For office use only10

T1 ______________

T2 ______________

T3 ______________

T4 ______________

Team Control Number

2020012

Problem Chosen

C

For office use only

F1 ______________

F2 ______________

F3 ______________

F4 ______________

2020

MCM/ICM

Summary Sheet

2019-nCoV infectious pneumonia SEIR model

Summary

China has once again encountered major difficulties after SARS. In order to better understand the spread and development of new coronary perceptive pneumonia, scholars' research is urgent. For Wuhan infectious diseases, we established the SEIR model.

For question one, we collected the outbreak data from January 10 to February 6 from Wuhan Health and Health Commission, and set the population on January 10 as the initial value. Because the number of infected people is a collective minority, it is a continuous optimization problem. Since the epidemic in Wuhan is known to be contagious, we established a new model of coronary infectious pneumonia SEIR. We set up equations based on the number of people in each category every day. After testing the true and predicted values, the model with a small error is better. The data we collected was small. We used grey prediction to find the four parameters on the 23rd, and brought the parameters back to the equation. We found that the number of infected people on the 23rd was **518**.

For predicting the first infection time of the pneumonia, we reversed the model based on the previous question and changed from backward prediction to forward prediction. We specified the corresponding parameters and predicted that the first infectious disease before the city's closure on January 23 would be **December 8**.

For question two, considering the changes before and after the closure of the city, we re-selected the contact rate and cure rate parameters of the first question model to reduce the daily contact rate and increase the cure rate. Based on the revised parameters (Figure 3), it can be seen that the inflection point will appear on February 25, and the number of patients will be about 34,270. The virus transmission will be basically eliminated on May 2.

For question three, in order to study the number and distribution of infected people in different counties of Huanggang City, we separately set up a set of differential equations for each county. Due to the movement of people during the Spring Festival, we added the term of movement of people, simulated the parameters in the equation from the data, predicted the numbers of the five types of population on February 6-9, and plotted the images. The model fits well, so the predictions for the five races from 6th to 9th are relatively accurate. We used MATLAB to find the 17-day epidemic progress of infected people, and the predicted value of the number was compared with the real value. Finally, the distribution of infected people in Huanggang City after the closure of Wuhan was obtained.

For question four, considering the differences in environmental policies between Tianjin and Wuhan, we changed the parameters again. First, we predicted the development image without opening school, and we will reach the inflection point on February 20. In the case of the beginning of the school on February 14, we are forecasting an increase in exposure. It can be seen that the inflection point will be postponed and the epidemic time will be prolonged, so we do not recommend starting school on February 14. We made predictions for the beginning of the school on different dates and obtained March 11 as a suitable start time. We tested the parameters with an accuracy of 88.6% and a certain degree of confidence.

Finally, we analyzed the advantages and disadvantages of the SEIR epidemic model and extended the model.

Kewords: 2019-nCoV Infectious Pneumonia, SEIR Model, Differential Equations, Wuhan

Content

1. Introduction

1.1. Background of the Problem

In China, when the Chinese New Year is approaching, an infectious disease has broken out. The SARS incident in 2003 not only took away many lives from the beginning to the end, but also caused losses to countries around the world and caused disorder in social order. New type of coronary pneumonia virus is coming. The death toll of new type of coronary infectious pneumonia in 2019 surpasses SARS, which has caught Chinese off guard.

On December 31, 2019, a patient was found in Wuhan, more patients were subsequently exposed to the South China Seafood Market in Hankou. Wuhan is located on the plain, and the transportation network extends in all directions, causing the virus to spread when people do not pay attention to the condition. With the sudden increase in the number of people seeking treatment, the structural shortage of related hospital resources, a large number of patients with pneumonia could not be admitted, resulting in a greater degree of infection. With the rapid construction of Wuhan Vulcan Mountain Hospital and Lei Shenshan Hospital and the renovation of other sites and facilities, Wuhan 2019-nCoV patients with pneumonia are gradually receiving hospital treatment. It is believed that with the determination of the government and the hard work of medical workers and scientists, people will certainly survive the difficult times.

1.2. Previous Works

There have been many large-scale infectious disease problems. Although the pathology of infectious diseases are different, it is not difficult to find common characteristics with the accumulation of pathological knowledge and knowledge. By establishing a mathematical model to study the spread, it is better for people to take measures to defeat the disease. The development of infectious disease transmission models has gone through the following processes:

Earlier, Kermack and McKendrick proposed the famous SIR warehouse model on the problem of infectious diseases. The SIS model was established in 1932, and a "threshold theory" for judging whether an infectious disease has returned, that is, an outbreak of an infectious disease can only occur if the number of susceptible persons in a group exceeds a threshold [2]. In 1978, Capasso and Serio introduced a saturated infection rate in the epidemic model to study the cholera epidemic in Bari in eastern Italy in 1973 [4].

With the extensive study of infectious disease models by experts and scholars from various countries, the emergence and development of infectious disease dynamics have been promoted. Among them, the dynamic model can be divided into infectious disease models with time delay, age structure, multi-group transmission, non-autonomy, impulsive and migration; Lin Guoji, etc. successfully fitted the 2003 SARS epidemic in Beijing with the World Network Model; Gong Jianhua and other kinetic models have studied the impact of the 2003 SARS epidemic in Beijing on the time period and implemented control, which has a certain impact on the control of the spread and spread of the epidemic; Marc Lipstch studied SARS propagation dynamics and control, using data from Singapore and other places. Local data. They estimated the number of basic regeneration in the case of re-control measures and future measures, highlighting the importance of control measures [2]; in 2019, Tang San led three college mathematics teams and established a propagation dynamics model. They predicted the risk of spreading the new coronavirus in Wuhan, and helped solve social problems in China in an epidemic situation.

1.3. Our Work

To present our conclusions, we went into work:

In question 1: We collected relevant data from Wuhan and made reasonable assumptions about infectious diseases. Based on the inter-population relationship, we established a 2019-

nCoV infectious pneumonia transmission model, the SEIR model. Unknown parameters in the equation were predicted from known data, and with the help of MATLAB, the number of infections before the city was closed on January 23 and the approximate time when the first infectious disease occurred were predicted.

In question 2: We collected patient data from Wuhan to Vulcan Mountain Hospital and calculated the number of infected patients who stayed in Wuhan after 23 days. With the study of medical conditions by physicians and scholars, the peak of new cases in Wuhan and the time to eliminate the spread of infectious diseases have been predicted.

In Question 3: We improved the model in Question 1, added the flow of people term, and set a set of differential equations for each county. Under the condition of human intervention, we adjusted the parameters to predict the number of pneumonia patients arriving in Huanggang, and plotted the distribution of Huanggang City.

In question 4: We compared the beginning of school on February 14 with the extension of school and predicted the development trend of the future condition. We also analyzed the rationality of the university's opening on February 14 from both economic and health perspectives. We predicted what we thought was the best start time.

2. Symbol Descriptions

Symbol	Definition
$S(t)$	Number of vulnerable people
$I(t)$	Confirm the number of patients
$E(t)$	Number of Lurkers
$R_1(t)$	Number of Healers
$R_2(t)$	Number of dead
α_1	Contact rate of vulnerable populations
α_2	Infection rate
m	Number of patients per patient per day
φ_1	Probability of curing an infected person
φ_2	Probability of death of infected person

Note: See the description of other symbols below.

3. Model Hypothesis

To simplify the problem, we make the following basic assumptions, each of which is justified.

● **The statistics of the new type of coronary infectious pneumonia provided by the Ministry of Health are true and reliable.**

In China, where modern media is fast-developed, major platforms have real-time reports of relevant disease data. Through national force censuses, national statistics are authentic and reliable.

● **Time is measured in days. The minority population has changed during the epidemic, and the change is small compared to the overall population.**

Time is discrete in days, but the infected population is small compared to the total population, which can be approximated as a continuity problem and a simplified problem.

● **Disregard the birth rate and natural mortality rate during this period.**

Calculate the number of people infected during this period. In order to simplify the model and ignore the effects of natural factors on the population, the birth rate and natural mortality of the population are not considered.

● **After recovering or dying, an infectious disease patient is non-infectious and has immunity to infection.**

According to current medical research and reports, there is no case of re-infection after the infected person is cured and there is no infectivity, which is in line with the facts.

● **People will go to the hospital for treatment as soon as they find out the disease. The infected people only include those who have been treated.**

When people find that they are unwell, they choose to go to the hospital the first time. This assumption is in line with people's actual thoughts and simplifies the model.

4. SEIR Infectious Disease Model

4.1. Analysis of the Problem

From the official website of Wuhan Health and the real-time big data report of the epidemic, we can collect the number of infections, suspected cases, cures, and deaths in Wuhan every day from January 10 to February 6. Infectious diseases are a continuous optimization problem. The new type of coronary infectious pneumonia has a latency period, so we have established a SEIR model. We list differential equations based on the relationship between population changes. According to the collected data, parameters were fitted, and the number of infections on the 10th was specified as the initial value. The initial value and parameters are used to predict the population number of 11-22, and the error analysis is performed to judge the reliability of the model. Because some information is unknown and the sample is small, the gray prediction model is used to predict the four parameters on the 23rd. Bring the parameters back to the equation and find the number of people infected on the 23rd.

For predicting the first infection time of the pneumonia, we reversed the model based on the previous question and changed from backward prediction to forward prediction. We specified the corresponding parameters and predicted the number of infections and the approximate time of the first infectious disease before the city was closed on January 23.

4.2. Data Collection

We can collect the number of infections, suspected cases, cures, and deaths from January 10 to February 6 in Wuhan from the official website of Wuhan Health and the real-time big data report of the epidemic.

Table 1: Five categories of people from January 10 to 23

Date	Normal person	Infected	Cured	Suspected
2020.01.23	495	31	23	3653
2020.01.22	425	28	17	2556
2020.01.21	363	28	9	1070
2020.01.20	258	25	6	988
2020.01.19	198	25	4	817
2020.01.18	121	24	3	784
2020.01.17	62	19	2	763
2020.01.16	45	15	2	763
2020.01.15	41	12	2	763
2020.01.14	41	7	1	763
2020.01.13	41	7	1	763
2020.01.12	41	7	1	763
2020.01.11	41	6	1	763
2020.01.10	41	2	1	739

4.3. Introduction to New Coronaviruses

The new Coronavirus, abbreviated as NCP in English, was discovered in 2019 due to Wuhan Viral Pneumonia cases and was named by the World Health Organization on January

12, 2020. It is known to cause more severe diseases such as colds and severe acute respiratory syndrome (SARS). The new coronavirus is a new coronavirus strain previously found in humans.

The common signs of people infected with coronavirus are respiratory symptoms, fever, cough, and dyspnea. There is no contagion. In severe cases, the infection causes pneumonia, kidney failure, and even death. At present, there is no special treatment for this disease. Therefore, people should take protective measures to avoid close personal contact as much as possible.

4.4. SIR Model

The SIR model is a simple infectious disease model. Kermack uses a dynamic method to establish a mathematical model of infectious diseases and proposes a threshold theory. That is, when the number of susceptible people in the population is higher than the threshold, the infectious disease will maintain [1]; below the threshold, the disease will tend to become extinct. Let $I(t), S(t), R(t)$ denote the number of infectious, susceptible and withdrawn persons (including the cured and the dead) at time t, respectively.

In practice, the number of infected people is a discrete variable and does not have continuous differentiability, but because a small number of people change in a short period of time, this change is small compared to the overall population, so a differential equation model can be established. Assume that the total population remains at a fixed level and that the total population is N.

$$\frac{dS}{dt} = -rSI \tag{4-1}$$

$$\frac{dI}{dt} = rSI - \delta I \tag{4-2}$$

$$\frac{dR}{dt} = \delta I \tag{4-3}$$

Among them, r means that a susceptible person is infected by any infectious person with a uniform probability r, δ means that an infectious person withdraws from the affected population with a constant proportion δ, and r, δ is a positive number.

The following conclusions can be drawn from the model:

The rate of change of excluded persons is approximately equal to

$$\frac{dR}{dt} = A\left(\frac{2}{e^{x} + e^{-x}}\right)^2 (\beta t - \varphi) \tag{4-4}$$

Among them, X is a normal number, which is a certain function of RE. On the basis of considering three individuals in the SIR model, a category of latent E was added, and the relationship between them constituted the SEIR model.

4.5. SEIR Model

4.5.1. Crowd Classification

According to people's illness and travel situation, we group the population into the following categories:

● Vulnerable persons: Use $S(t)$ to indicate the number of vulnerable persons. According to the travel situation, it is divided into the number of vulnerable people who are not out and out, and are represented by S_1, S_2 respectively.

● Lurkers: Suspected people who have been exposed to infection. The incubation period is 2-14 days and is contagious. Let $E(t)$ be the number of lurkers. According to the travel situation, the number of lurkers who are not out and out are expressed by E_1, E_2 respectively.

● Confirmed patients (infected): $I(t)$ indicates the number of confirmed patients. According to the travel situation, it is divided into the number of confirmed patients who did not go out and went out (including going to the hospital and not going to the hospital). Among them, those who did not go out were represented by I_1, those who did not go to the hospital were represented by I_2, and those who were infected by the doctor were represented by I_3.

● Exiters (the healers and the dead): Use $R(t)$ to indicate the number of withdrawals, and the specific R_1, R_2 respectively represent the number of cured and dead.

Let's visualize the relationship between people more clearly and intuitively, as shown in the figure below:

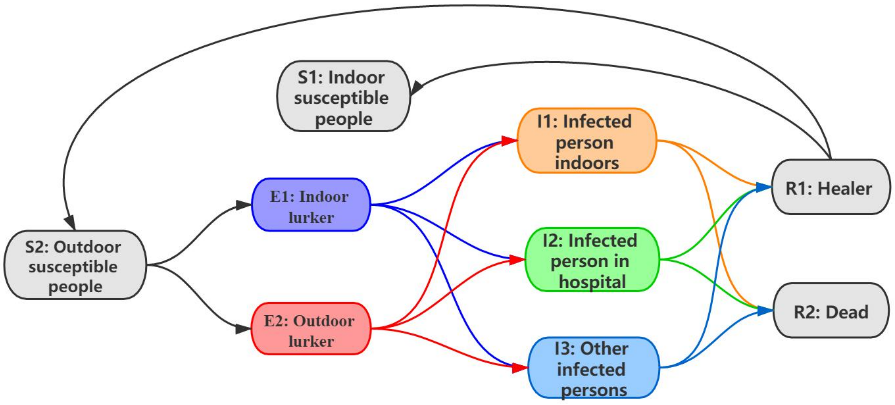


Figure 1: Relationship between infections among populations

4.5.2. Explanation of Symbols

α_1: Contact rate of vulnerable groups;

α_2: Infection rate of vulnerable groups;

β_1: Probability of the lurker indoors;

β_2: The probability that the lurker is outdoors.

4.5.3. Model Establishment

It is assumed that after recovering or dying, an infectious disease patient is non-infectious, and has immunity to infection.

On that day, the number of outdoor susceptible persons was equal to the number of confirmed patients in contact with outdoor susceptible persons. The equation is as follows:

$$\frac{dS_2}{dt} = -\alpha_1 S_2 (E_2 + I_2) \tag{4-5}$$

The number of indoor lurkers is equal to the number of outdoor susceptible patients in contact with confirmed patients and indoors, minus the number of indoor latent confirmed infections, that is:

$$\frac{dE_1}{dt} = -\alpha_2 E_1 + \beta_1 \alpha_1 S_2 (E_2 + I_2) \tag{4-6}$$

The number of outdoor lurkers is equal to the number of outdoor susceptible patients in contact with confirmed patients and being outdoors, minus the number of outdoor latent confirmed infections.

$$\frac{dE_2}{dt} = -\alpha_2 E_2 + \beta_2 \alpha_1 S_2 (E_2 + I_2) \tag{4-7}$$

The number of patients in the room is confirmed to be equal to the number of infections of the latent person in the room and minus the number of infected persons who have been cured or died.

$$\frac{dI_1}{dt} = r_{11} E_1 + r_{12} E_2 - \varphi_{11} I_1 - \varphi_{12} I_1 \tag{4-8}$$

The number of uninfected outdoor infections is equal to the number of latent infections outdoors minus cured or dead outdoor infections. The expression is:

$$\frac{dI_2}{dt} = r_{21}E_1 + r_{22}E_2 - \varphi_{21}I_2 - \varphi_{22}I_2 \tag{4-9}$$

The number of infected persons who seek medical treatment is equal to the number of infected persons identified as infected by the latent person minus the infected persons who have cured or died of medical treatment. The specific expression is as follows:

$$\frac{dI_3}{dt} = r_{31}E_1 + r_{32}E_2 - \varphi_{31}I_3 - \varphi_{32}I_3 \tag{4-10}$$

Among them, $r_{i1}, i = 1,2,3$ represents the probability that the indoor latent person becomes the three types of infected persons; $r_{i2}, i = 1,2,3$ the probability that the outdoor latent person becomes the three types of infected persons. Similarly, φ represents the probability of cure or death of the three infected persons.

The number of cured persons per day is equal to the cumulative number of confirmed patients of each type multiplied by the probability of cure, that is:

$$\frac{dR_1}{dt} = \varphi_{11}I_1 + \varphi_{21}I_2 + \varphi_{31}I_3 \tag{4-11}$$

The number of deaths per day is equal to the cumulative number of confirmed patients multiplied by the probability of death for each category, that is:

$$\frac{dR_2}{dt} = \varphi_{12}I_1 + \varphi_{22}I_2 + \varphi_{32}I_3 \tag{4-12}$$

Among them, the lurker is inside the house and the sum of the probability outside the house is one, which is an inevitable event. So there is:

$$\beta_1 + \beta_2 = 1 \tag{4-13}$$

Because there are too many unknown variables, the solution of the equation cannot be obtained, so we simplified the model and made the following assumptions:

Suppose people immediately go to the hospital for treatment, so only I_3 remains in the infected person, and zero is the number of infected people who have not been treated.

Assume that no infection occurs indoors, that is, the probability that the lurker is inside the house is zero, and the lurker exists only outdoors, so $\beta_2 = 1$.

From this, we simplified the Wuhan 2019-nCoV infectious pneumonia transmission model as:

$$\frac{dS_2}{dt} = -\alpha_1 m E_2 \tag{4-14}$$

$$\frac{dE_2}{dt} = -\alpha_2 E_2 + \alpha_1 S_2 E_2 \tag{4-15}$$

$$\frac{dI_3}{dt} = \alpha_2 E_2 - \varphi_1 I_3 - \varphi_2 I_3 \tag{4-16}$$

$$\frac{dR_1}{dt} = \varphi_1 I_3 \tag{4-17}$$

$$\frac{dR_2}{dt} = \varphi_2 I_3 \tag{4-18}$$

m : The number of patients each patient is exposed to per day;

φ_1 : The probability that they will cure an infected person;

φ_2 : The probability of death of an infected person.

To fit the parameters that best fit the known data, that is, the smaller the residual sum between the predicted value and the true value, the better, as follows:

$$\min \sum_{t=1}^{13} a_1(S_2 - S_2')^2 + a_2(E_1 - E_1')^2 + a_3(E_2 - E_2')^2 + a_4(I_3 - I_3')^2 + a_5(R_1 - R_1')^2 + a_6(R_2 - R_2')^2$$

4.5.4. Important Parameter Estimation

(1) Definition of error

The absolute error is the algebraic difference between the predicted population and the real population. The formula is as follows:

$$\text{Absolute Error} = |\text{Measured value} - \text{True value}|$$

Relative error refers to the ratio of absolute error to the actual population. The formula is as follows:

$$\text{Relative Error} = \frac{|\text{Measured value} - \text{True value}|}{\text{True value}}$$

(2) Parameter estimation

Based on the data from January 10 to January 22, we fit the $m\alpha_1, \alpha_2, \varphi_1, \varphi_2$ parameter. Since the data we collected began on February 10th, the number on the 10th was set as the initial value.

The initial values were: infected persons $I_0 = 41$, cured persons $R_{1,0} = 2$, deaths $R_{2,0} = 2$, and close contacts $E_0 = 739$. We substituted initial values and four parameters into the equation, and predicted the numbers of susceptible population, latent population, infected population, cured population, and death population in the five categories from February 11-22. The distribution of the population is shown in Appendix 1.

An error analysis was performed based on the predicted and true values in Appendix I, and the absolute and relative errors of the five groups of people were obtained. The results are shown in the following table:

Table 2: Parameter error value analysis

	Mean absolute error	Mean relative error
Susceptible people	14.4932	1.0219e-06
Latency	14.4959	0.0054
Infected	3.1291	0.0188
Number of cures	0.1094	0.0046
death toll	0.1795	0.0138

As can be seen from the above table, the relative error does not exceed 0.05, which indicates that the fitting effect is better. The model parameters are better. The corresponding parameter values can be obtained to predict the number of infections before the 23rd.

Because some of the information is unknown and the sample is small, if you want to use the generation and development of known information to understand, the gray prediction meets the conditions. Therefore, we use gray prediction to predict the four parameters, and get the 23rd parameter:

φ_1=0.0427, φ_2=0.009, α_2=0.0381, $m\alpha_1$=0.54046033.

4.5.5. Forecast Results and Analysis

Bringing the 23 parameter values into the simplified equation, we obtained the five categories of susceptible population, latent population, infected population, cured population, and death population of population 23, and performed residual and relative errors on the test results. Inspection, the situation is as follows:

Table 3: Predicted values of various types of people

	People	Absolute error	Relative error
Susceptible people	14182107	1.3442e-05	190.6325
Latency	3823	0.0466	170.4098
Infected	**518**	0.0471	23.3319
Number of cures	29	0.0362	1.1224
death toll	21	0.0864	1.9867

From the table above, we predict that the number of infections in Wuhan City before the seal on January 23 was **518**. The error is small, which is more in line with reality. It has certain reference value for the study of new coronary pneumonia.

This question asked us to predict the time of first infection of the pneumonia. We reversed the model on the basis of the previous question and changed from backward prediction to forward prediction. Therefore, the selection of parameters is very important. Considering that the flow of people before the onset of disease is the same as that of daily life, and there are very few patients with the disease, and the transmission efficiency is not high, so $m\alpha_1$ is 0.01 and α_2 is 0.002; and because it is at the beginning of the disease, no symptomatic drug φ_1 is taken as 0.001; At the beginning of the onset, the pneumonia had not reached the lethal level, so φ_2 was also taken as 0.001. Bring this parameter into the model and calculate the changes in the number of people in each category.

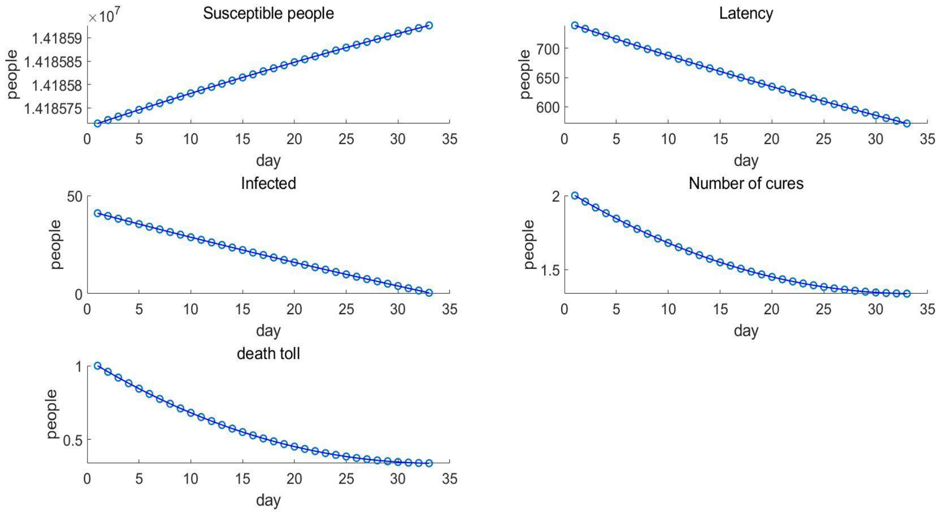


Figure 2: Five groups of people infected before the 23rd

As can be seen from the figure above, the number of populations in the five categories changed significantly. At T = 33 days, the number of patients decreased to one, that is, the approximate time when we predicted the first infectious disease in Wuhan was **December 8**. According to news reports, the earliest infectious disease was found on December 8. Therefore, the prediction is more accurate and has certain reference value.

5. Prediction of Infection in Wuhan

5.1. Analysis of the Problem

We used the patient data from Wuhan to Vulcan Mountain Hospital to calculate the number of infected patients who stayed in Wuhan after the closure. We predict the peak time of Wuhan pneumonia and still use the SEIR model. Considering that the city of Wuhan will be closed on January 23, we used the disease data on the 23rd as the initial data. The government controls population movements and strengthens citizens' awareness of protection, which effectively reduces the chance of virus infection. Therefore, the contact rate between susceptible people and infectious people in this model has gradually decreased. Over time, the number of hospitals and beds began to increase, so the cure rate for patients gradually increased. Because the lethality of the virus itself does not change, we do not change the mortality and the probability of the latent person becoming a patient.

5.2. Model Establishment

Vulcan Mountain Hospital: As the number of patients diagnosed and suspected is increasing every day on January 23, Wuhan decided to follow the model of Beijing Xiaotangshan Hospital to establish Wuhan Caidian Vulcan Mountain Hospital. Its construction area is 25,000 square meters and can accommodate 1,000 beds . Beginning on February 4th, it has officially become the main battlefield against new coronaviruses.

5.2.1. Modification of Parameters

We are required to predict when the pneumonia will reach its peak and when it will be substantially eliminated, so we used the SEIR epidemic model. And changed based on the model of the first question. Considering that Wuhan city was closed on January 23, we used the disease data on the 23rd as the initial data, that is, the 9 million population and the condition of the day as the initial conditions, and added external environmental changes to the model The impact of factors.

The government controls population movements and strengthens citizens' awareness of protection, which can effectively reduce the chance of virus infection. Therefore, the contact rate α_1 of susceptible people and infectious people in this model is gradually reduced. Over time, the number of hospitals and beds began to increase, so the cure rate φ_1 for patients gradually increased. Because the lethality of the virus itself does not change, we do not change the mortality φ_2 and the probability α_2 that the latent person becomes a patient.

5.2.2. Results Analysis

We substituted these four parameters into the model and obtained the following five categories of population changes after January 23:

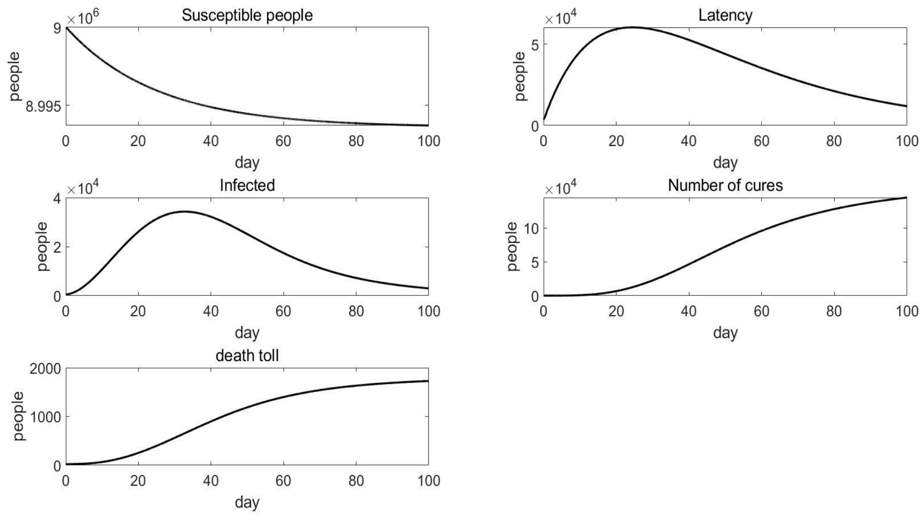


Figure 3: Population changes over time

As can be seen from the above figure, the number of infected persons reached a peak on the 32nd day, and decreased to substantially zero on the 100th day. We conclude that the number of confirmed diagnoses in Wuhan is expected to reach its peak **on February 25,** with 34,270 diagnosed cases. The spread of the virus will be largely eliminated **on May 2**.

6. Predict the Number of Patients in Huanggang

6.1. Analysis of the Problem

This article requires the number of patients who arrived in Huanggang before Wuhan was closed, but according to known data, the number of infections is basically zero, and there are a large number of latent people who should not be patients. Therefore, we have studied the number of infected people and the distribution of each province after the closure of Wuhan. In order to study the number and distribution of five types of people in different counties of the city, if the cellular automaton is used, the calculation amount is very large if accurate prediction is required. Therefore, we set up a separate set of differential equations for each county.

Due to the flow of people during the Spring Festival, we added a flow of people item. The parameters in the equation are simulated according to the number of people in the five categories from January 23 to February 5. We bring the corresponding parameters into the modified equation, use MATLAB to find out the 17-day epidemic situation of infected people, and explain the number of patients arriving in Huanggang provinces.

6.2. Data Sources

Huanggang City includes 10 counties and cities, namely Hong'an County, Macheng City, Luotian County, Yingshan County, Tuanfeng County, Huangzhou District, Qishui County, Hunchun County, Wuxue City, and Huangmei County. Due to incomplete data collection, we found the number of infected, cured and dead people in the five categories of people from January 23 to February 5. We randomly generate 4-7 as the number of latent patients.

Huanggang City is located in the easternmost part of Hubei. The locations of the provinces of Huanggang City are as follows:

Figure 4: Topographic distribution of Huanggang City

6.3. SEIR Model Improvement

First of all, in order to study the number and distribution of five types of people in different counties in the city, we followed the first differential equation and set up a set of differential equations for each county separately. However, a large number of people flocked to Huanggang City before Wuhan was closed. With the approaching of the Spring Festival, the flow of people between counties in Huanggang City gradually increased, and the factors that caused the entry of viruses due to the increase in foreign populations in each county cannot be ignored. Therefore, the flow equation R of the differential equation S_2 and E equation was added as:

$$R_1(i,t+1) = R_1(i,t) + \varphi_1 I_3(i,t) \tag{6-1}$$

R_1 : Is the number of inflows and outflows of susceptible people.

$$R_2(i,t+1) = R_2(i,t) + \varphi_2 I_3(i,t) \tag{6-2}$$

R_2 : Is the number of inflows and outflows lurking.

The parameters in the equation are simulated according to the number of people in the five categories from January 23 to February 5. As for the hospital reception capacity, with the continuous development of the epidemic, the hospital reception capacity has gradually increased. However, due to the limited number of hospital beds, the hospital wards for the diagnosis of patients have become urgent, resulting in the inability of many confirmed patients to receive effective treatment in the first place. To a certain extent, the cure rate has been reduced, and mortality and transmission rates have increased. So we adjusted the parameters for a period of time in the future. And predict the number of populations of the five categories from February 6-9 and draw the image, the model fits better, so the prediction of the five categories of races from 6 to 9 is relatively accurate.

6.4. Results Analysis

According to the adjusted model, MATLAB was used to obtain the 17-day epidemic progress of the infected people. The predicted value of the number of people is compared with the real value as shown below:

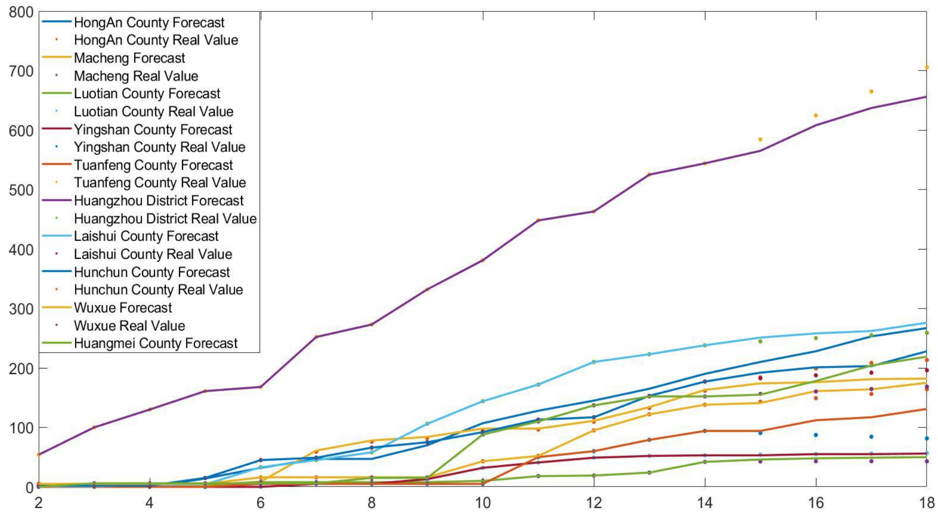


Figure 5: Epidemic progress of infected people in each county within 17 days

As can be seen from the figure above, the daily line segment represents the predicted value, and the point represents the true value. Observation shows that the predicted value of each county is close to the true value, and the fit is good. Among them, Huangzhou District has the largest number of infections, and the others are more evenly distributed.

Table 4: Predicted number of people infected in Huanggang

Area	February 8 Infected	February 9 Infected
Hong'an County	197	203
Macheng City	225	244
Luotian County	43	42
Yingshan County	56	57
Tuanfeng County	84	81
Huangzhou Distric	667	709
Qishui County	257	263
Hunchun County	192	197
Wuxue City	157	166
Huangmei County	165	169

As can be seen from the above table, the epidemic situation in Huangzhou District is the most serious.

Based on the forecast data from January 24 to February 9, in order to understand the number of patients in each county more clearly and intuitively, a map of the epidemic situation of each district and county in Huanggang City was drawn. After the closure of Wuhan, the situation of arriving in Huanggang City is shown in the figure:

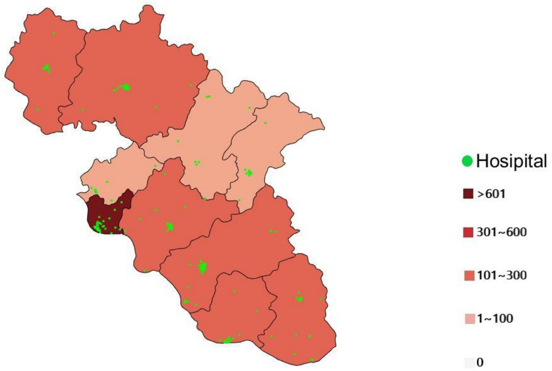


Figure 6: Epidemic map of Huanggang City

As can be seen from the above figure, on February 9th, the number of people who reached Huangzhou District was the highest, and the epidemic situation was the most severe. More than 600 people were infected; the epidemic situation in the middle area was relatively light, and the number of people infected was within 100; Hong'an County and Laishui County. In the west of Huanggang City, it is adjacent to Wuhan. As the epidemic spread center, Wuhan has the largest number of infections in the Huangzhou District, and the other counties and districts bordering it are also more infected. Therefore, our predictions are in line with reality.

7. Best start time of Tianjin University

7.1. Analysis of the Problem

This question allows us to analyze the appropriate start time, and we predict the development of the number of illnesses in the start and delayed start of the school, and make a comparative analysis. First, we predict the development of the epidemic situation in Tianjin based on known data, and modify the parameters of the model in question two. According to the policy strength of different cities and the implementation time is different, we have changed the parameters of contact rate and cure rate to obtain the trend of epidemic situation in two cases. From two perspectives, we will explain whether the start of the school on January 14 is appropriate, on the one hand, the students' physical health, and on the other hand, the economic benefits. Finally, we simulated what we thought was the best start time.

7.2. Delayed Infection Model

7.2.1. Changing Parameters

This question allowed us to analyze the appropriate start time. We decided to take the health of students as the primary condition and need to predict the development of the epidemic situation in Tianjin based on known data. We know that there are 2400 people from Wuhan to Tianjin. Currently Tianjin is a virus-importing city, but it cannot be ruled out that the virus will not spread in the future. Therefore, these 2400 people are likely to be hosts of the virus. Therefore, consider deleting the 400 people the next day, that is, the 400 people spread in Tianjin for only 1 day. Use these data to make appropriate changes to the model in question two.

According to the policy strength of different cities and the implementation time is different, we have made certain changes to the parameters of contact rate and cure rate. The cure rate $\varphi_1 = 0.002t$ and contact rate of infected people $\alpha_1 = \frac{0.06}{t}$, t represent days.

7.2.2. Analysis of Results

Starting from January 21, substituting the parameters into the equation, we get the following forecast curve.

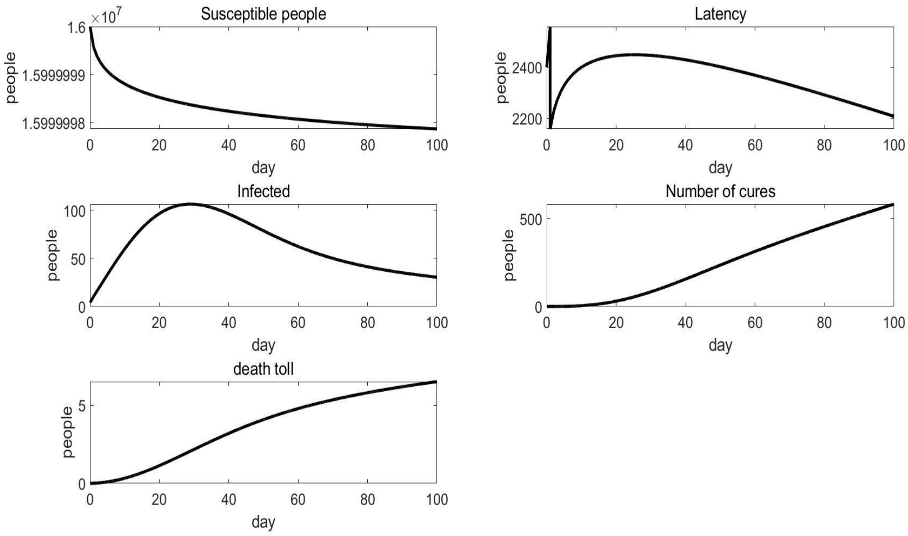


Figure 7: Changes in the number of five types of infections in Tianjin

From the predicted data, we know that the number of confirmed diagnoses on February 6 was 88. Compared with the actual situation of 78, the accuracy was 88.6%, and the model can be considered to have certain credibility. The infected person reached its peak on the 30th day, that is, February 20. After that day, the number of daily diagnoses will be less than the number of daily cures, and the epidemic will be relieved. However, February 14th is still in the development stage of the virus. At this time, the potential risk to people's health due to the start of the school is too large. Therefore, we do not recommend that the school start on February 14.

7.3. Infection Model for Normal School

7.3.1. Changing Parameters

Assume that the students start on the 14th. We will take the 14th data previously predicted for Tianjin as the initial data. The initial values on February 14 were 2449 suspected patients, 105 infected, 53 cured, and 2 dead. According to the Tianjin Statistical Yearbook, we have about 1.7 million enrolled students in all schools in Tianjin. We can see that this is not large among Tianjin's total population of 16 million. However, considering that the opening of students will inevitably lead to the opening of transportation, catering and other industries, it can be expected that the beginning of students will have a significant negative impact on the suppression of the epidemic. So we changed this epidemic model again.

It is assumed that the probability of exposure to transmission is increased at the start of the school. Therefore, the cure rate $\varphi_1 = 0.002t$ was changed and the exposure rate $\alpha_1 = \frac{0.6}{t}$ was changed, t represent days.

7.3.2. Results Analysis

The modified parameters are brought into the equation for prediction, and the changes in the number of infected persons are as follows:

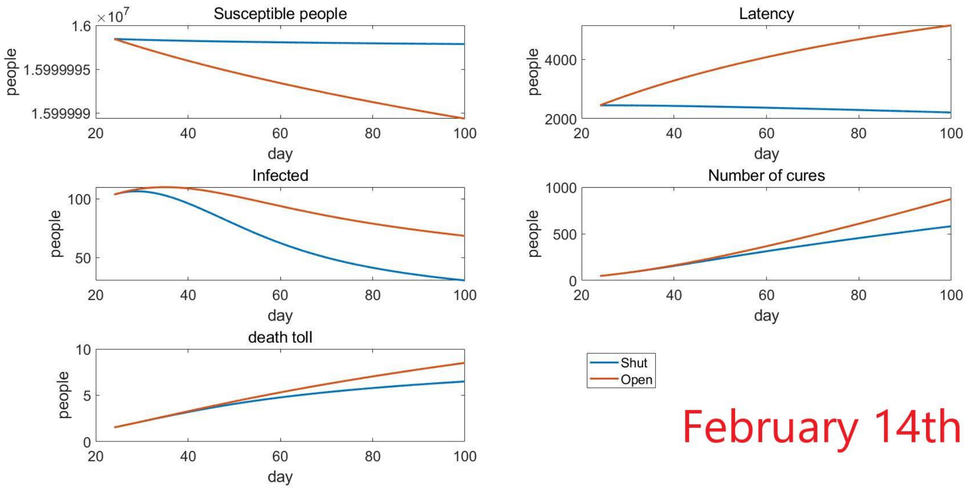


Figure 8: Changes in the epidemic on February 14

From the change trend of the number of infected people in the figure, it can be seen that the beginning of school on February 14 will cause the inflection point to move backwards, which will extend the duration of the epidemic and the number of infected people will increase dramatically. Therefore it is not recommended to start on February 14.

7.4. Predicting the Appropriate start time

In order to obtain a suitable start time, we simulated the start of the school on different days after the inflection point. After calculation, we found that the peak on February 20 did not change, so it was relatively suitable for the start of the school. The comparison of the epidemic situation between normal and delayed start is as follows:

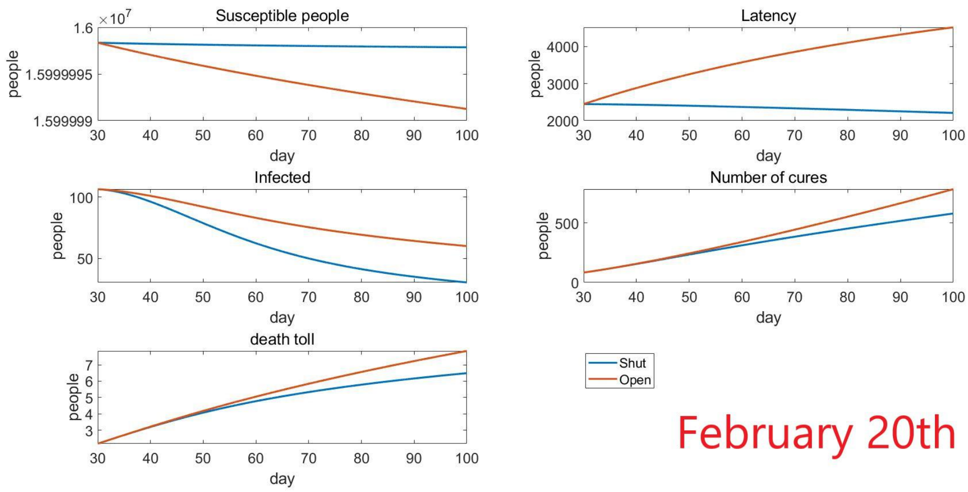


Figure 9: Changes in the epidemic situation on February 20

As can be seen from the figure above, the school will start on February 20th (the inflection point), and the inflection point will not move, it will only cause an extension of the epidemic time. In the infected person's chart, when the epidemic situation stabilizes, the number of infected people will drop to half of the inflection point, and it will be safer on March 11.

7.5. Comparative Summary Analysis

We show the impact of whether or not school starts on February 14 on people. We predict that school will start after February 20th. If we take the student's body as the first condition, we recommend that the school be started after the epidemic situation is stable, that is, the number of infected people will be reduced to half of the inflection point. If the economy is the first priority, you can start school early.

8. Model Extensions

(1) The values of infectious rate, cure rate, and mortality for a period of time in the future can be fitted based on Huanggang hospital data, traffic data, and population data.
(2) Intelligent algorithms can be used to simulate the flow of susceptible people and latent people in Huanggang City based on known data to obtain higher accuracy.

9. Conclusion

9.1. Strengths

(1) The SEIR model we set up is innovative in that it considers a group of latent people in combination with the disease. We divided the crowd into travellers who stayed indoors and outdoors in more detail, and the results were realistic.
(2) Under human intervention, we added a model to the flow of people to make the prediction more accurate.
(3) The model is fitted according to known data, and the prediction accuracy is high.
(4) The model has a wide range of considerations, which is more suitable for the characteristics of this new pneumonia.
(5) The long-term accuracy of the differential equation prediction is high.

9.2. Weaknesses

(1) The SEIR model only reflects the natural development process of the infectious disease itself, and does not consider the impact of human intervention on the disease.
(2) There is a certain degree of subjectivity in the contact rate of infected people, the cure rate and mortality of infected people, and there is a lack of certain data support. We have simplified the parameters and sacrificed some accuracy in exchange for runtime.

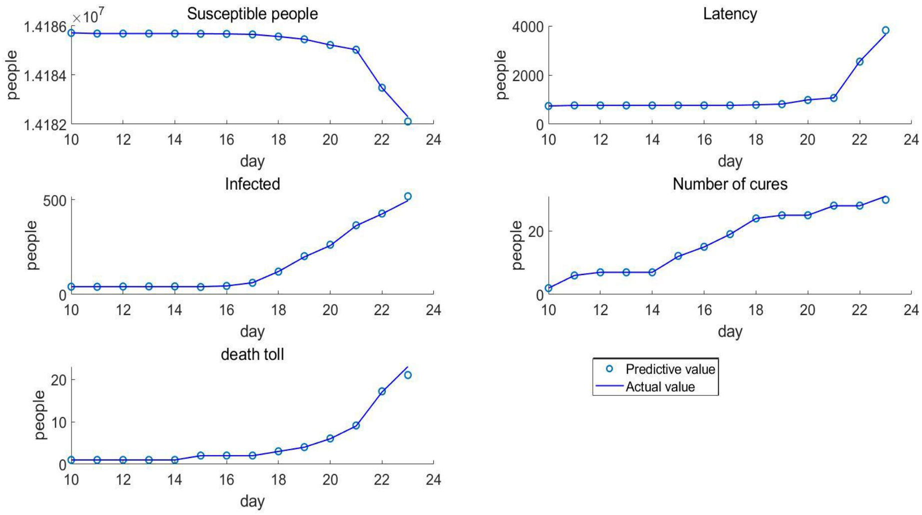
10. References

[1] Che Yanqiu. Study on the transmission model of atypical pneumonia [D]. Dalian University of Technology, 2005.
[2] Zhang Jia. SIR model and its application in SARS epidemic [D]. Shandong University, 2019.
[3] Ren Xiaoyu. Using mathematical models to analyze the effectiveness of prevention and isolation measures for atypical pneumonia [D]. Liaoning Normal University, 2004.
[4] Huang Fu Yannan. Asymptotic behavior of a stochastic SEIR epidemic model with a non-linear incidence rate [D]. Taiyuan University of Technology, 2019.
[5] Thäter Markus, Chudej Kurt, Pesch Hans Josef. Optimal vaccination strategies for an SEIR model of infectious diseases with logistic growth. [J]. Mathematical biosciences and engineering: MBE, 2018, 15 (2).
[6] Fatima-Zohra Younsi, Ahmed Bounnekar, Djamila Hamdadou, Omar Boussaid.SEIR-SW, Simulation Model of Influenza Spread Based on the Small World Network [J] .Tsinghua Science and Technology, 2015,20 (05): 460- 473.
[7] Zeng Derong, Cheng Xiaohong, Huang Ji'an. Study on the mathematical model of SARS epidemic prediction and control——an improved method based on general epidemic model [J]. Chinese Journal of Safety Science, 2004 (02): 44- 48.
[8] Qiu Wuqi, Chu Cordia, Mao Ayan, Wu Jing. Studying Communication Problems for Emergency Management of SARS and H7N9 in China. [J]. Journal of global infectious diseases, 2018, 10 (4).

[9] Goldizen Fiona C. From SARS to Avian Influenza: The Role of International Factors in China's Approach to Infectious Disease Control. [J]. Annals of global health, 2016, 82 (1).
[10] Du Wenju, Qin Shuang, Zhang Jiangang, Yu Jianning. Neimark-Sacker Bifurcation of a Discrete SEIR Epidemic Model with Infectivity in Latent Period [J]. Journal of Jinan University (Natural Science and Medicine Edition), 2016, 37 (06): 518-524.
[11] Chen Zhiqiang, Ma Zujun. Simulation Study on the Spread and Evolution of Major Infectious Diseases under Artificial Interference——Taking SARS as an Example [J] .Journal of Southwest Jiaotong University (Social Science Edition), 2011,12 (04): 131-136.
[12] Cao Zhidong, Wang Jinfeng, Han Weiguo, Gao Yige, Zeng Guang. Mathematical modeling and quantitative evaluation of SARS epidemic in Guangzhou [J]. Fudan Journal (Natural Science Edition), 2009, 48 (06): 793 -800.

11. Appendix

Appendix 1: Values of the five categories of people from February 11-22



Appendix 2: MATLAB program to solve problem one

```
clc, clear
close all
I = [425 363 258 198 121 62 45 41 41 41 41 41 41];
C = 14186500;
R1 = [28 28 25 25 24 19 15 12 7 7 7 6 2];
R2 = [17 9 6 4 3 2 2 2 2 1 1 1 1 1];
% cure = [0.066 0.077 0.097 0.126 0.198 0.306 0.333 0.293 0.171 0.171 0.171 0.146 0.049];% cure rate
% die = [0.040 0.025 0.023 0.020 0.025 0.032 0.044 0.049 0.024 0.024 0.024 0.024 0.024];%
E2 = [2556 1070 988 817 784 763 763 763 763 763 763 763 739];% suspected cases
I = fliplr (I);
R1 = fliplr (R1);
R2 = fliplr (R2);
cure = fliplr (cure);
die = fliplr (die);
E2 = fliplr (E2);
S2 = C * ones (1,13) -E2-I-R1-R2;% susceptible population
% Use the data of that day to predict the data of the next day
a1 = zeros (1,13);
a2_1 = zeros (1,13);
a2_2 = zeros (1,13);
cure = zeros (1,13);
die = zeros (1,13);
for i = 2: 13
% a_S = S2 (i-1) -a1 * E2 (i-1);
   a1 (i) = (S2 (i-1) -S2 (i)) / E2 (i-1);% a1 * m
   a2_1 (i) = (E2 (i-1) -E2 (i) + a1 (i) * E2 (i-1)) / E2 (i-1);
   a2_2 (i) = (I (i) -I (i-1) + cure (i-1) * I (i-1) + die (i-1) * I (i-1)) / E2 (i -1);
   cure (i) = (R1 (i) -R1 (i-1)) / I (i-1);
   die (i) = (R2 (i) -R2 (i-1)) / I (i-1);
% a_E2 = E2 (i-1) -a2 * E2 (i-1) + a1 (i) * E2 (i-1);
% a_I3 = I (i-1) + a2 * E2 (i-1) -cure (i-1) * I (i-1) -die (i-1) * I (i-1);
% a_R1 = R1 (i-1) + cure (i-1) * I (i-1);
% a_R2 = R2 (i-1) + die (i-1) * I (i-1);
```

```
% min_a2 = min_a2 + ((a_I3-I (i)) / I (i)) ^ 2;
% min_sum = min_sum + ((S2 (i) -a_S) / S2 (i)) ^ 2 + ((E2 (i) -a_E2) / E2 (i)) ^ 2;
end
a2 = (a2_1 + a2_2) / 2;
```

Appendix 3: MATLAB program to solve problem two

```
clc, clear
close all
% Cure = [0.066 0.077 0.097 0.126 0.198 0.306 0.333 0.293 0.171 0.171 0.171 0.146 0.049];% cure rate
% Die = [0.040 0.025 0.023 0.020 0.025 0.032 0.044 0.049 0.024 0.024 0.024 0.024 0.024];% Mortality
A1=[0,0.0378890392422192,0.00131061598951507,0,0,0.00786369593709043,0.00917431192660551,0.0275229357798165,0.112712975098296,0.142857142857143,0.285189718482252,0.195344129554656,1.45420560747664 0.54046033 ;;
% A2=[0, 0.00473139377537212, 0.00522280471821756, 0.00523918741808650, 0.00523918741808650, 0.00917103538663172, 0.0163971166448231, 0.0360190039318480, 0.0949908256880734, 0.116698341836735, 0.0923549571603427, 0.124979757085020, 0.0789841121495327];
A2=[0,0.00473139377537212,0.00327653997378768,0.000655307994757536,0,0.00393184796854522, 0.0111402359108781, 0.0270594252469392, 0.0848696665210427, 0.106957702435813, 0.0766662957605430, 0.110630597472430, 0.0656270376005216 0.0381];
% a2_1=[0,0.00541271989174560,0.00131061598951507,0,0,0.00786369593709043,0.00917431192660551,0.0275229357798165,0.0851900393184797,0.100765306122449,0.0758873929008568,0.112348178137137,0.0654205607476635];
% a2_2=[0,0.00405006765899865,0.00913499344692005,0.0104783748361730,0.0104783748361730,0.0104783748361730,0.0236199213630406,0.0445150720838794,0.104791612057667,0.132631377551020,0.108822521419829,0.137611336032389,0.0925476635514019;
% Cure = fliplr (Cure);
Cure=[0.0490000000000000,0.0975609756097561,0.0243902439024390,0,0,0.121951219512195,0.0731707317073171,0.0888888888888889,0.0806451612903226,0.00826446280991736,0,0.0116279069767442,0,0.00427];
Die=[0.0240000000000000,0,0,0,0,0.0243902439024390,0,0,0.0161290322580645,0.00826446280991736, 0.0101010101010101,0.0116279069767442,0.0220385674931130,0.009];
S2 = zeros (1,13);
E2 = zeros (1,13);
I3 = zeros (1,13);
R1 = zeros (1,13);
R2 = zeros (1,13);
S2 (1) = 14185717;
E2 (1) = 739;
I3 (1) = 41;
R1 (1) = 2;
R2 (1) = 1;
%%
I_t = [495 425 363 258 198 121 62 45 41 41 41 41 41 41];% infected
C = 14186500;% of total population
R1_t = [31 28 28 25 25 24 19 15 12 7 7 7 6 2];% healers
R2_t = [23 17 9 6 4 3 2 2 2 1 1 1 1 1];% death toll
E2_t = [3653 2556 1070 988 817 784 763 763 763 763 763 763 763 739];% suspected cases
I_t = fliplr (I_t);
R1_t = fliplr (R1_t);
```

```
R2_t = fliplr (R2_t);
E2_t = fliplr (E2_t);
S2_t = C * ones (1,14) -E2_t-I_t-R1_t-R2_t;%
%%
for i = 2: 14
    S2 (i) = S2 (i-1) -A1 (i) * E2 (i-1);
    E2 (i) = E2 (i-1) -A2 (i) * E2 (i-1) + A1 (i) * E2 (i-1);
    I3 (i) = I3 (i-1) + A2 (i) * E2 (i-1) -Cure (i) * I3 (i-1) -Die (i) * I3 (i-1);
    R1 (i) = R1 (i-1) + Cure (i) * I3 (i-1);
    R2 (i) = R2 (i-1) + Die (i) * I3 (i-1);
end
t = 10: 1: 23;
data = [S2; E2; I3; R1; R2];
data_t = [S2_t; E2_t; I_t; R1_t; R2_t];
% data_t (:, end) = [];
title_list = ('Vulnerable population', 'Lurged number', 'Infected number', 'Cure number',
'Death number');
figure (1)
subplot (3,2,1)
scatter (t, data (1, :));
hold on;
plot (t, data_t (1, :));
legend ('predicted value', 'true value');
title (title_list (1));
for i = 2: 5
    subplot (3,2, i);
    scatter (t, data (i, :));
    hold on;
    plot (t, data_t (i, :));
    title (title_list (i));
    legend ('predicted value', 'true value');
end
% C_m = mean (Cure);
% D_m = mean (Die);
% A2_m = mean (A2);
% A1 = 0.01;
me1 = abs (data-data_t);
m_error1 = mean (me1,2);% absolute error (10-22)
m_error2 = mean (me1./data_t,2);% relative error (10-22)
error1 = abs (data (:, end) -data_t (:, end)) ./ data_t (:, end);% relative error
error2 = abs (data (:, end) -data_t (:, end));% absolute error
```

Appendix 4: Huanggang patient distribution map

```
clc, clear
close all
load location.mat
% load E_yc.mat
load I_yc.mat
% load R1_yc.mat
% load R2_yc.mat
slCharacterEncoding ('UTF-8')
shp4 = shaperead ('CHN_adm3.shp');
id1 = [];
for i = 1: length (shp4)
```

```
    if (strcmp (shp4 (i) .NAME_1, 'Hubei') && strcmp (shp4 (i) .NAME_2, 'Huanggang')) ||
shp4 (i) .ID_1 == 13 && shp4 (i) .ID_2 == 133 && shp4 ( i) .ID_3 == 933)
        id1 = [id1 i];
    end
end
shp_hg = shp4 (id1);
shp_hg (1) .NAME_2 = 'Huanggang';
shp_hg (1) .NAME_3 = 'Huangzhou';
shp_hg (1) .ID_2 = 134;
shp_hg (2) .NAME_3 = 'Wuxue';
shp_hg (4) .NAME_3 = 'Tuanfeng';
red_max = max (max (I_yc));
t1 = 0; t2 = 0;
RGB = [245 245 245; 242 168 140; 226 101 84; 201 47 49; 118 22 27] / 255;
P (1: 1000,1) = 0: 1: 999;% matching matrix
P (1,2) = 1; P (2: 101,2) = 2; P (102: 301,2) = 3; P (302: 601,2) = 4; P (602: end, 2) = 5; 0
for i = 1: 17% 1-24 to 2-90
figure (i)
symspec = makesymbolspec ('Polygon', ...
        {'ID_3', 933, 'FaceColor', [RGB (P (round (I_yc (6, i)) + 1,2), :)]}, ...
        {'ID_3', 934, 'FaceColor', [RGB (P (round (I_yc (9, i)) + 1,2),:)]}, {'ID_3', 935,
'FaceColor', [RGB ( P (round (I_yc (1, i)) + 1,2),:)]}, {'ID_3', 936, 'FaceColor', [RGB (P
(round (I_yc (5, i)) + 1, 2),:)]}, ...
        {'ID_3', 937, 'FaceColor', [RGB (P (round (I_yc (10, i)) + 1,2),:)]}, {'ID_3', 938,
'FaceColor', [RGB ( P (round (I_yc (3, i)) + 1,2),:)]}, {'ID_3', 939, 'FaceColor', [RGB (P
(round (I_yc (2, i)) + 1, 2),:)]},...
        {'ID_3', 940, 'FaceColor', [RGB (P (round (I_yc (8, i)) + 1,2),:)]}, {'ID_3', 941,
'FaceColor', [RGB ( P (round (I_yc (7, i)) + 1,2),:)]}, {'ID_3', 942, 'FaceColor', [RGB (P
(round (I_yc (4, i)) + 1, 2), :)])));% hawaii and alaska are set to red
geoshow (shp_hg, 'SymbolSpec', symspec)
hold on;
scatter (location (1, :), location (2, :), 'g', '.')
axis off
title ('Huanggang City Epidemic Map');
set (gca, 'FontSize', 20);
end
slCharacterEncoding ('GBK')
```