Rain fall predict and comparing research based on Arcgis and BP neural network

Guan Zimo [1, a], Tian Ziyu [2,b], Xu Yushi [3,c] and Dai Han[4,d]

[1]Beijing Forestry University, Beijing, P.R.China

[2]Beijing Forestry University, Beijing, P.R.China

[3]Beijing Forestry University, Beijing, P.R.China

[4]Beijing Forestry University, Beijing, P.R.China

[a]zimoguan1013@gmail.com,[b]tianziyu1234@outlook.com,

[c]chelsea12138@gmail.com,[d]daihan188@163.com

Keywords: Arcgis; BP algorithms; Rainfall forecasting; Kriging interpolation; Error analysis

Abstract. Based on the data of the rainfall from 24 base stations on the area of Chao River Basin in the near 54 years(range from 1958 to 2012) , the estimation is carried out by using Arcgis kriging interpolation and BP algorithms. And try to do the error analysis and the consequences comparing of these two results in the use of statistical approach. It is surely an innovative study which applies the advanced mathematical method in the rainfall research in the environmental field. After a serious of complicated data processing, the final conclusion is that it is feasible to apply BP neural network model to the forecast of rainfall, which is the correct method with high accuracy rate.

Introduction

In recent years, as the rapid development of geographic information technology and the high quality require of spatial data, the method and use of interpolation of spatial data attract increasingly attention. The theory and model method of spatial analysis, like the spatial statistical analysis [1] is widely under study. People carry out the discussion and study of the Meteorological elements interpolation as the specific agenda. Considering the rainfall is the decisive weather factor influencing the existence and development of living organisms, it is significant to improve the rainfall spatial information, which has the great effect to the analysis of regional hydrology and water resource, regional water resource management, drought and flood disaster management and the ecological environmental management[3]. The artificial neural network is a new hot research area in the artificial intelligence from the 1980s. In the theory of information processing, it abstract the brain neural networks, and build a sort of naïve model to knit different network, according to the different ways of connections.

In these recent decades, with the development of the research of global change, and the continued exploitation in the ecosystem model area which involves landscape, region, even global point, the need of spatial rainfall data considering as the parameter of environmental divisor is definitely rising, showing its further significant as well. The key to technique of weather data spatialization is the Spatial interpolation, which forms another weather model. In this new model, you can see the section of time and space process, showing the state from one unit to a whole area during a specific time span. Accordingly the raw weather data is provided with the connotation of time and space, so the derived data have ability to reflect natural geographical characteristics of climate [4].

Based on this point, we calculate the slope and the exposure in the Chao River Basin, using the Kriging Interpolation method in the GIS program. Then the estimation of this basin comes out from the data of the rainfall from 24 base stations on the area. Afterwards we take the mathematical algorithm---BP network model as the tool to predict the precipitation in other area according the data of the rainfall from 24 base stations along with the consideration of latitude and longitude, slope

information as well. In the end, we do the compare between the results from Kriging Interpolation and mathematical algorithm.

Prediction of rainfall based on Arcgis

Introduction of Kriging Interpolation. Kriging Interpolation is also known as kriging spatial interposition , one of the main content in the geostatistics, which is built on Variation function theory and structural analysis, widely accepted as a method to do the unbiased optimal estimate to the regional variate in a limited area[5]. Because of the statistical characteristic of geostatistics, in order to get the prediction result, even the prediction deviation, the interpolation from Kriging Interpolation is beneficial to evaluate the uncertainty of the prediction result.

Kriging is available to Geological statistics grid handling, widely using in the area of groundwater simulation, soil mapping and so on. The spatial attribute's transitional positioning in the spatial position is the first stuff the Kriging have to take into account, with the purpose to confirm a distance range. It have the ability to affect a new vertex insertion, also in this range, we can estimate the attribute value of this vertex insertion. By this approach it can offer a best linear unbiased estimator (a certain value at a position) mathematically. It is still a smooth interpolation method. The confidence level of the result of interpolating will go up when the data maker are quantity. For all these reasons, Kriging is our choice to forecast the rainfall.

The calculation of slope. DEM(Digital Elevation Model), is the digital expression of terrain surface's shape attribute information, also regarded as the digital description with spatial position and elevation attribute character[6].

Input the basin DEM into GIS, and operate the slope computational tool to get the DEM result. As shown below:

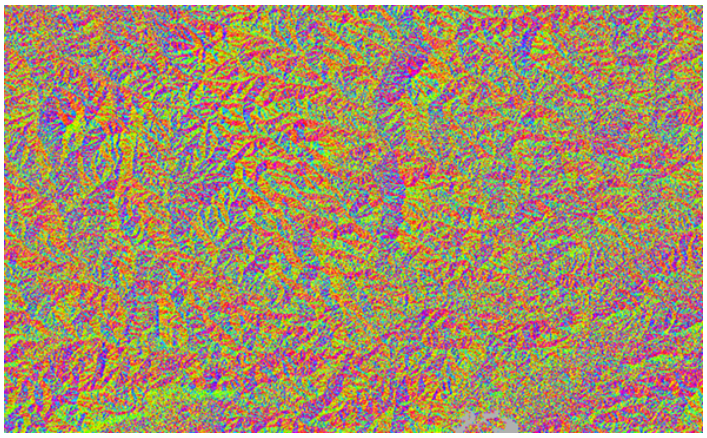

Image 1.Calculation result of slope.

The different colors in the picture represent the numerical range of slopes at different position.

Forecast of rainfall. It is similar to the slope calculation part. Use the Kriging interpolation toll in the toolbox to calculate the basin's rainfall year by year, based on the data from 24 base stations. We take the data from 2012 as an example, the calculation is shown below.

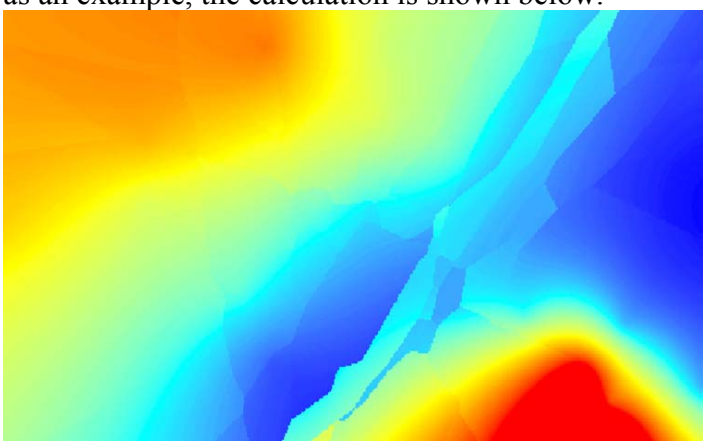

Image2. The predicted calculation result of rainfall in 2012

See the result of prediction of rainfall, by the recognition button in the every rain spot in the image. Different colors in the image represent the rainfall number range in different spot.

Prediction of rainfall based on BP neural network model

Introduction of model. Artificial Neural Network (ANN) is a new hot research area in the artificial intelligence from the 1980s. It abstracts the brain neural networks, and build a sort of naïve model to knit different network, according to the different ways of connections. It is also called neural network in engineering and academe. Neural Network (NN) is an operational model, consisted with a large number of nudes (or called neuron). Every nude stands for a kind of specific output function, called activation function. Every connection between two nudes stands for a weighted value through this connection signal, called weight, which is the equal of the memory of ANN. The output of ANN is varies according to the way of connection, weight value and activation function. And the network itself normally is the approach to some natural algorithm or function, or maybe an expression of a sort of logic strategy. The advantage of ANN is to realize the parallel processing of data, and the ability to learn on its own[7].

Construction of BP Artificial Neural Network and learning method. BP network is a kind of neural network with 3 or more than 3 stratums. Every neuron make the connection between higher-lower level. However, the neurons on the same level cannot contact each other.

The learning process of BP network can be divided into two phases:The first phase: input the known learning sample, and calculate the every output of every single neuron from the top level through network construction settings and the last time weight value and threshold value. The second phase: modify weight value and threshold value. Calculate the effect from the weight and the threshold to total error, from the last level, based on which, try to modify every couple of weight value and threshold value.These two phases is going to be repeated and alternative until converges.

The predicted model of rainfall. There are many accidents can affect rainfall. As we can know from the atmospheric water balance equation, the three factors, atmospheric precipitation and the evaporation from both land and sea, keep an atmospheric precipitation. The evaporation of an area is bound up with geographic location, and in the same area, it will change because of the different angle of sunshine. Therefore, we can get the predicted rainfall by handling the data which contains the latitude and longitude, the slope and the rainfall in surrounding areas, totally based on the assumption of long-run relative stability in this area.

The data of rainfall from 24 based stations is to predict other area's rainfall, with the consideration of the latitude and longitude, the slope. Take the mean annual precipitation as the example, and test repetitiously. We eventually be sure that the model is consisted by 4 input layer, 21 hidden layers of logarithmic S neurons and a output layer with linear neuron. It is a BP network with three layers.

Because the rainfall can be influenced much by accidents, and the data series is uncertainty, the error will go up if measured rainfall data is straightly adopted in this test. So the original data is normalized by the function premnmx firstly, in order to erase the difference between the data. After this process, the new data can be used to the test, having more chance to gain high predict accuracy. The mean annual precipitation appears through the function postmnmx by reverse calculating the predict result.[8]

Forecasting example. Take the rainfall data in 1967, 1988 and 2003 as the example, and the rainfall data from 24 observation station as the training method for the network, and the rainfall data from other 14 areas as the network examine data. The predict result is shown in chart 1. In the part of simulation value, only a few values are different from the measured data, while the others are totally same, with the accuracy of 80%. From this point, it is possible to predict rainfall with the use of BP network model, Artificial Neural Network.

Table 1.The measured rainfall and predict rainfall of 14 areas.

	Measured data			Forecast data		
Year	1967	1988	2003	1967	1988	2003
1	414.0948	499.6757	477.3333	499.6757	465.0357	496.5667
2	358.2639	373.9672	482.7583	373.9672	413.9322	457.6154
3	332.4504	389.2576	435.5	389.2576	394.6737	500.115
4	389.3555	362.5984	490.786	362.5984	378.7696	525.1631
5	385.7034	366.5277	491.1333	366.5277	385.6959	449.2287
6	378.7596	381.2922	485.2083	381.2922	475.5561	516.0523
7	402.692	449.8517	485.2833	449.8517	364.3562	401.8024
8	387.1674	389.8822	491.1333	389.8822	453.2094	461.6473
9	312.6726	369.1187	490.8167	369.1187	290.9194	386.0265
10	410.9438	468.9123	477.3333	468.9123	396.1096	556.2105
11	382.7365	344.2727	483.9917	344.2727	395.1114	510.9777
12	414.8417	423.1625	492.725	423.1625	449.4509	462.3449
13	419.6485	441.8912	477.3333	441.8912	458.1263	378.5098
14	415.96	445.2348	477.3333	445.2348	494.3951	526.761

Rainfall analysis to the result from Arcgis and BP neural network model

Analysis of computing result. After the interpolation calculus in Arcgis and the neural computation in Matlab, we can get the accuracy of the neural computation by comparing the two groups of data. We can get the comparatively exact predict rainfall in Arcgis, which has already built up a giant database with the progress of scientific methodology. And the neural computation also has a high accuracy compared with the data of Arcgis. We randomly choose 30 areas' rainfall data to compare the predict rainfall in the giant predict results. They are the data of 1967,1988and 2003, depicted by Excel.

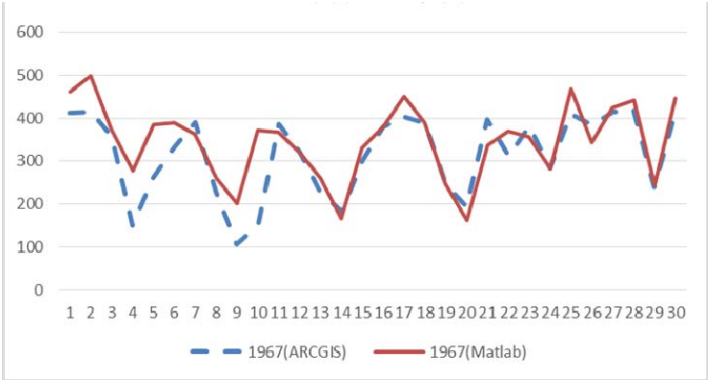

Image 3.Predict rainfall result in 1967.

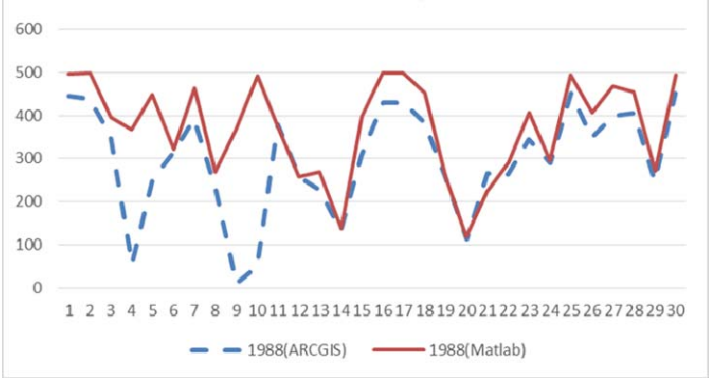

Image 4.Predict rainfall result in 1988.

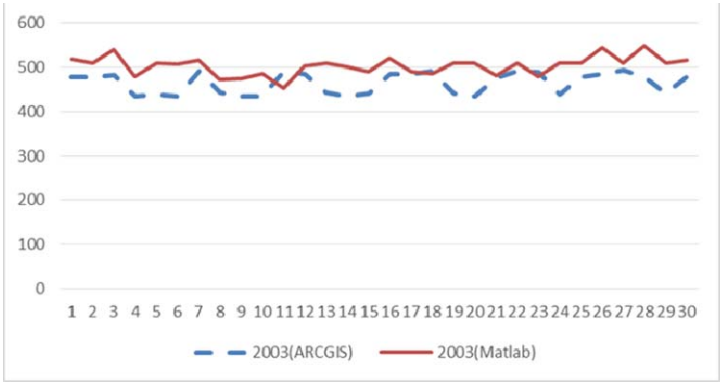

Image 5.Predict rainfall result in 2003.

From the images above, we can know that the result of interpolation calculus in Arcgis almost coincide with the result of the neural computation in Matlab, leaving out some unusual numbers.

The next step is to compare the accuracy of the result from neural network algorithm. The formula of accuracy is:

Accuracy=|value-real value|/real value×100% (11)

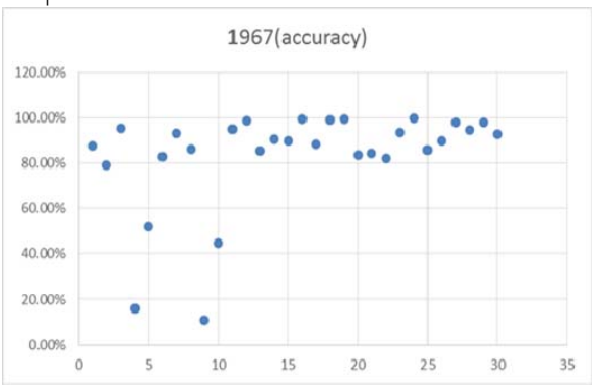

Image 6. Distribution of the accuracy result in 1967.

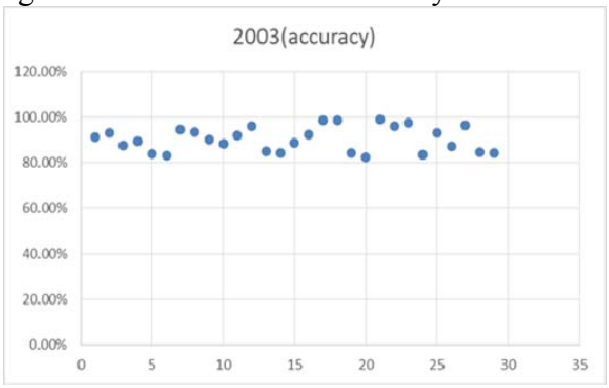

Image 7.Distribution of the accuracy result in 1988.

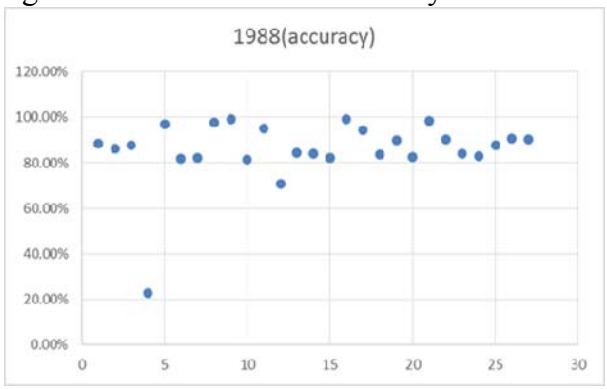

Image 8.Distribution of the accuracy result in 2003.

As we can see in these three images, the accuracy mostly distribute in the range of 80%-100%. Furthermore, after erasing the unusual values, the averages predict rainfall accuracy in these three years are 83.26%, 85.73% and 90.50%.

Consequently we can drop out the conclusion that neural network algorithm has a good effect and a high accuracy.

Error analysis. Neural network needs lots of training sample and testing sample as the guarantee. But the amount of data is limited, also without the consideration of extreme weather. So the network's predict result is sometimes unusual. Neural network has so much randomness. From this point, this model has some limitation in application.

Summary

Based on the calculation upon, we can draw the conclusion that: BP neural network can predict rainfall effectively. The adjustment of the number and type of factor on the input layer can influence the hidden layer straightly. In other words, to the all effective factors of rainfall prediction, the only thing we have to do is to input corresponding data into the input layer, so that we can build the hidden layer. By this approach, the purpose of rainfall prediction will realize.

While GIS has the foundation of environment, hydrology and soil analysis, it is a model with maturity and perfection. Compared with NN, it has a better control of all kinds of factors. More than that, GIS is easy to operate and less need of information, simplifying the process of earlier stage data reduction greatly.

Overall, it is feasible, accurate to apply mathematical algorithm into environmental research, with highly innovation. Although earlier stage data collection is complicated, the mathematical algorithm's intellectuality, which has the ability that weight value approximates the minimum error continually, is irreplaceable.

References

[1]Qin jianxin, & Zhang qingnian. Map visualization research[J]. (2000). 19(1):15-19.

[2]Houghton J. T. , Meira Filho L. G. , & Callander B.A.. et al.Climate Change 1995: the Science of Climate Change[M]. Cambridge: Cambridge University Press. (1995)

[3]ZHU Hui-yi, & JIA Shao-feng. Uncertainty in the spatial interpolation of rainfall data[J] Progress in Geography. (2004). 23(2):34 ～ 41.

[4]Cai Fu, Yu Huibo, Jiao Lingling, Tang Kai, Ming Huiqing, & Liu Bing. Comparison of Precision of Spatial Interpolation of Precipitation Factors: A Case Study in Northeastern China. (2006)

[5]Tang Guoan, & Yang Xin. Teaching method of Arcgis[M]. Beijing: Science Press. (2006)

[6]Wu Changlun, Liu Yu, & Zhang Jing. Geographic information system's theory, approach and application[M]. Beijing: Science Press. (2003)

[7]Niu Wenquan, & Li Jing. Rainfall's BP artificial neural network predict model and its application. (2001)

[8]Yu Jianli, & Li Ya. Grain Production predict model based on artificial neural network. (2005)