Prediction of Soil Moisture Based on BP Neural Network

Kewen Gao *, #, Ya Gao #, Miao Zhang #

School of Mathematical Sciences, Shanxi University, Taiyuan, Shanxi, 030006

* Corresponding Author Email: 252677587@qq.com

#These authors contributed equally.

Abstract. Grassland is the largest green ecological barrier in China, and reasonable grazing policies are key to ensuring people's livelihoods. The soil moisture data in this article is time series data and is related to multiple factors. On the premise of maintaining the same grazing strategy, first expand the collected data and use it as input. The L-BFGS algorithm was used to iteratively update the parameters, and the ReLU activation function was used to select the final model by comparing the Goodness of fit of models with different structures. A BP neural network model was established to predict the soil moisture at a depth of 40cm in 2022, providing theoretical reference for the research on grazing strategies of grasslands in pastoral areas in China

Keywords: BP Neural Network; L-BFGS Algorithm; Soil Moisture; Importance of Substitution.

1. Introduction

Grassland is the largest green ecological barrier in China, and it is the basic material for herdsmen to survive. Grassland plays an important ecological role in conserving soil and water, regulating soil erosion, and sandstorms. At present, the ecological situation of grasslands in China is still very severe. The problem of overloading and overgrazing on grasslands has restricted the development of grassland animal husbandry, and even affected the sustainable development of the entire national economy. Soil moisture is crucial for grassland grazing. Because soil moisture is an important factor affecting many aspects of grassland ecosystem, including community structure, soil carbon cycle and microbial growth, which have significant effects at different depths and time scales. There are currently some related studies, Fu Rong, Xie Luze, Liu Tao, Zheng Binbin, Zhang Yibo, Hu Shuai propose a soil moisture prediction model, based on the depth and water balance equation, which integrates the water balance equation with the seasonal ARIMA model, and introduces the depth parameter to consider the soil moisture at different depths.[1] Habiboullah Amina, Louly Mohamed Abdellahi used indices calculated from land surface spectral reflectances, namely NDVI Red, NDVI Green, NDVI Blue and NSMI to predict soil moisture content.[2] This paper uses a BP neural network model to predict future soil moisture.

2. Establishment and solution of the model

2.1. Analysis of Issue

Keeping the current grazing strategy unchanged, predict the soil moisture at a depth of 40cm from April to December 2022. Soil moisture itself is related to soil evaporation and precipitation (including precipitation, maximum daily precipitation and precipitation days). BP neural network model is built to predict the target.

2.2. Model Preparation

2.2.1. BP neural network model

The BP neural network is trained based on error backpropagation and learns the mapping relationship between inputs and outputs through input and output. From the perspective of topological structure, the BP neural network model includes input layer, hidden layer, and output layer:

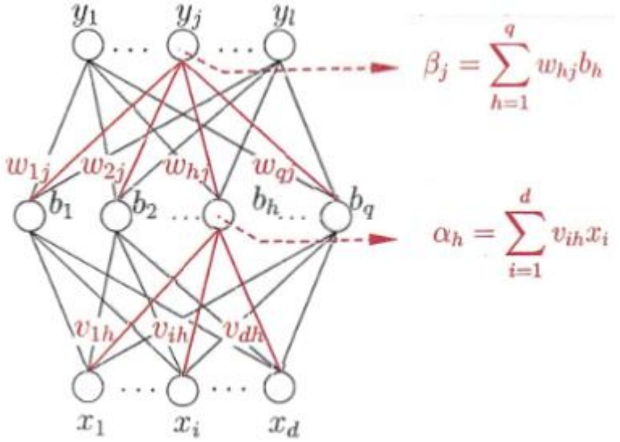

Figure 1. Neural network structure.

The network structure is shown in the Figure 1, given training set $D = \{(x_1, y_1), \ldots (x_m, y_m)\}$, $x_i \in R^d$, $y_i \in R^l$, v_{ih} represents the connection weight between the i-th neuron in the input layer and the h-th neuron in the hidden layer, w_{hi} and represents the connection weight between the h-th neuron in the hidden layer and the i-th neuron in the output layer. The input obtained by the h-th neuron in the hidden layer can be represented as:

$$\alpha_h = \sum_{i=1}^{d} v_{ih} x_i \tag{1}$$

The input received by the jth neuron in the output layer can be represented as:

$$\beta_j = \sum_{h=1}^{q} w_{hj} b_h \tag{2}$$

Where b_h is the output of the h-th neuron in the hidden layer [3] [4].

2.3. Model Building

2.3.1. Establishment of BP neural network model

SETP1: data preparation

First, the annual precipitation from 2012 to 2021, the soil evaporation of two different units, the maximum daily precipitation, the number of precipitation days and the precipitation on the corresponding time nodes are taken as inputs. Divide into training and testing sets in a 7:3 ratio. Normalize using the information from the training dataset, while normalize the test set according to the standards of the training set. During normalization, no information from the test set is applied to prevent loss of information. Some data is missing at certain time nodes, and exponential smoothing method is used to predict based on time, completing the data preprocessing stage.

SETP2: Determine neural network structure

The experiment shows that using too many hidden layers or neurons will make the problem solving more complicated, and will reduce the generalization ability of the model. Finally, the final model is selected only by comparing the Goodness of fit of the single hidden layer model and the double hidden layer model.

SETP3: Determination of weight

L-BFGS is the most commonly used method for solving unconstrained nonlinear programming problems. It is a quasi-Newton method, which has the advantages of fast convergence and low memory overhead [5] [6]. In this paper, L-BFGS algorithm is used to update parameters iteratively, and L2 regularization term is introduced, which is set to 0.0001, so that the second derivative is not zero and the overfitting phenomenon of BP neural network can be avoided simultaneously.

SETP4: Determine the activation function

The activation function in this article uses ReLU function (linear rectification function):

$$\mathrm{ReLU}(z) \equiv \max(0, z) = \begin{cases} z & if z \geq 0 \\ 0 & if z < 0 \end{cases} \tag{3}$$

The convergence speed of ReLU function is faster than that of Sigmoid and Tanh, without complex exponential calculation, which can improve operational efficiency, and there is no saturation region, which can alleviate the gradient vanishing problem in neural network training to a certain extent.

3. Results

First, preprocess the data and estimate a simple model, which contains only one hidden layer and only five neurons. Use the L-BFGS algorithm to iterate, use the ReLU activation function, and set the penalty parameter to 0.0001 by default L2. The results showed an accuracy of 70% with 27 iterations.

Draw a heat map of the neural network coefficient matrix:

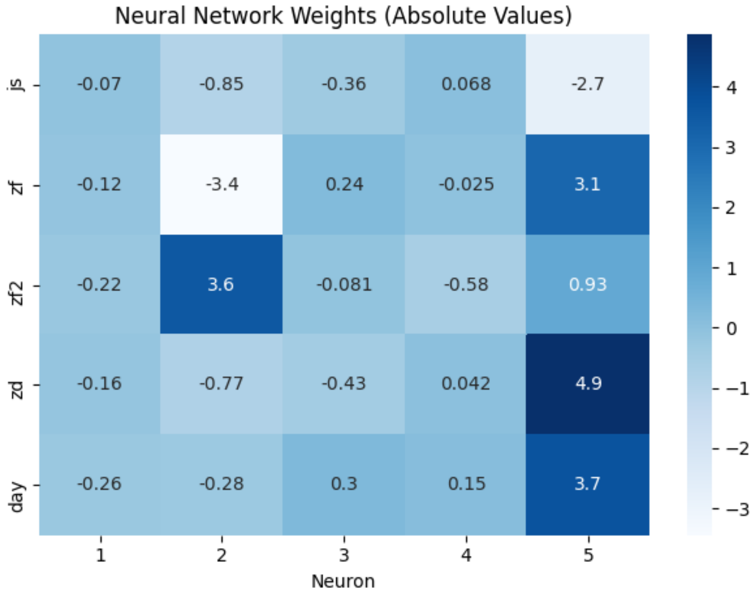

Figure 2. Neural network coefficient matrix heat map.

The heat map of the neural network coefficient matrix is shown in Figure 2, zf2 has the greatest impact on the second neuron, at 3.6.Heat maps can also be drawn based on the absolute values of the weight matrix:

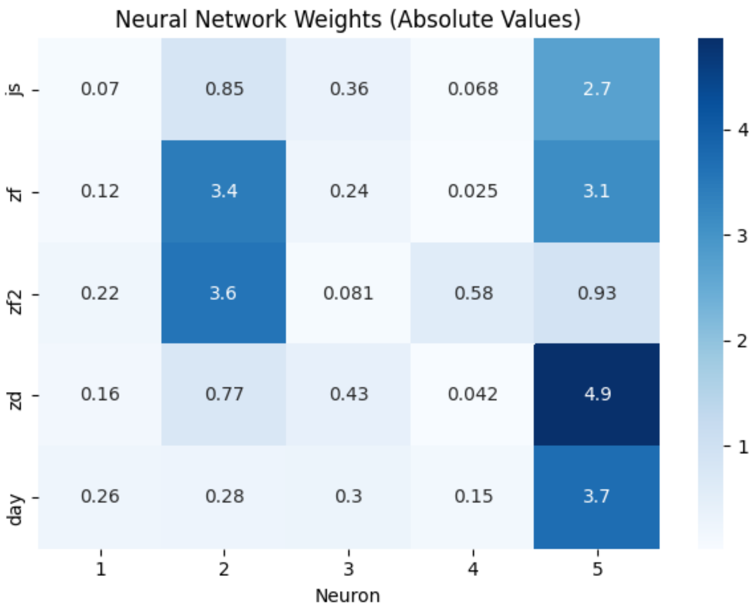

Figure 3. Neural network coefficient absolute value matrix heat map.

The heat map of the absolute value matrix of neural network coefficients is shown in Figure 3, the absolute weight values of the second and fifth neurons are relatively high. Below, we calculate the permutation importance of each feature variable. For a certain characteristic variable, the order of the observed values of this variable in the training set is randomly disordered to become a noise variable, which is called random replacement. Re estimate the model and calculate the Goodness of fit of the test set. Compared with the previous Goodness of fit, the difference is the replacement importance of this variable. The distribution of the replacement importance is investigated through the horizontal box graph [7].

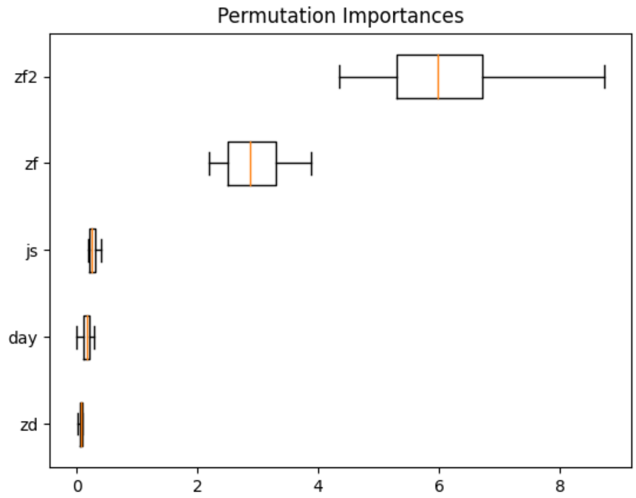

Figure 4. Variable Importance Box Chart.

The variable importance box chart is shown in Figure 4, it can be seen that the two most important variables are zf2 and zf, with values for each permutation importance being zd: 0.68, day: 0.162, js: 0.273, zf: 2.93, and zf2: 6.235, respectively.

Draw a partial dependency graph to understand the specific effects of feature variables on response variables [8]:

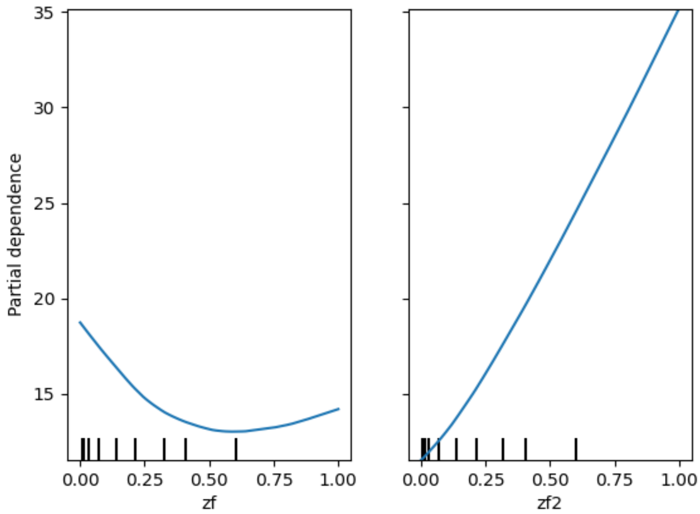

Figure 5. Partial dependence graph.

The partial dependency relationship is shown in Figure 5, it can be seen that the effect of zf on the response variable 40cm soil moisture shows a non-monotonic relationship of first decreasing and then increasing, while zf2 shows a monotonic increasing relationship with the response variable.

For the neural network with a single hidden layer, its network structure is determined by the number of neurons. By investigating the relationship between the number of neurons and the Goodness of fit of the test set, the optimal number of neurons is 27, which visually shows the impact of the number of neurons on the Goodness of fit of the test set [9]:

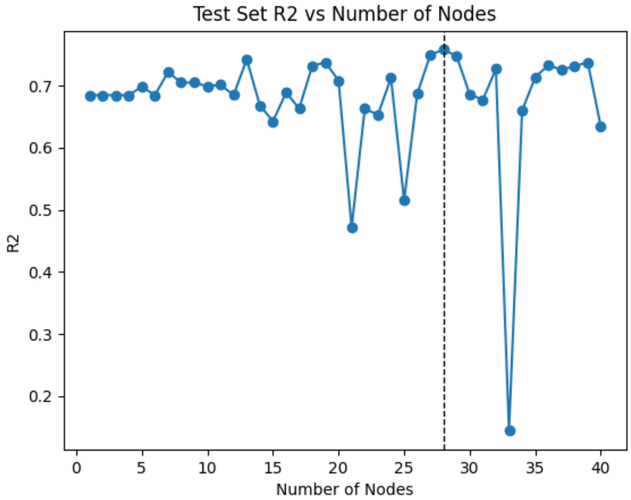

Figure 6. Relationship between number of neurons and Goodness of fit of test set.

The Goodness of fit between the number of neurons and the test set is shown in Figure 6.When there are fewer neurons, the Goodness of fit of the test set is low, and there is an under fitting phenomenon. When the number of neurons reaches 27, the Goodness of fit of the test set is the largest. Then continue to increase the number of neurons, and the Goodness of fit of the test set decreases instead, resulting in overfitting. Next, set the number of neurons to 27 and estimate again, with an accuracy of 85%. Next, consider a neural network model with two hidden layers. Using a loop to examine the number of neurons, the result is 4 neurons and 2 hidden layers [10]. Using a neural network with two hidden layers, the accuracy of the test set is 86.05%.

For JS, zf, zf2, zd, and day, the SARIMA time series model was used for prediction to obtain prediction data outside the sample points. By substituting it into the trained neural network, the predicted values of soil moisture at a depth of 40cm can be obtained, as shown in the table below.

Table 1. Forecast results.

Year	Month	40cm humidity(kg/m2)
2022	04	37.547
	05	39.692
	06	31.879
	07	37.454
	08	42.339
	09	46.251
	10	36.346
	11	47.439
	12	42.993

The predicted results of soil moisture from April to December 2022 are shown in Table 1.The Goodness of fit of the optimal model fitted by neural network reached 86.05%.

4. Conclusions

With the development of agriculture and animal husbandry, the grazing problem of grasslands has become increasingly severe. In order to better manage and protect the grassland ecosystem and provide important reference information for predicting soil moisture, the BP neural network model was used in this article to predict soil moisture at a depth of 40cm in 2022. In the process of constructing the neural network model, the data was divided into test sets and training sets, the sample data is normalized using the information from the training set, using the L-BFGS algorithm, which is a quasi-Newton algorithm that is superior to ordinary Newton iterative algorithms and suitable for smaller datasets. At the same time, penalty parameters are added in the process of building the network to effectively avoid the occurrence of overfitting. The optimal structure of the BP neural network is composed of two hidden layers and four neurons. The accuracy of its model is 86.05%. However, there is still room for improvement in this model. Different neural network architectures can be constructed, and different prediction models can be further explored by changing parameters. Multiple prediction models can also be combined or compared to construct ensemble learning to achieve the best prediction results.

References

[1] Fu Rong,Xie Luze,Liu Tao,Zheng Binbin,Zhang Yibo,Hu Shuai. A Soil Moisture Prediction Model, Based on Depth and Water Balance Equation: A Case Study of the Xilingol League Grassland [J]. International Journal of Environmental Research and Public Health, 2023, 20(2).

[2] Habiboullah Amina, Louly Mohamed Abdellahi. Soil Moisture Prediction Using NDVI and NSMI Satellite Data: ViT-Based Models and ConvLSTM-Based Model [J]. SN Computer Science, 2023, 4(2).

[3] Chang Daqing, Sun Shiliang, Zhang Changshui. An Accelerated Linearly Convergent Stochastic L-BFGS Algorithm [J]. IEEE transactions on neural networks and learning systems, 2019, 30(11).

[4] B. Liu, R. Wang, G. Zhao, X. Guo, Y. Wang, J. Li, S. Wang. Prediction of rock mass parameters in the TBM tunnel based on BP neural network integrated simulated annealing algorithm [J]. Tunnelling and Underground Space Technology incorporating Trenchless Technology Research, 2020, 95(C).

[5] Hongjian Wang, Hartmut Gemmeke, Torsten Hopp, Jürgen Hesser. Accelerating image reconstruction in ultrasound transmission tomography using L-BFGS algorithm [P]. Medical Imaging, 2019.

[6] Hou Yali, Wang Qunwei, Tan Tao. Prediction of Carbon Dioxide Emissions in China Using Shallow Learning with Cross Validation [J]. Energies, 2022, 15(22).

[7] Habibnia A., Rahimipour Gh. R., Ranjbar H. Introducing Improved Performance of Boxplot (New Method) in Estimating the Threshold (Separating Anomalies from the Background) [J]. Geochemistry International, 2022, 59(13).

[8] Shi Haoze, Yang Naisen, Yang Xin,Tang Hong. Clarifying Relationship between PM2.5 Concentrations and Spatiotemporal Predictors Using Multi-Way Partial Dependence Plots [J]. Remote Sensing, 2023, 15(2).

[9] Roy Jagabandhu,Saha Sunil. Ensemble hybrid machine learning methods for gully erosion susceptibility mapping: K-fold cross validation approach [J]. Artificial Intelligence in Geosciences, 2022, 3.

[10] Hou Yali, Wang Qunwei, Tan Tao. Prediction of Carbon Dioxide Emissions in China Using Shallow Learning with Cross Validation [J]. Energies, 2022, 15(22).