ISBN: 978-981-18-5852-9
WCSE 2022 Spring Event: 2022 9th International Conference on Industrial Engineering and Applications
doi: 10.18178/wcse.2022.04.085

GNSS-R Soil Moisture Retrieval Model Based on Elman Neural Network

Yuanhua Liu, Wenwen Chen and Xinliang Niu

[1,2] School of Communications and Information Engineering, Xi'an University of Posts & Telecommunications

[3] Xi'an Branch of China Academy of Space Technology, China Academy of Space Technology

Abstract. Global Navigation Satellite System-Reflectometry (GNSS-R) in soil moisture retrieval, the traditional method mainly includes linear regression and exponential regression methods. To address the defects of conventional methods such as poor prediction accuracy and sizeable computational effort, the Elman neural network with dynamic learning features is introduced. An Elman neural network-based soil moisture retrieval method is proposed to establish a multi-parameter retrieval model. Finally, the model is trained to validate the feasibility of this model. The results indicate that the soil moisture values estimated by the GNSS-R soil moisture retrieval method based on the Elman neural network have minor errors with the actual measured soil moisture values. Based on this model, the coefficient of determination (R^2) is 0.8988, and the Root Mean Square Error (RMSE) of soil moisture is 0.0207. When compared to the traditional linear regression model, the soil moisture values predicted by this method are more accurate and closer to the measured soil moisture values, demonstrating the method's validity and reliability.

Keywords: Global Navigation Satellite System-Reflectometry (GNSS-R), Elman neural network, soil moisture retrieval

1. Introduction

Soil moisture is an essential component of the ecosystem water vapor cycle [1]. Having accurate and stable soil moisture information plays a vital role in water resource management, climate change, agricultural production, and environmental testing in a region or globally [2-6]. Global Navigation Satellite System-reflectometry (GNSS-R) has the characteristics of wide signal coverage and strong penetration, which is a microwave remote sensing technology developed from the Global Navigation Satellite System (GNSS).The GNSS signal reflected from the ground surface carries information about the physical characteristics of the surface. By analyzing and studying reflected signal, some physical features of the surface can be extracted, such as soil moisture [7].

At present, a lot of research has been done on soil moisture retrieval using GNSS-R. Larson et al [8-10] proved that the Signal-to-Noise Ratio (SNR) showed a linear relationship between the amplitude, delay, and phase of the three characteristic parameters and soil moisture through the Plate Boundary Observatory (PBO) data in the United States, which can be used for obtaining soil moisture. Chew et al [11-12] investigated the relationship between the three parameters individually and soil moisture on this basis. The study showed the strongest correlation between the phase and soil moisture and established an empirical model for obtaining the soil moisture in the bare soil surface layer. Minsi Ao et al [13-14] conducted a comparative analysis test between SNR observations and soil moisture, showing a certain exponential relationship between the SNR phase and soil moisture. Jizhong Wu et al [15] addressed the parameter estimation problem of soil moisture and improved the prediction accuracy of the linear regression model by enhancing the reflected signal parameters accordingly.

The above studies have shown that SNR observations can effectively predict soil moisture by linear regression models. However, surface factors such as vegetation coverage and surrounding environment will generally affect GNSS reflected signal. The accuracy of soil moisture predicted by the linear regression model will inevitably be affected [16].

Artificial intelligence algorithms such as machine learning are becoming increasingly popular for remote sensing applications. They can build nonlinear relationships between data inputs and outputs and build complex regression models [17-19]. The Elman neural network is a representative dynamic recurrent neural network, which adds a takeover layer based on the feed-forward neural network. It is used to memorize and store the output data of the previous moment, ensure the full use of historical data, and promote the network to adapt to the dynamic information characteristics of the data.

This paper treats GNSS-R soil moisture retrieval as a nonlinear problem, taking into account the dynamic change characteristics of soil moisture observation data, introducing the Elman neural network with dynamic learning characteristics into GNSS-R soil moisture retrieval. A multi-parameter soil moisture retrieval model was established using SNR amplitude, phase, Normalized Difference Vegetation Index (NDVI), temperature and the satellite altitude angle as input terms and soil moisture as output term. It aims to restrain the influence of vegetation and the environment and improve the prediction accuracy of soil moisture model.

2. Theoretical Background

2.1. Retrieve Soil Moisture

GNSS-R is a remote sensing technique that estimates and retrieves the physical parameters of the earth's surface by processing GNSS signals reflected from the surface. When performing remote sensing surveys, GNSS receivers receive not only the direct signal, but also GNSS signal reflected from reflective surfaces. Due to the multipath effect, the signals of two paths are superimposed at the antenna to produce the interference signal, as shown in Fig.1. At the same time, the GNSS receiver will also record the signal strength, that is, *SNR* observation value. The relationship between the reflected signal, the direct signal and the SNR observation value of the interference signal is as follows:

$$SNR^2 = A_{\mathrm{d}}^2 + A_m^2 + 2A_d A_m \cos\phi \tag{1}$$

SNR represents the interference signal's SNR observation value, A_d stands the direct signal's the amplitude, A_m stands the reflected signal's the amplitude, and ϕ is the phase difference between the reflected signal and the direct signal.

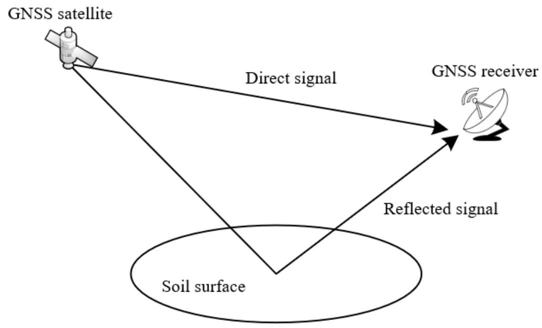

Fig. 1: GNSS signal superposition phenomenon.

Only the reflected signal carries the relevant information of the soil in the interference signal, and the direct signal component is much larger than the reflected signal component. Thus, using a low order polynomial fit remove the direct signal component and extract the reflected signal component, the reflected signal is expressed as [20]:

$$SNR_{\mathrm{m}} = A_m \cos\left(\frac{4\pi h}{\lambda}\sin\theta + \varphi\right) \tag{2}$$

where SNR_m represents the reflected signal's SNR observation value, A_m stands amplitude of the reflected signal, h is Antenna height, λ is GNSS signal wavelength, θ is the satellite altitude angle , and φ is the phase.

Then, the Lomb-Scargle algorithm is used to get the reflected signal's frequency and then calculate the amplitude and phase of the reflected signal according to the principle of least squares. The two characteristic parameters, amplitude and phase, were strongly correlated with soil moisture, and then a one-dimensional linear regression statistical model was developed to estimate soil moisture, respectively.

2.2. Elman Neural Network

Elman neural network is a dynamic recurrent neural network proposed by J. L. Elman [21], which is composed of an input layer, a hidden layer, a takeover layer and an output layer, as seen in Fig.2. Compared with the feed-forward neural networks, this network adds a takeover layer to the hidden layer, which is used to store the output of the hidden layer of the previous moment and return to the network's input layer to be fed into the hidden layer again at the next moment. This connecting the hidden and takeover layers gives the neural network a certain memory function. It can make the most of data, increase the network's ability to process dynamic information, and have stronger computing capabilities, thereby achieving the purpose of dynamic modeling.

In Fig.2, $u(t)$ represents the input vector, $y(t)$ represents the output vector, $x(t)$ represents the hidden layer vector, $x_c(t)$ represents the takeover layer vector. w_1, w_2 and w_3 stand weights of each layer.

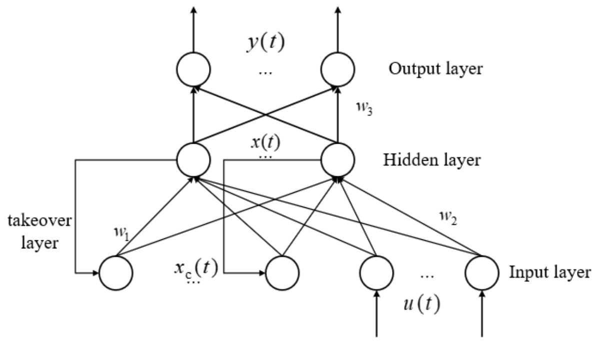

Fig. 2: Elman neural network structure.

3. Methodlogy

3.1. Elman Neural Network for GNSS-R Soil Moisture Retrieval Models

In this paper, when constructing the GNSS-R soil moisture retrieval model, apart from the two characteristic parameters of reflected signal amplitude and phase. Three other characteristic parameters related to soil moisture retrieval are also considered: temperature, Normalized Difference Vegetation Index (NDVI) and the satellite altitude angle, all of which directly or indirectly relate to soil moisture.

In the process of retrieving soil moisture by GNSS-R, vegetation coverage is inevitable, and vegetation will cause attenuation to the surface reflected signal and affect the SNR observation of the reflected signal. Hence, it is necessary to make corrections for the influence of vegetation. To a certain extent, NDVI can effectively reflect the vegetation growth and cover and characterize the vegetation information. In a certain area, the higher the NDVI value, the greater the attenuation of the reflected signal from the surface, and the GNSS receiver can receive the less reflected signal.

Temperature also affects soil moisture. Generally, there is a certain trend between temperature changes and soil moisture changes, with an increase in temperature decreasing soil moisture and vice versa; temperature changes also cause SNR observations. GNSS-R technology retrieve soil moisture uses the interference effect of the reflected signal and the direct signal. When the satellite altitude angle is different, the interference phenomenon also presents different states. The smaller the satellite altitude angle, the more pronounced the interference phenomenon and the more accurate the estimated soil moisture will be at this point

Table 1: Indicators and Calculation Formulae for Error Analysis

Evaluation indicators	The coefficient of determination (R^2)	The Root Mean Square Error (RMSE)
Calculation formula	$\frac{SSR}{SST}$	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y-y_{pred}\right)^2}$

According to the above analysis, this paper takes the amplitude and phase of the reflected signal, temperature, NDVI and the satellite altitude angle as the model's input targets, and the measured soil moisture value as the model's input target, to construct the Elman neural network-based soil moisture retrieval model. The soil moisture retrieval process is shown in Fig.3. After determining the characteristic observation parameters, establish a soil moisture retrieval model, train on the data set to select the optimal retrieval model, and perform the model evaluation.

R^2 and RMSE are used to evaluate the prediction capability and reliability of this model. The closer R^2 gets to 1, the more reliable the model. The closer the RMSE is to 0, the stronger the model's fit and the higher the prediction accuracy. The above indicators and their calculation formulae are shown in Table 1.

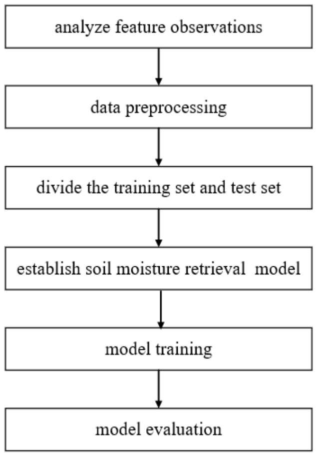

Fig. 3: Flowchart of the soil moisture retrieval algorithm.

In Table I, *SSR* represents the regression sum of squares, *SST* represents the total sum of squares, *n* is the number of test samples, y and y_{pred} represent the measured soil moisture values and predicted soil moisture values, respectively.

Different neural network structures can lead to differences in predicted results. When constructing an Elman neural network model, it is essential to determine the number of network layers, neurons in each layer, and the activation function. The input/output layer is determined by the input/output feature parameters. The number of neurons in the input layer is 5, the number of neurons in the output layer is 1, and the takeover layer is set the same as the hidden layer. In order to establish the optimal Elman neural network model, it is essential to select the suitable activation function, the number of hidden layer and the number of hidden layer neurons.

3.2. Elman Neural Network Structure

When the number of network layers and the number of neurons at each layer are different, the efficiency and accuracy of the Elman neural network predictions will also vary greatly. The number of neurons is too small, which will cause the network to acquire insufficient information to get the best training result. With too many neurons in the hidden layer, the network's functionality is improved, and the accuracy is higher. Still, the number of network training iterations increases and there is a possibility of overfitting. The number of hidden layers is also not fixed and requires constant experimentation to determine the optimal number of layers. This study randomly selected 30 sets of data from the sample data as a validation set to debug the network structure. Using the RMSE of the soil moisture prediction as a verification standard to measure the model's accuracy. When the RMSE is the smallest, the Elman neural network structure can be determined. Fig.4 shows the variation of RMSE with the number of hidden layers and the number of neurons.

As shown in Fig.4, when the hidden layer is one layer, the RMSE of soil moisture gradually reduces with the number of neurons, and reaches the minimum at 14 neurons. Meanwhile, when the number of neurons is fixed and select three hidden layers, the RMSE of soil moisture is smallest, and the prediction accuracy is higher. When the number of hidden layers is three and the number of neurons in the hidden layer is 14, the

RMSE of soil moisture reaches the minimum, so the structure of the Elman neural network model can be determined.

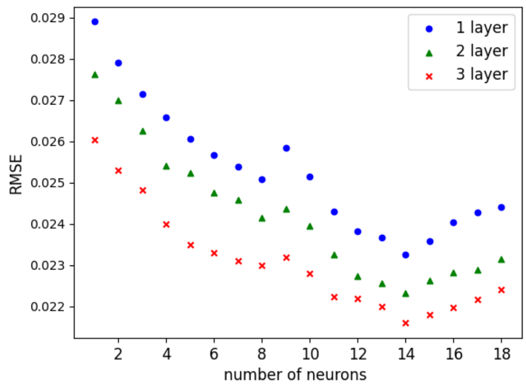

Fig. 4: RMSE of the predicted soil moisture with different number of hidden layers and neurons.

3.3. Select Activation Function

The activation function is a function that runs on the neuron and is responsible for mirroring the input of the neuron to the output. It introduces a nonlinear factor to the neuron, increases the neural network's learning ability to approximate any nonlinear function, and models the arbitrary relationship between input and output.

The ability and efficiency of the Elman neural network training and predicting sample data are related to the network structure and closely associated with the selected activation function. The commonly used activation functions of the Elman neural network mainly include the Sigmoid function, Tanh function, ReLU function, etc. Since there are differences in the performance and efficiency of each function, the corresponding activation functions need to be selected for different practical problems. Therefore, to optimize the performance and efficiency of the Elman neural network retrieval model, given the current practical issues, the predicted results under different activation functions are compared and analyzed.

The principle of the comparative analysis method is followed to ensure that the influencing factors are single. This study chooses to keep the learning rate, expectation error, and other influencing parameters consistent. According to the previous study, the hidden layer is three layers with 14 neurons in each layer. Different activation functions build the Elman neural network model and train it on the validation set. Table II shows R^2 and RMSE of the Elman neural network model under different activation functions.

Table 2 shows that the Elman neural network model has the largest R^2 and the smallest RMSE when the ReLU function is chosen with the remaining influencing parameters fixed. It indicates that the Elman neural network retrieval model constructed by the ReLU function as the activation function has the best performance and effect. Therefore, this study selects the ReLU function as the activation function to build the Elman neural network soil moisture retrieval model.

Table 2: Analysis Of Predicted Results of the Elman Neural Network Model

activation function	R^2	RMSE
Sigmoid function	0.8261	0.0289
Tanh function	0.8189	0.0261
ReLU function	0.8486	0.0226

4. Experiment and Results

4.1. DATASelect Activation Function

The experimental data in this study are all from a strawberry picking garden in Chang ' a District, Xi 'an , Shaanxi Province, with the site located at 34° 3' 40" N,108° 53' 47" E. When the satellite altitude Angle is between 5° and 25°, the interference phenomenon is most significant, so this study selects the sample data of the satellite altitude angle in this interval, and eliminates invalid values and abnormal values.

The numerical values and units in the sample data are not the same. Before the model training, the sample data should be normalized to all the data in the same dimension. The standardization method used in this study, that is, deviation standardization, linear transformation of the original data. The function expression for processing is as follows:

$$x^* = (x - \min) / (\max - \min) \quad (3)$$

x is the original data, x^* is the standardized data, max is the maximum value and min is the minimum in all sample data. After the normalization process, the sample data are processed to the range of 0~1, which is convenient for comprehensive comparison and evaluation and accelerates the convergence speed of the Elman neural network model.

After the above processing, 710 sets of sample data are finally selected. One set of sample data contains 6 observations. Using the random sampling method, 659 groups were selected as the training set and 51 groups as the test set. The training and test sets are mutually exclusive, and there is no intersection between them. Then 30 sets are randomly selected from the training set as the validation set, which is used to experimentally test and debug network parameters to determine the optimal network structure.

4.2. Results

After using the above data set for model training, to study the model's predictive ability, the model test is performed on the test set to obtain the predicted result, and the result is compared with the actual measured soil moisture value. Verify the accuracy of the Elman neural network model and perform the model evaluation. Fig.5 shows the comparison relationship between the soil moisture values predicted by the model and the measured soil moisture values. The RMSE value of this model is 0.0207. Fig.6 shows the correlation between the soil moisture values predicted by the model and the measured soil moisture values with an R^2 value of 0.8988 for this model.

As can be seen from Fig.5 and Fig.6, the predicted result of the Elman model is closer to the actual measured soil moisture value. Although there are still differences in the local ranges of the predicted values, the overall values are very close to the measured values, with a high degree of approximation, minor errors, and good agreement. This result indicates that the Elman neural network retrieval model can estimate soil moisture with high accuracy and reliability.

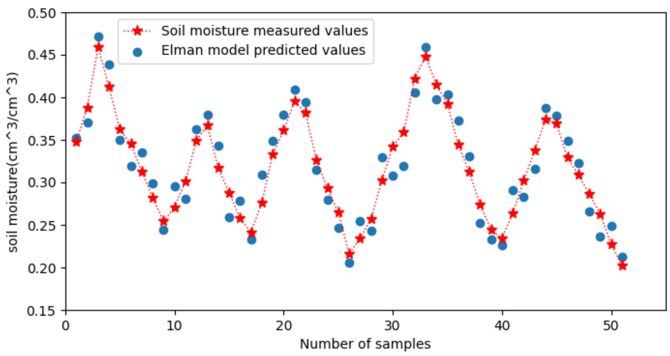

Fig. 5: Comparison between the predicted value of the Elman model and the measured soil moisture value.

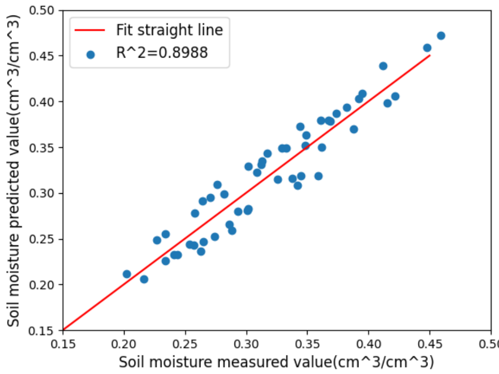

Fig. 6: Correlation between the predicted value of the Elman model and the measured soil moisture value.

To further verify the superiority of the Elman model, explain the model's predictive ability. Under the condition of using the same data set, this study also compares with two traditional linear regression methods, the amplitude linear model and the phase linear model. Fig.7 shows the comparison between the predicted value of the three models and the measured value. Table III3 shows RMSE and R^2 of the three models.

As can be seen from Fig.7, the predicted value of the three models is basically in line with expectations. Still, it is evident from Table III that the amplitude linear model has the worst predictive result and the most significant deviation from the measured value. The phase linear model has a better predictive effect than the amplitude linear model. The Elman neural network model has the best predictive effect. The predicted soil moisture value is more consistent with the measured soil moisture value, and the fitting result is good, which is significantly better than the two linear regression models. This result indicates that the Elman neural network model can be used for GNSS-R soil moisture estimation. It is more reliable and more accurate than the traditional model, which can evidence the prediction accuracy of soil moisture to some extent.

Table 3: Error Evaluation of the Three Models

model	R^2	RMSE
Elman neural network model	0.8988	0.0207
Amplitude linear model	0.6089	0.0418
Phase linear model	0.6784	0.0357

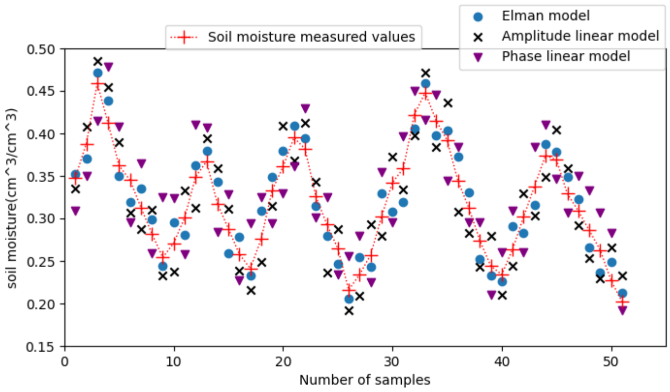

Fig. 7: Comparison between the predicted value of three models and the measured soil moisture value.

5. Conclusion

In this paper, a soil moisture retrieval model based on the Elman neural network with dynamic learning characteristics is introduced to address the shortcomings of the traditional linear regression model when performing GNSS-R soil moisture retrieval. First, the soil moisture retrieval characteristic quantities are determined, and a multi-parameter soil moisture retrieval model is established. Secondly, the validation set determined the optimal Elman neural network structure experimentally. Finally, the model was tested for validation and compared with the traditional linear regression statistical model. The result shows that the predicted result of the Elman neural network model is closer to the measured soil moisture value. The R^2 between the predicted and measured values is 0.8988, and the RMSE is 0.0207. Compared with the traditional linear regression model, the predicted value of this model is closer to the measured value with a minor error, which proves the validity and reliability of the Elman neural network model to predict soil moisture.

6. References

[1] S. Vey, A. Güntner ,J. Wickert ,T. Blume, and M. Ramatschi ,"Long-term soil moisture dynamics derived from GNSS interferometric reflectometry:a case study for Sutherland," GPS solutions, vol.39, no.4, pp.641-654, 2016.

[2] X.R. Wu, S.G. Jin, Y.Z. Song, L. Yang, X.R. Li, and R.L. Sha, "Current status of GNSS-R/IR soil moisture remote sensing research[J]," Journal of Geodesy and Geodynamics, vol.39. no.12. pp. 1277-1282,2019.

[3] A. Robock, K.Y. Vinnikov, G. Srinivasan, J.K.Entin, S.E.Hollinger, and N.A. Speranskaya,"The global soil moisture data bank[J],"Bulletin of the American Meteorological Society, vol.81,no.6,pp.1281-1300,2000.

[4] D.J. Zhang, R.L.Tang, W. Zhao, B.H. Tang, H. Wu, and K. Shao, "urface soil water content estimation from thermal remote sensing based on the temporal variation of land surface temperature[J]," Remote Sensing, vol.6, no.4, pp.3170-3187,2014.

[5] L. Zhang, H.Q. Lv, L.Y. Wang, and B.Y. Yang, "Temporal and Spatial Variation Characteristics of Soil Moisture in China[J]," Acta Geographica Sinica, vol.71, no.09, pp.1494-1508,2016.

[6] K. Chen, F. Shen, X.Y. Cao, and Y.F. Zhu, "GNSS-IR Soil Moisture Retrieval Based on Deep Belief Network[J]," Bulletin of Surveying and Mapping, September 2020,pp.100-105.

[7] D.K. Yang, and Q.S. Zhang, GNSS Reflected Signal Processing: Fundamentals and Applications, CA:Publishing House of Electronics Industry, 2012.

[8] K.M. Larson, J.J. Braun, E.E. Small, V.U. Zavorotny, E.D. Gutmann, and A.L. Bilich, "GPS multipath and its relation to near-surface soil moisture content[J]," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,vol.3,no.1,pp.91-99,2009.

[9] K.M. Larson, J.J. Braun, E.E. Small, V.U. Zavorotny, E.D. Gutmann, and A.L. Bilich, "Use of GPS receivers as a soil moisture network for water cycle studies[J]," Geophysical Research Letters, vol.35, no.24, pp.851-854,2008.

[10] K.M. Larson, J.J. Braun, E.E. Small, V.U. Zavorotny, E.D. Gutmann, and A.L. Bilich, "Using GPS multipath to measure soil moisture fluctuations: initial results[J]," GPS solutions, vol.12, no.3, pp.173-177,2008.

[11] C.C. Chew, E.E. Small, K.M. Larson, and V.U. Zavorotny, "Effects of near-surface soil moisture on GPS SNR data: Development of a retrieval algorithm for soil moisture[J]," IEEE Transactions on Geoscience and Remote Sensing, vol.52, no.1, pp.537-543,2013.

[12] C.C. Chew, E.E. Small, K.M. Larson, and V.U. Zavorotny, "Vegetation sensing using GPS-interferometric reflectometry: Theoretical effects of canopy parameters on signal-to-noise ratio data[J]," IEEE Transactions on Geoscience and Remote Sensing, vol.53. no.5, pp.2755-2764,2014.

[13] M.S. Ao, Y.J. Hu, Y.D. Liu, Y. Zeng, and Z. Wu, "Inversion of Soil Moisture Trend Based on GPS Signal-to-Noise Ratio Observations[J]," Journal of Geomatics Science and Technology,vol.29,no.02,pp.140-143,2012.

[14] M.S. Ao, J.J. Zhu, Y.J. Hu, Y. Zeng, and Y.D. Liu, "GPS soil moisture monitoring using SNR observations[J]," Geomatics and Information Science of Wuhan University, vol.40, no.01, pp. 117-120+127,2015.

[15] J.Z. Wu, T. Wang, and W. Wu , "The inversion model of using GPS-IR to monitor soil water content[J]," Geomatics and Information Science of Wuhan University,vol.43.no.6.pp.887-892,2018.

[16] Z.G. Zhang, Y.J. Liang, C. Ren, Y.B. Huang, and Y.L. Pan, "Research on GPS-IR Retrieval of Soil Moisture Based on Kalman Filtering[J]," Journal of Guilin University of Technology,vol.41,no.01,pp.149-154,2021.

[17] Y.H. Liu, R.G. Xu, and X.L. Niu, "ANN-based GNSS-R sea surface wind speed retrieval method," June 2021,pp.40-44.

[18] C.Deac, G.C. Deac, R.C. Parpală, C.L. Popa, and C.E. Cotet, "Vibration Anomaly Detection using Deep Neural Network and Convolutional Neural Network," *International Journal of Modeling and Optimization* vol. 11, no. 1, pp. 19-28, 2021.

[19] T. Zhang, W.H. Zeng, and F. L. Yang, "Applying a BP Neural Network Approach to the Evolution Stage Classification of China Rift Lakes, " *International Journal of Modeling and Optimization* vol. 4, no. 6, pp. 450-454, 2014.

[20] B. Sun ,C. Zhang, S.C. Yin, H. Xu, and W.J. Zhang, "Research on GNSS-IR Soil Moisture Inversion Method Based on PSO-RF [J]," Radio Engineering, vol.51, no.10, pp.1080-1085,2021.

[21] J.L. Elman, "Finding structure in time[J]," Cognitive science, vol.14, no.2, pp.179-211,1990.