

Application of Ground Settlement Prediction Based on EMD-RVM

Pengfei Wang

Fault of Information Technology, Beijing University of Technology, Beijing 100124, China

wpf19901012@163.com

Keywords: AIP Proceedings; Title Here; International Conference; Research Center of Engineering and Science.

Abstract. Ground settlement is a critical issue in underground construction. Ground settlement prediction is important to identify serious damage to adjacent structures caused by settlement exceed the standard. However, conventional methods have some limitations due to nonlinearity and non-stationarity of settlement data. This paper proposes a hybrid model based on empirical model decomposition (EMD) and relevance vector machine(RVM) regression optimized by particle swarm optimization (PSO) designated as EMD-RVM. EMD is used to decompose the ground measured settlement time series into several stationary components called intrinsic mode functions (IMFs). Then, SVR is applied to predict the components independently. At last, the expected prediction values are the sum of all components prediction value at the same time. In order to validate the performance of the proposed method., the signal support vector regression(SVR), signal relevance vector machine(RVM) regression, relevance vector machine based on wavelet transform(WT-RVM) are used to be compared with root mean square error (RMSE) and mean absolute percentage Error(MAPE) are used to evaluate these models. The evaluation results indicate the method proposed is effective and practical.

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.

Because the effectiveness of wavelet analysis in non-stationary time series, more and more prediction model are based on wavelet transform. However, Essentially, a wavelet transform is a linear and non-stationary approach. It re- presents a signal by a linear combination of wavelet basis functions. Therefore, it may not be suitable for nonlinear data [4]. In this paper, we try to establish a hybrid forecasting method based on EMD which is powerful and adaptive in analyzing nonlinear and non-stationary dataset and RVM which has a better adaptation and processing capacity to the complex dynamic system of the ground.

2. Methodology

Ground settlement is a complicated process of multi-factor action. 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. So EMD is used to decompose the ground measured settlement time series into several stationary components with different time scales called intrinsic mode functions (IMFs). Then, RVM is applied to predict the components independently. Hybridization procedure of EMD-SVR model is show in Figure1.

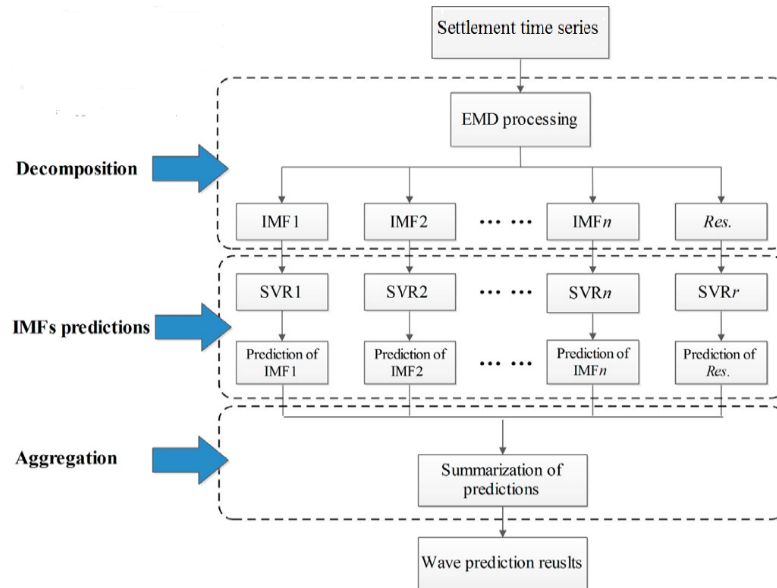


Figure 1. Hybridization procedure of EMD-SVR model

2.1 Empirical Model Decomposition

Empirical Mode Decomposition (EMD) is a new technique to process time frequency of signal, which was proposed by American Chinese N E Huang in 1998. It is also known as Hilbert-Huang Transform (HHT) [5].

By empirical mode decomposition, any data signal can be decomposed into several Intrinsic Mode Functions (IMFs) and a trend component representing the overall trend of the original signal. These intrinsic modal functions are in different scale features. Compared with the wavelet analysis and Fourier transform, the EMD method does not need to set any basis function, which is decomposed completely according to the time scale of the signal itself. Therefore, this decomposition result can reflect the physical process more accurately. It is outstanding in dealing with complex nonlinearity and nonstationary process. NE Huang defines the IMF needs to meet the next two conditions:

- (1) the number of zero and extreme points of the original signal should be equal, or at most one difference and
- (2) The upper and lower envelope at any time should be zero symmetric, ie, the mean of upper and lower envelope should be zero.

The decomposition process is as follows:

- (1) Identify all the maximum and minimum point of the original sequence $x(t)$. Use cubic spline interpolation function to fit the upper and lower envelope ,Take all the mean of upper and lower envelope as mean envelope recorded as $m_1(t)$.
- (2) Subtract $m_1(t)$ from the signal $x(t)$ and obtain a IMF candidate, i.e., $h(t) = x(t) - m_1(t)$
- (3) Judge whether $h(t)$ meet the IMF conditions, if satisfies, get the first IMF component $c_1(t)$, if not satisfies, regard $h(t)$ as a new sequence to repeat the above steps (1) and (2) until $h(t)$ is a IMF recorded as $c_1(t)$. The residue is recorded as: $r_1(t) = x(t) - c_1(t)$.
- (4) Regard $r_1(t)$ as a new sequence, repeat above steps (1)-(3),and the decomposition is completed until the residue of the nth stage is monotonically or less than the set value.
- (5) The sequence of the residue at this time can represent the overall trend of the original sequence.

The original sequence can be written by EMD decomposition:
$$x(t) = \sum_{i=1}^n c_i(t) + r_n(t)$$

2.2 Relevance Vector Machine

The RVM introduced by Tipping is a nonlinear machine learning algorithm based on positive definite kernel [6]. It is based on the sparse Bayesian learning theory, not only has the advantage of avoid over-fitting same as SVM , but also uses Less kernel function while achieves the same

performance than SVR. Furthermore, the kernel function does not need to meet the Mercer condition and has a better generalization ability.

Given a train dataset $\{\mathbf{x}_n, t_n\}_{n=1}^N$, \mathbf{x}_n is a n dimensions input vectors, t_n is output, then the RVM regression model can be regard as:

$$t_n = y(\mathbf{x}_n; \mathbf{w}) + \xi_n \quad (1)$$

$$y(\mathbf{x}_n; \mathbf{w}) = \sum_{i=1}^N w_i \Psi_i(\mathbf{x}) + w_0 \quad (2)$$

Where $\mathbf{w} = [w_0, w_1, \dots, w_n]$ is the weight vector, ξ_n is the independent white Gaussian noise with variance σ^2 , $p(\xi_n | \sigma^2) = \mathcal{N}(0, \sigma^2)$. $\{\Psi_i(x)\}$ is the nonlinear fundament function. $\Psi_i(\mathbf{x}) = K(\mathbf{x}, \mathbf{x}_i)$, $K(\cdot)$ is the kernel function.

Assuming t_n is independent, the likelihood of the dataset can be written as:

$$p(\mathbf{t} | \mathbf{w}, \sigma^2) = (2\pi\sigma^2)^{-N/2} \exp\left\{-\frac{1}{2\sigma^2} \|\mathbf{t} - \Phi\mathbf{w}\|^2\right\} \quad (3)$$

Where $\mathbf{t} = (t_1, \dots, t_n)^T$, Φ is $N \times (N+1)$ matrix and $\Phi = [\phi(\mathbf{x}_1), \dots, \phi(\mathbf{x}_N)]^T$,

$$\phi(\mathbf{x}_n) = (1, K(\mathbf{x}_n, \mathbf{x}_1), K(\mathbf{x}_n, \mathbf{x}_2), \dots, K(\mathbf{x}_n, \mathbf{x}_N))$$

Due to using maximize likelihood (3) to obtain \mathbf{w} and σ^2 to estimate t would lead to over-fitting, a zero-mean Gaussian prior distribution is adopted to \mathbf{w} based on sparse Bayes theory:

$$p(\mathbf{w} | \boldsymbol{\alpha}) = \prod_{i=0}^N \mathcal{N}(w_i | 0, \alpha_i^{-1}) \quad (4)$$

Where $\boldsymbol{\alpha}$ is $N+1$ dimensional hyper parameters vector. In this way, each super-parameter only corresponds to a weight respectively, which controls the influence of the prior distribution on each parameter to ensure the sparseness of the correlation vector model. After defining the prior distribution and the likelihood distribution, the posterior distribution of all unknown parameters can be obtained according to the Bayesian principle:

$$p(\mathbf{w} | \mathbf{t}, \boldsymbol{\alpha}, \sigma^2) = \frac{p(\mathbf{t} | \mathbf{w}, \sigma^2) p(\mathbf{w} | \boldsymbol{\alpha})}{p(\mathbf{t} | \boldsymbol{\alpha}, \sigma^2)} = (2\pi)^{\frac{(N+1)}{2}} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2} (\mathbf{w} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{w} - \boldsymbol{\mu})\right\} \quad (5)$$

Where posterior covariance matrix is $\boldsymbol{\Sigma} = (\mathbf{A} + \sigma^{-2} \Phi^T \Phi)^{-1}$, $\mathbf{A} = \text{diag}(\boldsymbol{\alpha})$. Mean vector is $\boldsymbol{\mu} = \sigma^{-2} \Phi^T \mathbf{t}$. In order to determine the model weight, we first need to obtain the optimal value of the hyper parameters, which can be obtained by the following iterative algorithm:

$$\alpha_i^{\text{new}} = \frac{1 - \alpha_i \sum_{ii}}{\mu_i^2} = \frac{\gamma_i}{\mu_i^2} \quad (6)$$

$$(\sigma^2)^{\text{new}} = \frac{\|\mathbf{t} - \Phi\boldsymbol{\mu}\|^2}{N - \sum_{i=1}^N (1 - \alpha_i \sum_{ii})} = \frac{\|\mathbf{t} - \Phi\boldsymbol{\mu}\|^2}{N - \sum_{i=1}^N \gamma_i} \quad (7)$$

Where μ_i is the i -th mean of posterior weight; γ_i can be interpreted as the variable that measures the optimal parameter n ; \sum_{ii} is the i -th diagonal element of the posterior covariance matrix; N is the number of samples.

If a new input x_* is given, the corresponding output probability distribution is

$$p(\mathbf{t}_* | \mathbf{t}, \boldsymbol{\alpha}_{MP}, \sigma_{MP}^2) = \int p(\mathbf{t}_* | \mathbf{w}, \sigma_{MP}^2) p(\mathbf{w} | \mathbf{t}, \boldsymbol{\alpha}_{MP}, \sigma_{MP}^2) d\mathbf{w} \quad (8)$$

Which obey the Gaussian distribution.

Where $\boldsymbol{\alpha}_{MP}, \sigma_{MP}^2$ is the optimal value used to prediction after iteration.

The prediction variance is. $y_* = \mu^T \phi(\mathbf{x}_*)$. Prediction mean is $\sigma_*^2 = \sigma_{MP}^2 + \phi(\mathbf{x}_*)^T \Sigma \phi(\mathbf{x}_*)$
 y_* can be regard as the prediction value of \mathbf{t}_* .

3. Case Analysis

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 surface monitoring point is shown Figure 2.

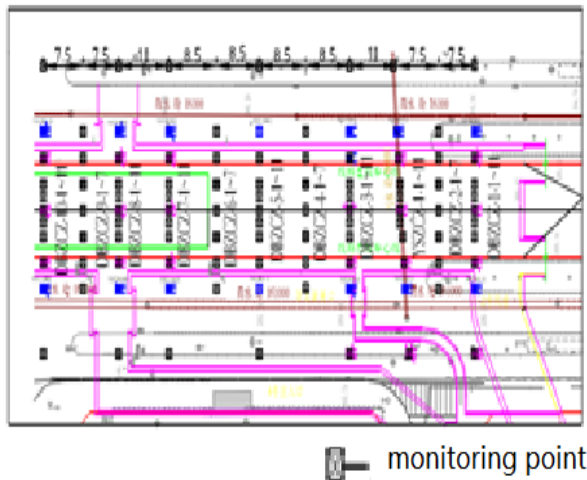


Figure 2. Surface Monitoring Points Location

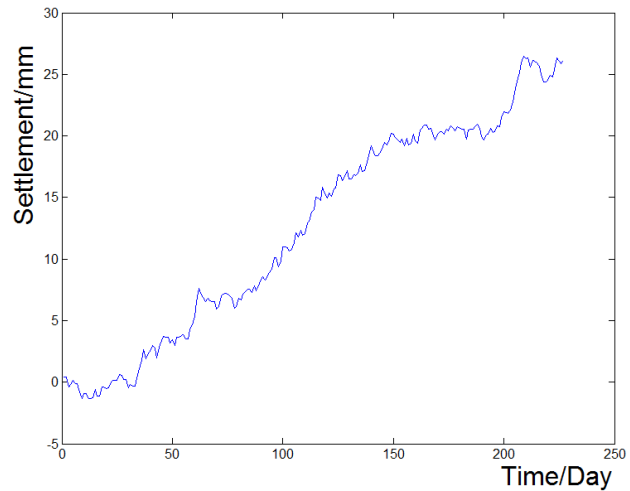


Figure 3. Curve of measured data

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 showed in Figure3 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. The former five days measured was used to predict the settlement value of expected day. Once five new measured datas was obtained, they are added to the training set to update the prediction madel. Using EMD, the train samples time sequence was decomposed into six IMFs as show in Figure4.

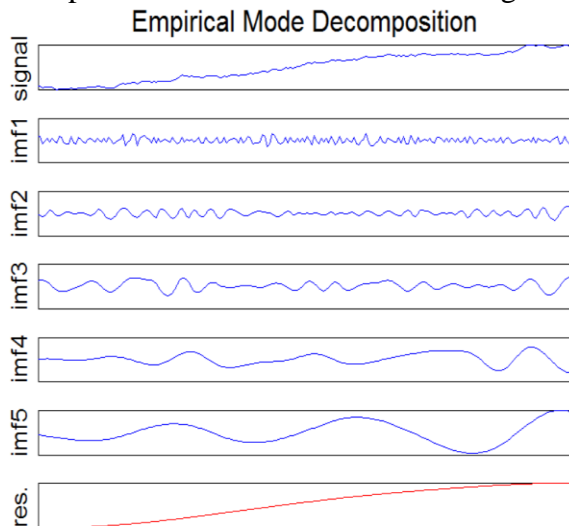


Figure 4. Result of EMD

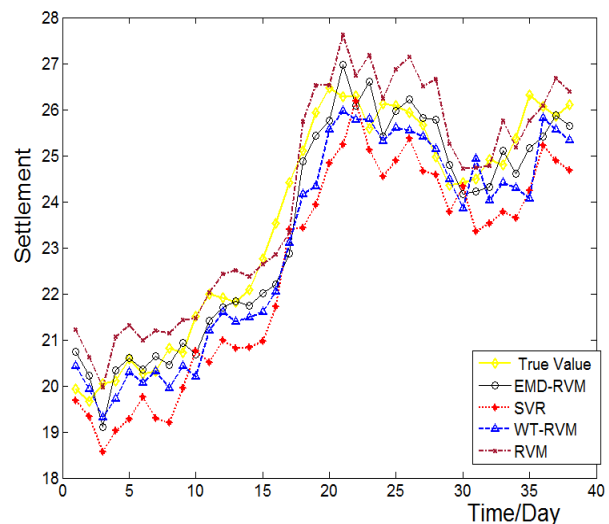


Figure 5. Results of prediction

TABLE 1. Result of Evaluation

| Prediction Model | Evaluation Criteria | | | | | |
|------------------|---------------------|---------|-------------|---------|-------------|---------|
| | DBZCZ-01-01 | | DBZCZ-11-03 | | DBZCZ-12-02 | |
| | RMSE(mm) | MAPE(%) | RMSE(mm) | MAPE(%) | RMSE(mm) | MAPE(%) |
| SVR | 1.2052 | 4.5938 | 1.3695 | 4.9847 | 1.3275 | 4.5275 |
| RVM | 0.7989 | 3.3947 | 0.8536 | 3.8936 | 0.7474 | 3.4807 |
| WT-RVM | 0.7268 | 2.6720 | 0.7951 | 3.3111 | 0.7119 | 3.2091 |
| EMD-RVM | 0.6284 | 1.9424 | 0.7123 | 2.2463 | 0.6734 | 2.3762 |

We use train samples to train RVM model. The widely-used RBF was chosen as the kernel function. However, the prediction effect of RVM model is overly dependent on its parameters. So Particle Swarm Optimization (PSO) algorithm can be chosen to optimize the parameters of RVM model. To avoid blind searching, the model parameter initialization range is:

$$c = [0, 1000] \quad \sigma = [0, 10]$$

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.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - x_i^*)^2} \quad (9)$$

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left(\frac{x_i - x_i^*}{x_i} \right) \times 100 \quad (10)$$

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 EMD-RVM model, a single SVR, a single RVM, a WT-RVM model is used to be compared with DBZCZ-01-01 monitoring point data. The prediction results are shown in Fig. 65 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 table 1 that two evaluation criteria of single RVM model are better compared with single SVR model. It proves that RVM model is better than SVR model. Also the hybrid model weather combined with WT or EMD are better than single RVM. and EMD-RVM model is better than WT-RVM. The formulation of decomposing the original ground settlement time series was right. The prediction accuracy of EMD-RVM model improves significantly.

4. Conclusion

(1) 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.

(2) Due to the nonlinear and non-stationarity of the ground settlement time sequence of subway construction, the sequence is decomposed into several stationary IMFs by EMD method. According to the different characteristics of the sequence, different models are established. The final prediction value is the sum of prediction values of the models at the same time which conforms to the physical change process of ground settlement.

(3) 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.

Acknowledgments

The author wishes to express his gratitude to colleagues, friend and classmates for help discussion and data collection and valuable comments.

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] S. S. Leu, H. C. Lo,, “Neural-network based regression model of ground surface settlement induced by deep excavation,” *Automation in Construction* 13,279–289
- [3] Cannas, B., Fanni, A., See, L., Sias, G., 2006. Data preprocessing for river flow fore- casting using neural networks, wavelet transforms and data partitioning. *Phys.Chem. Earth* 31 (18), 1164–1171.
- [4] Huang, N.E., Wu, Z.H., 2008. A review on Hilbert–Huang transform: method and its applications to geophysical studies. *Rev. Geophys.* 46, 1–23, RG 2006.
- [5] Huang, N.E., Shen, Z., Long, S.R., 1999. A new view of nonlinear water waves: the Hilbert spectrum. *Annu. Rev. Fluid Mech.* 31, 417–455.
- [6] Rasmussen, C. E., Williams, C., *Gaussian Process for Machine Learning*. Cambridge, MA: The MIT press, 2006