

Monitoring the evolution of the COVID-19 pandemic in China, South Korea, Italy and USA through the net relative rate of infection of the total number of confirmed cases

João Manoel Losada Moreira
Centro de Engenharia, Modelagem e Ciências Sociais Aplicadas
Programa de Pós-graduação em Energia
Universidade Federal do ABC
Joao.moreira@ufabc.edu.br

Abstract

Managing the COVID-19 pandemic in the middle of the events requires real-time monitoring of its evolution to perform analyses of containment actions and to project near future scenarios. This work proposes a scheme to monitor the temporal evolution of the COVID-19 pandemic using the time series of its total number of confirmed cases in a given region. The monitored parameter is the spread rate obtained from this time series (day^{-1}) expressed in %/day. The scheme's capability is verified using the epidemic data from China and South Korea. Its projection capability is shown for Italy and United States with scenarios for the ensuing 30 days from April 2nd, 2020. The spread rate (relative rate of change of the time series) is very sensitive to sudden changes in the epidemic evolution and can be used to monitor in real-time the effectiveness of containment actions. The logarithm of this variable allows identifying clear trends of the evolution of the COVID-10 epidemic in these countries. The spread rate calculated from the number of confirmed cases of infection is interpreted as a probability per unit of time of virus infection and containment actions. Its product with the number of confirmed cases of infections yields the number of new cases per day. The stabilization and control of the epidemic for China and South Korea appear to occur for values of this parameter below 0.77 %/day (doubling time of 90 days).

Keywords: coronavirus, COVID-19, projection of confirmed cases, monitoring of epidemic spread

Introduction

The COVID-19 exponential pandemic has spread across the world, affecting society, health care systems and governments in different manners.¹⁻⁴ Studies about monitoring the dynamics of the COVID-19 pandemic in China and other countries are found in the literature.⁵⁻⁹ Most of them assess the evolution of the pandemic transmission process through its reproduction number, R , inferred in different moment, and correlate variations of the results with interventions conducted to contain its spread. The decline of the reproduction number is associated to intervention actions.⁵⁻⁹ Studies focusing more directly on interventions to contain the pandemic performed simulations based on the SIR and SEIR models and make assumptions about parameters of interaction between the population compartments of susceptible, exposed, infected and recovered individuals.¹⁰⁻¹³ These studies allow understanding the pandemic evolution in possible scenarios but does not tackle directly the

problem of real-time monitoring.

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

Other studies try aim at approaches to account for spatial and temporal effects of the pandemic evolution.¹⁴⁻¹⁶ Despite the difficulty to perform real-time monitoring Chen and Yu¹⁵ propose a scheme for regions in China based on extracting information from the second derivative of the time series of the total number of infections which they interpreted as an indicator of the acceleration or deceleration of the infection process. They report success in identifying the impact of containment measures on the infection spread.¹⁵ For temporal monitoring purposes, Park et al¹⁶ suggest the use of the spread rate (relative rate of change) of the time series of the incidence instead of the reproduction number which, as a unitless quantity, does not contain any information about time.

This work proposes a scheme to monitor the exponential evolution of the pandemic through the spread rate considering Park et al¹⁶ suggestion which is coincident with the industry practice with processes of similar dynamics or kinetics.¹⁷⁻²² The approach is to monitor the exponent of a key variable of the dynamic process under study.^{29,20} We take the time series of the number of confirmed cases of infection with the new coronavirus as the variable that can provide information about the dynamic behaviour of the pandemic. We interpret the results using a scheme based on the SIR model.²³ The goal is to monitor the dynamic of the pandemic in real-time and being able to construct scenarios for assessing its evolution in the following 30 days.

This article begins presenting the data, methods and the theoretical basis for interpreting the results. Following, we present the results and discussions divided in two parts: in the first, we simulate the successful control of the pandemic carried out by China and South Korea attempting to verify how well the proposed scheme monitors these two pandemic processes. In the second part, we apply the scheme to Italy and United States, and present projection results of scenarios aiming at controlling the pandemic in these countries. Finally we present the conclusions.

Data and methods

The COVID-19 epidemic spreads across a region and is detected through tests applied to the population in varied conditions.^{9,10,13} The relationship between the total number or population of infected individuals in a region, I_n , at day t_n , and the number of infected individuals counted or detected by the tests applied to the population, C_n , is

$$C_n = \varepsilon I_n, \quad (1)$$

where ε is the overall sampling scale factor of the infection detection system (counts of confirmed cases in the region/true number of infected individuals in the region) and the sample has to be representative of the true infected population.¹⁹⁻²¹ Additionally, ratios such as lethality rate based on monitored confirmed cases can provide good estimates of hospital demands and other important time varying information. In this work we take the relative rate of change or exponential growth rate of C_n , to model the pandemic temporal evolution because it is the time series with more data available in the early phase of the epidemic.

The data were obtained from the health systems of China, South Korea, Italy, and the United States which are organized as time series in the Worldometer site.⁴ Figure 1 shows the number of confirmed cases, C_n , of COVID-19 for China, South Korea, Italy, and United States for the period from January 20th to April 15th, 2020. In the abscissa day 0 is 2020/01/31. For China and South Korea it is possible to observe the flattening of the confirmed cases curve

over time, with the number of cases in China being approximately an order of magnitude higher than the South Korean in late March. For Italy and US, the infection process is already showing signs of reduction and the same curve starts to flatten but the evolution in the number of cases is still unknown.

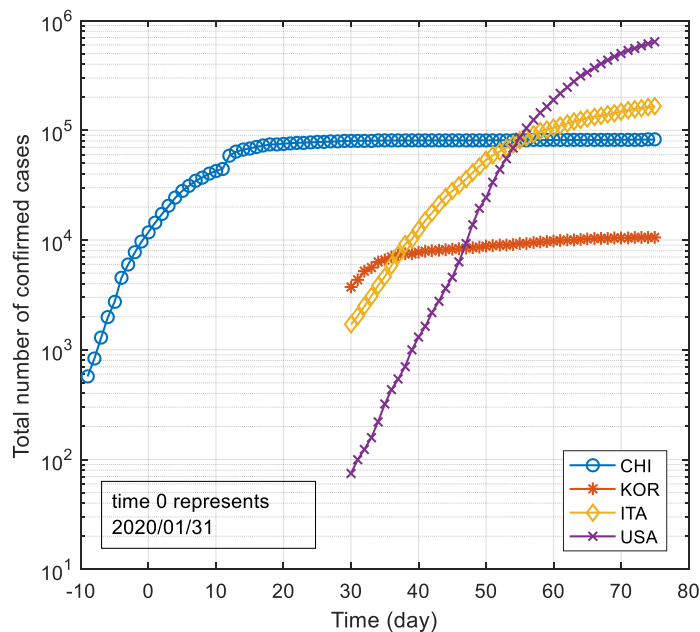


Figure 1 – Total number of COVID-19 cases detected over time (C_n) for China, South Korea, Italy, and United States.⁴

The problem posed is obtaining from these data dynamic information about the COVID-19 pandemic which is useful for those responsible to control it. The method adopted in this work follows the steps below:

- a) Construct a model to infer the temporal evolution of the pandemic using the relative rate of change over time, a_n , obtained from C_n ;
- b) Establish scenarios for the temporal evolution of the number of new cases of infection (NC_n) in the subsequent 30 days.

In the rest of the article we use the words pandemic and epidemic interchangeably.

Temporal model based on the relative rate of change of the number of cases of COVID-19

The balance equation for number of infected individuals according to the SIR model^{9,13,23} is

$$\frac{dI}{dt} = \text{rate of new cases of infection} - \text{rate of infected individuals recovered from the coronavirus} \quad (2)$$

The rate of new cases of infection depends on the action of the virus and the containment measures taken by society to avoid the spreading such as different of social isolation, quarantine, hygienic measures etc. This infection rate, $a_i(t)$, is assumed to be proportional to the number of infected individuals on day t , $I(t)$, who act as spreading agents of the infection. We can define $a_i(t)$ as a probability of new infection per unit time and per infected individual to determine the infection rate.

The rate of removal of infected individuals depends on recovery from the infection and death occurrences. If we assume that this removal rate, $a_r(t)$ is also proportional to the number of infected individuals, we can also define a probability of removal of infected individuals per unit time and per infected individual that takes into consideration all these removal processes. If we put together these two effects (new infection and removal) we obtain a probability of the net result of infection and removal per unit time and infected individual.^{21,22} Thus Eq. 2 can be written as

$$\frac{dI}{dt} = a(t) I(t) \quad (3)$$

where $a(t) = a_i(t) - a_r(t)$ is the probability of net infection and removal of infected individuals accounting for new infection of individuals, containment actions and removal of infected individuals from society. The net rate of infections appearing in the region, $NC(t)$, would be equal to the product $a(t)I(t)$.

Taking the time interval of one day, the solution of Eq. 3 is given by an exponential evolution of the number of infected individuals, that is,

$$I_n = I_{n-1} e^{\int_{t_{n-1}}^{t_n} a(u) du} = I_{n-1} e^{a_n \Delta t}, \quad (4)$$

where t_n is the n^{th} day, a_n is the average value of $a(t)$ on the n^{th} day, and I_n is the number of confirmed cases on day t_n , and $\Delta t = t_n - t_{n-1}$ or one day.

a_n is the spread rate of transmission of the epidemic, because when its value is positive I_n grows exponentially, when its value is negative, I_n falls exponentially and when it is null I_n is constant over time.¹² a_n can be estimated at t_n by

$$a_n = \frac{1}{\Delta t} \log\left(\frac{I_n}{I_{n-1}}\right) \quad (5)$$

This probability of net infection per unit of time and per infected individual would have to be obtained from a time series of the active number of infected individuals. Since this time series is not accurate in the middle of the epidemic process, we approximate the time variation of the active number of infected individuals as that of C_n . In addition, we assume that the scale factor, ϵ , is invariant over time and adequately covers the region affected by the epidemic. Thus we approximate a_n as

$$a_n \sim \frac{1}{\Delta t} \log\left(\frac{C_n}{C_{n-1}}\right), \quad (6)$$

and the number of new cases per day at t_n as

$$NC_n = a_n C_n. \quad (7)$$

It should be noted that a_n given by Eq. 6 is nothing more than the net relative rate of change in the total number of confirmed cases of infection calculated considering continuous time

variation¹¹⁻¹³ It is a partial net relative rate of infection because it accounts for the transmission action of the virus and measures to contain it such as social isolation or cleaning hands taken by the individuals. It does not account for those individuals who were infected and recovered from the illness.

The approximations of Eqs. 6 and 7 produce interesting consequences. First, since C_n is a cumulative variable it always grows. Thus a_n is always positive and the steady state condition for Eq. 3 is obtained only when $a_n = 0$. Second, due to Eq. 7, we have that a_n multiplied by C_n yields the number of new cases NC_n at time t_n .

In the rest of this article we refer to a_n interchangeably as the net probability of infection per unit of time or net relative rate of infection and express it in day^{-1} or $\%/ \text{day}$. The doubling time of infection, td_n , is related to the net rate of infection as $td_n = \frac{0,7}{a_n}$. To reduce the statistical variation of the data, a 3-day moving average was applied to the data before obtaining the net relative infection rates.

The main approximations of this monitoring scheme are:^{21,22} a) taking C_n as the active number of infected cases which is valid during the first months of the pandemic; and b) considering the scale factor, ϵ , constant throughout the study. Regarding the first approximation, it is a good approximation in the beginning of the epidemic when its growth is exponential and much bigger than the recovery rate. Regarding the second approximation, if it is not verified the analyst must consider the results with care.

Results

Monitoring of coronavirus containment actions in China and South Korea using the net relative infection rate

Figure 2 shows the net relative infection rate for China and South Korea obtained through Eq. 6 using the data presented in Figure 1. The x-axis shows the days and the day 0 corresponds to 2020/01/31. The y-axis is presented on a logarithmic scale to ease comparing the rates of change in China and South Korea along the studied period. The straight line represents a condition of net relative infection rate of 0.77 $\%/ \text{day}$ or 90 days of doubling time, a period sufficiently long to control it and plan and take actions to provide needed infrastructure to national health systems. In other words an epidemic condition in which the health systems can control it and manage its consequences.

Figure 2 shows that in China, before actions of social containment in January 2020, the net probability of infection or net relative rate of infection was 40 $\%/ \text{day}$. After the containment actions started it dropped to between 0.04 and 0.1 $\%/ \text{day}$. South Korea started rapidly its containment action (beginning of March) and its net relative rate of infection dropped and stabilized around 1 $\%/ \text{day}$ and then around 0.3 $\%/ \text{day}$. The epidemic condition of this two countries was one of very slow transmission rate (very low net relative rate of infection) with doubling times between 200 and 2000 days.

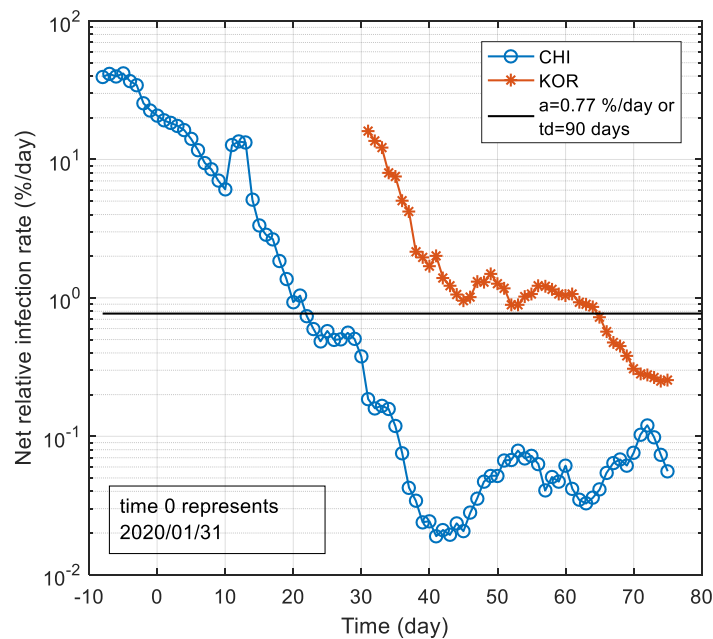


Figure 2 – Net relative infection rate as a function of time for China and South Korea.

Estimate of new cases of infection per day over time

Figure 3 compares the number of new cases per day estimated by Eq. 7 with the new cases obtained from the reported data. NC_n , given by Eq. 7, reproduces the overall behaviour of the true number of new confirmed cases per day. The root mean square of the relative discrepancy for China was 37 % and for South Korea, 51 %.

Figure 3 shows that after day 65 (mid-April), when the a_n for both countries were below the 90 day doubling time line (Figure 2), the number of new cases in China and South Korea oscillated around 60.

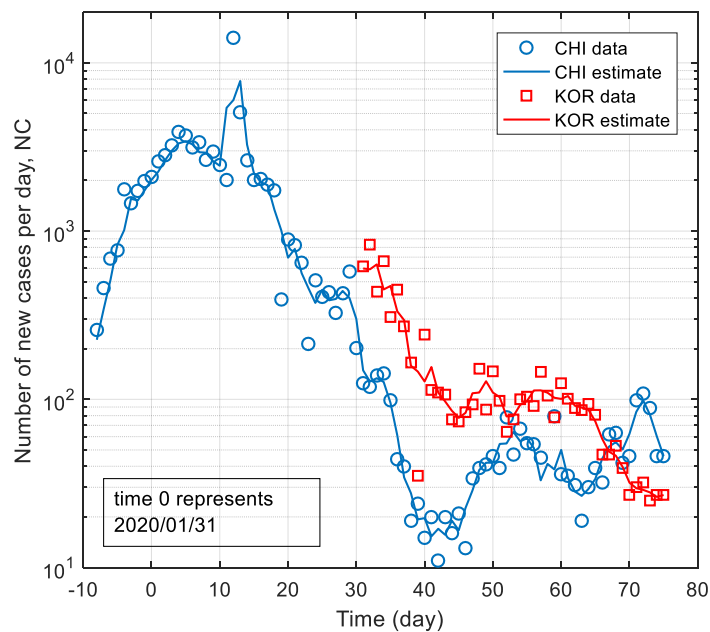


Figure 3 - Comparison between the number of detected new cases reported in the data and the rate of detected new cases estimated by Eq. 3.

Epidemic monitoring and projections for Italy and the USA

Figure 4 shows the net relative rate of infection by the coronavirus for Italy, the United States, and South Korea taken now as a successful benchmark process to control the epidemic. The curves for Italy and USA show similar behaviour when compared to China. They started with net relative rates of infection between 20 and 30 %/day and after introducing measures of social isolation their figures began to drop.

The figure also presents projections for Italy and the United States based on the linear extrapolation of the previous behaviour between the days 50 and 62, that is, between March 21st and April 2nd, 2020. The slopes of the extrapolations are -0.0856 \%/day^2 for Italy and -0.0730 \%/day^2 for the USA. The linear extrapolation of the net relative rate of infection of these two countries, considering that their respective behaviours in these 13 days are repeated, indicates that Italy and the United States cross the 90 day doubling time line around the days 80 and 100, respectively. It should be noted that such results represent a scenario that may occur only if the previous conditions persist.

The figure shows the projected estimations for C_n and NC_n for the next 30 days. The projected values for C_n and NC_n can be compared to actual data between days 62 and 75, or April 3rd and April 15th.

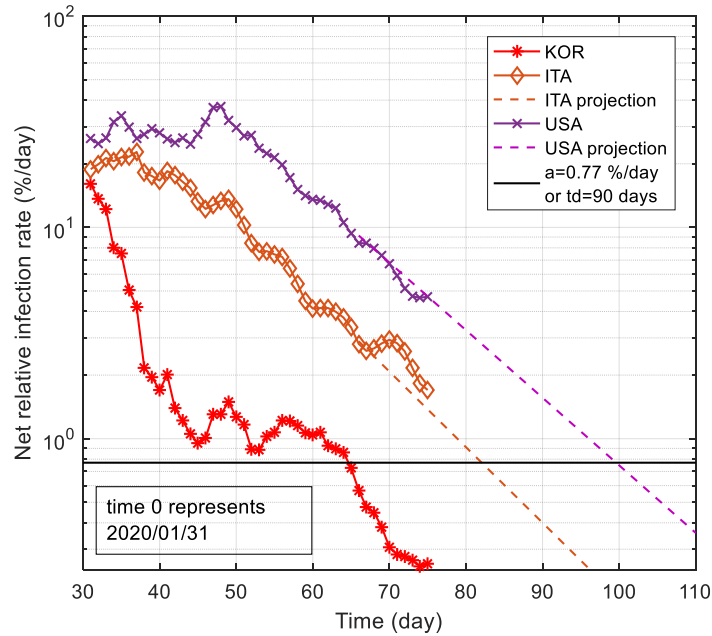


Figure 4 - Net relative infection rate as a function of time for South Korea, Italy and USA and linear projected values for Italy and USA. The projections were performed based on data from day 50 to day 63.

The projected net relative rates of infection allow estimation of future C_n and NC_n through Eqs. 4 and 7, and Figures 5 and 6 present these results for Italy and China, respectively. When Italy reaches the 0.77 %/day line on day 80 its figures for these two variables would be around 200,000 total cases and 1000 new cases per day, respectively. For the USA, on the day 100, the same figures would be around 1 million and 10,000, respectively.

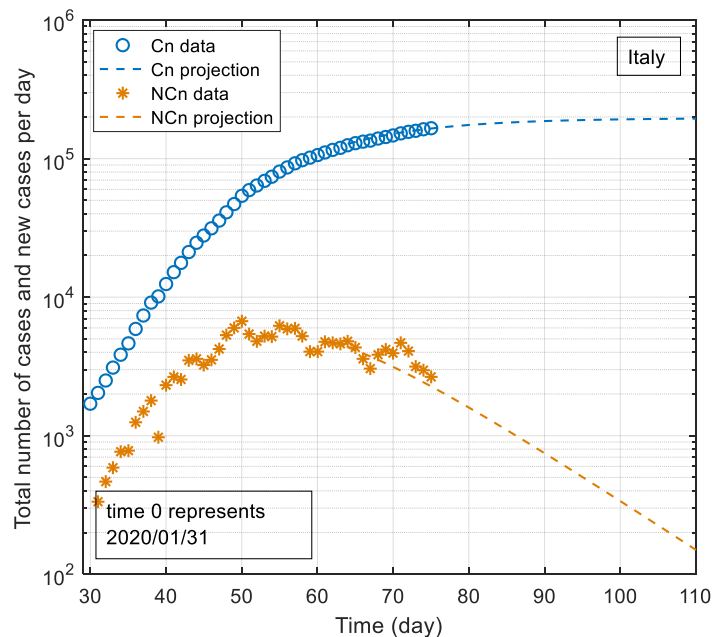


Figure 5 - Projections of the total number of confirmed cases and the number of new cases per day for Italy. The projections, in dashed lines, were calculated using Eqs. 4 and 7.

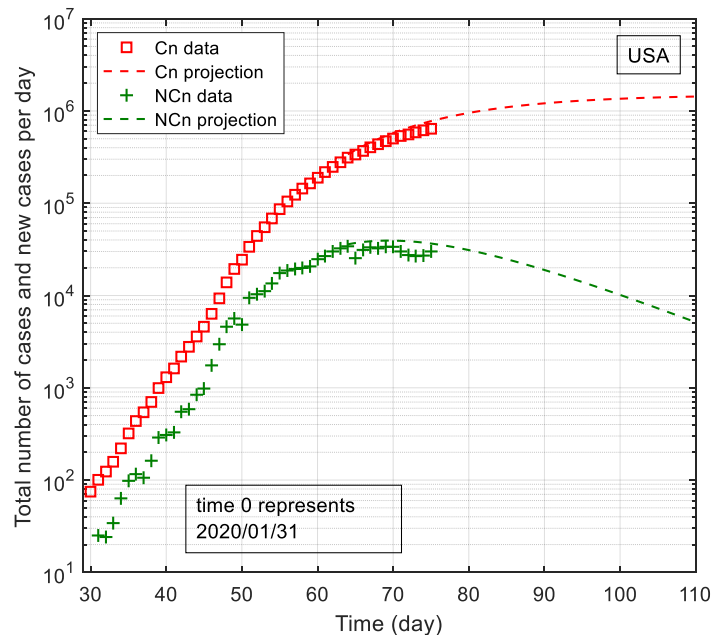


Figure 6 - Projections of the total number of confirmed cases and the number of new cases per day for the USA. The projections, in dashed lines, were calculated using Eqs. 4 and 7.

Discussion

Capability of the proposed scheme to monitor the coronavirus containment actions in China and South Korea

The net relative rates of infection for China and South Korea shown in Figure 2 allow us to monitor the effectiveness of measures to contain the spread of the virus in these two countries, i.e., how fast the relative rate of infection falls. In China, the net relative rate of infection was 40 %/day and that country started its measures to control the spread of infection in their society. At this time, the doubling time of the number of infected persons was between 2 and 3 days. The measures of social isolation were very strong and produced a continuous reduction of a_n . On February 25th, the rate was already less than 1 %/day (70 days of doubling time) and on March 8, less than 0.1 %/day (700 days of doubling time or about 2 years). On 03/27/2020 the confirmed number of infected people in China was 81394, the net relative infection rate was 0.06 %/day and the number of new cases per day, very small.

The infection reached the other countries strongly in March 2020. In South Korea, the infection process started a little earlier and the country took drastic measures to contain the spread of the virus, including social isolation. In early March, South Korea's net relative infection rate was just below 20 %/day (3.5 days of doubling time). After March 15th it stabilized around 1

%/day (70 days of doubling time) and on 03/27/2020 the confirmed number of infected persons was 9332. Why did this country decide to stabilize initially its infection rate around 1 %/day and not around 0.05 %/day as China did?

The answer to this question is the number of new cases per day, NC_n , in both countries which were between 50 and 100 and could be directly correlated to hospital demands. These figures seemed suitable to handle the consequences of the epidemic on their respective society.

The important variable is NC_n , the rate of new confirmed cases per day, obtained from the product $a_n C_n$, Eq. 7. Given that the total number of infected cases for China was around 10 times greater than that of South Korea, it need to lower its net relative rate of infection to levels 10 times lower to obtain similar numbers of new cases. It actually dropped more than that (see Figure 2).

More importantly, Figure 2 shows that the logarithm of the net relative rate of infection is very sensitive to sudden changes of the epidemic evolution and thus it can be used to monitor the effectiveness of containment actions of the epidemic. Additionally, it allows one to identify clear trends of the epidemic process and project future scenarios for its evolution.

Epidemic monitoring and projections for Italy and the USA

Figure 4 allows comparison of the evolution of the epidemic in Italy and USA with that of South Korea. The figure shows clearly that South Korea managed to contain the epidemic more rapidly than Italy and USA. It shows also that the signs of reduction of the epidemic in South Korea evidenced about 10 days before than in Italy and 15 days before than in the USA. Given the high net relative rates of infection (around 20 %/day) of Italy and USA, these delays caused important increases in their number of confirmed infected cases.

The results of Figures 2 and 4 show that the logarithmic display ease the identification of the course of the epidemic. While the number of confirmed cases and new cases per day are increasing (see Figure 1), these figures show clearly if the epidemic is being controlled or not.

Accuracy and utility of the proposed monitoring scheme

Regarding Figure 3, the relative discrepancies of 37 % and 51 % for China and South Korea data can be explained mostly by the scattering of the reported data of new cases per day. Comparing the results with 3-day average of the new cases per day data would have presented much better comparison. Figure 3 shows clearly that the estimation of NC_n for both China and South Korea suffices for understanding the dynamic or kinetics of the epidemic and for decision taking.

In Figures 5 and 6, the difference between the linear extrapolation and the actual reported data between days 63 and 75 shows typical pointwise discrepancies one may face in doing such exercises. They show that the accuracy of the scheme is sufficient to identify evolution trends of the epidemic and produces meaningful estimates for projected number of new cases of infection per day. The results show that the total number of infected cases in Italy will stabilize around 200,000 and in USA around 1,000,000.

Conclusions

The proposed scheme for monitoring the COVID-19 epidemic based on the net relative rate of change of the number of confirmed cases of COVID-19 infection presented good results. This parameter is very sensitive to sudden changes in the epidemic evolution and a plot of its logarithm as a function of time provides a good visual picture of the epidemic dynamic evolution over time. The effects of containment measures are clearly identified but it is not possible to identify which one influenced the infection dynamics. Additionally, it presents clear trends of the epidemic process which can be used to project future scenarios for its evolution.

The proposed scheme is not a model that simulates the epidemic but a monitoring scheme that infers the current net relative rate of infection based on reported confirmed cases of infection. Scenarios can be constructed by assuming future trends for net relative rate of infection.

The stabilization and control of the epidemic appears to occur for values of the net relative rate of infection below 0.77 %/day or 90 days of doubling time of the total number of confirmed cases. The specific value, however, depends on the capacity of each country or region to take care of its fellow citizens and the confirmed number of persons infected. If the latter grows significantly, the target lower net relative rate of infection has to be lower and, consequently, the effort to contain the spread of the epidemic is greater. South Korea stabilized its infection process in ~ 15 days and China in ~ 50 days.

The main approximations of this monitoring scheme, which will be taken as future studies, are taking the total number of confirmed cases as the active number of infected cases, which is valid during the first months of the pandemic; and considering the scale factor, relating the true number of infected individuals to the detected number of confirmed cases in a region, invariant throughout the evaluation.

Declaration of interests

We declare no competing interests.

Acknowledgments

The author thanks the Brazilian agency Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for financial support and acknowledges it had no interference on data collection, interpretation, or decision to submit the work for publication.

References

1. Walker PGT, Whittaker C, Watson O, et al. The global impact of COVID-19 and strategies for mitigation and suppression. WHO Collaborating Centre for Infectious Disease Modelling, MRC Centre for Global Infectious Disease Analysis, Abdul Latif Jameel Institute for Disease and Emergency Analytics, Imperial College London, 2020.
2. Khan S, Nabi G, Han G, et al. Novel coronavirus: how things are in Wuhan. Clin Microbiol Infec 2020; **26**: 399-400.

3. Lai CC, Shih TP, Ko WC, Tang HJ, Hsueh PR. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges. *Int J Antimicrob Ag* 2020; 55: 105924.
4. Worldometer. COVID-19 coronavirus pandemic. 2020. <https://www.worldometers.info/coronavirus/> (accessed in April 16, 2020).
5. Fang Y, Nie Y, Penny, M. Transmission dynamics of the COVID-19 outbreak and effectiveness of government interventions: A data-driven analysis. *J Med Virol*; 2020; **92**:645–659.
6. Kucharski AJ, Russell TW, Diamond C, et al. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. *Lancet Infect Dis* 2020; published online March 11. [https://doi.org/10.1016/S1473-3099\(20\)30144-4](https://doi.org/10.1016/S1473-3099(20)30144-4).
7. Mizumoto K, Chowell G. Transmission potential of the novel coronavirus (COVID-19) onboard the diamond Princess Cruises Ship, 2020. *Infect Dis Model* 2020; **5**: 264-270.
8. Zhang J, Litvinova M, Wang W, et al. Evolving epidemiology and transmission dynamics of coronavirus disease 2019 outside Hubei province, China: a descriptive and modelling study. *Lancet Infect Dis* 2020; published online April 2. [https://doi.org/10.1016/S1473-3099\(20\)30230-9](https://doi.org/10.1016/S1473-3099(20)30230-9).
9. Zhao S, Lin Q, Ran J, et al. Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak. *Int J Infect Dis* 2020; **92**: 214–217.
10. Berger D, Herkenhoff K, Mongey S. An SEIR Infectious Disease Model with Testing and Conditional Quarantine. Working paper. Washington Center for Equitable Growth, Washington DC, USA. 2020. <https://equitablegrowth.org/working-papers/an-seir-infectious-disease-model-with-testing-and-conditional-quarantine/> (accessed in April 7, 2020).
11. Kim S, Seo YB, Jung E. Prediction of COVID-19 transmission dynamics using a mathematical model considering behavior changes. *Epidemiol Health* 2020; published online April 13, 2020. doi: 10.4178/epih.e2020026.
12. Ming WK, Huang J, Zhang CJP. Breaking down of healthcare system: Mathematical modelling for controlling the novel coronavirus (2019-nCoV) outbreak in Wuhan, China. *bioRxiv* 2020; published online January 30. doi: 10.1101/2020.01.27.922443.
13. Read JM, Bridgen JRE, Cummings DAT, Ho A, Jewell CP. Novel coronavirus 2019-ncov: early estimation of epidemiological parameters and epidemic predictions. *medRxiv* 2020. published online January 28, 2020. doi: 10.1101/2020.01.23.20018549.
14. Wang L, Li X. Spatial epidemiology of networked metapopulation: an overview. *Chin. Sci. Bull* 2014; **59**: 3511-3522.
15. Chen X, Yu B. First two months of the 2019 Coronavirus Disease (COVID-19) epidemic in China: realtime surveillance and evaluation with a second derivative model. *Glob Health Res Policy* 2020; **5**: 7. doi: 10.1186/s41256-020-00137-4.
16. Park SW, Champredon D, Weitz JS, Dushoff J. A practical generation-interval-based approach to inferring the strength of epidemics from their speed. *Epidemics* 2019; **27**:12–18.

17. A. Molter, M. Rafikov. Nonlinear optimal control of population systems: applications in ecosystems. *Nonlinear Dynam* 2014; **76**: 1141–1150.
18. Wyse APP, Bevilacqua L, Rafikov M. Simulating malaria model for different treatment intensities in a variable environment. *Ecol Model* 2007; **206**, 322-330.
19. Moreira J, Lee JC. Space-Time Analysis of Reactor-Control-Rod-Worth Measurements. *Nucl Sci Eng* 1984; **86**: 91-105.
20. Lewis EE, *Fundamentals of Reactor Physics*. (kinetics chapter) pages 115-138. Academic Press, 1st edition, 2008.
21. Moreira, JML. Control dynamics of the COVID-19 pandemic in China and South Korea. medRxiv 2020. published online May 06, 2020. doi: 10.1101/2020.05.01.20087650.
22. Moreira, JML. Acompanhamento das medidas de contenção da infecção pelo coronavirus por meio da taxa relativa de infecção.
<https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbmxb2FvbW9yZWlyYXVmYWJfGd4OjY1MjY1ZWQxOTgzMGYxN21> (accessed in May 06, 2020).
23. Hethcote HW. Three Basic Epidemiological Models. In: Levin SA, Hallam TG, Gross LJ (eds) *Applied Mathematical Ecology*. *Biomathematics* 1989; **18**:119-144. Springer, Berlin, Heidelberg.