Summary

The La Nina event will cause drought and flood disasters in some parts of the world. It has lasted for two years and may continue until the end of this year or beyond. It is particularly important to establish models to evaluate the area affected and extent of La Nina influence and predict the possibility of triple La Nina occurrence.

For the first question, by analyzing the characteristics of La Nina phenomenon, the potential areas affected by La Nina are located, and the temperature, precipitation and disaster data of these areas in the past two years are statistically analyzed, so as to obtain the countries and regions directly related to the impact of La Nina. Then, according to the relationship between the **Southern Oscillation Index(SOI)** and La Nina, a neural network prediction model based on **Long Short Term Memory Module(LSTM)** is established. The model parameters are determined by multi-layer grid search algorithm. The weight matrix is optimized by Adam algorithm to obtain the SOI prediction data for the next ten years.

For the second question, the **AHP-fuzzy comprehensive evaluation system** is established to evaluate the losses caused by the high temperature and drought in the United States under the triple La Nina conditions by using various aspects of American industry data. In the three-tier evaluation system, the first-tier indexes include annual hydropower production, wildfire area, summer mortality of one million people, and agricultural production. Using the weight obtained by AHP method, the fuzzy comprehensive evaluation is carried out every three years, and it is concluded that the year from 2020 to 2022 under the effect of triple La Nina is the year with the most serious loss (0.262). And then through various indicators to put forward the solution to the situation.

For the third question, we sorted out and counted the flood loss indexes in China from 2011 to 2022, including infrastructure loss, population loss, economic loss, agricultural loss and eight sub-indexes, and then built a **flood loss evaluation model** based on coefficient of variation method and fuzzy comprehensive evaluation method. The **coefficient of variation method** was used to determine the weight of each index. The flood disaster level under the influence of triple La Nina events was determined by **fuzzy comprehensive evaluation method**, and the strategies to deal with triple La Nina events were given.

For the fourth question, we combined the results of the first three questions and put forward some targeted suggestions to the relevant management departments and the analysis report of the results we got.

Finally, we carried out **sensitivity analysis** and advantages and disadvantages evaluation of the established model. The results show that the model we built can well solve the problem of damage assessment under the influence of La Nina events, and has high stability.

Key word: Triple La Nina , LSTM , Disaster loss assessment model , Multi-layer grid search

Content

1. Introduction

1.1 Background

Climate disaster is one of the major natural disasters faced by mankind. Droughts, floods and temperature anomalies caused by extreme weather will have varying degrees of impact on ecology, agriculture, industry and human life. La Nina is a common climate phenomenon, which refers to the abnormal drop of water temperature in the eastern Pacific Ocean near the equator. It's manifested as the obvious cooling of the eastern Pacific Ocean, but also accompanied by global climate chaos. Its formation is related to the cooling of the central and eastern equatorial Pacific Sea temperature and the strengthening of trade winds. So, in fact, La Nina is the product of a combination of the tropical ocean and atmosphere. In the past two years, the Northern Hemisphere has been continuously affected by the La Nina phenomenon, and many places have experienced high temperature, drought and heavy rainfall rarely seen in history. For example, the temperature in Sichuan and Chongqing in China has broken the record high, and the river flow has decreased significantly, while the precipitation in southern China and Fujian has increased by 20% to 100%, causing a lot of agricultural, economic and human losses. In general, a La

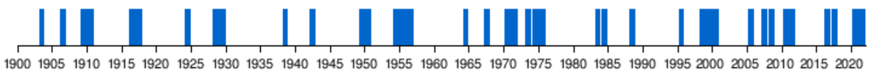

Figure 1 A timeline of all La Niña episodes between 1900 and 2022

Nina lasting two years is normal, but the current La Nina could last until the end of this year or even longer, and each La Nina will have different impacts, depending on how mild the event is and when it develops. In this context, it is particularly important to establish a data model to predict and evaluate the possibility of the occurrence of La Nina phenomenon in the future and the impacts it has already caused and will potentially cause, so as to avoid related disasters.

1.2 Work

At present, a wide range of countries and regions in the world have been greatly affected by the La Nina event. Based on the current development of the event, what we need to know is as follows:

1. First of all, we need to know which countries and regions in the world are mainly affected by La Nina events. Therefore, by comparing the changes of temperature and precipitation in all countries and regions before and after La Nina events, we can identify the countries and regions that are most affected. Collect and collate all the SOI indexes in the past 200 years, and predict the possibility of triple La Nina events by predicting the future SOI index.

2. For questions 2 and 3, we need to make statistics on the losses caused by disasters in these severely affected countries, including but not limited to: direct economic losses, the number of casualties, the area affected by crops, and the number of collapsed houses. At the same time, through the assessment of disasters, a series of measures are developed to reduce losses and targeted response strategies in line with disaster losses in different countries are proposed.
3. In view of the fourth question, the best prevention and control plan of La Nina event is quantitatively analyzed based on the statistical data we have collected, and submitted to the management department

2. Problem analysis

2.1 Analysis of question one

The topic calls for a statistical analysis of countries and regions affected by triple La Nina events and a prediction of the likelihood of future triple La Nina events. The answer to this question will also be divided into two parts.

Part 1: It is necessary to analyze the characteristics of La Nina phenomenon. We locate the areas easily affected by this phenomenon by combining relevant meteorological knowledge and news reports, and then search the meteorological information of the areas in the recent two La Nina years, compare and analyze them with the control group, and finally obtain the countries and regions affected by the triple La Nina, the aspects of the impact and the specific degree of the impact, and make a visual chart for presentation.

By analyzing the characteristics of La Nina phenomenon, it is known that this hydrological feature will make the water temperature in the eastern Pacific drop and lead to drought. On the contrary, the water temperature in the western Pacific rises and the precipitation is significantly higher than that in normal years, which will directly affect the eastern and western coasts of the Pacific. The eastern coast is drier and the western coast is wetter and easy to cause flooding. However, the influence of the middle and high latitudes, such as China, requires the help of multiple intermediate processes such as ocean and atmospheric circulation, which is generally indirect. The climate characteristics of these regions in La Nina years cannot be clearly and intuitively linked to La Nina phenomenon, so the investigation of the influence of La Nina phenomenon in the middle and high latitudes can be excluded. Mainly to the Pacific equatorial east and west coast countries and regions meteorological data statistical analysis.

The World Meteorological Organization says La Nina is responsible for periods of flooding, drought and other unusual weather. It affects the global climate and disrupts normal weather patterns, which can lead to intense storms in some places and droughts in others. Based on this information, the test indexes of regions that may be affected by La Nina are located on temperature and precipitation. By analyzing the temperature and precipitation data of each region in 2020 and 2021, the characteristics of drought and flood are found, and the characteristics of countries and regions affected by La Nina are finally determined by comparison and analysis.

Part 2: It is necessary to analyze and forecast the possibility of triple La Nina phenomenon in the future. First of all, it is necessary to define the prediction indicators. Through searching the data, we can see that the Southern Oscillation index (SOI) is closely related to La Nina phenomenon. In meteorology, the pressure difference between Tahiti and Darwin in the South Pacific is used to measure the active degree of El Nino phenomenon. If the southern Oscillation index is positive for a sustained period, there is a La Nina phenomenon in the year. Based on this, we collected the SOI data of the past 150 years, conducted data preprocessing to eliminate the data that were too long ago and the incomplete data in 2022. We used SPSS software for preliminary prediction to obtain the time series characteristics of SOI data, and then established the neural network model based on LSTM to calculate the output value of LSTM cells. The error term of each LSTM cell was calculated in reverse, including two reverse propagation directions at the time and network level. According to the corresponding error term, the gradient of each weight was calculated. The gradient based optimization algorithm was applied to update the weight and depict the future trend of SOI index, so as to predict the possibility of the occurrence of La Nina phenomenon.

The flow chart for predicting La Nina is as follows:

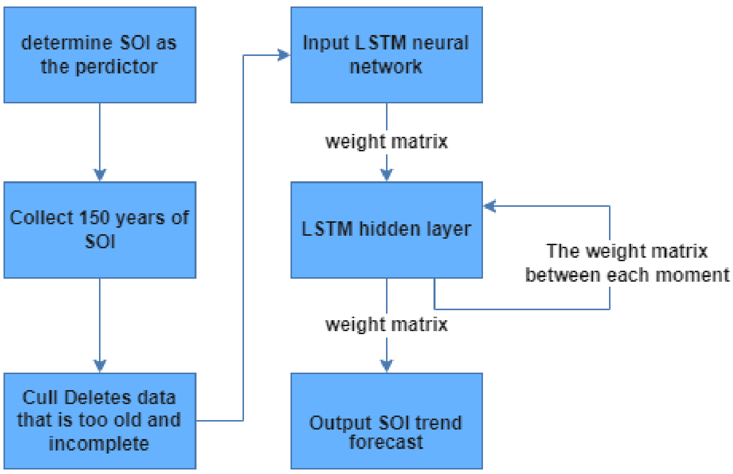

Figure 2 Flow chart for predicting La Nina

2.2 Analysis of question two

For question 2, it is required to assess and analyze the damage caused by high temperature and drought to a country and propose corresponding strategies. In the study of this issue, we choose the United States as the country for research and investigation. Based on the survey and analysis results of the first question, the southeast of the United States is one of the areas in the world seriously affected by La Nina events, which are characterized by high temperature and drought. But in the course of the data survey, few statistical agencies in the United States have counted the damage caused by heat and drought. A review of the extensive literature reveals four major impacts of heat and drought: hydropower generation, wildfire area, summer mortality, and agricultural production. We collected the data of the past 12 years. In order to highlight the different influences brought by triple La Nina, we divided every three years into one category for evaluation and analysis. As can be seen from the agricultural location distribution of the United States, subtropical crops dominate in the southeastern part of the United States, so we investigated rice, wheat and cotton. And through consulting the relevant data, using the analytic hierarchy process to assign value, get the score of agricultural production. In order to reflect the relative influence of each year more objectively, we use normalization to process each indicator, and then use analytic hierarchy process to get the weight of four indicators in the criterion layer, including hydraulic power generation 0.53, wildfire area 0.2575, summer population mortality 0.0665, agricultural production 0.1461. Then the fuzzy comprehensive evaluation method is used for scoring. The lower the score is, the greater the loss is. Finally, using the initial data, we conducted principal component analysis to analyze the meaning of each principal component and its contribution rate, so as to provide better advice to the government.

2.3 Analysis of question three

The third question requires the evaluation and analysis of disaster loss data caused by floods caused by triple La Nina events, and the establishment of flood disaster assessment model to comprehensively consider various factors caused by disasters within a certain spatio-temporal range. In this La Nina event, some parts of China suffered serious flood disasters, and Henan suffered unprecedented heavy rain in 2022. Therefore, we selected China as the primary object of this disaster assessment, and established the index layer of flood loss assessment, which integrated population, property and various economic losses. As well as the impact of the disaster on the ecological environment, infrastructure, food production and land, the final decision was made to summarize the evaluation indicators into the following four categories: **infrastructure loss, direct economic loss, population loss and agricultural loss**.

We established the evaluation model of flood disaster loss in China based on **fuzzy comprehensive evaluation method**, and adopted the relative deviation fuzzy matrix to establish the membership matrix. In order to reflect the relationship between indicators more objectively, we selected the coefficient of variation method to determine the weight of each indicator, among which the weight of infrastructure is **0.2065**, and the weight of direct economic loss is **0.2562**. Population loss weight **0.3195**, agricultural loss weight **0.2176**. We selected the flood loss data from 2011 to 2022 to evaluate the impact of triple La Nina events by fuzzy comprehensive evaluation method. Therefore, we also need to divide the data of the 12 years from 2011 to 2022 into four groups, and each of the three consecutive years is taken as one group. The average of three consecutive years of La Nina losses is used to more accurately distinguish the effects of triple La Nina events from the large losses of other single catastrophic years.

3. Symbol and Assumptions

3.1 Symbol description

Table 1 Symbol description

Symbol	Description
F_{te}	Partitioned test set time series
P_{te}	The fitting sequence of the test set
L	The segmentation window length
S_{state}	State vector size
η	Learning rate
f_t	Forget door
i_t	Input door
o_t	Output door
C_t	Current cell state
B	The judgment matrix
R	The fuzzy judgment matrix

3.2 Hypothesis

In order to make our model have better prediction and evaluation effect, we must make some assumptions to delineate its reasonable working range:

1. Assume that La Nina is only related to SOI: In this paper, we use SOI index to measure the occurrence of La Nina phenomenon, indicating that SOI index is the main indicator of La Nina phenomenon, ignoring other secondary indicators.

2. Assume that all indicators are only related to La Nina phenomenon, ignoring the influence of other conditions in the current year: In discussing the cost of heat and drought in the United States, we are using some grand statistics, and we are assuming that other factors do not account for fluctuations in the cost.
3. Assume that continuous positive SOI is La Nina phenomeno: The specific threshold of SOI and La Nina has not been determined in the academic circle. In this paper, it is assumed that the SOI index in most months of a year is positive, which means that La Nina will occur.
4. Assume that the areas affected by La Nina are all directly affected areas: Global atmospheric circulation is a complex system, so it is impossible to thoroughly analyze all the effects of La Nina. In this paper, only the direct effects of La Nina are considered.

4. Model

4.1 Question 1

4.1.1 Model preparation

- **SOI:**
  Southern Oscillation Index(SOI) effectively reflects the evolution of pressure increases and decreases on the east and west sides of the Pacific Ocean. When this index is positive, it means that Tahiti's pressure is higher than Darwin's normal. If the Southern Oscillation index is consistently positive, the year has an anti-El Nino, also known as a La Nina.
  Studies show that SOI can be a good predictor of La Nina events. Meanwhile, predecessors have developed different evaluation models of SOI index to study the causes and future trends of SOI changes.

- **LSTM:**
  Long Short-Term Memory(LSTM) is a temporal Recurrent Neural Network (RNN), which is suitable for processing and predicting important events with long intervals and delays in time series.
  Among many deep learning models, the Recurrent Neural Network introduces the concept of timing into network structure design, making it more adaptable in timing data analysis. Although RNN can deal with nonlinear time series effectively, it is still difficult to deal with time series with too long delay due to gradient vanishing and gradient explosion. Thus, the LSTM model was applied.
  LSTM was designed with a clear goal: to avoid long-term dependency problems. Compared with the chain form of RNN which contains a large number of repeated neural

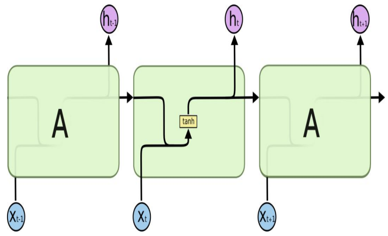

Figure 3 The standard RNN contains only duplicate modules for a single tanh layer

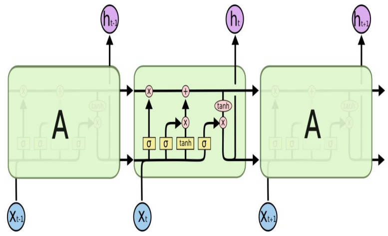

Figure 4 The schematic of LSTM

network modules, LSTM builds a chain structure of four neural networks that interact in a special way, making "remembering" information for a long time a default behavior rather than something difficult to learn.

The key to LSTM is the state of the cell, which is like a conveyor belt. It runs through the chain with only minor linear interactions. This is where the information is remembered, so it can easily flow through it in an unchanging form. To add/delete information in the cell, there are some control gates in the LSTM. They determine how information passes through and contain a sigmoid neural network layer and a pointwise dot product operation.

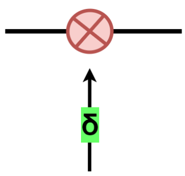

Figure 5 Gate of LSTM

After continuous evolution, the most widely used LSTM model's forward calculation process is as follows:

$$
\begin{aligned}
f_t &= \sigma(W_f \cdot [h_{t-1}, X_t] + b_f), \\
i_t &= \sigma(W_i \cdot [h_{t-1}, X_t] + b_i), \\
C_t &= f_t * C_{t-1} + i_t * \tanh(W_C \cdot [h_{t-1}, X_t] + b_C), \\
o_t &= \sigma(W_o \cdot [h_{t-1}, X_t] + b_o), \\
h_t &= o_t * \tanh(C_t),
\end{aligned}
\tag{1}
$$

Where $\sigma(x) = \frac{1}{1+e^{-x}}$ is the Logistic Sigmoid function with an output in [0,1], and $\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ represents the hyperbolic tangent function with an output in [-1,1]. W_f, W_i, W_C, W_o is the corresponding weight coefficient matrix in the process of cell information updating, and b_f, b_i, b_C, b_o is the corresponding bias matrix in the process of information updating.

4.1.2 Statistical analysis of countries affected by triple La Nina events

La Nina is mainly characterized by its influence on the eastern and western coasts of the Pacific Ocean, where the eastern coast is drier and the western coast is wetter and prone to flooding. It also has a certain impact on temperature and thunderstorms. As for natural disasters in the middle and high latitudes, due to the complex interaction of multiple intermediate processes such as ocean and atmospheric circulation, they cannot be clearly and directly connected with La Nina phenomenon. **Therefore, This paper mainly analyzes the meteorological data in 2020 and 2021, the two La Nina years, in the eastern and western equatorial countries of the Pacific Ocean.** Combined with relevant meteorological knowledge and news reports, the potential disaster areas were located, and the temperature, drought and flood conditions of the last two La Nina years in these areas were statistically analyzed. Compared with the control group, the countries and regions with climate anomalies were analyzed. If the geographical location matched the influence range and characteristics of La Nina phenomenon, the region was considered to be a triple La Nina affected area.

Through the investigation of relevant regional meteorological data, the situation of regions affected by triple La Nina events is shown in figures.

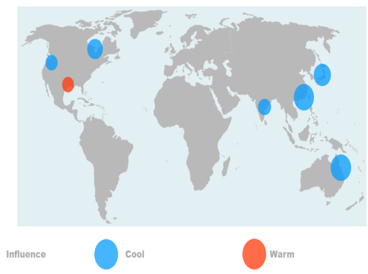

Figure 6 Areas affected by temperature

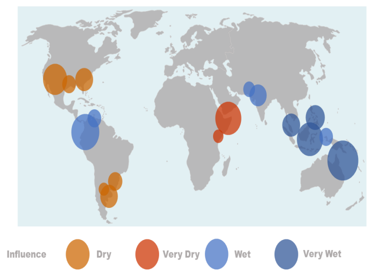

Figure 7 Areas affected by drought and flooding

Rainfall is heavy in Indonesia, eastern Australia and northeastern Brazil. Floods have hit many parts of Asia, such as Pakistan. Dry weather has hit eastern Argentina, southern Brazil, Uruguay and much of the southern United States, with equatorial Africa particularly affected. The devastating drought in the Horn of Africa continues to worsen and its consequences will affect millions of people.

4.1.3 Prediction of the possibility of the Triple La Niña events based on LSTM

First of all, we preprocessed the data. We collected the SOI of every month for nearly 200 years, eliminated the too old data and the incomplete data in 2022, and left 1,763 SOI records.

Then, the time series features of these data were explored, the data were initially fitted, and the correlation of the data was analyzed. The SOI showed obvious time series features and strong fluctuation in time. Accordingly, the SOI prediction and LSTM show a good fit The

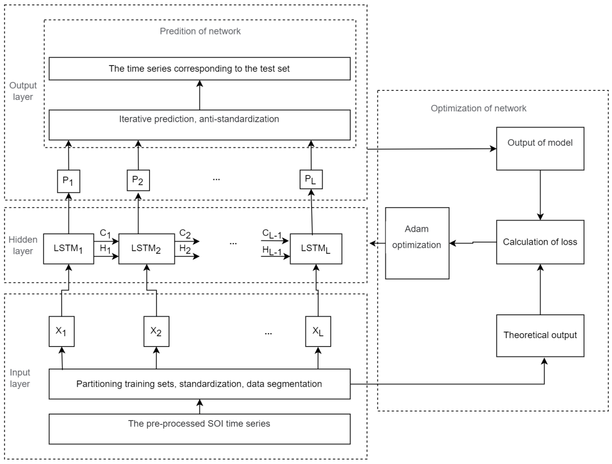

Figure 8 LSTM based framework

overall framework of the LSTM prediction model constructed in this paper is shown in the figure, including five functional modules: input layer, hidden layer, output layer, network training and network prediction. The input layer is responsible for the preliminary processing of the original SOI time series to meet the network input requirements; the hidden layer uses LSTM cells to build a single layer of circulating god meridian; the output layer provides prediction results; the Adam optimization algorithm is used for network training; the network prediction is made point by point by iteration.

Parameter optimization of LSTM prediction model is based on multi-layer grid search. Many parameters are involved in the construction of the LSTM prediction model, among which the segmentation window length L, state vector size Sstate and learning rate η are the most critical. In order to achieve better prediction results, this paper adopts the grid search method to optimize the three parameters. Compared with other hyperparameter optimization methods (such as genetic algorithm, random search algorithm, particle swarm optimization algorithm, Bayesian algorithm, etc.), grid search is a simple and practical optimization method that is easy to parallel computation and time-consuming controllable, and can well meet the task requirements and experimental requirements of fault time series prediction. Parameter optimization is based on the highest prediction accuracy at all test points in the test set, that is, the lowest prediction error $\varepsilon(P_{te}, F_{te})$, and the objective function can be expressed as

$$\min \varepsilon(P_{\text{te}}, F_{\text{te}})$$
$$\text{s.t.} \begin{cases} 2 \le L \le L_{\max} \le m, \text{step}_L \mid L \\ 2 \le S_{\text{state}} \le S_{\max}, \text{step}_{\text{state}} \mid S_{\text{state}} \\ \eta \in \{\eta_1, \eta_2, \cdots, \eta_r\}, \text{step}_\eta \mid r \\ L, S_{\text{state}}, r, \text{step}_L, \text{step}_{\text{state}}, \text{step}_\eta \in N \end{cases} \tag{2}$$

In the formula, $\text{step}_L, \text{step}_{\text{state}}$ and step_η are the mesh search step size of corresponding parameters respectively. The three parameters L, S_{state} and η constitute a three-dimensional search space, and the optimal parameter combination can be obtained by multi-layer grid search algorithm. The search process mainly consists of three layers, from the inside to the outside of the grid search for L, S_{state} and η. First, the number of random seeds and the number of training steps were fixed according to the preset value range of three parameters. Then, the values of the three parameters were traversed respectively, and the LSTM time series model was trained and predicted in the innermost layer, and the corresponding model parameters and model accuracy were saved. Finally, all the saved results are sorted according to the prediction accuracy from high to low, and the first parameter combination is the preferred model parameter.

According to the optimal parameters, **the number of hidden units and training times are determined to be 200 and 3000 respectively**, and the error of the test set is minimum

Then, the model based on LSTM is built and trained. In the process of model training, the first 90% of the data set is taken as the training set, and the last 10% is taken as the test set.

Input SOI data over the years and output SOI forecast data for the next ten years, compare the measured data with the predicted data, and verify the error range of the model.

For future La Nina projections, we first forecast SOI data for the second half of 2022 and the beginning of 2023, with the following results:

Table 2 SOI forecast for the next 9 months

Date	2022.7	2022.8	2022.9	2022.10	2022.11	2022.12	2023.1	2023.2	2023.3
Date	1.35	4.09	2.95	3.06	3.19	3.74	4.57	5.16	4.41

As can be seen from the data in the table, the occurrence probability of triple La Nina is relatively high due to continuous positive SOI in the next 9 months. We then analyzed SOI data

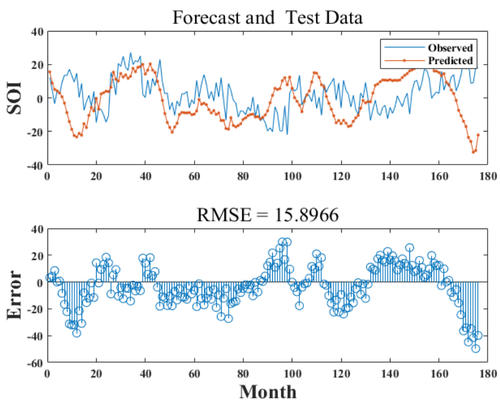

Figure 9 Comparison between predicted data and measured data

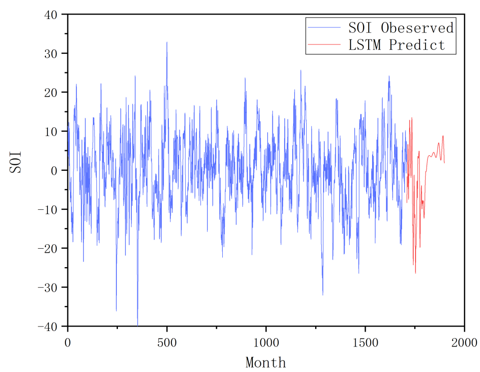

Figure 10 SOI forecast for the next ten years

for the next ten years to predict years that would be prone to La Nina events. According to the forecast results, **there is a high probability of La Nina in 2029 and 2023.**

By comparing the measured data with the predicted data, the two show a certain consistency, and the error is within the acceptable range, achieving a good effect.

4.2 Question 2

4.2.1 Drought related data preparation in the United States

In the process of reviewing the literature, we found that the high temperature and drought in the United States were closely related to four aspects, namely, the hydroelectric power generation, wildfire area, summer population mortality and agricultural production. Therefore, we collected these relevant data through the official websites of the United States Bureau of Statistics, the United States Bureau of Agriculture and the United States Electric Power Administration. In addition, we collected some economic indicators of that year and finally got nine indicators, including Twh net power generation from conventional hydroelectric power generation, number of wildfires, acres of wildfires, death rate per million people in summer, GDP growth rate, GDP index, rice yield, cotton yield and wheat yield.

In the process of reviewing the data, we found that some indicators were missing in 2022, and the forecast based on the first question is that 2022 is the month included in the triple La Nina, so it should be taken into account. For these missing values, we patiently conducted grade-ratio test for each indicator, and found that all of them had passed the grade-ratio test. Therefore, we used the gray model for prediction, completed the missing values, and finally got all the data in the table.

Table 3 High temperature and drought related data statistics in the United States

Year	Annual hydropower generation/Twh	Area of wildfire/acre	Number of wildfires per million people	Summer death rate / million dollars	GDP growth rate /million ton	GDP	Wheat /thousand bales	Rice/t	Cotton
2022	284.011	387296.197	57696.000	0.484	0.033	70365.332	43.406	10407003.562	18354.000
2021	260.230	382893.000	58985.000	0.476	0.057	69386.400	43.311	9743677.931	17852.000
2020	285.270	358447.000	58950.000	0.468	-0.034	63413.000	48.094	11558234.482	14061.000
2019	287.870	345962.000	50477.000	0.459	0.023	65279.000	50.829	9403708.939	19227.000
2018	292.520	364262.000	58083.000	0.565	0.029	63064.000	49.596	11371231.216	17566.000
2017	300.330	371096.000	71499.000	0.430	0.023	60109.000	45.801	9054392.324	20223.000
2016	267.810	384778.000	67743.000	0.495	0.017	58021.000	60.738	11387081.534	16601.000
2015	249.080	345506.000	68151.000	0.308	0.027	56863.000	54.247	9808908.084	12455.000
2014	259.370	361136.000	63312.000	0.138	0.023	55049.000	53.310	11289033.095	15753.000
2013	268.570	351316.000	47579.000	0.446	0.018	53106.000	56.169	9627035.850	12275.000
2012	276.240	273192.000	67774.000	0.618	0.023	51602.000	59.256	10157361.060	16534.000
2011	319.360	312789.000	74126.000	0.777	0.016	49882.000	52.436	9395428.164	14722.000
Ten years average	279.222	353222.766	62031.250	0.472	0.021	59678.311	51.433	10266924.687	16301.917

In addition, given the repetitive and difficult nature of these data, for example, GDP is influenced by a variety of factors, and given the widespread impact of COVID-19 in the last three years, it is difficult to use GDP to measure a country's drought degree. At the same time, between the number of acres of wildfire and the number of wildfires, the number of acres of wildfire is more representative of disaster losses. Meanwhile, the influence of the triple La Nina phenomenon is to be measured in this question. We take every three years as a class and take the mean value of each index for three years, as shown in the table. Among them, 2020 to 2022 are the years with the triple La Nina influence, and the rest are normal years, which can reduce the influence of other factors except La Nina on these indicators. It lays a foundation for establishing fuzzy comprehensive evaluation model.

Table 4 Drought related statistics of every three years in the United States

	Annual hydropower generation	Number of wildfires	Area of wildfire/acre	Summer death rate per million people	GDP growth rate	GDP/dollar	Rice/kiloton	Cotton/thousand bales	Wheat/million ton
From 2020 to 2022	276.504	58543.667	8410430.333	0.476	0.019	67721.577	10569.64	16755.667	44.94
From 2017 to 2019	293.573	60019.667	7819314.000	0.485	0.025	62817.333	9943.11	19005.333	48.74
From 2014 to 2016	288.553	55836.667	7851397.333	0.498	0.006	63918.667	10777.72	16951.333	49.51
From 2011 to 2013	272.407	69131.000	8553743.333	0.411	0.022	58331.000	10083.46	16426.333	53.60

4.2.2 Establishment and solution of drought disaster evaluation model

- **Review of the module:**
  To evaluate the impact of drought in the United States under the triple La Nina phenomenon, considering the large number of indicators, we established a three-tier evaluation model under the condition of combining and processing similar indicators. The target layer is the severity of the disaster, and the first-tier index is the annual hydroelectric power generation, wildfire area, summer mortality per million people, and agricultural production. Agricultural production also includes a secondary index, with the annual production of wheat, rice and cotton as factors. It is worth noting that the annual hydropower generation and agricultural production are negative indicators, that is, the more serious the disaster, the less electricity generation and production will be, as shown in Table. We use analytic hierarchy process (AHP) to score agricultural production and determine the weight of each factor. Then the evaluation set is determined and the loss degree of the four three-year disasters is given by fuzzy comprehensive evaluation method.
- **Establishment of fuzzy comprehensive evaluation method:**
  - **Determination of factor set of evaluation index:**
    For the four evaluation objects in the data, we determined the evaluation set U as:
    U=Very serious loss, relatively serious loss, serious loss, little loss
    We then determine the specific score in the three years in which the triple Lanina occurred and in other general years to get an approximate estimate of the extent of the loss.
  - **Determination of the weight of the first-level indicator layer and the second-level**

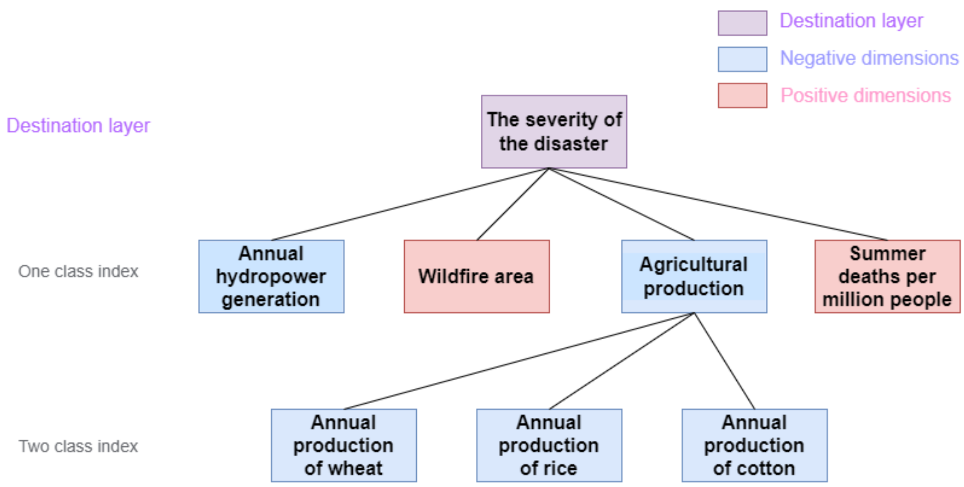

Figure 11 Disaster level diagram

indicator layer:

Since there are many factors affecting the data, we determined the weights by referring to literature and using the analytic hierarchy process, hoping to highlight the losses caused by La Nina to the United States in all aspects.

First of all, for each factor of the secondary index under agricultural production, by consulting various literatures, referring to expert opinions and the actual situation, we made a pairwise comparison of the annual production of wheat, rice and cotton, and concluded that cotton and rice should be the main crops affected, so as to construct the judgment matrix B:

$$B = \begin{bmatrix} 1.00 & 5.00 & 3.00 \\ 0.20 & 1.00 & 0.33 \\ 0.33 & 3.00 & 1.00 \end{bmatrix} \tag{3}$$

Secondly, to establish the weight of each index of the first-level indicator layer, by referring to various literatures and various data and reports, we conducted pairwise comparison and scoring on the annual hydropower generation, wildfire area, mortality rate per million people in summer and agricultural production. Among them, hydropower generation should be an important index for direct evaluation of high temperature and drought, so the judgment matrix C is obtained:

$$B = \begin{bmatrix} 1 & 3 & 7 & 3 \\ \frac{1}{3} & 1 & 3 & 3 \\ \frac{1}{7} & \frac{1}{3} & 1 & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & 3 & 1 \end{bmatrix} \tag{4}$$

Then, the matrix B and C are used for consistency test:

First, the consistency index is defined as:

$$CI = \frac{\lambda - n}{n - 1} \tag{5}$$

Where λ is the eigenvalue of matrix B and n is the order of matrix A. In order to judge the consistency of *CI* and test whether the judgment matrix is reasonable, it is necessary to introduce the consistency index *RI*:

$$RI = \frac{CI_1 + CI_2 + \cdots + CI_n}{n} \tag{6}$$

At the same time, considering that the deviation of consistency may be caused by random reasons, it is necessary to compare CI and random consistency index RI to obtain the test coefficient CR when verifying whether the judgment matrix has satisfactory consistency. The formula is as follows:

$$CR = \frac{CI}{RI} \tag{7}$$

Through matlab calculation, it can be seen that matrix B and matrix C have passed the consistency test, thus the weight is obtained:

$$Wb = [0.6376, 0.1046, 0.2578]$$
$$Wc = [0.530, 0.2575, 0.0665, 0.1461]$$

- **Establishment of fuzzy judgment matrix:**

  In this paper, the relative deviation fuzzy is used to establish the membership matrix, and the fuzzy judgment matrix of each index u_{1i} is established, namely:

$$R = \left[r_{ij}\right]_{4\times4} = \begin{bmatrix} r_{11} & r_{12} & r_{33} & r_{14} \\ r_{21} & r_{22} & r_{23} & r_{24} \\ r_{31} & r_{32} & r_{33} & r_{34} \\ r_{41} & r_{42} & r_{43} & r_{44} \end{bmatrix} \tag{8}$$

$$r_{ij} = \frac{\left|x_{ij} - u_{1j}\right|}{\max\left\{x_{ij}\right\} - \min\left\{x_{ij}\right\}} \tag{9}$$

  Where, r_{ij} represents the membership degree of factor u_i to grade v_j, x_i is the value of evaluation factor i, and a_{ij} is the classification threshold of grade j of evaluation factor i.

- **Comprehensive judgment:**

  Using the calculated weight and fuzzy judgment matrix, a fuzzy transformation can be obtained:

$$T_R : F(U) \to F(V) \tag{10}$$

Through this transformation, the comprehensive evaluation result S = Wc·R can be obtained, and then the damage degree of the disaster can be obtained by corresponding the score to the evaluation language set.

- **The solution of fuzzy comprehensive evaluation method**

  We have used the analytic hierarchy process to calculate the weights of the first and second indexes

$$Wb = [0.6376, 0.1046, 0.2578]$$
$$Wc = [0.530, 0.2575, 0.0665, 0.1461]$$

  Then we should solve the fuzzy judgment matrix. Since the AHP was applied to score agricultural production, a score value between 0 and 1 was obtained. In order to unify the dimension, we normalized the column vector of each index

$$\bar{u}_{1i} = \frac{u_{1i}}{\sum_{i=1}^{4} u_{1i}} \tag{11}$$

  On this basis, the fuzzy matrix is solved as follows:
  Firstly, we distinguish that the annual hydropower generation and agricultural production are negative indicators, and the fire area and the death rate of one million people in summer are positive indicators. For the negative indicators, we take the minimum value, and for the positive indicators, we take the maximum value.
  Thus, the fuzzy judgment matrix R can be written as:

$$R = \begin{bmatrix} 0.193 & 0.831 & 0.308 & 0 \\ 0.234 & 1 & 0.464 & 0 \\ 0.616 & 0.299 & 0 & 4.234 \\ 0.397 & 0.059 & 0.236 & 0 \end{bmatrix} \tag{12}$$

  The comprehensive score S is obtained by using the fuzzy transform S = Wc·R

$$S=[0.262\ 0.726\ 0.317\ 0.281]$$

  Where S_i respectively corresponds to the score of the i-th year, and the lower the score is, the greater the disaster loss of the high temperature and drought in the current year. Therefore, we can know that the disaster loss of the high temperature and drought in the three years under the triple La Nina is the most severe

4.2.3 Analysis of the results of evaluation of drought model

It can be seen from the above table that the triple La Nina phenomenon is the most serious, so we can show that the triple La Nina phenomenon has caused serious losses to the power system, agricultural production and natural environment of the United States. By referring to relevant literature, in 2021 and 2022, the total economic losses of heat waves and droughts in the United States were \$9.4 billion and \$9.3 billion, respectively, ranking among the top 10 in the history of disaster losses in the United States. Given the severe impact of La Nina, the U.S. government should pay more attention to heat and drought protection measures. The

Table 5 U.S. Heat Hazard assessment scores of every three years

	2011-2013	2014-2016	2017-2019	2020-2022
Score	0.281	0.317	0.726	0.262
Evaluation	relatively serious loss	serious loss	little loss	Very serious loss

weights obtained by fuzzy comprehensive evaluation should pay attention to the two aspects of hydropower and forest fires.

In order to further explore the damage impact in a more detailed direction, we carried out principal component analysis on the initial data and obtained three principal component values, among which the cumulative contribution of the first two indicators reached 85%, and the first index of the two indicators was mainly composed of hydraulic power generation (0.3546) and wildfire area (0.4028). The second index mainly affects rice yield (0.4567), which indicates that the U.S. government should pay more attention to disaster prevention in these two aspects.

We further checked the regional distribution map of disaster levels in the United States under the La Nina phenomenon. According to the winter drought distribution map in the United States, the drought in the central and western parts of the United States is more serious. By comparing the distribution in summer and winter, we found that due to the strengthening of the La Nina phenomenon in winter, the eastern part of the country also has different degrees of drought. So the U.S. government should tighten its drought watch in the Southeast this coming winter.

Based on the above data analysis, we propose the countermeasures as follows: develop a variety of power generation methods to promote the development of photovoltaic power generation industry, so as to make up for the lack of hydroelectric power generation in hot and dry weather. At the same time, meteorological monitoring should be strengthened, and various agricultural activities should be rationally arranged to reduce the impact of rice and other crops. In addition, public publicity should be strengthened to popularize the causes of heat emission disease and protective measures to reduce the impact of high temperatures.

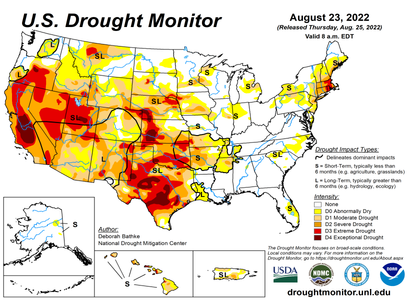

Figure 12 Regional distribution of drought in the United States in summer

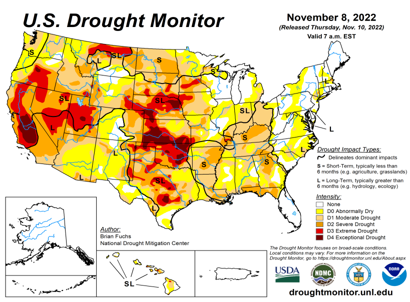

Figure 13 Regional distribution of drought in the United States in winter

4.3 Question 3

4.3.1 Establishment of evaluation index system of flood disaster loss in China

The index selection of the disaster grade evaluation system is very important in the fuzzy evaluation model, which often needs to weigh all aspects of the comprehensive quantitative analysis. The evaluation of flood disaster level in China needs to consider various indexes comprehensively to reflect the authenticity of the disaster situation. Therefore, considering the economic loss, the number of population loss, the number of infrastructure and the area of agricultural land loss caused by flood, the evaluation indexes can be summarized into four categories: infrastructure loss, population loss, agricultural loss and economic loss.

For the loss of infrastructure, we mainly consider the number of damaged reservoirs and collapsed houses. In terms of population loss, we divided it into three sub-indicators: the number of affected people, the number of dead people and the number of missing people; The agricultural losses mainly counted the affected area and failed area of crops in the flood, and took them as indicators. Area of crop failure refers to the area sown with more than 80% of the crop yield due to the disaster. In this model, we only consider the direct economic loss, that is, the sum of personal and immovable property losses directly caused by primary natural disasters. The damage of post-flood infrastructure requires a large amount of human and material resources, including the damage of some historical relics and the cost of repairing them.

Indicators of each layer are represented by a tree-like hierarchy as shown in the figure below: Due to the triple La Nina events from 2020 to 2022, the data of 2022 are indispensable. Since

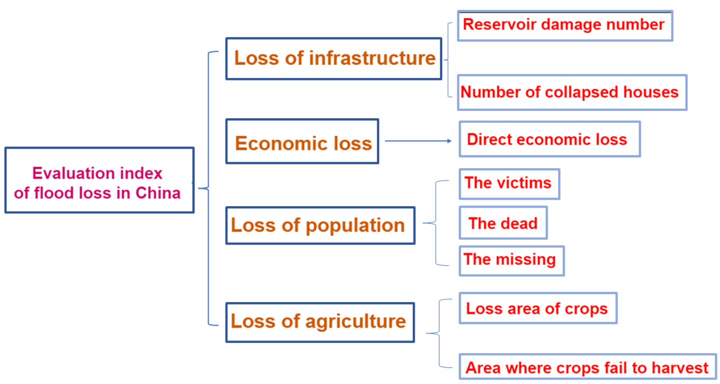

Figure 14 Evaluation index system of flood disaster loss in China

it is now November 2022, we cannot get the data of flood disaster loss for the whole year of 2022. However, through the investigation of the data of the first half of 2022, and based on the ratio of flood losses in the first half of the last decade to those in the whole year, and taking into account the frequent flood season in autumn, we make a reasonable estimate of flood losses in 2022. Considering the continuity of triple La Nina events, the data from 2011 to 2022 were statistically analyzed and divided into four groups, that is, the first group was from 2020 to 2022, the second group was from 2017 to 2019, the third group was from 2014 to 2016, and the fourth group was from 2011 to 2013. In this way, to a great extent, the assessment of disaster losses caused by La Nina events, such as 2012,which happened the largest flood peak since the Three Gorges reservoir was built , and the impact of a single La Nina event on disaster losses in that year, such as: The single La Nina events in 2013 and 2016 focused on the flood impacts of the triple La Nina events. After visualization of all the preprocessed indicator data, the following figure can be obtained: It is not difficult to see from the data in the figure that under the influence

	Affected population (10,000)	Death Population (10,000)	Missing Population(10,000)	Loss area of crops (10,000 hectares)	Area of crop failure (10,000 hectares)	Number of collapsed houses (10,000)	Direct economic loss (100 million yuan)	Number of damaged reservoirs
Average of 2020-2022	6838.10	371.00	81.33	638.49	112.91	12.43	2442.24	1810.33
Average of 2017-2019	5286.02	358.67	52.00	610.13	92.29	10.86	1893.57	727.00
Average of 2014-2016	8372.69	497.00	126.33	716.49	100.82	28.00	2292.52	1117.67
Average of 2011-2013	11094.36	655.67	218.00	1006.24	146.36	57.42	2377.44	1086.67

Figure 15 Flood damage in China

of La Nina events, direct economic losses and damage to infrastructure in 2020-2022 will reach the maximum. In terms of the affected population, due to the improvement of China's flood prevention and control ability, more and more people can survive the flood disasters, so the disaster loss performance is not outstanding intuitively.

4.3.2 To build a fuzzy set

According to the index system of flood disaster loss evaluation in China established in Figure, the factor set of comprehensive flood disaster evaluation is established

$$\begin{aligned} u &= [u_{21}, u_{22}, u_{23}, u_{24}] \\ &= [Infrastructureloss, populationloss, agriculturalloss, economicloss] \end{aligned} \tag{13}$$

4.3.3 Coefficient of variation method

In order to more objectively and truly reflect the relationship between the evaluation indexes, another parameter representing the difference between the characteristic values of the evaluation indexes, namely the coefficient of variation, is used to determine the weight of the evaluation indexes, so as to avoid the overly balanced weight distribution.

- Construct the evaluation index matrix. 4 disaster loss conditions to be evaluated were established by using the 4 groups of data. The number of indexes involved in the evaluation was 4, and then the evaluation index matrix was as follows:

$$X = [x_{ij}]_{4x4} \quad (i = 1, 2, 3, 4; j = 1, 2, 3, 4) \tag{14}$$

Where x_{ij} is the relative index evaluation value of the i-th index of the data in the j group.

- Calculate the coefficient of variation of the evaluation index i

$$\varepsilon = D/\overline{x}_i \tag{15}$$

- Calculate the weight of the i-th evaluation index

$$W_i = \varepsilon_i / \sum_{i=1}^{m} \varepsilon_i \tag{16}$$

Where ε_i is the variation coefficient of the ith indexes. D is the mean square deviation of the index eigenvalue of the i-th term

$$D = \frac{1}{n} \sum_{j=1}^{n} (x_i - \overline{x}_i)^2 \tag{17}$$

$$X_i = \frac{1}{n} \sum_{j=1}^{n} X_{ij} \tag{18}$$

Where x_i is the mean value of the eigenvalue of the ith evaluation index

The final index weight set W is:

$$\begin{aligned} W &= [W_1, W_2, W_3, W_4] \\ &= [0.2065, 0.3195, 0.2176, 0.2562] \end{aligned} \tag{19}$$

4.3.4 Optimization index data

The direct purpose of disaster assessment is to determine the level of disaster and the index of the great impact of centralized management on the disaster, and then put forward a series of targeted measures, so as to reduce the loss caused by the disaster to a great extent. Due to the great differences in the frequency and affected area of floods in different regions, as well as the great differences in the level of economic development, absolute indicators are not suitable to evaluate the flood disaster situation. In order to make the evaluation results more conducive to the formulation of disaster prevention and reduction policies, this paper adopts relative indicators, namely, the loss rate of various indicators caused by floods in China, to evaluate the disaster situation

$$I_i = \left(\frac{\text{Number of damaged reservoirs}}{\text{Total reservoir number}} + \frac{\text{Number of collapsed houses}}{\text{Total number of house}} \right) \times 10^4 \tag{20}$$

$$E_i = \frac{\text{Direct economic loss}}{\text{GDP}} \times 10^4 \tag{21}$$

$$P_i = \left(\frac{\text{Population affected}}{\text{Total population}} + \frac{\text{Dead + missing persons}}{\text{Population affected}} \right) \times 10^4 \tag{22}$$

$$A_i = \frac{\text{Area of crop disaster + area of crop failure}}{\text{Total crop area}} \times 10^4 \tag{23}$$

Where Ii represents the mean value of three-year infrastructure losses in each group, Ei represents the mean value of three-year economic losses in each group, Pirepresents the mean value of three-year population losses in each group, and Ai represents the mean value of three-year agricultural losses in each group

After sorting out the statistical historical data and integrating various experts and authoritative literatures, based on the total loss caused by floods, China's flood disaster loss level is divided into 4 levels, s.t. ,the disaster loss evaluation set is

$$\begin{aligned} v &= [v_1, v_2, v_3, v_4] \\ &= [Lossisveryserious, lossisserious, lossisrelativelyserious, lossissmall] \end{aligned} \tag{24}$$

4.3.5 Establish the membership matrix of evaluation index

In this paper, the relative deviation fuzzy matrix is used to establish the membership matrix, namely

$$R = \left[r_{ij} \right]_{4\times 4} = \begin{bmatrix} r_{11} & r_{12} & r_{33} & r_{14} \\ r_{21} & r_{22} & r_{23} & r_{24} \\ r_{31} & r_{32} & r_{33} & r_{34} \\ r_{41} & r_{42} & r_{43} & r_{44} \end{bmatrix} \tag{25}$$

$$r_{ij} = \frac{\left| x_{ij} - u_{2j} \right|}{\max \left\{ x_{ij} \right\} - \min \left\{ x_{ij} \right\}} \tag{26}$$

Where, r_{ij} represents the membership degree of factor Ui to grade Vj, xi is the value of evaluation factor i, and a_{ij} is the classification threshold of grade j of evaluation factor i.

4.3.6 Establish a flood damage assessment model

According to the fuzzy comprehensive evaluation method, disaster comprehensive factor set U, evaluation set V and membership matrix R can form a fuzzy comprehensive evaluation model. The input of this model is the evaluation data of each index, and the output is the comprehensive evaluation of disaster situation B=(b1, b2, b3, b4), namely, the evaluation model of flood disaster loss in China is:

$$\begin{aligned} \mathrm{B} &= (b_1, b_2, b_3, b_4, b_5) = W \cdot \mathrm{R} \\ &= (W_1, W_2, W_3, W_4) \begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} \\ r_{21} & r_{22} & r_{23} & r_{24} \\ r_{31} & r_{32} & r_{33} & r_{34} \\ r_{41} & r_{42} & r_{43} & r_{44} \end{bmatrix} \end{aligned} \tag{27}$$

Where, B is the evaluation value of disaster loss, b_i is the comprehensive evaluation result of factor u_i, W_i is the weight value of u_i, and r_{ij} is the membership degree of factor u_i to grade V_i

4.3.7 Analysis of the results of evaluation of flood model

Based on the fuzzy evaluation model established by us, we first normalize the index data obtained comprehensively, import the data into MATLAB, and use the coefficient of variation method to establish the weight of each first-level index, and get:

$$W = \{W_1, W_2, W_3, W_4\} = \{0.2065, 0.3195, 0.2176, 0.2562\}$$

The fuzzy synthesis matrix is determined according to the index data:

$$R = \begin{bmatrix} 0 & 1 & 0.6045 & 0.5102 \\ 0.8613 & 1 & 0.4863 & 0 \\ 1 & 0.9868 & 0.6249 & 0 \\ 0 & 0.7548 & 1 & 0.2752 \end{bmatrix}$$

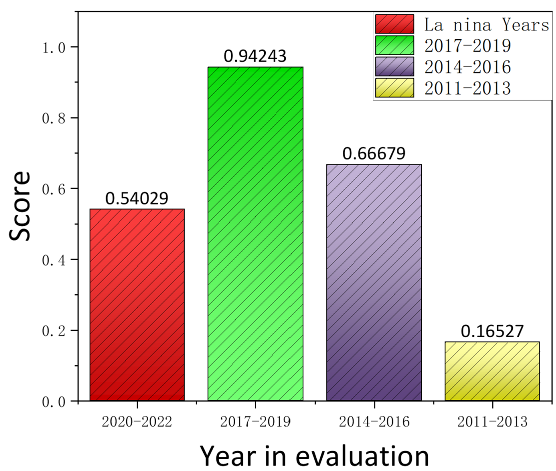

Figure 16 Evaluation of years

Thus, we input the results into the flood disaster loss assessment model and get the figure.

By analyzing the disaster grade score, it can be seen that the lower the score is, the more serious the disaster loss is, and the three years directly affected by La Nina events belong to V3={loss is serious} in the disaster grade evaluation set.

By comparing W_1, W_2, W_3, W_4, it can be seen that direct economic losses account for a large weight in the comprehensive disaster loss evaluation model, but direct economic losses are not affected by human factors. Therefore, if the policy establishment in this aspect has little effect, in terms of flood losses in China, under the influence of triple La Nina events, The losses caused by floods in 2020-2022 belong to the evaluation of ”serious losses”, among which the population loss evaluation index has a larger weight, indicating that the flood has a wider impact scope, so the number of people affected by floods increases significantly compared with other years. Coupled with three La Ninas in three years and a global outbreak, infrastructure maintenance is not as frequent as before, resulting in more infrastructure damage when floods occur because maintenance is not done in time. Frequent flooding also leads to the rapid spread of infectious diseases, which greatly magnifies the impact and population losses indirectly caused by disasters. The floods have also had a big impact on agriculture. In the past three years, the affected area reached 19.15463 hectares, and the grain output decreased significantly.

Therefore, we have put forward very targeted measures to deal with this phenomenon, because the La Nina phenomenon in China is characterized by frequent floods in the north and frequent high temperature and drought in the south. The decrease in crop yields caused by floods is devastating to the citizens who have been farming for a long time, so the government needs to improve relevant policies to ensure the basic standard of living of farmers. At the same time, the frequency of maintenance of levees and DAMS during the epidemic should be increased, and the construction of infrastructure should be strengthened. For areas vulnerable to La Nina events, the construction of flood control and disaster resistance in designated areas should be strengthened.

4.4 Question 4

Dear relevant management departments:

As an enthusiast with strong interest in meteorological knowledge, I pay close attention to this triple La Nina event. In this regard, I would like to provide some countermeasures with relevant mathematical knowledge. We built three different mathematical models to deal with three aspects of the problem, namely LSTM model, AHP-fuzzy comprehensive evaluation model, coefficient of variation method - fuzzy comprehensive evaluation model, and then predicted the possibility of triple La Nina events in the future, as well as the loss assessment of various disasters associated with La Nina events on a global scale. Through this series of processes, we have obtained interesting and interesting results, which I hope can help you to make appropriate strategies for La Nina events.

First of all, we determined the relevant influencing factors for judging the occurrence of La Nina events. In the United States, the 3-month sliding average absolute value of NINO3.4 index exceeded 0.5℃ for at least 5 consecutive months, and the occurrence of La Nina events was identified as the root of the occurrence of La Nina events. According to the Southern Oscillation theory of scholar Walker, Tahiti Station and Ervin Station were selected. The difference of sea level pressure between the two stations is processed by mathematical statistics and the Southern Oscillation index (SOI) is obtained. According to the value of SOI, we can judge whether the La Nina event occurs in the current year.

Therefore, we came up with an effective way to predict La Nina, that is, to predict the future changes of the SOI index, and finally to predict the possibility of triple La Nina events in the future. In our analysis, a La Nina event is judged to be highly likely in 2022 and is forecast to last until March 2023. In addition, using nearly 200 years of SOI data, we predict that there is a high probability of a La Nina event occurring again in 2029 and 2030, while a triple La Nina event is unlikely to occur again in the next decade.

In addition, we established AHP-fuzzy comprehensive evaluation model and coefficient of Variation method-fuzzy comprehensive evaluation model. We selected the drought-severely affected areas in the United States and the flood-severely affected areas in China under the influence of La Nina events as samples to establish the high-temperature and drought loss assessment model in the United States and the flood disaster loss assessment model in China.

Then, we analyzed the data of the drought loss evaluation results in the United States, and found that the main indicators affecting the disaster loss level were mainly hydropower generation (0.3546) and wildfire area (0.4028). We therefore suggest that the US government should pay more attention to disaster prevention in these two aspects. We further examined the regional distribution of La Nina hazard levels in the United States. According to the winter drought Map of the United States, the drought is more severe in the central and western United States. By comparing the distribution in summer and winter, we found that eastern China also had different degrees of drought in winter due to the strengthening of La Nina. So the U.S. government should step up its drought watch for the Southeast this coming winter.

In fact, we did encounter unprecedented high temperature weather in July and August this year. Based on the above data analysis, we believe that we should develop a variety of power generation methods to promote the development of photovoltaic power generation industry, so as to make up for the shortage of hydropower power generation in terms of heat and dry weather. At the same time, it is necessary to strengthen the meteorological monitoring, and carry out multiple monitoring reasonable arrangement of agricultural production activities, reduce the impact of rice and other crops. In addition, public awareness of the causes of heat emission and protective measures to reduce the impact of heat on disease should be strengthened.

In terms of flood losses in China, under the influence of triple La Nina events, the losses caused by floods in 2020-2022 belong to the evaluation of ”serious losses”, among which the evaluation index of population loss has a relatively large weight, which indicates that the flood has a wide range of impacts, so the number of people affected by floods increases sharply compared with other years. Combined with the three years of triple La Nina and the global outbreak, infrastructure maintenance is not as frequent as before, resulting in more infrastructure damage due to lack of timely maintenance when floods occur. Frequent floods will also lead to the rapid spread of infectious diseases, which will greatly expand the damage of the disaster and the loss of population. The impact of the floods on agriculture is also significant. The total area affected by the floods in the past three years has reached 19.15463 hectares, and the grain output has decreased significantly.

Thank you for taking time out of your busy schedule to read my letter. I hope my suggestions will be helpful to you.

Yours sincerely

Team 2022092012570

5. Sensitivity Analysis

In the discussion of Model One, we have analyzed the model. For Model 2 and Model 3, in order to prevent the process of determining the weight from being too subjective, we carried out the sensitivity analysis of the model.

Aiming at the AHP-fuzzy comprehensive evaluation method in question 2, we extracted the weight of the annual hydropower generation index for sensitivity analysis. When the weight of other indicators remained unchanged, we moved the weight of the annual hydropower generation index (0.53) up or down by 0.1, so as to obtain the relationship between the score and the weight in three years of La Nina phenomenon, as shown in the figure. It can be seen that the score has a good sensitivity to the index of the annual hydropower station, indicating that the annual hydropower generation has a greater impact on the drought evaluation system in the United States, and its weight is larger, in line with the expected results. As for the three coefficient of variation methods of the problem - fuzzy comprehensive evaluation method, the impact of flood on population has a higher index (0.32). As mentioned above, we let it fluctuate up and down by 0.1, thus finding that the sensitivity of the model is good, indicating that this index has a higher weight in the evaluation system and is in line with the result expectation.

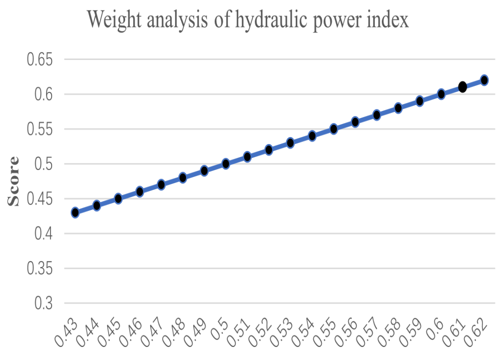

Figure 17 Weight analysis of hydraulic power index

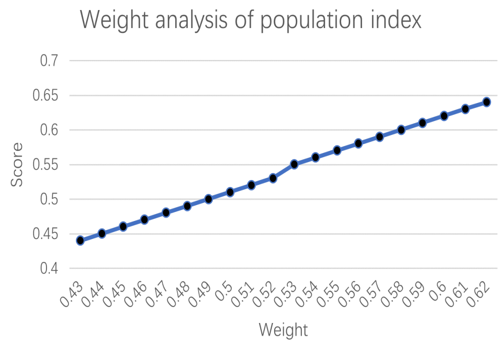

Figure 18 Weight analysis of population index

6. Strengths and Weakness

Advantages of the model:

For model 1, the data we collected is a long-term series of time oscillations. If Arima model is used for analysis and prediction, the prediction effect is not ideal, and it is difficult to predict the future for a long time, while LSTM model can do this, and its prediction for a long time series is more accurate.

For model 2, we collected various industrial data of the United States in each year. Since there was no comparison between the data itself, we divided the data into normal years and triple La Nina years in terms of data processing. In this way, the impact of ordinary La Nina years on the indicators could be reduced, and the impact of the high temperature and drought caused by the triple La Nina phenomenon in the United States could be highlighted to a large extent.

For model 3, we can accurately collect the data caused by floods in China every year. The data source is reliable and accurate, and the losses caused by disasters can be quantitatively and qualitatively analyzed. The relative index is introduced to increase the objectivity of the index.

Disadvantages of the model:

In the LSTM model, there are many oscillations in the data that are difficult to eliminate, which may cause some deviation to the results. In the second model, the data of American industries are relatively macro and have little influence on the results.

References

[1] Zhang Jiquan, Zhang Hui, Tong Zhijun, Song Zhongshan & Wu Xiaotian.(2007). Evaluation and classification of grassland fire disaster in Northern China. Journal of Prataculture (06),121-128.

[2] Wang Xin, Wu Ji, Liu Chao, Yang Haiyan, Du Yanli & Niu Wensheng.(2018). Fault time series prediction based on LSTM recurrent neural network. Journal of Beijing University of Aeronautics and Astronautics (04),772-784. doi:10.13700/j.bh.1001-5965.2017.0285.

[3] Philander, S. G. H. (1985). El Niño and La Niña, Journal of Atmospheric Sciences, 42(23), 2652-2662. Retrieved Nov 20, 2022, from https://journals.ametsoc.org/view/journals/atsc/42/23/1520-0469_1985_042_2652_enaln_2_0_co_2.xml

[4] Zhang Chi,Song Xumei & Li Wei.(2008). Application of variable fuzzy evaluation method in flood disaster evaluation. Journal of Natural Disasters (05),34-39.

[5] Baohua Wang. (2008). Flood loss analysis, and evaluation model (a master's degree thesis, northeast agricultural university). https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD2009&filename=2008145316.nh

[6] Zhou H X. (2012). Genetic analysis of El Nino and La Nina phenomena and their impacts on climate change. The Computer Age (08),1-4.

[7] Bulletin of Flood and Drought Disaster Prevention in China 2021-2011, Water Resources Bureau of China,from: http://www.mwr.gov.cn/?type=1

[8] Lu Cong, Wei Yiming, Fan Ying, Xu Weishuan.(2002). Quantitative analysis model of disaster impact on national economy and its application. Journal of Natural Disasters (03),15-20

[9] Economic data, National Bureau of Statistics, from: http://www.stats.gov.cn/

[10] Drought data, U.S. Bureau of Economic Statistics,from: https://www.bea.gov/

[11] Drought data,NOAA,from:https://www.noaa.gov/

[12] Meteorological Data, NCEI,from: https://www.ncei.noaa.gov/maps/daily/

[13] Agricultural data, United States Bureau of Agriculture,from: https://www.usda.gov/

Appendix

Listing 1: AHP code

```
clear;clc
A =[1 3 7 3; 1/3 1 3 3;1/7 1/3 1 1/3;1/3 1/3 3 1]
[x,x] = size(A)
[V,D] = eig(A)
Max = max(max(D))
D == Max
[r,c] = find(D == Max , 1)
disp('Weight result');
disp( V(:,c) ./ sum(V(:,c)) )
CI = (Max - x) / (x-1);
RI=[0 0 0.52 0.89 1.12 1.26 1.36 1.41 1.46 1.49 1.52 1.54 1.56 1.58 1.59];
CR=CI/RI(x);
disp('CI=');disp(CI);
disp('CR=');disp(CR);
if CR<0.10
disp('CR < 0.10yes!');
else
disp('CR >= 0.10adjust!');
end

AHP- Fuzzy comprehensive evaluation method:
clc
clear
J = xlsread('D:\Desktop\hhc\TEMP.xlsx','D9:G12');
E=J';
[m,n]=size(E);
maxA = max(E,[],2);
minA = min(E,[],2);
G = maxA-minA;
A1=min(E(1 , :),[],2);
A2=max(E(2:3 , :),[],2);
A3=min(E(4 , :),[],2);
P=[A1',A2',A3'];
R=zeros(m,n);
for i=1:m
for j=1:n
```

```
R(i,j)=abs(E(i,j)-P(i))/G(j)
end
end
c = [ 0.5300 0.2575 0.0665 0.1461]
first =c*R;

Coefficient of variation-- Fuzzy comprehensive evaluation method:
B=[1,0,0.395532552438028,0.489784799782076;0.138750000000000,0,0.513750000000000,1;0,0.
WA=B;
[m,n]=size(WA);
maxA = max(WA,[],2);
minA = min(WA,[],2);
G = maxA-minA;
A1=max(WA(1:4 , :),[],2);
S=[A1'];
R=zeros(m,n);
for i=1:m
for j=1:n
R(i,j)=abs(WA(i,j)-S(i))/G(j)
end
end
x=mean(WA,2);
s=std(WA,0,2);
v=s ./ x;
v2=sum(v);
c=zeros(1,4);
for i=1:4
c(i)=v(i)/v2;
end
x = c
temp = c(3)
for j = 1:10
c(3)= c(3)-0.01
FF=c*R;
b(j)= FF(1)
end
c = x
for j = 1:10
c(3)= c(3)+0.01
```

```
FF=c*R;
a(j)= FF(1)
end
FF=c*R;

Part of Sensitive analysis code:
for j = 1:10
c(3)= c(3)-0.01
FF=c*R;
b(j)= FF(1)
end
c = x
for j = 1:10
c(3)= c(3)+0.01
FF=c*R;
a(j)= FF(1)
end
FF=c*R;

Principal component analysis:
clc
clear all
PA=xlsread('D:\Desktop\hhc\TEMP.xlsx','B2:K5');
a=size(PA,1);
b=size(PA,2);
for i=1:b
CA(:,i)=(PA(:,i)-mean(PA(:,i)))/std(PA(:,i));
end
CM=corrcoef(CA);
[V,D]=eig(CM);
for j=1:b
DS(j,1)=D(b+1-j,b+1-j);
end
for i=1:b
DS(i,2)=DS(i,1)/sum(DS(:,1));
DS(i,3)=sum(DS(1:i,1))/sum(DS(:,1));
end
T=0.9;
for k=1:b
```

```
if DS(k,3) >= T
com_num=k;
break;
end
end
for j=1:com_num
PV(:,j)=V(:,b+1-j);
end
newscore=CA*PV;
for i=1:a
score(i,1)=sum(newscore(i,:));
score(i,2)=i;
end
result=[newscore,score];
result=sortrows(result,-4);
```

Listing 2: The lingo source code

```
clc;
clear;
filename='D:\qq file\SOIreshape.xlsx';
Y=xlsread(filename,'Sheet1');
out=myMapminmax(Y,-42.6,34.8)
data = out'
figure
plot(data)
xlabel("Month")
ylabel("Cases")
title("Monthly Cases of SOI")
numTSTrain = floor(0.9*numel(data));
data_Train = data(1:numTSTrain+1);
dataTest = data(numTSTrain+1:end);
m_mean = mean(data_Train);
sig = std(data_Train);
dataTrainStandardized = (data_Train - m_mean) / sig;
XTrain = dataTrainStandardized(1:end-1);
YTrain = dataTrainStandardized(2:end);
Num_F = 1;
num_R = 1;
Num_Hidden = 200;
```

```
layers = [ ...
sequenceInputLayer(Num_F)
lstmLayer(Num_Hidden)
fullyConnectedLayer(num_R)
regressionLayer];
op = trainingOptions('adam', ...
'MaxEpochs',3000, ...
'GradientThreshold',1, ...
'InitialLearnRate',0.005, ...
'LearnRateSchedule','piecewise', ...
'LearnRateDropPeriod',125, ...
'LearnRateDropFactor',0.2, ...
'Verbose',0, ...
'Plots','training-progress');
net = trainNetwork(XTrain,YTrain,layers,op);
dataTestStandardized = (dataTest - m_mean) / sig;
XT = dataTestStandardized(1:end-1);
net = predictAndUpdateState(net,XTrain);
[net,ypred] = predictAndUpdateState(net,YTrain(end));
numTimeStepsTest = numel(XT);
for i = 2:numTimeStepsTest
[net,ypred(:,i)] =
    predictAndUpdateState(net,ypred(:,i-1),'ExecutionEnvironment','cpu');
end
ypred = sig*ypred + m_mean;
YTest = dataTest(2:end);
rmse = sqrt(mean((ypred-YTest).^2))
figure
plot(data_Train(1:end-1))
hold on
index_1 = numTSTrain:(numTSTrain+numTimeStepsTest);
plot(index_1,[data(numTSTrain) ypred],'.-')
hold off
xlabel("Month")
ylabel("SOI")

title("Forecast")
legend(["Observed" "Forecast"])
figure
```

```
subplot(2,1,1)
plot(YTest)
hold on
plot(ypred,'.-')

hold off
legend(["Observed" "Forecast"])
ylabel("SOI")
title("Forecast")
grid on;
subplot(2,1,2)
stem(ypred - YTest)
xlabel("Month")

ylabel("Error")
title("RMSE = " + rmse)

net = resetState(net);
net = predictAndUpdateState(net,XTrain);

ypred = [];
numTimeStepsTest = numel(XT);
for i = 1:numTimeStepsTest
[net,ypred(:,i)] =
   predictAndUpdateState(net,XT(:,i),'ExecutionEnvironment','cpu');
end
ypred = sig*ypred + m_mean;
rmse = sqrt(mean((ypred-YTest).^2))

figure
subplot(2,1,1)
plot(YTest)
hold on
plot(ypred,'.-')
hold off
legend(["Observed" "Predicted"])
ylabel("SOI")
title("Forecast")
subplot(2,1,2)
```

```
    stem(ypred - YTest)
    xlabel("Month")
    ylabel("Error")
    title("RMSE = " + rmse)
```