i	k_r	K	i_0
India	0.895	0.716	10566605	2
China	0.967	0.7736	413210	27
S. Korea	1.0120	0.8096	51420	4

Then the above parameters are used to solve the Master equation (Eq. 3), and we perform Monte-Carlo simulation to get stochastic trajectories up to 15 April. We present the simulated stochastic trajectories in Fig. 7. For each country, we have five trajectories. For China and South Korea, we find that our stochastic trajectories are consistent with the real time-series of the number of infected people. However, for India, our result shows that on 15 April 2020, the number of infected people can go up to approximately 13500 (which is an average of final values of the five simulated trajectories).

The problem of predicting the spreading of COVID-19 is a complex one and depends on many factors like social distancing, an early detection of the disease, the detection of major hubs of the disease, etc. Here, we have provided a minimal kinetic model which uses the trends of the available data and may work only for short term prediction.

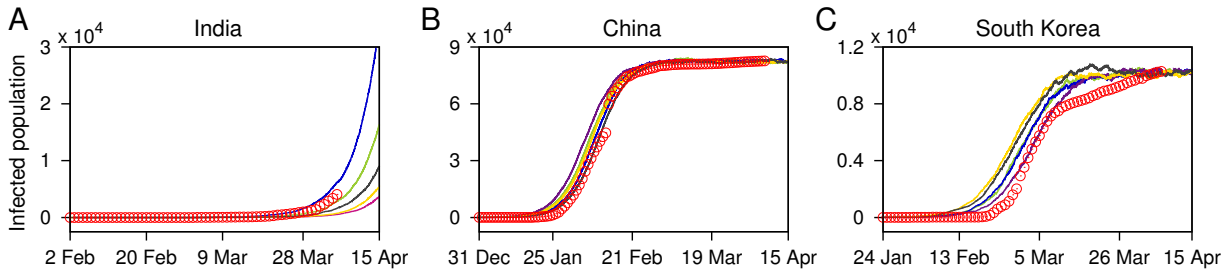

FIG. 7. Stochastic trajectories (marked with pink, yellow, grey, green, and blue curves) of the infected population generated using the Gillespie algorithm for: A. India, B. China, and C. South Korea. The original data sets up to 6 April for the respective countries are depicted by circular (red) points, for a comparison.

III. DISCUSSION

The COVID-19 pandemic revealed an exponential rise in the reported number of cases and affected the public health ranging from mild to severe conditions. Countries across the world are combating the spread of the coronavirus through various social distancing/intervention measures such as closure of schools and universities, banning of public events and large gatherings, isolation of symptomatic COVID-19 cases, mass quarantines etc. For national as well as international control of public health, it is crucial to understand the significance of the onset timing of such measures [29].

The World Health Organisation lately reports new cases being reported into several new countries across the globe [30–32]. Our study can provide an insight to tackle the ongoing pandemic and its associated growth curve in the context of the timing and strength of the interventions. We use the data of the number of COVID-19 cases in nine different countries to investigate some statistical patterns in the growth curves. The number of cases covers a small fraction of the population during the initiation of the epidemic, and the fraction remains nearly stagnant for around 20 days from the arrival of the first case. After the threshold of around 20 days, the number of cases starts increasing rapidly, and in a relatively shorter span a significant fraction of the population can be affected. This trend is analogous to the idea that the growth curve remains close to one stable state for sufficient time and, crossing a time threshold invokes a sudden shift/transition to another stable state, where a significant fraction of the population gets affected. In our work, we employ statistical indicators of critical slowing down to check if such transitions can be signaled beforehand and how the anticipation of such transitions can help mitigate such crisis at a policy level.

We observe that the 35-40 days window from the arrival of the first case in each country signaled an impending transition. In the vicinity of 35-40 days, an increase in the short term memory (at lag 1) of the data as well as a decrease in the return rate indicates the phenomenon of critical slowing down. Our work suggests that while non-pharmaceutical interventions are necessary to mitigate such epidemic, the timing of initiation of concerned actions can strongly influence the outcome of the situation. Owing to the time lag in the detection of symptomatic cases, the statistical indicators suggest that the initial 20-25 days being crucial to suppress the loss of public health. The controlled response of the epidemic growth curve for China and South Korea can be associated with the time distance between implementation of interventions and the transition point. Thus, EWSs analyses is crucial while defining the onset of the interventions and suppress the rise of daily cases. Importantly, another crucial aspect is the proportion of affected cases in each country, i.e., a measure of the fraction of the country's population, and not the absolute numbers, which is infected at the time of interventions such as a strict lockdown. As probability of the propagation of disease can be thought of as mostly similar or equal amongst individuals across the globe, it depends upon the fraction of infected cases in each country during the beginning of interventions. For instance, a decrease in the return rate and increased autocorrelation anticipated the upcoming rise in the growth curves for both India as well as Italy, and interestingly, both the countries imposed individual nationwide lockdown near the situation close to the transition (Table I). However, the control in India depicts better results in altering the growth curve than that in Italy. The alterations in the growth curve is most likely to be a consequence of a difference in the proportion of cases affected by the epidemic at the beginning of mitigation strategies. India resembling China in terms of the total population density accounted for approximately 2.36×10^{-7} cases of the total population, while Italy with relatively less population density crossed 1.22×10^{-4} cases of their total population. Thus, even with imposition of the public health measures near the signals of CSD, the outcome for both the countries can dramatically vary. Furthermore, the disruptive situation in the US is indicated by EWSs, as the EWSs indicators were visible quite early for the US. Thus, sharp rise in the number of cases for the country is a consequence of both the delay in effective social distancing interventions

as well as a large proportion of affected cases at that time. Overall, our work suggests that in almost all the countries, a sharp upcoming rise in the growth curve can be captured using statistical measures, before the actual transition.

Another horizon the infectious coronavirus put forwards is on the quality of air pollution in countries where social distancing/lockdown is enforced. NO_2, which is majorly emitted from anthropogenic activities like land transportation, industries, and energy sectors, was estimated to decrease in consequence of lockdown measures implemented by the government of respective countries. Population weighted average column NO_2 was found to decrease with amplification in a number of cases across most of the countries. Apart from this, NO_2 column quantities may be used as a proxy to estimate the effectiveness of a lockdown on air quality. We find that NO_2 column quantities started following a decreasing trend during the last week of February in Italy and US which indicate a partial unofficial closure of anthropogenic activities, taking into consideration that the official COVID-19 induced lockdown was enforced on 09 March and around 25 March in Italy and the US, respectively. Whereas, in the UK, an increasing trend in NO_2 column till the last week of March indicates no such public awareness to restrict anthropogenic activities (the government declared the lockdown from 23 March). We acknowledge that the reduction in NO_2 is also associated with the compliance of the population of the individual nation to abide by the lockdown measures.

Furthermore, we suggest that the interventions employed by India may not be at the time when the curve is very far from reaching the transition; however, the smaller proportion of affected cases may be the determining factor in limiting the disease spread in India. Implementation of a nationwide lockdown in India may have better prepared the country for taking measures to control the epidemic spread and bend the curve. However, our analyses also suggest that the period beyond the signals of CSD also needs efficient monitoring. The result of our minimal stochastic model predicts that on 15 April, the number of infected people can go up to approximately 13500. Thus, the extended period of such measures are needed and likely to be effective [5].

We envision that it is fundamental to identify the situation of such a crisis across the world and make use of the lead time. The EWSs can keep track of the changes in the trend statistics in the number of reported cases and warn when a threshold is reached. The statistical tools used can be beneficial to identify whether the features of shift in a system are suppressed by the intervention strategies being adopted. In particular, while different combinations of strategies are adopted to overcome such crisis, the information of an upcoming transition and its threshold is important to formulate the degree of such interventions. However, special care should be taken in the choice of rolling window size and the filtering bandwidth while estimating the signals of slowing down. Inappropriate choices may give weak and/or diminish the signals of an imminent transition, which may deviate from understanding the urgency of the situation. Another aspect to consider is that the varying extent of testing for COVID-19 across the countries may have affected the total number of reported cases; thus, our results here hold specifically for the number of reported cases.

MATERIALS AND METHODS

The COVID-19 Data Source. We have used the COVID-19 data set provided by the European Centre for Disease Prevention and Control (ECDC): An agency of the European Union (https://www.ecdc.europa.eu/en/publications-data/download-todays-data-geographic-distribution-covid-19-cases-worldwide). Initially we extract the data of the daily number of reported cases up to 25 March, 2020 and in general mark the first date of the reported cases as the day of the beginning of the epidemic in the respective countries. Regardless of the affected person recovers or dies, the virus contraction occurs once; thus, we consider cumulative data of the daily number of the confirmed cases for nine different countries for our analyses.

Data selection. We use the available time-series to test the predictability of an upcoming transition for each country. The generic indicators are examined using the time-series segments of an initial $35-40$ days (35 days for the US) of the epidemic in each country (shaded regions in Fig. 1).

Detrending. Often, non-stationarities in the data lead to false indications of impending transitions. To overcome this, we obtain the residual time-series by subtracting a Gaussian kernel smoothing function from the empirical time-series [14]. Further, we estimate the return rate and autocorrelation at first lag for the residual time-series choosing a rolling window of $\frac{3}{4}$ the size of the time-series data (i.e., 75 %) for Italy and 50% for the other countries. We choose the filtering bandwidth using Silverman's thumb rule to avoid any overfit.

Autocorrelation at first lag and return rate. The fluctuations in the time-series reveal different novel phenomena such as sudden transition, flickering and stochastic switching, etc. It is established that followed by a perturbation, the rate of return of the system slows down near an impending transition or a tipping point. This phenomenon of slow return rate or recovery from a perturbation in the vicinity of a sudden transition is known as critical slowing down (CSD). We capture the signals of CSD by estimating short term autocorrelation (at lag-1) and return rate of the time-series. Furthermore, CSD also increases the short term memory of the time-series, which is observed through the correlation structure of the time-series before a transition. We compute autocorrelation at lag-1 by fitting an autoregressive model of order 1 (of the form $z_{t+1} = \alpha_1 z_t + \epsilon_t$) using an ordinary

least-squares fitting method. Return rate is calculated as the inverse of the first-order term of fitted autoregressive model, given by $\frac{1}{\alpha_1}$ [14, 33]. The time-series analysis has been performed using the "Early Warning Signals Toolbox" (http://www.early-warning-signals.org/).

Surrogates. To test the significance of our statistical analyses, we estimate Kendall rank correlation-τ test statistic for both the generic indicators. We generate 1000 surrogate time-series of the same length as the analysed real data sets to test the likelihood of obtaining the computed trends by chance. The surrogate records are obtained on bootstrapping the real datasets by shuffling the original residual time-series and sampling the data with replacement. This method generates the surrogate time-series with a similar distribution of the original time-series [18]. For each surrogate, we consider the Kendall-τ estimate as the test statistic to measure the robustness of the outcomes. Further, we calculate the fraction of the surrogates having the same or higher (lower, for return rates) test static value than the original data and measure the probability $P(\tau^* \leq \tau)$ to calculate that the observed test statistic is by chance.

Sensitivity Analyses. The predictability of each of the indicator depends upon the datasets investigated as well as the choices made for processing the data. Thus, it is essential to check the efficacy of our results to such choices. In particular, we analyse the sensitivity of our observations to the choice of rolling window size and degree of smoothing (filtering bandwidth) used during the calculation of indicators and detrending/filtering the data sets, respectively. We estimate the return rate using window sizes ranging from 40% to 90% of the time-series length in an increment of 2 points and for bandwidths ranging from 5% to 100% with the increment of 2 points. We quantify the robustness of the outcomes towards the range of window sizes and bandwidth using the distribution of the Kendall-τ test statistic.

Satellite retrieved NO_2 emissions data source. Worldwide, the lockdown response to the onset and spread of COVID-19 caused a decrease in daily and economic activities, which in turn is expected to cause a reduction in ambient air pollution. This can also be used as an indicator to determine whether government policies of lockdowns/restricted human movements are successful or not. To further examine this, we use the Ozone Monitoring Instrument (OMI) retrieved total column NO_2 (available from https://aura.gsfc.nasa.gov/omi.html) as a proxy to infer the change in anthropogenic air pollution for the time-period of our study. OMI flies onboard the EOS Aura sun-synchronous polar-orbiting satellite. It has a swath length of 2600km and a level-2, spatial resolution of $13 \times 24 \text{km}^2$ [34]. The OMI NO_2 column was satisfactorily validated against surface spectrometer measurements in recent studies [35, 36]. To roughly obtain a global coverage, we consider 3-day time slices (triads) within which the overlapping swath overpasses were averaged. Thereafter, we perform a population-weighted average of the grids that lie within the political boundaries of the countries considered in this study. Gridded population data was obtained for 2015 from SEDAC (https://sedac.ciesin.columbia.edu/data/collection/gpw-v4).

ACKNOWLEDGMENTS

S.S. acknowledges the financial support from the Department of Science & Technology (DST), Govt. of India under the scheme DST-Inspire [Grant No.: IF160459]. P.S.D. acknowledges financial support from the Science & Engineering Research Board (SERB), Govt. of India [Grant No.: CRG/2019/002402].

[1] World Health Organization. Coronavirus disease 2019 (covid-19): situation report, 59. Technical report, World Health Organization, 2020.

[2] Eurosurveillance editorial team. Note from the editors: World health organization declares novel coronavirus (2019-ncov) sixth public health emergency of international concern. *Eurosurveillance*, 25(5), 2020.

[3] Covid-19 coronavirus pandemic. `https://www.worldometers.info/coronavirus/`, 2020.

[4] Joseph T Wu, Kathy Leung, Mary Bushman, Nishant Kishore, Rene Niehus, Pablo M de Salazar, Benjamin J Cowling, Marc Lipsitch, and Gabriel M Leung. Estimating clinical severity of covid-19 from the transmission dynamics in wuhan, china. *Nature Medicine*, pages 1–5, 2020.

[5] Rajesh Singh and R Adhikari. Age-structured impact of social distancing on the covid-19 epidemic in india. *arXiv preprint arXiv:2003.12055*, 2020.

[6] Elisabeth Mahase. Covid-19: Uk starts social distancing after new model points to 260 000 potential deaths, 2020.

[7] Coronavirus in new york: Lunar new year events canceled over fears. `https://www.nytimes.com/2020/01/29/nyregion/coronavirus-nyc.html`, 2020.

[8] Neil Ferguson, Daniel Laydon, Gemma Nedjati Gilani, Natsuko Imai, Kylie Ainslie, Marc Baguelin, Sangeeta Bhatia, Adhiratha Boonyasiri, ZULMA Cucunuba Perez, Gina Cuomo-Dannenburg, et al. Report 9: Impact of non-pharmaceutical interventions (npis) to reduce covid19 mortality and healthcare demand. Technical report, Imperial College London, 2020.

[9] Roy M Anderson, Hans Heesterbeek, Don Klinkenberg, and T Déirdre Hollingsworth. How will country-based mitigation measures influence the course of the covid-19 epidemic? *The Lancet*, 395(10228):931–934, 2020.

[10] Kiesha Prem, Yang Liu, Timothy W Russell, Adam J Kucharski, Rosalind M Eggo, Nicholas Davies, Stefan Flasche, Samuel Clifford, Carl A B Pearson, James D Munday, Sam Abbott, Hamish Gibbs, Alicia Rosello,

Billy J Quilty, Thibaut Jombart, Fiona Sun, Charlie Diamond, Amy Gimma, Kevin [van Zandvoort], Sebastian Funk, Christopher I Jarvis, W John Edmunds, Nikos I Bosse, Joel Hellewell, Mark Jit, and Petra Klepac. The effect of control strategies to reduce social mixing on outcomes of the covid-19 epidemic in wuhan, china: a modelling study. *The Lancet Public Health*, 2020.

[11] M. Scheffer. *Critical Transitions in Nature and Society*. Princeton University Press, 2009.

[12] Tobias S Brett, John M Drake, and Pejman Rohani. Anticipating the emergence of infectious diseases. *Journal of The Royal Society Interface*, 14(132):20170115, 2017.

[13] M. Scheffer, S. R. Carpenter, T. M. Lenton, J. Bascompte, W. A. Brock, V. Dakos, J. van de Koppel, I. A. van de Leemput, S. A. Levin, E. H. van Nes, M. Pascual, and J. Vandermeer. Anticipating critical transitions. *Science*, 338:344–348, 2012.

[14] V. Dakos, S. R. Carpenter, W. A. Brock, A. M. Ellison, V. Guttal, A. R. Ives, S. Kéfi, V. Livina, D. A. Seekell, E. H. van Nes, and M. Scheffer. Methods for Detecting Early Warnings of Critical Transitions in Time Series Illustrated Using Simulated Ecological Data. *PLoS One*, 7:e41010, 2012.

[15] Rong Wang, John A Dearing, Peter G Langdon, Enlou Zhang, Xiangdong Yang, Vasilis Dakos, and Marten Scheffer. Flickering gives early warning signals of a critical transition to a eutrophic lake state. *Nature*, 492(7429):419, 2012.

[16] Sonia Kefi, Vasilis Dakos, Marten Scheffer, Egbert H. Van Nes, and Max Rietkerk. Early warning signals also precede non-catastrophic transitions. *Oikos*, 122(5):641–648, 2013.

[17] Marten Scheffer, Jordi Bascompte, William A Brock, Victor Brovkin, Stephen R Carpenter, Vasilis Dakos, Hermann Held, Egbert H Van Nes, Max Rietkerk, and George Sugihara. Early-warning signals for critical transitions. *Nature*, 461(7260):53–59, 2009.

[18] Vasilis Dakos, Marten Scheffer, Egbert H van Nes, Victor Brovkin, Vladimir Petoukhov, and Hermann Held. Slowing down as an early warning signal for abrupt climate change. *Proceedings of the National Academy of Sciences U.S.A.*, 105(38):14308–14312, 2008.

[19] Mallory J Harris, Simon I Hay, and John M Drake. Early warning signals of malaria resurgence in kericho, kenya. *Biology Letters*, 16(3):20190713, 2020.

[20] Marten Scheffer, Steve Carpenter, Jonathan A Foley, Carl Folke, and Brian Walker. Catastrophic shifts in ecosystems. *Nature*, 413(6856):591, 2001.

[21] Vasilis Dakos, Egbert H Van Nes, Paolo d'Odorico, and Marten Scheffer. Robustness of variance and autocorrelation as indicators of critical slowing down. *Ecology*, 93(2):264–271, 2012.

[22] Partha Sharathi Dutta, Yogita Sharma, and Karen C Abbott. Robustness of early warning signals for catastrophic and non-catastrophic transitions. *Oikos*, 127(9):1251–1263, 2018.

[23] Sukanta Sarkar, Sudipta Kumar Sinha, Herbert Levine, Mohit Kumar Jolly, and Partha Sharathi Dutta. Anticipating critical transitions in epithelial–hybrid-mesenchymal cell-fate determination. *Proceedings of the National Academy of Sciences U.S.A.*, 116(52):26343–26352, 2019.

[24] Wei-jie Guan, Wen-hua Liang, Yi Zhao, Heng-rui Liang, Zi-sheng Chen, Yi-min Li, Xiao-qing Liu, Ru-chong Chen, Chun-li Tang, Tao Wang, Chun-quan Ou, Li Li, Ping-yan Chen, Ling Sang, Wei Wang, Jian-fu Li, Cai-chen Li, Li-min Ou, Bo Cheng, Shan Xiong, Zheng-yi Ni, Jie Xiang, Yu Hu, Lei Liu, Hong Shan, Chun-liang Lei, Yi-xiang Peng, Li Wei, Yong Liu, Ya-hua Hu, Peng Peng, Jian-ming Wang, Ji-yang Liu, Zhong Chen, Gang Li, Zhi-jian Zheng, Shao-qin Qiu, Jie Luo, Chang-jiang Ye, Shao-yong Zhu, Lin-ling Cheng, Feng Ye, Shi-yue Li, Jin-ping Zheng, Nuo-fu Zhang, Nan-shan Zhong, and Jian-xing He. Comorbidity and its impact on 1590 patients with covid-19 in china: A nationwide analysis. *European Respiratory Journal*, 2020.

[25] Joan T Matamalas, Sergio Gómez, and Alex Arenas. Abrupt phase transition of epidemic spreading in simplicial complexes. *Physical Review Research*, 2(1):012049, 2020.

[26] Ruiyun Li, Sen Pei, Bin Chen, Yimeng Song, Tao Zhang, Wan Yang, and Jeffrey Shaman. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2). *Science*, 2020.

[27] Daniel J Jacob and Darrell A Winner. Effect of climate change on air quality. *Atmospheric environment*, 43(1):51–63, 2009.

[28] Daniel T Gillespie. Stochastic simulation of chemical kinetics. *Annu. Rev. Phys. Chem.*, 58:35–55, 2007.

[29] Jodie Lynn Guest, Carlos del Rio., and Travis Sanchez. The 3 steps needed to end the covid-19 pandemic: Bold public health leadership, rapid innovations, and courageous political will. *JMIR Public Health and Surveillance*, 2020.

[30] Coronavirus disease 2019 (covid-19) situation report 72. `https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200401-sitrep-72-covid-19.pdf?sfvrsn=3dd8971b_2`, 2020.

[31] Coronavirus disease 2019 (covid-19) situation report 74. `https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200403-sitrep-74-covid-19-mp.pdf?sfvrsn=4e043d03_14`, 2020.

[32] Coronavirus disease 2019 (covid-19) situation report 76. `https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200405-sitrep-76-covid-19.pdf?sfvrsn=6ecf0977_2`, 2020.

[33] SR Carpenter, WA Brock, JJ Cole, JF Kitchell, and ML Pace. Leading indicators of trophic cascades. *Ecology Letters*, 11(2):128–138, 2008.

[34] Nickolay A. Krotkov, Lok N. Lamsal, Edward A. Celarier, William H. Swartz, Sergey V. Marchenko, Eric J. Bucsela, Ka Lok Chan, Mark Wenig, and Marina Zara. The version 3 OMI NO_2 standard product. *Atmospheric Measurement Techniques*, 10(9):3133–3149, September 2017.

[35] Yang Wang, Steffen Beirle, Johannes Lampel, Mariliza Koukouli, Isabelle De Smedt, Nicolas Theys, Ang Li, Dexia Wu, Pinhua Xie, Cheng Liu, Michel Van Roozendael, Trissevgeni Stavrakou, Jean-François Müller, and Thomas Wagner. Validation of OMI, GOME-2A and GOME-2B tropospheric NO_2, SO_2 and HCHO products using MAX-DOAS observations from 2011 to 2014 in Wuxi, China: investigation of the effects of priori profiles and aerosols on the satellite products. *Atmospheric Chemistry & Physics*, 17(8):5007–5033, 2017.

[36] LN Lamsal, NA Krotkov, EA Celarier, WH Swartz, KE Pickering, EJ Bucsela, JF Gleason, RV Martin, S Philip, H Irie, et al. Evaluation of omi operational standard no2 column retrievals using in situ and surface-based no2 observations. *Atmospheric Chemistry and Physics*, 14(21):11587, 2014.

[37] Curfews and lockdowns related to the 2019–20 coronavirus pandemic. `https://en.wikipedia.org/wiki/Curfews_and_lockdowns_related_to_the_2019-20_coronavirus_pandemic`, 2020.

TABLE I. Country wise dates of the onset of restricted public movement or respective nation-wide lockdown [37], along with the fraction of infected cases reported at the mentioned dates. Note that the fraction of infected cases is calculated over the total population of each country, which is considered to be constant.

Country	**Lockdown dates**	**Fraction of the affected population during lockdown**
India	22 March (Janata curfew)	$\approx 2.36 \times 10^{-7}$
China	23 January	$\approx 4.47 \times 10^{-7}$
South Korea	Early February (social distancing measures)	$\approx 3.09 \times 10^{-7}$
US	After 25 March (partial lockdown)	More than 1.7×10^{-4}
Singapore	27 March (partial lockdown)	$\approx 9.89 \times 10^{-5}$
Germany	23 March	$\approx 2.98 \times 10^{-4}$
Italy	9 March	$\approx 1.22 \times 10^{-4}$
UK	23 March	$\approx 8.54 \times 10^{-5}$
Spain	14 March	More than 9.05×10^{-5}

TABLE II. Probability of, by chance, obtaining the observed trend statistic of the original data for the set of 1000 surrogates having similar distribution. The likelihood of randomness in the estimated return rates and ACF(1) is mentioned for the data sets of each country studied in the work.

Country	**Kendall-τ (Return rate)**	**Kendall-τ (ACF(1))**
China	0.26	0.36
South Korea	0.24	0.16
US	0.17	0.34
Singapore	0.05	0.10
Germany	0.36	0.20
Italy	0.27	0.49
UK	0.19	0.12
Spain	0.26	0.35