International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2015): 78.96 | Impact Factor (2015): 6.391

Time Sequence Forecast of Ground Settlement Based on WT-SVR-ARMA

Pengfei Wang

[1]Beijing University of Technology,Beijing University of Technology,No.100,shuanglong Road, Chaoyang District, Beijing,100124, China

Abstract: ***Ground settlement caused by the subway underground excavation is a very critical issue. Considering the important significance of an accurate ground settlement prediction model to underground construction, a time series prediction method based on wavelet transform (WT) and support vector regression(SVR) and autoregressive moving average (ARMA) model to predict ground settlement is prosed. Wavelet transform is used to decompose and reconstruct ground settlement time sequence into trend sequence and random sequence. Trend sequence is forecasted by SVR model. Random sequence is forecasted by ARMA model. The final prediction values are the sum of trend sequence and random sequence prediction values. This method is used to the data sampled from sensors located at Ziyou Road station in Changchun and it shows that this method is valid and applicable.***

Keywords: time sequence, wavelet transform, support vector regression, autoregressive moving average, ground settlement

1. Introduction

Ground settlement have direct impacts on safety and stability of adjacent structures and underground facilities. Choosing a long-term settlement observation to test the amount of future settlement not only needs consuming lots of time and resources and cannot predict the final settlement, but also delays the implementation of preventive measures[1]. Accurately predicting future settlement of the ground surface can effectively prevent accidents caused by excessive settlement, and achieve dynamic design and construction control. Therefore, ground settlement prediction during construction is very important.

At present, there are three main methods to predict ground settlement: The first method is based on the layer-wise summation method to calculate the final settlement and the simplified consolidation formula to calculate the consolidation degree, Then the regularity of settlement can be estimated; The second method is numerical method. That is, according to the consolidation theory combined with the constitutive model of various soils, the final settlement and its development regularity be calculated. In theory, these two methods are reasonable and feasible, but the calculation parameters involved in these two methods must be obtained experimentally. In the sampling process, the soil samples will inevitably be disturbed, leading to the experimental parameters difference between the actual parameters. So the predictive results of both methods usually have large errors. The third method is the commonly used method based on the measured data which build a forecasting model using pre-engineering measured deformation data, through a certain statistical method, according to the settlement itself change with time and space. This method is simple and has certain theoretical basis, also can make full use of the scene measured data to obtain satisfactory results

In the third method, ANN method, gray prediction theory, regression analysis method and time series model are commonly used[2-3]. Due to these methods and models are mostly single and have their own characteristics and applicable condition, the measured data information cannot be fully excavated. The prediction accuracy also needs to be improved. The combination of different single models can be used to construct a hybrid forecasting model, which can greatly exploit and make full use of forecasting sample information. It also can reduce the impact by random factors effectively to improve the prediction accuracy and enhance the prediction stability .In this paper, we try to establish a hybrid forecasting method based on wavelet transform and SVR and ARMA which has a better adaptation and processing capacity to the complex dynamic system of the ground

2. WT-SVR-ARMA Model

Ground settlement is a complicated process of multi-factor action. Conventional theoretical calculation method can only consider the main factors such as the nature of the reinforcement, the geological conditions and the type of the upper structure. For the factors such as construction conditions, technical level and climatic conditions which can not be complicated quantitative is not be considered. In addition, because the rock and soil is a heterogeneous anisotropic elastoplastic viscous body, the complexity geological conditions make its mechanical parameters and mechanical phenomena with a strong randomness and uncertainty.

For the above reasons, the measured data of ground settlement are in addition to the theoretical trend of the soil mechanics deformation caused by load, but also with some randomness. Therefor the deformation sequence of the ground settlement can be decomposed into the trend sequence and the random sequence. The trend sequence reflects the theoretical regularity of ground settlement, which is the main basis of deformation data and belongs to nonstationary series. Random sequence belongs to noise series and has certain stability, which is one of the main reasons for the prediction accuracy of single model. Therefore, in the process of ground settlement prediction, we should aim at the potential characteristics of trend and random sequence to establish two models for the two sequence respectively. Wavelet transform is a kind of time sequence analysis method developed in recent years[5-6], and it is applied in the field of noise reduction. In this paper, the ground settlement data are reconstructed into trend time sequence and random time sequence using wavelet transform.

The Support Vector regression (SVR) proposed by Suykens et al. has shown some advantages in time series prediction[7], which is characterized by the structural risk minimization principle, good generalization capacity, and make sure the global optimal solution is obtained. In this paper, we establish a SVR model to forecast the trend time sequence. Due to the random time series has some degree of stationarity and randomness, and ARMA model can effectively deal with the characteristics. So ARMA model is used for random time sequence prediction. Finally, the prediction values of the two sequences is summed as the final prediction result. The idea of this paper is shown in Figure1.

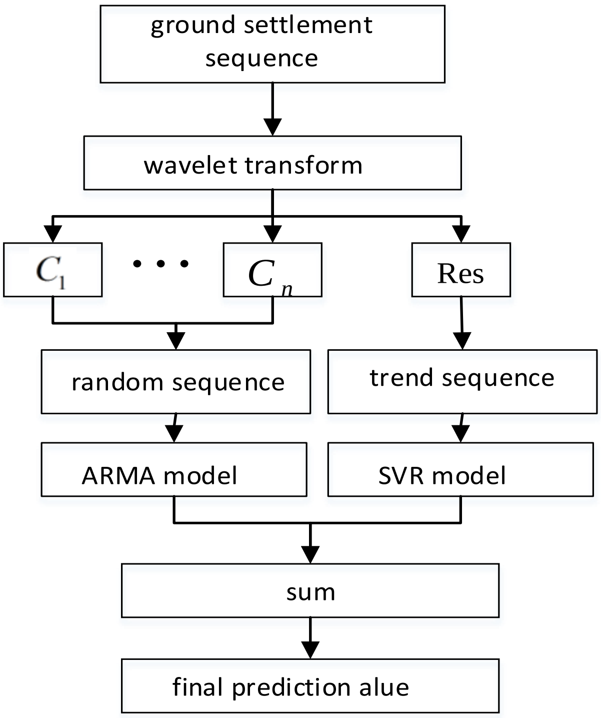

Figure 1: Flow chart

Background on WT-SVR-ARM

3. Background on WT-SVR-ARM

3.1 Wavelet Transform

Let the function $\psi(t)$ be a square integrable function, and if its Fourier transform $\psi(w)$ satisfies the following admissible condition:

$$\int_R \frac{|\psi(w)|^2}{w} dw < \infty \qquad (1)$$

Then According to the different scale α and the different translation β the function family $\psi(t)$ can produce a function family :

$$w_{\alpha,\beta}(t) = |a|^{-1/2}\psi\left(\frac{t-\beta}{\alpha}\right) \qquad (2)$$

Where α is the scale factor, β is the translation factor.

The essence of wavelet transform is to express the original time series $f(t)$ as the weighted sums of the family of functions:

$$w_f(\beta,\alpha) = \int f(x)\psi_{\alpha,\beta}(\mathrm{t})\,\mathrm{dt} \qquad (3)$$

Reconstruction algorithm is used to reconstruct the time sequence after wavelet transform The reconstruction algorithm is:

$$f(t) = \iint \frac{1}{a^2} w_f(\beta,\alpha)\,\psi_{\alpha,\beta}\,d\alpha\,d\beta \qquad (4)$$

Wavelet decomposition layers should not be too little or too much. In order to ensure the smoothness of the trend time sequence, this paper uses dB4 orthogonal wavelet to decompose the measured deformation time sequence.

3.2 Support Vector Regression

Given a sample set (x_i, y_i), $i = 1, 2, \cdots, n$, $x_i \in R^n$, $y_i \in R$, the sample set is mapped from the input space to the high-dimensional feature space through the nonlinear mapping $\varphi(x)$ defined by the inner product function Then make linear regression in the high-dimensional space, so that the nonlinear function estimation problem in the input space is transformed into the liner estimation problem in high-dimensional feature space , the estimation function is expressed as:

$$f(x) = w^T\varphi(x) + b \qquad (5)$$

The linear lossless function ε is used as the loss function of SVR, defined as follows:

$$L_\varepsilon = \begin{cases} |f(x) - y| - \varepsilon, & |f(x) - y| \geq \varepsilon \\ 0, & |f(x) - y| < \varepsilon \end{cases} \qquad (6)$$

$f(x)$ is the regression function predictive value, y is the true sample value. According to the principle of structural risk minimization, the above regression problem is transformed into the following:

$$\min \ \frac{1}{2}\|w\|^2 + c\sum_{i=1}^{n}(\xi_i + \xi_i^*) \qquad (7)$$

Constraints are:

$$\begin{cases} y_i - w^T\varphi(x) - b \leq \varepsilon + \xi_i \\ w^T\varphi(x) - y_i + b \leq \varepsilon + \xi_i^* \\ \xi_i \geq 0, \xi_i^* \geq 0 \end{cases} \qquad (8)$$

The above equation represents the structural risk, where w a is the weight coefficient, $\|w\|^2$ reflects the model complexity, b is the offset amount, c is the penalty factor, ξ_i and ξ_i^* are the relaxation factor, ε specifies the error of the regression function.

The support vector regression equation can be obtained by introducing the Lagrangian function and the kernel function technique which satisfy the Mercer condition which can transformed into the linear operation of the high dimension space in the functional theory. ξ_i^* are the relaxation factor, ε specifies the error of the regression function .

The support vector regression equation can be obtained by introducing the Lagrangian function and the kernel function technique in the functional theory which satisfy the Mercer condition. It can transform nonlinear operation in the low dimension space into the linear operation in the high dimension space.:

$$y(x) = c\sum_{i=1}^{n}\delta_i K(x, x_i) + b \tag{9}$$

Where δ_i is the Lagrangian multiplier and $K(x, x_i)$ is the kernel function.

Since the radial basis function has a good generalization ability, this paper uses the Gaussian radial basis function as the kernel function, namely:

$$K(x_i, x_j) = \exp\left(-\|x_i - x_j\|^2 / 2\sigma^2\right) \tag{10}$$

The selection of SVR parameters have significant effects on the generalization capacity and prediction accuracy of the model, this paper considers the PSO algorithm, which can better optimize the penalty factor and kernel parameters ,to improve the prediction accuracy of the model .

3.3 Autoregressive Moving Average Model

The basic idea of the ARMA model is that time series is a set of random variables that depend on time t, and the individual values that make up the time sequence are uncertain, but the change of the whole sequence has regularity. An approximate mathematical model can descript the time sequenc e[8]. Three types are used commonly : autoregressive (AR) model, moving average (MA) model and autoregressive moving average (ARMA) model. The first two are the special cases of the latter. In the ARMA model, the time series value X_t is a linear function of the stochastic error term and the previous value of the current and pre-period, denoted as $ARMA(p,q)$, expressed as

$$X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \omega_t + \theta_1\omega_{t-1} + \cdots + \theta_q\omega_{t-q} \tag{11}$$

$\phi_1, \phi_2 \cdots \phi_p$ is the autoregressive coefficient;. $\theta_1, \theta_2 \cdots \theta_q$ is the moving average. They are all model parameters to be estimated. ω_t is a white noise sequence independent of each other. If the original sequence is not smooth, and become smooth after d stepwise difference, the original sequence can be expressed as ARMA(p,d,q) model.

4. Case Analysiss

Ziyou Road station is located in the People's Street in Changchun. The station constructed by the PBA method to is a double digging station, the total length is 194.6m, the net width is 20.2m, the depth is 8.0 ~ 11.3m. The geological conditions are complex. The soils in the exploration area of the project are divided into three class quaternary new artificial soil filling layer (Q4ml), the quaternary pleistocene eruption and soil (Q2al+pl), cretaceous mudstone (K) . The plane layout of the pilot hole is shown in Figure 2. The scheme of monitoring point is shown in Figure 3. The surface monitoring point is shown Figure 4.

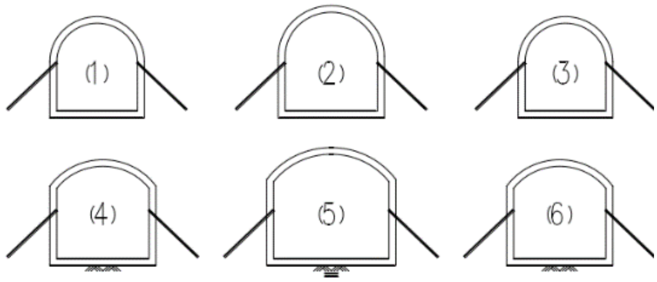

Figure 2: Plane Layout of Pilot

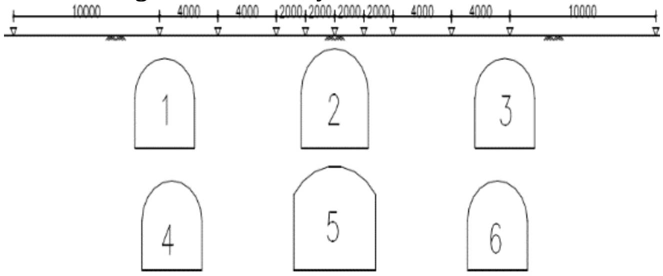

Figure 3: Monitoring Points for pilot

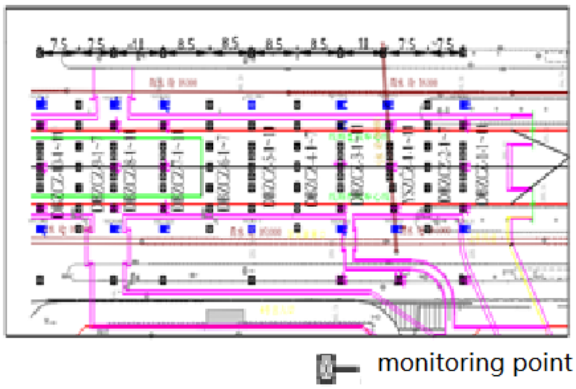

Figure 4: Surface Monitoring Points

The data are monitored from March 2014 to October 2014, once every day and lasted for 228 days. In this paper, data of monitoring point DBZCZ-01-01 is used as an example to establish the prediction model .Then the method is applied to DBZCZ-11-03 and DBZCZ-12 -02 monitoring point to prove the model validity further.

The data of the first 180 groups of DBZCZ01 were used as training samples, and last 48 sets of measured data were used as test samples. Using the Db4 orthogonal wavelet, the train samples time sequence was deconpsed into a trend sequence and two random sequences as show in Figure5.

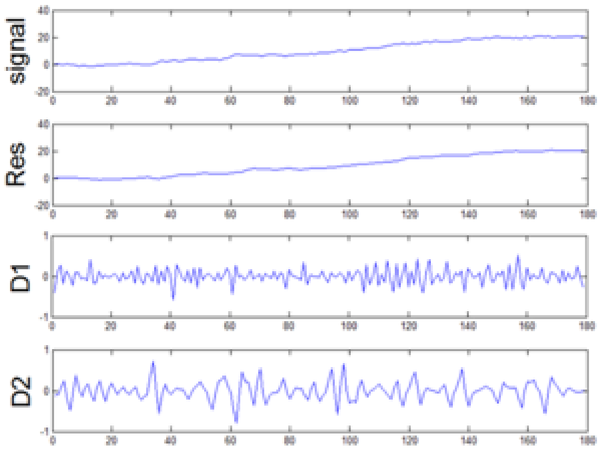

Figure 5: Results of WT

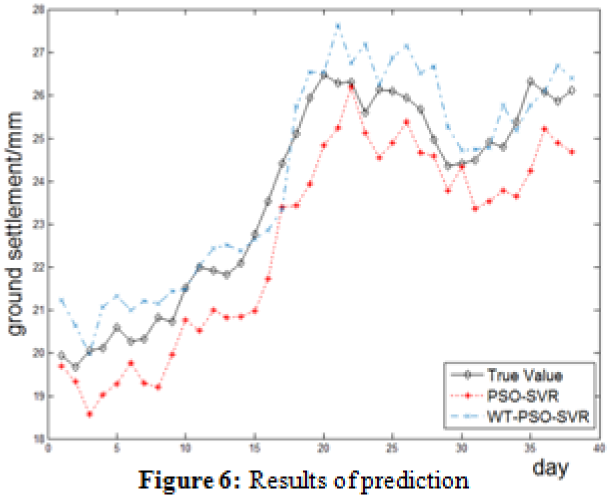

Figure 6: Results of prediction

Table 1: Result of Evaluation

	Evaluation Criteria					
Prediction Model	DBZCZ-01-01		DBZCZ-11-03		DBZCZ-12-02	
	RMSE(mm)	MAPE(%)	RMSE(mm)	MAPE(%)	RMSE(mm)	MAPE(%)
PSO-SVR	1.2052	4.5938	1.3695	4.9847	1.3275	4.5275
WT-PSO-SVR	0.7989	3.3947	0.8536	3.8936	07474	3.4807

We use trend time sequence of train samples to train SVR model . However, the prediction effect of SVR model is overly dependent on its parameters. So Particle Swarm Optimization (PSO) algorithm can be choosed to optimize the parameters of SVR model. To avoid blind searching, the model parameter initialization range is:

$$c=\left[0,1000\right] \ \sigma=\left[0,10\right]$$

PSO optimal number of iterations is M = 200, initial population number is N = 30.

The Mean Predefined Percentage Error (RMSE) and Mean Absolute Percentage Error (MAPE) were used to evaluate the prediction performance of the model.

$$RMSE=\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(x_i-x_i^*\right)^2} \qquad (12)$$

$$\mathrm{MAPE}=\frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i-x_i^*}{x_i}\right)\times 100 \qquad (13)$$

Where N is the number of samples, x_i is the measured value, and x_i^* is the prediction value.

In order to evaluate the predictive ability of WT-SVR-ARMA model, a single PSO-SVR- model is used to be compared with DBZCZ-01-01 monitoring point data. The prediction results are shown in Fig. 6, and the data of DBZCZ-11-03 and DBZCZ-12-02 are used to further verify the model. The evaluation results of the three monitoring points are shown in Table 1. It can be seen from the evaluation results in Table 1 that compared the single PSO-SVR model, the prediction accuracy of WT-SVR-ARMA improves significantly.

5. Conclusion

(1) According to the nonlinear and non-stationarity of the ground settlement time sequence of subway construction, the sequence is decomposed into random sequences and a trend sequence by wavelet transform method. According to the different characteristics of the sequence, two models are established. The final prediction value is the sum of prediction values of two models at the same time which conforms to the physical change process of ground settlement.

(2) The hybrid prediction method decomposes the ground settlement random time sequence into stationary sequences at different scales, which reduce the influence of the non-stationary on the prediction accuracy. The ARMA model can take the randomness of the uncertainty factors into account. Due to the advantages of the ARMA model has on stationary time sequence, the hybrid model avoids the limitations of a single model, and improve the accuracy of prediction.

(3) The SVR model parameters are optimized by particle swarm optimization, which can improve the generalization ability and learning performance of SVR model.

(4) The application of the proposed approach in the data of Ziyou Road station in Changchun indicates its effectiveness and practicability. According to the monitoring standard of the monitoring and control of shallow tunneling method, when the forecast value exceeds the standard, an advance support can be made. It has a guiding significance to construction.

References

[1] Wang F C, Zhu F S, Kang Y M, et al, "Modeling prediction of composite foundation settlement," Journal of

Northeastern University(Natural Science), 2004, 25(10):1010-1012.
[2] Leu,S.S.,Lo,H.C,"Neural-network based regression model of ground surface settlement induced by deep excavation,"Automation in Construction 13,279–289
[3]]Zhong L , Guo C C, Yuan J L, et al, "Application of predicting composite foundation settlement based on fuzzy BP neural networks," J. Wuhan Univ.(Nat. Sci. Ed.), 2007, 53(5): 548-552
[4] Ding L Y,Ma L, "Wavelet Analysis for tunneling-induced ground settlement based on a stochastic model," Tunnelling and Underground Space Technology, v 26, n 5, p 619-628.
[5] Zhang Q W, "Wavelet analysis of monitoring dam settlement considering effects of meteorological factors," Yantu Lixue/Rock and Soil Mechanics, v 31, n 3, p 811-814+826,
[6] Sun Z B, Zhao Q, Zhao H B, "An SVR based hybrid modeling method," Beijing Hangkong Hangtian Daxue Xuebao/Journal of Beijing University of Aeronautics and Astronautics, v 43, n 2, p 352-359
[7] ChenX F, Wang S T, Kang Y M, et al, "SVR with adaptive error penalization" Dianzi Yu Xinxi Xuebao, Journal of Electronics and Information Technology, v 30, n 2, p 367-370
[8] Zerubia, Josiane ,Alengrin, Gerard , "Estimation of ARMA (p, q) parameters," Signal Processing, v 22, n 1, p 53-60.

Author Profile

Pengfei Wang is reading the M.S. degrees in Information Departme nt of Beijing University of Technology .