

## Research on Multi-star Nonlinear Regression Estimation model of Soil Moisture Based on wavelet analysis

Yalong PAN<sup>1</sup>, Yueji LIANG<sup>1,2</sup>, Chao REN<sup>1,2\*</sup>, Zhigang ZHANG<sup>1</sup>, Yibang HUANG<sup>1</sup>, Yajie SHI<sup>1</sup>

<sup>1</sup> College of Geomatics Engineering and Geoinformation, Guilin University of Technology, 541006 Guilin, China – panyl0617@qq.com

<sup>2</sup> Guangxi Key Laboratory of Spatial Information and Geomatics, 541006 Guilin, China – renchao@glut.edu.cn

**KEY WORDS:** GPS-IR, Wavelet Analysis, Soil Moisture, SNR, Multi-star Nonlinear Regression

### ABSTRACT:

Soil moisture is an important parameter for studying meteorological, hydrological, and agricultural research. How to obtain accurate soil moisture has become a key issue. The traditional method of obtaining soil moisture has the disadvantages of high cost, complicated operation and limited application range. In recent years, using GPS-IR (GPS-Interferometric Reflectometry) technology to estimate soil moisture has become a Research hot. Traditional GPS-IR technology uses low-order polynomial to separate GPS satellite direct and reflected signals, but the separation effect is poor. And single-star inversion accuracy is low. In order to solve these problems, this paper proposes a multi-star nonlinear regression model based on wavelet analysis. Firstly, the satellite direct reflection signal is fitted by wavelet analysis, then the relative phase delay is solved by the nonlinear least squares method. Finally, the multi-star nonlinear regression model is established to estimate the soil moisture. The experiment used the observation data of the P041 station provided by the US Plate Edge Observation Program PBO in 2012. The results show that the wavelet analysis separates the reflected signal better than the low-order polynomial. The model can fully combine the advantages of wavelet analysis and multi-star fusion inversion, and effectively improve the abnormal jumping phenomenon of single-star. The inversion result is significantly upgrade than the traditional method. When the model combination reached double-star and triple-star, the better results were obtained. The R reached 0.922 and 0.948, respectively. The test results increased by 18.6% and 20.9% compared with the traditional method.

### 1. INTRODUCTION

In recent years, with the rapid development of GNSS (Global Navigation Satellite System) reflection remote sensing technology, it has been widely used in the fields of sea breeze, sea level, soil moisture, snow depth (Zhang, et al, 2016). Among them, the use of GPS-IR technology developed based on measurement receivers to monitor soil moisture has become a low-cost, high-efficiency, high-resolution method.

The key data of GPS-IR inversion of soil moisture is the SNR (signal-to-noise ratio) data in the GPS observation file. The reflected signal in SNR can reflect the soil moisture change within 5cm of the soil surface. There is a certain correlation between signal phase, satellite elevation angle and soil water content, but the correlation is weak when the soil water content is less than 10% (Larson, et al, 2008; Larson, et al, 2010). In an experiment using an electrodynamic model to simulate the effects of soil moisture changes on the phase and amplitude of multipath interference, it was found that the multipath interference phase of the reflected signal is linearly related to surface soil moisture and is the best parameter for estimating (Zavorotny, et al, 2010; Chew, et al, 2014). Further research found that there is an exponential relationship between SNR multipath interference phase and soil moisture (Aao, et al, 2010). In terms of satellite elevation angle variation, studies have shown that low satellite elevation angles (5° ~30°) and high satellite elevation angles (30° ~70°) SNR data fusion and observation and soil The correlation of humidity is higher than the correlation between the original observation and soil moisture (Roussel, et al, 2010).

Based on the existing research, the use of GPS-IR technology to invert soil moisture is more limited to single-star inversion, and less research on multi-star combined inversion of soil moisture. Considering the use of multiple satellites, machine learning can achieve the initial fusion of multi-stars, and multi-star fusion inversion of soil moisture has a significant improvement compared to single-star inversion, but the establishment of models is limited to machine learning method, not easy to visualize the model (Ren, et al, 2018; Liang, et al, 2018). Therefore, based on the multi-star combination idea, this paper considers the visualization of the model and establishes a multi-star nonlinear regression model based on wavelet analysis. The feasibility and effectiveness of the model inversion of soil moisture were verified by comparison and analysis with the inversion results of a single satellite.

### 2. SOIL MOISTURE INVERSION PRINCIPLE

#### 2.1 GPS-IR basic principle

The SNR observation and the multipath interference phase are a sin(cos) relationship, and the GPS-IR monitoring of soil moisture is only related to the reflected signal, then the SNR after removing the direct component is only the multipath reflection component, which is There is still a sin(cos) function relationship of a fixed frequency between sin( $\theta$ ) (Chwe, et al, 2015). Under the one-reflection assumption, the reflection component  $SNR_r$  can be expressed as:

$$SNR_r = A \cdot \cos\left(\frac{4\pi h}{\lambda} \sin(\theta) + \varphi_r\right) \quad (1)$$

\* Corresponding author: renchao@glut.edu.cn

Where  $\theta$ ,  $\lambda$ ,  $h$  represent the satellite incident angle height, carrier frequency, and antenna height, respectively;  $A_r$ ,  $\varphi_r$  represent the amplitude of the multipath reflection component and the multipath interference phase, respectively. If  $t = \sin(\theta) f = 2h/\lambda$ , then equation (1) can be expressed as  $SNR_r = A_r \cos(2\pi ft + \varphi_r)$ .

## 2.2 Wavelet separation and reflection signal principle

Let the SNR observation  $f(t)$  be:

$$f(t) = [x_1^0, x_2^0, \dots, x_i^0] \quad (t = 1, 2, \dots, i) \quad (2)$$

Where  $t$  represents an epoch.

Then, when the reflected signal is separated by the wavelet principle,  $f(t)$  represents the original signal. The focus of wavelet analysis is to gradually multi-scale the original signal by wavelet-based function expansion and translation operations (Hou, et al, 1997; Ai, et al, 2002). Using the wavelet decomposition signal is:

$$Wf(\alpha, \beta) = \int f\varphi(\alpha, \beta)(t) dt \quad (3)$$

Where  $\alpha$  is a scale factor,  $\beta$  is a translation factor, and  $\varphi(\alpha, \beta)$  is a continuous wavelet generated by a wavelet master function. Then use the reconstructed  $f(t)$  as:

$$f'(t) = C_{\varphi}^{-1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{Wf(\alpha, \beta)}{\alpha^2} \varphi(\alpha, \beta)(t) d\alpha d\beta \quad (4)$$

In the formula, it is a wavelet basis function. Then  $f'(t)$  is the low-frequency component reconstructed by the signal in multiple layers. Decomposing 6 layers with *coif5* wavelet base can effectively improve the resolution and accuracy of reflection signal separation. The correlation between the acquired interference phase and soil moisture is higher than the correlation between the interference phase obtained by low-order polynomial and soil moisture (Zhang, et al, 2019). Therefore, this paper uses the *coif5* wavelet function to decompose and reconstruct the SNR signal.

## 2.3 Multi-star nonlinear regression model

There is an exponential relationship between the interference phase and the soil moisture (Aao, et al, 2015), then there are other nonlinear relationships between the two, which has become the starting point of this paper. The multivariate nonlinear regression analysis method is suitable for explaining the nonlinear relationship between a dependent variable and multiple independent variables. In order to explore whether it can also be applied to soil moisture inversion, this paper establishes a soil moisture  $y$  and each satellite. Multi-star nonlinear regression model between the inverse phase  $x_i (i=1, 2, \dots, n)$  of the inverse performance (Yang, et al, 2014):

$$y = f(x_1, x_2, \dots, x_n) = b_1 + b_2 x_1 + b_3 x_2 + \dots + b_j x_j x_2^2 + b_m x_n^3 \quad (5)$$

Where  $b_1$  is the regression coefficient and  $b_m$  is the partial regression coefficient ( $m$  is a positive integer).

The regression coefficients of multi-linear nonlinear regression equations are solved by the Levenberg-Marquardt algorithm (L-M) (Wen, et al, 2016; Zhao, et al, 2017).

## 2.4 Inversion process

- 1) Separate direct and reflected signals. The TEQC software is used to solve the GPS observation data to obtain the SNR value of the L2 carrier, and the *coif5* wavelet function is used to decompose and reconstruct the SNR signal, and separate the direct and reflected signals;
- 2) Signal resampling. Re-sampling the reflected component with the epoch change into a relationship with the sinusoidal value  $\sin(\theta)$  of the incident angle of the satellite;
- 3) Parameter estimation. The sinusoidal fitting is performed on the resampled component by nonlinear least squares fitting method to obtain the interference phase.
- 4) Establish a model. The sample data was input, and the regression coefficients of the multi-star nonlinear regression model were solved by L-M method. A multi-star nonlinear regression model was established, and the test data was used to test the inversion accuracy of the model.

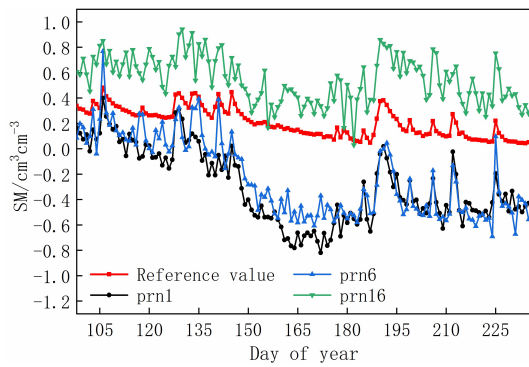
## 3. EXPERIMENT ANALYSIS

The GPS observing data and soil moisture reference values from 98 to 236 day of year in 2012 worked out by P041 station under the framework of PBO, a US Plate Edge Observation Program, are going to be used for analysis in this experiment. The station receiver is TRIMBLE NERT9, adopting SCIT radome, and the antenna model is TRM59800.80, the data sampling frequency is 15Hz. It can be seen from Figure 1 that the terrain around the station is flat and the vegetation is sparse, which is convenient for GPS-IR inversion of soil moisture experiments.

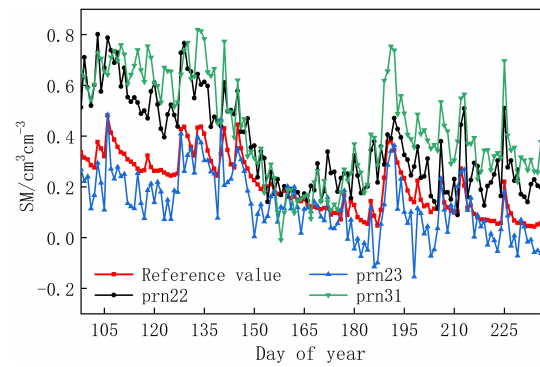


Figure 1. Surround environment of P041 station

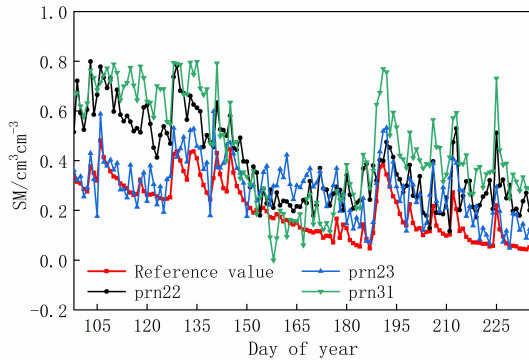
In this paper, the SNR data between the satellite height angle  $5^\circ \sim 20^\circ$  is used to estimate the reflected signal parameters. In order to verify the feasibility and accuracy of the model, triple experimental schemes are set: 1) using a second-order polynomial to fit the reflected signal and using a single-star data to establish a linear soil moisture inversion model; 2) using wavelet analysis to separate the reflected signal and A single-star linear soil moisture inversion model was established by using single-star data. 3) The model proposed in this paper uses wavelet analysis to separate the reflected signals and establish a multi-star nonlinear regression model, and select the interference phase of 98-216 days of year product as the input sample of the model, select the interference phase of 217-236 days of year product as the test sample of the model. In this paper, the second-order polynomial is used to solve the correlation analysis between the interference phase of each satellite and the soil moisture reference value, and the PRN1, 6, 16, 22, 23, 31 satellites are selected according to the correlation. The multipath interference phases solved by the first scheme and the second scheme are shown in Fig. 2 and Fig. 3.



(a) Multipath interference phase of PRN1、6、16

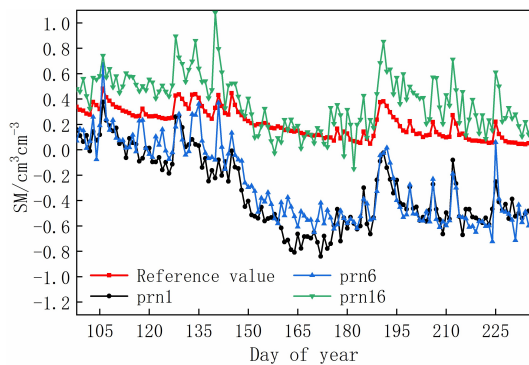


(b) Multipath interference phase of PRN22、23、31



(b) Multipath interference phase of PRN22、23、31

Fig 2. Multipath interference phase of scheme 1



(a) Multipath interference phase of PRN1、6、16

Fig 3. Multipath interference phase of scheme 2

It can be seen from Fig. 2 and Fig. 3 that whether using a second-order polynomial or a multi-path interference phase calculated by wavelet analysis, it can respond to changes in soil moisture, but the interference phase of different satellites has different reflections on soil moisture and there are errors. There are many abnormal hopping values, which may be related to the different running orbits of each satellite and the different time of occurrence. To further compare the accuracy of the first and second schemes, a linear regression equation of multipath interference phase and soil moisture is established, and the R (Correlation coefficient) is compared, as shown in Fig. 4.

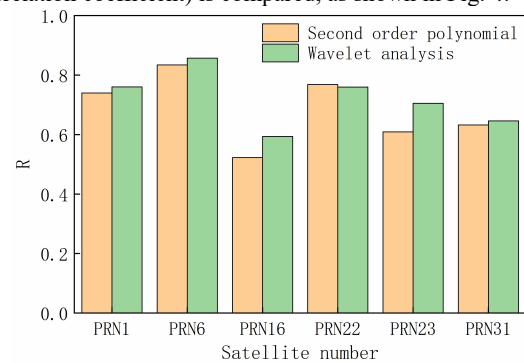


Fig.4 R between scheme 1 and scheme 2

It can be seen from Fig. 4 that the overall effect of the second scheme is better than that of the first scheme, and the wavelet analysis is better than the second-order polynomial for separating the reflected signals. As can be seen from Fig. 2, Fig. 3 and Fig. 4, the R between the multipath interference phase and the soil moisture calculated by a single satellite is low, and the error is large. Considering the feasibility of multi-star fusion, this paper makes full use of the advantages of each satellite by establishing a multi-star nonlinear regression model based on wavelet analysis to obtain more accurate and reliable soil moisture values. The regression coefficients of each combination of Scheme 3 are shown in Tables 1, 2 and 3. Figure 5 shows the results of soil moisture inversion for different combinations.

Combination	Satellite	b1	b2	b3	b4
1	PRN 1	0.329	0.344	-0.178	-0.327
2	PRN 6	0.305	0.299	-0.059	0.025
3	PRN16	0.118	-0.197	1.519	-1.035
4	PRN22	0.150	-0.562	2.696	-2.003
5	PRN23	0.112	0.472	2.255	-3.893
6	PRN31	0.179	-0.670	2.187	-1.261

Tab. 1 Single star nonlinear regression model regression coefficient

Combination	Satellite	b1	b2	b3	b4	b5
7	PRN1、6	0.306	0.176	0.144	0.879	0.530
8	PRN6、16	0.158	-0.180	0.327	-0.498	-0.052
9	PRN16、22	0.191	-0.398	-0.464	-0.0002	0.558
10	PRN16、23	0.191	-0.259	-0.513	-0.274	0.952
11	PRN22、23	0.112	-0.243	0.168	1.105	-0.734
12	PRN22、31	0.226	-0.828	-0.298	1.355	-0.774

Combination	Satellite	b6	b7	b8	b9	b10
7	PRN1、6	-1.329	1.386	-0.819	2.884	-3.255
8	PRN6、16	0.917	-0.063	-0.061	-0.346	0.494
9	PRN16、22	2.650	-0.087	0.079	-2.167	-0.006
10	PRN16、23	1.759	0.240	-0.676	-0.517	-0.402
11	PRN22、23	1.821	-0.297	-1.695	4.346	-4.004
12	PRN22、31	3.155	0.617	1.865	-2.407	-2.685

Tab. 2 Double star nonlinear regression model regression coefficient

Combination	Satellite	b1	b2	b3	b4	b5	b6	b7
13	PRN1、6、23	0.288	0.604	-0.079	-0.384	1.289	0.032	2.669
14	PRN16、22、31	0.229	-0.494	-0.662	-0.012	-0.365	1.007	-1.032
15	PRN22、23、31	0.096	-0.004	0.446	-0.128	-0.528	-1.424	-0.425

Combination	Satellite	b8	b9	b10	b11	b12	b13	b14
13	PRN1、6、23	-0.673	-3.613	1.077	0.479	0.626	-2.756	0.901
14	PRN16、22、31	2.395	1.096	0.843	0.237	-0.463	2.211	-0.597
15	PRN22、23、31	2.012	1.687	-0.995	-1.735	-1.413	2.598	-2.036

Combination	Satellite	b15	b16	b17	b18	b19
13	PRN1、6、23	-2.487	-1.517	-0.023	6.484	-1.530
14	PRN16、22、31	-0.022	-1.399	1.290	-0.584	-3.416
15	PRN22、23、31	5.898	0.593	4.616	-6.725	-1.357

Tab. 3 Triple star nonlinear regression model regression coefficient

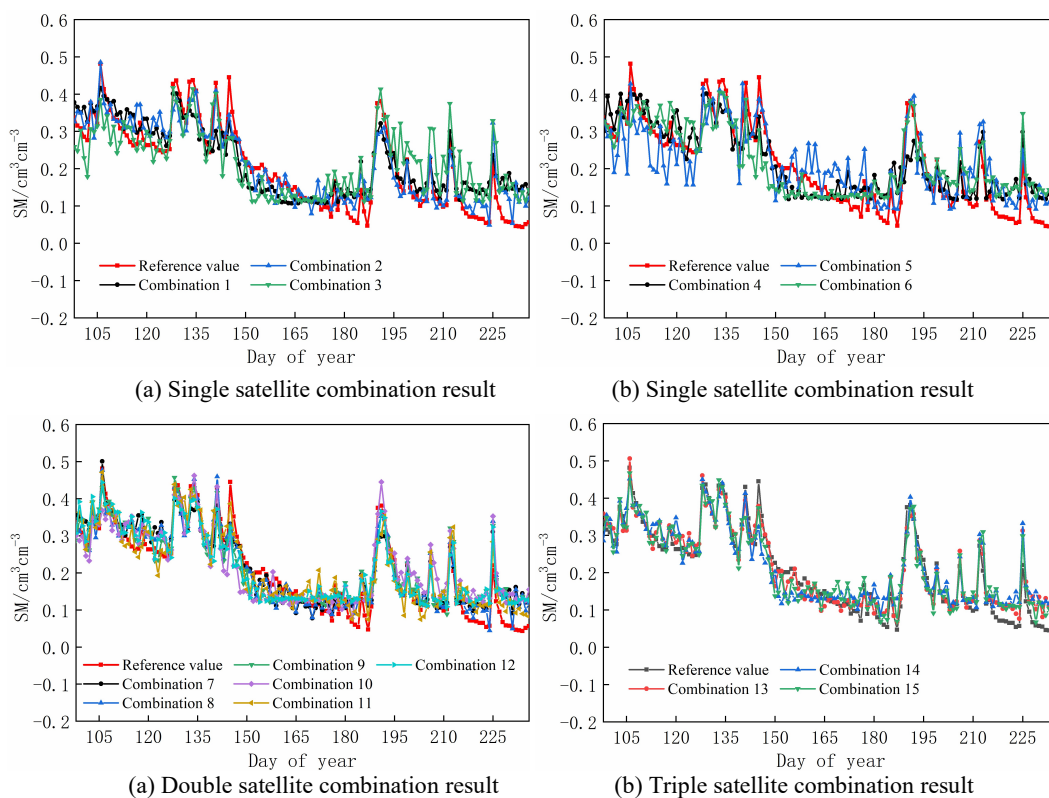
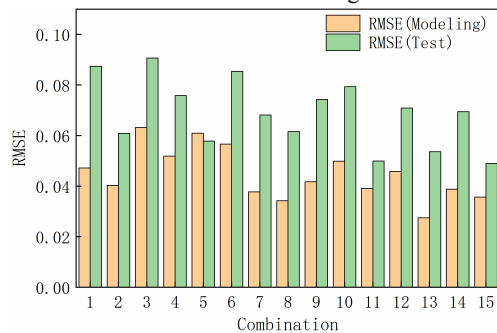
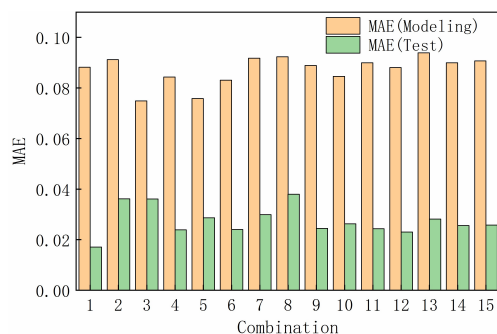


Fig.5 Multipath interference phase of scheme 3

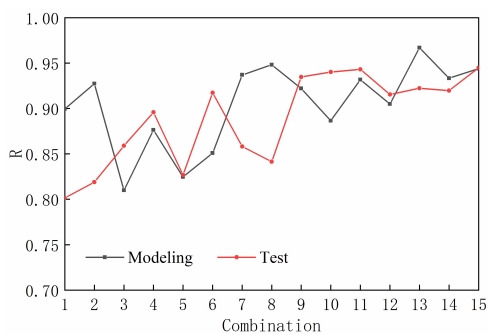
It can be seen from Fig. 5(a) and (b) that the result of the scheme three single-star is better than that of scheme one and scheme two. The obtained multipath interference phase can better reflect the change of soil moisture, but there is still a large inversion error, and there are some abnormal jump values. Such as combination 3 is during the annual accumulation period of 195-210 day of year, the combination 5 has an abnormal jump value during the annual accumulation of 98-127 and 150-180 day of year. It can be seen that although single-star can improve the effect of inversion of soil moisture, the accuracy is still low. It can be seen from Fig. 5(c) that the effect of using double-star modeling is obviously better than that of single-star, which can more accurately reflect the change of soil moisture. It can be seen from Fig. 5(d) that Samsung's modeling has the best effect, and the inversion results have a strong correlation with soil moisture, and the abnormal jump value is effectively improved. The modeling effect and test results have been improved. In order to further comprehensively evaluate the feasibility and effectiveness of each scheme, this paper uses the correlation coefficient (R), Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) Comprehensive assessment of the modeling and test accuracy of model Option 3. The modeling and test accuracy are evaluated comprehensively. The accuracy indicators of each scheme are shown in Figure 6.



(a) Root mean squared error



(b) Mean absolute error



(c) Correlation coefficient

Fig.6 Scheme 3 precision indicators

It can be obtained from the analysis in Figure 6 that the modeling results and the test results obtained by using the third scheme are higher than those of the first and second schemes. However, due to the difference in performance of each satellite, the inversion results are incomplete when using single-star modeling to invert the soil moisture. The same, and the accuracy index is poor. The double-star and triple-star modeling can achieve better results. The R between the binary star modeling result and the test result and the soil moisture reached 0.922 and 0.906, respectively, which was 22.2% higher than that of the scheme 1 and scheme 2 and 18.6%; the R between Samsung modeling results, test results and soil moisture reached 0.948 and 0.929, respectively, which was 24.5% and 20.9% higher than that of scheme 1 and scheme 2. The average of RMSE of the double-star and triple-star modeling test sections were 0.0673 and 0.0573, respectively.

In summary, the effect of the third scheme is obviously better than the scheme one and the second scheme. Multi-star nonlinear regression model based on wavelet analysis can fully integrate the performance of each satellite and can better invert soil moisture. When the combination reaches double-star and triple-star, the model works best, and the inversion results have a high correlation and accuracy with soil moisture. Therefore, the model proposed in this paper is feasible and effective.

#### 4. CONCLUSION

In this paper, a multi-satellite nonlinear regression model based on wavelet analysis is proposed. Through theoretical and experimental analysis, it is concluded that: 1) Wavelet analysis can better separate the reflected signals in SNR, and the result is better than the traditional second-order polynomial. 2) Based on the wavelet analysis to effectively decompose the reflected signal, the proposed model fully integrates the advantages of each satellite, effectively improves the single-star anomaly jump phenomenon, and the inversion results and accuracy are significantly improved. 3) When the combination reaches double-star and triple-star, the model has the best effect, and the R reaches 0.922 and 0.948 respectively. The test results are increased by 18.6% and 20.9% compared with the traditional method, and the average RMSE is 0.067 and 0.057 respectively. Therefore, the multi-star nonlinear regression model based on wavelet analysis proposed in this paper is feasible.

#### ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (Grant Nos.41461089 and 41901409); the Guangxi Young and Middle-aged Teacher Basic Ability Improvement Project (Grant No. 2018KY0247); the Guangxi Spatial Information Key Laboratory of Surveying and Mapping (Grant Nos. 16-380-25-22 and 15-140-07-34). The authors would like to thank the PBO of the US Plate Edge Observations Program for providing observational data and soil moisture reference values.

#### REFERENCES

- Dai K. Y., Zhang S. C., Li Z. Y., et al, 2016. Snow Depth Detection with GPS SNR Observations. *Journal of Geodesy and Geodynamics*, 36(06), 525-528.
- Larson, K. M., Small, E. E., Gutmann, E., et al, 2008. Using GPS multipath to measure soil moisture fluctuations: initial results. *GPS solutions*, 12(3), 173-177.

- Zavorotny, V. U., Larson, K. M., Braun, J. J., et al, 2010. A physical model for GPS multipath caused by land reflections: Toward bare soil moisture retrievals. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 3(1), 100-110.
- Larson, K. M., Braun, J. J., Small, E. E., et al, 2009. GPS multipath and its relation to near-surface soil moisture content. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 3(1), 91-99.
- Chew, C. C., Small, E. E., Larson, K. M., et al, 2013. Effects of near-surface soil moisture on GPS SNR data: development of a retrieval algorithm for soil moisture. *IEEE Transactions on Geoscience and Remote Sensing*, 52(1), 537-543.
- Ao, M. S., Zhu, J. J., Hu, Y. J., et al, 2015. GPS soil moisture monitoring using SNR observations. *Geomatics and Information Science of Wuhan University*, 40(01), 117-120.
- Roussel, N., Frappart, F., Ramillien, G., et al, 2016. Detection of soil moisture variations using GPS and GLONASS SNR data for elevation angles ranging from 2 to 70. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(10), 4781-4794.
- Ren, C., Liang, Y. J., Lu, X. J., et al, 2019. Research on the soil moisture sliding estimation method using the LS-SVM based on multi-satellite fusion. *International Journal of Remote Sensing*, 40(5-6), 2104-2119.
- Liang, Y., Ren, C., Huang, Y., 2018. Research on multi-satellites fusion inversion model of soil moisture based on sliding window. In *China Satellite Navigation Conference (pp. 85-95)*. Springer, Singapore.
- Chew, C. C., Small, E. E., Larson, K. M., et al, 2014. Vegetation sensing using GPS-interferometric reflectometry: Theoretical effects of canopy parameters on signal-to-noise ratio data. *IEEE Transactions on Geoscience and Remote Sensing*, 53(5), 2755-2764.
- Hou, Z. Z., Yang, W. C., 1997. Wavelet Transform and Multiscale Analysis of Gravity Anomalies in China. *Chinese Journal of Geophysics*. 1997(01):85-95.
- Ai, L., Liu, Z. S., Ge, Y. J., 2002. Several methods of wavelet denoising. *Journal of Hefei University of Technology*, 2002(02):167-172.
- Zhang, Z. G., Liang, Y. J., Ren, C., 2019. Research on the preferred method of wavelet base based on GPS-IR inversion of soil moisture. *Progress in Geophysics*. 1-9[2019-09-18]. <http://kns.cnki.net/kcms/detail/11.2982.p.20190624.1542.128.html>.
- Yang, F., Liu, C., Tang F. P., et al, 2014. Numerical Simulation on Hydraulic Performance of Axial-flow Pumping System with Adjustable Inlet Guide Vanes. *Transactions of the Chinese Society for Agricultural Machinery*. 5(05):51-58.
- Wen, H. F., He, H., 2016. A Nonlinear 3D Rectangular Coordinate Transformation Method Based on Levenberg-Marquardt Algorithm. *Journal of Geodesy and Geodynamics*. 36(08):737-740.
- Zhao, D. H., Zhao, D. B., Fan, W. Y., 2017. Based on the Calculation of the Mine Surface Subsidence Parameters Berg Levenberg-Marquardt(L-M) Algorithm. *Bulletin of Surveying and Mapping*. (S2):91-96.