Forecasting the 2019-ncov Epidemic in Wuhan by SEIR and Cellular Automata Model

SiYou Wang¹, Hui Fang¹, ZhiHong Ma² and Xiufen Wang⁴

¹ College of Computer and Information Engineering, Tianjin Agricultural University, Tianjin, China
² College of Basic Science, Tianjin Agricultural University, Tianjin, China

*Corresponding author e-mail: 22787940@qq.com

Abstract. Since December 2019, Wuhan has become an outbreak epicentre of the novel coronavirus pneumonia. This study inferred the epidemic spread in Wuhan, which is of great importance for government immediately deployed mitigation interventions. In view of the fact that the virus is infectious during the incubation period, we added new variables to the SEIR model. We used the data published by Wuhan government from Dec 8, 2019, to Feb 4, 2020 to forecast the spreading trend and the number of infected people, then simulated the results of different government control efforts by changing the control variables. Furthermore, we used mathematical formulas to infer the basic reproduction number, and then calculated the probability of the neighbor being infected at each time step, simulated the epidemic course of epidemic by using the cellular automata model. Our finding suggested that the epidemic outbreak earlier under the government's powerful control measures and the outbreak period begin on mid-January, then reach a peak on the mid-February. The cumulative number of infections in Wuhan would be controlled at 58310. More importantly, the basic reproduction number is 3.13, which indicated that the virus is high risk. Simultaneously, the cellular automata model is very suitable for simulating the spread course after the public health interventions.

Keywords: SEIR, Cellular Automata, Novel Coronavirus Pneumonia

1. Introduction

Wuhan, the capital city of Hubei Province, which is an important transportation hub in China. Since the outbreak of Novel Coronavirus Pneumonia in December 2019, because of the Spring Festival, a period with extremely high traffic across China, the epidemic quickly spread to different countries and regions. On January 23, Wuhan closed the out of the city and established new hospital to treat people infected by the virus. Recently, many scientists have done research on the epidemic. Among them, a paper published by the Lancet represents that as of January 25, 75800 people may be infected in Wuhan [1]. Most of the papers also predicted the basic reproduction number is between 2-4. In this study, we predicted the probable size epidemic in Wuhan and simulated the spread course of this epidemic, which is of great importance for government immediately deployed preparedness plans and mitigation interventions at both the population and individuals in Wuhan.
2. The SEIR Model
Novel Coronavirus Pneumonia has a certain incubation period. Healthy people who have been in contact with patients do not immediately get infected but become carriers of the pathogen. Most researches of infectious diseases used the SIR or SEIR model.

Compared with the SIR model, the SEIR model further considers the factors that people who contacted patients might be contagious, thus the transmission period longer, which is more in line with the characteristics of the epidemic. SEIR consists of susceptible, exposed, infectious, and recovered. S, susceptible, refers to those who have not been ill, but lack immunity, and have an infection easily after contacting with infected persons. E, exposed, refers to those who have been in contact with infected persons but do not have ability to transmit it to other. I, infectious, refers to those who were infected with an infectious disease and transmitted it to S to make them become E or I. R, recovered, refers to those who have been isolated or who became immune after recovered.

Figure 1. SEIR Model.

2.1. Revising Model
Novel coronavirus pneumonia is contagious during incubation period, which can make healthy susceptible persons become exposed persons [2]. Thus, we add $\beta_i$ into the the model as follows.

$$\frac{dS}{dt} = -\beta IS - \beta_1 ES \tag{1}$$

$$\frac{dE}{dt} = \beta IS - \alpha E + \beta_1 ES \tag{2}$$

$$\frac{dI}{dt} = \alpha E - \gamma I \tag{3}$$

$$\frac{dR}{dt} = \gamma I \tag{4}$$

2.2. Basic Parameters and Data Processing

2.2.1. Symbol Definition. In order to analyze the epidemic spread in Wuhan, we need to estimate parameters.

Table 1. Symbol Definition.

| Symbol | Definition |
|--------|------------|
| $T$    | Average incubation period |
| $C$    | Mean duration of infection |
| $N$    | The total population of Wuhan |
| $\gamma$ | Average rate of recovery of infectious person, $\gamma = 1/C$ |
| $k$    | Average number of contactted by infectious person per day |
| $\beta_i$ | Average number of infected made by exposed person per day, $\beta_i \approx \beta$ |
| $\alpha$ | Probability of exposed person conversion to infection person, $\alpha = 1/T$ |
All the epidemiological information was provided by publicly available data sources. We assumed that the average number of infected made by exposed or infectious person were basically same, the value of \( k \) is 5, so we only need to estimate \( b \) based on the corresponding data.

### 2.2.2. Calculation of \( b \).

Through sorting and screening the data published by the Wuhan Municipal Heath Commission. December 8, 2019 as Day 0. The related data were shown in Table 2.

**Table 2. Official Data.**

| Date  | Day | Infectious | Date  | Day | Infectious | Date  | Day | Infectious |
|-------|-----|------------|-------|-----|------------|-------|-----|------------|
| 12.8  | 0   | 1          | 1.21  | 44  | 318        | 1.29  | 52  | 1905       |
| 12.31 | 23  | 27         | 1.22  | 45  | 478        | 1.30  | 53  | 2261       |
| 1.3   | 26  | 44         | 1.23  | 46  | 548        | 1.31  | 54  | 2837       |
| 1.17  | 40  | 62         | 1.24  | 47  | 572        | 2.1   | 55  | 4109       |
| 1.18  | 41  | 121        | 1.25  | 48  | 618        | 2.2   | 56  | 5142       |
| 1.19  | 42  | 198        | 1.26  | 49  | 698        | 2.3   | 57  | 6384       |
| 1.20  | 43  | 258        | 1.27  | 50  | 1590       | 2.4   | 58  | 8351       |

Since the clinical manifestation of the exposed person were not obvious and it is hard to count the population, this study include the number of the exposed person in the susceptible person. Due to the incubation period is 14 days, the data we used was from the day 0 to day 58 (14 days after the government took control measures). This paper assumed that \( N \approx S \) at the initial phase of epidemic spread in Wuhan, we estimated the parameter by the following formula.

\[
\frac{dI}{dt} = \beta \frac{IS}{N} - \gamma I = (\beta - \gamma) I
\]

When \( I(t = 0) = 1 \), the solution of above formula is

\[
I(t) = e^{(\beta - \gamma)t}
\]

The result (with 95% confidence bounds) calculated by data fitting is

\[
b = 0.04472 \quad (0.04432, 0.04511)
\]

### 2.3. Result

In our baseline scenario, we found that if there is no large-scale public health interventions, the simulation results are shown in Figure 2. Day 0 in the figure corresponds to December 8, 2019.

It shows the epidemic curves in Wuhan, the outbreak began on the 55th day (about February 1), reached its peak around 85 days (early March), and came to an end 130 days later (mid-April). The number of people infected will eventually exceed 4.05 million.
Figure 2. The number of four populations in the SEIR model.

To mitigate the spread of the virus, the Wuhan government has progressively implemented quarantine and other public health interventions since Jan 20, 2020. Besides, Wuhan was completely blocked on January 23. The related department of Wuhan restricted public transportation, closed unnecessary entertainment places and cancelled various gathering activities [3].

We simulated the intensity of government control by changing the value of $k$ [4]. If the government intervention is very strong ($k = 1$) (Figure 3), the cumulative number of infections in Wuhan will be controlled at 58310. If the government intervention is slightly strong, ($k = 1.5$) (Figure 4), the cumulative number of infections in Wuhan will exceed 466600. If the government intervention is weak ($k = 2$) (Figure 5), the cumulative number of infections in Wuhan will more than 1269000. If government control is extremely weak ($k = 3$) (Figure 6), the cumulative number of infections in Wuhan will exceed 2589000.

Figure 3. Under strong intervention.  
Figure 4. Under slightly strong intervention.
3. Cellular Automaton Simulation

3.1. $R_0$

$R_0$, that is, the basic reproduction number, is the average number of people infected by a person who have certain infectious disease with no external intervention on the circumstance that all individuals do not have immunity. We estimated the $R_0$ of the novel coronavirus pneumonia is:

$$R_0 = \frac{\beta}{\gamma} = \frac{kb}{1/C} = \frac{5 \times 0.04472}{1/14} = 3.13(3.10 \sim 3.16)$$

3.2. Cellular Automata

CA, that is cellular automata. Compared with the SEIR model, this model can simulate the characteristic of epidemic spread in space. Nowadays, the public transportation in Wuhan have stopped, the communities were also strictly controlled, and people rarely went out. The virus can only spread around the infectious individuals, which is very suitable for the CA model.

On the basic of SEIR, we build the cellular automata model. Cell means that all people in the particular space. Cell Space means a two-dimensional space with $N = 500 \times 500$. Neighbor Form means that the Moore type that neighbor radius is 1. Cell Space State represents that the state of the cell. $X(x = 1, 2, 3, 4)$ respectively represents S, E, I, R. Rule is that if there are cell E in the four neighbors of east, south, west and north of cell S, the cell S will change to cell E with probability P at next time, and cell E transformed into cell I after T time steps. Cell I became cell R after C time steps.

In this modelling study, we randomly set the incubation and healing period of each cell, and respectively obey the normal distribution of $N(T, \frac{T}{2})$ and $N(C, \frac{C}{2})$. We assumed that each exposed cell could infect 3.13 of the 4 neighbors on average during the incubation period. The probability of the neighbor being infected at each time step is

$$P = R_0 \frac{T}{4T}$$

3.3. Simulation Result
In this model, we use the above CA model to simulate the epidemic spread in cell space. A certain percentage of cells are exposed at the initial moment. The simulation results are shown in Figures 7, 8 and 9. The curve of the number of four kind people change over time shown in Figure 10 is very similar to the modified SEIR model, which proves that the cellular automaton can forecast the probable course of epidemic spread in Wuhan.

4. Conclusion
The basic reproduction numbers of NCP and other infectious diseases are compared as follows.

| Table 3. The $R_0$ of different epidemic |
|---|
| Name                                    | $R_0$       |
| HIV                                     | 3.65-4.14   |
| Ebola virus                             | 1.3-2.7     |
| H1N1 Flu                                | 1.4-3.1     |
| Novel Coronavirus Pneumonia             | 3.10-3.16   |

The larger the $R_0$, the more difficult it is to control the epidemic. If $R_0$ is greater than 1, the infectious disease will spread exponentially and become an epidemic. If $R_0$ less than 1, the infectious diseases will gradually disappear. Compared $R_0$ with other infectious diseases [6], it is obviously to find that the risk of NCP is higher and it is not easily to control.

Undoubtedly, the government's immediate and robust intervention and the self-isolation of the
masses have played a huge role in the control of the epidemic by comparing the Figure 2 to Figure 3. Our estimates suggested that the inflection point of this epidemic is greatly advanced. We inferred that the outbreak period begin on the 35th day (mid-January), and reach a peak on the 64th day (mid-February), then the epidemic would gradually fade out. Finally, the number of infectious people in Wuhan would eventually be controlled within 58310. However, if the government adopt further measures, such as the study of new medicine and vaccines, the actual number of infected people might be less than the model predicts.

During the middle period of this epidemic, people of each community do not go out and would isolate themselves at home. So, the spread of the NCP can only made by infectious people in a limited space. The cellular automata model simulates the spread of the epidemic vividly in the limited space.

In this study, we have estimated the outbreak size of 2019-nCoV in Wuhan. However, we have limited knowledge of this virus at the initial period of the epidemic, the number of infectious people during the pre-epidemic period might be significantly more than the confirmed case number. Besides, Wuhan is the epicenter of this outbreak, the limited medical resources and materials in this city cannot guarantee 100% treatment of infected people, the number of infectious people that we estimated might be slightly larger than the official data.

5. Conclusion
The virus have complex characteristics, and due to the lunar new year, Wuhan has extremely huge population movements [7]. The published case numbers were rising drastically during this period. Thus, forecasting is of crucial importance for public health control and planning. The modelling techniques that we used in this paper are more applicable to the characteristics of the NCP than the basic model. Furthermore, the $R_0$ is closer to the predictions of other international scholars, the transmission trend and the number of infectious people are more in line with the actual situation.

Nonetheless, our study has several limitations. First, because the number of exposed people is difficult to count, we included the exposed people in the susceptible people. Actually, there might be a lot of exposed people in Wuhan because the vague understanding of clinical symptoms in the initial period of the epidemic.

Second, before that traffic management took place on January 23, large population movements have occurred in major railway stations and airports in Wuhan, however, we have neglected the impact of the per-spring Festival on the expansion of the epidemic.

Besides, the data we used was published by the official department. Due to the complex system of classification of patients in the early age of the epidemic, many infectious people with mild symptoms have not been treated. The official data cannot precisely reflect the infectious people per day in the initial epidemic period.

Acknowledgments
This work was financially supported by Tianjin Agricultural University student innovation and entrepreneurship training project (201910261056).

References
[1] Joseph T Wu, Kathy Leung, Gabriel M Leung. "Nowcasting and forecasting the potential domestic and international spread of the 2019- nCoV outbreak originating in Wuhan, China: a modelling study", The Lancet, 2020
[2] Experts: Incubation period of new coronavirus may be contagious on http://www.ce.cn/xwzx/gnsz/gdxw/202001/26/t20200126_34189665.shtml
[3] New Coronavirus Infected Pneumonia Prevention and Control Headquarters Circular (No. 1) on http://www.xinhuanet.com/politics/202001/23/c_1125495557.htm
[4] Xiaohua Yu. Estimating infected population of wuhan coronavirus in different policy scenarios by sir model, 2020. on http://uni-goettingen.de/en/infectious+diseases/619691.html
[5] (2019-ncov) data on https://github.com/globalcitizen/2019-wuhan-coronavirus-data
[6] Basic reproduction number on https://en.wikipedia.org/wiki/Basic_reproduction_number.

[7] 5 million people left Wuhan on http://news.china.com.cn/2020-01/26/content_75650784.htm