Epidemic Transmission Characteristics, Epidemic Risk Assessment, Prevention and Control in the Post-Epidemic Era: Methods and Demonstration

Hao Chang

Corresponding author: 1536761480@qq.com

School of Statistics, Capital University of Economics and Business, 121 Zhangjia Road, Huaxiang, Fengtai District, Beijing 100070

Abstract: The author selected the epidemic data of Beijing in the first half year of 2022 and studied the transmission characteristics of the epidemic through the Logistic model. It was found that the model was well fitted according to the goodness of fit, but from the calculation results of the data, there was a large deviation. Subsequently, the author used the modified SEIR model for analysis. According to the results of the current epidemic in Beijing, if the infection rate $\beta < 0.07$ and daily contact $r < 11$ can be controlled effectively, the spread of the epidemic can be controlled effectively. Then, the author selected a region for empirical analysis, and obtained information about the construction index by issuing questionnaires and field research. Subsequently, the author analyzed the above survey results through the cloud model, and the total epidemic prevention and control assessment results were between low risk and medium risk. Finally, the above research results were summarized, and conclusions and suggestions were given, including optimizing the policy duration and making dynamic adjustments according to the epidemic situation.

Keywords: The spread of the epidemic, Logistic model, SEIR model, Social research, Cloud model

1. INTRODUCTION

1.1 Research background and significance

The novel coronavirus (SARS-CoV-2) has developed a series of variants since it was detected in late 2019, with several variants, including Alpha, Delta and Omicron, becoming the main variants circulating around the world. At present, the research on mutant viruses mainly focuses on the detection and diagnosis of mutant strains, identification and analysis of genomic characteristics, influence on vaccine efficacy, epidemiological investigation and so on, but the research on its transmission mechanism is still very limited. Therefore, it is of great significance for the prevention and control of the virus in the future to understand the differences of virus transmission under different conditions and its influencing factors.

At the same time, the different characteristics of different regions lead to some differences in epidemic prevention and control work. Therefore, it is necessary to carry out relevant epidemic prevention policies flexibly and apply the medicine to the case, so as to ensure the resumption

of work and production of enterprises while strictly implementing relevant epidemic prevention policies.

1.2 Research status

In response to studies on the transsexual properties of mutated viruses, Hoyte Ashley et al. suggested that there was evidence that some of these variants were less sensitive to in vitro antibody neutralization[1]. Makoto Koizumi and others argued that when mutated viruses compete, infections with highly infectious viruses rise sharply, while infections with less infectious viruses converge[2]. According to Fan Lingling, In the context of continuous virus mutation, the protection rate of the vaccine against infection with the mutated virus is unclear at this stage[3]. Domestic scholars have carried out some studies on risk assessment under the novel coronavirus epidemic. Bi Jia et al. combined the SEIR model with risk assessment to realize the trend prediction and dynamic risk prediction of the evolution of the novel coronavirus epidemic[4]. Kong Bo et al. established an appropriate normal epidemic control system based on the key points of prevention and control, such as unit control and emergency response, and combined with the dynamic changes of the epidemic[5]. Liu Wuzhong et al, analyzed and evaluated key risk points such as centralized office and collective dormitory for employees, and put forward corresponding prevention and control suggestions[6]. He Zhihui et al. used epidemic risk index to rapidly assess epidemic risk levels in Wenzhou in different periods[7].

2. DATA SOURCE AND DESCRIPTION ANALYSIS

According to the epidemic data of Beijing in the first half of 2022, the peak of the epidemic was concentrated in May to June, and the growth was stable after June. Although the epidemic was fierce, after gaining some experience, we were able to make appropriate responses to the epidemic, which made the duration of the epidemic short and the number of infections small. It can be seen from the Figure 1. that the cumulative growth curve approximately fits the Logistic model curve, which will be fitted in the subsequent analysis, and the fitting coefficient will be observed.
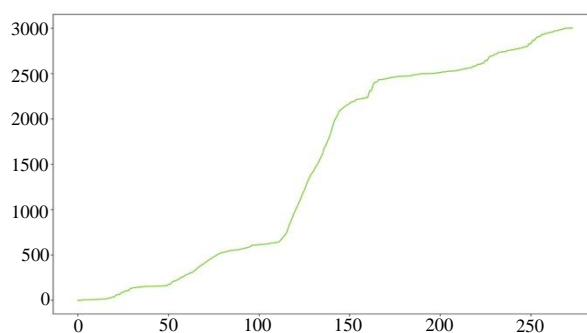

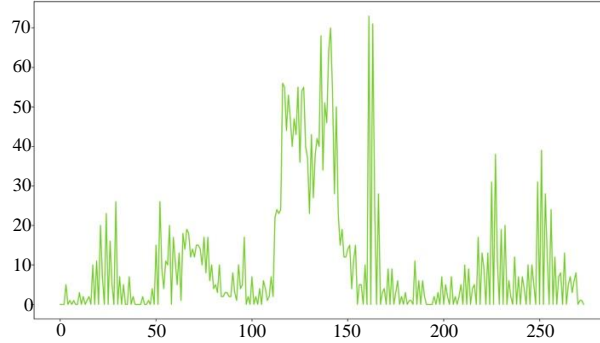**Figure 1.** Cumulative Infection Data

Figure 2. Daily New Data

According to Figure 2. the epidemic data of Beijing in the first half of 2022, the peak of the epidemic was concentrated in May to June, and the growth was stable after June. Although the epidemic was fierce, after gaining some experience, we were able to make appropriate responses to the epidemic, which made the duration of the epidemic short and the number of infections small. There were no more than 100 new cases in a single day, and the number of new cases in a single day fluctuated and was difficult to be constant, indicating the recurrence of the epidemic. During August and September, there were more new cases in this epidemic, which can also be used to judge the uncertainty of the epidemic. It can be seen that the traditional model can no longer adapt to the actual situation under multi-factor conditions. The model needs to be adjusted to fit the real problem as soon as possible.

3. LOGISTIC MODEL WAS USED TO ANALYZE THE CHARACTERISTICS OF EPIDEMIC TRANSMISSION

3.1 Regression coefficient was estimated by least square method[8]

$curve\_fit$ function in Python was used for least squares fitting of the data, and the initial results were as follows: K capacity 2816.902, P_0 initial value 18.962, and r growth rate 0.038. That is, $x_0 = 19, x_m = 2817, r_0 = 0.03892$ At the same time, in order to prevent data distortion, the author also uses the approximation principle to carry out the least squares approximation. In R 4.1.2, firstly, according to the data characteristics, the initial values are given, namely, $x_0 = 1$, $x_m = 2723, r_0 = 0.04203$. Through 20 times of least squares approximation, the least squares estimation of the three parameters can be obtained as follows: $x_0 = 18$, $x_m = 2817$, $r_0 = 0.03892$, the error is almost negligible, indicating that the fitting parameters are better, conducive to the subsequent goodness of fit test. The author will use the determination coefficient: R^2 to test. Using Python 3.10. The specific value can be obtained, which is 0.98709, indicating that the model has a good goodness of fit, See Figure 3 for details.

3.2 Final model

After plugging in the data, the final model results are as follows (1):

$$x(t) = \frac{2817}{1+(\frac{2817}{18}-1)\cdot e^{-0.03892\cdot t}} \tag{1}$$

The detailed results after regression are as follows:

The above shows that the fitting errors of the three parameters are all below 0.001, and the standard error of fitting residual is 126. The regression effect is tested as follows, as well as the actual number of infected people and the fitting situation in the figure above. Three inflection points are calculated at the same time: The initial peak period of transmission is: $t_1 = \frac{\ln\left(\frac{x_m}{x_0}-1\right)-1.317}{r_0} = \frac{\ln\left(\frac{2817}{18}-1\right)-1.317}{0.03892} = 95.9932$. The peak of transmission is: $t_0 = \frac{\ln\left(\frac{x_m}{x_0}-1\right)}{r_0} = \frac{\ln\left(\frac{2817}{18}-1\right)}{0.03892} = 129.8318$. The peak period of transmission is: $t_2 = \frac{\ln\left(\frac{x_m}{x_0}-1\right)+1.317}{r_0} = \frac{\ln\left(\frac{2817}{18}-1\right)+1.317}{0.03892} = 163.6705$. The three times numbers correspond to April 6, 2022, May 9, 2022 and June 12, 2022. The gradual increase period of the original virus was from January 1, 2022 to April 6, 2020, the rapid increase period was from April 6, 2022 to May 9, 2022, and the slow increase period was after June 12, 2022. Now the author uses the above output results and related models to get the propagation duration. It is known that when the growth rate is equal to 0, it means the end of the current epidemic. First, our team used Mathematica 9.0 to draw a transmission model image: $x(t) = \frac{2817}{1+\left(\frac{2817}{18}-1\right)\cdot e^{-0.03892\cdot t}}$. And then using Mathematica 9.0 to differentiate it. Get $v(t) = \frac{2817\cdot 0.03892\cdot\left(\frac{2817}{18}-1\right)\cdot e^{-0.03892\cdot t}}{\left[1+\left(\frac{2817}{18}-1\right)\cdot e^{-0.03892\cdot t}\right]^2}$.

```
Formula: confirmed ~ xm/(1 + (xm/x0 - 1) * exp(-r0 * date))

Parameters:
    Estimate Std. Error t value Pr(>|t|)
xm 2.817e+03  1.713e+01 164.456   <2e-16 ***
x0 1.824e+01  2.058e+00   8.866   <2e-16 ***
r0 3.892e-02  9.119e-04  42.681   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 126.3 on 271 degrees of freedom

Number of iterations to convergence: 5
Achieved convergence tolerance: 4.473e-06
```

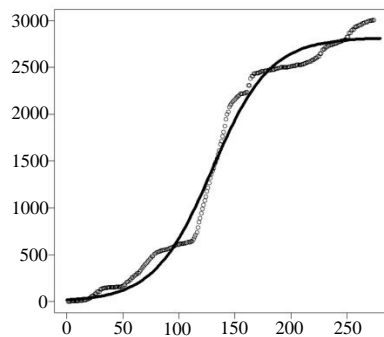**Figure 3.** Model fitting test

Figure 4. Scatter plot of actual infection and fitted cumulative population

You can see very clearly in Figure 4. tends to zero over time, and it tends to zero very much between 320 and 350, so the final transmission time of this outbreak caused by the original virus is between 320 and 350 days. And, on average time series $\bar{t} = \dfrac{\frac{(320+350)}{2}+1}{2} = 168$ substitute into $\dfrac{2817 \cdot 0.03892 \cdot \left(\frac{2817}{18}-1\right) \cdot e^{-0.03892 \cdot t}}{\left[1+\left(\frac{2817}{18}-1\right) \cdot e^{-0.03892 \cdot t}\right]^2}$ Find the propagation rate is $v(t) = 16 \ ^n/_d$, That's 16 infections a day.

3.3 Summary of this chapter

The logistic model was used to fit the infection situation, and the least squares and other nonlinear regression methods were used to fit the infection situation in detail, and the time nodes of different periods and the average transmission rate were obtained. Yet the model fits well according to goodness of fit, there are large deviations from the reality and the calculation results of data. For example, the model fitting effect is poor after 250 days. This is because the infection situation is affected by many factors, and the simple basic theoretical model can no longer adapt to complex environmental conditions, so it is necessary to optimize the parameters of the model. In the next section, the author will use the improved SEIR model for in-depth analysis and research.

4. SEIR MODEL WAS USED TO ANALYZE THE INFLUENCING FACTORS OF EPIDEMIC TRANSMISSION

Therefore, SEIR model is used for supplementary discussion. SEIR model can be used to analyze the difference of epidemic transmission in terms of different factors affecting the epidemic. However, according to the characteristics of the novel coronavirus, the traditional SEIR model should be revised first.

4.1 Model modification process

According to the differential equations of different rates of change, the specific growth model is listed, which is denoted as SEIR Model I[9] in (2) ~ (5):

$$S_{t+1} = S_t - \frac{r\beta S_t I_t}{N} \quad (2) \qquad\qquad S_{t+1} = S_t - \frac{r\beta S_t I_t}{N} - \frac{r_2 \beta_2 S_t E_t}{N} + \frac{\beta_3 S_t E_t}{N} \quad (6)$$

$$E_{t+1} = E_t + \frac{r\beta S_t I_t}{N} - \alpha E_t \quad (3) \qquad E_{t+1} = E_t + \frac{r\beta S_t I_t}{N} + \frac{r_2 \beta_2 S_t E_t}{N} - \frac{\beta_3 S_t E_t}{N} - \alpha E_t \quad (7)$$

$$I_{t+1} = I_t + \alpha E_t - \gamma I_t \quad (4) \qquad\qquad I_{t+1} = I_t + \alpha E_t - \gamma I_t \quad (8)$$

$$R_{t+1} = R_t + \gamma I_t \quad (5) \qquad\qquad\qquad R_{t+1} = R_t + \gamma I_t \quad (9)$$

Among them, r is the number of contacts of infected persons, β is the coefficient of infection, α is the probability of the incubation period, γ is recovery rate. There are major defects in the above model, mainly because it ignores that the sleeper also has transmission characteristics, but it does not take into account the possibility that the nucleic acid test result of the sleeper is negative, that is, the sleeper is divided into two categories, one nucleic acid result is positive, the other nucleic acid result is negative, which is denoted as Model II in (6) ~ (9). Among them, where r_2 is the number of contacts of the sleeper, and β_2 is the probability of infection by the

sleeper. the new parameter β_3 is the probability that the latent person is misjudged as susceptible to infection, that is, the probability that the nucleic acid result of the latent person is negative.

4.2 Model parameter assignment

During this period, the 14-day isolation policy was still adopted in Beijing, so it was estimated that the negative conversion rate of the sleeper was $\beta_3 = 1/14 = 0.071$. And by referring to the relevant literature, it is found that the probability of susceptible infection being infected by latent infection is relatively low, and the probability of susceptible infection being infected by latent infection is set here as $\beta_2 = 0.03$. In addition, since there were no human deaths as a result of this outbreak, mortality rates were not considered. It is assumed that the initial number of people infected with COVID-19 is 1, and the total population of Beijing is 30 million. However, according to the effective division of epidemic areas and the strong influence of government policies, the actual total population in the model is about 120,000 people in the built warehouse model. Based on the literature, the infection rate of direct contact with infected persons and latent persons can be estimated as $\beta = 0.2$. Therefore, the initial values of parameters in the model are assigned as follows Table 1.

Table 1. The initial value of each parameter of the model

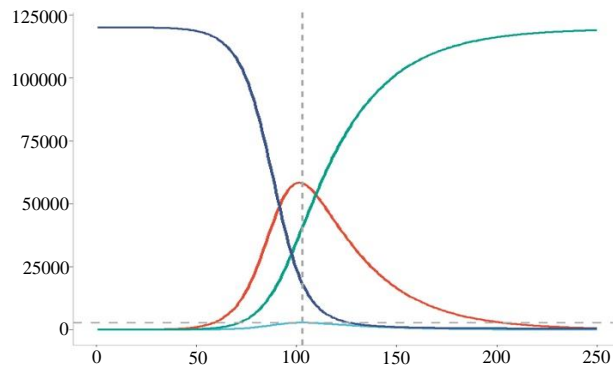| Parameter	Num	Parameter	Num
N	120000	r_2	20
I	1	β_2	0.003
E	0	β_3	0.071
R	0	k	0
D	0	γ	0.7
β	0.2	α	0.034
r	20		

Figure 5. Final fitting model

In order to study the impact of different epidemic prevention and control measures on the development of the epidemic, two parameters of daily contact rate and infection rate were selected for simulatio study. At the same time, visualization of SEIR model obtained by the

above parameters is given here. In Figure 5. the results of the fitting model are good, which can well reflect the trend of the current epidemic. The rapid increase began in 100-150 days, and the growth rate approached 0 after 250 days. According to the results of the model, the maximum number of confirmed cases was 2,833, which is similar to the fitting results of 2,817 cases in the previous chapter, indicating that the model has a good effect.
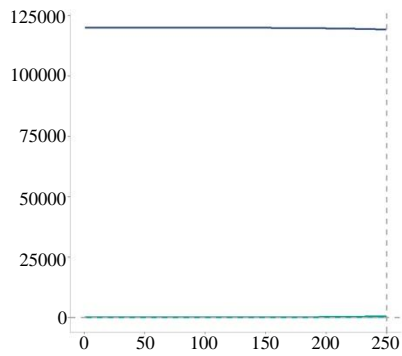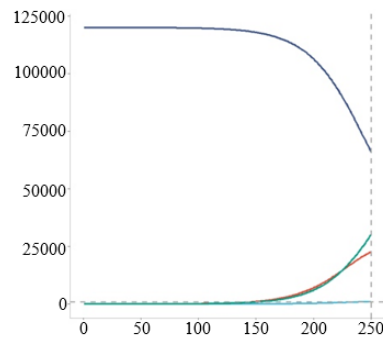

Figure 6. $r = r_2 = 10$
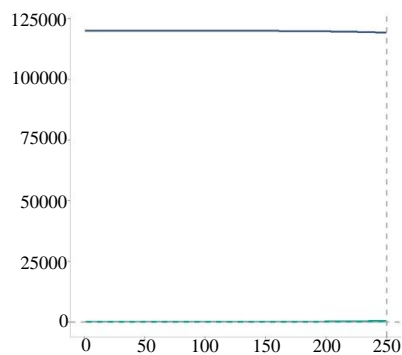

Figure 7. $r = r_2 = 12$
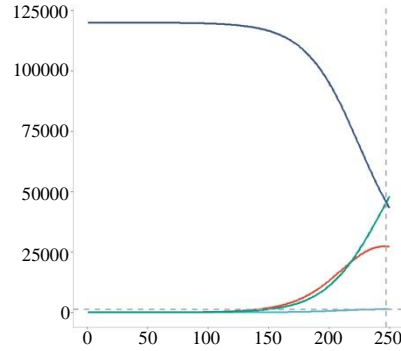

Figure 8. $\beta = 0.07$

Figure 9. $\beta = 0.1$

Based on the model given, the author conducted simulation of this epidemic. If the infection rate $\beta < 0.07$ and daily contact $r < 11$ can be controlled, the spread of the epidemic can be effectively controlled, show in Figure 6. to Figure 9.

5. ANALYSIS OF EPIDEMIC PREVENTION AND CONTROL MODEL

According to early research ideas, the first two problems to be solved in epidemic prevention and control are the daily contact rate and infection rate of infected people. The author lists the indicators based on the epidemic prevention and control situation, and gives the estimation range of the study area. α_{11}:The epidemic risk of the region, α_{12}: The proportion of people moving inside and outside the region, α_{21}: Preparation of epidemic prevention and control plans and related systems, α_{22}: Stocks of epidemic prevention materials, α_{23}: Regional personnel's knowledge of epidemic prevention and control, α_{24}: Cleaning and disinfection of key areas, α_{25}:Population density in public areas, α_{26}: Garbage disposal in public areas, β_{11}: Regional population size, β_{21}: Preparation of emergency plans for public health emergencies, β_{22}: Epidemic prevention and control exercises and assessment, β_{23}: Construction of the quarantine zone, β_{31}: Reporting rate of public health emergencies of infectious diseases, γ_{11}: The size of the city where the region is located, γ_{21}: Support rate of regional personnel for epidemic prevention and control policies in Table 2.

Table 2. Weight of each index, cloud digital characteristics, assignment interval

Index name	Weight	Characteristics	Interval	Index name	Weight	Characteristics	Interval
α_{11}	0.2280	(0.110,0.033,0.021)	(0,0.3)	β_{11}	0.2185	(0.690,0.023,0.008)	[0.5,0.75)
α_{12}	0.2250	(0.320,0.031,0.017)	[0.25,0.5)	β_{21}	0.2514	(0.310,0.030,0.009)	[0.25,0.5)
α_{21}	0.1616	(0.210,0.029,0.016)	(0,0.25)	β_{22}	0.1899	(0.290,0.033,0.013)	[0.25,0.5)
α_{22}	0.0852	(0.510,0.041,0.003)	[0.5,0.75)	β_{23}	0.1713	(0.130,0.033,0.005)	(0,0.25)
α_{23}	0.0503	(0.335,0.029,0.009)	[0.25,0.5)	β_{31}	0.1789	(0.110,0.033,0.006)	(0,0.25)
α_{24}	0.1030	(0.317,0.038,0.013)	[0.25,0.5)	γ_{11}	0.2500	(0.910,0.041,0.021)	[0.75,1)
α_{25}	0.1581	(0.821,0.041,0.017)	[0.75,1)	γ_{21}	0.7500	(0.390,0.033,0.003)	[0.25,0.5)

Index name	Weight	Characteristics	Interval	Index name	Weight	Characteristics	Interval
α_{26}	0.0278	(0.113,0.033,0.001)	(0,0.25)				

First, according to relevant literature data, the author determined the weight of each indicator by applying CRITIC method of experts. The weight of all experts is the same. The weight of possibility index, severity index and sensitivity index in the risk evaluation system of universities under the major epidemic situation is (0.4745, 0.4573, 0.0682). The weight of each indicator in the three dimensions is shown in the table. The sample mean was calculated according to xi, the titration value of n clouds[10]:$Ex = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$,Calculate the entropy of cloud droplets:$En = \sqrt{\frac{\pi}{2}} \frac{1}{n}\sum_{i=1}^{n} |x_i - Ex|$, Calculate the sample variance of cloud droplets:$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - Ex)^2$,Calculate the super-entropy of cloud droplets: $He = \sqrt{|S^2 - En^2|}$. According to the cloud digital characteristics of each index and the weight weighting operation, the evaluation cloud digital characteristics of three dimensions of possibility, severity and sensitivity are obtained. The expectation of each dimension index cloud model is:$Ex_i = \sum_{j=1}^{m} Ex_{ij}W_{ij}$ .The entropy of each dimension index cloud model is:$Ex_i = \sum_{j=1}^{m} En_{ij}W_{ij}$.The supere-ntropy of each dimension index cloud model is:$Ex_i = \sum_{j=1}^{m} He_{ij}W_{ij}$.Cloud droplet evaluation was generated by the forward cloud generator algorithm to form the cloud map for the evaluation of three dimensions of epidemic risk in colleges and universities. Find the random numbers En_i and En_{2i} respectively and find the cloud drops: $\mu_i(x) = e^{\frac{-(x_i - Ex)^2}{2(En_i)}}$. Here the number of cloud drops is set to 1000. The results are as follows:
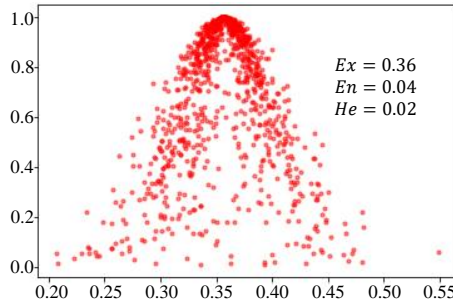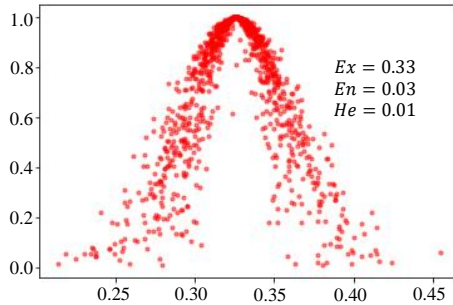

Figure 10. Probability map
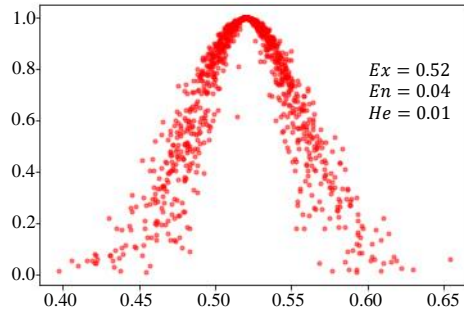

Figure 11. Severity map
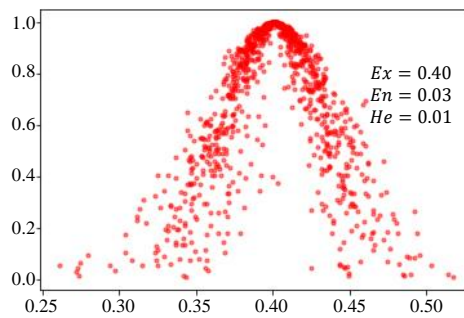
Figure 12. Sensitivity map

Figure 13. General cloud map

It can be seen that the total epidemic prevention and control assessment results are between low risk and medium risk. And there is a certain volatility, show in Figure 10. to Figure 13.

6. CONCLUSION

6.1 Research conclusion

Through data observation, it was found that the infection curve was basically consistent with the logistic growth curve, with the peak beginning, peak and peak end as the time nodes. The author simply used the logistic growth curve to predict the development trend of the epidemic. Although model fitting results are better, $R^2 = 0.98709$. However, through the observation of the actual data, it is found that there is a great uncertainty in the epidemic. In order to solve the above problems, the author added two factors into the model that the sleeper may also be infectious and the nucleic acid test result of the sleeper is negative, and obtained the SEIR model II, which can better consider the actual development of the epidemic situation. The author reproduced the final epidemic situation and simulated it. According to the results, if the infection rate $\beta < 0.07$ and the daily contact rate $r < 11$ can be effectively controlled, the spread of the epidemic can be effectively controlled. Based on this, the first two problems to be solved in epidemic prevention and control are the daily contact rate and the infection rate. According to the epidemic prevention and control situation, the indicators are listed, and the estimated range of the study area is given. Cloud models combined with weights were used for risk assessment

6.2 Research significance and future direction

Under the current situation, the epidemic situation is still not very optimistic. In view of the deficiencies in the current research on the transmission characteristics of the novel coronavirus, the author uses a variety of models to simulate the transmission characteristics, find out the relevant important transmission factors, and provide relevant assessment methods, so as to explore the transmission mechanism of the epidemic and provide reference for the formulation of reasonable policies. The future research direction is mainly to establish an appropriate mechanism research system of infectious diseases, which can be used as a reference for the study of other infectious diseases.

REFERENCES

[1]   Hoyte Ashley,Webster Mark,Ameiss Keith,Conlee Douglas A.,Hainer Nicole,Hutchinson Kendra,Burakova Yulia,Dominowski Paul J.,Baima Eric T.,King Vickie L.,Rosey Everett L.,Hardham John M.,Millership Jason,Kumar Mahesh. Experimental Veterinary SARS-CoV-2 Vaccine cross neutralization of the Delta (B.1.617.2) variant virus in cats[J]. Veterinary Microbiology,2022,268(prepublish).

[2]   Koizumi, M., Utamura, M., & Kirikami, S. (2021). Infection spread simulation technology in a mixed state of multi variant viruses. AIMS public health, 9(1), 17–25.

[3]   Fan Lingling, Liu Dan, Xu Hongbo, Wei Yongchang. Epidemic risk assessment in universities considering virus mutation and vaccine protection [J]. Journal of Wuhan University of Technology (Information and Management Engineering Edition),2022,44(04):519-525.

[4]   Bi Jia, Wang Xianmin, HU Yue, et al. A dynamic risk assessment and prediction method for public health emergencies based on improved SEIR model: a case study of COVID-19 in ten European countries [U], Journal of Global Information Science, 201,23(2):259-273.

[5]   Kong Bo, Zhao Gang. Discussion on key points and measures of railway prevention and control under the normalization of novel coronavirus pneumonia Shan China Safety Science Journal, 20,30(S1):16-20.

[6]   Liu Wuzhong, Wang Zubing, Zhang Xuetao, et al. Assessment of key risk points of novel coronavirus pneumonia in employers returning to work [U] Occupational Health and Emergency Rescue, 20,38(2):138-141

[7]   He Zhihui, Song Tie, Huang Qiong, et al. Exploration and application of rapid risk assessment method in the prevention and control of COVID-19 in cities: A case study of Wenzhou [U] South China Preventive Medicine, 20,46(2):101-105.

[8]   Mo Jiaqi, Wang Hui, Xie Feng. Perturbation solution of ecological model of epidemic disease dynamics [J]. Journal of Lanzhou University: Natural Science, 2007, 43(5):4.

[9]   Fan Ru-Guo, WANG Yi-bo, Luo Ming, ZHANG Ying-qing, ZHU Chao-ping. Transmission model and inflection point prediction of COVID-19 based on SEIR [J]. Journal of University of Electronic Science and Technology of China,2020,49(03):369-374.

[10]   Guo Qiang, Wen Weilu, Wang Yani, Qi Lian-Gang,Kaliuzhny Mykola. Basic Probability assignment generation Method based on cloud model and its application [J/OL]. Journal of Electronics and Information Technology :1-8[2022-10-20].