Exploring the spatiotemporal distribution of measles in China and predicting the level of measles risk in each province

Wang Tan, Rencong Nie, Qianqian Yang¹ and Xiang Gao

School of Mathematical Sciences, Ocean University of China

¹Email: 17806257396@163.com

Abstract. Objective To explore the spatiotemporal distribution of national measles epidemic from 1950 to 2014 and establish a model to predict the measles epidemic hot spots in the next five years. Methods Using national measles monitoring data from 1950 to 2014, a geographic information database is established. Firstly, the basic statistical analysis is done by using Matlab software. Then GeoDa 1.4.6 geographic information software is used to draw the risk level of measles in each province. Finally, the gray Verhulst model of measles risk level in each province is set up by Gtms3.0 software to predict the risk level of measles in the next five years. Results The basic statistical results indicate that there is a certain periodicity in the time dimension, and the local extremum also shows a downward trend. The provinces with the highest incidence of measles in the past 20 years are Tibet and Qinghai. The analysis of gray Verhulst model shows that the risk indices of Beijing and Chongqing respectively will reach 18.65 and 10.88 in the next five years, far exceeding the national average incidence of measles 1.96. This requires the relevant departments of these two provinces to focus more on the prevention and control. Conclusions The incidence of measles has declined overall, but it still needs to be highly regarded. Using the Verhulst model can effectively predict the risk of measles in each province, and help the relevant departments to control the measles epidemic.

1. Introduction
Measles is one of the most common acute respiratory infections in children. It is highly contagious and prone to epidemics in densely populated areas where there is no vaccine. Globally, measles is the leading cause of death in children. As a WHO target, significant reductions in measles morbidity and mortality have been achieved since the 1980s, due to the widespread use of the measles vaccine [1-4]. The coverage of measles vaccine before children’s first birthday increased from 72% in 2000 to 84% in 2011, mainly due to the improvement of daily health services [5]. As a result, Measles mortality has declined in different degrees in different years. However, the recurrence of measles cases has received increasing attention in recent years when large-scale outbreaks occurred in vaccinated people [6-8].

At the same time, the risk of measles is also geographically different. This disease is distributed throughout the world and there are significant differences between regions. The world is working together to achieve the elimination of measles and rubella. Some scholars have targeted studies on the distribution of measles cases in a certain province. For instance, Wang Wei [9] studied spatiotemporal distribution characteristics of measles in Anhui Province from 2010 to 2015, and conducted statistical analysis of spatial autocorrelation and spatiotemporal scanning to find popular areas for measles, providing a basis for the prevention and control of measles. Another example is Yu Feng's [10] analysis of the spatiotemporal characteristics of the measles epidemic in different populations in
Minhang District of Shanghai. The morbidity and relative risk in different years and different populations were calculated.

Different from previous scholars, we have innovations in research methods, contents, scope and depth. This paper selects data from measles cases in each province of the country from 1950 to 2014, and studies the temporal and spatial distribution of measles cases in the country and provinces through basic statistical methods, and analyzes the distribution characteristics by combining natural factors and social factors. At the same time, a gray Verhulst model is established to predict the measles risk index for each province in the next five years, and identify high-risk areas and provide scientific evidence for the scientific decision-making of relevant departments, thus effectively reducing or eliminating measles.

2. Materials

2.1. Data sources

2.1.1. Measles epidemic data. Data is from the national scientific data sharing platform for population and health. The research subjects are the confirmed measles cases reported by provinces from 1950 to 2014. The classifications of the years involved in the article are based on the date of occurrence. Based on the real and reliable epidemic data, this paper conducts in-depth analysis of the spatiotemporal characteristics of morbidity and mortality and future development trends.

2.1.2. Basic geographic data. Geographic information data is obtained from the State Department of surveying and mapping.

2.2. Construction and analysis of database

The provincial geographic information database and population measles information database are established. The epidemiological method is used to analyze the regional names, regional codes, years, number of cases and deaths of measles cases, and the morbidity and mortality are calculated.

3. Analysis methods and results

Firstly, we use Matlab 2017a to conduct the basic statistical analysis of the national incidence to draw the image of the change trend of the national incident, then use GeoDa1.4.6 to map the risk level of the measles in each province [11-12]. Finally, this paper conducts the spatiotemporal sequence analysis to explore the space-time clusters of the measles epidemic in the next five years [13-15], based on the grey Verhulst model.

3.1. Basic statistical analysis method

The basic statistical analysis method is a highly efficient and intuitive data analysis method with a wide range of applications. According to the indices established in the database, with the time as an independent variable and the incidence as a dependent variable, the trend image of national morbidity and mortality is plotted. We observe the line chart, analyze the outliers and trend of change, then study the reasonable year data range for subsequent processing.

Figures 1 and 2 respectively show the time trend of the incidence and mortality of measles in each province from 1950 to 2014. Through observation, it can be seen that the incidence and mortality rate of all provinces reached the peak from 1959 to 1961, which may be related to the "three years of natural disaster" at that time. The reported incidence of measles in all provinces fell to a low value from 1988 to 2014. The incidence of each province in each year is basically less than 50/100,000, but it still shows a certain periodic outbreak trend. In addition, in Tibet 1980 and 1985, the incidence of morbidity and mortality was higher than normal in the provinces, which may be related to the "wind, rain and drought" in Tibet from 1980 to 1985[16-17].
3.2. Distribution of incidence rates in each province across the country from 2005 to 2014
Since the incidence of measles varies greatly among the provinces around 2005, in order to better predict the morbidity risk in the provinces, we divide the sum of incidence data for each province and municipality from 2005 to 2014 by the national incidence, and their proportions are divided into five levels from high to low. The higher the level, that is, the greater the proportion of incidence in each province, the greater its impact on the incidence of national measles. Then we use GeoDa1.8.10 software to make the figure 3.

The figure shows that Tibet and Qinghai are ranked at the highest level, and the total incidence of measles in recent ten years is still at a relatively high level. The reason is speculated that their economic and medical level is lower than other provinces and cities in the country, and the people’s health awareness is not strong enough. The economically developed areas in Jiangsu and Zhejiang are at the second level, and the incidence of measles is not high, but the economic development does not represent a definitely low incidence of measles. The prominent examples are Beijing and Guangdong which both are prosperous, but there is still a high incidence of measles. Probably because the two provinces have a very high flow of people and mobility, and are prone to spreading infectious diseases.
3.3. Spatiotemporal sequence analysis of measles and risk assessment of each province based on gray Verhulst model

3.3.1. Gray Verhulst model introduction. Gray forecasting is to find and master the rules of system development through the processing of raw data and the establishment of a gray model, and to make a scientific quantitative prediction of the system state in the future. The GM (1, 1) model is suitable for sequences with strong exponential laws and can only describe monotonous changes. However, the incidence of measles is a dynamic time-varying system. The incidence, as a behavioral feature of epidemic infectious diseases, has a certain degree of random volatility. Its development presents a nonstationary random process with a certain trend of change. Therefore, a Verhulst prediction model for the incidence of measles in each province can be established to improve prediction accuracy. The modeling process is as follows:

Step1. The raw sequence and residual sequence: Suppose that the raw sequence composed of n raw data is \( x^{(0)}(k) = \{x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(n)\} \), adjacent terms do residual operations to generate the sequence \( y^{(0)}(k) = x^{(0)}(k-1)x^{(0)}(k-1) \). If there is a residual sequence \( y^{(0)}(t) < 0 \), an appropriate positive number should be added to the sequence so that the minimum value becomes 0. The formula is \( y_{\text{min}}^{(0)}(t) = \{y^{(0)}(1), y^{(0)}(2), \ldots, y^{(0)}(n)\} \). Add \( |y_{\text{min}}| \) to the residual sequence, making the non negative sequence \( y^{(0)}(k) = y^{(0)}(k) + |y_{\text{min}}| \).

Step2. Calculate the mean value to generate the sequence:
\[
\varepsilon^{(1)}(t) = \frac{1}{2}[y^{(1)}(k) + y^{(1)}(k-1)]
\]

Step3. Establish the gray Verhulst model:
\[
y^{(1)}(k+1) = \frac{ax^{(1)}(0)}{bx^{(1)}(0) + [a - bx^{(1)}(0)]e^{ak}}
\]

Where \( y^{(1)}(k+1) \) is accumulated generating predictive value. \( a \) and \( b \) are the pending parameters, which can be calculated by the formula:
\[
B = \begin{bmatrix}
-\varepsilon^{(2)}(2) & (\varepsilon^{(2)}(2))^2 \\
-\varepsilon^{(3)}(3) & (\varepsilon^{(3)}(3))^2 \\
-\varepsilon^{(4)}(4) & (\varepsilon^{(4)}(4))^2 \\
M & M \\
-\varepsilon^{(1)}(n) & (\varepsilon^{(1)}(n))^2
\end{bmatrix}
\]
\[
Y_n = \begin{bmatrix}
Y^{(0)}(2) \\
Y^{(0)}(3) \\
Y^{(0)}(4) \\
M \\
Y^{(0)}(n)
\end{bmatrix}
\]

Step4. Calculate theoretical predictive value which corresponds to the gray Verhulst model theoretical predictive value of raw data:
\[
y^{(0)}(k) = y^{(1)}(k) - y^{(1)}(k-1) - |y_{\text{min}}|
\]

By applying data from each province's annual morbidity rate data into the model to iterate the data, it is possible to obtain predictive values for the incidence in each province in the next five years, which is taken as the average risk index for measles in the province. An estimate of the risk of measles outbreaks in the provinces can be achieved.

3.3.2. Solution and analysis of forecasting results
3.3.2.1. National incidence forecast. With the great improvement of medical conditions in China, every child will be vaccinated against measles virus, which will greatly reduce the mortality rate and control the incidence. Using the data before 2000 to predict the prevalence of measles in the provinces in the next five years is unreasonable, so the incidence of measles from 2005 to 2014 is selected as the study object.
The gray Verhulst model is used to predict the incidence of measles in the country. And the model equation is:

$$Y = \frac{1.484}{0.11 + 0.028e^{0.1177t}} - \frac{1.484}{0.11 + 0.028e^{0.1277(t-1)}} - 10.778$$ \hspace{1cm} (5)

The relative error of the model is 37.98%. Compared with other gray prediction models, the overall relative error is smaller and performs best. The relative error rate of the model in the past two years is less than 15%, and the incidence from 2015 to 2019 is also decreasing year by year. The average incidence of measles is 1.9644, indicating that the incidence of measles is in the low period of cyclical outbreaks. The specific results are shown in the table 1:

**Table 1.** The incidence of measles in China from 2005 to 2014 and the forecast from 2015 to 2019.

| Year | Actual value | The theoretical value / predictive value | Error correction | The relative error /% |
|------|--------------|----------------------------------------|------------------|----------------------|
| 2013 | 3.9503       | 3.3853                                 | 0.565            | 14.30                |
| 2014 | 2.8589       | 3.0484                                 | 0.1895           | 6.63                 |
| 2015 | /            | /                                     | 2.4480           | /                    |
| 2016 | /            | /                                     | 2.1842           | /                    |
| 2017 | /            | /                                     | 1.9438           | /                    |
| 2018 | /            | /                                     | 1.7259           | /                    |
| 2019 | /            | /                                     | 1.5291           | /                    |

3.3.2.2. **Provincial risk assessment.** We use the measles incidence data of each province from 2005 to 2014 to predict the incidence of measles in 2015-2019, and the average value is taken as the index of disease risk in each province. The specific values are shown in the table 2:

**Table 2.** Measles Incidence Risk Index in each Province in 2015-2019.

| Province  | The disease index | Province  | The disease index | Province  | The disease index |
|-----------|-------------------|-----------|-------------------|-----------|-------------------|
| Sichuan   | 0.00678           | Fujian    | 0.45052           | Jilin     | 4.1378            |
| Shaanxi   | 0.00758           | Hunan     | 0.55522           | Guangdong| 4.24902           |
| Hainan    | 0.01512           | Liaoning  | 0.69012           | Tianjin  | 4.3361            |
| Henan     | 0.03346           | Sichuan   | 0.9507            | Hebei    | 4.41734           |
| Jiangxi   | 0.0363            | Inner Mongolia | 1.05804     | Qinghai  | 4.43622           |
| Jiangu    | 0.03848           | Anhui     | 1.67568           | Shanghai | 4.7672            |
| Shandong  | 0.273             | Tibet     | 1.69476           | Ningxia  | 5.85038           |
| Gansu     | 0.36734           | Shanxi    | 1.75224           | Xinjiang | 6.55266           |
| Heilongjiang | 0.40654 | Yunnan    | 1.85748           | Chongqing| 10.87992          |
| Hegui     | 0.42486           | Zhejiang  | 1.90362           | Beijing  | 18.64974          |
| Guizhou   | 0.44386           | /         | /                 | /        | /                 |

As can be seen directly from the above table, the incidence of measles in Beijing and Chongqing will reach 18.65 and 10.88 respectively in 2015-2019, which is far higher than the national average incidence of measles 1.96, indicating that they have the highest risk and need to take effective measures to strictly control the outbreak of measles by relevant departments. In addition, the incidence of measles in Xinjiang and Ningxia Autonomous Region is three times the national average, which requires strict prevention. In order to make the data more intuitive, the risk of morbidity index is divided into five levels. The higher the grade, the greater the probability of the onset of measles, as shown in the figure 4.

As can be seen from the figure, most provinces and cities are at the lowest level of measles incidence, indicating that the overall situation of the incidence of measles has been controlled by
comprehensive prevention and control of infectious diseases. But the incidence of measles is still at a relatively high level, so the relevant departments need to be vigilant at all times to control the increase.

Figure 4. The Incidence of Measles Risk Index Rank Map by Province in 2015-2019.

4. Conclusions
All in all, this is a study to analyze the spatial cluster of measles based on measles data from all provinces and estimate the future risks of measles in all provinces.

The spatiotemporal distribution of the measles epidemic is the basis for disease investigation and analysis. Through the analysis of the time and space of measles epidemics in each province, we can draw the following conclusions: In the time dimension, the incidence of measles nationwide shows a downward trend overall, with the characteristics of periodic outbreaks. The morbidity and mortality rates of all provinces reached the peak from 1959 to 1961, which may be related to the "three years of natural disasters" at that time. The reported incidence of measles in all provinces fell to a low value from 1988 to 2014. The incidence in each province is basically less than 50/100,000. In terms of spatial dimensions, in Tibet 1980 and 1985, the incidence of morbidity and mortality was higher than normal in the provinces, which may be related to the "wind, rain and drought" from 1980 to 1985[16].

With the increasing importance of the measles epidemic and the improvement of modern medical science, the incidence and mortality have been reduced. This paper establishes a gray Verhulst model to predict the risk of measles in each province in the next five years, identifies disease clusters that are related to the prevalence of the disease in order to provide scientific basis for prevention and treatment. According to the results, we found that the incident of measles is at the lowest level in most provinces and cities, but respectively reach 18.65 and 10.88 in Beijing and Chongqing in 2015-2019, which is far more than the national average incidence of measles 1.96. It is also 3 times higher than the national average in Xinjiang and Ningxia Autonomous Region. Accordingly, relevant departments of high-risk provinces should attach great importance to these two regions, and ensure that measles vaccination and surveillance of the epidemic are done well to control the increase in the incidence.

The limitation of this research is that we only consider the characteristics of spatiotemporal distribution, and do not analyze the different measles epidemics caused by the characteristics of the population, like the population's gender, age, and other factors. We also ignore the difference in incidence of measles between urban and rural areas. Besides, despite the special national measles surveillance program, financial support and work supervision inspection for the monitoring work, there may be cases of missing reports due to different sensitivity in different regions.

It can be seen that the epidemic of measles is caused by many factors, which are related to the distribution of time, space, population, and economic level. Eliminating measles is a common goal of global efforts. Relevant departments should attach great importance to reducing morbidity, mortality,
and disease aggregation by increasing the measles vaccination rate, developing new measles vaccines, carrying out publicity and education campaigns for measles, strengthening epidemic surveillance and improving medical standards. They should make full use of research data, and use scientific methods to guide the elimination of measles.

Acknowledgements
This paper is supported by Natural Science Foundation of Shandong (ZR2018MA006), Shandong Graduate Education Innovation Program (SDYY15129) and Shandong Graduate Tutor Mentoring Ability Enhancement Project (SDYY17009).

References
[1] World Health Organization 2013 Progress in global control and regional elimination of measles, 2000-2011. Weekly epidemiological record 88 29-36. Available: http://www.who.int/wer/en. Accessed 2013 May 6
[2] Aspects G 2012 Global measles and rubella strategic plan 2012-2020 Genetic Aspects
[3] Chao M, Lixin H, Jing M A et al 2011 Measles Epidemiological Characteristics and Progress of Measles Elimination in China, 2010 Chinese Journal of Vaccines and Immunization 3 015.
[4] Ma C, Hao L, Zhang Y et al 2014 Monitoring progress towards the elimination of measles in China: an analysis of measles surveillance data Bulletin of the World Health Organization 92(5) 340-347
[5] World Health Organization Media Centre 2013 Measles: Fact sheet 286. Geneva: WHO. Available: http://www.who.int/mediacentre/factsheets/fs286/en. Accessed 2013 May 4
[6] Sugerman D E, Barskey A E, Delea M G Ortega-Sanchez IR, Bi D, et al 2010 Measles outbreak in a highly vaccinated population, San Diego, 2008: role of the intentionally undervaccinated. Pediatrics 125 747-755
[7] Anis E, Grotto I, Moerman L, Warshavsky B, Slater PE et al 2009 Measles in a highly vaccinated society: the 2007-08 outbreak in Israel. J Infection 59 252-258
[8] Poland G A, Jacobson R M 2012 The re-emergence of measles in developed countries: Time to develop the next-generation measles vaccines? Vaccine 30 103
[9] Wang Yi, Wang Jianjun, Tang Jihai, et al 2016 Spatial and temporal distribution of measles in Anhui Province from 2010 to 2015 Chinese Journal of Vaccines and Immunization 6 627-631 (in Chinese with English abstract)
[10] Yu Feng, Ni Zheng, Zhang Liping, Du Yan, Jin Baofang and Li Shuhua 2009 Temporal and Spatial Characteristics of Measles in Different Populations in Minhang District of Shanghai Environmental and Occupational Medicine 26(03) 255-258 (in Chinese with English abstract)
[11] Wang J, Cao Z, Zeng D D et al 2014 Epidemiological analysis, detection, and comparison of space-time patterns of Beijing hand-foot-mouth disease (2008-2012) PLoS one 9(3) e92745
[12] Wu W, Guo J, Guan P et al 2011 Clusters of spatial, temporal, and space-time distribution of hemorrhagic fever with renal syndrome in Liaoning Province, Northeastern China BMC infectious diseases 11(1) 229
[13] Kayacan E, Ulutas B and Kaynak O 2010 gray system theory-based models in time series prediction Expert systems with applications 37(2) 1784-1789
[14] Sun B, He S 2011 Journal of Preventive Medicine Information 8 015
[15] Li Q, Gao Y, Cui Z et al 2016 Comprehensive Evaluation of Yantai’s Air Quality Based on gray Verhulst Model Environmental Science & Management
[16] Yang B, Kang X, Bräuning A et al 2010 A 622-year regional temperature history of southeast Tibet derived from tree rings The Holocene 20(2) 181-190
[17] Duan A, Wang M, Lei Y et al 2013 Trends in summer rainfall over China associated with the Tibetan Plateau sensible heat source during 1980-2008 Journal of Climate 26(1) 261-275