

# Research on the Analysis and Prediction of the Coronavirus Pandemic Based on the TOPSIS Model and SEIR Model

**Peng Liangzheng**

*School of Statistics and Mathematics, Guangdong University of Finance & Economics, Guangzhou, Guangdong, 510220, China  
pengliangzheng1999@163.com*

**Abstract:** *In this paper, using the statistical data set of the global coronavirus epidemic of Johns Hopkins University, at the beginning of the third round of epidemic outbreak, the TOPSIS model was established, and the epidemic prevention and control of five governments were comprehensively evaluated and analyzed. And finally, it calculates their prevention and control scores within 100 days and carries out the corresponding analysis. The outbreak in Changchun, China, in early March 2022, was affected by the third round of the global epidemic. In this paper, an SEIR model was developed to predict and analyze the outbreak in Changchun. The model predicts that the current round of the epidemic will last two months, and the number of infected people will reach more than 25,000.*

**Keywords:** *TOPSIS model; SEIR model; Analysis of forecast; Epidemic*

## 1. Introduction

Since the world's first case of coronavirus was reported in Wuhan at the end of 2019, the epidemic has not subsided for more than two years but has become much more severe. Presently, the third round of the epidemic has broken out worldwide. Western countries' prevention and control measures, represented by the United Kingdom, have shifted from eliminating the epidemic to coexisting with the virus. However, the Chinese government adheres to the principle of safeguarding people's right to life and health and still adopts stricter measures for regional control. It is vital to establish a reasonable model to predict the development of the epidemic in China and the world under the pressure of internal prevention of spillover and external prevention of importation.

## 2. Research Background

Abroad, Ghosal [2] et al. built a model to predict weekly deaths from COVID-19 in India. They used linear and multiple regression for prediction and deployed autoregression to improve the model's predictive power. The projected model is based on an analysis of data from 15 high-infection countries. Jaiswal [3] et al. deployed DenseNet for classification. The proposed method achieves high accuracy. Singh et al. [4] analyzed time-series data and predicted the number of reported cases and deaths (mortality) per case based on World Health World Population data from COVID-19. The study concluded that routine mortality from COVID-19 positively correlates with the number of confirmed cases.

Classical SEIR and other epidemiological mathematical models and their variants are widely used in various studies. In general, susceptibility-exposure-infection-recovery (SEIR) type epidemiological models are commonly used to simulate epidemics with incubation periods. Yuan Min [6] modeled the incidence and death of COVID-19 in different countries through the Prophet, XGBoost, and LSTM models and then analyzed and predicted. In other to verify the prediction effect of different models, the variance was compared. The variance flashback method is used to analyze and predict different models to predict the daily new cases better. Zhang Qing [7] established a SIR infection model by arranging relevant data. Taking into account the characteristics of virus mutation, she used the temporal convolutional neural network TCX to predict new daily cases. Geng Hui [8] and others established the SEIR model. For analyzing the effectiveness of prevention and control, parameters such as the change rate of the infected population and the infection rate during the incubation period were innovatively added to predict the development of the epidemic. They put forward the view that the inflection point of the

new crown outbreak in Wuhan will appear in March due to the practical measures taken by the Chinese government.s

### 3. Model Principle

#### 3.1 TOPSIS Model

This paper uses the Johns Hopkins University Global New Crown Epidemic Statistical Dataset [1], from which the epidemic data of China, India, Russia, the United Kingdom, and Mexico are obtained. It includes multiple indicators, such as the number of newly diagnosed patients in each country, the cumulative number of confirmed cases, the cumulative number of deaths, the cumulative number of cured patients, and the number of doctors per million people.

Excerpts are from the dataset for the 100 days from April 27, 2021, to August 4, 2021, for the five countries mentioned above. On the one hand, the period before and after the outbreak of the third round of epidemics has great reference significance. On the other hand, as many western countries no longer update the daily death toll from COVID-19 infection, the data set also stops updating the death toll.

After that day, some countries could not calculate the mortality rate and lacked an indicator, so the data from these 100 days were selected.

The indicator's peak value is obtained through the relevant panel data, and the peak value is regarded as the data to be processed. Then the processed data is standardized, and the closest information is found through principal component dimensionality reduction. Through principal component analysis, it is found that the cumulative contribution rate of the first three indicators in this study exceeds 80%, indicating that the first three principal components contain the information of all indicators. The first three main components are the cumulative number of confirmed cases, the growth rate of newly diagnosed patients, and the death rate [9].

The required formula is as follows:

To calculate the growth rate  $n_1$  of the number of newly diagnosed patients, two variables,  $p_1$  and the cumulative number of confirmed patients  $p_4$ , need to be introduced:

$$n_1 = \frac{p_1}{p_4} \quad (1)$$

To calculate the death rate  $n_2$ , we need to introduce two variables: the cumulative number of deaths  $p_2$  and the cumulative number of confirmed cases  $p_4$ :

$$n_2 = \frac{p_2}{p_4} \quad (2)$$

To calculate the cure rate  $n_3$ , we need to introduce two variables: the cumulative number of people  $p_3$  and the cumulative number of confirmed patients  $p_4$ :

$$n_3 = \frac{p_3}{p_4} \quad (3)$$

A standardized evaluation matrix is obtained for the comprehensive evaluation matrix of the utilization degree of epidemic development variables.

$$X = (x_{ij})_{m \times n} \quad (4)$$

Then the judgment matrix is normalized, and a weighted decision matrix is constructed. The weight of each indicator is multiplied by the dimensionless matrix formed to calculate the weighted decision matrix:

$$z = (r_{ij})_{mn} \quad (5)$$

And this is the range.  $(i = 1, 2, \dots, n; j = 1, 2, \dots, m), r_{ij} = 2_i x_{ij}^*$

The positive ideal solution vector  $Z^+$  and the negative ideal solution vector  $Z^-$  of the matrix  $Z$  are determined - the distances between each evaluated unit and the ideal and negative ideal solutions are calculated. The Euclidean distance is generally used to calculate the relative proximity of each evaluation unit to the optimal value:

$$C_i = \frac{D_i^-}{D_i^+ + D_i^-} \times 100\% \tag{6}$$

Among it,  $i=1, 2, \dots, n$ .

Sorted by relative proximity, the higher the  $C_i$  is, the closer it is to 100, which means that the  $i$ -th evaluation unit is closer to the optimal level. The lower the  $C_i$ , the worse the  $i$ -th evaluation unit's level.

In the evaluation process, weight is an important influencing factor that directly affects the evaluation results. Traditional weight determination methods, such as expert survey, Delphi method, binary comparison, and AHP, are usually subjective. Because people are subjective, it is inevitable to bring personal subjective factors when evaluating the results, leading to deviations or even errors. The entropy weight method can more objectively reflect the information order of the data itself. It determines the index's weight by evaluating a judgment matrix formed by the index values. In this way, each factor's subjectivity of the weights can be eliminated as much as possible, making the evaluation results more realistic. According to the definition of entropy, there are  $m$  schemes and  $n$  evaluation indicators, and the entropy of the evaluation indicators can be determined [5] as:

$$H_j = \frac{1}{-\ln m} \left( \sum_{i=1}^m f_{ij} \ln f_{ij} \right) \tag{7}$$

$i=1, 2, \dots, n$ .

$$f_{ij} = \frac{b_{ij}}{\sum_{i=1}^m b_{ij}} \tag{8}$$

In order to make  $\ln f_{ij}$  meaningful, it is generally assumed that when  $f_{ij}=0$ ,  $f_{ij} \ln f_{ij}=0$ .

But when  $f_{ij}=1$ , the result of  $\ln f_{ij}$  is also zero, but this does not conform to the actual situation and is contrary to the meaning expressed by entropy, so  $f_{ij}$  needs to be corrected, which can be:

$$f_{ij} = \frac{(1 + b_{ij})}{\sum_{i=1}^m (1 + b_{ij})} \tag{9}$$

Calculate the entropy weight  $W$  of the evaluation index:

$$w = (w_j)_{1 \times n} \tag{10}$$

Among them:

$$w_j = \frac{1 - H_j}{n - \sum_{j=1}^n H_j}; \sum_{j=1}^n w_j = 1 \tag{11}$$

### 3.2 SEIR Model

By the SEIR model theory, human beings in the world are divided into five types. People who are not infected are regarded as susceptible. Those infected with the virus but with no clinical manifestations are regarded as Infiltrators. And those who are clinically diagnosed with the virus are considered sick. Divide the removers into two categories: the recovered and the dead. In these categories, except the dead cannot be transformed into each other. The rest can be transformed into each other.

Infectious diseases are complex. Through the big data analysis of SARS cases, it can be known that different categories of people are converted to each other according to a certain proportion in the new coronary pneumonia virus infectious disease process.

Because patients infected with new coronary pneumonia will have no symptoms of infection, they will be contagious and infect others. However, due to individual physical differences, some asymptomatic patients who do not have the disease will heal themselves in time and not develop the condition, so this model introduces the self-healing patient  $c$  as the self-healing rate [10].

At the same time, considering that the Chinese government has adopted zero-tolerance measures in the face of the epidemic, it has set up quarantine areas in the fever clinics of hospitals and adopted a proactive approach to treating infected people. Hence, the equation introduces a new medical quarantine personnel  $Q$ , while for those infected, its treatment rate is  $q$ . Therefore, the mortality rate  $k_1$  and the cure rate  $r_1$  of patients in medical isolation should be divided into two different variables and considered separately. Because of effective isolation measures, the patient is isolated, so he is no longer infectious after isolation [11].

China has taken strong measures, has sufficient medical resources to treat patients, and the advantages of the governance system can take regional control measures, which can significantly reduce the transmission capacity of infected and latent people in the population. Therefore, the model introduces the infection coefficient  $h$ , which is reduced due to the practical measures of the Chinese government. In addition, due to the medical isolation of the population, every day, the susceptible population  $S$  becomes less sensitive to the medical isolate  $F$ , the ratio of  $f$ . Those vulnerable people who become self-isolating are now considered perfectly safe.

Based on the model assumptions above as shown in Figure 1.

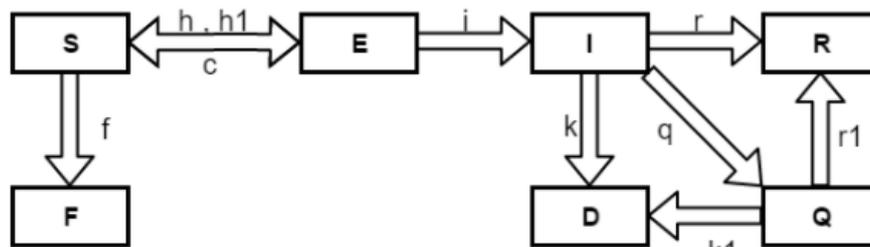


Figure 1: SEIR model schematic

So, establishing a new coronavirus transmission model based on the SEIR model is as follows:

$$\left\{ \begin{aligned} \frac{ds}{dt} &= -\frac{hIS}{N} - \frac{h_1ES}{N} + cE - fS \\ \frac{dE}{dt} &= \frac{hIS}{N} + \frac{h_1ES}{N} - iE - cE \\ \frac{dI}{dt} &= iE - rI - kI - qI \\ \frac{dR}{dt} &= rI + r_1Q \\ \frac{dD}{dt} &= kI + k_1Q \\ \frac{dQ}{dt} &= qI - r_1Q - k_1Q \end{aligned} \right. \quad (12)$$

Among them, letter  $I$  represents the number of infected persons.  $S$  represents the number of susceptible persons.  $N$  means the total population of China, while  $h$  means the infection coefficient.

Taking SARS as the standard, the average infection rate for every 20 infected persons is 0.03.  $i$  represents the probability of the latent person getting sick;  $r$  represents the probability of recovery;  $k$  represents the probability of death.  $h_1$  means the latent person's infection coefficient,  $c$  represents

patients' negative conversion rate after an illness. Q means those who are in medical isolation, q represents the rate of patients being treated in hospital after Covid-19 infection, k1 represents the mortality rate of patients isolated after infection, and r1 represents the cure rate of patients isolated.

After iterative calculation of the differential equation, the following equation is obtained:

$$\begin{cases} S_{n+1} = S_n - \frac{hI_n S_n}{N} - \frac{h_1 I_n S_n}{N} + cE_n - fS_n \\ E_{n+1} = E_n + \frac{hI_n S_n}{N} - iE_n - cE_n \\ I_{n+1} = I_n + iE_n - rI_n - kI_n - qI_n \\ R_{n+1} = R_n + rI_n + r_1 Q_n \\ D_{n+1} = D_n + kI_n + k_1 Q_n \\ Q_{n+1} = qI_n - r_1 Q_n - k_1 Q_n \end{cases} \quad (13)$$

#### 4. Model Solution

##### 4.1 TOPSIS Model Solution

The TOPSIS method is a comprehensive evaluation method widely used at present, and this method does not have strict requirements for the number and data form of the evaluation objects and indicators. An intuitive analysis result can be achieved through relatively simple calculations. In addition, during the calculation process, the original data is unified and treated in the same direction, eliminating the effects of different indicators due to different dimensions and making the evaluation results more scientific and objective. Model data organized by the above articles are the number of new diagnoses, people's mortality, growth rate, and treatment rate, build a TOPSIS comprehensive evaluation model, rating the daily epidemic prevention and control of governments [12].

Indian government hundred-day scoring chart in figure 2.

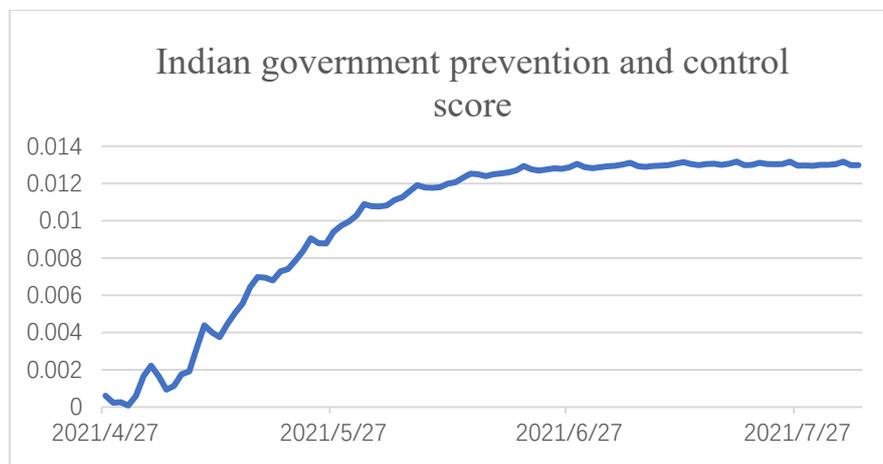


Figure 2: Government of India Prevention and Control Score

After the Indian government experienced a like-disaster April, at 11:30 pm on May 5, the number of diagnosed patients in India exceeded 400,000 cases that day, and the number of dead patients was 3971. These records broke the highest record of COVID-19 in India. Subsequently, although some people said that the epidemic in India had faded, the number of cases is greatly reduced, but this is not a sign of India being out of the haze of the epidemic. From the current situation, India's epidemiology data is no longer like a thriller, and the reason has a lot to do with the failure to continue the testing activities. As we all know, in rural areas in India, many are not isolated. No asymptomatic infected people. As far as the current vaccination in India is concerned, the coverage is far from satisfactory. Not only is the number of inoculating people less and the coverage rate low, but there is still a long way to go toward group immunity. India has a more developed medical industry; if the vaccine formula and raw materials are

supported, the vaccine can be quickly generated. However, the Indian authorities failed to prevent the outbreak of the epidemic. As a result, many vaccine factories were forced to close because they did not have workers, which seriously impacted vaccines' global supply. In the case where this epidemic is still spreading in India, India cannot guarantee vaccination used to respond to the spread of the next wave of the virus.

British government hundred-day scoring chart in figure 3.

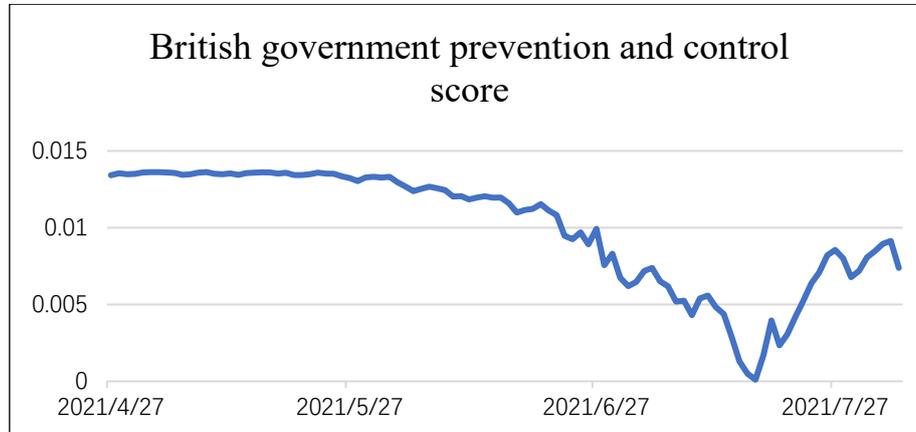


Figure 3: UK Government Prevention and Control Score

As the number of vaccinations in the UK continued to rise, on July 5, 2021, for the epidemic situation in the UK, the British government made a judgment it is considered to cancel the corresponding epidemic prevention measures. British Prime Minister Boris Johnson announced at the subsequent evening press conference that many measures are to be canceled on July 19 for Britain's COVID-19 diseases. Therefore citizens in the UK do not need to maintain social distancing and are forced to wear masks. But he also noticed by then, the number of infections in COVID-19 may be as high as 50,000 people per day. What is even more exaggerated is that the British Prime Minister encourages the people and viruses to coexist. This series of solutions resulted in British daily confirmed cases rising, most of which are "Delta" strains.

Mexican government hundred-day scoring chart in figure 4.

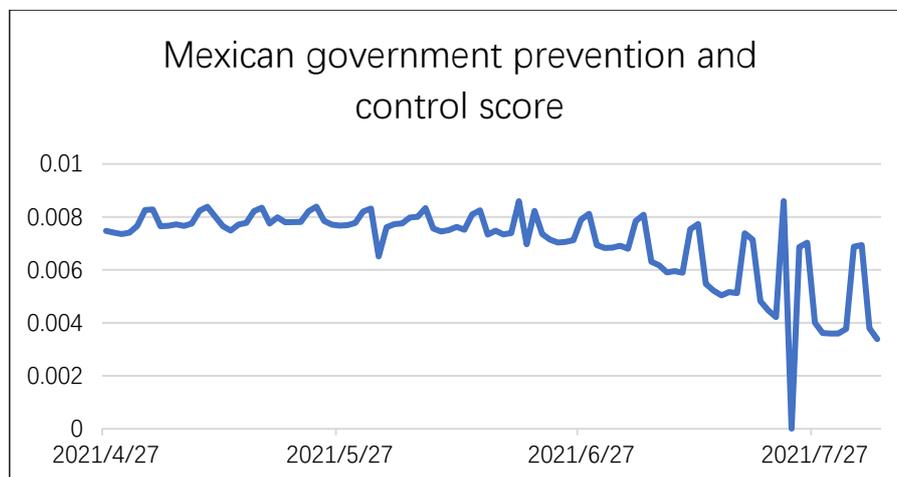


Figure 4: Mexican government prevention and control score

Due to the paralysis of early medical infrastructure and the mutation of the new coronal virus, a large number of cases of illness have occurred, improving the mortality rate in Mexico. Therefore it leads to a lower score, but with the arrival of the vaccine and vaccination with Mexican people, the Mexican government successfully spent the first and second waves of the epidemic in its domestic outbreak, tending to stabilize.

However, after entering mid-June, a significant rebound in the epidemic of COVID-19 occurred in Mexico. Subsequently, the Ministry of Health of Mexico announced the arrival of the third wave of COVID-19 on July 6, 2021, reminding the public and relevant units to prepare accordingly. More than a month after it announced the arrival of the third wave of the epidemic, Mexico ushered in the third test

of COVID-19.

Russian government hundred-day scoring chart in figure 5.



Figure 5: Russian government prevention and control score

In June, the Russian government lifted most of its limited restrictions and announced that the epidemic would end. Restaurants, bars, and shops are open as usual, most workers have returned to the office, and people are packing for the event. Some regions are scrambling to re-implement measures following the June surge. In Moscow, the mayor's office has adopted a more positive attitude, and the company has been instructed to allow some employees to work from home. At 11 pm, the bar closed. Several areas with severely affected areas have been re-implemented. However, in most places, restrictions remain very loose, life did not change fundamentally, and as a result, the Delta virus has ravaged Russia.

Chinese government hundred -day scoring chart in Figure 6:

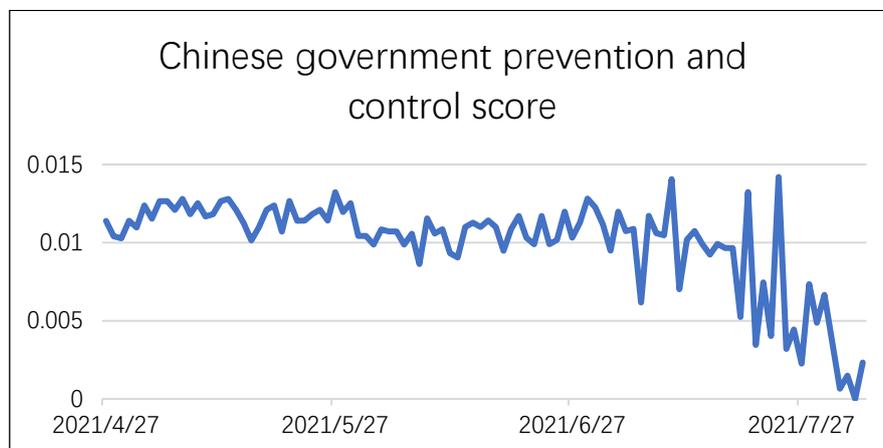


Figure 6: Chinese government prevention and control score

The Chinese government has chosen more active epidemic prevention measures than Western countries, people have highly complied with social and physical restrictions, and the command of wearing masks is derived from the cultural value that emphasizes common interests. China has already cleared zero, and new cases are mostly foreign input cases. Against the Chinese government for foreign input cases, take corresponding measures in time, the virus being rejected outside the country, give the Chinese people a safer living environment than Western countries. But after the virus has changed many times abroad and become more "cunning", it is more difficult to be discovered. There is even an incubation period of patients as high as 21 days. Although many times tests during the isolation period, there is still no timely investigation to do further isolation, there are many exfoliating cases after July, presenting regional characteristics.

#### 4.2 SEIR TOPSIS Model Solution

The SEIR model consists of four variables: Susceptible variable, Exposed variable, Infected variable, and Remove variable. Its meaning is susceptible person, lurker, an infected person, and removal. No infected virus is regarded as a susceptible person, the infected virus but no clinical manifestations is regarded as a lurker, the crowd who confirmed through clinical manifestations are regarded as an infected person, and patients who had cured or died are regarded as the removal.

We establish the infectious dynamics SEIR model by analyzing a new round of epidemics in Chinese Mainland in 2022. When we establish a parameter, conduct amendments to the Chinese epidemic prevention measures, and solve the model, a reasonable prediction of China's future epidemic is made.

At present, the coronary virus is still raging globally, with more than 30 million confirmed cases. Many countries have not seen the turning point of this epidemic disease; China has already entered globalization, and the epidemic has influenced it. In March 2022, China Changchun was attacked by the epidemic, which caused many personnel infections. This model is established in accordance with China's national conditions in figure 7 and can be used to predict the development of this Changchun epidemic.

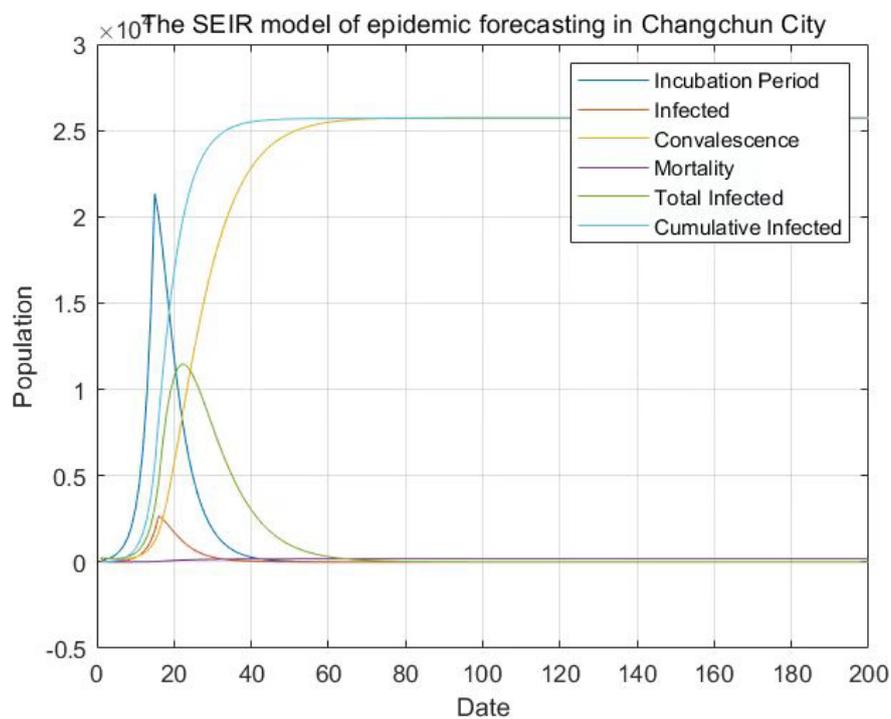


Figure 7: Changchun outbreak forecast map

It can be seen from the prediction that the outbreak of this round in Changchun, if effective epidemic prevention measures can be taken in time, will last for two months. The number of daily new infections will peak at 18 days, the number of daily new added infected will decline, and asymptomatic infected will also reach its peak on the 18th, but China has a high vaccination rate. The new coronary pneumonia virus mainly adapts auxiliary therapy, and the most important thing is to rely on the human self-immune system. Therefore there will be a considerable part of the asymptomatic infection with strong self-immunity capacity, before being detected, will not report it. This model regards it as the person who does not infect with COVID-19. The number of infections in the final report reached about 25,000.

#### 5. Conclusion

Since it was discovered in Wuhan at the end of 2019, it has not disappeared like SARS as summer approaches. As a result, it has become more intense over time. Through the TOPSIS model, we can timely monitor epidemic prevention in various countries in light of the third round of outbreaks worldwide. By observing which national government has poor prevention and control measures, we can modify them in time to avoid the spread of the epidemic, avoiding large-scale property and life losses. The outbreak of the world epidemic has inevitably affected China, which is experiencing cases in many

areas across the country. Using this article, we can forecast and analyze epidemic situations throughout the country according to the SEIR model that has been revised following China's national conditions, thus allowing the local government to take more targeted measures to prevent and control the epidemic and prevent further spread.

In summary, in the face of the intensified epidemic, the analysis and Prediction of COVID-19 is a field of sustainable research and deserves in-depth research.

## References

- [1] Ensheng Dong, Hongru Du, Lauren Gardner. *An interactive web-based dashboard to track COVID-19 in real time*[J]. *The Lancet Infectious Diseases*,2020,20(5):533-53.
- [2] Ghosal S, Sengupta S, Majumder M, et al. *prediction of the number of deaths in India due to SARS-CoV-2 at 5–6 weeks*[J]. *Diabetes and Metabolic Syndrome Clinical Research and Reviews*, 2020, 14(4):311-315.
- [3] Jaiswal Aayush, Gianchandani Neha, Singh Dilbag, Kumar Vijay, Kaur Manjit. *Classification of the COVID-19 infected patients using DenseNet201 based deep transfer learning*[J]. *Journal of biomolecular structure & dynamics*,2020,39(15):5682-5689.
- [4] Singh D, Kumar V, Yadav V, et al. *Deep Convolutional Neural Networks based Classification model for COVID-19 Infected Patients using Chest X-ray Images*[J]. *International Journal of Pattern Recognition and Artificial Intelligence*, 2020,35(03):35-58.
- [5] Fang Chong, Chunyue Zhang, Mingben Lu. *Entropy-based TOPSIS model application in the comprehensive evaluation of water-saving transformation benefits in Youjiang Irrigation District*[J]. *Saving Water Irrigation*,2011(02):52-54+65.
- [6] Yuan Min. *Analysis and predictive study of new coronary pneumonia infection cases*[D]. *Qingdao University*,2021.
- [7] Zhang Qing. *Research and predictions on the epidemic of new coronary pneumonia*[D]. *Shandong University*,2021.
- [8] Geng Hui, Anding Xu, Xiaoyan Wang, Zhang Yong, Xiaomei Yin, Ma Mao, LV Jun. *The role of related intervention measures based on SEIR model analysis in the epidemic of new coronary pneumonia virus*[J]. *Journal of Jinan University(Natural Science and Medicine Edition)*, 2020, 41(02):175-180.
- [9] Tian Hao. *COVID-19 epidemic visualization monitoring and trend prediction research based on the improvement of the SEIR model*[D]. *Beijing Agricultural College*,2021.
- [10] JIng Yi Liang. *The establishment of the new coronary pneumonia epidemic predictive model and its optimization*[D]. *Guangzhou Medical University*,2021.
- [11] Li Peng. *Research and implementation of the new coronary epidemic information tracking and trend prediction system*[D]. *Liaoning University*,2021.
- [12] Hang Sheng, Lingling Xia. *Under the background of the outbreak of new coronary pneumonia, reviewed the Indian democratic system and government governance issues* [J]. *Indian Ocean Economy Research*,2021(06):97-113+153.