Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company’s public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Forecasting of COVID-19: spread with dynamic transmission rate

Yiping Zeng, Xiaojing Guo, Qing Deng, Shengfeng Luo, Hui Zhang∗

Institute of Public Safety Research, Department of Engineering Physics, Tsinghua University, Beijing 100084, China

A R T I C L E   I N F O

Keywords:
COVID-19
SEIR
Emergency management
Dynamic transmission rate
Basic reproduction number

A B S T R A C T

The COVID-19 was firstly reported in Wuhan, Hubei province, and it was brought to all over China by people travelling for Chinese New Year. The pandemic coronavirus with its catastrophic effects is now a global concern. Forecasting of COVID-19 spread has attracted a great attention for public health emergency. However, few researchers look into the relationship between dynamic transmission rate and preventable measures by authorities. In this paper, the SEIR (Susceptible Exposed Infectious Recovered) model is employed to investigate the spread of COVID-19. The epidemic spread is divided into two stages: before and after intervention. Before intervention, the transmission rate is assumed to be a constant since individual, community and government response has not taken into place. After intervention, the transmission rate is reduced dramatically due to the societal actions or measures to reduce and prevent the spread of disease. The transmission rate is assumed to follow an exponential function, and the removal rate is assumed to follow a power exponential function. The removal rate is increased with the evolution of the time. Using the real data, the model and parameters are optimized. The transmission rate without measure is calculated to be 0.033 and 0.030 for Hubei and outside Hubei province, respectively. After the model is established, the spread of COVID-19 in Hubei province, France and USA is predicted. From results, USA performs the worst according to the dynamic ratio. The model has provided a mathematical method to evaluate the effectiveness of the government response and can be used to forecast the spread of COVID-19 with better performance.

1. Introduction

On December 8, 2019, a case of unexplained pneumonia was reported in Wuhan. The virus was brought to Hubei province and China by people travelling for Chinese New Year. After the outbreak of new pneumonia, the coronavirus disease was named as COVID-19 by the World Health Organization on February 11, 2020 [1]. In order to control and stop the spread of COVID-19, Chinese National Health Commission took strict measures to lockdown the Wuhan city and all transportation was suspended to prevent human-to-human contact on February 23, 2020 [2].

Epidemiological modeling plays a vital role in the early warning and prevention of outbreaks, such as Severe Acute Respiratory Syndrome (SARS-Cov) [3], Middle East Respiratory Syndrome (MERS-Cov) [4], Ebola Virus [5, 6], Zika virus [7, 8] and so on. Until now, a lot of researches were performed on COVID-19 [9-11]. Yang et al. [10] calculated the basic reproduction number and the death rate by analyzing the data from infected people, and they found that the basic reproduction number was about 3.77 and the mean incubation period was estimated to be 4.75 days. Li et al. [12] analyzed the data for the first 425 confirmed cases in Wuhan for the purpose of investigating the epidemiologic characteristics of COVID-19. They found that the mean incubation period was 5.2 days (95% confidence interval, 4.1 to 7.0), with the 95th percentile of the distribution at 12.5 days. Zhong et al. [13] collected data from 1099 patients confirmed with COVID-19 in China. Through data analysis, they found that the median incubation period was 4 days (interquartile range, 2–7). All studies help us better understand the COVID-19 and find the suitable methods to prevent virus and cure individuals.

Researchers mentioned above focused on the clinic characteristics of COVID-19, while others paid attention to modelling and prediction, which could also provide reference for the management of anti-virus. A majority of researchers [1,14] modelled and reproduced the spread process of virus using the original or modified SIR and SEIR models. Natsuko et al. [15] estimated the potential number of novel coronavirus cases in Wuhan. From results, there were a total of 1723 cases of COVID-19 with onset of symptoms by 12th January 2020. Based on the susceptible-exposed-infected-removed (SEIR) compartment model, Zhou et al. [14] found that the basic reproduction number ranged from 2.8 to 3.5 with the help of dataset reported on the People’s Daily in China. The predicted value fell between 3.2 and 3.9. Xiong et al. [16] analyzed the infected population and spread trend of COVID-19 under different policy with the help of SEIR model, and they found that the epidemic spreading was dominated by the quarantine rate and starting
date of intervention. Ming et al. [17] explored the effect of COVID-19 on healthcare system using mathematical modelling, and found that if there was no effective intervention, the healthcare system burden would be increased with the increased confirmed cases.

However, few researchers have taken the dynamic transmission rate into consideration because of varied preventable measures by authorities. In reality, after actions taken by authorities, cities are in lockdown and individuals need to stay at home, resulting in the decrease of the transmission rate. In this paper, the spread of COVID-19 is divided into two stages: before and after intervention. Before intervention, the transmission rate is assumed to be a constant since individual, community and government response has not taken into place. After intervention, the transmission rate is reduced dramatically due to the societal actions or measures to reduce and prevent the spread of disease. The transmission rate is assumed to follow the exponential function. In this paper, the original and modified SEIR models are briefly introduced in terms of the transmission rate and removal rate in Section 2. In Section 3, based on the least square method, the improved model is optimized by considering accumulated number of infected individuals and daily new cases. Then we compare and discuss the performance between the original and modified models. Using the modified model, the spread of COVID-19 in Hubei province, France and USA is predicted and compared. Conclusions are made in the last section.

2. Mathematical method

2.1. The original SEIR model

The original SEIR (Susceptible Exposed Infectious Recovered) model is widely used to predict the spread of epidemic disease. S is the susceptible individuals, and the susceptible individuals S have a probability β to enter the exposed class E after they meet individuals with epidemiological virus by close contact. After some days without any obvious features (the incubation period), some exposed individuals E have a probability of α to show some characteristics of epidemic disease, that is infected individuals. Infected individuals I either recover or die, which will be removed from the system. The rate to remove is called γ. The following formula describe the spread process of epidemic disease:

\[
\begin{align*}
\frac{dS}{dt} &= -\beta IS / N \\
\frac{dE}{dt} &= \beta IS / N - aE \\
\frac{dI}{dt} &= aE - \gamma I \\
\frac{dR}{dt} &= \gamma I
\end{align*}
\]

where r is the average number of contacts per person per day.

2.2. The modified SEIR model

In the case of COVID-19, individuals S enter the exposed class E because of exposed individuals E and infected individuals I. However, the original model does not consider the fact that exposed individuals have an ability to infect susceptible individuals, which is one of the most important factors for epidemic disease spread. In the modified model, the individuals are affected by exposed individuals E with a probability of β₂ and infected individuals I with β₁.

\[
\begin{align*}
\frac{dS}{dt} &= -r₁β₁ IS / N - r₂β₂ ES / N \\
\frac{dE}{dt} &= r₁β₁ IS / N - aE + r₂β₂ ES / N \\
\frac{dI}{dt} &= aE - \gamma I \\
\frac{dR}{dt} &= \gamma I
\end{align*}
\]

where r₁ is the average number of individuals with whom an infected individual is confronted by close contact. r₂ is for the exposed individuals.

In order to prevent the spread of epidemiological virus, authorities took various measures, such as to lock down city, to control traffic, to wear facemask, to educate the public on the knowledge of the disease, to require potential patients to stay in hospital or stay at home, and so on. The implemented measures can control and stop the spread of virus and save the life. On the other words, after measures, the transmission rates β₁ and β₂ begin to decrease. In our model, we assume that the transmission rate following the following formulation:

\[
\beta(t) = \begin{cases} 
\beta_0 & t < r \\
\beta_0 e^{-(k-r)} & t \geq r
\end{cases}
\]

where r is the time when the measures are taken to intervene the virus and k is a constant value to control the transition rate.

The removal rate in the original model is kept as a constant. However, in reality, as the time progresses, medicine and therapy for curing patients are found, the death rate is reduced gradually. The removal rate will increase with the passage of the time. At the same time, from the data collected, we can find the recovery rate is larger than the death rate. Therefore, the following formulation is proposed to show the relationship between the removal rate γ(t) and time:

\[
γ(t) = a + b^t
\]

where a and b are constant values in the model.

2.3. Reproduction number

The basic reproductive number R₀ measures the average number of secondary people infected by a primary patient in a pool of mostly susceptible individuals in absence of controlling measures [18] and it is a parameter to estimate the epidemic spread in a sealed group [19]. For any initial level of epidemic disease, it is going to disappear from the population in the infected area when R₀ is smaller than 1. R₀ is larger than 1, which implies that disease is spreading in the population [20]. There are many ways [21,22] to calculate the R₀ in terms of formula derivation and model fitting. In our model, the basic reproduction number R₀ is estimated by the formula of R₀ = β₁/γ for the purpose of simplifying the model. Furthermore, because of measures taken by authorities, the basic reproduction number R₀ may vary with the passage of the time. By considering transmission rate β(t) and removal rate γ(t), the effective reproduction number Rₑ(t) is estimated by the following formulation:

\[
Rₑ(t) = \frac{β(t)}{γ(t)}
\]

2.4. Model hypothesis

In order to simplify the model, some hypotheses are made, as shown in Fig. 1:

1) The exposed individuals and infected individuals have same probability to infect susceptible individuals, that is β₁=β₂;
2) There is no pedestrian flow between Hubei and outside Hubei, and COVID-19 spreads in the corresponding area;
3) Removed individual from the system has no ability to infect others;
4) The transmission rate β is assumed to follow an exponential function considering the fact that fewer individuals are infected after measures are in place;
5) The removal rate γ is supposed to follow a power exponent function, and the removal rate increases as the time processes due to the better treatment.
6) The basic reproduction number R₀ changes with the time because of measures taken by authorities. The effective reproduction number is calculated by Rₑ(t)= β(t)/γ(t).

3. Results and discussion

The National Health Commission of the People’s Republic of China published the accumulative number of the confirmed cases and daily new cases on the official website. By the use of R package [23], we acquired the datasets of Hubei province in respect of accumulative confirmed cases and daily new cases from January 11 to March 19, 2020.
Table 1
Parameters setting up of the models.

| Parameters                                      | Indicators | Hubei Province | Outside Hubei Province |
|-------------------------------------------------|------------|----------------|------------------------|
| Initial total number of individuals             | N          | 591 million    | 1.34 billion           |
| The number of susceptible                       | S          | N – I – E – R  | N – I – E – R          |
| Initial number of exposed individuals           | E          | 410            | 210                    |
| Initial number of infected individuals          | I          | 41             | 21                     |
| Initial removal individuals                     | R          | 0              | 0                      |
| Daily transmission rate from infected individuals| \( \beta_1 \) |                |                        |
| Daily transmission rate from exposed individuals | \( \beta_2 \) |                | \( \beta(t) = \left\{ \begin{array}{ll} \beta_0 & t < t_c \\ \beta_0 e^{\beta(t_c-t)} & t \geq t_c \end{array} \right. \) |

Fig. 1. Illustration of the hypotheses.

For out of Hubei province, the first case was reported on January 20. The data of outside Hubei province in China was obtained from January 20 to March 19, 2020 (Table 1).

The interest is shifted to the possible range of parameters in the model. First of all, the number of infected patients is needed to be confirmed. According to datasets from the National Health Commission of China and reports by the Health Commission of Hubei province, the number of infected people was 41 in Hubei province and the first day was set up as January 11. For the data out of Hubei province, January 20 is set up as the first day. Furthermore, the number of infected individuals on the first day is 21. For the initial exposed individuals, it is difficult to estimate the number of exposed individuals due to medical techniques. Some researchers thought that the rate (infected: exposed) is about 13%. According to the experience, the rate is set to be 10%. Initial exposed number of Hubei province is therefore set to be 210 and it is equal to 410 for outside Hubei province. No individual recovers from the disease on the first day. All initial removal individuals are therefore set to be zero. There is no way or method to define the average number of individuals who an infected individual meets, so \( r_1 \) is set to be a value ranging from 0 to 10 according to the value used in SARS. The exposed individuals have a big opportunity to meet others by close contact because of the incubation period, \( r_2 \) is set to be within [0, 35]. At last, the transmission rate from the exposed individuals to infected individuals \( \alpha \) is set from 0 to 0.6. In addition, with the help of the least square method, other unknown parameters are optimized to fit the real data, which is shown in Table 2.

Because there is a big difference between Hubei and outside Hubei provinces. Based on the dataset, the model is fitted by simulation. Figs. 2 and 3 show the fitted and predicted data of Hubei and outside Hubei provinces, respectively. From Fig. 2, the spread of epidemiological virus is controlled at the start of March because of fewer new cases in Hubei. It is also found that on January 20, there are fewer infected individuals from Fig. 3. Furthermore, the fitted parameters are obtained and shown in Table 2. Firstly, the transmission rate without measures \( \beta_0 \) is equal to 0.033 and 0.030 for Hubei and outside Hubei provinces. Thus the virus in Hubei has a larger probability to in-
Table 2
Estimation of model parameters by use of datasets from website.

| Parameters                                                                 | Indicators | Range                        | Hubei Province | Outside Hubei Province |
|---------------------------------------------------------------------------|------------|------------------------------|----------------|------------------------|
| Average number of contacts per infected individual per day                | $r_1$      | $[0,10]$                     | 4.10E-13       | 1.45E-11               |
| Average number of contacts per exposed individual per day                 | $r_2$      | $[0,35]$                     | 4.10E-13       | 1.45E-11               |
| Transmission rate from exposed individuals to infected individuals        | $\alpha$  | $[0,0.06]$                   | 20             | 24                     |
| Transmission rate without measures                                       | $\beta_0$ | $[0.05]$                     | 0.53           | 0.400                  |
| The time when measures are taken                                          | $\tau$    | $-4$                         | 13             | 4                      |
| A constant value to control the transition rate.                         | $k$        | $[0.2,0.01]$                 | 0.014          | 0.059                  |
| Constant value for removal rate $\gamma(t)$                             | $a$        | $[0.01,0.006]$               | 0.006          | 0.006                  |
| Constant value for removal rate $\gamma(t)$                             | $b$        | $[0.01,0.006]$               | 0.006          | 0.005                  |

![Fig. 4. Effective reproduction number (a) Hubei province; (b) Outside Hubei province.](image)

![Fig. 5. Comparison of daily new cases in Hubei province between the original and modified models.](image)

![Fig. 6. Comparison of daily new cases outside Hubei province between the original and modified models.](image)

...fect other individuals. The same rule is suitable for the transmission rate from the exposed individuals to infected individuals $\alpha$.

From Figs. 2 and 3, accumulated infected individuals and daily new cases can be predicted. Variation of effective reproduction number with time can be obtained. Fig. 4 shows the effective reproduction number $R_e$ with passage of the time in Hubei and outside Hubei provinces, respectively. From Fig. 4, $R_e$ is decreased slightly before January 23 due to the increase of the treatment. Then $R_e$ declines suddenly because authorities implement several control measures and individuals are alerted to prevent epidemiological virus. Measures and treatments result in decreasing effective reproduction number with the passage of the time. Fig. 4 shows that virus is about to fade way in May in Hubei province.
Also from Fig. 4, virus outside Hubei province is going to die out earlier due to secondary or third generation of virus.

The results from the original and modified models are compared in terms of $R_e$. From Fig. 4, $R_e$ from the original model is larger than that from the modified model. $R_e$ is decreased slightly as the time progresses, which cannot perform well because of preventable measures taken by authorities. The daily new cases are compared in Hubei and outside Hubei provinces, respectively. From Figs. 5 and 6, it is found that daily new cases by the original model are smaller than those by the modified model at the early period of virus spread. Underestimation in virus contributes to less attention to the virus, which can result in large damage and casualties for any countries and regions.

Our model is used to fit the data in France (dataset from [24]) and USA (dataset from [25]). The ratio (Daily new cases: accumulated infected individuals) is used to estimate the spreading of COVID-19. The ratio is smaller than 0.1, which means that the virus is under control. In Fig. 7, the label (actual infected population) means the number of cases confirmed with COVID-19 in reality. It is found that, in real life, USA performs the worst. From Fig. 7, the spread of COVID-19 in Hubei province is under control firstly.

4. Conclusion

In this paper, a modified model is developed to better predict the spread of COVID-19 considering the dynamic change of the control measures and treatment. In our model, the transmission rate $\beta$ is assumed to follow exponential function by considering the fact that fewer individuals are infected after measures to prevent the virus spread. Then the removal rate $\gamma$ is supposed to follow a power exponent function. The removal rate is increased with time because of better cure for disease. Based on real data, we optimize the model parameters using the least square method. Transmission rate without measures $R_0$ is equal to 0.033 and 0.030 for Hubei and outside Hubei. The results from the original and modified models are compared in terms of effective reproduction number $R_e$ and daily new cases. It is found that daily new cases obtained by the original model are smaller than those by the modified model at the starting spread of epidemiological virus. Fewer infected individuals contribute to less attention to the virus, which might result in large damage and casualty. Furthermore, the model is used to evaluate the coronavirus spread of Hubei province, France and USA. USA performs the worst according to the ratios. The model has provided a mathematical method to evaluate the effectiveness of the government response and can be used to forecast the spread of COVID-19 with better performance.

Declaration of Competing Interest

The authors declare that they have not known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work is supported by National Key R and D Program of China (No. 2017YFC0803300), National Science Foundation of China (Grant Nos. 7204100828, 91646201, U1633203), High-tech Discipline Construction Funding for Universities in Beijing (Safety Science and Engineering) and Beijing Key Laboratory of City Integrated Emergency Response Science.
References

[1] J.T. Wu, K. Leung, G.M. Leung, Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study, Lancet N. Ed. 395 (10225) (2020) 689–697.

[2] http://www.nhc.gov.cn/xcs/yqfkdt/gzbd_index.shtml

[3] Y. Zhou, Z. Ma, F. Brauer, A discrete epidemic model for SARS transmission and control in China, Math. Comput. Model. 40 (13) (2004) 1491–1506.

[4] G. Chowell, S. Blumberg, L. Simonena, M.A. Miller, C. Viboud, Synthesizing data and models for the spread of MERS-CoV, 2013: key role of index cases and hospital transmission, Epidemics 9 (2014) 40–51.

[5] W.E.R. Team, Ebola virus disease in West Africa—the first 9 months of the epidemic and forward projections, N. Engl. J. Med. 371 (16) (2014) 1481–1495.

[6] S.E. Bellan, J.R. Pulliam, J. Dushoff, L.A. Meyers, Ebola control: effect of asymptomatic infection and acquired immunity, Lancet N. Am. Ed. 384 (9953) (2014) 1499–1500.

[7] S. Towers, F. Brauer, C. Castillo-Chavez, A.K. Falconar, A. Mubayi, C.M. Romero-Vivas, Estimate of the reproduction number of the 2015 Zika virus outbreak in Baranquilla, Colombia, and estimation of the relative role of sexual transmission, Epidemiics 17 (2016) 50–55.

[8] L.R. Petersen, D.J. Jamieson, A.M. Powers, M.A. Honein, Zika virus, N. Engl. J. Med. 374 (16) (2016) 1552–1563.

[9] N. Chen, M. Zhou, X. Dong, J. Qu, F. Gong, Y. Han, Y. Qiu, J. Wang, Y. Liu, Y. Wei, Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study, Lancet N. Am. Ed. 395 (10225) (2020) 507–513.

[10] Y. Yang, Q. Lu, M. Liu, Y. Wang, A. Zhang, N. Jalali, N. Dean, I. Longini, M.E. Halloran, B. Xu, Epidemiological and clinical features of the 2019 novel coronavirus outbreak in China, medRxiv (2020).

[11] S. Darougar, M. Wishart, N. Vinwalingam, Epidemiological and clinical features of primary herpes simplex virus ocular infection, Br. J. Ophthalmol. 69 (1) (1985) 2–6.

[12] Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K.S. Leung, E.H. Lau, J.Y. Wang, Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia, N. Engl. J. Med. (2020).

[13] W.-J. Guan, Z.-y. Ni, Y. Hu, W.-h. Liang, C.-q. Ou, J.-x. He, L. Liu, H. Shan, C.-l. Lei, D.S. Hui, Clinical characteristics of coronavirus disease 2019 in China, N. Engl. J. Med. (2020).

[14] T. Zhou, Q. Liu, Z. Yang, J. Liao, K. Yang, W. Bai, X. Lu, W. Zhang, Preliminary prediction of the basic reproduction number of the Wuhan novel coronavirus 2019-nCoV, J. Evid.-Based Med. (2020).

[15] N. Imai, I. Dorigiatti, A. Costi, S. Riley, N.M. Ferguson, Estimating the potential total number of novel Coronavirus cases in Wuhan City, China, 2020.

[16] H. Xiong, H. Yan, Simulating the infected population and spread trend of 2019-nCov under different policy by EIR model, Available at SSRN 3537083 (2020).

[17] W.-K. Ming, J. Huang, C.J. Zhang, Breaking down of healthcare system: Mathematical modelling for controlling the novel coronavirus (2019-nCoV) outbreak in Wuhan, China, bioRxiv (2020).

[18] F. Brauer, C. Castillo-Chávez, Mathematical models in population biology and epidemiology.

[19] G. Chowell, N.W. Hengartner, C. Castillo-Chavez, P.W. Fenimore, J.M. Hyman, The basic reproductive number of Ebola and the effects of public health measures: the cases of Congo and Uganda, J. Theor. Biol. 229 (1) (2004) 119–126.

[20] A. Korobeinikov, Global properties of SIR and SEIR epidemic models with multiple parallel infectious stages, Bull. Math. Biol. 71 (1) (2009) 75–83.

[21] A. Khan, M. Naved, M. Dur-E-Ahmad, M. Imran, Estimating the basic reproductive ratio for the Ebola outbreak in Liberia and Sierra Leone, Infect. Dis. Poverty 4 (1) (2015) 13.

[22] C. Kribs, G. Chowell, C. Castillo-Chavez, P.W. Fenimore, L. Arriola, J.M. Hyman, Model parameters and outbreak control for SARS, (2004).

[23] T. Wu, X. Ge, G. Yu, An R package and a website with real-time data on the COVID-19 coronavirus outbreak, medRxiv (2020).

[24] https://gisanddata.maps.arcgis.com/apps/opsdashboard/index.html#/bda7594740d40299423467b848f0eaec6, (2020).

[25] https://www.worldometers.info/coronavirus/country/us/, (2020).