Settlement Prediction of Road Soft Foundation Using a Support Vector Machine (SVM) Based on Measured Data

Huiling YU[1] and Yunlong SHANGGUAN[2,*]

[1]JiLin Communications polytechnic, Changchun, 130021, China

[2]JiLin Provincial Transport Scientific Research Institute, Changchun, 130001, China

Abstract. The suppor[1]t vector machine (SVM) is a relatively new artificial intelligence technique which is increasingly being applied to geotechnical problems and is yielding encouraging results. SVM is a new machine learning method based on the statistical learning theory. A case study based on road foundation engineering project shows that the forecast results are in good agreement with the measured data. The SVM model is also compared with BP artificial neural network model and traditional hyperbola method. The prediction results indicate that the SVM model has a better prediction ability than BP neural network model and hyperbola method. Therefore, settlement prediction based on SVM model can reflect actual settlement process more correctly. The results indicate that it is effective and feasible to use this method and the nonlinear mapping relation between foundation settlement and its influence factor can be expressed well. It will provide a new method to predict foundation settlement.

1 Introduction

The engineering projects based on the soft soil are also developed rapidly with the speedily developing of the economy in China. It becomes more and more a vital problem that how we can design and schedule reasonably in a soft soil engineering in the condition that certain work time and certain settlement after construction must be satisfied. And the only way to solve the problem is the settlement prediction.

Research about settlement prediction has been carried out broadly and profoundly by the scholars all over the world, which as a result gains a lot. But the mechanism of the soil settlement is very complicated since the settlement is influenced by the loading and the soil-properties are so complexed.1fie methods for predicting foundation settlement in existence are always limited by the simplified presumptive conditions, which as a result causes the inconsistence between the calculated and the observed value. It is necessary to make more profound research in the field of settlement prediction.

At present, there are mainly three methods which are Asaoka method [1], hyperbolic method [2], grey theory method [3] are used for prediction of foundation settlement. With the development of artificial intelligence, artificial neural network method and support

---

* Corresponding author: 5918445@126.com

vector machine have been widely used in soft foundation settlement prediction. This paper uses the SVM model to predict the settlement of soft foundation and compared with BP artificial neural network model and traditional hyperbola method. The prediction results indicate that the SVM model has a better prediction ability than BP neural network model at the same training set mean-square error.

2 Methodology

The support vector machine model (SVM) is a universal approach for solving the problems of high-dimensional function estimation. This method was introduced in the 1990s based on Vapnik–Chervonenkis (VC) statistical learning theory [4, 5].

This method employs the structural risk minimization (SRM) principle to minimize errors in the model, whereas other methods, such as ANNs, apply empirical risk minimization (ERM) [6, 7].

The formulation is based on Structural Risk Minimization (SRM) which has been shown to be superior to the Empirical Risk Minimization (ERM) used in conventional neural networks [4].This section of the paper serves an introduction to this relatively new technique. Briefly, SVM model estimates the regression by using a set of linear functions, although, it can be developed using non-linear regression [8].

A regression problem is considered having a set of training vectors (D) belonging to the SVM applied to a dataset. Suppose that we are given a training set $(x_1, y_1), \ldots, (x_l, y_l), x \in R^m, y \in R$. For a set of data, we can use is a linear function *f(x)* to solve the regression problem as given below:

$$f(x) = w \cdot x + b \tag{1}$$

where x is the input vector; y is the output value; w is the weight vector; b is the bias and $< w \cdot x >$ shows the inner product of the two vectors, w and x. And the main objective is to find a function *f(x)*, that has almost ε deviation from the actual targets y_i given by the training data and at the same time, is as flat as possible [4].

A loss function with an ε-insensitive zone is also defined as:

$$N_\varepsilon(y) = |y - f(x)|_\varepsilon = \begin{cases} 0 & |y - f(x)| \le \varepsilon \\ |y - f(x)| - \varepsilon & \text{otherwise} \end{cases}$$

$$\tag{2}$$

where $N_\varepsilon(y)$ is the loss function and $\varepsilon > 0$ is a constant value.

The goal is to identify a function that has the greatest ε deviation from the target actually obtained for all the training data and at the same time as flat as possible. It has been shown that, in order to achieve this objective, the following function must be minimized:

$$T(w, \alpha^*, \alpha) = \frac{\|w\|^2}{2} + C(\sum \alpha^* + \sum \alpha) \qquad \text{Subject to} \begin{cases} y_i - ((w.x) + b) \le \varepsilon + \alpha_i \\ ((w.x) + b) - y_i \le \varepsilon + \alpha_i^* \\ \alpha_i, \alpha_i^* \ge 0 \end{cases}$$

$$\tag{3}$$

Where $\|w\|^2$ is the Euclidian norm of weight vector, C is a constant known as the penalty parameter which is larger than 0; α_i and α_i^* are slack variables.

Finally, non-linear regression within a SVM can be achieved by mapping the training patterns into a higher dimensional feature space (where the linear regression is feasible) and then applying the linear regression algorithm in that feature space.

3 Case Study

The load-settlement time series curves are obtained through the test of loading for the selected road and the measurement of surface settlement. The subgrade of selected road for test mainly is soft soil with high water content, low permeability which belong to high compressible subgrade soil. Construction of subgrade were divided into three stages. The first layer of geogrid and broken stone hardcore were built in 1 April 2013, and the settlement measurement was carried out on the next day (t=0d), and the subgrade was loaded to the designed elevation in 5 November 2013(t=243d). The load-settlement time series curves can be seen in Fig.1.
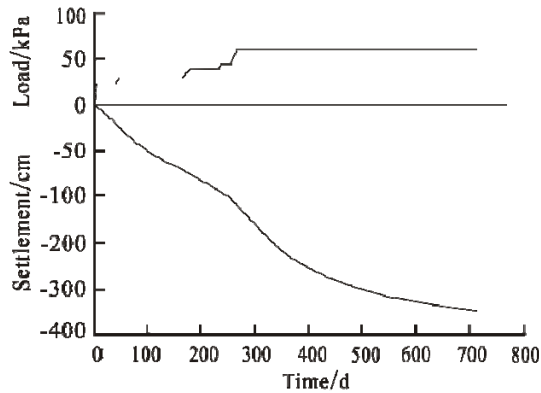

Fig. 1 The load-settlement time series curves

There are 80 measured settlement data of the selected road for test during nearly two years after the subgrade was constructed. And we choose 10 settlement data from 400d to 720d which lasted nearly one year as the tested sample data set, the rest data from the beginning to 400d were regarded as the training sample data set. Generally, the principle factors affected subgrade settlement are engineering properties of soft soil, load characteristics and subgrade characteristics. Due to the properties and the basic engineering properties of the foundation soil for a typical cross section can be regarded as constant, therefore this article selects the current embankment load p and the current time t as input variables, and takes the value of settlement at current time t as output variables. The prediction and results contrast were carried out by SVM and BP neural network model. This paper also chose the traditional hyperbolic curve method which is commonly used in practical engineering to validate the veracity of SVM model.

Koski kernel function was used in the SVM model, the values of parameters C, γ and ε are 250, 1 and 0.001, respectively. The value of ε controls the size of function fitting error so that it can govern the numbers of support vector and it's generalization capability, and the value of ε is 0.001 is suitable. The values of parameters C, γ have direct influence on predication accuracy. The parameters in this paper are selected through testing a large number of parameters. Fig.2 represents that the effect of parameter C on average relative error of test set when γ equals to 1 constantly. Fig.3 represents that the effect of parameter γ on average relative error of test set when C equals to 1 constantly. ME represents the mean relative error.

The settlements prediction was conducted based on the same training data using SVM model, BP neural network model and the traditional hyperbolic curve, respectively. Modeling of BP neural network used MATLAB neural network toolbox, the number of neurons in the input layer, hidden layer and output layer are 2, 3 and 1 respectively. Termination condition of training the BP neural network is that the mean square error (MSE) is up to the same with the SVM model. The traditional method of settlement prediction is

based on the analysis of the settlement observation data with the constant load, and settlement observation data from 250d to 400d are chosen in the hyperbolic curve method.
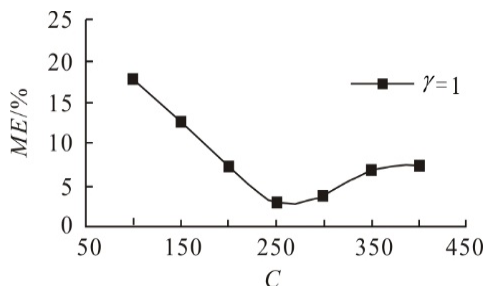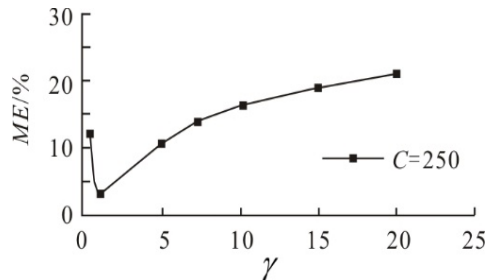
Fig. 2 Effect of *C* on ME of test set   Fig. 3 Effect of γ on ME of test set

4 Results

The results are shown in Table 1 based on three method. The mean relative error (ME) and maximum absolute error are used as measures of prediction accuracy and the result can be seen in Table 2.

Table 1 Relative Error And Absolute Error Of Different Prediction Methods

Time /d	Measured values /cm	SVM Model		BP neural network		Hyperbolic curve	
		Relative error/%	Absolute error/%	Relative error/%	Absolute error/%	Relative error/%	Absolute error/%
400	166.4	-1.94	-2.18	-1.67	-2.79	0.84	1.74
405	166.9	-3.45	-3.62	-2.93	-3.79	0.45	0.68
410	169.4	-1.99	-2.44	-1.43	-2.81	1.03	2.29
420	173.5	-2.34	-2.51	-3.45	-4.51	2.52	2.98
450	174.6	-3.12	-3.97	-4.62	-5.91	2.59	3.17
480	182.5	1.28	2.07	2.61	4.27	3.30	3.74
520	186.4	-0.60	-1.44	-3.57	-4.47	1.63	2.63
570	196.9	-3.54	-4.09	-5.67	-5.99	4.69	6.19
630	209.3	-2.12	-5.45	-5.62	-5.79	5.94	7.16
700	213.5	2.43	5.27	-4.83	-5.27	7.17	9.98

The comparation among three methods are shown in Table 1 and Table 2. The mean relative error and maximum absolute error of the prediction results used BP neural network model are greater than the SVM model averagely in the case of the same training set mean square error. It illustrates that the generalization (prediction) capability of SVM is better than BP neural network model.

The prediction results used the hyperbolic curve method is greater than measured settlement values and this situation is coincidence with the practical engineering experience. The prediction capability of SVM is better than the hyperbolic curve method by comparing these two methods. And SVM model used settlement observation data of hierarchical loading process as training sample set comprehensively, while the hyperbolic curve method

only used the constant load after loading. This reflects the engineering mechanical properties of soft soil foundation with the gradual improvement of the loading process, so the SVM model can reflect the deformation trend of soft soil foundation.

The prediction accuracy of these three method are high in early prediction period which relative error are under 5%. And the mean relative error and maximum absolute error of the prediction results increase after 570d, but the results can meet the actual needs of the project. It also illustrates that the prediction accuracy is greatly affected by the selection of prediction time, and the way to select the optimal prediction time is to be further studied. The prediction accuracy of SVM model is best by comparing the three methods, so the SVM model is applied to predict the settlement of soft foundation.

Table 2 Me And Maximal Absolute Error

Parameter	SVM Model	BP neural network	Hyperbolic curve
Mean relative error/%	2.16	3.64	3.02
Maximum absolute error/%	5.27	5.99	9.98

5 Conclusions

The application of the support vector machine (SVM) model in settlement of soft foundation prediction is presented in this study. Calculation results of engineering projects show that the prediction capability of SVM is better than BP neural network model in the case of the same training set mean square error. And SVM model used settlement observation data of hierarchical loading process as training sample set comprehensively, so it is more rational than the hyperbolic curve method which only considers the constant load after loading. The SVM model is simpler to apply than the conventional methods and it is feasible to apply the support vector machine to predict the settlement of highway soft foundation.

References

1. Asaoka A, Observational procedure of settlement prediction. Soils & Foundations, 1978, 18(4): 87－101.
2. Qian YL, Ding Y, Forecasting settlement of roadbed and its engineering application. Journal of Yangzhou University (Natural Science Edition), 2001, 4(2): 75－78.
3. Chen XD, Xia J. Xu Q, Differential hydrological grey model (DHGM) with self-memory function and its application to flood forecasting. Science China Technological Sciences, 2009, 39 (2), 341-350
4. Cortes, C., Vapnik, V., Support vector network. Mach. Learn, 1995, 20, 273–297.
5. Cherkassky, V., Mulier, F., Learning from Data: Concepts, Theory and Methods. Wiley, New York, 1998.
6. Dibike YB, Velikov S, Solomatine D, Abbott MB. Model induction with support vector machines: introduction and applications. J Comput Civ Eng ASCE, 2001, 15(3):208–16.
7. Cristianini N, Shaw-Taylor J. An introduction to support vector machines and other kernel-based learning methods. Cambridge: Cambridge University Press, 2000.
8. Samui P. Support vector machine applied to settlement of shallow foundations on cohesionless soils. Comput Geotech, 2008, 35:419–27.