Problem Chosen	2021	Team Control Number
B	ShuWei Cup Summary Sheet	202111046836

Summary

Extreme precipitation will pose a serious threat to people's lives, safety, and property. At present, the extreme precipitation events in China have brought great losses to the local people, so it is very necessary to establish prediction models and loss quantitative analysis models of extreme precipitation events.

In task 1, the data preprocessing to reduce or eliminate the impact of outliers in the data for subsequent detection and analysis firstly.Because of the lack of several days of data in the year, directly replacing the corresponding annual total precipitation with the total precipitation in the data will result in a decrease of total precipitation and distortion. In order to reduce the impact of data distortion, we then used the Matlab to calculate the average daily precipitation of the annual data replacing the average daily precipitation of that year, with that average daily precipitation multiplying the days of that year to obtain the total precipitation of the corresponding year. Later, mutation analysis of precipitation in Zhengzhou region using Mann-Kendall (M-K) method obtained the results of mutational sites in 2021, which is consistent with the fact of Zhengzhou rainstorm in 2021. Moreover, Wavelet analysis was used to analyze the periodic relationship of precipitation in Zhengzhou, drew the contour sheet of the real part , the module and the square of module of the Wavelet coefficient in Zhengzhou, the Wavelet variance diagram and the Wavelet real process line of the characteristic time scale, and further analyzed the precipitation in 2021.

In task 2, the precipitation data of Beijing and Taiyuan from 1951-2019 were collected and compiled in National Oceanic and Atmospheric Administration (NOAA) and Daily data set of surface climate data in China. We then and calculated the average daily precipitation and the total amount per year.Later, the precipitation of Beijing and Taiyuan was analyzed by M-K method, and the periodic information of Taiyuan precipitation was analyzed by Wavelet analysis, and further analyzed the precipitation trend.

In task 3, gray prediction and Back Propagation (BP) neural network prediction were used to predict the precipitation for Zhengzhou, Beijing, and Taiyuan in 2022, 2023, with the gray prediction value of 38.04, 38.77, 24.35, 25.56, 17.08, 17.07, and the corresponding percent absolute error of 6.8832%, 3.9433% and 3.4808%, respectively. Images were also plotted. For prediction using the BP neural network, We divided the precipitation set into a training set and a test set, the latter was used to detect the generalization ability of the neural network. The relative error is obtained by dividing the absolute value of the difference between the predicted value and the real value by the real value. After taking the average value of the relative error of the predicted precipitation, the average relative error of the BP neural network prediction test set is obtained, which is 16.37%, 13.42% and 20.59% respectively, The predicted values are 23.77, 31.72, 25.64, 27.45, 17.89 and 17.02 respectively. Finally, the analysis and comparison prediction results infer that the precipitation in Zhengzhou is still high in the future. From the image, Beijing may appear extreme rainfall in the future.

In task 4, we sorted out the characteristics of heavy rainfall and flood loss of rainstorm in the inundation event of Zhengzhou and Shanxi Province, and found that the rainfall characteristics of the two places were different, which was related to environmental factors such as geographical environment and climate. At the same time, the differences in rainfall characteristics and natural environment also caused the difference of flood losses between the two places.

In task 5, we summarized the characteristics of the inundation event of Zhengzhou, and found that although the planning and construction of most cities in China could withstand extremely strong rainstorm weather to a certain extent, it had insufficient ability to withstand extremely heavy rainfall in a short period of time.Take Taking Beijing as an example, if the characteristics of Rainstorm in Zhengzhou were transferred to Beijing, serious waterlogging would occur in Beijing. In this regard, we gave the long-term construction plan for cities.

Finally, this paper evaluated and generalizes the model.

Key word: precipitation; Mann-Kendall method; Wavelet methods; gray prediction; Back Propagation neural network; Quantitative loss analysis

Content

1. Introduction

1.1 Background

In the past two years, Henan, Shaanxi, Hubei and other places in China have encountered extremely rare heavy rainfall weather. Meanwhile, some northern cities have been hit by the worst snowfall in history. These rainstorms and snowfalls have caused huge losses to the safety of local people's lives and property, and seriously affected production and life. In July, Zhengzhou, Henan province suffered a severe rainstorm. According to preliminary statistics, the province's emergency resettlement population was 395989, the affected area of crops was 44209.73 hectares, and the direct economic loss was 655 billion yuan. The floods and secondary disasters caused by the rainstorm have killed hundreds of people.

Global warming has exacerbated the instability of the climate system, and extreme weather and climate events occur frequently. In the past, "once in decades" or even "once in a century" extreme weather and climate events are becoming a new normal, seriously affecting people's production, work and life. For China, due to the large land area and the comprehensive influence of various landform and other factors, the precipitation characteristics of different cities show different characteristics. At the same time, due to the complexity of the climate system, the amount, intensity, frequency and type of precipitation in China will be directly affected. Therefore, it is imperative to establish the prediction model of cities with different potential extreme precipitation events and the quantitative analysis model of their losses.

1.2 Work

Our ultimate goal is to establish prediction models for cities with different potential extreme precipitation events and to quantitatively analyze the losses caused by rainstorms. For purpose 1, we need to collect a large number of rainfall data from other cities, analyze the annual change characteristics of the precipitation characteristics, establish at least two models, predict which cities in China will have extreme rainfall events in the future, and compare and analyze the results of different models. For the second purpose, we need to collect local loss data from major official websites, establish analysis models and conduct specific quantitative analysis.

2. Problem analysis

2.1 Data analysis

After observing the data, we found that the data of observation station 1 began in 1957, the data of observation station 2 began in 1983 and the data of observation station 3 began in 1961. There were only December data in 1957. There were many missing data from 1958 to 1964, which would have a great impact on the results. In addition, the lack of precipitation data from 1965 to 1972 would also have a certain impact on the next research. Therefore, we selected the data from 1972 to 2021 for this study. At the same time, we need to consider the authenticity of the data. If it is not real data, we need to replace or delete the abnormal data for smoothing.

2.2 Analysis of question one

Question one can be divided into the following three questions:

1) The annual average precipitation is calculated first, and then the correlation analysis can be carried out.
2) For the analysis of the annual change characteristics of precipitation characteristics in Zhengzhou, we carry out mutation analysis and periodic analysis. Firstly, M-K method is used to detect and diagnose whether there is a mutation point in the annual rainfall series, and find the mutation point to verify the fact of Rainstorm in Zhengzhou in 2021. Secondly, Wavelet methods for time series analysis is used to fully understand whether the rainfall series shows a periodic trend with time.
3) Heavy rainfall event is a complex system. Aiming at the heavy rainfall event in Zhengzhou, we focus on the disaster causing factors and disaster pregnant environment. Disaster causing factor is a rainstorm disaster risk index system established based on AHP. Disaster pregnant environment is the specific natural environment of rainstorm. Through quantitative analysis of heavy rainfall events in Zhengzhou, it provides reference for disaster prevention and reduction in urban planning in China.

2.3 Analysis of question two

In question two, we need to analyze the precipitation trend of more cities in China over the years. First, we need to have accurate and real data sources. For reference, the websites of the National Meteorological Science and mathematics center and NOAA are proposed to obtain the observed values of rainfall, maximum wind speed, sea-level pressure and visibility in Beijing and Shanxi over the years. Then draw a broken line chart using the annual rainfall to reflect the change trend, and analyze the change characteristics by M-K test and Wavelet analysis.

2.4 Analysis of question three

In question three, it is necessary to establish prediction models for cities with different potential extreme precipitation events. There are many kinds of prediction models to choose, such as Logistic, Markov, Data Mining, Deep Learning, etc., but it still needs to be combined with the rainfall data of this topic. At the same time, the acquisition and processing of data are still the focus. In view of the limitations of data volume, we chose to use Grey prediction method and BP neural network to predict the rainfall data, and compare and analyze the results to find the cities with potential extreme precipitation.

2.5 Analysis of question four

Through the National Bureau of Statistics, People's Daily Online, Chinanews.com, Zhengzhou and other official media, we collected the data of rainstorm characteristics and rainstorm losses in Zhengzhou and Shanxi Province, summarized the rainstorm characteristics of the two places, and compared and analyzed the differences of flood losses between the two places. Combined with the geographical environment and rainstorm characteristics of the two places, the causes of the difference in flood loss were briefly analyzed.

2.6 Analysis of question five

Rainstorm is not only the cumulative effect of long-lasting precipitation, but also caused by short-term heavy precipitation. China's rainstorm process, especially the catastrophic Extreme Rainstorm Process, is generally accompanied by extreme short-term heavy precipitation weather[1], such as the severe rainstorm in Jinan on July 18, 2007, Beijing on July 21, 2012, Nanchong on August 17, 2015, etc. Overall, heavy rainfall events in Zhengzhou will cause huge losses in any city. It is very necessary to carry out in-depth research on Extreme Rainstorm and extreme short-term heavy precipitation, especially the urban planning research on disastrous Extreme Rainstorm in large cities such as Beijing and Zhengzhou.

3. Model Assumptions

1. Most of the data of observation stations 1, 2 and 3 in the question can truthfully reflect the precipitation in Zhengzhou, that is, there will be no observation station failure, resulting in most of the data distortion.
2. The average value calculated from the existing non abnormal daily precipitation in the year with few missing data can represent the average daily precipitation in that year.
3. The previous precipitation can be used to predict the future precipitation.

4. Symbol description

No.	Symbol	Meaning
1	$\bar{a}$	Average daily precipitation
2	$S^{'}$	Sum of precipitation in one year based on available data
3	S	Total precipitation in a year
4	α	Significance level in M-K test
5	ψ	Basic Wavelet function
6	Var	Variance
7	x_i^*	The i prediction value of BP neural network
8	x_i	ist true value
9	r_i	Relative error of the i predicted value
10	$\bar{r}$	Average relative error

5. Model establishment and solution

5.1 Question one

5.1.1 Data processing

In order to calculate the annual precipitation in Zhengzhou, the data in the table must be processed first. According to international standards, the precipitation of 24-hour severe rainstorm is more than 7.87 inches. Therefore, it is speculated that the rainfall of 99.99 in the data is an abnormal value. Firstly, the abnormal value needs to be removed. Since observation station 1 was established in 1957, observation station 2 in 1983 and observation station 3 in 1961, the idea of replacing the data is to use the data of observation station 1 to calculate the annual average precipitation. If there is an abnormal value of 99.99, replace it with the data of observation station 3 on the same date. If observation station 3 is also an abnormal value, replace it with the data of observation station 2. For the remaining outliers after replacement, take the average value of the nearest number that is not an outlier for estimation and replacement.

Because there is a lack of precipitation for several days in individual years, the calculated precipitation will be too small, which is inconsistent with the facts. In order to reduce this impact, the existing data of the year are used to calculate the average daily precipitation, which is approximately used to replace the real average daily precipitation of the year, and then multiplied by 365 or 366 to replace the precipitation of the year. The calculation formula is:

$$\bar{a} = S'/T \#(1)$$

$$S = \bar{a} \times 365(\text{Ordinary year}),\ S = \bar{a} \times 366(\text{Leap year}) \ \#(2)\#$$

Where $\bar{a}$ represents the average daily precipitation, I.e. average daily precipitation, S' represents the total precipitation of the existing data in a year, T represents the number of data , and S represents the total precipitation of the year.(in Appendix 12)

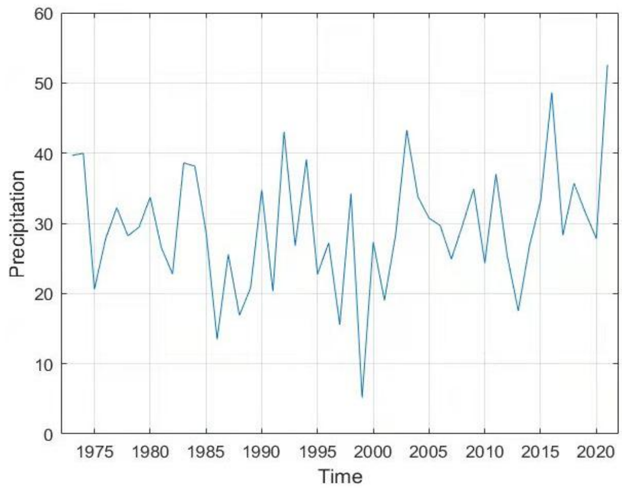

Figure 1 Annual total precipitation in Zhengzhou from 1973 to 2021

5.1.2 Establishment of M-K Model

We used catastrophe analysis and periodic analysis to study the annual change characteristics of precipitation characteristics in Zhengzhou. For mutation analysis, we used M-K test. M-K test is a climate diagnosis and prediction technology. The application of M-K test can judge whether there is a climate mutation in the climate sequence. If so, the time of the mutation can be determined. M-K test is also often used to detect the frequency trend of precipitation and drought under the influence of climate change.

5.1.2.1 Theoretical basis

For a time series x with n sample sizes, construct a one rank series:

$$s_k = \sum_{i=1}^{k} r_i \quad (k = 2,3,\ldots,\ n),\ r_i = \begin{cases} 1 & x_i > x_j \\ 0 & else \end{cases} \quad (j = 1,2,\ldots,i)$$

It can be seen that the rank series s_k is the cumulative number of values at the time i greater than the number of values at time j. Under the assumption of random independence of time series, define statistics:

$$UF_k = \frac{s_k - E(s_k)}{\sqrt{Var(s_k)}} \quad (k = 1,2,\ldots,n)$$

Where $UF1 = 0$, $E(s_k)$, $Var(s_k)$ are the mean and variance of the cumulative number s_k . When $x_1, x_2, \ldots, x_n$ are independent of each other and have the same continuous distribution, they can be calculated by the following formula:

$$E(s_k) = \frac{n(n+1)}{4}, \quad Var(s_k) = \frac{n(n-1)(2n+5)}{72}$$

UF_i is the standard normal distribution, which is the sequence of Statistics calculated according to the sequence $x_1, x_2, \ldots, x_n$ of time series x . given the significance level α , check the normal distribution table. If $|UF_i| > U_\alpha$, it indicates that there is an obvious trend change in the sequence.

Repeat the above process according to the reverse order $x_n, x_{n-1}, \ldots, x_1$ of time series x , and make

$$UB_k = -UF_k, \quad k = n, n-1, \ldots, 1, \quad UB = 0$$

To solve this problem, first determine the significance level, and then calculate the corresponding series UF and UB from the precipitation series from 1973 to 2021. See Appendix 3 for the calculation code. Then use MATLAB to draw.

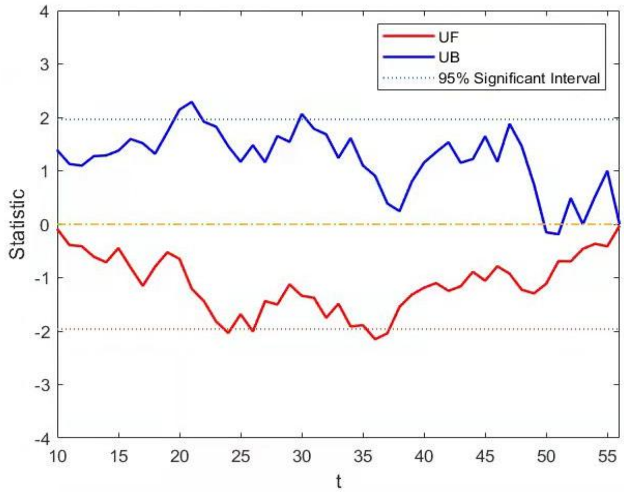

Figure 2 M-K test image in Zhengzhou from 1965 to 2021

This is the M-K test image. Abscissa 0-57 represents 1965 to 2021. In the figure 2, positive values represent growth and negative values represent decline. The middle part of the two red dotted lines represents the confidence interval, and the intersection of the two lines in the confidence interval represents the point where the rainfall may suddenly change. It can be seen from the figure that there will be intersections in 2021, that is, there may be abrupt points, which is basically consistent with the fact of precipitation in Zhengzhou in 2021.

5.1.3 Establishment of Wavelet methods for time series analysis

For periodic analysis, we use Wavelet analysis method based on time series. The change of rainstorm with time is often affected by many factors, most of which belong to non-stationary series. They not only have the characteristics of trend and periodicity, but also have randomness, mutation and "multi time scale" structure, with multi-level evolution law. For the study of this kind of non-stationary time series, it usually needs the time information corresponding to a certain frequency band or the frequency domain information of a certain period. The proposal of Wavelet analysis makes it possible to better study the problem of time series. It can clearly reveal a variety of change cycles hidden in the time series, fully reflect the change trend of the system in different time scales, and qualitatively estimate the future development trend of the system. In this paper, we use Wavelet analysis to obtain Wavelet transform coefficients, draw Wavelet coefficient real part contour map, Wavelet coefficient modulus and modulus square contour map, Wavelet variance map, and Wavelet real part process lines with different time scales.

5.1.3.1 Theoretical basis

The basic idea of Wavelet analysis is to use a cluster of Wavelet function systems to represent or approximate a signal or function. Therefore, Wavelet function is the key of Wavelet analysis. It refers to a kind of function with oscillation and can quickly decay to zero. The Wavelet function is:

$$\psi(t) \in L^2(R), \quad \int_{-\infty}^{+\infty} \psi(t)dt = 0$$

where, $\psi(t)$ is the base Wavelet function, which can form a cluster of functions through the expansion and contraction of scale and the translation on the time axis:

$$\psi_{a,b}(t) = |a|^{-1/2} \psi(\frac{t-b}{a}) \qquad (a,b \in R, a \neq 0)$$

Where, $\psi_{a,b}(t)$ is sub Wavelet; a is the scale factor, which reflects the period length of Wavelet; b is the translation factor, which reflects the translation in time.

After obtaining the sub Wavelet, for a given energy limited signal $f(t) \in L^2(R)$, it's continuous Wavelet transform is:

$$W_f(a,b) = |a|^{-1/2} \int_R f(t)\overline{\psi}(\frac{t-b}{a})dt$$

Where, $W_f(a,b)$ is the Wavelet transform coefficient; $f(t)$ is a signal or square integrable function; a is the expansion scale; b is the translation parameter, $\overline{\psi}(\frac{x-b}{a})$ is the complex conjugate function of $\psi(\frac{x-b}{a})$.

Since the observed time series data are discrete, let the function be

$$f(k\Delta t), \quad (k = 1,2,\ldots n; \quad \Delta t \text{ is the sampling interval})$$

then the discrete Wavelet transform form is:

$$W_f(a,b) = |a|^{-1/2} \Delta t \sum_{k=1}^{N} f(k\Delta t)\overline{\psi}\left(\frac{k\Delta t - b}{a}\right)$$

The low-frequency or high-frequency information of the signal is obtained by increasing or decreasing the telescopic scale a, and then the overview or details of the signal are analyzed to realize the analysis of different time scales and spatial local characteristics of the signal.

The Wavelet variance can be obtained by integrating the square value b of the Wavelet coefficient in the domain

$$Var(a) = \int_{-\infty}^{\infty} \left|W_f(a,b)\right|^2 db$$

5.1.3.2 Specific operation process

Step 1: (Eliminate the boundary effect) because the limited time series may have "boundary effect", in order to eliminate or reduce the "boundary effect", first use the "Signal Extension" in the Wavelet toolbox in MATLAB to extend both ends of the original time series (i.e. the average daily precipitation series from 1973 to 2021), 7 on the left and 8 on the right, The extended new time series is obtained.

Step 2: (Calculate Wavelet coefficients) select "Complex Continuous Wavelet 1-D" in MATLAB Wavelet toolbox, take the period as 1 and the maximum scale as 32, and save the Wavelet coefficient file after analysis. See Supporting documents.

Step 3: (Calculate the real part, module, module square and variance) directly execute the Appendix 6 code in the command line window. After running the code, the real part, modulus, modulus square and variance of Wavelet coefficients are obtained.

Step 4: (Draw contour map of real part, module and module Square). Surfer software is a software compiled by golden software company to draw three-dimensional map. The software has powerful interpolation function and drawing ability. It is a professional mapping software commonly used by geologists. Therefore, surfer is selected to draw the image, but the drawing image has certain requirements for the file format. In order to make the data to be drawn (i.e. real part, module and module) meet the requirements, Appendix 5 function is used to convert the previously calculated real part, module and The data arrangement of modular square shall be arranged according to Surfer's drawing requirements. Then it is converted into Excel file, and Surfer is used to draw the contour map of real part, module and module square.

Step 5: (Draw the Wavelet variance map and find the maximum points) use the plot function of MATLAB to draw the Wavelet variance map for the variance obtained in step 3. Observe the image and variance data to obtain four maximum points, which are 4, 9, 17 and 27 respectively, corresponding to the time scale of 4, 9, 17 and 27 years in turn.

Step 6: (Draw the main cycle trend chart) draw the main cycle trend chart by using the plot function of MATLAB based on the four time scales obtained in step 5 and the Wavelet coefficients obtained in step 2.

5.1.3.3 Result analysis

Using the obtained Wavelet coefficients, the contour map of the real part of the Wavelet coefficients is drawn.

From Fig. 3, we can clearly see the multi time scale characteristics in the evolution process of annual rainfall. The abscissa is the year, the ordinate is the time scale, and the equivalent curve in the figure is the real part of the Wavelet coefficient. The real part contour map of Wavelet coefficients can reflect the periodic changes of annual rainfall series at different time scales and their distribution in time domain, and then judge the future change trend of annual rainfall at different time scales.

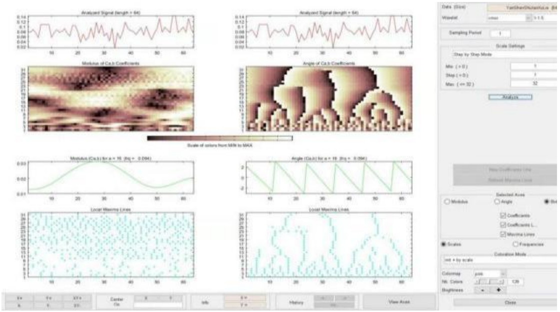

Figure 3 Wavelet analysis software process diagram

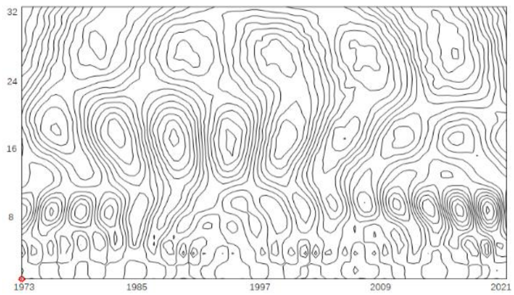

Figure 4 Contour map of real part of Wavelet coefficient in Zhengzhou

The multi time scale of annual precipitation time series means that there is no real change cycle in the evolution process of annual precipitation, but its change cycle changes accordingly with different research scales. This change is generally manifested in the change cycle of small time scale, which is often nested in the change cycle of large scale. In other words, the annual precipitation change has multi-level time scale structure and local variation characteristics in the time domain.

It can be seen from the figure that in the evolution process of annual rainfall, there are three types of periodic variation laws of 25-32 years, 15-22 years and 6-10 years. Among them, there are five quasi oscillations with alternating abundance and drought on the scale of 25-32 years; There are quasi-8 shocks on the time scale of 15-19 years. At the same time, it can be seen that the periodic changes of the above two scales are very stable and global in the whole analysis period; The periodic change on the scale of 6-10 years has an unstable period from 1982 to 1997.

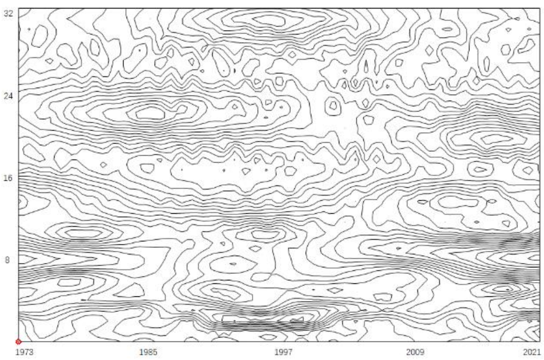

Figure 5 Contour map of Wavelet coefficient modulus in Zhengzhou

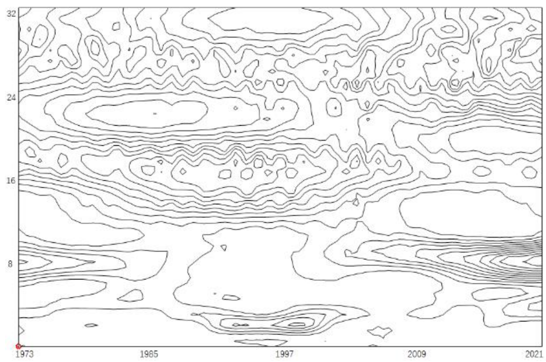

Figure 6 Contour map of modulus square of Wavelet coefficients

The modulus of Wavelet coefficient is the reflection of the distribution of energy density corresponding to the change period of different time scales in the time domain. The larger the modulus of coefficient, the stronger the periodicity of its corresponding period or scale. As can be seen from Figure 13, in the evolution process of annual rainfall, the time scale modulus of 18-26 years is the largest, indicating that the periodic change of this time scale is the most obvious, and the periodic change of other time scales is small.

The modulus square of Wavelet coefficient is equivalent to Wavelet energy spectrum, which can analyze the oscillation energy of different periods. As can be seen from Figure 14, the energy of 12-18 years time scale is the strongest and the cycle is the most significant, followed by 20-26 years; Although the energy on the time scale of 4-10 years is weak, the periodic distribution is obvious, occupying almost the whole research time domain.

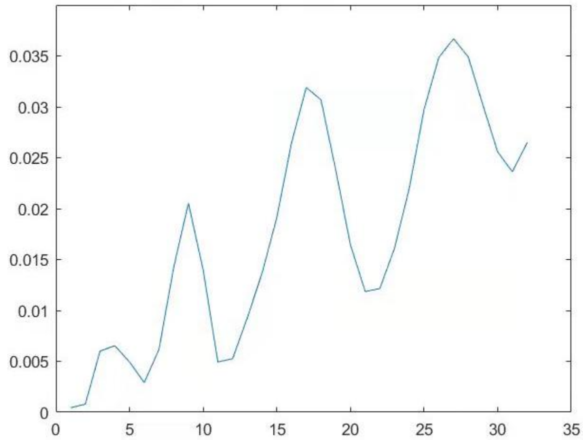

Figure 7 Wavelet variance diagram

Wavelet variance diagram can reflect the distribution of fluctuation energy with scale α of annual rainfall time series. It is used to determine the main period in the evolution of annual rainfall. As shown in the figure7, there are four obvious peaks in the wavelet variance diagram of annual rainfall, which are 4, 9, 17 and 27 respectively. Among them, the maximum peak value corresponds to the time scale of 27 years, indicating that the cycle oscillation of about 27 years is the strongest, which is the first main cycle of annual rainfall change; The 17 year time scale corresponds to the second peak, which is the second main cycle of annual rainfall change. The third and fourth peaks correspond to the time scales of 9 and 4 years respectively, which are the third and fourth main cycles of annual rainfall change in turn. This shows that the fluctuation of the above four cycles controls the variation characteristics of annual rainfall in the whole time domain.

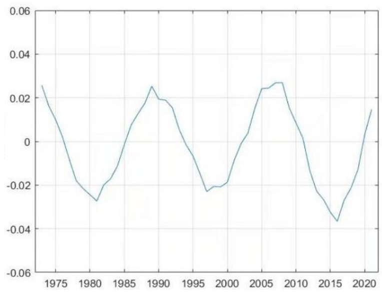

Figure 8 The 17 characteristic time scale wavelet real part hydrograph

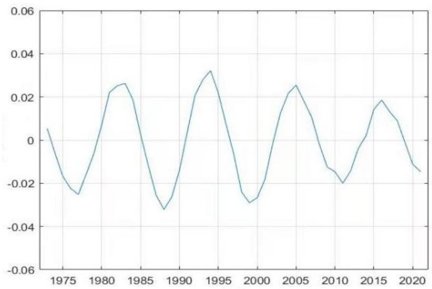

Figure 9 The 27 characteristic time scale wavelet real part hydrograph

According to the results of Wavelet variance test, we drew the Wavelet coefficient diagrams of the first and second main periods of annual rainfall evolution. From the main cycle trend chart, we can analyzed the average cycle and Runoff variation characteristics of annual rainfall under different time scales. Figure 1 shows that the average cycle of annual rainfall variation on the 10-year characteristic time scale is about 6 years, undergoing approximately four Runoff transition periods.

However, on the 21-year characteristic time scale, the average change cycle of annual rainfall is about 13 years,which has experienced about two Runoff transition periods.

5.1.4 The quantitative analysis of the inundation event of Zhengzhou riging method

Kriging method consists in estimating the grade of a panel by computing the weighted average of available samples attempting to evaluate the unknown grade of the panel with a linear estimator[4].

This is an important method of spatial analysis. Based on the variogram theory, it determines the weight value of unknown sample points through known sample points. The formula is:

$$Z(x_0) = \sum_{i=1}^{k} \lambda_i Z(x_i)$$

In this equation：$Z(x_0)$ is the meteorological value of the estimation point；λ_i is the weight value of the ith known station relative to the unknown meteorological point；$Z(x_i)$ is the value of known meteorological points[5].

Analytic Hierarchy Process（AHP）

For heavy rainfall events in Zhengzhou, the maximum daily rainstorm directly related to rainstorm, average annual rainstorm days and average annual rainstorm are selected as three indicators of disaster causing factors. Firstly, we normalize the three evaluation indexes of disaster causing factors through range standardization, and then calculate the contribution weight of each index to the target based on analytic hierarchy process. The normalized evaluation indexes are multiplied by their respective weights relative to disaster causing factors to obtain the multiple regression model of disaster causing factors. For the pregnant disaster environment, we use AHP analytic hierarchy process and weighted synthesis method to normalize the four evaluation indexes of the pregnant disaster environment through range standardization, and then multiply their respective weights relative to the pregnant disaster environment to obtain the multiple regression model of the pregnant disaster environment.

5.2 Question 2

Through the combination of the website of the National Center for meteorological science and mathematics and the website of the National Oceanic and Atmospheric Administration, we downloaded the observed values of rainfall, in Beijing and Shanxi from 1951 to 2021. Then we use the annual average rainfall to draw a broken line diagram, analyze their change trend, and use m-k test and wavelet analysis to analyze the change characteristics.

As can be seen from the figure11, the average annual rainfall in Taiyuan area is abundant-drying trend, but the range of fluctuations is small, in 1996 and 2009 rainfall is more, 1997 and 2005 rainfall is less. Rainfall has continued to rise in recent years, but further analysis of the amount of rainfall after 2021 requires further analysis of forecast models.

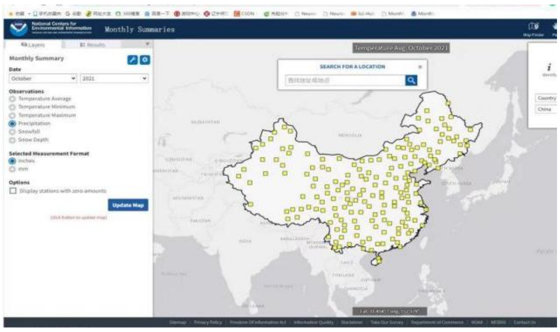

Figure 10 Data source website display

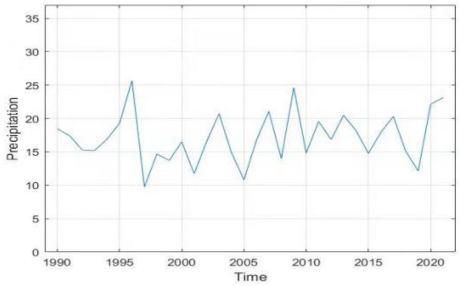

Figure 11 Broken line chart of annual rainfall from 1990 to 2021

Next, we use M-K method to tested the Average Annual Rainfall Data for the Taiyuan Region for 1951-2021, and from the graph, we can see that the main mutation point is 71 years, which is consistent with the fact that the Shanxi rainstorm in October 2021.

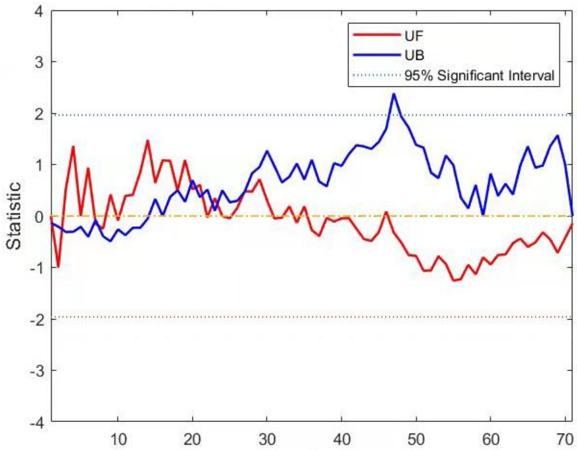

Figure 12 Taiyuan Mann Kendall test image

Finally, we carry out wavelet analysis on the average annual rainfall in Shanxi from 1990 to 2021.

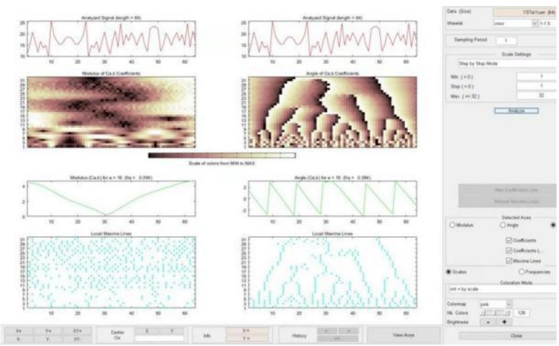

Figure 13 Taiyuan wavelet analysis software process diagram

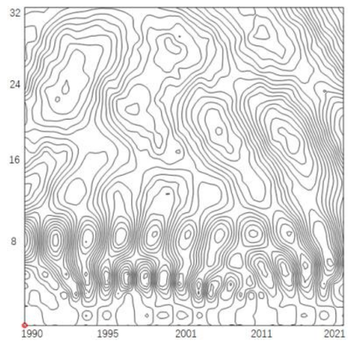

Figure 14 Contour sheet of real part of wavelet coefficient

From Figure 12, we can clearly see the multi time scale characteristics in the evolution process of annual rainfall in Taiyuan. In the evolution process of annual rainfall, there are three kinds of periodic variation laws of 16-24 years, 6-10 years and 3-5 years. Among them, there are quasi-8 oscillations with abundant dry alternation on the scale of 6-10 years; There are quasi 10 shocks on the time scale of 3-5 years. At the same time, it can be seen that the periodic change of 6-10 scale is very stable and global in the whole analysis period.

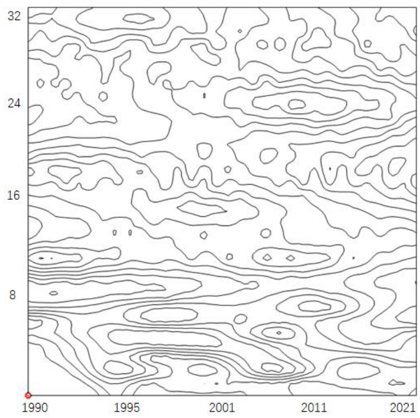

Figure 15 Wavelet coefficient modulus contour map

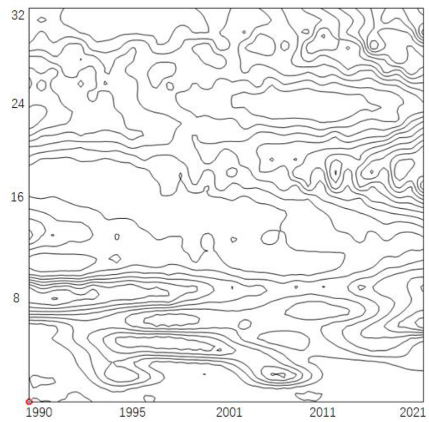

Figure 16 Wavelet coefficient modulus square contour map

It can be seen from Figure 15 that in the evolution process of annual rainfall, the modulus of time scale of 20-26 years is the largest, indicating that the periodic change of this time scale is the most obvious, and the periodic change of other time scales is small.

As can be seen from Figure 16, the energy on the 6-10-year time scale is the strongest and the cycle is the most significant, followed by 3-5; Although the energy of 16-24 year time scale is weak, the periodic distribution is obvious, occupying almost the whole research time domain.

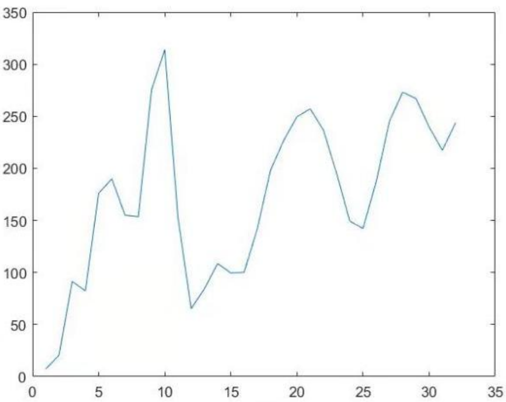

Figure 17 Wavelet variance diagram

As shown in the figure, by observing the image and combining the square wavelet difference data, the maximum points are 6, 10, 21 and 29. Among them, the maximum peak corresponds to the time scale of 10 years, indicating that the cycle oscillation of about 10 years is the strongest, which is the first main cycle of annual rainfall change; The 29 year time scale corresponds to the second peak, which is the second main cycle of annual rainfall change. The third and fourth peaks correspond to the time scales of 21 and 6 years respectively, which are the third and fourth main cycles of annual rainfall change in turn. This shows that the fluctuation of the above four cycles controls the variation characteristics of annual rainfall in the whole time domain.

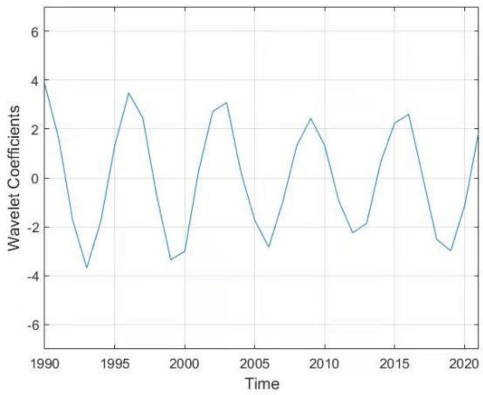

Figure 18 The 10 characteristic time scale wavelet real part hydrograph

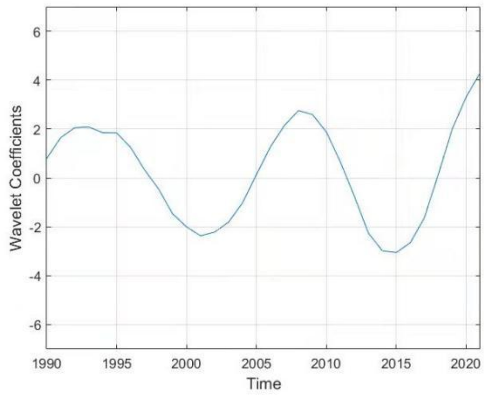

Figure 19 The 21 characteristic time scale wavelet real part hydrograph

According to the results of wavelet variance test, we draw the wavelet coefficient diagrams of the first and third main periods of annual rainfall evolution. From the main cycle trend chart, we can analyze the average cycle and Runoff change characteristics of annual rainfall under different time scales. Figure 1 shows that on the 10-year characteristic time scale, the average period of annual rainfall change is about 6 years, which has experienced about 4 Runoff transition periods; On the characteristic time scale of 21 years, the average change cycle of annual rainfall is about 13 years, with about 2 cycles of Runoff change.

Next, we analyze the rainfall trend in Beijing, and the process is as follows:

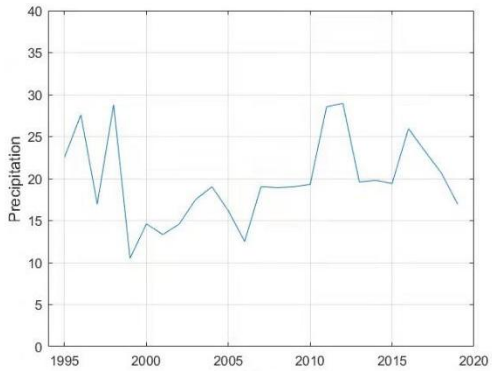

Figure 20 Broken line chart of annual rainfall from 1995 to 2019

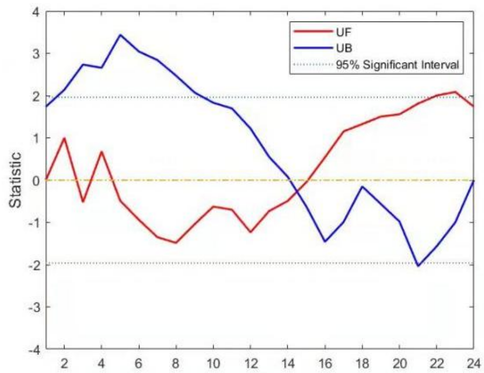

Figure 21 Mann Kendall test image of precipitation in Beijing

We conducted Mann Kendall test on the 25-year average annual rainfall data in Beijing. It can be seen from the figure that there are intersections between the 14th and 15th years, which may be abrupt points. Combined with the above figure, we believe that the annual rainfall in Beijing changes periodically with time, with a rich dry trend. Extreme rainfall occurred in 2012, resulting in the emergence of abrupt points. The emergence of extreme rainfall in Beijing is also closely related to the intensification of greenhouse effect and global warming in recent years. Extreme rainfall weather has become a potential hidden danger threatening people's production, work and life, which also reflects the necessity and urgency of establishing prediction models to accurately predict extreme rainfall events.

5.3 Question three

Different methods are used to predict the cities that may have extreme rainfall in the future. Here, we choose grey prediction and neural network to predict whether extreme rainfall will occur in Zhengzhou, Taiyuan and Beijing in two years, and compare and analyze the results.

5.3.1 Grey prediction

Grey prediction model is a prediction method to establish a mathematical model and make prediction through a small amount of incomplete information [6]. By calculating the correlation degree between various factors, the difference degree of development trend among various factors of the system is identified. Its core system is grey model (GM), in which GM (1,1) model is the most commonly used grey model.

5.3.1.1 Model establishment

Record the original sequence as：

$$X_0 = \left(x_0(1), x_0(2), \ldots, x_0(n)\right)$$

One time accumulation generation sequence (1-AGO) is：

$$X_1 = \left(x_1(1), x_1(2), \ldots, x_1(n)\right)$$

among：

$$x_0(k) = \sum^{k} x_0(i) = x_1(k-1) + x_0\,, \quad (k = 2,3,\ldots,n)$$

Let Z_1 be the nearest mean generating sequence of X_1

$$Z_1 = \left(z_1(2), z_1(2), \ldots, z_1(n)\right)$$

It meets the following relationships：

$$z_1(k) = \frac{1}{2}\left(x_1(k) + x_1(k-1)\right), \quad (k = 1,2,\ldots,n)$$

We get the basic form of GM (1,1) model：

$$x_0(k) + a z_1(k) = b$$

Where a is the development coefficient, $z_1(k)$ is the whitening background value, and b is the amount of ash.

List parameters as：

$$u = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} x_0(2) \\ x_0(3) \\ \cdots \\ x_0(n) \end{bmatrix}, \quad B = \begin{bmatrix} -z_1(2) & 1 \\ -z_1(3) & 1 \\ \cdots & \\ -z_1(n) & 1 \end{bmatrix}$$

At this time, the model is transformed into $Bu = Y$, and the least squares estimation parameter column of GM (1,1) model $x_0(k) + a z_1(k) = b$ satisfies:

$$u = \left(B^T B\right)^{-1} B^T Y$$

Thus, we get the whitening equation of the grey equation:：

$$\frac{dx_1(t)}{dt}+ax_1(t)=b$$

The above formula satisfies the equation $x_1(t_0)=x_1(1)$, Eventually we got:

$$x_1(t)=\left(x_0(1)-\frac{b}{a}\right)e^{-a(t-1)}+\frac{a}{b}$$

In this paper, the matlab function that can automatically carry out grey prediction by inputting precipitation series and days to be predicted in MATLAB has been written. in Appendix 9. The results can be obtained by prediction by this program.

5.3.1.2 Result analysis

We use grey prediction to predict the rainfall data of Zhengzhou in 2022 and 2023 according to the rainfall data of Zhengzhou from 1995 to 2021, and the following results are obtained.

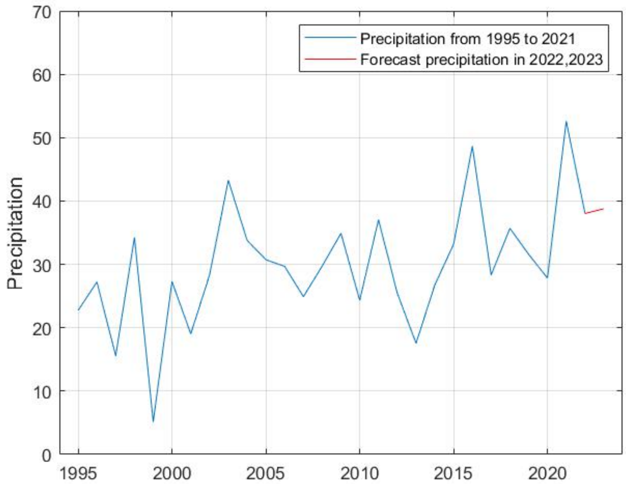

Figure 22 Forecast of precipitation in Zhengzhou in the next two years

Table 2 Grey prediction output results

Year	2022	2023
Estimate	38.0354	38.7665
Percent absolute error	6.8832%	

The red line in the figure is the prediction curve. From the grey prediction results, it can be seen that there is little possibility of extreme precipitation in Zhengzhou in the next two years, but the rainfall still tends to rise. At the same time, although the rainfall value is lower than that in 2021 and 2016, it is significantly higher than that in Zhengzhou six years ago. It can also be seen that the climate in Zhengzhou has been greatly affected by global warming in recent years. Therefore, relevant departments need to improve their awareness of flood control and waterlogging prevention in order to better control extreme weather and climate and minimize the risk.

Next, we use grey prediction to predict the rainfall data of Taiyuan in 2022 and 2023 according to the rainfall data of Taiyuan from 1990 to 2021 in Shanxi Province, and the following results are obtained.

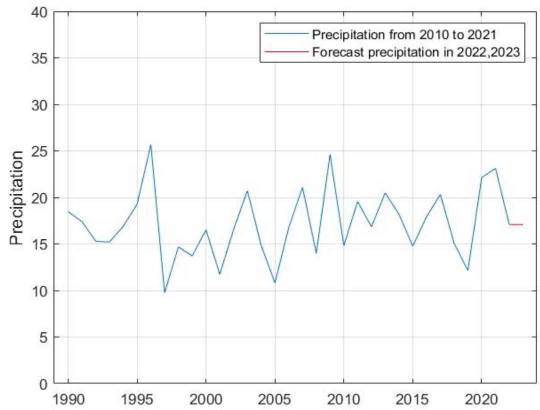

Figure 23 Forecast of precipitation in Taiyuan in the next two years

Table 3 Grey prediction output results

Year	2022	2023
Estimate	17.0831	17.0709
Percent absolute error	3.4808%	

From the grey prediction results, it can be seen that the rainfall trend of Taiyuan in the next two years is gentle, the change is small, and there is no extreme rainfall event. From this point of view, the climate is suitable.

Finally, we use grey prediction to predict the rainfall data of Beijing in 2020, 2021, 2022 and 2023 according to the rainfall data of Beijing from 1995 to 2019, and the following results are obtained.

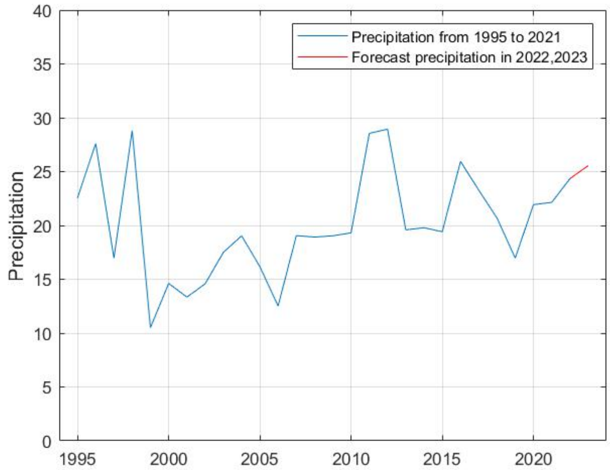

Figure 24 Forecast of precipitation in Beijing in the next two years

Table 4 Grey prediction output results

Year	2020	2021	2022	2023
Estimate	21.9348	22.1401	22.1401	22.5564
Percent absolute error	3.9433%			

It can be seen from the grey prediction results that the rainfall in Beijing will continue to rise in the next two years, and the rainfall in 2023 is close to that in 2016. Considering the possibility of extreme precipitation, relevant departments should take more precautions and deploy the emergency treatment of extreme precipitation in advance.

5.3.2 BP neural network

5.3.2.1 Basic principles

BP network is called back propagation neural network. Through the training of sample data, the network weight and threshold are continuously modified to make the error function decline along the negative gradient direction and approach the expected output. It is a widely used neural network model, which is mostly used for function approximation, model recognition and classification, time series prediction and so on [7]. The information processing units are generally divided into three categories: input unit, output unit and implicit unit, that is, the input unit receives external signals and data, the output unit realizes the output of system processing results, the implicit unit is between the input and output units, and the structure of the implicit unit cannot be observed from outside the network system. In addition to the above three information processing units, the connection strength between neurons is determined by parameters such as weight.

5.3.2.2 Specific operation

The relevant precipitation data is divided into training set and test set. The neural network correlation function of MATLAB is used for training with the training set. The input layer is the rainfall of the previous four years, represented by x_1, x_2, x_3, x_4, the number of hidden layers is set to 9, represented by $\mathrm{k}_1, \cdots, \mathrm{k}_9$, and there is only one output layer, that is, the value of rainfall to be predicted, as shown in Figure.

The test set is used to quantify the generalization ability of the neural network, and the average relative error is calculated. The calculation formula of the average relative error is

$$r_i = \frac{|x_i^* - x_\mathrm{i}|}{x_\mathrm{i}}, (i = 1,2) \quad \bar{r} = \frac{\sum_{i=1}^{2} r_i}{2}$$

Where x_i^* represents the ith predicted value, x_i represents its corresponding real value, r_i represents the relative error of the ith predicted value, and $\bar{r}$ represents the average value of each relative error, that is, the average relative error.

If the average relative error is large, the number of hidden stratification points and training times are re adjusted, and the predicted values of precipitation in Zhengzhou, Beijing and Taiyuan in 2022，2023 with small average relative error in the training set are obtained. In this paper, the matlab code for training neural network and calculating the average relative error has been written. See Appendix 10.

Finally, the images of precipitation and predicted precipitation in previous years are drawn by MATLAB.

5.3.2.3 Result analysis

From the results of BP neural network, it can be seen that the rainfall in Zhengzhou in the next two years is small, so there is no need to worry about the emergence of extreme precipitation.

Next, we train the rainfall data of Taiyuan to predict the rainfall in 2022 and 2023, and ensure that the relative error of the neural network training set is 20.59%.

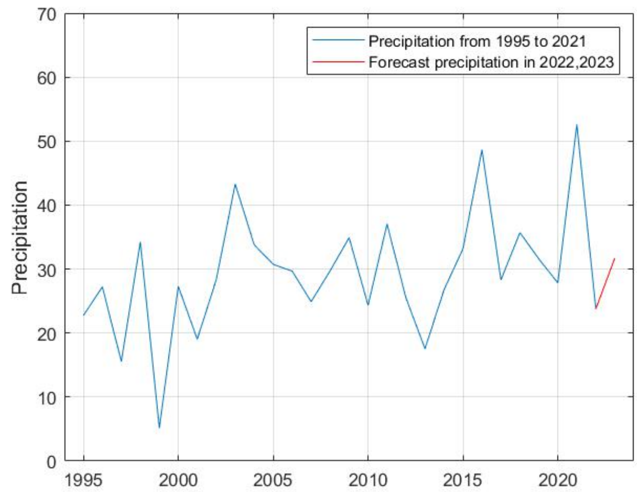

Figure 25 Forecast of precipitation in Zhengzhou in the next two years

Table 5 Neural network output results

Year	2022	2023
Estimate	23.7673	31.7225
Average relative error	16.37%	

As can be seen from the figure26, the rainfall in Taiyuan will show a downward trend in the next two years, so there is no need to worry about extreme weather.

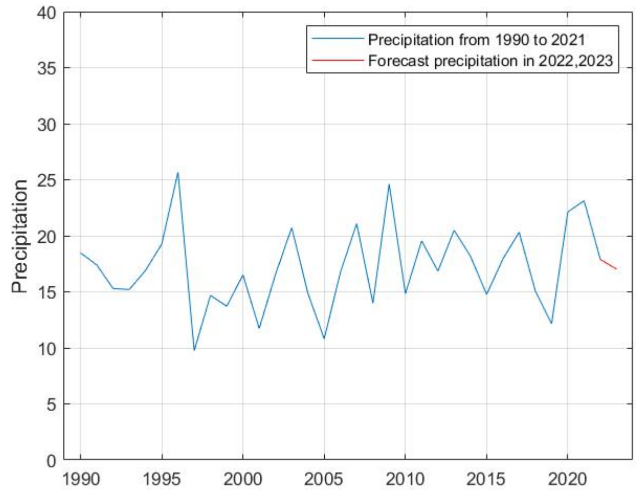

Figure 26 Forecast of precipitation in Taiyuan in the next two years

Table 6 Neural network output results

Year	2022	2023
Estimate	17.8884	17.0213
Average relative error	20.59%	

Finally, we train the rainfall data in Beijing to predict the rainfall in 2020, 2021, 2022 and 2023, and ensure that the relative error of the neural network training set is 13.42%.

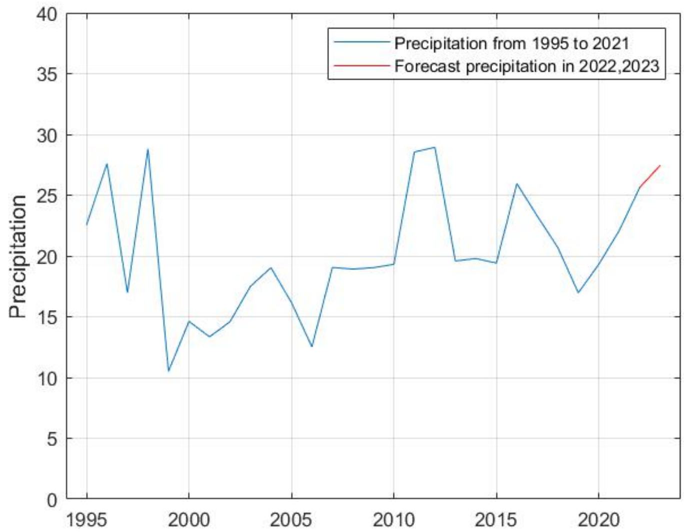

Figure 27 Forecast of precipitation in Beijing in the next two years

Table 7 Output results of BP neural network

Year	2020	2021	2022	2023
Estimate	19.3152	22.0904	25.6442	27.4550
Average relative error	13.42%			

It can be seen from the prediction results of BP neural network that the rainfall in Beijing will continue to rise in the next two years, and the rainfall in 2023 will reach 27.4550, which can be regarded as an extreme rainfall event. From the broken line chart, it can be seen that although the average annual rainfall in 2023 is less than the "721" rainstorm in Beijing in 2012, it is also the highest rainfall in the past decade, which needs considerable attention.

According to the statistics of the National Climate Center, since the autumn of 2021, the rainfall in the north of China has obviously exceeded that in the south, and there has been a phenomenon of "waterlogging in the north and drought in the South". While the rainfall in China has increased significantly on a large scale, the rainfall distribution has been uneven. According to the data, the rainfall in the north is 1.4 times more than that in the same period of the year, It is the most in the same period since the complete meteorological records were made in 1961. This also provides evidence for the 2023 extreme rainfall event in Beijing.

5.3.3 comparative analysis of prediction results

We have established two models of grey prediction and BP neural network to predict the rainfall in Zhengzhou, Taiyuan and Beijing in the next two years. From the results, we can see that the prediction results of the two models for the three regions are basically not different. Within the allowable range of error, we can see that firstly, for the change trend of rainfall, the two prediction results are the same. Zhengzhou has a continuous upward trend, but there is no extreme rainfall; The rainfall in Taiyuan has a downward trend and the speed is relatively gentle; In contrast, there is a great potential for extreme rainfall in Beijing in 2023. The grey prediction results are close to 2016 and the BP neural network results are close to 2012. These are the years when extreme rainfall occurs in Beijing. Therefore, it can be seen from the comparison that it is urgent to establish an extreme precipitation prevention and control plan for Beijing.

5.4 Question four

5.4.1 Overview of the Heavy Rainfall Characteristics in Zhengzhou and Shanxi

In the first three questions, we have analyzed the mutational sites and the periodic information of precipitation in Zhengzhou and Taiyuan, Shanxi Province, and made certain predictions for the future precipitation. We knew that the precipitation in Zhengzhou in 2021 was the precipitation mutational site by M-K method. In fact, since July 2021, there had been continuous rainstorms in Henan Province, from the central and western regions to the northwest. Among them, there had been a severe rainstorm in Zhengzhou. From 4 p.m. to 5 p.m. on July 20 alone, the rainfall in Zhengzhou had reached 201.9 mm, which meant that the rainfall in this hour in Zhengzhou had exceeded the average rainfall of one month in previous years. This extreme rainfall was also verifying this conclusion. At the same time, we also knew that the precipitation in Taiyuan, Shanxi in 2021 was also a mutational site. Moreover, the heavy rainfall in Shanxi was from October 2 to October 7, with an average precipitation of 119 mm. Therefore, it was not difficult to find the difference between the characteristics of heavy rainfall in Zhengzhou and Shanxi. Although it was a mutational site in 2021, the heavy rainfall in Zhengzhou was concentrated, the accumulated rainfall was large and extreme, while the heavy rainfall in Shanxi was widely distributed and lasted for a long time.

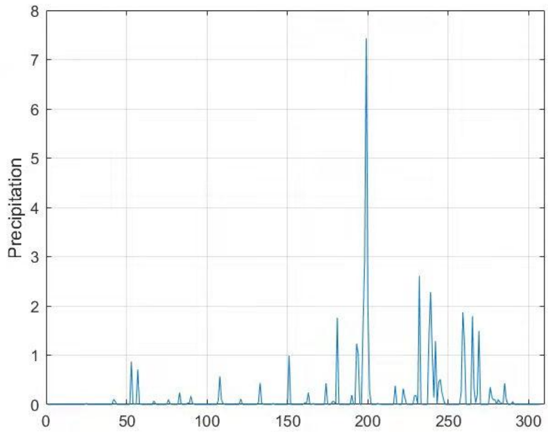

Figure 28 Line plot of precipitation change in Zhengzhou in 2021

5.4.2 Overview of the Flood Losses in Zhengzhou and Shanxi Province

According to the press conference on flood control and disaster relief of Henan provincial government, in this disaster, the transportation, communication, Internet, water supply, power supply, gas supply, garbage treatment, sewage treatment, drainage, flood control and other systems in the city were damaged at key points or surfaces, resulting in major losses, including vehicle flooding, subway water inflow, building water inflow, facility water inflow, bridge and road damage, hospital closure and transfer, ground traffic shutdown and so on, Affect the normal operation of the city and the normal life of the people. The scope of disaster is far beyond the scope of waterlogging, and the indirect loss even exceeds the direct loss [8].

According to news reports, 37 rivers in Shanxi Province were flooded, affecting the operation of expressways, national and provincial trunk lines and railways. Crevasses occurred in many places, such as Xinjiang section of Fenhe River, Fenyang section and Xiaoyi section of Ciyao river. The abutment of Changyuan River Bridge in Qi county of South Tongpu line was washed out, and the sleepers were suspended, resulting in the shutdown of the train. There were 33 geological disasters with collapse and landslide as the main types, resulting in economic losses of 51.249 million yuan, which generally showed three characteristics of "high frequency, wide distribution and large losses".

Table8 Comparison of the losses situation in two places

LOCATION	Zhengzhou	Shanxi
flood-hit population	1.736million	1.7571million
municipal administration path	2718	37
trunk highway	1183	72
country road	6181	3238
collapsed building	52500	17000
Damaged crop	930.67km^2	2384.6km^2
direct economic loss	53200 million YUAN	5029 million YUAN

5.4.3 Cause of Difference

Using the quantitative analysis method of the inundation event of Zhengzhou, this paper analyzed the inundation event of Shanxi, we initially derived the different causes of the characteristics of heavy rainfall events in the two places.

Zhengzhou is located in the climate transition zone between North and south of China and the landform transition zone from mountains, hills to plains. The complex weather system, terrain and urban effects may have a significant impact on the precipitation in Zhengzhou[1]. The high-frequency in the western mountainous area of Zhengzhou and the extremely strong short-term heavy rainfall in the main urban area on the afternoon of the 20th far exceeded the urban water storage capacity, which are the main reasons for serious mountain torrents and urban waterlogging disasters in Zhengzhou.

Shanxi Province is located in the middle reaches of the Yellow River and the east of the Loess Plateau. It belongs to the monsoon climate area of the middle temperate zone and warm temperate zone, that is, the temperate continental climate.Due to the influence of atmospheric circulation and topography, the precipitation distribution in Shanxi Province is more in the east than in the west, more in the South than in the north, more in mountains than in basins, more in windward slope than leeward slope, with large seasonal changes and high precipitation intensity. The terrain of Shanxi Province is a parallelogram inclined from northeast to southwest. It is a typical mountain plateau covered with loess. The terrain is high in Northeast and low in southwest. The plateau is rugged with vertical and horizontal river valleys. The landform includes mountains, hills, platforms and plains. The mountainous area accounts for 80.1% of the total area. Due to the landform and long-term coal mining, the geological disasters in the region are high and prone[9].

5.5 Question five

The inundation event of Zhengzhou in 2021 has great reference significance for the prevention and control of waterlogging in other cities. Take the Beijing municipal problem below as an example. Although China's urban water conservancy projects and waterlogging prevention and control work can resist the heavy rainfall weather rarely encountered in decades, it is still difficult to deal with the situation of rapid rainfall and great rainfall in a short time[10].

According to the prediction results of our second question, we find that there is a possibility of severe rainstorm in Beijing in 2023. According to the existing research findings, the main problems of urban waterlogging prevention and control in Beijing are: lack of large drainage channel and storage space, unclear data of rainwater pipe network, failure to give full play to the regulation and storage function of pipe network, some breakpoints or water blocking points of drainage and waterlogging prevention still need to be opened, and the monitoring, identification and intelligent management system of drainage and waterlogging prevention need to be further strengthened, The source emission reduction system of rainfall runoff is not perfect[8].

Combined with the rainstorm in Zhengzhou, we give the following suggestions: 1)strengthen the forecast and early warning of disastrous weather, formulate corresponding work plans according to different weather conditions, and strive for time for flood control; 2)strengthen the joint operation of water conservancy engineering facilities and municipal waterlogging drainage facilities at the river basin, regional and urban levels, and give overall consideration to flood control and internal waterlogging; 3)establish a high-level flood prevention and control management system . During the flood in Zhengzhou, non-governmental rescue organizations used online documents to collect and sort out the information of rescuers and volunteers. These documents played a key role in the flood and saved a lot of time for the rescue work, which is called "life-saving documents". This enlightens us that we can use 5G, big data and other information means to realize flood and disaster data sharing between relevant professional operation units at the municipal and district levels.

6. The assessment and promotion of model

6.1 Advantages

1)In the specific quantitative analysis of the inundation event, we use analytic hierarchy process. AHP is more systematic, practical and concise. The application of AHP to establish the hierarchy of all elements can more clearly show the relationship of various elements, which is convenient for us to directly understand and master the flood event in Zhengzhou.

2)On the prediction of potential extreme precipitation, we establish grey prediction model and BP neural network model. The grey prediction model can predict the approximate results even if the amount of data is relatively small; BP neural network model is the most commonly used neural network in nonlinear prediction. It has a wide range of adaptability, learning ability and mapping ability, and has a high degree of fault tolerance, which provides us with a tool for accurate prediction.

6.2 Disadvantages

1)Due to the time limit of this study, we can not find the rainfall data of more cities and years in China, and make a more accurate and comprehensive prediction of the rainfall in the future, which also makes the research results imperfect. At the same time, we also hope to have the opportunity to conduct more in-depth research on this.

2)In the analysis of extreme precipitation events in Beijing, there is a lack of data in 2020 and 2021. Therefore, the precipitation in these two years can only be predicted through grey prediction and neural network, and then the precipitation in 2022 and 2023. Although the application of the two models can draw a conclusion that there will be potential extreme precipitation events in Beijing in 2023, more accurate rainfall data can be obtained if the data are perfect.

6.3Promotion

At the application area level, the M-K method, Wavelet methods, grey prediction, BP neural network and other methods used in this paper can analyze and predict the precipitation trend all over the country outside Zhengzhou, Beijing and Taiyuan. In particular, the matlab function file of M-K method and grey prediction has been written, and only the precipitation sequence of the city to be analyzed and predicted needs to be input, The drawing and prediction can be carried out automatically by running the function. BP neural network only needs to modify the code slightly and arrange the sequence according to a certain format. It can also predict the precipitation in other areas.In the application field, the tide, stock market trend and disease transmission can also be analyzed and predicted with reference to the method of this paper.

7. Conclusion

The climate system is a complex system, which is affected by many factors. In recent years, the greenhouse effect has intensified, and extreme weather and climate occur frequently. Since 2020, the novel coronavirus pneumonia has been raging in a certain degree, which has a certain impact on our country's livelihood. However, the epidemic is still unreported and the rainstorm is coming again. This year, Zhengzhou, Shanxi and other places are attacked by rainstorms. Therefore, it is of great practical significance to master the objective law of rainfall, accurately grasp the rainfall characteristics, quantitatively analyze the rainfall loss, and establish a model to predict extreme rainfall weather.

In this paper, we processed the rainfall data of Zhengzhou and collect the rainfall data of Taiyuan and Beijing; AHP is used to quantitatively analyze rainstorm events, and M-K method and Wavelet analysis are used to accurately grasp the variation characteristics of rainstorm; Analyze and compare the disaster data of Rainstorm in Zhengzhou and Shanxi, and analyze the difference of flood loss; Grey prediction and BP neural network model are established to predict the rainfall data of three regions in the next two years, and important conclusions are drawn: there are potential extreme rainfall events in Beijing in 2023; Finally, measures and plans are put forward.

References

[1] Aifang Su, Xiaona Lv, Liman Cui. Prediction and test of optimal integrated precipitation based on similar spatial distribution of precipitation, Torrential Rain and Disasters, 40(5): 445-454, 2021.

[2] Qi Hu, Xueqing Ma. Application of Matlab in Meteorological Teaching, Research and Exploration in Laboratory, 38(12):48-51, 2019.

[3] Ping Mu, Shengfa Yang, Sichen Tong. Periodical Analysis on Annual Runoff in Jialing River Basin Based on the Morlet Wavelet Analysis, Journal of Chong Qing Jiao Tong University. 35(1):76-79,2019.

[4]Matheron, G. Principles of geostatistics, Economic geology, 58(8):1246—1266, 1963.

[5]Sha Wang, Xiaofeng Wang, Xianchao Huang. Rainstorm Disaster Risk Assessment Based on GIS in Qinling Region, HENAN SCIENCE, 33(9):1647-1652, 2015.

[6]Sifeng Liu. Emergence and Development of Grey System Theory and Its Forward Trends, Journal of Nanjing University of Aeronautics, 36(2):267-272, 2004.

[7]Juanjuan Lu, Chen Hong. Researching Development on BP Neural Networks, Control Engineering of China, 13(5):449-456, 2006.

[8]Shuhan Zhang, Fandong Zheng, Suchuang Di. Thoughts on urban waterlogging control in Beijing from the rainstorm and flood of "2021. 7.20" in Zhengzhou, China Flood & Drought Management, 31(9):5-11, 2021.

[9]Xinkai Shen, Yiqing Lu, Jing Zhang. Risk Assessment of Rainstorm Disaster in Shanxi Province Based on ArcGIS, Scientific and Technological Management of Land and Resources, 37(1):61-73, 2020.

[10]Yuanyuan Liu, Jingwei Zheng. Warning and thoughts of extreme rainstorm on urban flood prevention, China Flood & Drought Management, 39(9):21-24, 2021.

Appendix

Appendix 1

```
% This is matlab code,.
%It is used to handle outliers
% There are three variables in the work area required for operation, namely Observation station1, Observation station2 and Observation station3. The first column is the date, and the second column is the cell matrix of precipitation on that date. The dates and precipitation of observation stations 1, 2 and 3 are stored respectively. The way to obtain them is to import excel into Matlab, select the corresponding column and import them into the cell matrix.
% Firstly, detect the value when the precipitation of observation station 1 is 99 . 99, and replace it with the value of observation station 3 (because there are more data of observation station 3 than 2). If observation station 3 is still 99.99 or not, replace it with the value of observation station 2.
for i = 1:19447
    if GuanCeZhan1{i,2}==99.99
        for j = 1:18665
            % Check whether observation station 3 has this date.
```

```
            %If there is such a date, the data will be replaced
            if GuanCeZhan3{j,1}==GuanCeZhan1{i,1}
              GuanCeZhan1{i,2}=GuanCeZhan3{j,2};
            end
        end
    end
end
```

%For the data of observation station 1 replaced by observation station 3, use the same method to replace observation station 1 with the data of observation station 2 (because it may still be an abnormal value after replacing the abnormal value with 3)

```
for i = 1:19447
    if GuanCeZhan1{i,2}==99.99
        for j = 1:13973
            %Check whether observation station 2 has this date. If so, replace the data
            if GuanCeZhan2{j,1}==GuanCeZhan1{i,1}
                GuanCeZhan1{i,2}=GuanCeZhan2{j,2};
            end
        end
    end
end
```

%For the data that cannot be replaced (such as the data not available at the other two observation stations in front of observation station 1), use the average value of the nearest rainfall to replace it

```
for i = 1:19447
    if GuanCeZhan1{i,2}==99.99
        for j=1:min(i-1,19447-i) %In order to keep the array from crossing the boundary, find the nearest number in the maximum searchable radius that is not 99.99

            if (GuanCeZhan1{i+j,2}~= 99.99)&&(GuanCeZhan1{i-j,2}~= 99.99)
                GuanCeZhan1{i,2}=(GuanCeZhan1{i-j,2}+GuanCeZhan1{i+j,2})/2;
            end
        end
    end
end
%Check whether 99.99 still exists. If a appears in the variable, it still exists
for i = 1:19447
    if GuanCeZhan1{i,2}==99.99
        a=1;
    end
end
```

Appendix 2

```
%This is matlab code.
%It can calculate the annual average precipitation
```

```
% It needs to run the code in Appendix 1 before it can run
%After converting the double quoted string in the first column to a single quoted string (which can index the period elements), extract the year and finally convert it to a number
%The first column becomes a single quote string (in order to extract the year, double quotes cannot be indexed, so they cannot)
ZuiZhongGuanCeZhan1=cell(19447,2);
for i = 1:19447
    ZuiZhongGuanCeZhan1{i,1} = char(JGGuanCeZhan1{i,1});
end
%Extraction year
for i = 1:19447
    ZuiZhongGuanCeZhan1{i,1} = ZuiZhongGuanCeZhan1{i,1}(1:4);
end
%Convert year to number
for i = 1:19447
    ZuiZhongGuanCeZhan1{i,1} = str2num(ZuiZhongGuanCeZhan1{i,1});
end
%Make up the precipitation in the second column and convert it into a digital matrix as a whole
ZuiZhongGuanCeZhan12=zeros(19447,2);
for i = 1:19447
    ZuiZhongGuanCeZhan12(i,1) = ZuiZhongGuanCeZhan1{i,1};
    ZuiZhongGuanCeZhan12(i,2) = JGGuanCeZhan1{i,2};
end
%Calculate the average annual precipitation
NianZong = zeros(19447,2); %The total annual precipitation, but due to the lack of individual days, it will be very small in individual years, so it will be more convincing to use the average precipitation
NianPingJun = zeros(19447,2); %Average annual precipitation
NianZong(1,2) = ZuiZhongGuanCeZhan12(1,2);
j = 1; %j is used to count how many times it is added every year
k = 1; %k is used to put the number of years of change on the next line, starting with the first line
for i = 1:19447-1%-1 is because 19447 will be added to 19446
    MuQianNianFen = ZuiZhongGuanCeZhan12(i,1); %What year is it now, the current year
    XiaYiGeNianFen = ZuiZhongGuanCeZhan12(i+1,1); %What year is to be added, the next year
    if MuQianNianFen ==    XiaYiGeNianFen
    %If the years are the same, the number of rows remains the same, and the data of the next year is added
        NianZong(k,2) = NianZong(k,2)+ZuiZhongGuanCeZhan12(i+1,2);
        j = j+1; %How many numbers have been added to the calculation
    else
        %If the years are different, cut down one line and add the data
        NianPingJun(k,2) = NianZong(k,2)/j;
        %If the years are different, start calculating the previous average
```

```
        j = 1; %Initialize j
        k=k+1;
        NianZong(k,2) = NianZong(k,2)+ZuiZhongGuanCeZhan12(i+1,2);
    end
    %Finally, the average precipitation in 2021 is calculated
     NianPingJun(k,2) = NianZong(k,2)/j;
end
nianfen = [1957:1964,1973:2021]';
NianZong(1:57,1) = nianfen;
NianPingJun(1:57,1) = nianfen;
NianPingJun = NianPingJun(1:57,:);
```

Appendix 3

```
%This is matlab code.
%It is used to do the m-k method on the sequence.
%The input is time series, and the output is vectors and images of UF and UB.
function [UFk,UBk] = MKTest(x)
n=length(x);
%Calculate UF
r=zeros(n,1);
for i=2:n
    for j=1:i
        if x(i)>x(j)
            r(i)=r(i)+1;
        end
    end
end
sk=zeros(n,1);
for i=2:n
    for j=1:i
        sk(i)=sk(i)+r(j);
    end
end
UFk=zeros(n,1); % Define statistics UFk=(sk-E)/sqrt(var)
for i=2:n
    E=i*(i-1)/4; % Sk(i)的 mean value
    Var=i*(i-1)*(2*i+5)/72; % Sk(i) variance
    UFk(i)=(sk(i)-E)/sqrt(Var);
end
%Calculate UK
y2=flipud(x);
r2=zeros(n,1);
for i=2:n
```

```
    for j=1:i
        if y2(i)>y2(j)
            r2(i)=r2(i)+1;
        end
    end
end
sk2=zeros(n,1);
for i=2:n
    for j=1:i
        sk2(i)=sk2(i)+r2(j);
    end
end
UFk2=zeros(n,1);
for i=2:n
    E=i*(i-1)/4; % Sk(i) mean value
    Var=i*(i-1)*(2*i+5)/72; % Sk(i) variance
    UFk2(i)=(sk2(i)-E)/sqrt(Var);
end
UBk=zeros(n,1);
UBk=flipud(-UFk2);
%Next, draw
plot(1:n,UFk,'r-','linewidth',1.5);
hold on;
plot(1:n,UBk,'b-','linewidth',1.5);
plot(1:n,1.96*ones(n,1),':','linewidth',1);
plot(1:n,-1.96*ones(n,1),':','linewidth',1);
plot(1:n,0*ones(n,1),'-.','linewidth',1);
axis([1,n,-4,8]);
legend('UF','UB','95% Significant interval ');
xlabel('t','FontName','TimesNewRoman','FontSize',12);
ylabel('Statistic','FontName','TimesNewRoman','Fontsize',12);
end
```

Appendix 4

```
%This is matlab code.
%It can plot the average daily precipitation of each year.
%Before running this code, you need to run codes 1 and 2
t = [1973:2021];
y = NianPingJun(9:1:57,2)';
plot(t,y);
axis([1972,2022,0,0.2]); %Set the abscissa range and ordinate range
grid on
xlabel('Time','FontName','TimesNewRoman','FontSize',12); %Set X-axis label
```

```
ylabel(' Average daily precipitation ','FontName','TimesNewRoman','Fontsize',12);% %Set Y-axis label
```

Appendix 5

```
%This is matlab code.
%Wavelet analysis result data is a function for preparing and processing in order to draw with surferfunction
B = JGCL(A)
B2 = reshape(A,[32*49,1]);
H = [1:32];
H2 =
[H,H,H,H,H,H,H,H,H,H,H,H,H,H,H,H,H,H,H,H,H,H,H,H,H,H,H,H,H,H,H,H,H,H,H,H,H,H,H,H,
H,H,H,H,H,H,H,H,H]';
B1 = [H2,B2];
B = zeros(32*49,3);
B(:,2:3) = B1;
k = 0;
for i = 1:49
    for j = 1:32
        B((i-1)*32+j,1) = 1973+k;
    end
    k = k+1;
end
end
```

Appendix 6

```
%This is matlab code.
%It can calculate the relevant data of wavelet analysis
%Run the function that needs to save code 5 first, and save the coefs obtained by using the wavelet processing toolbox
shibu=real(coefs);
mo=abs(coefs);
mofang=(mo).^2;
fangcha=sum(abs(coefs).^2,2);
CLHshibu = JGCL(shibu); %The processed real part is used for drawing
CLHmo = JGCL(mo);
CLHmofang = JGCL(mofang);
```

Appendix 7

```
%This is matlab code.
%Code 6 needs to be run before it can run successfully
```

```
%27 time scale wavelet real part hydrograph of precipitation change in Zhengzhou
y1 = shibu(27,:);
x = 1973:2021;
plot(x,y1);
axis([1972,2022,-0.06,0.06]); %Set the abscissa range and ordinate range
grid on
xlabel('Time','FontName','TimesNewRoman','FontSize',12); %Set X-axis label
ylabel(' Wavelet coefficients ','FontName','TimesNewRoman','Fontsize',12); %Set Y-axis label

%17 time scale wavelet real part hydrograph of precipitation change in Zhengzhou
y2 = shibu(17,:);
x = 1973:2021;
plot(x,y2);
axis([1972,2022,-0.06,0.06]); %Set the abscissa range and ordinate range
grid on
xlabel('Year','FontName','TimesNewRoman','FontSize',12);% %Set X-axis label
ylabel('小波系数','FontName','TimesNewRoman','Fontsize',12);% %Set Y-axis label
```

Appendix 8

```
%This is matlab code.
%This program can draw the broken line chart of daily precipitation in Zhengzhou in 2021, and
can be run directly
zhengzhou2021 =
[0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0.0200000000000000;0;0;0;0;0;0;0;0;0;0;0;0;0;
0;0.100000000000000;0.0600000000000000;0;0;0;0;0;0;0;0;0.870000000000000;0.020000000
0000000;0;0.0100000000000000;0.710000000000000;0.0200000000000000;0;0;0;0;0;0;0;0.07
00000000000000;0.0100000000000000;0;0;0;0;0;0;0.100000000000000;0.0200000000000000;
0;0;0;0;0.0100000000000000;0.240000000000000;0.0100000000000000;0;0;0;0.0300000000000
000;0.0200000000000000;0.170000000000000;0.0200000000000000;0;0;0;0;0;0;0;0;0;0;0;0;
0;0.0700000000000000;0.570000000000000;0.0900000000000000;0.0200000000000000;0;0;0;0;
0;0;0;0;0;0;0.110000000000000;0;0;0;0;0;0;0;0;0;0.0400000000000000;0.440000000000000;0.
0300000000000000;0;0;0;0;0;0;0.0200000000000000;0;0;0;0;0;0;0;0;0.990000000000000;0;0;
0;0;0;0;0;0;0;0.0400000000000000;0.0300000000000000;0.240000000000000;0;0;0.0200000000
000000;0;0;0;0;0;0;0.430000000000000;0;0;0;0.0600000000000000;0.0600000000000000;0;1.
76000000000000;0;0;0;0;0;0;0;0.190000000000000;0;0;1.23000000000000;1.02000000000000;
0;0;1.61000000000000;2.91000000000000;7.43000000000000;2;0.310000000000000;0;0;0;0;0.0
200000000000000;0.0100000000000000;0;0;0;0;0;0;0;0;0.380000000000000;0.010000000000
0000;0;0;0;0.320000000000000;0.140000000000000;0.0100000000000000;0;0;0;0;0.180000000
000000;0.180000000000000;0;2.61000000000000;0.0100000000000000;0;0;0;0;1.400000000000
00;2.28000000000000;1.06000000000000;0.140000000000000;1.29000000000000;0;0.45000000
0000000;0.510000000000000;0.240000000000000;0.100000000000000;0;0;0;0;0;0;0;0;0;0;0.250
000000000000;1.87000000000000;1.28000000000000;0;0;0;0;1.79000000000000;0.3100000000
00000;0.0300000000000000;0.200000000000000;1.49000000000000;0;0;0;0;0;0;0.35000000000
```

```
0000;0.150000000000000;0.0900000000000000;0.100000000000000;0;0.100000000000000;0.04
00000000000000;0.0300000000000000;0;0.430000000000000;0.100000000000000;0;0;0;0.0600
000000000000;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0];
x = [1:307];
plot(x,zhengzhou2021);
axis([0,310,0,8]);
grid on
xlabel('Time','FontName','TimesNewRoman','FontSize',12);
ylabel('Precipitation','FontName','TimesNewRoman','Fontsize',12);
```

Appendix 9

```
%This is matlab code.
%It can be used for grey prediction
y=input(' Please enter data ');
n=length(y);
yy=ones(n,1);
yy(1)=y(1);
for i=2:n
    yy(i)=yy(i-1)+y(i)
end
B=ones(n-1,2);
for i=1:(n-1)
    B(i,1)=-(yy(i)+yy(i+1))/2;
    B(i,2)=1;
end
BT=B';
for j=1:(n-1)
    YN(j)=y(j+1);
end
YN=YN';
A=inv(BT*B)*BT*YN;
a=A(1);
u=A(2);
t=u/a;
t_test=input(' Enter the number to be predicted ');
i=1:t_test+n;
yys(i+1)=(y(1)-t).*exp(-a.*i)+t;
yys(1)=y(1);
for j=n+t_test:-1:2
    ys(j)=yys(j)-yys(j-1);
end
x=1:n;
xs=2:n+t_test;
```

```
yn=ys(2:n+t_test);
plot(x,y,'^r',xs,yn,'*-b');
det=0;
for i=2:n
    det=det+abs(yn(i)-y(i));
end
det=det/(n-1);
disp([' Percent absolute error is： ',num2str(det),'%']);
    disp([' The predicted value is:',num2str(ys(n+1:n+t_test))]);
```

Appendix 10

```
%This is matlab code.
%Enter a two column matrix. The first column is the year and the second column is the daily average precipitation. This function can convert it into a two column matrix. The first column is the year and the second column is the total precipitation of the year
function [x,y] = PJtoZ(A)
y = zeros(size(A,1),1)
for i =1:size(A,1)
    Ly = A(i,1);
switch true
    case mod(Ly,400)==0
        y(i) =   A(i,2)*366;
    case mod(Ly,4)==0 & mod(Ly,400)~=0
        y(i) =   A(i,2)*366;
    case mod(Ly,4)>0
        y(i) =   A(i,2)*365;
    otherwise
end
end
x = A(:,1);
end
```

Appendix 10

```
%This is matlab code.
%This code uses the precipitation of Zhengzhou to train the neural network. The previous four years are input and the fifth year is output for training, and calculates the relative error and average relative error of the test set.
%It can run directly, but the neural network trained in the text is in the supporting materials. Running this code may not get the same neural network.
%The neural network prediction of Beijing and Taiyuan is similar, which will not be repeated here, but Beijing needs to continue to predict with the predicted value
P=[18.4724409400000,17.3976378000000,15.2952755900000,15.2007874000000,16.905511810
```

```
0000,19.2598425200000,25.6692913400000,9.75590551200000,14.6811023600000,13.71259843
00000,16.5078740200000,11.7322834600000,16.5393700800000,20.7204724400000,14.8503937
000000,10.8149606300000,16.7244094500000,21.0787401600000,13.9881889800000,24.610236
2200000,14.8267716500000,19.5511811000000,16.8425196900000,20.4826608400000
17.3976378000000,15.2952755900000,15.2007874000000,16.9055118100000,19.259842520000
0,25.6692913400000,9.75590551200000,14.6811023600000,13.7125984300000,16.50787402000
00,11.7322834600000,16.5393700800000,20.7204724400000,14.8503937000000,10.8149606300
000,16.7244094500000,21.0787401600000,13.9881889800000,24.6102362200000,14.826771650
0000,19.5511811000000,16.8425196900000,20.4826608400000,18.1863242900000
15.2952755900000,15.2007874000000,16.9055118100000,19.2598425200000,25.669291340000
0,9.75590551200000,14.6811023600000,13.7125984300000,16.5078740200000,11.73228346000
00,16.5393700800000,20.7204724400000,14.8503937000000,10.8149606300000,16.7244094500
000,21.0787401600000,13.9881889800000,24.6102362200000,14.8267716500000,19.551181100
0000,16.8425196900000,20.4826608400000,18.1863242900000,14.7643022000000
15.2007874000000,16.9055118100000,19.2598425200000,25.6692913400000,9.7559055120000
0,14.6811023600000,13.7125984300000,16.5078740200000,11.7322834600000,16.53937008000
00,20.7204724400000,14.8503937000000,10.8149606300000,16.7244094500000,21.0787401600
000,13.9881889800000,24.6102362200000,14.8267716500000,19.5511811000000,16.842519690
0000,20.4826608400000,18.1863242900000,14.7643022000000,17.9263312000000];
T =
[16.9055118100000,19.2598425200000,25.6692913400000,9.75590551200000,14.681102360000
0,13.7125984300000,16.5078740200000,11.7322834600000,16.5393700800000,20.72047244000
00,14.8503937000000,10.8149606300000,16.7244094500000,21.0787401600000,13.9881889800
000,24.6102362200000,14.8267716500000,19.5511811000000,16.8425196900000,20.482660840
0000,18.1863242900000,14.7643022000000,17.9263312000000,20.3133106900000];
PC =
[18.1863242900000,14.7643022000000,17.9263312000000,20.3133106900000,15.06769604000
00
14.7643022000000,17.9263312000000,20.3133106900000,15.0676960400000,12.146032730000
0
17.9263312000000,20.3133106900000,15.0676960400000,12.1460327300000,22.133858270000
0
20.3133106900000,15.0676960400000,12.1460327300000,22.1338582700000,23.125984250000
0];
ZZ =
[15.0676960400000,12.1460327300000,22.1338582700000,23.1259842500000,17.08310000000
00];
net = newff(minmax(P),[9,1],{'tansig','purelin'},'trainlm');
net.trainParam.show=50;
net.trainParam.lr=0.05;
net.trainParam.epochs=300;
net.trainParam.goal=0.001;
[net,tr]=train(net,P,T);
net.iw{1,1};
```

```
net.b{1};
net.lw{2,1};
%Calculate the deviation and relative error of each column of the test set
yc = sim(net,PC);
wc = ZZ'-yc;
%Calculate the error, WC represents the error, and calculate the deviation between the predicted value and the actual value
wcl = zeros(1,size(wc,2)) %Create relative error vector
for j = 1:size(wc,2)
    wcl(j) = abs(wc(j)/ZZ(j))
end
pj=mean(wcl); %Calculate average relative error
DR =[12.1460327300000
22.1338582700000
23.1259842500000
17.0831000000000];
%Bring in data for prediction
yc2022=sim(net,DR);
yc2023=sim(net,[22.1338582700000,23.1259842500000,17.0831,yc2022]');
```

Appendix 11

```
%This is matlab code. It can run directly
%The code can draw the broken line map of precipitation over the years and the broken line map of BP neural network prediction in Taiyuan, and display them in different colors. The practices of Zhengzhou and Beijing are similar and will not be repeated here
%The rendering of grey prediction image is similar
hold off
y = [18.47244094
17.3976378
15.29527559
15.2007874
16.90551181
19.25984252
25.66929134
9.755905512
14.68110236
13.71259843
16.50787402
11.73228346
16.53937008
20.72047244
14.8503937
10.81496063
```

```
16.72440945
21.07874016
13.98818898
24.61023622
14.82677165
19.5511811
16.84251969
20.48266084
18.18632429
14.7643022
17.9263312
20.31331069
15.06769604
12.14603273
22.13385827
23.12598425
17.0831]';
plot([1990:2022],y)
axis([1989,2024,0,40]); %Set the abscissa range and ordinate range
grid on
xlabel('Time','FontName','TimesNewRoman','FontSize',12); %Set X-axis label
ylabel('Precipitation','FontName','TimesNewRoman','Fontsize',12); %Set Y-axis label
hold on
yy = [17.0831,17.0709];
xx = [2022,2023];
plot(xx,yy,'r');
legend('Precipitation from 1990 to 2021','Forecast precipitation in 2022,2023');
```

Appendix 12

Table 1 Summary of annual total precipitation from 1957 to 2021

Year	S	Year	S	Year	S	Year	S	Year	S
1957	1.659091075	1977	32.22000015	1989	20.79999987	2001	19.02999989	2013	17.54000018
1958	40.44051007	1978	28.21999982	1990	34.76000011	2002	28.31758257	2014	26.74999992
1959	20.77023813	1979	29.45000018	1991	20.32000005	2003	43.27999983	2015	33.18000015
1960	69.723	1980	33.71210957	1992	43.07000014	2004	33.78999986	2016	48.64000005
1961	13.90775852	1981	26.54000009	1993	26.81000008	2005	30.73000007	2017	28.31000006
196	30.0150	1982	22.78000	1994	39.08999	2006	29.67999	2018	35.68777

2	0011		004		985		982		463
1963	32.41499994	1983	38.61000003	1995	22.73000015	2007	24.89999997	2019	31.57000012
1964	45.45000005	1984	38.1399999	1996	27.22999996	2008	29.72999993	2020	27.82999987
1973	39.67870872	1985	28.85000004	1997	15.53999983	2009	34.89999987	2021	52.59804578
1974	39.96999989	1986	13.48999996	1998	34.23379129	2010	24.32999991		
1975	20.60645606	1987	25.55999991	1999	5.144093555	2011	37.03999998		
1976	27.76000017	1988	16.89999986	2000	27.29000017	2012	25.44999995		