Article

Soil Moisture Prediction Based on the LSTM Neural Network Integrating Particle Swarm Optimization

Xiaoyi Huang [1], Jianlong Wu [2] and Jizheng Chu [3,*]

[1] Beijing University of Chemical Technology; romonh@163.com
[2] Beijing University of Chemical Technology; buctwujl@mail.163.com
[2] Beijing University of Chemical Technology; chujz@mail.buct.edu.cn
* Correspondence: chujz@mail.buct.edu.cn; Tel.: +86(010)64435711

Abstract: Soil moisture is an important factor affecting the plant growth. For a long time, the convenience, timeliness and accuracy of soil moisture monitoring have been limited due to the backward of observation methods and equipment. Therefore, the quantitative prediction of soil moisture has become a difficult problem. Aiming at the problems of high erection cost, easily damaged sensors and low measurement accuracy of the existing fixed sensor soil moisture monitoring system, a soil moisture prediction model based on the long short term memory neural network (LSTM) integrating the particle swarm optimization (PSO) (PSO-LSTM) is designed and implemented. The hyperparameters of the LSTM network can be obtained based on the excellent global search ability of the PSO algorithm. According to the meteorological data and soil moisture data of Haidian Park in 2019, the long short term memory(LSTM) neural network based prediction model is constructed with input vectors of surface temperature, average temperature, evaporation, sunshine hours, precipitation and average wind speed, and the output vector of soil relative humidity. The results show that compared with the back propagation(BP) neural network, the Elman neural network and the LSTM neural network, the proposed PSO-LSTM model has higher prediction performance.

Keywords: Soil moisture prediction; LSTM; PSO

1. Introduction

There is a shortage of water resources in China, and plant irrigation requires a large amount of water. In order to reduce unnecessary water waste, it is particularly important to irrigate plants reasonably. Soil moisture is a key factor affecting plant growth, and its monitoring and prediction is of great significance to the healthy growth of plants Kerr[1]. Yihyun Kim et al. [2] used remote sensing technology to inverse modeling for obtaining farmland moisture information, but remote sensing technology can only collect soil moisture information at a depth of about 5 cm, and the cost is too high, so it is difficult to Large-scale promotion. Dumedah[3] et al. used soil hydrodynamics and soil water balance method to estimate the change of moisture content, which required a large amount of data to measure, more parameters and more complicated calculation process. Seneviratne[4] et al. used the characteristics of surface soil water coupling to root zone soil water through diffusion process for predicting the profile soil water through surface soil state measurement. But this method had strong regional characteristics, which was less versatile and stable. Al-Mukhtar[5] used the soil and water assessment tool(SWAT) model based on the comprehensive soil moisture data of the upper reaches of the Spree River to simulate the soil moisture in the root zone of the upper Spree River. Zeng Li[6] used the penalized linear regression integration method to predict the drought index in the Northeast region. Due to the strong limitations of the linear regression method. Moreover, the factors affecting crop water demand are very complex, so its prediction accuracy is not high. In practical application of the above methods, various parameters that meet the conditions need

 doi:10.20944/preprints202212.0141.v1

to be obtained through experimental measurement or statistical analysis. Due to the temporal and spatial variability of natural conditions, the practical extension of the model is limited.

Soil moisture is a nonlinear, temporal and spatial heterogeneous dynamic uncertain variable, so neural networks have been widely used in soil moisture prediction in recent years. At present, the static BP neural network is mostly used for the soil moisture prediction[7-9]. However, the complexity of static BP neural networks increases with the increase of the system order, which makes the network learning more efficient. The convergence speed is slowed down, resulting in too many network input nodes and difficulties in training. Tabari et al.[10] studied the adaptive fuzzy inference system and the support vector machine (SVM) to establish the nonlinear relationship between ET 0 and meteorological factors, and the simulation accuracy was better. Fung[11] et al. used the wavelet decomposition and the fuzzy support vector to predict soil moisture. However, the calculation of the SVM parameters and kernel functions is complex, which affects the prediction accuracy. Al-Mukhtar et al.[12] used the Elman dynamic neural network to predict the soil moisture content in the Spree River Basin. Sankhadeep et al.[13] used the improved flower pollination algorithm(MFPA) to train artificial neural network for predicting soil water content. However, the flower pollination algorithm is easy to fall into the local extreme value, and the stability needs to be further improved.

The LSTM neural network model has the characteristics of self-organization, self-adaptation and self-learning, which is suitable for the prediction of highly nonlinear and dynamic time series. It only needs less preparation to map the internal rules between inputs and outputs, and it is not easy to fall into local minimum, and has good generalization ability. Compared with the ordinary neural network, each hidden layer unit of the LSTM neural network is not independent of each other. Each hidden layer is not only related to each other, but also related to the time series input before the time received by the hidden layer unit. This feature is extremely helpful for processing time series-related data. When facing long sequences, the RNN is prone to gradient explosion and gradient disappearance problems. However, the LSTM network can remember long-term dependencies, which can overcome this problem [17].

The parameter settings of the LSTM model have a direct impact on the prediction accuracy of soil relative humidity. Because the addition of the gate structure increases the number of parameters, the LSTM model needs to constantly adjust the parameters to achieve the optimal effect of the prediction model. The PSO algorithm is a biologically inspired swarm intelligence optimization algorithm, which originated from the research on bird predation[14]. In this paper, the PSO algorithm is used to optimize the parameters with the number of hidden layer units, the number of iterations and the dropout coefficient in the LSTM model to obtain the optimal hyperparameters of the LSTM model and determine the optimal parameters of the soil relative humidity prediction model. This process has no influence of artificial parameter adjustment, small randomness and relatively stable optimization effect.

2. The LSTM neural network

The internal structure of the LSTM is mainly to control the transmission state through the gated state. It can remember key information and forget the unnecessary information in long sequence information. This is a very effective network for long sequence data such as network security situation prediction. The structural unit of the LSTM is shown in Fig. 1. It mainly includes the input gate, the output gate, the forget gate and the self-connected memory cell state value (candidate value). The main functions of the input gate, the output gate and the forget gate in the LSTM model are to control the transmission of information, how much information can be transmitted to the current neuron, and how much information of the current neuron is allocated to the next neuron. The values of x_t, h_{t-1} and C_{t-1} are related.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2022 doi:10.20944/preprints202212.0141.v1

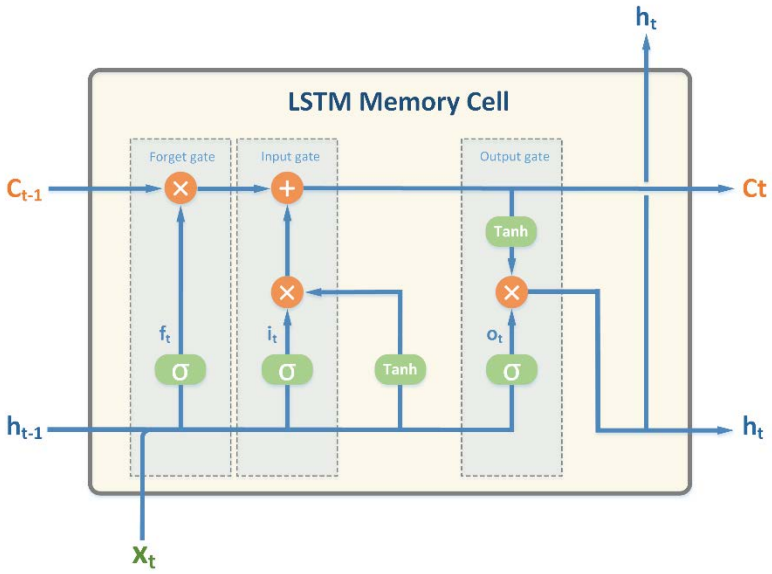

Figure 1. LSTM network structure diagram.

1. The forgetting gate is the output of the previous unit and the input of this unit. Through the sigmoid function, the value in [0, 1] generated by the function controls the degree to which the state of the previous unit is forgotten as shown in Eq.(1).

$$f_t = \sigma(W_f \bullet [h_{t-1}, x_t] + b_f), \tag{1}$$

2. The input gate determines what values to keep and what values to update. After appling the sigmoid function to decide what to update, the tanh function is applied to create a new vector of candidate values.

$$i_t = \sigma(W_i \bullet [h_{t-1}] + b_i), \tag{2}$$

$$\tilde{C}_t = tanh(W_c \bullet [h_{t-1}, x_t] + b_c), \tag{3}$$

Updating the old cell state C_{t-1} to the new cell state C_t implements this process with the following formula:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t, \tag{4}$$

where $f_t * C_{t-1}$ is the information that needs to be discarded, and $i_t * \tilde{C}_t$ is the new candidate value, which changes according to the extent of updating each state.

3. The output gate use a sigmoid layer to determine the output part of the unit state, then process the unit state through tanh (to get a value between -1 and 1) and multiply it with the output of the sigmoid gate, and finally output the section to determine the output.

$$O_t = \sigma(W_o \bullet [h_{t-1}, x_t] + b_o), \tag{5}$$

$$h_t = O_t * \tanh(C_t), \tag{6}$$

Where x_t and h_t represent the input vector and output vector, respectively. f, i, and O represent the forget gate, input gate and output gate, respectively. C_t and C_{t-1} represent the previous moment and the current unit state, respectively. h_{t-1} and h_t represent the outputs of the previous and current hidden layer units respectively. σ represents the sigmoid activation function, tanh represents the tangent function, and W and b represent the weight matrix and bias vector, respectively.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2022 doi:10.20944/preprints202212.0141.v1

Since the selection of key parameters of the LSTM has a great influence on the accuracy of load forecasting, it is necessary to select these parameters reasonably. The PSO algorithm has the advantages of simple structure, high accuracy and fast convergence , which has certain advantages in dealing with nonlinear and multi-variable problems and can be used to select parameters of the LSTM model[18].

3. The prediction model based on the PSO-LSTM

3.1. The PSO algorithm

The PSO algorithm is used to optimize the hyperparameters of the LSTM for obtaining the optimal parameter combination of the entire network.

The principle of PSO updating weights: The D parameters including weights and thresholds that need to be updated in the neural network are formed into a vector and considered to be a particle and represent a "position" [15].

Initializing the particle to randomly generate N particles is expressed as:

$$X_i = (x_{i1}, x_{i2}, \dots\dots, x_{iD}), \mathrm{i} = 1,2, \dots\dots \mathrm{N}, \tag{7}$$

The "velocity" of each particle is also a D-dimensional vector, which is denoted as:

$$V_i = (v_{i1}, v_{i2}, \dots\dots, v_{iD}), i = 1,2, \dots\dots, N, \tag{8}$$

The optimal position searched by the i-th particle is called the individual extremum, which is recorded as:

$$\mathrm{Pbest} = (p_{i1}, p_{i2}, \dots\dots, p_{1D}), i = 1,2, \dots\dots, N, \tag{9}$$

The optimal position searched by the entire particle swarm is the global extremum, which is recorded as:

$$\mathrm{gbest} = (p_{g1}, p_{g2}, \dots\dots, p_{gD}), \tag{10}$$

Upon finding the individual extrema and the global extremum, the particle updates velocity and position according to the following formula:

$$\begin{cases} V_{id}(t+1) = V_{id}(t) + c1 * r1\big(P_{id}(t) - x_{id}(t)\big) + c2 * r2(p_{gd}(t) - x_{id}(t)) \\ x_{id}(t+1) = x_{id}(t) + v_{id}(t+1) \end{cases}, \tag{11}$$

The flow process of the PSO is shown in Fig.2.

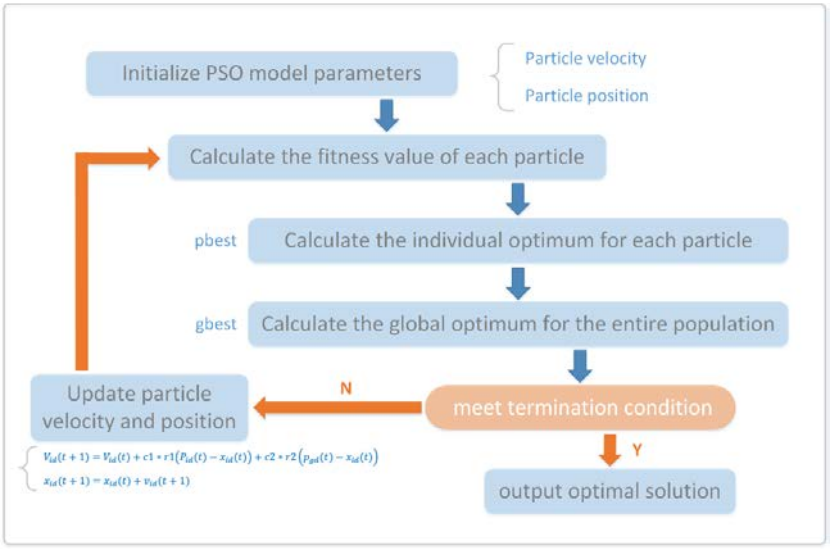

Figure 2. The flow chart of the PSO algorithm.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2022 doi:10.20944/preprints202212.0141.v1

3.2. The PSO-LSTM model

The structure of PSO-LSTM model is shown in Fig. 3.

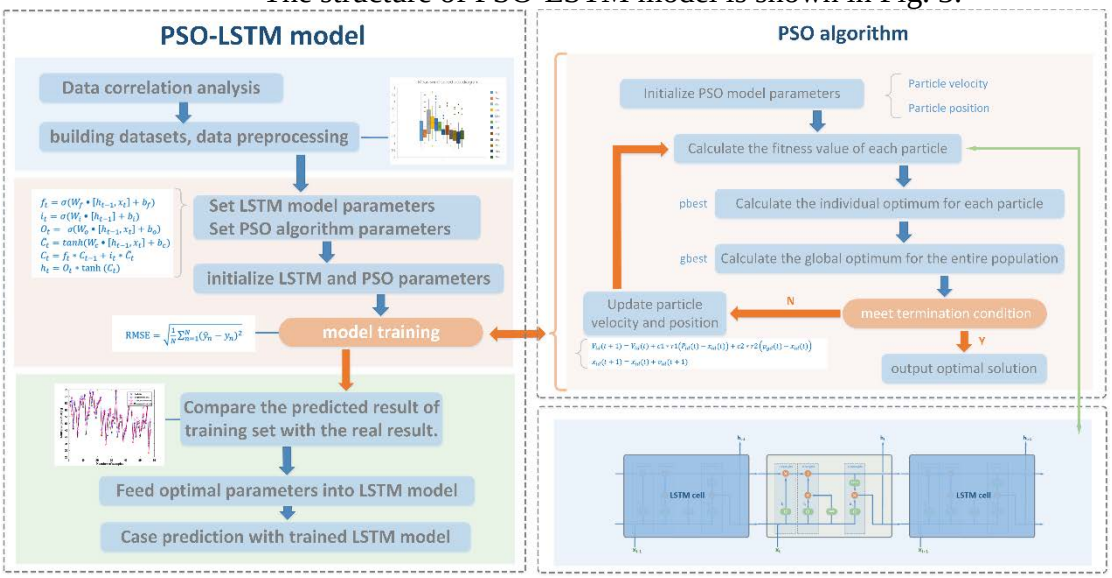

Figure 3. the structure of PSO-LSTM.

The process of the PSO-LSTM is described as follows:

Step 1. Taking the number of LSTM hidden layer units, the number of iterations, and the dropout coefficient as the optimization object, and initializing the position information of the particle according to the preset range.

Step 2. Initialize the particle swarm, divide the training set and the test set, and input the initialization parameters in step 1 into the LSTM network for training, and ue the model prediction error is used as the fitness value of the particle.

Step 3. Compare the fitness value of each particle and the best position it has experienced, determine the optimal position of the particle, update the speed and position of the particle according to Eq. (5), and calculate the fitness value of a new round of particles.

Step 4. When the search process reaches the preset maximum number of iterations, or the fitness value of the particle no longer changes significantly with the number of iterations, the updating is stopped to obtain the sample batch of the LSTM model, the number of hidden layer units, the learning rate and the number of iterations value.

Step 5. Input the values obtained in Step 4 into the LSTM model for training and prediction.

4. Case Study

4.1. Data processing and analysis

In order to verify the effectiveness of the PSO-LSTM model, the daily meteorological data and soil moisture of Haidian Park in 2019 are used to predict the land moisture. The land moisture data comes from the installed land information sensors on the ground. At the same time, the meteorological data is comprehensively collected. The meteorological data includes the daily average wind speed, average air pressure, sunshine hours, surface temperature, precipitation, evaporation in 2019, volume and average temperature. Due to the variety of meteorological data, it is not conducive to the optimization of the neural network. In order to determine the degree of correlation between each meteorological data and soil moisture, the grey relational analysis (GRA) method is introduced.

4.1.1. Data processing

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2022 doi:10.20944/preprints202212.0141.v1

Data preprocessing can improve the mining value of soil moisture and reduce the impact of sample errors on the prediction model. The data preprocessing method adopted in the experiment mainly include screening outliers and data supplementation. Among them, SPSS is used to draw soil moisture box plot, and outliers are screened according to the box plot. Fig. 4 is the boxplot in the process of data cleaning for the average wind speed, in which the discrete points are the points with large errors in the data, and these data are eliminated.

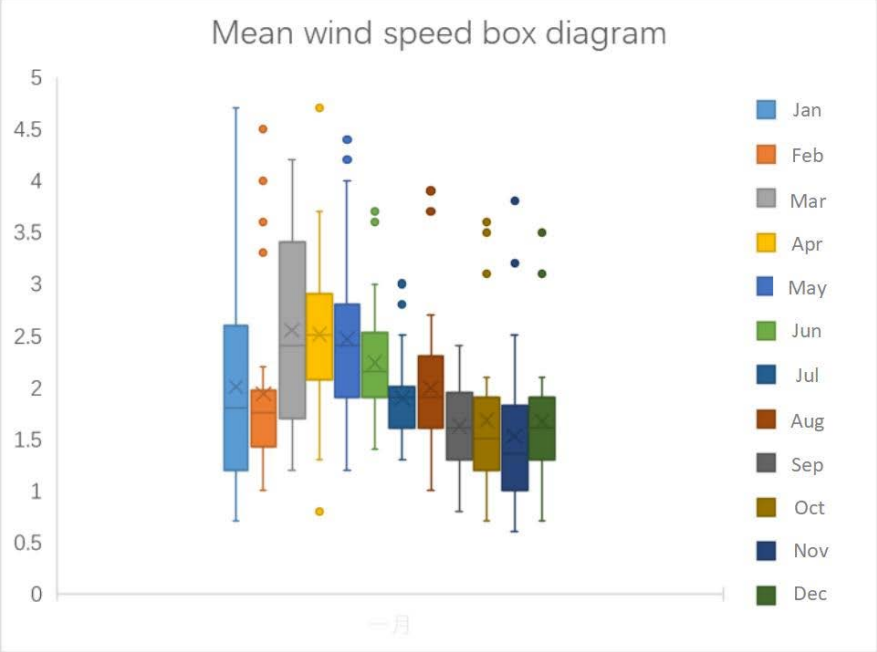

Figure 4. Box plot of selected data.

Since soil moisture data is a continuous time series, which is closely related to the adjacent values, kriging interpolation is used to fill in the vacancies in the data to make the data complete. Similarly, the data of average air pressure, sunshine hours, average surface temperature, precipitation, evaporation and average air temperature, and soil relative humidity are also subject to the same process of data cleaning. The abnormal data processing of meteorological data is shown in Table 1.

Table 1. Exception data handling list.

Case	Causal variable	Variable influence	Variable	Variable norm	Anomaly indicator
224	average wind speed	0.330	3.9	1.663	3.431
181	Small evaporation	0.451	10.0	4.653	3.369
139	average wind speed	0.338	2.8	1.039	3.301
221	precipitation	0.787	88.5	9.565	2.892
108	average wind speed	0.480	4.1	1.663	2.716
204	sunshine hours	0.489	10.9	1.349	2.616
228	precipitation	0.410	9.6	0.405	2.499
180	Small evaporation	0.336	8.4	4.653	2.451
80	average wind speed	0.557	4.1	1.663	2.340
30	average wind speed	0.439	3.0	1.471	2.333
45	Relative humidity	0.409	73	40.41	2.108

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2022 doi:10.20944/preprints202212.0141.v1

225	average wind speed	0.685	2.6	1.231	2.021

In order to avoid the inconsistency of the input data dimension, it also needs to be normalized as is shown in Eq.(12).

$$x' = \frac{x-\min(x)}{\max(x)-\min(x)}, \tag{12}$$

4.1.2. The Grey relation analysis (GRA)

The GRA refers to a method of quantitative description and comparison of the development and change of a system [19]. The closer the curves are, the greater the correlation between the corresponding sequences and vice versa. TheGRA method makes up for the shortcomings caused by using mathematical statistics methods for systematic analysis. The GRA is applied in the size of the sample and the regularity of the sample. The calculation amount is small, and there is no mismatch between the quantitative results and the qualitative analysis results.

1. Determine the system reference sequence and comparison sequence.

   To analyze an abstract system or phenomenon, it is necessary to select the data sequence that reflects the characteristics of the system behavior as the mapping quantity of the system behavior, and use the mapping quantity to indirectly represent the system behavior. According to the experimental requirements, the relative soil humidity is selected as the reference sequence of the system, and seven groups of data including average wind speed, average air pressure, sunshine hours, surface temperature, precipitation, evaporation and average temperature are respectively used as the comparison sequence in order.

   The reference sequence is set to

   $$\mathrm{Y} = Y(t)|t = 1,2,\dots,365;, \tag{13}$$

   The reference sequence is set to

   $$X_i = X_i(t)|t = 1,2,\dots,365, i = 1,2,\dots,7, \tag{14}$$

2. Normalized data.
   The dimensions of data in each factor column in the system are different, so correct conclusions cannot be drawn during comparison. Therefore, in the analysis of grey relational degree, it is generally necessary to carry out dimensionless processing of the data. There are two main methods of dimensionless including initial value processing and mean value processing. This paper adopts the initial value processing method as shown in Eq.(15).

   $$x_i(t) = \frac{x_i(t)}{x_i(1)}, t = 1,2,\dots,365, i = 1,2,\dots,7, \tag{15}$$

3. The correlation coefficient is calculated as follows.

   $$\xi_i(t) = \frac{\min_i \min_t |y(t)-x_i(t)| + \rho \max_i \max_t |y(t)-x_i(t)|}{|y(t)-x_i(t)| + \rho \max_i \max_t |y(t)-x_i(t)|}, \tag{16}$$

   Among them, ρ is called the resolution coefficient. The smaller the ϱ, and the greater the resolution. Generally, the value range of ρ is (0, 1), usually ρ=0.5.

4. Calculate the degree of correlation.
   The mean value of the correlation coefficient at each moment is used as the correlation degree between the reference sequence and the i-th group comparison sequence, which is shown in Eq.(17).

   $$\mathrm{ri} = \frac{1}{n}\sum_{t=1}^{n} \xi_i(t), \ t = 1,2,\dots 365, \tag{17}$$

5. Relevance ranking

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2022 doi:10.20944/preprints202212.0141.v1

The degree of association is sorted by size. If r1>r2, it means that the association between X1 and the reference sequence Y is stronger than that of X2.

The correlation between each group of data and soil relative humidity after the GRA analysis is shown in Table 2 and Fig.5.

Table 2. Correlation of meteorological data with soil relative humidity.

Reference sequence (X)	Average wind speed	Average air pressure	Sunshine hours	Surface temperature	Precipitation	Evaporation	Average temperature
Relevance (r)	0.9025	0.3465	0.9119	0.9297	0.9030	0.9239	0.9271

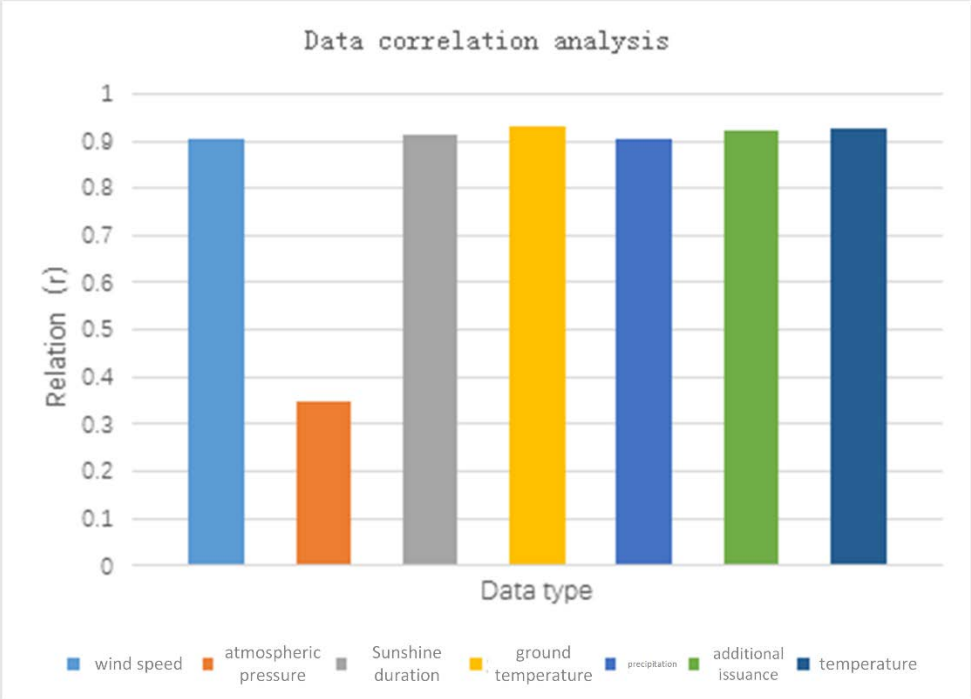

Figure 5. Data correlation analysis.

It can be seen form Table 3 and Fig.5 that the highest correlation with soil relative humidity is surface temperature, while the correlation between average air pressure and soil relative humidity is the lowest, only 0.35, indicating that the average air pressure has little effect on the change of soil moisture. For other meteorological data with average wind speed, sunshine hours, precipitation, evaporation, average temperature and soil relative humidity, the correlation is above 0.9. Therefore, the average air pressure data can be removed in the next experiments to simplify the neural network. structure.

4.2. Simulation Analysis

Divide the processed data into two parts with selecting the data from January 1, 2019 to October 31, 2019 as the training set, and the data from November 1, 2019 to December 31, 2019 as the test set. The time interval for meteorological data statistics is 1 day.

This paper selects root mean square error (RMSE), mean square error (MSE), mean absolute percentage error (MAPE) and mean absolute error (MAE) as evaluation indicators to quantitatively evaluate the prediction effect of the prediction model. Among them, the smaller the values of the RMSE, the MSE, the MAPE, and the MAE, the smaller the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2022 doi:10.20944/preprints202212.0141.v1

deviation between the model prediction result and the true value, and the more accurate the result [20]. The specific formula is defined as:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(\hat{y}_n - y_n)^2}, \tag{18}$$

$$\mathrm{MSE} = \frac{1}{N}\sum_{n=1}^{N}(\hat{y}_n - y_n)^2, \tag{19}$$

$$\mathrm{MAPE} = \frac{1}{N}\sum_{n=1}^{N}\frac{|\hat{y}_n - y_n|}{y_n}, \tag{20}$$

$$\mathrm{MAE} = \frac{1}{N}\sum_{n=1}^{N}|\hat{y}_n - y_n|, \tag{21}$$

In the formula: N is the number of experimental predictions, $\hat{y}_n$ is the model predicted value, y_n is the true value and $\bar{y}$ is the average value of the true value.

4.2.1. The prediction model

In order to test the prediction performance of the LSTM model, different models are compared. The most widely used the BP model, the Elman model and the LSTM prediction model are selected for prediction, and the prediction error is analyzed.

The relevant parameters are set as: the number of neurons in the input layer of the BP model and the Elman model is 6, the number of hidden layers is 1, the number of nodes in the hidden layer is 5, the number of neurons in the output layer is 1, the required step size for prediction is 50, and the learning rate is set is 0.005 and the number of iterations is 500. The number of neurons in the input layer of the LSTM model is 6, the number of hidden layers is 1, the number of nodes in the hidden layer is 5, the number of neurons in the output layer is 1, and the dropout coefficient is 0. The prediction results are shown in Fig. 6, and the prediction errors of each model are shown in Table 3.

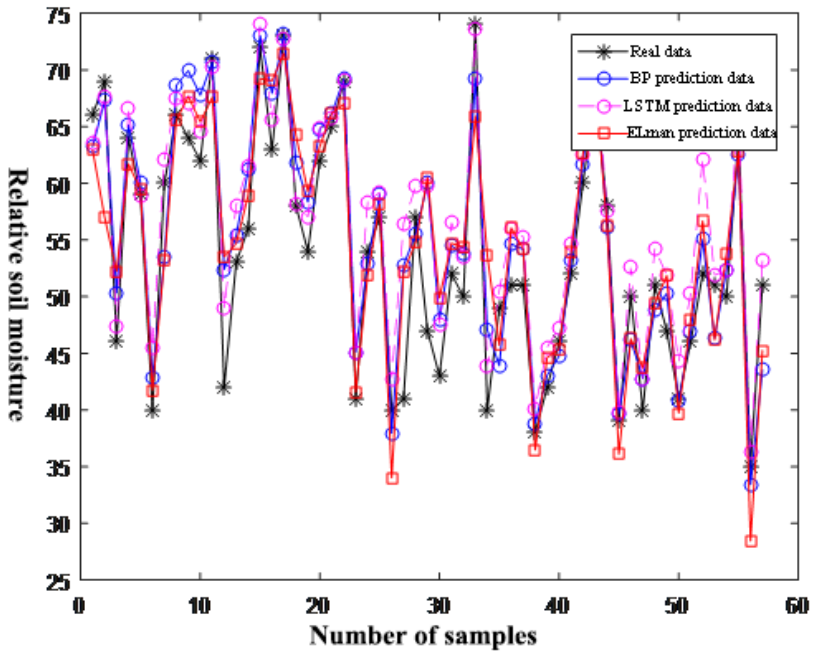

Figure 6. Forecast model prediction results graph.

Table 3. Prediction errors across models.

Model \ Error	RMSE	MSE	MAPE	MAE	R^2
BP	4.9653	24.6545	0.0805	4.2254	0.9266

 doi:10.20944/preprints202212.0141.v1

Elman	3.8261	14.639	0.0578	2.8947	0.9267
LSTM	3.7411	13.9957	0.0582	3.0843	0.9584

As can be seen from Fig. 6 and Table 3, the errors of the three models are all within the required accuracy. Among them, the BP based prediction model has the largest error and the worst accuracy. The prediction accuracy of the Elman based prediction model is higher than that of the BP based prediction model, and the error between the predicted value and the actual value is relatively small. The four errors of the LSTM based prediction model are the smallest among several models, and its determination R2 is also closer to 1 than other models. Therefore, LSTM network model has more advantage than the other two models in moisture predition content.

4.2.2. The analysis of the PSO-LSTM modeling

In order to further improve the prediction accuracy of the LSTM model, the PSO algorithm is used to optimize the LSTM parameter model.

The relevant parameters are set as follows: the inertia weight w=0.8 in the basic PSO, the learning factors c1=2 and c2=2; the maximum number of evolution iterations in the PSO algorithm is set to 100, and the population size N=50. The number of two hidden layer units in LSTM ranges from [1, 15], the dropout coefficient is [0, 1], and the number of training times ranges from [1, 100]. The fitness change curve in the optimization process is shown in Fig. 7.

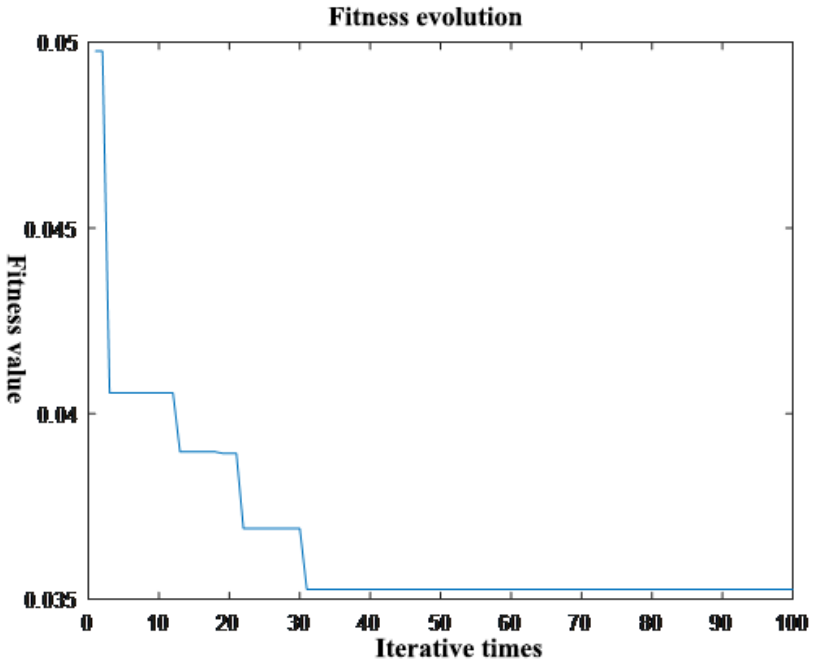

Figure 7. PSO adaptation evolutionary curves.

It can be seen from tht the curve in Fig.7that the PSO can converges quickly in the early stage and has a good global search ability.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2022 doi:10.20944/preprints202212.0141.v1

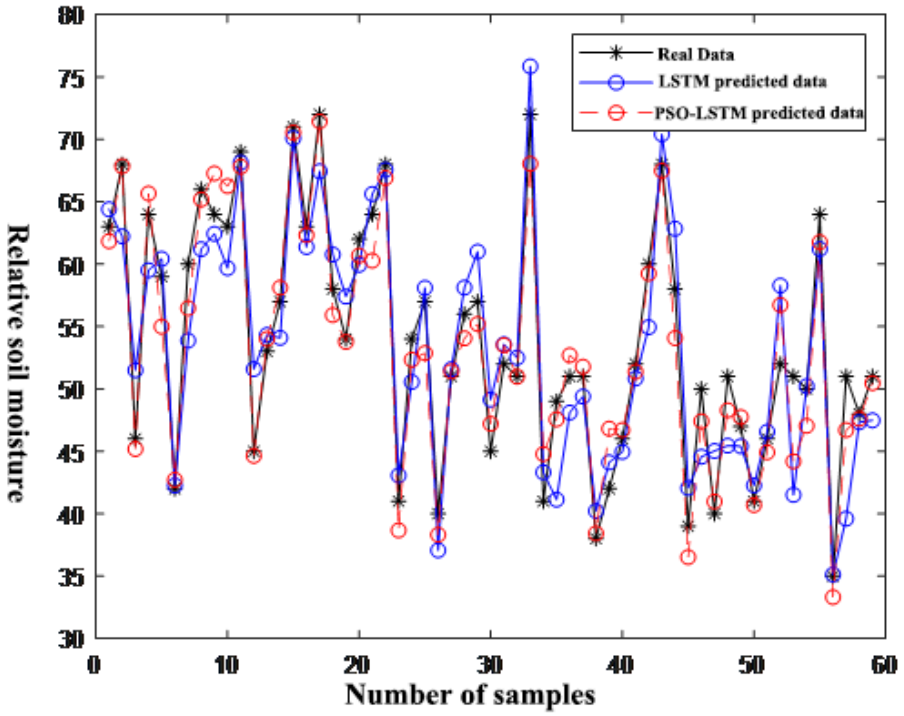

Figure 8. PSO-LSTM model prediction results graph.

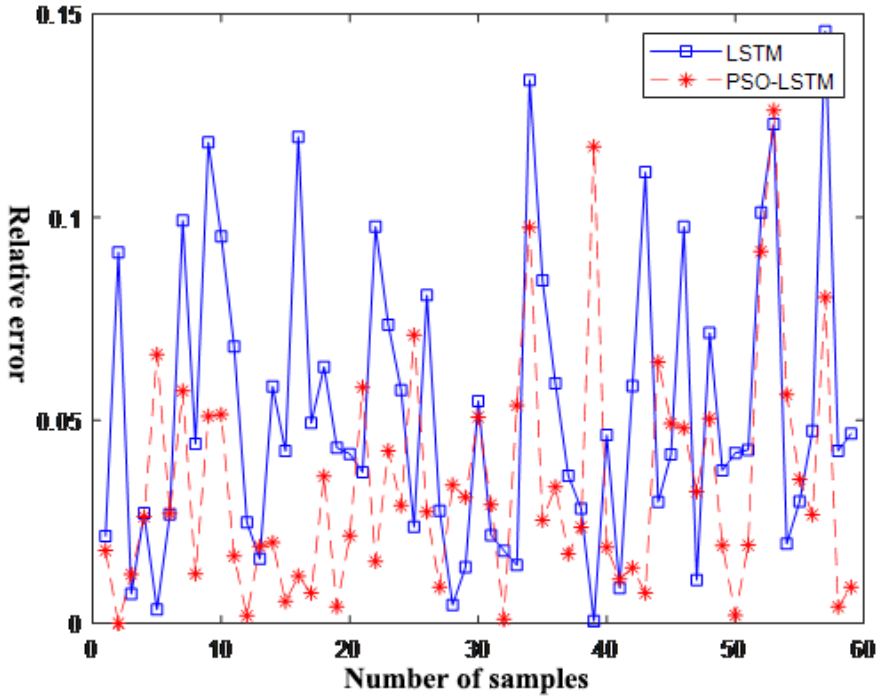

Figure 9. PSO-LSTM model relative error.

The results of the PSO-LSTM prediction model are shown in Figs, 8, and the relative error of the prediction is shown in Fig. 9. It can be seen from Figs.8-9 that the error of the PSO-LSTM prediction model is mostly within 0.05, and the error value of most prediction points is smaller than that of the LSTM network, indicating that the automatic parameter

 doi:10.20944/preprints202212.0141.v1

adjustment of the LSTM model through the PSO can be optimized. The prediction accuracy of the proposed model is greatly improved.

Table 4. Prediction errors across models.

Model \ Error	RMSE	MSE	MAPE	MAE	R^2
BP	4.9653	24.6545	0.0805	4.2254	0.9266
Elman	3.8261	14.639	0.0578	2.8947	0.9267
LSTM	3.7411	13.9957	0.0582	3.0843	0.9584
PSO-LSTM	2.2911	5.2493	0.0339	1.7814	0.9725

The errors predicted by these prediction models are shown in Table 4. It can be seen from Table 4 that the RMSE of the PSO-LSTM based soil moisture prediction model is 2.2911, the MSE is 5.2493, the MAPE is 0.0339, and the MAE is 1.7814. Compared with the single LSTM model, the RMSE of the PSO-LSTM model is reduced by 1.45%, the MSE is reduced by 8.7464%, the MAPE is reduced by 2.43%, the MAE is reduced by 1.3029%, and the R^2 is also improved by 0.0141. It shows that the PSO algorithm has a good application effect on the parameter adjustment of the LSTM model, and can improve the accuracy of the soil moisture prediction.

5. Conclusion

In this paper, the collected meteorological data are analyzed by the grey correlation to determine the main variables affecting soil moisture. Then, combined with the current mainstream time series model, the PSO algorithm is used to optimize the hyperparameters of the LSTM model, Then the PSO-LSTM model is constructed to study the changing law of soil moisture. Through experimental comparison, it is found that: (1) Correlation analysis of data can effectively reduce the dimension of training data, reduce training time, and improve prediction accuracy. (2) The LSTM network model can better solve the hysteresis problem of the BP model and improve the accuracy. (3) The accuracy of the PSO-LSTM model is significantly improved compared to the single LSTM, indicating that the optimization of parameters has a greater impact on the model prediction accuracy. (4) Compared with the BP, the Elman, and the LSTM, the deviation between the predicted value and the real value of the PSO-LSTM model is small, and the accuracy and stability are improved significantly. The prediction accuracy of the PSO-LSTM model at one time point and multiple time point in the future is better than other models, and it can be better applied in the prediction of soil moisture.

In summary, the PSO-LSTM model has good prediction accuracy in predicting soil moisture, which can play a role in water-saving irrigation and intelligent irrigation. In future practical applications, more comprehensive integration and analysis of data is required to further improve the accuracy and stability of predictions.

Author Contributions: Conceptualization, Xiaoyi Huang and Jizheng Chu; Software, Jianlong Wu; Supervision, Jizheng Chu; Writing – original draft, Jianlong Wu; Writing – review & editing, Xiaoyi Huang.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kerr, Y., et al. (2010), The SMOS mission: New tool for monitoring key elements of the global water cycle, *Proc. IEEE*, **98**(5), 666– 687, doi:10.1109/JPROC.2010.2043032.
2. Y. Kim, T. Jackson, R. Bindlish, H. Lee and S. Hong, "Radar Vegetation Index for Estimating the Vegetation Water Content of Rice and Soybean," in IEEE Geoscience and Remote Sensing Letters, vol. 9, no. 4, pp. 564-568, July 2012, doi: 10.1109/LGRS.2011.2174772.
3. Dumedah, G., and P. Coulibaly (2012), Evolutionary assimilation of streamflow in distributed hydrologic modeling using in-situ soil moisture data, *Adv. Water Resour.*,**53**, 231– 241, doi:10.1016/j.advwatres.2012.07.012.
4. Seneviratne, S. I., T. Corti, E. L. Davin, M. Mirschi, E. B. Jaeger, I. Lehner, B. Orlowsky, and A. J. Teuling (2010), Investigating soil moisture-climate interactions in a changing climate: A review, *Earth Sci. Rev.*, **99**(1), 125– 161, doi:10.1016/j.earscirev.2010.02.004.
5. Al-Mukhtar, M. Modelling the root zone soil moisture using artificial neural networks, a case study. *Environ Earth Sci* **75,** 1124 (2016).
6. Li, Z., Chen, T., Wu, Q. *et al.* Application of penalized linear regression and ensemble methods for drought forecasting in Northeast China. *Meteorol Atmos Phys* **132,** 113–130 (2020).
7. Yang S H,Wang Y M.An artificial neural network model for soil moisture prediction responding to weather parameters. 34th International Symposium on Agricultural Engineering . (2006).
8. J. Xie et al., "Irrigation Prediction Model with BP Neural Network Improved by Genetic Algorithm in Orchards," 2019 Eleventh International Conference on Advanced Computational Intelligence (ICACI), 2019, pp. 108-112, doi: 10.1109/ICACI.2019.8778528.
9. M. Beyki and M. Yagoobi, "Chaotic logic gate: A new approach in set and design by genetic algorithm", *Chaos Solitons & Fractals*, vol. 77, pp. 247-252, (2015).
10. Fung, K.F., Huang, Y.F. & Koo, C.H. Coupling fuzzy–SVR and boosting–SVR models with wavelet decomposition for meteorological drought prediction. *Environ Earth Sci* **78,** 693 (2019).
11. Al-Mukhtar, M. Modelling the root zone soil moisture using artificial neural networks, a case study. Environ Earth Sci 75, 1124 (2016).
12. Flower pollination algorithm with dimension by dimension improvement Math. Probl. Eng., 2014 (2014).
13. Hochreiter S, Schmidhuber J. Long Short-Term Memory[J]. Neural Computation,1997,9(8):1735-1780.
14. Kennedy J, Eberhart R. Particle swarm optimization[C]//Proceedings of ICNN'95-international conference on neural networks. IEEE, 1995, 4: 1942-1948.
15. Zhao W, Tao T, Zio E, et al.A novel hybrid method of parameters tuning in support vector regression for reliability predictio n: particle swarm optimization combined with analytical selection[J]. IEEETransactions on Reliability,2016,65(3):1393-1405.
16. Xiangyun Qing, Yugang Niu, Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM[J]. Energy,2018,148:461-468.
17. Hochreiter,S. and Schmidhuber,J.: Long Short-Term Memory, Neural Computation,9,1735-1780, https://doi.org/10/bxd65w,(1997).
18. Tommy W S, Yong Fang. A recurrent neural network based real time learning control strategy applying to nonlinear system with unknown dynamics. IEEE Transaction on Industry Appltication. (1998).
19. Peijie Lin;Zhouning Peng;Yunfeng Lai;Shuying Cheng;Zhicong Chen;Lijun Wu. Short-term power prediction for photovoltaic power plants using a hybrid improved Kmeans-GRA-Elman model based on multivariate meteorological factors and historical power datasets[J]. Energy Conversion and Management. (2018).
20. Moriasi DN, Arnold JG, Van Liew MV, Binger RL, Harmel RD, Veith TL (2007) Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Am Soc Agric Biol Eng 50(3):885–900.