The Epidemic Prediction of Coronavirus disease 2019 Based on a SEIR(Susceptible-Exposed-Infected-Removed) Model Considering Population Migration between Wuhan and other Provinces in China

Anran Wang, Ruobing Zhan, Fen Yu, Jie Mi and Kaiye Gao*

1School of Economics and Management, Beijing Information Science and Technology University, Beijing, 100192, China

*Corresponding author e-mail: kygao@foxmail.com

Abstract. As the first city detected the novel coronavirus in China, Wuhan was quarantined since Jan 23, 2020. The measure effectively prevented the epidemic spreading between provinces in China. This article proposes the SEIR infectious disease dynamics model introduced the inter-provincial migration factor, which is called the SEIRM model. The proposed model analyses the epidemic situation of 7 kinds of people when they can migrate freely between Wuhan and other provinces in China. Using the proposed model, we predict that until March 1st, 2020, the top five most severely epidemic-affected provinces are Hubei Province, Guangdong Province, Henan Province, Hunan Province, and Jiangxi Province; it is also predicted that more than 60.25 million people will be infected and approximately 837 thousand people will die due to this epidemic, which will exceed the medical resources load in China. According to our results, the population migration during the Spring Festival will quickly lead to a large-scale epidemic outbreak in China, which implies that locking down Wuhan and quarantine measures nationwide have a positive effect on China's epidemic prevention. Epidemic prevention measures will effectively prevent the spread of the epidemic, quickly restore social and economic order, and protect the health of the people of the country.

1. Introduction
Corona Virus Disease 2019 (hereinafter called COVID-19) is an emerging respiratory disease that is caused by a novel coronavirus and was first detected in December 2019 in Wuhan, China. The ongoing COVID-19 epidemic has spread rapidly[1]. Until January 23th, government authorities had locked down the whole city of Wuhan, and Chinese residents, both inside and outside of Wuhan, were also required to just stay at home to avoid contacting with others. In this way, the population movement was totally frozen. From the practical results, it seems that the effect of limiting population migration was powerful in preventing the transmission of the epidemic from Wuhan to other provinces. However, there is still room to know the true effect of banning trans provincial migration compared with no banning measures. Thus, it is needed to predict the situation of the COVID-19 epidemic under the condition that people can freely migrate between Wuhan and other provinces in China.

Mathematical models of infectious disease transmission dynamics are now ubiquitous. Such
models play an important role in helping to quantify possible infectious disease control and mitigation strategies [2]. Epidemical models are classified into single population, meta-population, and microscopic models [3]. In the single population method, the population is regarded as a whole, and the epidemic process is reflected in the changes in the number of susceptible, infected, and other groups; The meta-population model considers the spatial heterogeneity of the population and divides the population into several subgroups. The subgroups are coupled due to the flow of personnel to form a complex dynamic system; The starting point of microscopic models is the individuals in the crowd. Individuals have their own attributes and behavior rules. The contact network is formed between individuals. The contact between the infected and susceptible leads to the change of susceptible individuals’ state, which forms the transmission dynamic process on the network.

Since the single population model can not reveal the impact of human interaction on the spread of disease, and the microscopic model requires too much data accuracy and calculation complexity, this paper uses the meta-population method to establish the warehouse model, and the subgroups are coupled with full connection structure. Therefore, we need to use the classical compartment model with latent infectious diseases, the SEIR model.

There have been some researches in the academic circle concerning the effect of limiting population migration in Wuhan to the COVID-19 situation in China. Jinghua Li used five independent mathematical models to accurately describe the epidemic characteristics of the virus, which proved that the basic reproduction number (R0) of the virus decreased significantly after the closure of Wuhan [4]. Tian used a univariate linear regression model combined with epidemiological and population mobility data and obtained COVID-19 can be delayed 2.9 days from Wuhan city to other cities by restricting travel [5]. Shao established the susceptible-exposed-infected-removed-dead model (SEIRD model) which considers population mobility [6]. By taking Hubei Province as an example, they found that isolating the infected on the initial stage of pneumonia, reducing individuals’ across communities move, improving the level of treatment is effective to control the large-scale spread of the epidemic in a short time. Kiesha Prem used the SEIR model of age stage to simulate the track of outbreak in Wuhan. It was pointed out that many interventions taken in Wuhan have curbed the spread of the virus when the COVID-19 outbreak [7]. The epidemic subsides quickly won over critical time for medical treatment and significantly reduced the total number of infections finally. Besides, Wei Chen also agreed that the containment measures reduced the risk of continuous community spread in other regions of China except for Wuhan, delayed the arrival time of epidemic peak, and relieved the pressure of medical resources [8]. But the containment measures caused huge social costs and economic costs, it harmed people’s life and social production in Wuhan and other epidemic areas. All the above studies proved the positive significance of Wuhan lockdown and limitation of population migration to control the epidemic situation of COVID-19. However, they didn’t consider the impact of the large-scale provincial migration during the Spring Festival. This paper proposed a new model based on the susceptible-exposed-infected-removed model (SEIR model) [9] and considering the inter-provincial migration factor, which is named SEIRM model. In this model, people are divided into seven groups according to the infection stage, and they are assumed to move freely between Wuhan and other provinces.

This paper is organized as follows: Section 2 introduces the model hypothesis; Section 3 proposes the model; Section 4 presents the empirical analysis using the proposed model; Section 5, concludes this study.

### Table 1. Nomenclature.

| Symbol | Symbolic Description       | Symbol | Symbolic Description       |
|--------|---------------------------|--------|---------------------------|
| $S$    | Susceptible Population    | $R$    | Recovered Population      |
| $E$    | Exposed Population        | $P$    | Protected Population      |
| $I$    | Infected Population       | $A$    | Absorbed Group            |
2. Problem description and assumptions

2.1. Problem description and data
To predict the situation of the epidemic in China as of March 1, 2020, after the Spring Festival migration, under the assumption that Wuhan was not closed and the population moved freely across provinces, we searched kinds of data. The cumulative number of confirmed, cured, and death cases of COVID-19 were obtained from the official websites of health committees at all levels. The migration data from Wuhan was obtained and calculated from January 10 to January 23, 2020(half a month before Wuhan was shut down) by Baidu's migration platform [10]. By introducing the factors of inter-provincial migration into the SEIR epidemic dynamics model, we formulated a new model, called the SEIRM model, which can simulate the epidemic situation in the non-epidemic source areas.

2.2. Assumptions
According to the problem description, the following assumptions are proposed for the SEIRM model proposed in this study:

1. There is no movement of the population between the other provinces.
2. The final destination of the floating population is fixed, and can reach the final destination, does not involve sudden stagnation
3. The mortality of COVID-19 from January 29, 7 days after the city closure of Wuhan, has real significance.

The assumption (1) is proposed because in the early stage of virus transmission, there was no
widespread community transmission in other areas except Wuhan. To simplify the model, the population flow between other provinces outside Wuhan of Hubei Province is not considered temporarily. The reason behind assumption (2) is that the number of people migrating in the Spring Festival is huge and the purpose is complex. This assumption helps to construct a clearer model for analysis. Assumption (3) comes from that the extensive nucleic acid testing by the Wuhan government began after the city was shut down. Because of the small scale of new coronavirus detection before the closure of the city, the number of infected people and the number of deaths during this period cannot reflect the actual situation very well, so we need to use the data after the closure of the city to calculate the death rate of new coronavirus. According to statistics, a new pneumonia patients treatment time for an average of 6-8 days, if the patient from the first day of onset of hospital treatment, also need at least 6-8 days after discharge. Therefore, we take the statistics after 7 days (that is, January 29) as a reference[11].

3. SEIRM model
According to the COVID-19 epidemic characteristics, this paper constructs the revised SEIR model with inter provincial Migration factors as follows.

In the model, 9 population groups are considered: \( S \) (susceptible population), \( E \) (exposed population), \( I \) (infected population), \( Q \) (quarantined population), \( R \) (recovered population), \( D \) (dead population), \( P \) (protected population), \( N \) (number of population) and \( A \) (absorbed group). The transformation relationship of the population is shown in Figure 1.

![Figure 1. Schematic diagram of modified SEIR infectious disease dynamics model.](image)

According to the above diagram, the following differential equations can be constructed:
Among them: A is the number of Wuhan population moving in of the corresponding population.

In this paper, Runge Kutta method is used to solve the partial differential equation in formula (1), as follows:

First, build the parameters to calculate assemblage \( Pa \), and move in population (absorbed group) assemblage calculate \( A \), as follows:

\[
P_{a} = \begin{bmatrix} -\alpha & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\gamma & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \gamma & -\delta & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \delta & -\lambda - \kappa & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \kappa & 0 & 0 & 0 \\ \alpha & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} A_{S} \\ A_{E} \\ A_{I} \\ A_{Q} \\ A_{R} \\ A_{D} \\ A_{P} \end{bmatrix}
\]  

(2)

Among them, \( A_{D} = 0 \).

Then, build an assemblage \( Y \) of the groups except the total population, as well as the assemblage \( F \), as follows:

\[
Y = \begin{bmatrix} S \\ E \\ I \\ Q \\ R \\ D \\ P \end{bmatrix} \quad F = \begin{bmatrix} -\beta \cdot S \cdot (I + E) / N \\ \beta \cdot S \cdot (I + E) / N \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}
\]  

(3)

Finally, the Runge-Kutta differential solution method [12] is used, and the calculation method is as follows:

\[
k_{1} = \text{Fun}(Y(t), A, F); \\
k_{2} = \text{Fun}(Y(t) + 0.5k_{1} \cdot dt, A, F); \\
k_{3} = \text{Fun}(Y(t) + 0.5k_{2} \cdot dt, A, F); \\
k_{4} = \text{Fun}(Y(t) + 0.5k_{3} \cdot dt, A, F); \\
Y(t + dt) = Y(t) + (1/6) \cdot (k_{1} + k_{2} + k_{3} + k_{4}) dt + dA / dt.
\]  

(4)

And in these, \( \text{Fun}(Y, A, F) = A \cdot Y + F \).

The parameters taken above are as follows:
Table 2. Parameter value of Runge-Kutta differential equation

| Parameter value | Value     | Source                                      |
|-----------------|-----------|---------------------------------------------|
| $\alpha$        | 0.04      | Only a few people will take protective measures without active prevention and control |
| $\beta$         | 2.04/5.01 | Ratio of $R_0$ to mean latency [13]         |
| $\gamma$        | 1/14      | The reciprocal of the longest incubation period |
| $\delta$        | 1/5.1     | Reciprocal of the mean incubation period    |
| $\lambda$       | $F_d(br,Q,N)$ | The number of beds per capita in the area, the infected population under isolation treatment, and the function of the total population in the area |
| $\kappa$        | $F_c(br,Q,N)$ | The number of beds per capita in the area, the infected population under isolation treatment, and the function of the total population in the area |

The specific form of $F_d(br,Q,N)$ is:

$$F_d(br,Q,N) =
\begin{cases}
0.97 / T_{zd} & Q / N \leq br \\
0.97(br \cdot N) + 0.2(Q - br \cdot N) / T_{zd} & Q / N > br
\end{cases}
$$

(5)

In the formula, $T_{zd}$ is the average treatment time and $br$ is the number of beds per capita. We take $T_{zd} = 20$ here according to relevant data. The meaning of the formula is: when the number of beds can meet the needs of the isolated population, the cure rate will be at a higher level, here we take 0.97 according to the relevant reports; However, when the number of beds is less than the needs of the isolated population, the cure rate of the population without treatment will be relatively low. Here we take 0.2 according to the relevant report [14].

Accordingly, the specific form of cure rate function $F_c(br,Q,N)$ can be obtained:

$$F_c(br,Q,N) =
\begin{cases}
(1 - 0.97) / T_{zd} & Q / N \leq br \\
(1 - 0.97(br \cdot N) + 0.2(Q - br \cdot N) / T_{zd} & Q / N > br
\end{cases}
$$

(6)

4. Empirical analysis

By running the program code, we predicted the five provinces with the most serious epidemic situation are Hubei Province (except Wuhan), Guangdong Province, Henan Province, Hunan Province, and Jiangxi Province. The changes of the exposed population, infected population, cured population, and dead population are as follows.

Table 3. Analog data table of epidemic situation in five provinces and China.

| Province                  | S     | E     | I     | Q     | R     | D     | P     |
|---------------------------|-------|-------|-------|-------|-------|-------|-------|
| Henan                     | 1828100 | 268120 | 79884 | 126220 | 56607 | 1791  | 7723600 |
| Hunan                     | 1308000 | 231340 | 69000 | 109160 | 49000 | 1555  | 55450000 |
| Jiangxi                   | 8766000 | 191210 | 57114 | 90506  | 40680 | 1298  | 37334000 |
| Guangdong                 | 21941000| 311920 | 92916 | 146780 | 65824 | 2076  | 92650000 |
| Hubei(except Wuhan)       | 4962400 | 3089800| 1075100 | 2058200 | 511400 | 564630 | 35697000 |
| China                     | 18177000| 6573400| 2165500 | 3860300 | 1088500 | 837180 | 107300000 |
Table 4. Actual data of epidemic situation in five provinces and China.

| Province          | Confirmed | Suspected | Cured | Dead |
|-------------------|-----------|-----------|-------|------|
| Henan             | 1272      | 0         | 1205  | 22   |
| Hunan             | 1018      | 0         | 876   | 4    |
| Jiangxi           | 935       | 0         | 850   | 1    |
| Guangdong         | 935       | 0         | 850   | 1    |
| Hubei(except Wuhan)| 17981    | 0         | 14223 | 608  |
| China             | 80174     | 715       | 44518 | 2915 |

Figure 2. Forecast map of Hubei Province (except Wuhan).

Figure 3. Forecast of Guangdong Province.

Figure 4. Forecast of Henan Province.

Figure 5. Forecast of Hunan Province.
2.165 million COVID-19 patients are expected to be unable to receive medical treatment by March 1, 2020. Since some patients with COVID-19 will suddenly develop to critical condition without warning, and the treatment of critically ill patients is extremely difficult [15], the strategy of "group immunity" will face a huge ethical problem when the medical resources in China cannot meet such a huge demand of patients. So the strategy of "group immunity" does not work in China.

According to the simulation development of this paper, there are some differences between China's actual epidemic development data and the simulated epidemic situation development of the paper. The main reason is that China had taken active measures to prevent the novel coronavirus pneumonia epidemic. The first significant measures are that Wuhan City closure measures were issued in time, from January 23, 2020, to April 8, 2020, the further spread of the epidemic was effectively blocked by closure measures and the closure measure also controlled the migration of high-risk infection center population migration and the spread of infection sources. The second significant measure is the application of big data tracking, which can collect data and intelligently analyze and evaluate data through the epidemic monitoring system so that the government can quickly track the migration routes of infected people and timely prevent and control vulnerable areas. The third significant measures are the systematic allocation of medical resources and living materials, it is an important support for China's effective tracking of the epidemic situation and it strengthens the effectiveness of epidemic prevention measures. Due to the above important measures, China's novel coronavirus pneumonia outbreak had not been uncontrolled and continues to develop, the health of the Chinese people has been effectively guaranteed.

Based on the simulation of the development of the epidemic situation and the actual situation of the epidemic situation in China, this paper believes that taking active epidemic prevention measures will quickly restore social and economic order and protect the health of the people of the country.

5. Conclusion
Based on the modified SEIR model, it can be predicted that the five most serious provinces in mainland China on March 1 are Hubei Province (except Wuhan), Guangdong Province, Henan Province, Hunan Province and Jiangxi Province. In this case, the mass migration of the population during the Spring Festival will quickly lead to a massive outbreak of the disease in China, with more than 6.025 million people across the country becoming infected with COVID-19, and about 837,000 people dying from the disease, far exceed the carrying capacity of Chinese medical resources. It can also be seen that the "herd immunity" strategy is not feasible in China. From the current situation of China's actual epidemic situation, the closure of the city of Wuhan measures effectively intervened the Spring Festival population migration caused by the spread of the epidemic.
Acknowledgement
This work was supported in part by the Project of model based reliability design and optimization method for mechanical mechanism of reactor equipment under grant number S2026025, the Capacity building of science and technology innovation service basic research business cost (Scientific Research) No. 71E2010946 and Special Project for Improving Public Service Capacity in the Field of Industry and Information technology under Grant No. 2019-00909-2-1.

References
[1] Faïçal Ndaïrou, Iván Area, Juan J. Nieto, Delfim F.M. Torres 2020 Mathematical modelling of COVID-19 transmission dynamics with a case study of Wuhan Safety Science
[2] F Zhang, Lu Li and Huiyv Xuan 2011 Survey of transmission models of infectious diseases Systems Engineering-Theory & Practice 09 1736-1744
[3] B L Zhong, W Luo and HM Li 2020 Knowledge, attitudes, and practices towards COVID-19 among Chinese residents during the rapid rise period of the COVID-19 outbreak: a quick online cross-sectional survey INTERNATIONAL JOURNAL OF BIOLOGICAL SCIENCES 16(10) 1745-52
[4] J Li, Y Wang and S Gilmour 2020 Estimation of the epidemic properties of the 2019 novel coronavirus: A mathematical modelling study The Lancet THE LANCET-D-20-01420
[5] H Tian, Y Liu Y Li and et al 2020 The impact of transmission control measures during the first 50 days of the COVID-19 epidemic in China Preprint DOI/10.1101/2020.01.30.20019844
[6] S Peng and J S Ying 2020 Beware of asymptomatic transmission: Study on 2019-nCoV prevention and control measures based on extended SEIR model bioRxiv (Preprint DOI/10.1101/2020.01.28.923169)
[7] Kiesha Prem, Yang Liu, Timothy W Russell, Adam J Kucharsk and et al 2020 The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study The Lancet 5 261–70
[8] W Chen, Q Wang, Y Li and et al 2020 Early containment strategies and core measures for prevention and control of COVID-19 in China Chinese Journal of Preventive Medicine 03 1-6
[9] S Cao, P Feng and P Shi 2020 Study on the epidemic development of corona virus disease-19 (COVID-19) in Hubei province by a modified SEIR model Journal of Zhejiang University (Medical Sciences) 02 178-84
[10] H Li, Z Cheng, Y Huang and et al 2020 Study on the relationship between the 2019 Novel Coronavirus Disease epidemic in China and population migration from Wuhan Chinese Journal of Preventive Medicine 33 007
[11] Z Ding, Y Liu, J Kong and et al 2020 A probability model for estimating the expected number of the newly infected and predicting the trend of the diagnosed Operations Research Transactions 24(01) 1-12
[12] N Shao, M Zhong, J Cheng and et al 2020 Modelling for COVID 19 and the Prediction of the Number of the Infected Based on FUDAN CCDC Mathematical Modelling and Its Applications 9(01) 29-32
[13] Q Li 2008 Numerical analysis (Beijing: Tsinghua University Press)
[14] Neil M Ferguson Daniel Laydon, Gemma Nedjati-Gilani and et al 2020 Imperial-College-COVID19-NPI-modelling Imperial College COVID-19 Response Team 9 1-12
[15] B Tang, X Wang, Q Li and et al 2020 Estimation of the Transmission Risk of 2019-nCov and Its Implication for Public Health Interventions Preprintssrn.3525558