

Estimations of the coronavirus epidemic dynamics in South Korea with the use of SIR model

Igor Nesteruk

Institute of Hydromechanics, National Academy of Sciences of Ukraine,
Zheliabova St, 8/4, UA-03680 Kyiv, Ukraine

National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”,
Peremohy Av, 37, UA-03056, Kyiv, Ukraine
inesteruk@yahoo.com

ABSTRACT

The epidemic outbreaks in many countries caused by coronavirus COVID-19 are of great concern. A detailed scientific analysis of this phenomenon is still to come, but now it is urgently needed to evaluate the parameters of the disease dynamics in order to organize the appropriate quarantine measures, to estimate the required number of places in hospitals, etc. The corresponding mathematical models must be simple enough, since their parameters are unknown and have to be estimated using limited statistical data sets. The known SIR model, statistical approach to the parameter identification and the official WHO daily data about the confirmed cumulative number of cases in the Republic of Korea were used to calculate the SIR curves and make some estimations. New cases (connected with local transmission) could stop to appear after March 20, 2020, and the final number of such accumulated cases could be around 8120.

Keywords: coronavirus epidemic, South Korea, Europe, China, coronavirus COVID-19, coronavirus 2019-nCoV, mathematical modeling of infection diseases, SIR model, parameter identification, statistical methods.

Introduction

Here, we consider the development of epidemic outbreak in the Republic of Korea caused by coronavirus COVID-19 (2019-nCoV) (see e.g., [1]). Some investigations of the epidemic spreading in mainland China [2–7] could be useful to understand the epidemic outbreak in other countries, since we deal with the same pathogen. A preliminary comparison of the epidemic dynamics in Italy and in mainland China has been done in [8, 9]. In [10] the global coronavirus epidemic dynamics was analyzed. In this paper we will use the official WHO daily data [1] for the confirmed accumulated number of patients (victims) $V(t)$ (number of persons who caught the infection and got sick; t is time measured in days), the SIR model [11–14] and the statistics-based

method of parameter identification [14] in order to calculate the epidemic characteristics and to make some estimations.

Data

We will use official data about the accumulated number of confirmed cases in the Republic of Korea V_j from the WHO daily situation reports [1]. Since the reports show the numbers accumulated by 10 a.m. CET, we assumed that every number V_j corresponds to the previous day. The values V_j and corresponding moments of time t_j are shown in Table 1.

Day in February 2020	Time moment t_j	Accumulated number of cases in the Republic of Korea V_j , [1]	Day in March 2020	Time moment t_j	Accumulated number of cases in the Republic of Korea V_j , [1]
17	-5	31	1	8	4212
18	-4	51	2	9	4812
19	-3	104	3	10	5328
20	-2	204	4	11	5766
21	-1	346	5	12	6284
22	0	602	6	13	6767
23	1	763	7	14	7134
24	2	977	8	15	7382
25	3	1261	9	16	7513
26	4	1766	10	17	7755
27	5	2337	11	18	7869
28	6	3150	12	19	7979
29	7	3736	-	-	-

Table 1. Official cumulative numbers of confirmed cases in the Republic of Korea

The information from [1]. The corresponding time moments t_j and the accumulated confirmed numbers of cases V_j in South Korea.

SIR model and optimal values of its parameters

The SIR model for an infectious disease [6, 11–14] relates the number of susceptible persons S (persons who are sensitive to the pathogen and **not protected**); the number of infected is I (persons who are sick and **spread the infection**; please don't confuse with the number of still ill persons, so known active cases) and the number of removed R (persons who **no longer spread the infection**; this number is the sum of isolated, recovered, dead, and infected people who left the region); α and ρ are constants.

$$\frac{dS}{dt} = -\alpha SI \quad (1)$$

$$\frac{dI}{dt} = \alpha SI - \rho I \quad (2)$$

$$\frac{dR}{dt} = \rho I \quad (3)$$

To determine the initial conditions for the set of equations (1–3), let us suppose that at the moment of the epidemic outbreak t_0 , [6, 14]:

$$I(t_0) = 1, R(t_0) = 0, S(t_0) = N - 1, N = S + I + R \quad (4)$$

The analytical solution for the set of equations (1–3) was obtained by introducing the function $V(t) = I(t) + R(t)$, corresponding to the number of victims, [14]:

The solution for the SIR set of differential equations depends on four parameters N , α , $\nu = \rho / \alpha$, t_0 , which were identified with the use of the statistical approach developed in [14]. This method and V_j data set presented in Table 1 were used to define the optimal (the most reliable) values of four parameters and calculate numbers of infected I , susceptible S , removed R persons and the number of victims $V = I + R$. Corresponding dependences versus time are shown in Fig. 1.

Results and discussion

The calculated values of the optimal values of parameters are:

$$N = 68037.824; \nu = 63900.7878956716; \alpha = 8.71198727144046e-05; t_0 = -9,19521866246193$$

Now every person familiar with differential equations can use this data to integrate Eqs. (1–3) with initial conditions (4) to obtain the SIR curves and to check the results of calculations (it is also possible to use the analytical solution available in [6]). We have used fractional values of the parameter N to avoid accuracy losses, although only integer numbers of persons are of practical

importance. The values of final numbers susceptible persons $S_\infty \approx 59921$ and final number of victims (final accumulated number of cases) $V_\infty \approx 8117$ were calculated.

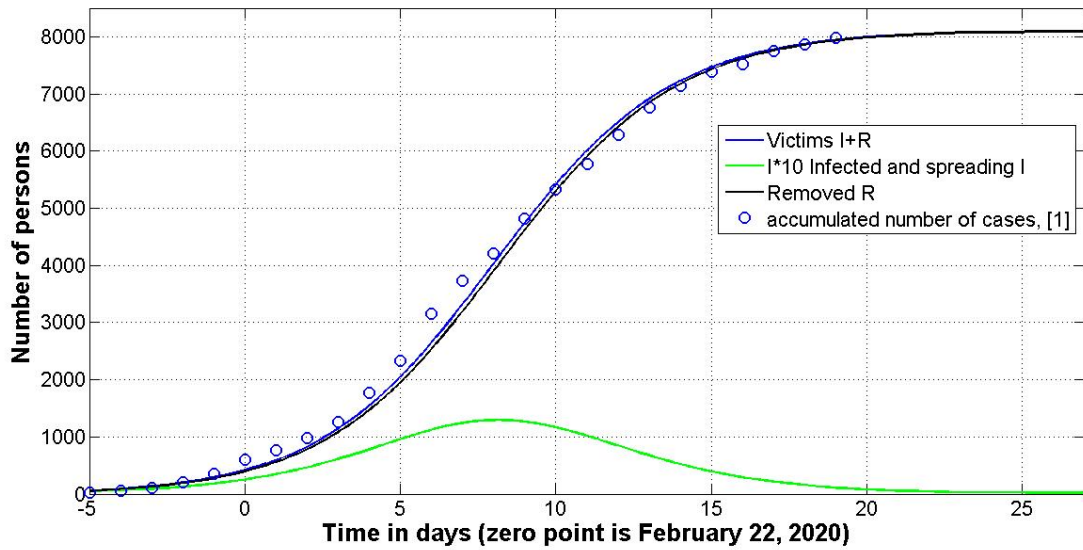


Fig. 1. Results of calculations.

Numbers of infected I multiplied by 10 (the green line), removed R (the black line) and the number of victims $V=I+R$ (the blue line). The “circles” correspond to the confirmed accumulated number of cases in the Republic of Korea shown in Table 1, [1].

To estimate the duration of the coronavirus epidemic outbreak in South Korea, we can use the condition $V(t_{final})=1$ which means that after this moment less than one person still spread the infection. The calculations give us the value $t_{final} \approx 26.07$. According to this estimation, we can expect that local transmission of the epidemic in South Korea could stop after March 20, 2020, provided that existing quarantine measures and patient isolation rates continue. Unfortunately, too many infected persons are now in many countries. Some of them could come to South Korea and worsen these optimistic estimations.

It is also possible to calculate the value of parameter $\rho = \nu\alpha = 5.567$ and the inverse value $1/\rho = 0.1796$. Thus, the average time of spreading the infection in South Korea is approximately 4.3 hours. By comparison, in mainland China it was approximately 2.5 days [9]. That is why the $V(t)$ and $R(t)$ curves are very close in Fig. 1 and the maximal number of persons spreading the infection $I(t)$ was less than 130 (to make the corresponding curve more visible in Fig. 1, this value is multiplied by 10).

Probably due to the rapid isolation of infected persons the expected final number of cases in South Korea ($V_\infty \approx 8117$) is ten times smaller than the estimation for mainland China ($V_\infty \approx 81257$, [9]). Rapid isolation allows patients to start treatment more quickly and reduce mortality. According to the WHO report [1] (March 10, 2020), the mortality rate in South Korea was 0.77%. By

comparison, it was 6.22% in Italy. Such a huge difference in mortality rate can be explained by not so rapid isolation in Italy. It may also be that a more virulent coronavirus strain is spreading in Italy.

All the parameters in SIR model are supposed to be constant. If new infected persons are still coming in the country and spread the infection, the value N is not constant. This fact can reduce the accuracy of prediction. If these “new imported” and connected with them cases can be calculated and extracted from the total number, the accuracy of the prediction could be improved.

Conclusions

The characteristics of the coronavirus epidemic outbreak in the Republic of Korea were estimated using the SIR model and statistical methods of the parameter identification. The expected final number of cases is approximately 10 times smaller than predicted for mainland China. It seems that the secret of South Korean success lies in the rapid isolation of patients.

References

1. World Health Organization. “Coronavirus disease (COVID-2019) situation reports”. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/>. Retrieved Mar. 14, 2020.
2. Li, Q., Guan, X., Wu, P., et al. “Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia.” *The New England journal of medicine*, 2020 Jan. 29. doi:10.1056/NEJMoa2001316.
3. Wu, J. T., Leung, K., & Leung, G. M. “Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study.” *Lancet*, vol. 395, no. 10225, pp. 689–697. 2020 Jan. 31. doi:10.1016/S0140-6736(20)30260-9.
4. 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.” *International journal of infectious diseases*, vol. 92, pp. 214–217. 2020 Jan. 30. doi:10.1016/j.ijid.2020.01.050.
5. Tang, B., Bragazzi, N. L., Li, Q., et al. “An updated estimation of the risk of transmission of the novel coronavirus (2019-nCoV).” *Infectious disease modelling*, vol. 5, pp. 248–255. 2020 Feb. 11. doi:10.1016/j.idm.2020.02.001.
6. Nesteruk, I. “Statistics-based predictions of coronavirus epidemic spreading in mainland China.” *Innovative biosystems and bioengineering*, vol. 4, no. 1, pp. 13–18. 2020. doi:10.20535/ibb.2020.4.1.195074.
7. Batista, M. “Estimation of the final size of the COVID-19 epidemic.” [Preprint.] *medRxiv*. 2020 Feb. 28. doi:10.1101/2020.02.16.20023606.
8. Nesteruk, I. “How can we estimate the dangers of the coronavirus epidemic in Europe?” [Preprint.] *ResearchGate*. 2020 Feb. doi:10.13140/RG.2.2.32072.06402.

9. Nesteruk, I. "Characteristics of coronavirus epidemic in mainland China estimated with the use of official data available after February 12, 2020." [Preprint.] *ResearchGate*. 2020 Mar. doi:10.13140/RG.2.2.19667.32804.
10. Li, Y., Liang, M., Yin, X., et al. "COVID-19 epidemic outside China: 34 founders and exponential growth." [Preprint.] *medRxiv*. 2020 Mar. 5. doi:10.1101/2020.03.01.20029819.
11. Kermack, W. O. & McKendrick, A. G. "A contribution to the mathematical theory of epidemics." *Proceedings of the Royal Society, Ser. A*, vol. 115, pp. 700–721. 1927.
12. Murray, J. D. *Mathematical biology*, 3rd ed. 2 v. New York : Springer, 2002–2003.
13. Langemann, D., Nesteruk, I. & Prestin, J. "Comparison of mathematical models for the dynamics of the Chernivtsi children disease." *Mathematics in computers and simulation*, vol. 123, pp. 68–79. 2016. doi:10.1016/j.matcom.2016.01.003.
14. Nesteruk, I. "Statistics based models for the dynamics of Chernivtsi children disease." AMMODIT Conference, Kyiv, Ukraine, January 2017. *Naukovi visti NTUU KPI*, 2017, no. 5, pp. 26–34. doi:10.20535/1810-0546.2017.5.108577.