Evaluation and Prediction of COVID-19

Summary

By studying the changes in the number of patients. Forecast the crest period and duration of the infection of infectious diseases, so that we can take corresponding quarantine to control the spread of infectious diseases.

We need to establish a dynamic model to analyze, predict, and study the spread of COVID-19. By consulting relevant information, we know that there are many models of infectious diseases, and the differential equation model is the most representative one. Based on traditional differential equations, in this article we analyze the spread of COVID-19 from a mathematical view, establish a *SEIR* model, dividing the sample into four categories: *Susceptible, Exposed (patient with incubation period), Infected, Recovered (immune, together with Dead)*. The SEIR model in the differential equation is used to attribute the dead and cured persons to the system removers, collectively referred to as the recovered population.

On this basis, we find out the changes in the number of these four groups of people per unit time to establish a differential equation and get a model. Taking Wuhan pandemic-related data, the following indicator parameters for evaluating the level of COVID-19 in the city are given as *Number of contacts per patient, Probability of patient with incubation period turning positive, Probability of infection, Probability of cure, Case fatality rate,* and the indicators were calculated using Excel. The Wuhan pandemic is divided into four stages: *Stage 1 – the early outbreak, Stage 2 – the crest outbreak, Stage 3 – the later outbreak, Stage 4 – the effective control period.*

Through the set of indicators, we used MATLAB to predict the pandemic situation in 2021 in Wuhan, Shenzhen, Moscow, and New York in a time series individually, and draw a trend curve. We believe that the level of pandemic control in Shenzhen in 2021 is very high and no patients will appear. If Moscow adopts compulsory control measures, the number of infected people will gradually decrease, and the number of cured patients will gradually increase. If New York increases medical investment, the number of patients will decrease.

Therefore, we provide a report on current situation in Shenzhen and how-to slowdown the pandemic: Call on citizens to wear masks while traveling, reduce throng contact, and limit large-scale gatherings, etc. Organizing virus detection such as nucleic acid test. Grant subsidies to hospitals and institute of research on the virus. Issue policies, demanding the managers of the activities with throng to install hand sanitize and other disinfection devices in the venues and strengthen ventilation.

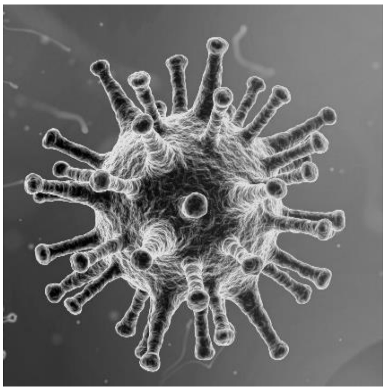
Keywords: COVID-19 pandemic; Differential equations; Data fitting; SEIR model; multistage time-lag discrete dynamic model

Contents

I. Introduction

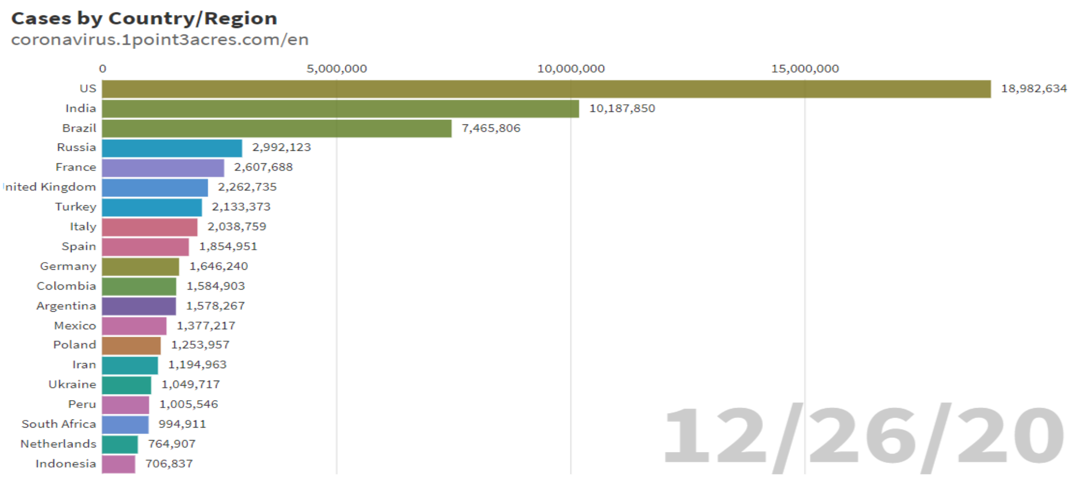
In order to indicate the level of COVID-19, the following background is worth mentioning.

1.1 Background



Nowadays, COVID-19 is of great concern to researches, governments, and all people because of the high rate of the infection spread and the significant number of deaths that occurred. In December 2019, coronavirus first reported in Wuhan, China, is an infectious disease caused by a newly discovered coronavirus. The virus that causes COVID-19 is mainly transmitted through droplets generated when an infected person coughs, sneezes, or exhales. These droplets are too heavy to hang in the air and quickly fall on floors or surfaces. Coronavirus-confirmed cases reached nearly four million in 187 countries, and approximately 295,000 people have lost their lives due to this virus.

Although this virus is extremely contagious, as long as strict control methods and effective treatment methods are adopted, the epidemic can be easily controlled.



Therefore, we must understand and analyze the spread of COVID-19, and find ways to reverse this trend, which is very important. In this report, we also adopted a mathematically rigorous model. To sum up, we will use the data of Wuhan to establish reference indicators and select relevant data of three cities-Shenzhen, Russia, and New York. And develop a prediction model to predict the epidemic situation in 2021. Then, we will try to find a solution to the same problem using this model.

1.2 Problem Summary

1. Build a mathematical model for COVID-19. Hint: all discrete model, continuous model and statistical model are available according to your understanding and specific purpose.
2. Based on your model, introduce an indicator for a city, which can be used for evaluating its level of COVID-19.
3. According to your introduced indicator, predict the level of COVID-19 for several cities in 2021.
4. Provide a one-page report to the local government. In this report, you should firstly describe the current situation of COVID-19 of the city. Then, you need to offer a recommendation on how to organize the activities in the city in order to reduce/slow down the spread of COVID-19.

1.3 Data Sources

Our model is based on the SARS virus infection model. According to the relevant epidemic data of the World Health Organization, we have collected the epidemic data of New York and Moscow in 2020, and collected the epidemic data of Shenzhen on the Shenzhen health and medical website, and collected epidemic data in Wuhan from the GitHub website.

1.4 Our work

According to the meaning of the question, it is required to establish a mathematical model of COVID-19, that is, to establish a process of the evolution of infectious viruses over time. It is necessary to study the changes in the number of people in various groups during the spread of infectious diseases. By studying the changes in the number of patients. Forecast the crest period and duration of the infection of infectious diseases, so that we can take corresponding quarantine to control the spread of infectious diseases.

We need to establish a dynamic model to analyze, predict, and study the spread of COVID-19. By consulting relevant information, we know that there are many models of infectious diseases, and the differential equation model is the most representative one. Therefore, the population is divided into four categories: ***Susceptible***, ***Exposed*** *(patient with incubation period)*, ***Infected***, ***Recovered*** *(immune, together with Dead)*, and the SEIR model in the differential equation is used to attribute the dead and cured persons to the system removers, collectively referred to as the recovered population. On this basis, we find out the changes in the number of these four groups of people per unit time to establish a differential equation and get a model. Then use MATLAB programming to draw the graph, change the quarantine strength and re-draw the graph for comparison, analyze the results, and use this model to make suggestions for controlling the spread of the pandemic.

To further present our solutions, we arrange our paper as follow.

- In section 1, On the basis of the original SIR model, we established the SEIR model and gave differential equations to calculate the indicators later.
- In section 2, Taking Wuhan as an example, we divided the pandemic situation in Wuhan into four stages, established a multistage time-lag discrete dynamic model, and calculated ***Number of contacts per patient, Probability of patient with incubation period turning positive, Probability of infection, Probability of cure, Case fatality rate*** as indicators，which can be used for evaluating level of COVID-19.
- In section 3, We use the indicators of Wuhan to predict the COVID-19 levels in Shenzhen, Moscow, and New York in 2021. We believe that the level of pandemic control in Shenzhen in 2021 is very high and no patients will appear. If Moscow adopts compulsory control measures, the number of infected people will gradually decrease, and the number of cured patients will gradually increase; if New York increases medical investment, the number of patients will decrease.
- In section 4, We provide a report to the local government that describes the current situation of COVID-19 and provides some practical ways to control the epidemic.

II. Nomenclature

Symbol	Definition
$\boldsymbol{S}$	Susceptible person
$\boldsymbol{R}$	Recovered
$\boldsymbol{I}$	Infected
$\boldsymbol{D}$	Dead patient
$\boldsymbol{E}$	Patients with incubation period
$\boldsymbol{N}$	Total population
$\boldsymbol{r}$	Number of contacts per patient
$\boldsymbol{a}$	Probability of *PIP turning positive
$\boldsymbol{b}$	Probability of infection
$\boldsymbol{y}$	Probability of cure
$\boldsymbol{k}$	Case fatality rate
$\boldsymbol{T}$	Total time
$\boldsymbol{t}$	Time from onset to diagnosis

*PIP – Patients with incubation period

Table 1: Variables and functions

Ⅲ. Assumptions

1. Assume that the number of people infected with the virus per unit time is proportional to the existing infected people;
2. Assume that the number of people cured per unit time is directly proportional to the current infected;
3. Assume that the number of deaths per unit time is directly proportional to the existing infected persons;
4. Assume that the patient will not be infected with the same virus again after healed and recovered, and has strong immunity, that is, be removed from the system;
5. Assume that normal people enter a period of incubation period after being infected, people in the incubation period will not show symptoms;
6. Assume that if a patient is admitted to the hospital, it means that the patient is treated in quarantine and is considered unable to contact others, so it will not infect healthy people;
7. Assume that after the actual cure cycle, if the patient is not cured, the patient is considered dead, that is, after the actual cure cycle, the patient is removed from the system anyway;
8. Assume that the population dynamics such as births, deaths, and mobility are ignored during the period of pandemic in the investigation area. That is, the total number of people remains unchanged, which is recorded as N;

IV. Model Construction

In order to establish the best model for predicting COVID-19, we collected the SARS virus spreading model, and established the SEIR model on the basis of the most primitive existing model.

4.1 Existing Model

First consider the I model, which is the basic model that is constructed from the number of infected people and effective contact rate, but does not distinguish between infected people (patients) and uninfected people (healthy people). It turns out that as time grows, the number of patients will increase indefinitely, which is obviously unrealistic.

The SI model is an improved model of the I model which distinguishes the infected and the uninfected. However, this model does not take into account the curability of the patient. As a result, healthy people become patients and patients cannot recover. This is obviously unrealistic.

The following is the formula, and β is infection rate. In the duration of the pandemic， the total population of the investigated area remains $S(t)+I(t) = K$. Based on the conservation relation, we have:

$$\frac{dS}{dt}=-\beta SI,\quad \frac{dI}{dt}=\beta SI$$

$$\frac{dI}{dt}=rI\left(1-\frac{I}{K}\right),\quad r=\beta K$$

Considering that patients have stronger immunity after recovery, the SIR model has improved the SI model, that is, the dead and the cured are removed from the system. It is assumed here that the patient gets permanent immunity after recovery, and therefore can be removed from the system, and the incubation period of SARS is very short and can be ignored, that is, a person becomes an infectious person immediately after being sick. For fatal infectious diseases, dead patients can also be classified. Therefore, the model has only two independent dynamic variables, and their phase trajectories satisfy:

$$\frac{dS}{dt}=-\beta IS,\quad \frac{dI}{dt}=\beta IS-\gamma I$$

$$\frac{dI}{dS}=\frac{\gamma}{\beta S}-1\Rightarrow I+S-\frac{\gamma}{\beta}\ln S=const$$

γ、β - proportional constant, γ - rate of sample removed from system, β - rate of infection.

In this case, sample is divided into 3 categories:

The first category is composed of those who are able to transmit SARS to others, and $I(t)$ is used to represent the proportion of the first category of people in the total population at the moment.

The second category is composed of those who are not ill but can become patients with the disease. Use $S(t)$ to represent the proportion of the second category in the total population at the moment.

The third category includes people who have died from the disease and those who will not be infected again after the disease recovers. Use $R(t)$ to indicate the proportion of the third category in the total population at the moment.

4.2 Our Model

In actual circumstances, susceptible patients and patients with incubation period will appear. So, we select SEIR model to evaluate the pandemic.

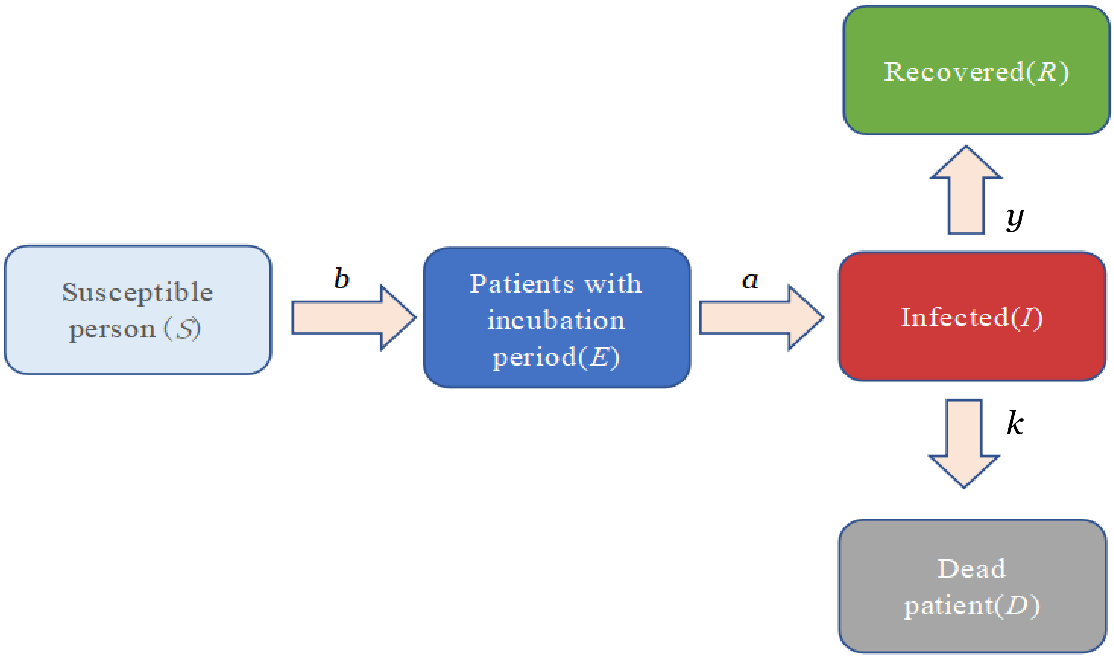


Fig.1 SEIR model diagram

Re-classifying the sample based on the SEIR model, we have the result below：

The healthy population is the susceptible. Record its number as S(t), which means people who are not infected at time t but may be infected with the disease;

Confirmed patients, that is, people infected with the disease, record their number as I(t), which means the number of patients who have been diagnosed as patients at time t;

Suspected patients, that is, people who are admitted to the hospital for isolation, including some normal people and some infected persons in the incubation period, record their number as E(t), which means the number of people admitted to the hospital who may have been infected with the disease at time t; People who have been infected with the virus but are in the incubation period, record their number as Q(t), which means the number of people who have been infected with the virus but did not show symptoms at time t, that is, the number of people in the incubation period;

The number of recovered people is R(t), which means the number of people who have been removed from the infected person at time t, including the dead and cured. This part of the population is neither a confirmed patient nor a susceptible population, and does not have They are infectious and will not be infected again, they have been removed from the infected system.

V. Indication for a city

Having established the SEIR model, in order to ascertain the indicators, we take Wuhan as an example, collect the daily new case data from January to April in 2020, and draw the image of Daily new case of Wuhan.

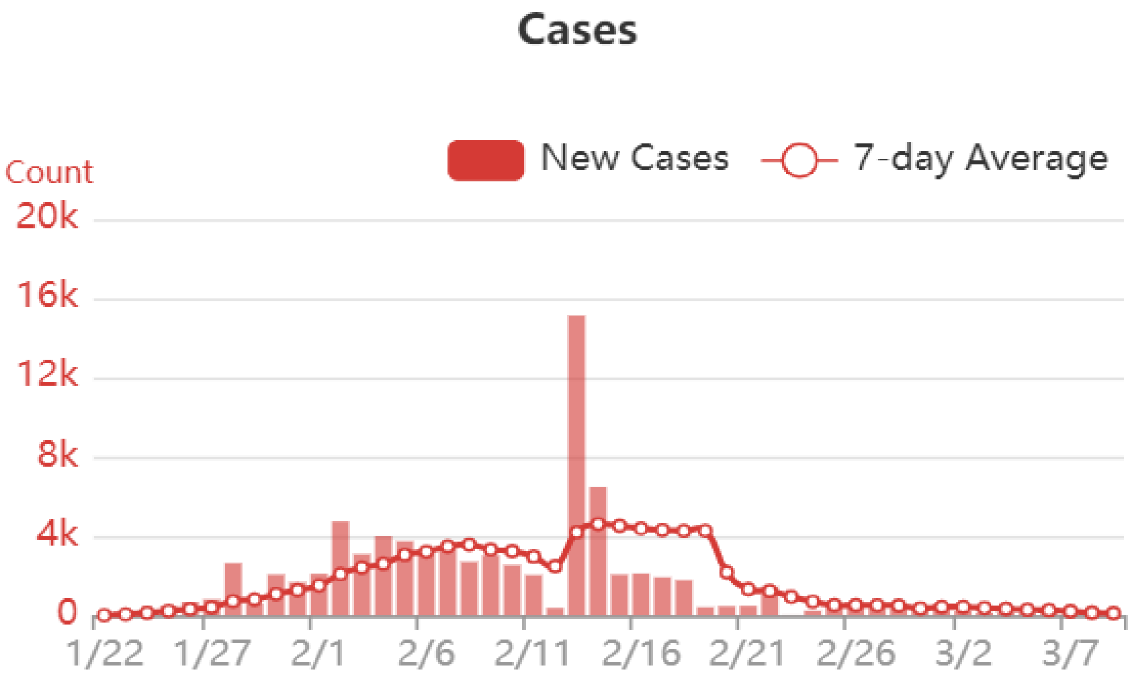


Fig.2 Images of new daily cases in Wuhan

5.1 Differential equations for COVID-19

Taking Wuhan pandemic-related data, the following indicator parameters for evaluating the level of COVID-19 in the city are given as *Number of contacts per patient, Probability of patient with incubation period turning positive, Probability of infection, Probability of cure, Case fatality rate*. Taking into account that the growth of COVID is affected by factors such as government control, medical resources, personnel flow, and per capita quality, we perform data fitting in stages and establish the following multi-stage time-lag discrete dynamics model (unit: d)

Suppose *Probability of infection* at this stage is b, *Number of contacts per patient* is r, and the daily change of susceptible persons from time t to time t+1 is the sum of the number of patients with incubation period and confirmed infections at time t, the calculation is as follows:

$$S(t+1)=S(t)-r\cdot b\cdot I(t)\cdot S(t)/N$$

From time t to time $t+1$,

Daily change of patients with incubation period equals to number of *Susceptible* that is infected by the *Infected* at time $t+1$ minus the sum of number of *patients with incubation period* and *patient with incubation period that turn positive.*

$$E(t+1)=E(t)+r\cdot b\cdot S(t)\cdot I(t)/N-a\cdot E(t)$$

From time t to time $t+1$,

The daily change in the number of confirmed cases is the number of patients converted from the incubation period to confirmed infections minus the sum of the number of deaths and the number of cured people

$$I(t+1)=I(t)+a\cdot E(t)-(y+k)\cdot I(t)$$

From time *t* to time *t+1*,

The daily change in the number of cured people is converted into the number of cured people

$$R(t+1)=R(t)+y\cdot I(t)$$

From time t to time t+1,

The daily change in the number of deaths is the number of confirmed infections converted into deaths.

$$D(t+1)=D(t)+k\cdot I(t)$$

5.2 Indicator establishment

The incubation period of COVID-19 is 7-14 days. In order to facilitate modeling and calculation, it is assumed that the average incubation period is 8 days, and the *Probability of patient with incubation period turning positive* is 0.125; the time required for COVID-19 patients from onset to diagnosis continues with the development of the epidemic reduces, so we take 2 days comprehensively. The Wuhan's pandemic is divided into four stages: *Stage 1 – the early outbreak, Stage 2 – the crest outbreak, Stage 3 – the later outbreak, Stage 4 – the effective control period.* Using excel data analysis, we calculated two parameters: *Probability of cure y, Case fatality rate k*

Since the *Probability of infection* b of *Number of contacts per patient* r are of special significance in this model environment and lack of experimental data, this paper uses the method of parameter inversion to determine the values of b and r in various situations. Set up following differential equations:

$$\begin{cases} D(t+1)=D(t)+k\cdot I(t) \\ S(t+1)=S(t)-r\cdot b\cdot I(t)\cdot \dfrac{S(t)}{N} \\ I(t+1)=I(t)+a\cdot E(t)-(y+k)\cdot I(t) \\ R(t+1)=R(t)+y\cdot I(t) \\ E(t+1)=E(t)+r\cdot b\cdot S(t)\cdot \dfrac{I(t)}{N}-a\cdot E(t) \end{cases}$$

Stage 1 – the early outbreak (2020 Jan 14~ Jan 28): The whole country is at the peak of the Spring Festival travel season and complete isolation has not been implemented. In Jan 22, Wuhan shut down the city, forbidding personnel flow in the city, and medical goods are in short supply. Analyzing data from Jan 14 to 28, we know that *Probability of cure* y_1 *is 0.23, Case fatality rate* k_1 is 0.09025，b_1 is 0.6, r_1 is 11.5.

Stage 2 – the crest outbreak (Jan 29 ~ Feb 19 2020): compulsory quarantine was adopted; all gathering was forbidden. *Number of contacts per patients* decreased to $r_2 = 1.8$. Shortage of medical goods improved as hospitals *Mount Huoshen* and *Mount Leishen* came into use. *Case fatality rate* k_2 decreased to 0.05375, *Probability of cure* y_2 increased to 0.9, and *Probability of infection* b_2 decreased to 0.6.

Stage 3 – the later outbreak (Feb 20 ~ Mar 9 2020), as the country increased its quarantine efforts and the infected and suspected infections had been accurately isolated and treated, *Number of contacts per patient* was reduced to r_3 to 1.2. Medical goods were relatively sufficient, *Probability of cure* was raised to y_3 to 0.95, people's awareness of virus prevention was improved, and *Probability of infection* was reduced to b_3 to 0.015. At this stage, the number of existing infections began to decline, and the number of people infected during the incubation period remained low. The number of new infections per day dropped to the order of a hundred. The pandemic was basically under control, and *Case fatality rate* was reduced to k_3 to 0.035373.

Stage 4 – the effective control period (Mar 10 ~ Apr 16 2020). With sufficient medical resources, after maintaining the highest level of quarantine and prevention and control methods for a period of time, the number of COVID-19 infections dropped rapidly, *Case fatality rate* k_4 was reduced to 1.8, and *Number of contacts per patient* r_4 was reduced to 1.0. *Probability of infection* b_4 was 0.015 and *Probability of cure* y_4 remained 0.0173.

Following is the indicators:

	The first stage	The second stage	The third stage	The fourth stage
Number of contacts per patient(*r*)	11.5	1.8	1.2	1.0
Probability of infection(*b*)	0.60	0.09	0.015	0.015
Probability of recovery(*y*)	0.23	0.90	0.95	0.95
Case fatality rate(*k*)	0.090	0.053	0.035	0.017

Table 2: Indicators and values

Based on the situation of 4 stages, draw the following image：

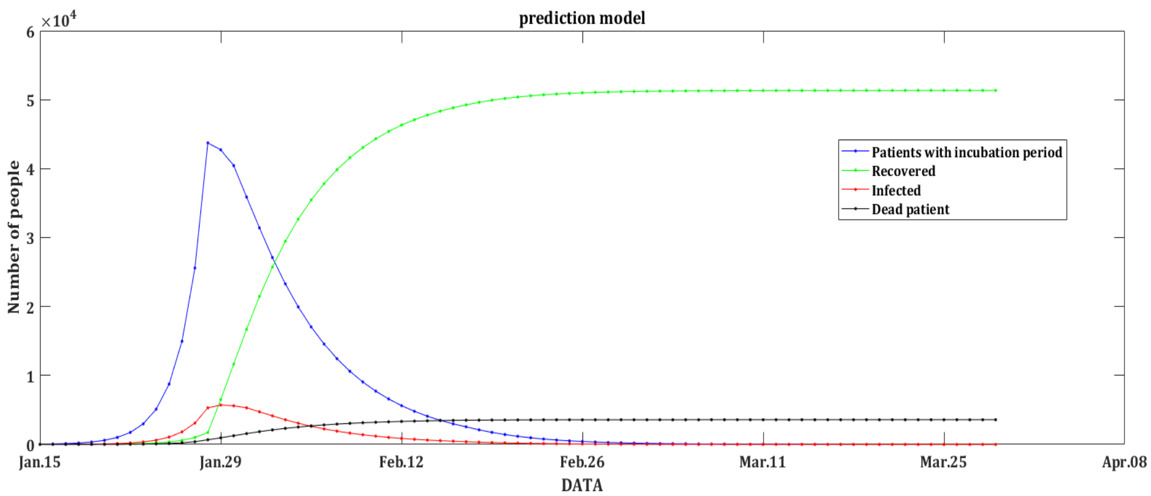


Fig.3 Fitting curve of Wuhan epidemic situation

The simulation is approximately consistent to the real situation.

VI. Prediction

According to indicator in Wuhan, we predict the level of COVID-19 for several cities in 2021 as follows: Shenzhen, Moscow, New York.

6.1 Predict the level of COVID in Shenzhen

According to the Shenzhen pandemic data statistics [1]， we drew the graph 4.1 through excel charts

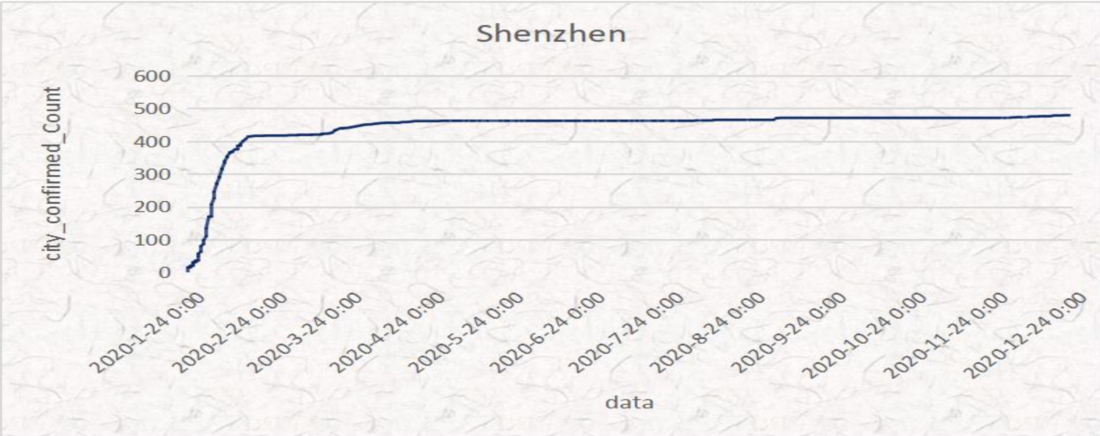


Fig4.1 The number of new cases of COVID-19 daily in Shenzhen from January to December

We select the fourth stage indicators of Wuhan City to predict pandemic in Shenzhen. Through MATLAB curve fitting, we draw the image 4.2

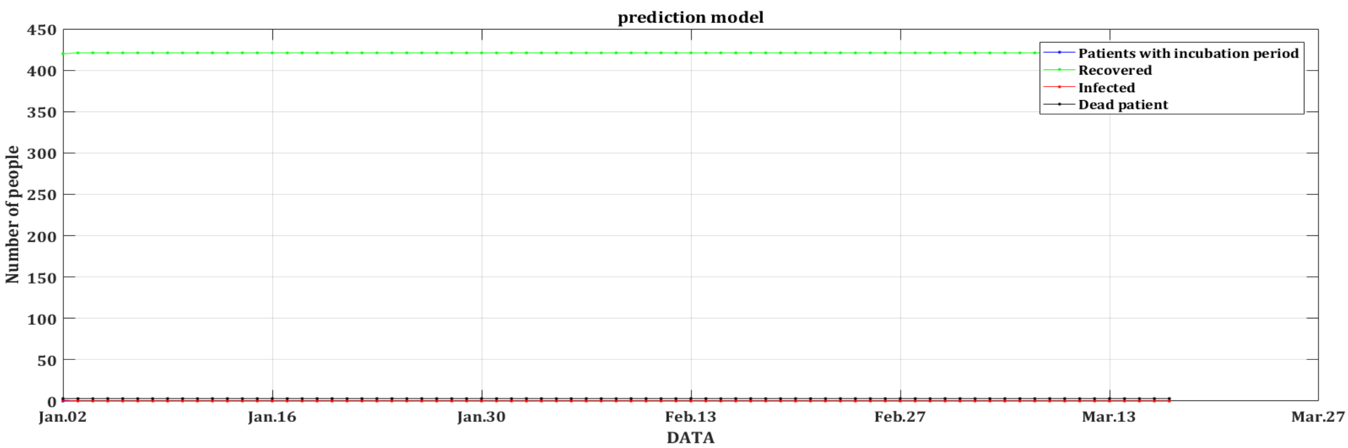


Fig4.2 Forecast image of the epidemic in Shenzhen

It can be seen from the image that thanks to the excellent pandemic prevention measures in Shenzhen and abundant medical resources, the predicted image shows that there will be very few or no COVID infection patients in Shenzhen in 2021.

6.2 Predicting the level of COVID in Moscow

By collecting the number of daily new cases of COVID in Moscow in 2020， we draw image 5.

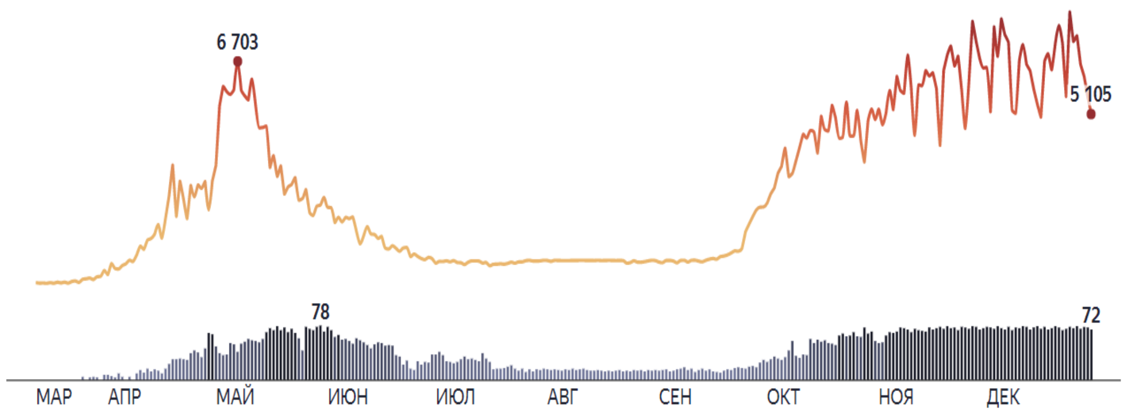


Fig.5 The number of new cases of COVID-19 daily in Moscow from March to December

According to Moscow's latest epidemic data statistics[2]， as shown in Figure 4, although the number of new people added every day in Moscow is flattened and distributed in a straight line, the epidemic is still not effectively controlled. The overall number of new cases per day is around 6000 Volatility is still a very large number.

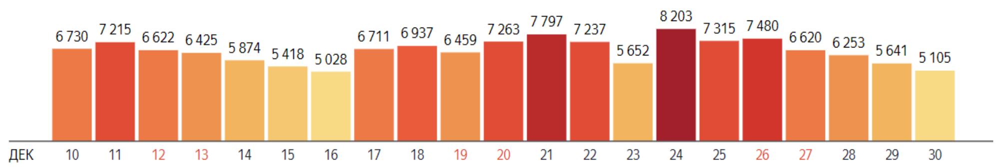


Fig.6 The latest data about new cases of COVID-19 daily in Moscow

Therefore, we consider the following two assumptions to predict the level of COVID in Moscow in 2021

Assumption 1：If Moscow take the same stringent measures as Wuhan, adopt compulsory isolation measures, stop all gathering activities, increase the isolation of patients, and invest a lot of manpower and financial resources in the medical industry to widely publicize the harm of COVID-19 and improve Per capita awareness of prevention, then we select the second stage indicators of Wuhan City to predict Moscow. Through MATLAB curve fitting, we draw the image 7.1

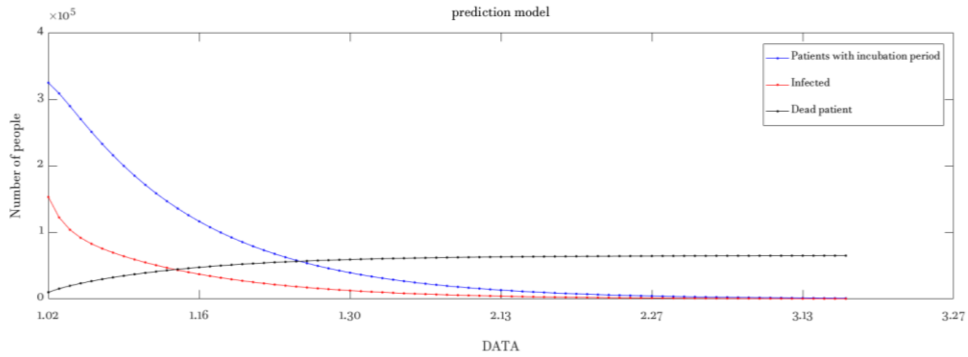


Fig.7.1 Forecast image of the Moscow epidemic

It can be seen from Figure 7.1 that if Moscow takes the above measures, the number of patients will drop quickly. At the end of January, the trend will be flat, and COVID-infected patients will gradually approach 0.

Assumption 2: If Moscow does not adopt strict measures and compulsory isolation, the isolation of patients is insufficient, medical resources are still insufficient, and the per capita awareness of prevention is low, then we will select the *Stage 1* indicators of Wuhan to predict Moscow. Through MATLAB curve fitting, we draw the image 7.2

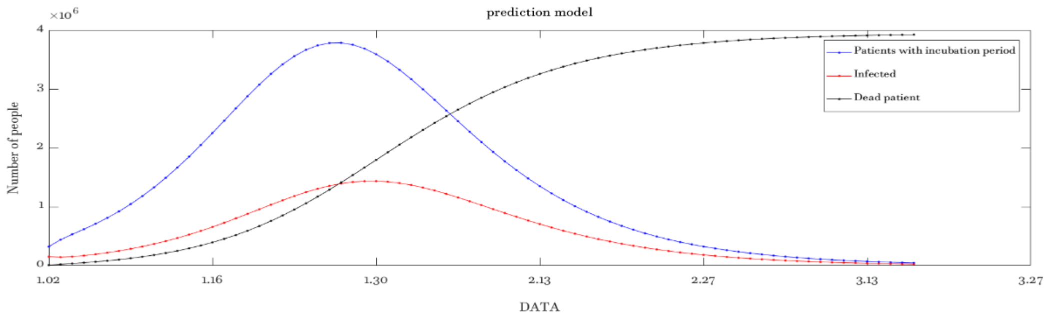


Fig7.2 Forecast image of the Moscow epidemic

It can be seen from Figure 7.2 that if Moscow does not take compulsory measures, the number of patients will gradually rise and reach a peak, and the number of deaths will continue to increase. It will reach a peak at the end of January. As the death increases, the probability of per capita contact with the sick in Moscow will decrease, so the number of patients will also drop, and the death will eventually reach its peak.

6.3 Predicting the level of COVID in New York

Based on the real-time pandemic data of New York [3], we have drawn the following urban regional epidemic distribution map, daily new patient image and daily death toll image, as shown in Figure 8.1:

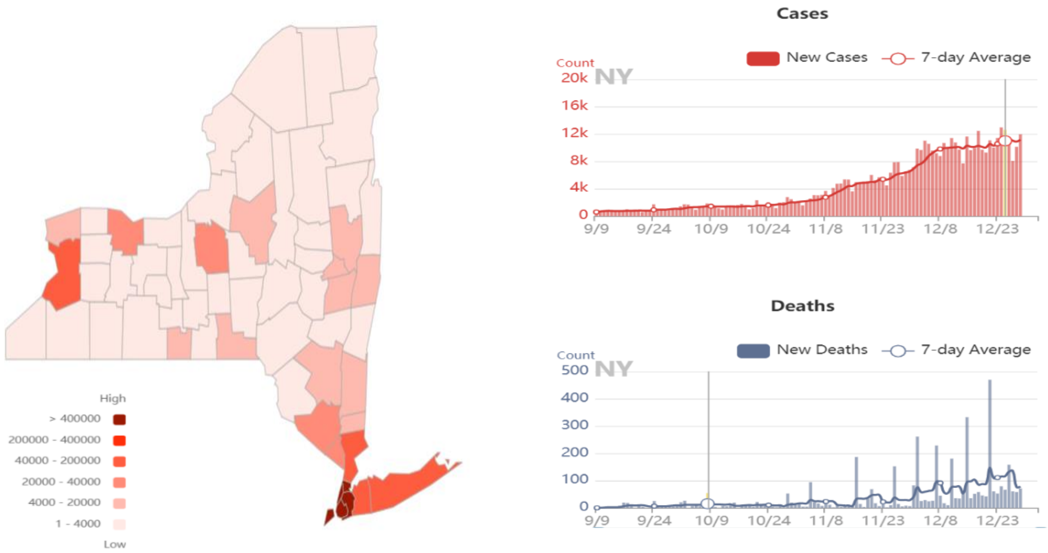


Fig8.1 The latest data about new cases of COVID-19 daily in New York

It can be seen from Figure 8.1 that the pandemic in New York is very serious and has not been effectively controlled at all. The number of deaths continues to hit new highs. The number of new patients per day is basically maintained at about 10,000. The government's management measures are not strict enough and medical resources are lacking.

However, considering the research and development of the COVID-19 vaccine and the economic strength and scientific research capabilities of the United States, we believe that if New York can adopt strong control measures and increase investment in medical undertakings in 2021, the number of infected people should be controlled. As the healing capacity increases, the mortality rate will also decrease. Therefore, we select Wuhan's second and third stage indicators to predict New York. Through MATLAB curve fitting, we draw the image 8.2

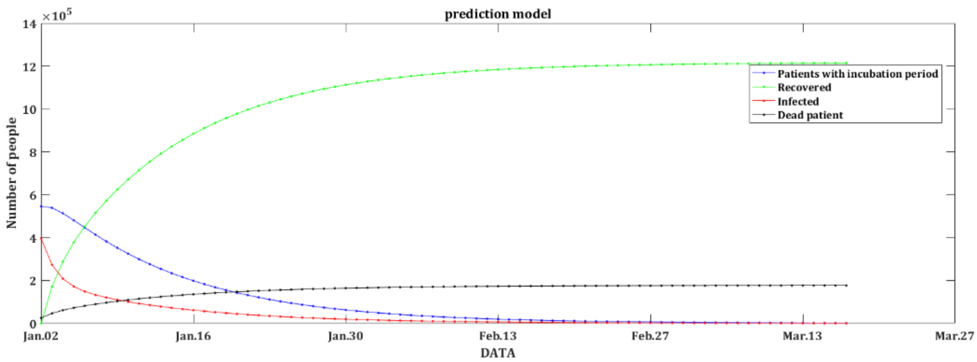


Fig8.2 Forecast image of the New York epidemic

VII. Report on situation and recommendation

To: Chief Administrator,
From: Team 017
Date: 3 January 2021
Subject: Evaluation and Prediction of COVID-19

MCM
Modelers Countering Methylfentanyl

Introduction

COVID-19, which broke out in 2019, has had an extreme impact all over the world. The government should attach great importance to COVID-19 and take measures to lower the level of COVID-19 and minimize the loss. Our team evaluate the city's level of COVID-19, referring to the related data of Wuhan city and using indicators of *Number of contacts per patient, Probability of patient with incubation period turning positive, Probability of infection, Probability of cure, Case fatality rate.* In addition, we set up a SEIR(Susceptible-Exposed-Infected-Recovered) and a multistage time-lag discrete dynamic model to predict the trend of COVID-19 in the Shenzhen city.

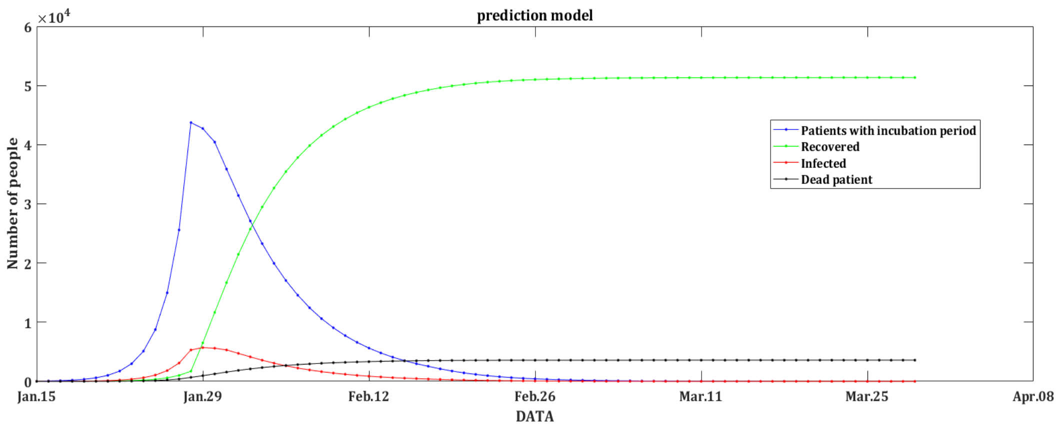
The following is our report based on the model we set up,

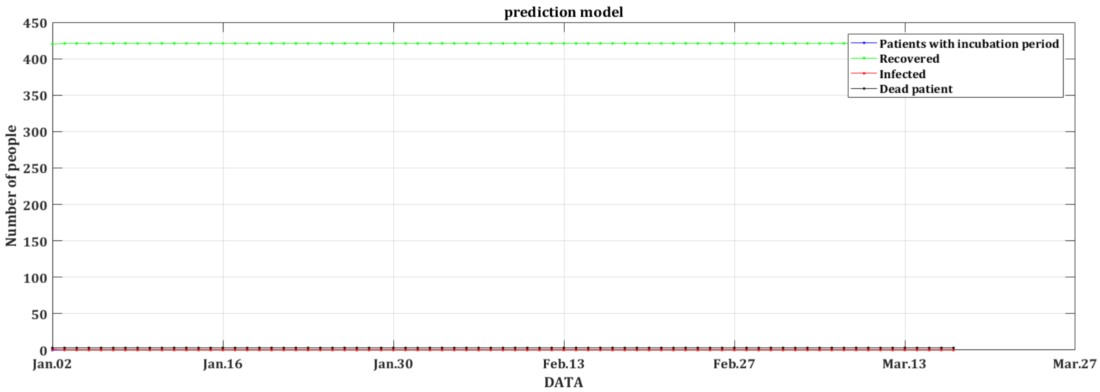
Report

Take Wuhan as an example

	The first stage	The second stage	The third stage	The fourth stage
Number of contacts per patient(r)	11.5	1.8	1.2	1.0
Probability of infection(b)	0.60	0.09	0.015	0.015
Probability of recovery(y)	0.23	0.90	0.95	0.95
Case fatality rate(k)	0.090	0.053	0.035	0.017

Table : Indicators and values





Insights

What can be inferred from the obtained images of the model is that:

1. The higher the Number of contacts per patient, Probability of patient with incubation period turning positive, Probability of infection, Case fatality rate are, the worse the pandemic will be.

2. The higher the *Probability of cure* is, the lower the level of pandemic will be.

Therefore, recommendation is given below:

Recommendation

1. Call on citizens to wear masks while traveling, reduce throng contact, and limit large-scale gatherings, etc. In that way, the chance of contact between possible virus carriers *(Exposed)* and the healthy people *(Susceptible)* is reduced, so are the *Number of contacts per patient, Probability of infection.*

2. Organizing virus detection such as nucleic acid test, which can detect the possible carriers at an early stage, and the government should provide (set up if necessary) venues for accommodation and quarantine of detected *Infected.* It can lower the *Number of contacts per patient, Probability of patient with incubation period turning positive, Probability of infection.*

3. Grant subsidies to hospitals and institute of research on the virus. It makes it possible for hospitals to take in and treat more *Infected* and conduct larger-scale nucleic acid testing, and the institute is more likely to develop a more efficient method of cure. It adds on *Probability of cure,* and lower the *Case fatality rate.*

4. Issue policies, demanding the managers of the activities with throng to install hand sanitize and other disinfection devices in the venues and strengthen ventilation, In case of unavoidable throng contact activities such as schools and public transportation. It can lower the *Probability of infection.*

Ⅷ. Strengths and Weaknesses

8.1 Strengths

- Any region only needs to provide the boundary values of the model, that is, the total number of people, the number of recovered people, the number of latent people, the number of deaths and the number of patients at a certain time. Once these five data are provided, the entire subsequent epidemic trend can be predicted.

- This model is modeled by differential equations, and the Euler forward formula can be used to quickly predict the trend. The curve drawn by MATLAB is more continuous and smoother. The parameter setting of this model is based on past data and experience. The relationship between the parameters is clear. For the calculation of fitting and expected value, each parameter can be easily obtained. The model in Annex 1 adopts a segmented approach. Different parameters need to be set at different stages, and parameters need to be adjusted continuously in different regions. This brings considerable difficulty to scientific and reasonable decision-making.

- The epidemic situation, population density, sanitary conditions and public awareness of different places are different. When designing parameters, each place can select reasonable samples according to local actual conditions for data fitting, and determine the parameters that meet the local conditions to make more scientific predictions. And truly become a model that can provide reliable and sufficient information for prevention and control.

- Because the data is known to be affected by many random factors and the regularity is disturbed, its changes cannot well express the overall regularity, and thus cannot accurately predict the epidemic situation. In response to this problem, we have performed fitting and statistical analysis on the known data, starting from the overall law, while using the existing data, we still have to accept the test of the existing data. Judging from the prediction graph obtained from the previous solution process, the predicted curve does better represent the overall change law of the existing data.

8.2 Weaknesses

- Accurate and detailed data is essential to enhance the predictive evaluation capabilities of the model. However, the detection capacity of COVID-19 this time is greatly restricted by the number of test kits. This problem was particularly prominent after Wuhan was "closed" for a week. Many patients cannot be effectively diagnosed because they are not tested in time, resulting in the number of confirmed cases lagging behind the actual number of real cases, so there will be a relatively large underestimation in the data. Since 2020-02-12, after the Wuhan Municipal Government changed the method of nucleic acid detection using kits to clinical methods (including CT methods), the cumulative number of confirmed cases has suddenly increased, and the existing confirmed cases have fluctuated greatly.

- The model has not considered the influence of susceptible people (such as doctors and nurses) on various parameters.
- The model has not considered the impact of the floating population on entering and leaving the city and the spread of the COVID epidemic caused by transportation. At the same time, the seasonal influence is ignored.
- The setting of each parameter in the model takes Wuhan as a sample, which is inconsistent with the actual situation, because the epidemic situation in various places has different characteristics, and only using Wuhan parameters to represent global parameters will inevitably bring considerable errors.

References

[1] Shenzhen data: https://ncov.dxy.cn/ncovh5/view/pneumonia

[2] Moscow data: https://yandex.ru/covid19/stat?geoId=213#statistics-table

[3] New York City data: https://coronavirus.1point3acres.com/en

[4] Model introduction: https://baike.baidu.com/item/%E4%BC%A0%E6%9F%93%E7%97%85%E6%A8%A1%E5%9E%8B/5130035?fr=aladdin

[5] Model reference paper: https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&dbname=CJFDLAST2020&filename=WHCH202005004&v=F8VFDZDxdDUSIKhy7FNEAUcqHmqd%25mmd2BZmxLrUykJty24qKOFrTFIuAtHJPFLiP3g0Z

A.Appendix

A.1. Index determination

```
clear;clc;
I=11;
R=7;
D=1;
E=0;
N=10890000;
S=N-I;
```

```
r1=11.5;
r2=1.8;
r3=1.2;
r4=1;
a=0.125;
b1=0.6;
b2=0.09;
b3=0.015;
b4=0.015;
y1=0.23;
y2=0.90;
y3=0.95;
y4=0.95;
k1=0.09025;
k2=0.05373;
k3=0.035373
k4=0.0173;
t_1=2;
T=1:75;
t0=datetime(2020,01,14);
data1=[];
for i=1:length(T)
    data1=[data1,t0+caldays(i)];
    % caldays
end
%%
for idx =1:length(T)-1
    if idx<14
        S(idx+1)=S(idx)-r1*b1*I(idx)*S(idx)/N;
        E(idx+1)=E(idx)+r1*b1*S(idx)*I(idx)/N-a*E(idx)
        I(idx+1)=I(idx)+a*E(idx)-(y1+k1)*I(idx);
        R(idx+1)=R(idx)+y1*I(idx);
        D(idx+1)=D(idx)+k1*I(idx)
    elseif idx>=14&idx<34
        S(idx+1)=S(idx)-r2*b2*I(idx)*S(idx)/N;
        E(idx+1)=E(idx)+r2*b2*S(idx)*I(idx)/N-a*E(idx-t_1);
        I(idx+1)=I(idx)+a*E(idx)-(y2+k2)*I(idx);
        R(idx+1)=R(idx)+y2*I(idx);
        D(idx+1)=D(idx)+k2*I(idx);
    elseif idx>=34&idx<42%第 3 阶段
        S(idx+1)=S(idx)-r3*b3*I(idx)*S(idx)/N;
        E(idx+1)=E(idx)+r3*b3*S(idx)*I(idx)/N-a*E(idx-t_1);
```

```
        I(idx+1)=I(idx)+a*E(idx)-(y3+k3)*I(idx);
        R(idx+1)=R(idx)+y3*I(idx);
        D(idx+1)=D(idx)+k3*I(idx);
    elseif idx>=42&idx<76
        S(idx+1)=S(idx)-r4*b4*I(idx)*S(idx)/N;
        E(idx+1)=E(idx)+r4*b4*S(idx)*I(idx)/N-a*E(idx-t_1)
        I(idx+1)=I(idx)+a*E(idx)-(y4+k4)*I(idx);
        R(idx+1)=R(idx)+y4*I(idx);
        D(idx+1)=D(idx)+k4*I(idx);
    end


end
figure
plot(data1,E,'.-b',data1,R,'.-g',data1,I,'.-r',data1,D,'.-k');
grid on;

xlabel('DATA');
ylabel('Number of people');
legend('Patients with incubation period','Recovered','Infected','Dead
patient');
title('prediction model');
```

A.2. Forecast Moscow

```
Moscow(1)
clear; clc;
I=153005;
R=545430;
D=10095;
E=324570;
N=14150000;
S=N-I;
r1=11.5;
r2=1.8;
r3=1.2;
r4=1;
a=0.125;
b1=0.6;
b2=0.09;
b3=0.015;
```

```
b4=0.015;
y1=0.23;
y2=0.43;
y3=0.95;
y4=0.95;
k1=0.09025;
k2=0.05373;
k3=0.035373;
k4=0.0173;
t_1=2;
T=1:75;
t0=datetime(2021,01,01);%mocke
data1=[];
for i=1:length(T)
   data1=[data1,t0+caldays(i)];

end
for idx =1:length(T)-1

      S(idx+1)=S(idx)-r2*b2*I(idx)*S(idx)/N;
      E(idx+1)=E(idx)+r2*b2*S(idx)*I(idx)/N-a*E(idx);
      I(idx+1)=I(idx)+a*E(idx)-(y2+k3)*I(idx);
      R(idx+1)=R(idx)+y2*I(idx);
      D(idx+1)=D(idx)+k3*I(idx);

end

figure
plot(data1,E,'.-b',data1,I,'.-r',data1,D,'.-k');
grid on;
xlabel('DATA');
ylabel('Number of people');
legend('Patients with incubation period','Infected','Dead patient');
title('prediction model');

Moscow (2):
clear;clc;
%²ÎÊýÉèÖÃ
I=153005;
R=545430;
D=10095;
E=324570;
```

```
N=14150000;
S=N-I;
r1=11.5;
r2=1.8;
r3=1.2;
r4=1;
a=0.125;
b1=0.6;
b2=0.09;
b3=0.015;
b4=0.015;
y1=0.23;
y2=0.43;
y3=0.95;
y4=0.95;
k1=0.09025;
k2=0.05373;
k3=0.035373;
k4=0.0173;
t_1=2;
T=1:75;
t0=datetime(2021,01,01);%mocke
data1=[];
for i=1:length(T)
    data1=[data1,t0+caldays(i)];

end
for idx =1:length(T)-1

        S(idx+1)=S(idx)-r1*b2*I(idx)*S(idx)/N;
        E(idx+1)=E(idx)+r1*b2*S(idx)*I(idx)/N-a*E(idx);
        I(idx+1)=I(idx)+a*E(idx)-(y1+k1)*I(idx);
        R(idx+1)=R(idx)+y1*I(idx);
        D(idx+1)=D(idx)+k1*I(idx);

end

figure
plot(data1,E,'.-b',data1,I,'.-r',data1,D,'.-k');
grid on;
xlabel('DATA');
ylabel('Number of people');
```

```
legend('Patients with incubation period','Infected','Dead patient');
title('prediction model');
```

A.3. Forecast Shenzhen

```
%SEIRÄ£ÐÍ
clear;clc;
%²ÎÊýÉèÖÃ
I=1;
R=420;
D=3;
E=0;
N=13430000;
S=N-I;
r1=11.5;
r2=1.8;
r3=1.2;
r4=1;
a=0.125;
b1=0.6;
b2=0.09;
b3=0.015;
b4=0.015;
y1=0.23;
y2=0.90;
y3=0.95;
y4=0.95;
k1=0.09025;
k2=0.05373;
k3=0.035373;
k4=0.0173;
t_1=2;
T=1:75;
%% Ê±¼ä
t0=datetime(2020,01,14);
data1=[];
for i=1:length(T)
   data1=[data1,t0+caldays(i)];

end
%%
```

```
for idx =1:length(T)-1

       S(idx+1)=S(idx)-r4*b4*I(idx)*S(idx)/N;
       E(idx+1)=E(idx)+r4*b4*S(idx)*I(idx)/N-a*E(idx);
       I(idx+1)=I(idx)+a*E(idx)-(y4+k4)*I(idx);
       R(idx+1)=R(idx)+y4*I(idx);
       D(idx+1)=D(idx)+k4*I(idx);

end
figure
plot(data1,E,'.-b',data1,R,'.-g',data1,I,'.-r',data1,D,'.-k');
grid on;

xlabel('DATA');
ylabel('Number of people');
legend('Patients with incubation period','Recovered','Infected','Dead patient');
title('prediction model');
```

A.4 Forecast New York

```
clear;clc;
I=396785;
R=0;
D=25055;
E=545781;
N=8510000;
S=N-I;
r1=11.5;
r2=1.8;
r3=1.2;
r4=1;
a=0.125;
b1=0.6;
b2=0.09;
b3=0.015;
b4=0.015;
y1=0.23;
y2=0.43;
y3=0.95;
y4=0.95;
k1=0.09025;
```

```
k2=0.05373;
k3=0.035373;
k4=0.0173;
t_1=2;
T=1:75;
%% Ê±¼ä
t0=datetime(2021,01,01);
data1=[];
for i=1:length(T)
    data1=[data1,t0+caldays(i)];

end
%%
for idx =1:length(T)-1
    if idx<70
        S(idx+1)=S(idx)-r2*b2*I(idx)*S(idx)/N;
        E(idx+1)=E(idx)+r2*b2*S(idx)*I(idx)/N-a*E(idx)
        I(idx+1)=I(idx)+a*E(idx)-(y2+k2)*I(idx);
        R(idx+1)=R(idx)+y2*I(idx);
        D(idx+1)=D(idx)+k2*I(idx);
    elseif idx>=70
        S(idx+1)=S(idx)-r3*b3*I(idx)*S(idx)/N;
        E(idx+1)=E(idx)+r3*b3*S(idx)*I(idx)/N-a*E(idx-t_1);
        I(idx+1)=I(idx)+a*E(idx)-(y3+k3)*I(idx);
        R(idx+1)=R(idx)+y3*I(idx);
        D(idx+1)=D(idx)+k3*I(idx);
    end
end
figure
plot(data1,E,'.-b',data1,R,'.-g',data1,I,'.-r',data1,D,'.-k');
grid on;

xlabel('DATA');
ylabel('Number of people');
legend('Patients with incubation period','Recovered','Infected','Dead patient');
title('prediction model');
```