

# Structural health monitoring of concrete dams using least squares support vector machines

\*Fei Kang<sup>1)</sup>, Junjie Li<sup>2)</sup>, Shouju Li<sup>3)</sup> and Jia Liu<sup>4)</sup>

<sup>1), 2), 4)</sup> *School of Hydraulic Engineering, Dalian University of Technology, Dalian 116024, China*

<sup>3)</sup> *State Key Laboratory of Structural Analysis for Industrial Equipment, Dalian University of Technology, Dalian 116024, China*

<sup>1)</sup> [kangfei@dlut.edu.cn](mailto:kangfei@dlut.edu.cn) or [kangfei2009@163.com](mailto:kangfei2009@163.com)

## ABSTRACT

This study presents a least squares support vector machine (LSSVM) based displacement prediction model for health monitoring of concrete dams. LSSVM is a novel machine learning technique. The model can produce similar good generalization performance and learns faster than the basic support vector machines in engineering problems. The advantages such as high prediction accuracy, fast training speed of the LSSVM model are verified using monitoring data of a real concrete dam. Results are also compared with that of the multiple linear regression and stepwise regression models for dam health monitoring.

## 1. INTRODUCTION

Dam safety is a critical issue to avoid huge losses of personal and property of residents living in the downstream of the dams caused by failure of dams. Therefore, it is necessary to adopt reasonable and effective methods to monitor and estimate running states of dams, and simultaneously to forecast the dam future behaviors with high accuracy. Structural health monitoring based on quantities that can reflect behaviors of concrete dams, like horizontal and vertical displacements, rotations, stresses and strains, seepage, etc., is an important method to evaluate operational states of concrete dams correctly and predict the future structural behaviors accurately.

Predictive models are an important element in dam health monitoring (DHM) system. They provide an estimate of the dam response faced with a given load combination, which can be compared with the actual measurements to draw conclusions about dam safety (Salazar et al. 2015). The statistical, deterministic and hybrid models are three classical approaches to describe and evaluate the behaviors of

---

<sup>1)</sup> Associate Professor

<sup>2)</sup> Professor

<sup>3)</sup> Professor

<sup>4)</sup> Master candidate student

concrete dams by analyzing the real-time measured data of concrete dams. In these numerical simulation models, the main variables considered include hydrostatic pressure, environment temperature and time effects when analyzing the structural responses of the concrete dams (Leger and Leclerc 2007; Tatin et al. 2015). The statistical model has several advantages of simple formulation, the faster speed of execution and the availability of any type of correlation between effective variables and response variables. Nevertheless, the MLR suffers from defects of the multicollinearity existing among independent variables and near-linear dependence of regressors. Multicollinearity can dramatically influence the variances and co-variances for the least square estimators of the regression coefficients.

Recently, computational intelligence methods, artificial neural networks (Kang et al. 2016b), Gaussian process regression (Kang et al. 2015), support vector machines (SVM) (Kang and Li 2015; Kang et al. 2016a; Kang et al. 2016c) are verified to be powerful approaches for engineering prediction problems. Conventional models for dam health monitoring have been enriched with some various artificial intelligence techniques, such as artificial neural networks (Mata 2011; Kao and Loh 2013), neuro-fuzzy models (Ranković et al. 2012), artificial immune algorithms (Xi et al. 2011), support vector machines (Ranković et al. 2014; Su et al. 2016). Salazar et al. (2015) has made a comparison on several machine learning techniques for dam behavior modeling. Least squares support vector machines (LSSVM) is introduced by Suykens et al. (2002) as a variation of the conventional support vector machines (Vapnik 2000). LSSVM not only has the same advantages of SVM, such as retaining the principle of SVM and having good generalization capability, but also has additional advantages over the SVM method. LSSVM uses least square errors instead of nonnegative errors to the loss function. As a result, LSSVM expresses model training in terms of solving a linear system instead of a quadratic programming problem, which thus reduces the computation time significantly. The LSSVM model showed the excellent performance in the various prediction fields (Kang et al. 2016a).

In this study, the LSSVM algorithm is adopted to establish the health monitoring model of concrete dams. The performance of the proposed model is demonstrated on a real concrete gravity dam. Results are also compared with that of the multiple linear regression and stepwise multiple regression models. The rest of the paper is organized as follows. In Section 2, concrete dam health monitoring is stated. In Section 3, LSSVM based model is introduced. Simulation results are presented in Section 4, and in Section 5, conclusions are provided.

## **2. DAM HEALTH MONITORING PROBLEM**

Dam health monitoring generally known as dam surveillance is a process of observation and assessment for running states of concrete dams with the intent to gain insights into behaviors of concrete dams, to detect anomalies, and to enable a timely response either in the form of repairs, reservoir management by means of analyzing the collected data of concrete dams (Ahmadi-Nedushan 2002; Bukenya et al. 2014). Data obtained from static health monitoring system are generally analyzed by using statistical models based on correlations between environmental factors and dam responses (Mata et al. 2014).

The statistical model is a classical approach for data analysis and is employed in various domains. The model is also known as quantitative interpretation model, used to predict the horizontal displacements of concrete dams based on the gained long-term monitoring data of concrete dams and mathematical statistics analysis methods. The external effective factors, such as the water level difference between the upstream and the downstream ( $H$ ), ambient temperature ( $T$ ) and time ( $t$ ) are taken as the independent variables, and the monitored effect quantity like horizontal displacement  $y$  is taken as the dependent variable. The effects of external loads caused from effective factors for behaviors of concrete dams mainly consist of two parts: a part of reversible effects ( $y_H$ ,  $y_T$ ) deriving from the variations of the hydrostatic load because of changes of reservoir levels and the load on account of temperature changes, and another part of irreversible effect ( $y_\theta$ ) due to the evolution of the concrete dam response over time (Salazar et al. 2015). Therefore, it is considered that the sum of effects with respect to a limited period can be approximated as,

$$y_D = y_H + y_T + y_\theta + E_0 \quad (1)$$

where  $y_D$  is the observed horizontal displacements,  $y_H$  is the horizontal displacements due to the hydrostatic load,  $y_T$  is the horizontal displacements caused from the influence of temperature changes,  $y_\theta$  is the horizontal displacement on account of the evolution of the dam response over time and constant  $E_0$  is introduced as the zero point of measurement.

The displacements  $y_H$  due to the hydrostatic load can be described by a polynomial function of the upstream reservoir water depth ( $H$ ) as

$$y_H = a_0 + a_1H + a_2H^2 + a_3H^3 \quad (2)$$

where  $a_i$  ( $i = 0, 1, 2, 3$ ) are unknown coefficients need to be estimated.

the effect of temperature variations can be described as a sinusoidal functions with one year period and six months period (Leger and Leclerc 2007; Mata 2011), and the relationship of them can be described by trigonometric functions as

$$y_T = b_0 + b_1 \sin(d) + b_2 \cos(d) + b_3 \sin^2(d) + b_4 \sin(d) \cos(d) \quad (3)$$

where  $d = 2\pi g/365$ ,  $g$  ( $1 \leq g \leq 365$ ) is the time in days from the observation date to the beginning of the year.

Aging component  $y_\theta$  is a kind of irreversible component evolving toward a certain direction over time, the link between  $y_\theta$  and time follows a curvilinear relationship, therefore, displacement contributed by aging can be modeled as

$$y_\theta = c_0 + c_1\theta + c_2 \ln \theta \quad (4)$$

where  $c_k$  ( $k = 0, 1, 2$ ) are unknown coefficients;  $\theta = t/100$ ,  $t$  is the number of days since the beginning of the analysis. These undetermined coefficients can be computed by minimizing the difference between the real value of measurement and the calculated value of Eq. (1) by using the least squares minimum method.

### 3. LSSVM MODEL FOR DAM HEALTH MONITORING

#### 3.1 LEAST SQUARES SUPPORT VECTOR MACHINES

Support vector machines (Vapnik 2000) are powerful machine learning techniques based on the statistical learning theory. The structural minimization adopted by SVM could provide better generalization ability than the empirical risk minimization used by the traditional methods.

Given a data set  $\{(\mathbf{x}_i, y_i), i = 1, \dots, l\}$ , where  $\mathbf{x}_i$  is a  $D$ -dimensional input vector, and  $y_i$  is a scalar output or target. The nonlinear relationship between the input and the output can be described by a regression function as

$$f(\mathbf{x}) = \mathbf{w}^T \varphi(\mathbf{x}) + b, \quad (5)$$

where  $f(\mathbf{x})$  denotes the forecasting values,  $\varphi(\mathbf{x})$  is a nonlinear mapping function, and  $\mathbf{w}$  and  $b$  are the coefficients to be adjusted.

The commonly used SVM regression model is called  $\varepsilon$ -support vector regression (SVR). The primal optimization problem for SVR can be defined as follows

$$\text{Minimize } R(\mathbf{w}, \boldsymbol{\xi}, \boldsymbol{\xi}^*) = \frac{1}{2} \|\mathbf{w}\|^2 + \gamma \sum_{i=1}^l (\xi_i + \xi_i^*), \quad (6)$$

$$\text{Subject to } \begin{cases} y_i - \mathbf{w}\varphi(\mathbf{x}_i) - b \leq \varepsilon + \xi_i \\ \mathbf{w}\varphi(\mathbf{x}_i) + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases}, \quad (7)$$

where  $\gamma$  is the regularization parameter that balances model complexity and approximation accuracy, and  $\xi_i, \xi_i^*$  are the slack variables.

Compared with the standard SVM, LSSVM applies linear least squares criteria for the loss function instead of inequality constraints with equality ones. The following optimization problem is formulated:

$$\text{Minimize } R(\mathbf{w}, \mathbf{e}) = \frac{1}{2} \|\mathbf{w}\|^2 + \frac{1}{2} \gamma \sum_{i=1}^l e_i^2, \quad (8)$$

$$\text{Subject to } y_i = \mathbf{w}^T \varphi(\mathbf{x}_i) + b + e_i, \quad i = 1, \dots, l, \quad (9)$$

where  $e_i$  is the error between the actual output and the predictive output of the  $i$ th sample.

The Lagrange function can be constructed as

$$L(\mathbf{w}, b, \mathbf{e}, \boldsymbol{\alpha}) = R(\mathbf{w}, \mathbf{e}) - \sum_{i=1}^l \alpha_i (\mathbf{w}^T \varphi(\mathbf{x}_i) + b + e_i - y_i), \quad (10)$$

where  $\alpha_i$  is the  $i$ th Lagrange multiplier. The Karush-Kuhn-Tucker (KKT) conditions for optimality are given by

$$\begin{aligned} \frac{\partial L}{\partial \mathbf{w}} = 0 &\rightarrow \mathbf{w} = \sum_{i=1}^l \alpha_i \varphi(\mathbf{x}_i), & \frac{\partial L}{\partial b} = 0 &\rightarrow \sum_{i=1}^l \alpha_i = 0, \\ \frac{\partial L}{\partial e_i} = 0 &\rightarrow \alpha_i = \gamma e_i, & \frac{\partial L}{\partial \alpha_i} = 0 &\rightarrow \mathbf{w}^T \varphi(\mathbf{x}_i) + b + e_i = y_i. \end{aligned} \quad (11)$$

After eliminating  $e_i$  and  $\mathbf{w}$ , the following linear equation set is obtained:

$$\begin{bmatrix} 0 & \mathbf{I}_l^T \\ \mathbf{I}_l & \boldsymbol{\Omega} + \gamma^{-1} \mathbf{I} \end{bmatrix} \begin{bmatrix} b \\ \boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{y} \end{bmatrix}, \quad (12)$$

where  $\mathbf{I}_l = [1, 1, \dots, 1]^T$  is an  $l$ -dimensional column vector,  $\mathbf{I}$  is the identity matrix  $\mathbf{y} = [y_1, y_2, \dots, y_l]^T$  is the output vector of the training set;  $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \dots, \alpha_l]^T$  and  $b$  are the parameters of the LSSVM model;  $\boldsymbol{\Omega}$  is a  $l \times l$  symmetric matrix, the elements of which are given by:

$$\Omega_{ij} = \varphi(x_i)^T \varphi(x_j) = K(x_i, x_j). \quad (13)$$

By solving Eq. (12), LSSVM model parameters can be obtained as

$$\begin{cases} \boldsymbol{\alpha} = (\boldsymbol{\Omega} + \gamma^{-1} \mathbf{I})^{-1} (\mathbf{y} - b \mathbf{I}_l) \\ b = \frac{\mathbf{I}_l^T (\boldsymbol{\Omega} + \gamma^{-1} \mathbf{I})^{-1} \mathbf{y}}{\mathbf{I}_l^T (\boldsymbol{\Omega} + \gamma^{-1} \mathbf{I})^{-1} \mathbf{I}_l} \end{cases}. \quad (14)$$

The resulting LSSVM model for function estimation becomes

$$y(\mathbf{x}) = \mathbf{w}^T \varphi(\mathbf{x}) + b = \sum_{i=1}^l \alpha_i K(\mathbf{x}, \mathbf{x}_i) + b, \quad (15)$$

where  $\boldsymbol{\alpha}$  and  $b$  are the solutions to Eq. (12),  $K(\mathbf{x}, \mathbf{x}_i)$  is the kernel function. The architecture of LSSVM is shown in Fig. 1.

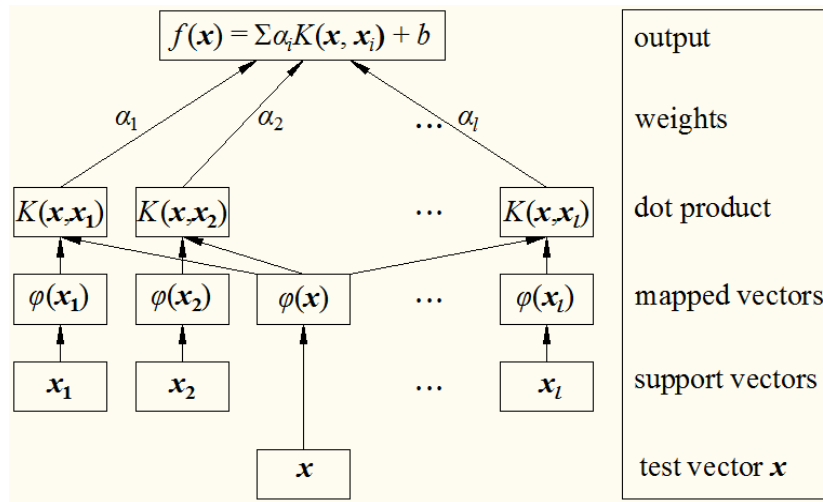


Fig. 1. Architecture of LSSVM

The most applicable kernel function for nonlinear regression is the Gaussian radial basis function (RBF) kernel

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right), \quad (16)$$

where  $\sigma$  is the kernel parameter. RBF kernel is adopted in this study and there are two user-determined parameters  $\gamma$  and  $\sigma$ . The selection of these parameters is crucial to the performance of LSSVM. Optimization algorithms can be adopted to obtain the optimal hyper-parameters for the regression model (Kang and Li 2015; Kang et al. 2016a).

### 3.2 PROCEDURE OF THE PROPOSED APPROACH

LSSVM is adopted to establish the displacement prediction model for dam health monitoring. The procedure of the LSSVM model for DHM can be described as follows:

Step 1: Determine the research problem, and select the effective factors  $\mathbf{x}$  that correlated with response variable  $t$ .

Step 2: Collect data from concrete dam monitoring system to build sample sets for the LSSVM model establishment, including the training sample set and testing sample set. The data should be normalized before establish the model.

Step 3: Turning parameters of the LSSVM model.

Step 4: Establish the LSSVM model. The LSSVM response surface is established using the optimized parameters on the datasets generated in the previous steps.

Step 5: Verify whether the training accuracy obtained from step 4 meets the training error threshold. If the training accuracy achieves the goal, the procedure goes to the next step, otherwise repeat steps 3 to step 5.

Step 6: Test the prediction performance of the LSSVM model established in Step 5 based on the testing samples prepared in Step 2.

The flowchart of the LSSVM model for DHM is shown in Fig. 2.

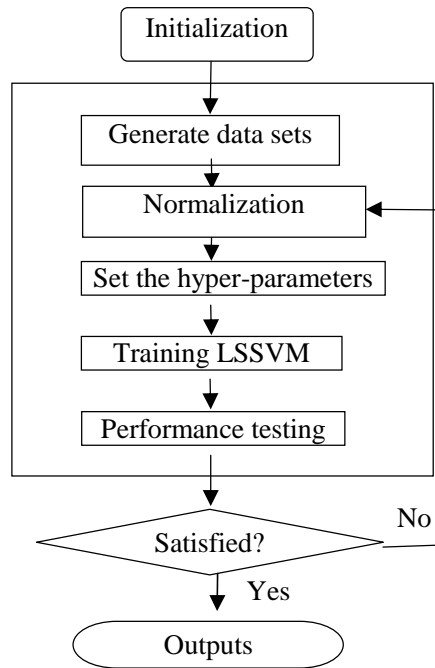


Fig. 2. Flowchart of the LSSVM model for dam health monitoring.

Different health monitoring models are evaluated by several performance criterions, such as the mean absolute error ( $AE_{\text{mean}}$ ), the root mean square error ( $RMSE$ ), the maximum absolute error ( $AE_{\text{max}}$ ) and the correlation coefficient ( $R$ ) shown as follows.

$$AE_{\text{mean}} = \frac{1}{N} \sum_{i=1}^N |y_D(i) - y(i)| \quad (17)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_D(i) - y(i))^2} \quad (18)$$

$$AE_{\max} = \max(y_D(i) - y(i)), \quad i = 1, 2, \dots, N \quad (19)$$

$$R = \frac{\sum_{i=1}^N (y_D(i) - \bar{y}_D)(y(i) - \bar{y})}{\sqrt{\sum_{i=1}^N (y_D(i) - \bar{y}_D)^2 \sum_{i=1}^N (y(i) - \bar{y})^2}} \quad (20)$$

where  $y_D$  and  $\bar{y}_D$  represent the simulated values and the average value;  $y$  and  $\bar{y}$  denote the measured values and the average value; and  $N$  is the number of observations.

## 4. Examples

### 4.1 Data and parameter settings

The proposed procedure is verified on a concrete gravity dam as shown in Fig. 3. It is a 91m high and 1080m long concrete gravity dam which is averagely divided into 60 sections. The horizontal displacement is one of the important measurement items. In this study, the horizontal displacement prediction model of the 18th dam section will be studied. The data between the years 1997 to 2006 is adopted for training the models and the data between the years 2007 to 2009 is adopted for testing the accuracy of the obtained models. The hyper-parameters are set as  $(\gamma, \sigma^2) = (100, 6)$ , which are chosen by a trial-and-error method.

The horizontal displacement at the crest of the 18th dam section is shown in Fig. 4 and the water level of the reservoir is shown in Fig .5.



Fig. 3 Upper stream view of the Fengman dam

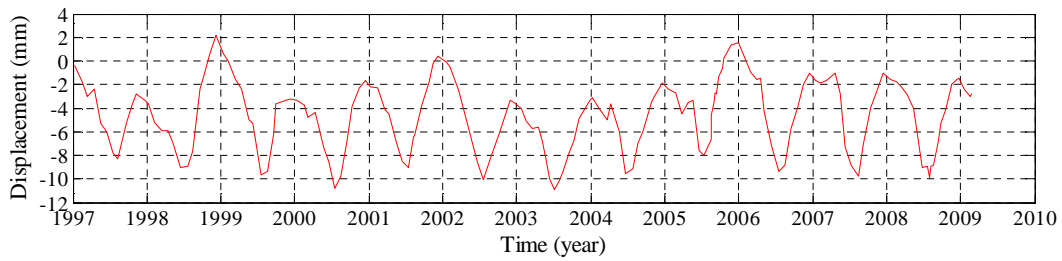


Fig. 4 Horizontal displacement at the crest of the 18th dam section

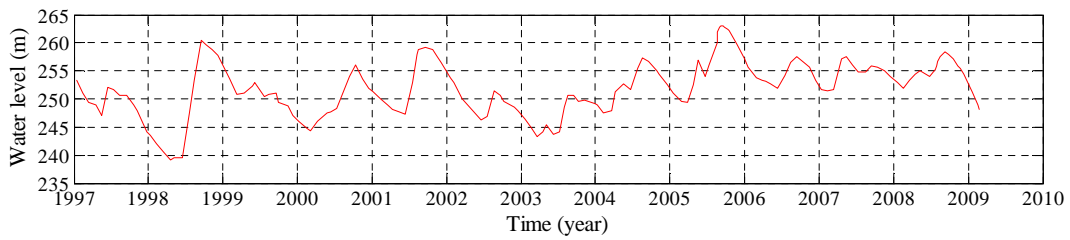


Fig. 5 Water level of the reservoir

Except to the LSSVM model, MLR and stepwise regression (SR) model are established for comparison. The variables are selected according to experience for MLR model. In statistics, stepwise regression includes regression models in which the choice of predictive variables is carried out by an automatic procedure. Stepwise regression can select the suitable variables and delete the others without appreciably increasing the residual sum of squares. The adopted variables for several models are listed in Table 1. The last variable is not adopted since the nonlinear term of aging effect is not prominent considering the dam has been operated for many years.

Table 1 Variables adopted in different forecasting models

Variable	MLR	SR	LSSVM-1	LSSVM-2
$H^1$	√		√	
$H^2$	√		√	
$H^3$	√	√	√	√
sin(d)	√	√	√	√
cos(d)	√	√	√	√
sin <sup>2</sup> (d)	√	√	√	√
sin(d)cos(d)	√	√	√	√
$\theta$	√	√	√	√
ln $\theta$				



## 4.2 Results and performance comparison

Outlier detection is performed with different number of variables. The standardized residual is adopted to judge whether the observation is an outlier. The points outside the boundary of  $\pm 3$  are outliers, since the probability of observations within the range of  $[-3, +3]$  is 97.7%. As shown in Fig. 6, there is only one outlier and we can delete it from the data for modelling.

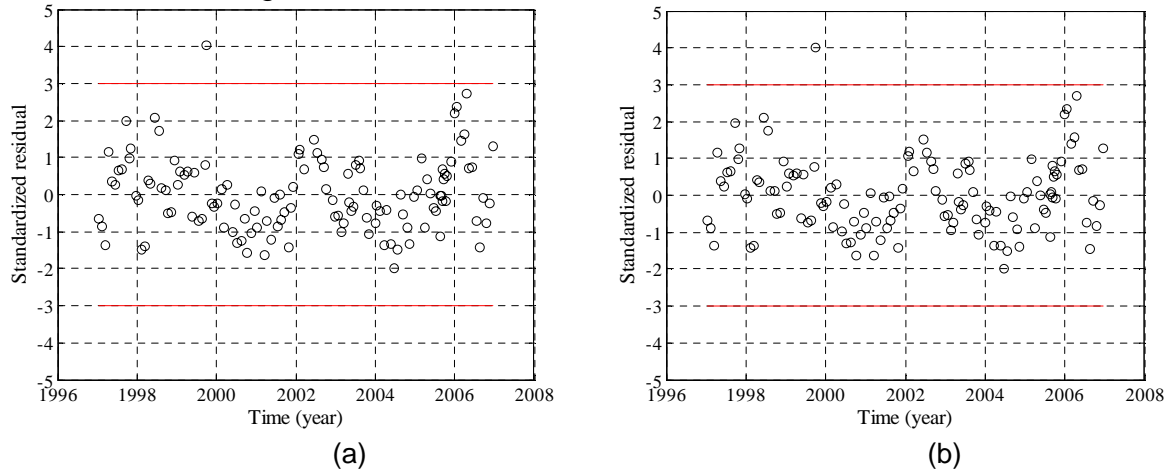


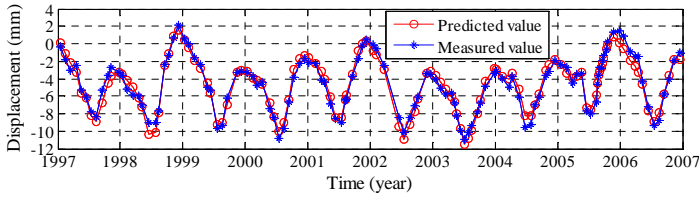
Fig. 6 Outlier detection results using different variables in Table 1: (a) using the first 8 variables; (b) using the variables 3-8.

The accuracy of different prediction models are shown in Table 2. It can be seen that the MLR and SR models have nearly the same accuracy. That means the first two variables in Table 1 is not needed in the regression model. The LSSVM models perform much better than the MLR and SR models considering the error criterion of LSSVM is much smaller than that of the MLR and SR models. The two LSSVM models with different number of variables perform nearly the same. It seems the LSSVM-1 model with more variables according to experience performs better than the LSSVM-2 model. From this point, select the most informative variables and delete the others are not needed for the LSSVM models.

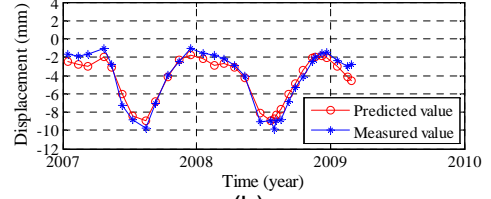
The comparison of predicted values and measured values of crest horizontal displacement are shown in Fig. 7 ~ Fig. 10. The four models can fit the measured values well, and the LSSVM model performs the best. Histograms of residuals for different models are shown in Fig. 11 and Fig. 12. The residuals nearly normal distributed and the residuals of LSSVM models are smaller than that of the statistical models.

Table 2 Accuracy of several prediction models

Algorithm	Training				Test			
	$AE_{max}$	$AE_{mean}$	$RMSE$	$R$	$AE_{max}$	$AE_{mean}$	$RMSE$	$R$
MLR	1.7896	0.4960	0.6145	0.9800	1.7557	0.6706	0.7941	0.9787
SR	1.7496	0.4965	0.6157	0.9799	1.7509	0.6678	0.7907	0.9788
LSSVM-1	1.2771	0.3559	0.4410	0.9898	1.4318	0.5301	0.6540	0.9871
LSSVM-2	1.3604	0.3917	0.4780	0.9880	1.4989	0.5429	0.6736	0.9880

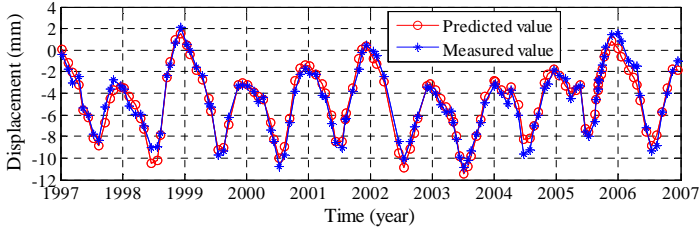


(a)

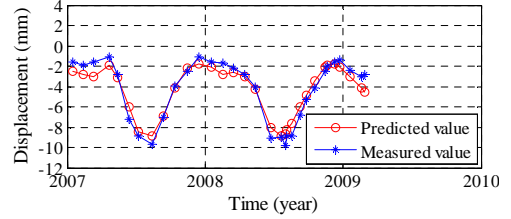


(b)

Fig. 7 Performance of regression model: (a) training; (b) test.

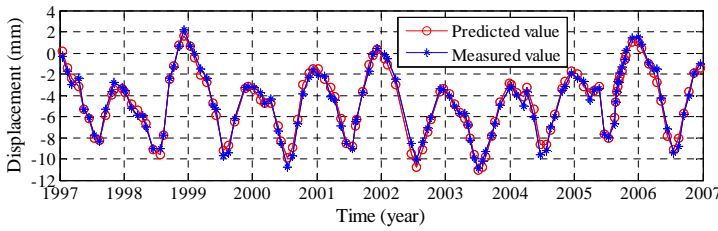


(a)

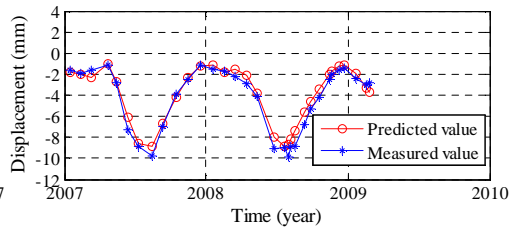


(b)

Fig. 8 Performance of stepwise regression model: (a) training; (b) test.

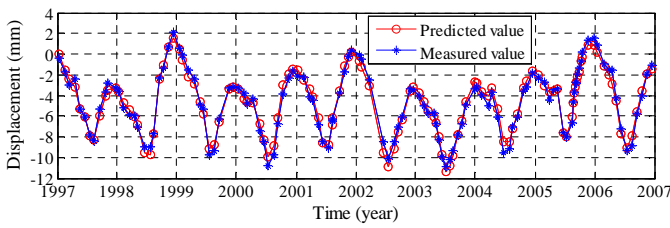


(a)

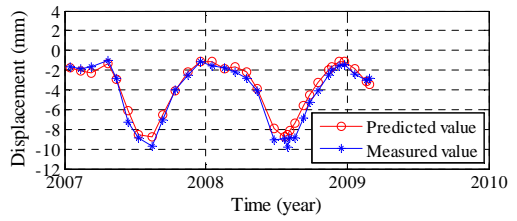


(b)

Fig. 9 Performance of LSSVM-1 model: (a) training; (b) test.



(a)



(b)

Fig. 10 Performance of LSSVM-2 model: (a) training; (b) test.

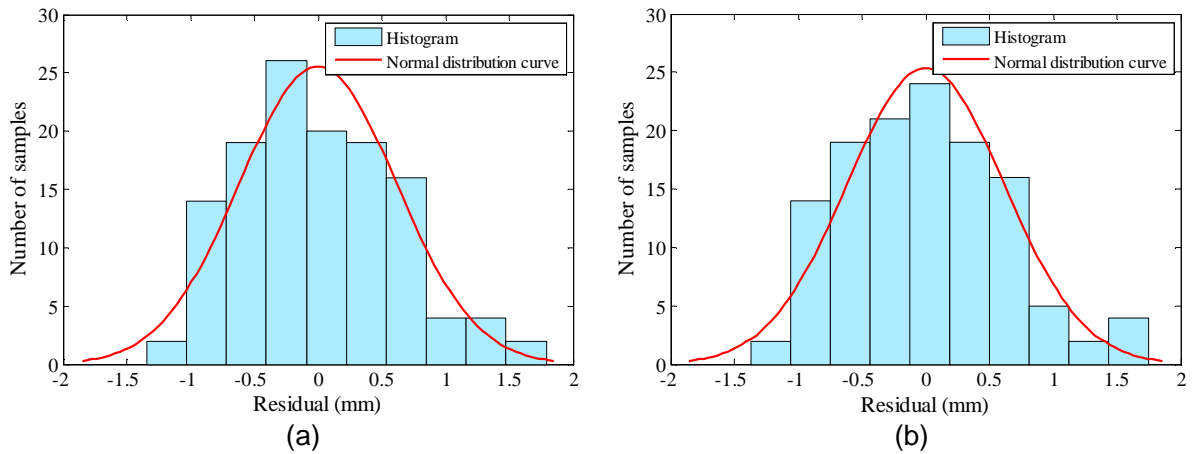


Fig. 11 Histogram of residuals for regression models: (a) MLR; (b) SR.

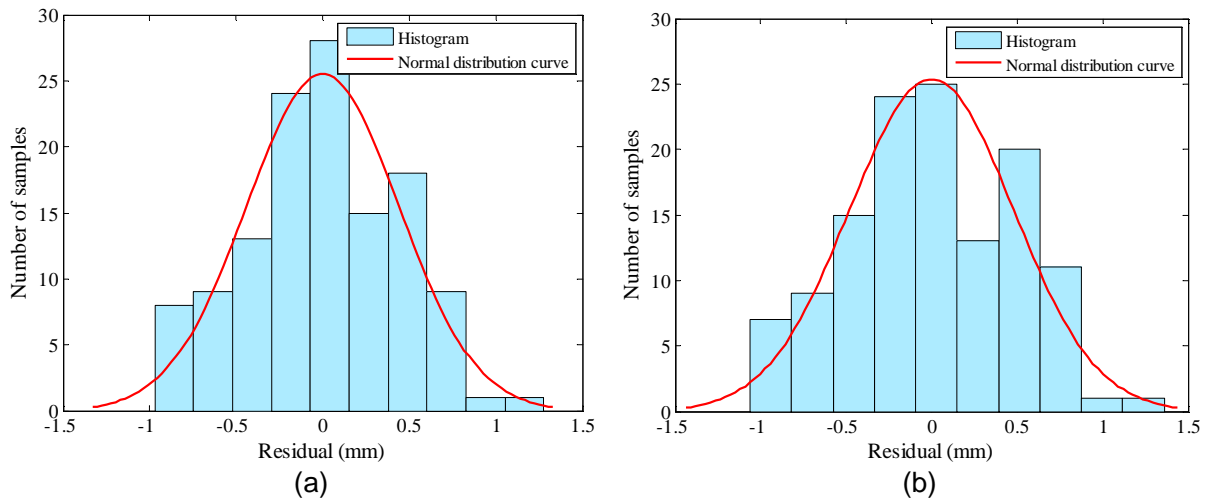


Fig. 12 Histogram of residuals for LSSVM models: (a) LSSVM-1; (b) LSSVM-2.

## 5. CONCLUSIONS

Because of the complexity of dam health monitoring modeling, it is difficult to establish an accurate prediction model. A methodology based on LSSVM is proposed to solve this problem. Results show that the proposed approach performs much better than the MLR and SR models in terms of accuracy. Meanwhile, it seems no need to delete variables which should be deleted in a SR procedure, since the performance are not different much for the two LSSVM models. The proposed intelligent procedure is shown to be an efficient scheme with respect to both computational effort and accuracy.

For the future work, the proposed methodology for DHM can also be extended to deal with health monitoring problems of other structures.

## Acknowledgement

This research was supported by the Fundamental Research Funds for the Central Universities under Grant no. DUT15LK11, the Open Research Fund of the State Key Laboratory of Structural Analysis for Industrial Equipment, Dalian University of Technology, under Grant no. GZ15207, National Natural Science Foundation of China under Grant no. 51109028, and the State Scholarship Fund of China to pursue study in the USA as a visiting scholar under Grant no. 201208210208.

## REFERENCES

- Ahmadi-Nedushan B. Multivariate statistical analysis of monitoring data for dams. Dissertation. McGill University, Montreal 2002.
- Bukenya P, Moyo P, Beushausen H. (2014), "Health monitoring of dams: a literature review," *Journal of Civil Structural Health Monitoring*, 4 (4), 235-244.
- Kang, F., Han, S., Salgado, R., & Li, J. (2015), "System probabilistic stability analysis of soil slopes using Gaussian process regression with Latin hypercube sampling," *Computers and Geotechnics*, 63, 13-25.
- Kang, F., & Li, J. (2015), "Artificial bee colony algorithm optimized support vector regression for system reliability analysis of slopes," *Journal of Computing in Civil Engineering*, 30(3), 04015040.
- Kang, F., Li, J. S., & Li, J. J. (2016a), "System reliability analysis of slopes using least squares support vector machines with particle swarm optimization," *Neurocomputing*, <http://dx.doi.org/10.1016/j.neucom.2015.11.122>.
- Kang, F., Li, J. S., Wang, Y., & Li, J. (2016b), "Extreme learning machine-based surrogate model for analyzing system reliability of soil slopes," *European Journal of Environmental and Civil Engineering*, DOI: 10.1080/19648189.2016.1169225.
- Kang, F., Xu, Q., & Li, J. (2016c), "Slope reliability analysis using surrogate models via new support vector machines with swarm intelligence," *Applied Mathematical Modelling*, 40(11), 6105-6120.
- Kao, C. Y., & Loh, C. H. (2013), "Monitoring of long - term static deformation data of Fei - Tsui arch dam using artificial neural network - based approaches," *Structural Control and Health Monitoring*, 20(3), 282-303.
- Léger, P. and Leclerc, M. (2007), "Hydrostatic, temperature, time-displacement model for concrete dams," *Journal of engineering mechanics*, 133(3), 267-277.
- Mata, J. (2011), "Interpretation of concrete dam behaviour with artificial neural network and multiple linear regression models," *Engineering Structures*, 33(3), 903-910.
- Mata, J., de Castro, A. T., & da Costa, J. S. (2013), "Time-frequency analysis for concrete dam safety control: Correlation between the daily variation of structural response and air temperature," *Engineering Structures*, 48, 658-665.
- Mata, J., Tavares de Castro, A., & Sá da Costa, J. (2014), "Constructing statistical models for arch dam deformation," *Structural Control and Health Monitoring*, 21(3), 423-437.
- Ranković, V., Grujović, N., Divac, D., et al. (2012), "Modelling of dam behaviour based on neuro-fuzzy identification," *Engineering Structures*, 35, 107-113.
- Ranković, V., Grujović, N., Divac, D., & Milivojević, N. (2014), "Development of support vector regression identification model for prediction of dam structural behaviour," *Structural Safety*, 48, 33-39.
- Salazar F, Morán R, Toledo M Á, et al. (2015), "Data-based models for the prediction of dam behaviour: a review and some methodological considerations," *Archives of Computational Methods in Engineering*, DOI: 10.1007/s11831-015-9157-9.
- Salazar, F., Toledo, M. A., Oñate, E., & Morán, R. (2015), "An empirical comparison of machine learning techniques for dam behaviour modeling," *Structural Safety*, 56, 9-17.

- Su, H., Chen, Z., & Wen, Z. (2016), "Performance improvement method of support vector machine-based model monitoring dam safety," *Structural Control and Health Monitoring*, 23(2), 252-266.
- Suykens JAK, Van Gestel T, De Brabanter J, De Moor B, Vandewalle J. *Least Squares Support Vector Machines*, World Scientific, Singapore, 2002.
- Tatin, M., Briffaut, M., Dufour, F., Simon, A. and Fabre, J. P. (2015), "Thermal displacements of concrete dams: accounting for water temperature in statistical models," *Engineering Structures*, 91, 26-39.
- Vapnik V. *The nature of statistical learning theory*. Springer, 2000.
- Xi, G. Y., Yue, J. P., Zhou, B. X., & Tang, P. (2011), "Application of an artificial immune algorithm on a statistical model of dam displacement," *Computers & Mathematics with Applications*, 62(10), 3980-3986.
- Xu, C., Yue, D., & Deng, C. (2012), "Hybrid GA/SIMPLS as alternative regression model in dam deformation analysis," *Engineering Applications of Artificial Intelligence*, 25(3), 468-475.