Fire frequency analysis and prediction based on back propagation neural network model and Long-Short Term Memory model

Xinyue Hu
Faculty of Science, Beijing Forestry University, Beijing, China
Email:1198159259@qq.com

Abstract. It is reported that there have been many fires in many places this year. This article wants to explore the factors affecting the occurrence of fires and make predictions to prevent the occurrence of forest fires. The correlation between forest indexes and fire frequency in Fujian, Guangxi, Heilongjiang, Hunan, Sichuan, Yunnan and Zhejiang provinces was analyzed based on forest indexes and fire frequency prediction in seven provinces from 1998 to 2018. According to the six relevant indexes of seven provinces in 20 years, the Pearson correlation coefficient between them and the number of fires was analyzed to find the correlation between the variables. The Grey system theory Model (GM (1, 1)) was established to predict the six related index values. The BP neural network model was established based on the historical and the predicted values were substituted into the Back Propagation (BP) neural network model to predict the number of fires. Then, a prediction model based on Long-Short Term Memory (LSTM) was constructed through deep learning library Keras, and the predicted value was substituted into the model to predict the number of fires. Zhejiang Province and Hunan Province were taken as examples to compare the accuracy of the two models' predictive values. It was found that the LSTM model had better fitting effect, and the LSTM model was finally used to predict the number of future fires. The forest ecological risk identification and prediction system based on Graphical User Interface system (GUI) was developed, the phase space reconstruction of one-dimensional data with time series characteristics was carried out, and the high-dimensional data were trained by using fuzzy neural network to realize data prediction and evaluation. This prediction can be applied to the forest prevention and control departments, and the annual fire situation can be predicted based on historical data, so as to prevent and reduce losses in advance.

1. Introduction
A forest fire broke out in Xichang, Liangshan Prefecture, Sichuan province in China on March 30, 2020, which seriously affected the lives of residents in Xichang. A forest fire also broke out in Liangshan on March 30, 2019. Several fire fighters not only died, but also suffered huge damage to social property and ecological environment. Forest fire is one of the natural enemies of forest construction and has uncertain and unstable influence on the social and economic development of forest areas. In China, not only Sichuan province is facing the problem of forest fire outbreak, but other provinces such as Fujian, Hei Longjiang and other densely forested provinces are also facing the test of forest risk. Therefore, it is of great significance to analyze and predict forest fires to reduce economic losses, protect forest resources and maintain ecological balance.
At present, fire prediction has been studied in various fields [1] and different methods. For example, Adaptive Neurofuzzy Reasoning System prediction model is used to predict electrical fires to ensure production safety.[2] However, most researches on forest fires are based on back propagation (BP) neural network model, which adopts linear regression model and makes prediction analysis based on machine learning, or fit influencing factors and fire frequency based on weighted Logistic regression model.[3] In the field of forest fire research, this paper explores more accurate prediction methods. The Long-Short Term Memory (LSTM) neural network model is adopted and compared with BP neural network model to produce higher accuracy of the prediction results.

2. Correlation analysis between forest ecological index and fire

Through data analysis and cleaning, six related indicators (forest land area, forest area, forest area artificial forest, forest coverage rate, total stock of living trees and forest stock) were collected and used as independent variables.[4] The number of fires and the affected area, which is taken as the dependent variable, measures Forest disaster.

2.1. Pearson's correlation coefficient between index and fire frequency

Pearson correlation analysis was carried out on independent variables and dependent variables using SPSS software to calculate the correlation coefficient between the index and forest fire.[5] The results are shown in Table 1.

| Province | Pearson correlation coefficient |
|----------|--------------------------------|
|          | Forestry land area | Forest area | Plantation area forest | Forest cover | Total stock of living trees | Forest stock |
| Fu Jian  | -0.920            | -0.999      | -0.993                  | -0.999       | -0.959                     | -0.957       |
| Guang Xi | -0.600            | -0.753      | -0.603                  | -0.753       | -0.443                     | -0.451       |
| Hei Longjiang | -0.508        | -0.863      | -0.941                  | -0.863       | -0.587                     | -0.603       |
| Hu Nan  | -0.224            | -0.456      | -0.219                  | -0.456       | -0.251                     | -0.154       |
| Si Chuan | -0.543            | -0.661      | -0.481                  | -0.543       | -0.537                     | -0.532       |
| Yun Nan | -0.894            | -0.974      | -0.909                  | -0.973       | -0.903                     | -0.895       |
| Zhe Jiang | -0.400           | -0.995      | 0.345                   | -0.995       | -0.95                      | -0.959       |

2.2. Significance analysis of correlation coefficient

Correlation coefficient is the quantity to study the degree of linear correlation between the number of fires and six related indicators. According to whether correlation coefficient is less than 0.05, the correlation between six indexes in different provinces and the number of fires is judged significant.

3. Establish grey prediction model

Forest fire is a time-related grey process, we establish grey model which we call it GM(1,1) model. According to the behavior characteristic data of the specific gray system, variables and the number of fires, the quantity and size of time series are predicted, and a grey prediction model based on the past to predict the future is established to predict the six indexes of the future time.[6]

3.1. Establishment of model

Suppose the original data sequence:[7]

\[ X(0)=(X(0)(1),X(0)(2),...,X(0)(n)) \]

Establish 1-AGO sequence to get
\[ X^{(1)} = (X^{(1)}(1), X^{(1)}(1) + X^{(0)}(2), \ldots, X^{(1)}(n-1) + X^{(0)}(n)) \]

GM(1,1) model equation:
\[ \frac{dX^{(1)}}{dt} + aX^{(1)} = u \]  

(1)

Assuming that \( \alpha = (a, \mu)^T \) by the least square method:
\[ \hat{\alpha} = (B^T B)^{-1} B^T Y_1 \]

\[ B = \begin{bmatrix} \frac{1}{2}(X^{(1)}(1) + X^{(0)}(2)) & 1 \\ \frac{1}{2}(X^{(0)}(2) + X^{(1)}(3)) & 1 \\ \vdots & \vdots \\ \frac{1}{2}(X^{(0)}(n-1) + X^{(1)}(n)) & 1 \end{bmatrix} \]

\[ Y_1 = \begin{bmatrix} X^{(0)}(2) \\ X^{(0)}(3) \\ \vdots \\ X^{(0)}(n) \end{bmatrix} \]

The solution of model equation (1) is obtained as:
\[ \hat{X}^{(1)}(k+1) = (X^{(0)}(1) - \frac{u}{a})e^{-ak} + \frac{u}{a} \]

3.2. Grey prediction results

Through residual prediction and correlation prediction, we predict the value of six indicators for the next few years. The results are shown in Table 2.

**Table 2.** Predicted values of indicators in each province.

| Province     | Year | Forestry land area | Forest area | Plantation area forest | Forest cover | Total stock of living trees | Forest stock |
|--------------|------|--------------------|-------------|------------------------|--------------|----------------------------|--------------|
| Fu Jian      | 2023 | 931.6144           | 838.8710    | 401.1746               | 69.0472      | 97218.4719                 | 89172.4050   |
|              | 2028 | 936.4593           | 862.8442    | 415.5155               | 71.0214      | 118525.4676                | 109018.2503  |
| Guang Xi     | 2023 | 1689.7138          | 1528.0155   | 874.8625               | 64.3130      | 88528.8032                 | 80212.7211   |
|              | 2028 | 1764.3370          | 1632.2188   | 1039.3316              | 68.7015      | 108168.1229                | 97584.8363   |
| Hei Longjiang| 2023 | 2568.8433          | 2024.1215   | 249.4229               | 44.5170      | 218803.2348                | 202524.3890  |
|              | 2028 | 2727.7898          | 2057.1544   | 253.3244               | 45.2400      | 241081.8536                | 223428.8769  |
| Hu Nan       | 2023 | 1271.7035          | 1112.5379   | 518.9099               | 52.5208      | 49496.4108                 | 42706.0612   |
|              | 2028 | 1283.6475          | 1171.6562   | 539.6789               | 55.3123      | 54820.4472                 | 46433.8922   |
| Si Chuan     | 2023 | 1923.7537          | 2512.2537   | 549.5345               | 39.7622      | 211719.2088                | 199664.8180  |
|              | 2028 | 2027.2757          | 2589.8969   | 604.6553               | 41.8982      | 229181.7744                | 215896.0112  |
| Yun Nan      | 2023 | 2651.8693          | 2254.0933   | 629.1816               | 58.9097      | 236784.5030                | 220550.7530  |
|              | 2028 | 2717.7356          | 2428.8811   | 781.6828               | 63.4789      | 264587.0730                | 249132.3677  |
| Zhe Jiang    | 2023 | 654.6540           | 617.6967    | 234.9477               | 60.6767      | 39536.5908                 | 35556.2635   |
|              | 2028 | 650.6107           | 628.3914    | 224.7868               | 61.7269      | 50393.3133                 | 45480.9526   |
4. Back propagation neural network predicts the number of future fires

The BP neural network model was established based on the data of six historical indexes and the number of fires. Then, according to the six index data obtained from the grey prediction in 2023 and 2028, the number of fires in 2023 and 2028 can be predicted by plugging in the established BP neural network model.[8][9]

4.1. Establishment of input-output layer model

1. For the input layer:

\[ net_j = \sum_{i=0}^{n} v_{ji} x_i, j = 1, 2, ..., m \]

output \( y_j = f(net_j), j = 1, 2, ..., m \)

2. For the output layer:

\[ net_k = \sum_{j=0}^{n} v_{kj} x_j, k = 1, 2, ..., i \]

output \( o_j = f(net_k), j = 1, 2, ..., i \)

where \( f(*) \) is excitation function:

\[ f(a) = \frac{1}{1+e^{-a}} \]

and \( f(a) \) is continuously derivable, \( f'(a) = f(a)(1-f(a)) \).[10]

4.2. Prediction results

The BP neural network model is used to predict the number of fires in each province in 2023 and 2028. The results are shown in Table 3.

Table 3. The number of fires Predicted by BP neural network model.

| Year | Fujian | Guangxi | Heilongjiang | Hunan | Sichuan | Yunnan | Zhejiang |
|------|--------|---------|-------------|-------|--------|--------|---------|
| 2023 | 49     | 502     | 87          | 45    | 240    | 69     | 41      |
| 2028 | 16     | 493     | 102         | 10    | 221    | 42     | 19      |

Based on the results of the GM model above, a model was established based on BP neural network model to predict the sub-images of fires in 2013 and 2018, and the results were compared with the real values. And then, we take Zhejiang province and Hunan province as examples, as shown in Figure 1 and Figure 2.

Figure 1. BP Model of Zhejiang Province.  
Figure 2. BP Model of Hunan Province.
4.3. Model accuracy analysis
Although the results are good, there are still errors between the predicted number of fires in 1998-2018 and the real value, and the accuracy of the prediction is not as high as expected in some provinces. Therefore, new neural networks such as LSTM neural network model are explored to predict the number of fires.

5. Long-Short Term Memory (LSTM) neural network model

5.1. Establish LSTM prediction model
The deep learning library Keras was used to construct the LSTM model. LSTM is more suitable for processing and predicting very long intermediate delay in time series and is suitable for complex nonlinear units to construct larger deep neural networks. The hidden layers of the LSTM model were five layers, and the number of neurons in each layer was 6, 64, 64, 64, 64 and 1, respectively. In the time step, the recursion number is set to 2.0; the training batch size is set as 20, which means that five groups of training samples are selected from the training set for training in each round. The number of training rounds is set to 1000, indicating that the training process repeats 1000 times; the optimization algorithm adopts Adam algorithm and dropout is set as 0.2.[11]

5.2. Prediction results
The LSTM neural network model[12] is used to predict the number of fires in each province in 2023 and 2028. The results are shown in Table 4.

| Year | Fujian | Guangxi | Hei Longjiang | Hunan | Sichuan | Yunnan | Zhejiang |
|------|--------|---------|---------------|-------|---------|--------|----------|
| 2023 | 47     | 508     | 92            | 97    | 178     | 55     | 47       |
| 2028 | 18     | 496     | 106           | 35    | 154     | 39     | 25       |

At the same time, the number of fires based on LSTM model is predicted. Similarly, take the provinces of Zhejiang province and Hunan province. Compared with the reality, the prediction results of LSTM neural network model are better than those of BP neural network model. The result is shown in Figure 3 and Figure 4.

[Figure 3. LSTM Model of Zhejiang Province.]
[Figure 4. LSTM Model of Hunan Province.]
6. Comparative analysis of the two models
We use determination of coefficient

\[ R^2 = \frac{\sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \]

Comparing the \( R^2 \) values of BP neural network model and LSTM neural network model to analyze the accuracy of their prediction. The \( R^2 \) closer to 1.0, the better the model fits the data. The comparison results are shown in the Table 5.

| Neural network model | Fujian | Guangxi | Hei Longjiang | Hunan | Sichuan | Yunnan | Zhejiang |
|----------------------|--------|---------|---------------|-------|---------|--------|----------|
| BP                   | 0.8925 | 0.6431  | 0.8542        | 0.7462| 0.6597  | 0.9414 | 0.9798   |
| LSTM                 | 0.9341 | 0.9193  | 0.9425        | 0.8239| 0.7385  | 0.9653 | 0.9801   |

In each province, the value of \( R^2 \) based on LSTM neural network model is higher than that based on the BP neural network model, and it is closer to 1.0. Therefore, the LSTM neural network model has a better fitting effect on the data, and the predicted value is closer to the real value.

7. Development prediction system
The forest ecological risk identification and prediction system based on MATLAB GUI has completed the development of forest ecological risk identification and prediction system, which can carry out phase space reconstruction of one-dimensional data with time series characteristics, and realize data prediction through RBF network training.[13] The fuzzy neural network[14] is used to train the high-dimensional data and realize the data prediction and evaluation. At the same time, the pSO-BP network quantity prediction system based on phase space reconstruction was developed, and the one-dimensional sequence was reconstructed in phase space, and the BP network training of particle swarm optimization was carried out. Renderings of the results are shown in figure 5 and figure 6.

![Figure 5](image1.png)  ![Figure 6](image2.png)

**Figure 5.** RBF network training model based on phase space reconstruction.

**Figure 6.** A risk assessment model based on fuzzy neural network.
8. Conclusions
The forest area, forest coverage rate, total stock of living trees and total stock of forests in Fujian province are significantly negatively correlated with the number of fires. The forest area and the total stock of living trees in Zhejiang province were negatively correlated with the number of fires. The indexes of other provinces are negatively correlated with the number of fires.

Since the LSTM neural network model has a better fitting effect, we use this model for prediction and get the prediction results. The model predicts that the number of fires in Fujian, Guangxi, Heilongjiang, Hunan, Sichuan, Yunnan and Zhejiang are 47, 508, 92, 97, 178, 55 and 47 respectively in 2023 and 18, 496, 106, 35, 154, 39 and 25 respectively in 2028.

After predicting the number of fires, all provinces can strengthen the prevention and control of fires according to forest indicators, and relevant departments can also appropriately predict the loss to reduce as much as possible according to the predicted results. This prediction technology can be applied to the actual prevention and control, and it can be made into a client for forest-related departments to carry out real-time prediction of fire to avoid excessive losses when fire occurs in various regions.

References
[1] Li Qing-gong, WANG Yue, LI Ji-fan, GAO You-qiang, Song Wen-hua 2020 Research on Fire Early Warning Method Based on Big Data Analysis [J] Journal of Nankai University (Natural Science edition) 53(04) 108-112
[2] L P Maguire, B Roche, T M McGinnity, L J McDaid 1998 Predicting a chaotic time series using a fuzzy neural network [J] Information Sciences 112(1)
[3] Yang Jinru 2019 Forest fire prediction based on linear regression algorithm [J] Communications World 26(04) 227-228
[4] Wang Wenye, Liu Jing, Jia Nan 2020 Forest fire prediction analysis based on neural network [J] Journal of armed police academy 36(06) 15-20
[5] Liu Xiaoshuang, XIONG Bin, ZENG Qingsong, Zhang Liang 2013 A study on fire and fire safety in China based on SPSS [J] Henan science 31(01) 105-107
[6] Ruishan Li, Yu Bi, Ke Yang, Cuihuan Ren 2019 Prediction Model of Forest Fire Based on Large Data Gra-Bp Neural Network [J] International Journal of Computational and Engineering 4(4)
[7] Yu Yong, Wu Qiong, Ding Bingqing 2017 Application of GM(1,1) Model in National Natural Disaster Prediction and Assessment Project -- A Case study of forest fire prediction [J] Project management technology 15(03) 24-26
[8] Wu Qiaoying, Chen Lihua, Li Xiaofeng, Wang Peng 2011 Forest ecosystem health prediction based on BP neural network Beijing Forestry University, 20111000-288X, 02-0150-05
[9] Ruishan Li, Yu Bi, Ke Yang, Cuihuan Ren 2019 Prediction Model of Forest Fire Based on Large Data Gra-Bp Neural Network [J] International Journal of Computational and Engineering 4(4)
[10] Fu Tianju 2016 Forest fire image recognition Algorithm and Implementation based on Deep learning [D] Beijing Forestry University
[11] Qu Yue 2018 Research on Machine Learning Algorithm for Environmental Data Prediction [D]. China University of Mining and Technology (Beijing)
[12] Xu Chunfang, Qiao Yuanjian, Li Jun 2020 Fire prediction method based on LSTM and RBF-BP Deep learning model [J] Journal of gilu university of technology 34(03) 53-59
[13] Wang Jichao, Zhang Jian, Ruan Zongli 2020 Design and Implementation of experimental system based on Matlab GUI calculation Method [J] Experimental technology and management 37(09) 130-134+138
[14] Ren Jieying 2019 Research on image-based fire identification Method based on convolutional neural network [D]. Xi ‘an University of Science and Technology