Computers and Concrete, Vol. 9, No. 3 (2012) 215-226

DOI: http://dx.doi.org/10.12989/cac.2012.9.3.215

Utilization of support vector machine for prediction of fracture parameters of concrete

Pijush Samui*[1] and Dookie Kim[2]

[1]*Centre for Disaster Mitigation and Management. VIT University, Vellore-632014, Tamilnadu, India*
[2]*Department of Civil Engineering, Kunsan National University, Kunsan, Jeonbuk, South Korea*

(Received July 19, 2010, Revised April 12, 2011, Accepted May 26, 2011)

Abstract. This article employs Support Vector Machine (SVM) for determination of fracture parameters critical stress intensity factor (K_{Ic}^{s}) and the critical crack tip opening displacement ($CTOD_c$) of concrete. SVM that is firmly based on the theory of statistical learning theory, uses regression technique by introducing ε-insensitive loss function has been adopted. The results are compared with a widely used Artificial Neural Network (ANN) model. Equations have been also developed for prediction of K_{Ic}^{s} and $CTOD_c$. A sensitivity analysis has been also performed to investigate the importance of the input parameters. The results of this study show that the developed SVM is a robust model for determination of K_{Ic}^{s} and $CTOD_c$ of concrete.

Keywords: concrete; fracture mechanics; support vector machine; sensitivity analysis; artificial neural network; two-parameter model.

1. Introduction

Engineers use different non-linear fracture mechanics approach for modeling concrete such as fictitious crack model (Hillerborg *et al.* 1976), the crack band model (Bazant and Oh 1983) the Two-Parameter Model (TPM) (Jenq and Shah 1985), the effective crack model (Nallathambi and Karihaloo 1986), the size effect model (Bazant and Kazemi 1990) and the peak load method (Tang *et al.* 1996). The determination of fracture parameters of concrete structure is a difficult task (Ince 2004). Several methods are being used for determination of fracture parameters of concrete structure such as experimental techniques (Hillerborg *et al.* 1976, Jenq and Shah 1985, Bazant and Kazemi 1990, Tang *et al.* 1996), regression models (Bazant and Oh 1983, John and Shah 1989, Hilsdorf and Brameshuber 1991). However, the above mentioned methods have some limitations (Ince 2004). Recently, Artificial Neural Network (ANN) has been successfully used for prediction of fracture parameters of concrete structure. However, ANN model has some limitations. The limitations are listed below:

- Unlike other statistical models, ANN does not provide information about the relative importance of the various parameters (Park and Rilett 1999).
- The knowledge acquired during the training of the model is stored in an implicit manner and hence it is very difficult to come up with reasonable interpretation of the overall structure of the network (Kecman 2001).
- In addition, ANN has some inherent drawbacks such as slow convergence speed, less generalizing

* Corresponding author, Professor, E-mail: pijush.phd@gmail.com

performance, arriving at local minimum and over-fitting problems.

This paper adopts Support Vector Machine (SVM) for determination of fracture parameters of concrete following the TPM approach which is based on the critical stress intensity factor (K^s_{Ic}) and the critical crack tip opening displacement ($CTOD_c$) as fracture parameters. It provides a new, efficient novel approach to improve the generalization performance and can attain a global minimum. In general, SVM has been used for pattern recognition problems. Recently it has been used to solve non-linear regression estimation and time series prediction by introducing ε-insensitive loss function (Mukherjee *et al*. 1997, Muller *et al.* 1997, Vapnik 1995, Vapnik *et al*. 1997). The SVM implements the structural risk minimization principle (SRMP), which has been shown to be superior to the more traditional Empirical Risk Minimization Principle (ERMP) employed by many of the other modelling techniques (Osuna *et al.* 1997, Gunn 1998). SRMP minimizes an upper bound of the generalization error whereas, ERMP minimizes the training error. In this way, it produces the better generalization than traditional techniques. The paper has the following aims:

- To investigate the capability of SVM for prediction of K^s_{Ic} and $CTOD_c$
- To develop equations for determination of K^s_{Ic} and $CTOD_c$ based on the developed SVM model
- To make a comparative study between ANN, experimental method and the developed SVM model
- To do sensitivity analysis for determination of effect of the each input parameter

2. The general information of SVM

Support Vector Machine (SVM) has originated from the concept of statistical learning theory pioneered by Boser *et al.* (1992). This study uses the SVM as a regression technique by introducing a ε-insensitive loss function. In this section, a brief introduction on how to construct SVM for regression problem is presented. More details can be found in many publications (Boser *et al.* 1992, Cortes and Vapnik 1995, Gualtieri *et al.* 1999, Vapnik 1998, Samui 2008, Samui *et al.* 2008). There are three distinct characteristics when SVM is used to estimate the regression function. First of all, SVM estimates the regression using a set of linear functions that are defined in a high dimensional space. Secondly, SVM carries out the regression estimation by risk minimization where the risk is measured using Vapnik's ε-insensitive loss function. Thirdly, SVM uses a risk function consisting of the empirical error and a regularization term which is derived from the structural risk minimization (SRM) principle. Considering a set of training data, $\{(x_1,y_1),\ldots,(x_l,y_l)\}, x \in R^n,\ y \in r$. Where x is the input, y is the output, R^N is the N-dimensional vector space and r is the one-dimensional vector space. The four input variables used for the SVM model in this study are water-cement ratio (w/c), maximum aggregate size ($d_{\max}$), and compressive strength of concrete (f_c). The output of the SVM model is K^s_{Ic} or $CTOD_c$. So, in this study, $x = [w/c,\ d_{\max},\ f_c']$ and $y = [K^s_{Ic}\ \&\ CTOD_c]$.

The ε-insensitive loss function can be described in the following way

$$L_\varepsilon(y) = 0 \quad \text{for} \quad |f(x)-y| < \varepsilon \quad \text{otherwise} \quad L_\varepsilon(y) = |f(x)-y| - \varepsilon \tag{1}$$

This defines an ε tube (Fig. 1) so that if the predicted value is within the tube the loss is zero, while if the predicted point is outside the tube, the loss is equal to the absolute value of the deviation minus ε. The main aim in SVM is to find a function $f(x)$ that gives a deviation of ε from the actual output and at the same time is as flat as possible. Let us assume a linear function

$$f(x) = (w \cdot x) + b \quad w \in R^n,\ b \in r \tag{2}$$

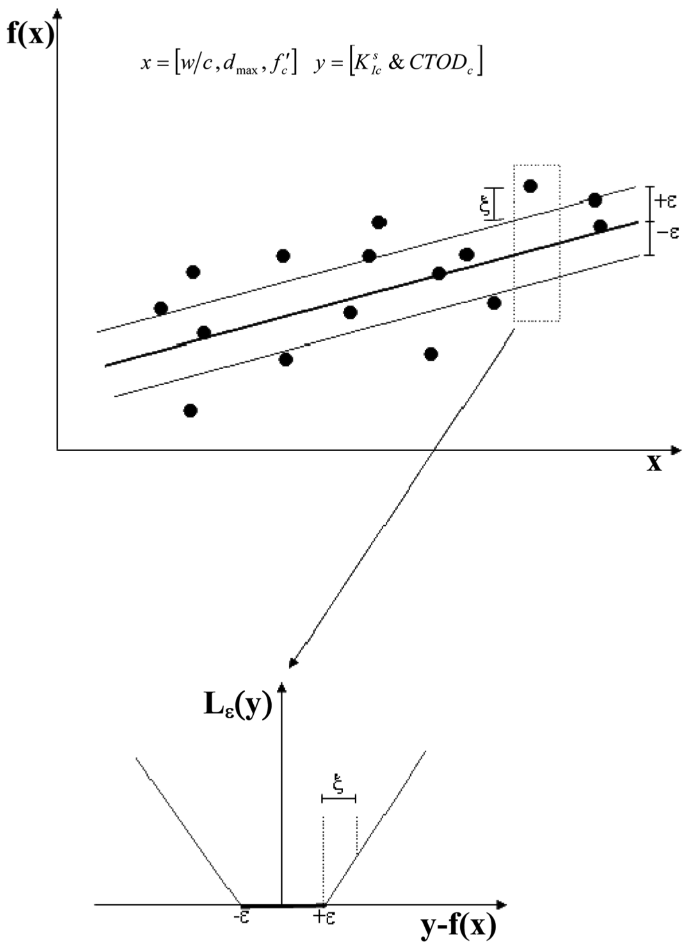

Fig. 1 Prespecified Accuracy ε and Slack Variable ξ in support vector regression (Scholkopf 1997)

Where, w = an adjustable weight vector and b = the scalar threshold.

Flatness in the case of (3) means that one seeks a small w. One way of obtaining this is by minimizing the Euclidean norm $\|w\|^2$. This is equivalent to the following convex optimization problem

Minimize: $\frac{1}{2}\|w\|^2$

Subjected to :$y_i - ((w \cdot x_i) + b) \le \varepsilon$, i = 1, 2, ..., 1

$$((w \cdot x_i) + b) - y_i \le \varepsilon,\ \text{i} = 1, 2, ..., 1 \qquad (3)$$

The above convex optimization problem is feasible. Sometimes, however, this may not be the case, or I also may want to allow for some errors. Analogously to the "soft margin" loss function (Bennett and Mangasarian 1992) which was used in SVM by Cortes and Vapnik (1995). As shown in the Fig. 1, the parameters ξ_i, ξ_i^* are slack variables that determine the degree to which samples with error more than ε be penalized. In other words, any error smaller than ε does not require ξ_i, ξ_i^* and hence does not enter the objective function because these data points have a value of zero for the loss function. The slack variables (ξ_i, ξ_i^*) has been introduced to avoid infeasible constraints of the optimization problem (3).

$$\text{Minimize: } \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{1}(\xi_i + \xi_i^*)$$

$$\text{Subjected to: } y_i - ((w\cdot x_i) + b) \le \varepsilon + \xi_i,\ \text{i} = 1, 2, ..., 1$$

$$((w\cdot x_i) + b) - y_i \le \varepsilon + \xi_i^*,\ \text{i} = 1, 2, ..., 1$$

$$\xi_i \ge 0 \text{ and } \xi_i^* \ge 0,\ \text{i} = 1, 2, ..., 1 \tag{4}$$

The constant $0 < \text{C} < \infty$ determines the trade-off between the flatness of f and the amount up to which deviations larger than ε are tolerated (Smola and Scholkopf 2004). This optimization problem (4) is solved by Lagrangian Multipliers (Vapnik 1998), and its solution is given by

$$f(x) = \sum_{i=1}^{N}(\alpha_i - \alpha_i^*)(x_i\cdot x) + b \tag{5}$$

Where $b = -\left(\frac{1}{2}\right)w.[x_r + x_s]$, α_i, α_i^* are the Lagrangian Multipliers and N is the number of data. An

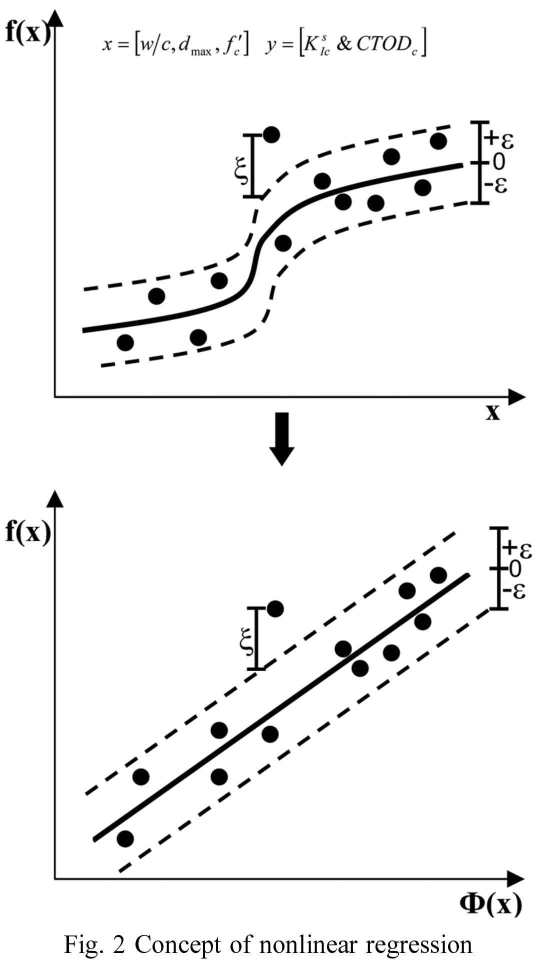

Fig. 2 Concept of nonlinear regression

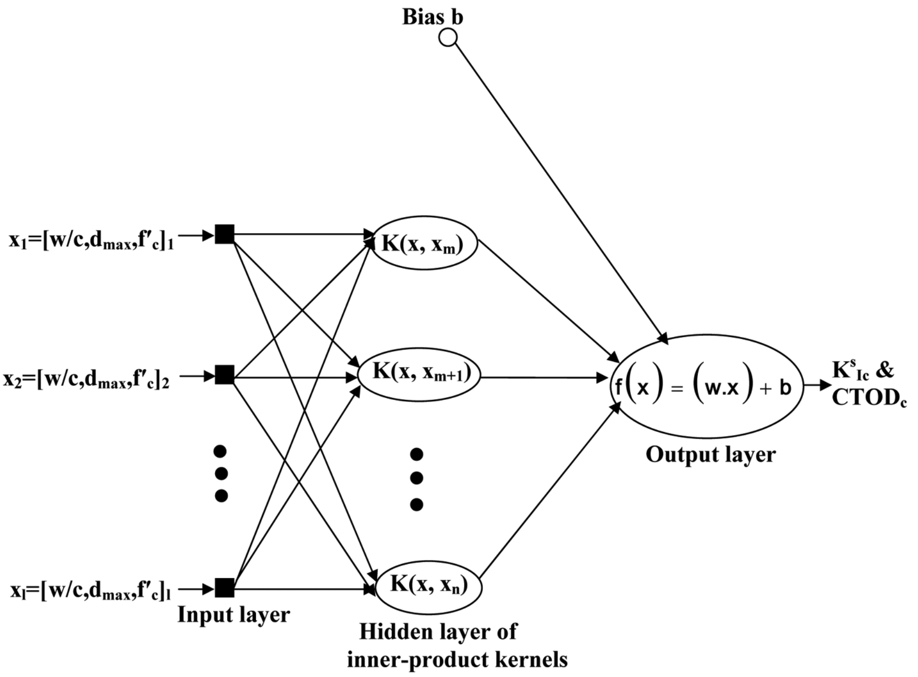

Fig. 3 SVM architecture for K_{Ic}^{s} and $CTOD_c$ prediction

important aspect is that some Lagrange multipliers (α_i, α_i^*) will be zero, implying that these training objects are considered to be irrelevant for the final solution (sparseness). The training objects with nonzero Lagrange multipliers are called support vectors.

When linear regression is not appropriate, then input data has to be mapped into a high dimensional feature space through some nonlinear mapping (Boser *et al.* 1992) (see Fig. 2). The two steps that are involved are first to make a fixed nonlinear mapping of the data onto the feature space and then carry out a linear regression in the high dimensional space. The input data is mapped onto the feature space by a map Φ(see Fig. 2). The dot product given by $\Phi(x_i)\cdot\Phi(x_j)$ is computed as a linear combination of the training points. The concept of kernel function $[K(x_i,x_j) = \Phi(x_i)\cdot\Phi(x_j)]$ has been introduced to reduce the computational demand (Cristianini and Shwae-Taylor 2000, Cortes and Vapnik 1995). So, Eq. (5) becomes written as

$$f(x) = \sum_{i=1}^{N}(\alpha_i - \alpha_i^*)K(x_i \cdot x_j) + b \tag{6}$$

Some common kernels have been used such as polynomial (homogeneous), polynomial (nonhomogeneous), radial basis function, Gaussian function, sigmoid etc for non-linear cases. Fig. 3 shows a typical architecture of SVM for K_{Ic}^{s} and $CTOD_c$.

3. Support vector machine-based analysis of fracture parameters of concrete

This study employs above methodology for prediction of K_{Ic}^{s} and CTOD of concrete. This study uses the database collected by Ince (2004). The data has been further divided into two sub-sets; a

Table 1 Training dataset

w/c	d_{max}(mm)	f_c' (MPa)	K_{Ic}^s (MPa $\sqrt{m}$)	CTOD$_c$(mm)
0.65	19	25.2	0.976	0.017
0.65	4.8	27.2	0.707	0.0093
0.45	4.8	39.4	0.958	0.0097
0.45	0	27.7	0.618	0.0069
0.4	3	33.8	0.883	0.0075
0.22	8	110	2.13	0.0338
0.44	6	60.7	1.141	0.0145
0.52	12.7	45.5	1.475	0.022
0.52	12.7	43.4	1.53	0.0169
0.52	12.7	30.7	1.036	0.0115
0.53	25.4	15.4	0.857	0.0169
0.64	20	39	1.269	0.026
0.5	20	49.4	1.381	0.026
0.36	20	65.7	1.509	0.024
0.2	20	78.2	1.847	0.026
0.4	9	34.5	0.72	0.0076
0.4	9	39.3	0.826	0.0132
0.4	9	55.3	1.483	0.0143
0.29	9	45	0.978	0.0091
0.29	9	57.3	1.168	0.0107
0.29	9	87.2	1.491	0.0116
0.5	8	38.8	0.867	0.0055

training dataset, to construct the model, and a testing dataset to estimate the model performance. The same training data (see Table 1), and testing data (see Table 2) has been used in this study as used by Ince (2004). The data is normalized between 0 and 1. Radial basis function, polynomial and spline have been used as kernel function. In training process, a simple trial-and-error approach has been used to select the design value of C, ε and width (σ) of radial basis function.

In this study, a sensitivity analysis has been done to extract the cause and effect relationship between the inputs and outputs of the SVM model. The basic idea is that each input of the model is offset slightly and the corresponding change in the output is reported. The procedure has been taken from the work of Liong *et al.* (2000). According to Liong *et al.* (2000), the sensitivity(S) of each input parameter has been calculated by the following formula

$$S(\%) = \frac{1}{N}\sum_{j=1}^{N}\left(\frac{\%\text{ change in ouput}}{\%\text{ change in input}}\right)_j \times 100 \tag{7}$$

Where N is the number of data points. In this study, $N = 22$. The analysis has been carried out on the trained model by varying each of input parameter, one at a time, at a constant rate of 20%. In the present study, training, testing and sensitivity analysis of SVM has been carried out by using MATLAB.

Table 2 Testing dataset

w/c	d_{max}(mm)	f_c' (MPa)	K_{Ic}^s (MPa $\sqrt{m}$)	$CTOD_c$(mm)
0.25	4.8	54.8	1.059	0.01
0.77	20	26.8	0.923	0.0242
0.4	9	51.8	1.208	0.0122
0.29	9	58.7	1.175	0.0123
0.29	9	66.3	1.346	0.0094
0.53	3.15	59.1	0.779	0.0018
0.54	32	31	1.018	–
0.54	2	35	0.645	–
0.5	19	54.4	0.755	0.0102
0.5	19	55.8	0.56	0.0067
0.5	19	53.1	1.124	0.0222
0.57	25	41.3	0.65	–
0.62	25	35.9	0.88	–
0.27	25	52.3	0.97	–
0.47	20	51	0.76	0.0054
0.48	12.5	55	1	0.0062
0.5	9	30.5	0.793	0.0078
0.5	9	30	0.838	0.0133

4. Results and discussion

This study employs coefficient of correlation(R) to assess the performance of SVM model. Table 3 shows the performance of the different kernels. It is observed from Table 3 that the performance of radial basis function is best. For this reason, we have presented only the result of radial basis function. For prediction of K_{Ic}^s, the design value of C, ε and σ is 50, 0.01 and 0.7 respectively. Fig. 4 depicts the performance of SVM model for training dataset. The value of R ($R = 0.956$) is close to one. For good model, the value of R should be close to one. So, the developed SVM model has successfully captured input and output relationship for training dataset. The developed SVM model has been also used to determine the performance of testing dataset. Fig. 5 illustrates the performance of testing dataset. It also shows that the value of R($R = 0.935$) is close to one. The value of R is close to one for both training as well as testing dataset. Therefore, the developed SVM model has the

Table 3 Performance of the different kernels

Kernel function	K_{Ic}^s		$CTOD_c$	
	Training performance (R)	Testing performance (R)	Training performance (R)	Testing performance (R)
Radial basis function	0.956	0.935	0.940	0.938
Polynomial	0.874	0.789	0.769	0.721
Spline	0.628	0.601	0.714	0.702

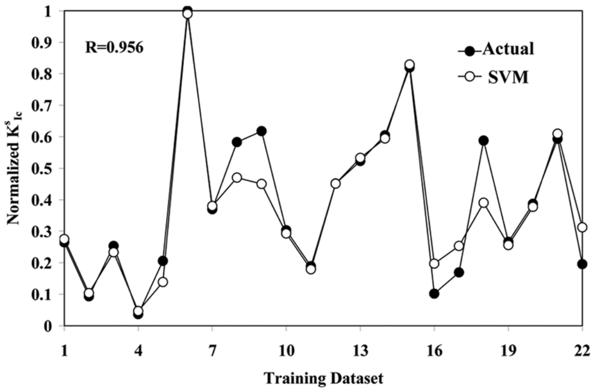

Fig. 4 Performance of training dataset for K^s_{Ic} prediction

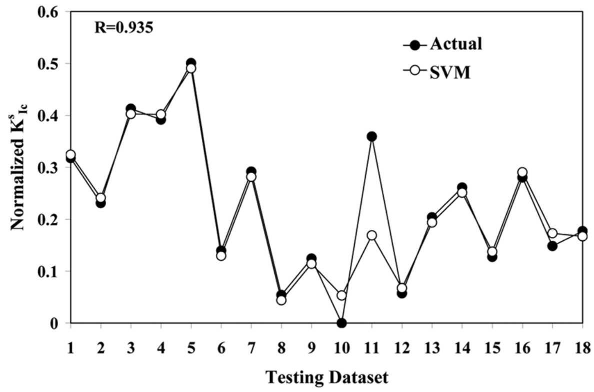

Fig. 5 Performance of testing dataset for K^s_{Ic} prediction

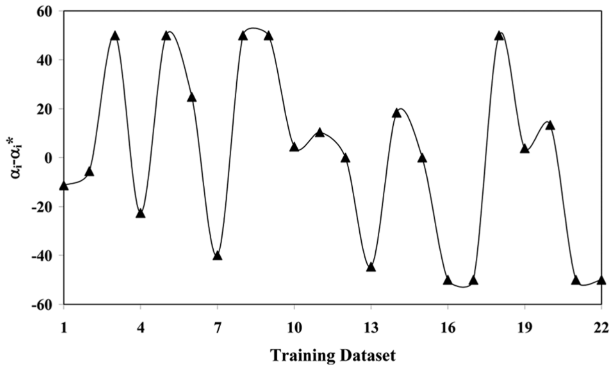

Fig. 6 Values of $(\alpha_i - \alpha_i^*)$ for K^s_{Ic}

ability to predict K^s_{Ic}. The following equation (by putting

$$K(x_i, x) = \exp\left\{-\frac{(x_i - x)(x_i - x)^T}{2\sigma^2}\right\},$$

$N = 22$, $\sigma = 0.7$ and $b = 0$ in Eq. (6)) has been developed for the prediction of K^s_{IC} based on the developed SVM model.

$$K^s_{Ic} = \sum_{i=1}^{22} (\alpha_i - \alpha_i^*) \exp\left\{-\frac{(x_i - x)(x_i - x)^T}{0.98}\right\} \tag{8}$$

Fig. 6 shows the values of $(\alpha_i - \alpha_i^*)$ for K^s_{Ic}.

For $CTOD_c$, the design value of C, ε and σ is 60, 0.009 and 0.8 respectively. The performance of training dataset has been determined by using the design value of C, ε and σ and it has been shown in Fig. 7. Fig. 7 also shows that the value of R is close to one. The performance of testing dataset has been also determined same way as for training dataset. Fig. 8 illustrates the performance of testing dataset. Form these results, it is confirmed that the developed SVM has the ability to predict $CTOD_c$. The developed SVM model gives the following equation (by putting

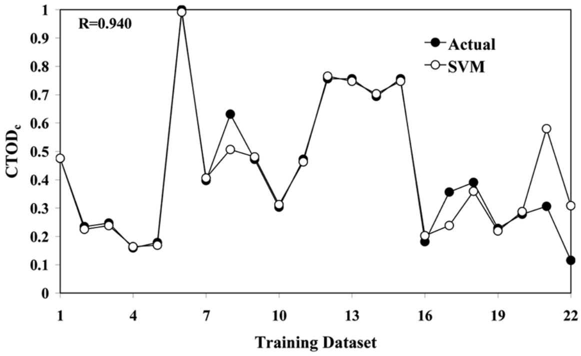

Fig. 7 Performance of training dataset for $CTOD_c$ prediction

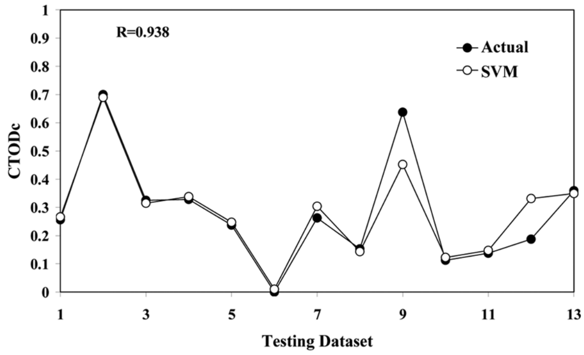

Fig. 8 Performance of testing dataset for $CTOD_c$ prediction

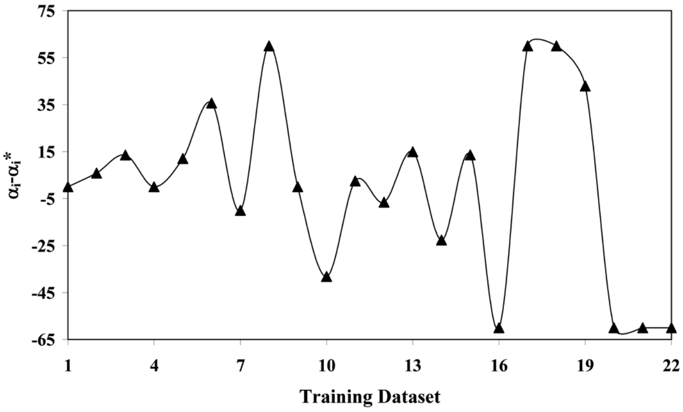

Fig. 9 Values of $(\alpha_i - \alpha_i^*)$ for $CTOD_c$

$$K(x_i, x) = \exp\left\{-\frac{(x_i - x)(x_i - x)^T}{2\sigma^2}\right\},$$

$N = 22$, $\sigma = 0.8$ and $b = 0$ in Eq. (6)) for prediction of $CTOD_c$.

$$CTOD_c = \sum_{i=1}^{22} (\alpha_i - \alpha_i^*) \exp\left\{-\frac{(x_i - x)(x_i - x)^T}{1.28}\right\} \tag{9}$$

Fig. 9 shows the values of $(\alpha_i - \alpha_i^*)$ for $CTOD_c$.

The developed SVM uses 20 and 19 training data as support vector for prediction of K_{Ic}^s and $CTOD_c$ respectively. These support vectors have only used for final prediction. So, there is real advantage gained in terms of sparsity. Sparseness is desirable in SVM for several reasons, namely (Figueiredo 2003):

- Sparseness leads to a structural simplification of the estimated function.
- Obtaining a sparse estimate corresponds to performing feature/variable selection.
- The generalization ability improves with the degree of sparseness.

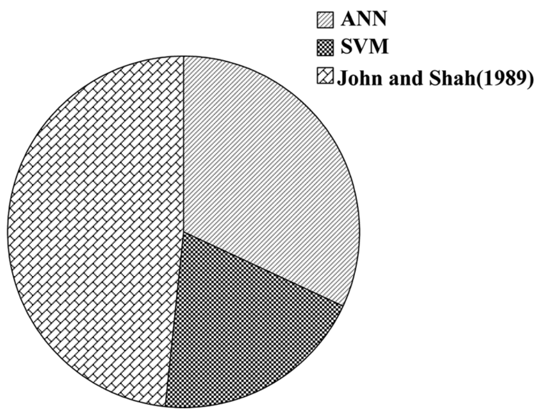

Fig. 10 Comparison between different methods for prediction of K_{Ic}^{s} in terms of RMSE

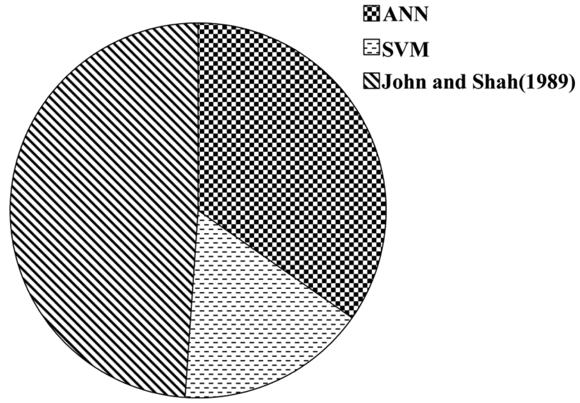

Fig. 11 Comparison between different methods for prediction of K_{Ic}^{s} in terms of MAE

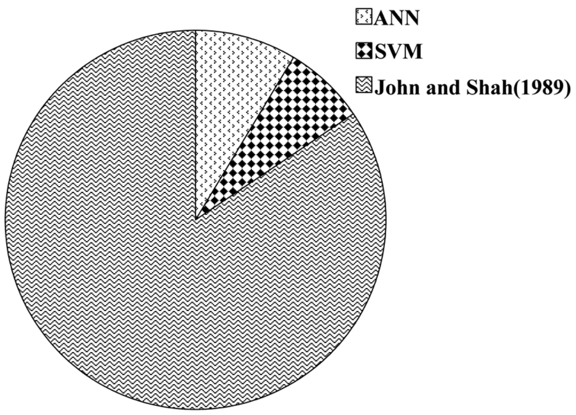

Fig. 12 Comparison between different methods for prediction of $CTOD_c$ in terms of RMSE

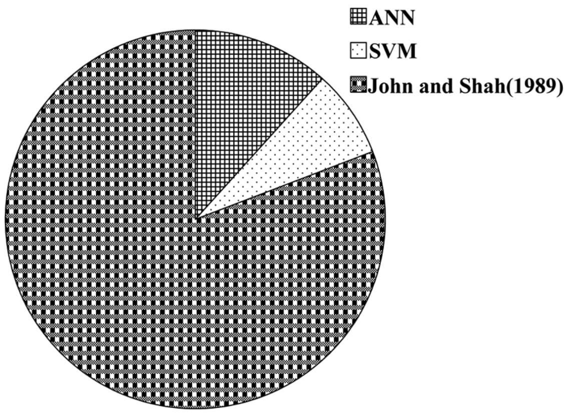

Fig. 13 Comparison between different methods for prediction of $CTOD_c$ in terms of RMAE

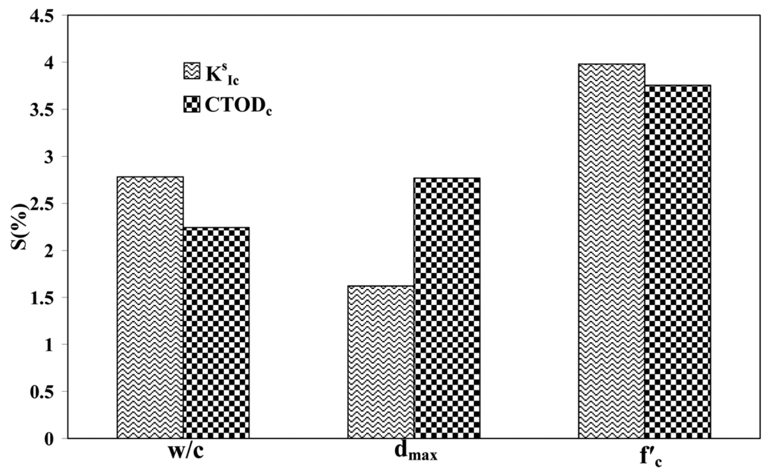

Fig. 14 Sensitivity analysis of input parameters

Sparseness means that a significant number of the weights are zero (or effectively zero), which has the consequence of producing compact, computationally efficient models, which in addition are simple and therefore produce smooth functions.

Figs. 10, 11, 12 and 13 present a comparative study between the developed SVM and other methods (Ince 2004, John and Shah 1989) for determination of K_{Ic}^{s} and $CTOD_c$. The comparisons have been carried out in terms of Root-Mean-Square-Error (RMSE) and Mean-Absolute-Error (MAE). Figs. 10, 11, 12 and 13 confirm that the developed SVM is better than the available methods for prediction of K_{Ic}^{s} and $CTOD_c$. SVM employs only three parameters (C, ε and σ). Whereas, ANN uses number of hidden layers, number of hidden nodes, learning rate, momentum term, number of training epochs, transfer functions and weight initialization methods.

Fig. 14 presents the results of sensitivity analysis. For K_{Ic}^{s}, f_c' has the most significant effect on the predicted K_{Ic}^{s} followed by *w/c* and $d_{\max}$. Sensitivity analysis also shows that f_c' has the most significant effect on the predicted $CTOD_c$ followed by $d_{\max}$ and *w/c.*

5. Conclusions

This study has successfully applied SVM for prediction of K_{Ic}^{s} and $CTOD_c$ of concrete. The performance of the developed SVM is better than the available methods. SVM training consists of solving a – uniquely solvable – quadratic optimization problem and always finds a global minimum. User can use the developed equations for determination of K_{Ic}^{s} and $CTOD_c$ of concrete. The sensitivity analysis indicates that f_c' is the most important factor affecting facture parameters of concrete. The proposed SVM is not a substitute but may be a viable alternative for prediction of the facture parameters of concrete.

References

Bazant, Z.P. and Kazemi, M.T. (1990), "Determination of fracture energy, process zone length and brittleness number from size effect with application to rock and concrete", *Int. J. Fracture*, **44**(2), 111-31.

Bazant, Z.P. and Oh, B.H. (1983), "Crack band theory for fracture concrete", *Mater. Struct.*, **16**(93), 155-177.

Bennett, K.P. and Mangasarian, O.L. (1992), "Robust linear programming discrimination of two linearly inseparable sets", *Optim. Method. Softw.*, **1**(1), 23-34.

Boser, B.E., Guyon, I.M. and Vapnik, V.N. (1992), "A training algorithm for optimal margin classifiers", *In D. Haussler, editor, 5th Annual ACM Workshop on COLT, Pittsburgh, PA, ACM Press*, 144-152.

Cortes, C. and Vapnik, V.N. (1995), "*Support vector networks machine learning*", **20**(3), 273-297.

Cristianini, N. and Shawe Taylor, J. (2000), "*An introduction to support vector machine*", Cambridge University press, London.

"*Fuzzy Logic Models*", The MIT press, Cambridge, Massachusetts, London, England.

Gualtieri, J.A., Chettri, S.R., Cromp, R..F. and Johnson, L.F. (1999), "Support vector machine classifiers as applied to AVIRIS data", *In the Summaries of the Eighth JPL Airbrone Earth Science Workshop.*

Hillerborg, A., Modeer, M. and Petersson, P.E. (1976), "Analysis of crack formation and crack growth in concrete by means of fracture mechanics and finite elements", *Cement Concrete,* **6**(6), 773-781.

Hilsdorf, H.H. and Brameshuber, W. (1991), "Code-type formulation of fracture mechanics concepts for concrete", *Int. J. Fracture*, **51**(1), 61-72.

Ince, R. (2004), "Prediction of fracture parameters of concrete by Artificial Neural Networks", *Eng. Fract. Mech.,* **71**(15), 2143-2159.

Jenq, Y.S. and Shah, S.P. (1985), "Two-parameter fracture model for concrete", *J. Eng. Mech.-ASCE*, **111**(10), 1227-1241.

Kecman, V. (2001), *Learning and Soft Computing: Support Vector Machines, Neural Networks, and Fuzzy logic Models, The MIT Press, Cambridge, MA, 2001.*

Liong, S.Y., Lim, W.H. and Paudyal, G.N. (2000), "River stage forecasting in Bangladesh: neural network approach", *J. Comput. Civil Eng.*, **14**(1), 1-8.

Mukherjee, S., Osuna, E. and Girosi, F. (1997), "Nonlinear prediction of chaotic time series using support vector machine", *Proc., IEEE Workshop on Neural Networks for Signal Processing 7*, Institute of Electrical and Electronics Engineers, New York, 511-519.

Muller, K.R., Smola, A., Ratsch, G., Scholkopf, B., Kohlmorgen, J. and Vapnik, V. (1997), "Predicting time series with support vector machines", *Proc., Int. Conf. on Artificial Neural Networks*, Springer-Verlag, Berlin.

Nallathambi, P. and Karihaloo, B.L. (1986), "Determination of the specimen size independent fracture toughness of plain concrete", *Mag. Concrete Res.*, **38**(135), 67-76.

Park, D. and Rilett, L.R. (1999), "Forecasting freeway link travel times with a multi-layer feed forward neural network", *Comput.-Aided Civ. Inf.*, **14**(5), 358-367.

Samui, P. (2008), "Support vector machine applied to settlement of shallow foundations on cohesionless soils", *Comput. Goetech.*, **35**(3), 419-427.

Samui, P., Kurup, P.U. and Sitharam, T.G. (2008), "OCR prediction using support vector machine based on piezocone data", *J. Geotech. Geoenviron*., **134**(6), 894-898.

Tang, T., Ouyang, C. and Shah, S.P. (1996), "A simple method for determining material fracture parameters from peak loads", *ACI Mater. J.*, **93**(2), 147-157.

Vapnik, V.N. (1995), *The nature of statistical learning theory*, Springer, New York.

CC

Nomenclature

SVM	= Support Vector Machine
d_{max}	= Maximum aggregate size (mm)
f_c'	= Compressive strength of concrete (MPa)
CTOD_c	= Critical crack tip opening displacement (mm)
K_{Ic}^s	= Critical stress intensity factor based on two-parameter model (MPa $\sqrt{m}$)
ε	= Error insensitive zone
σ	= width of the radial basis function
R	= Coefficient of correlation
R^n	= N-dimensional vector space
r	= one dimensional vector space
ξ_i, ξ_i^*	= slack variable
α_i, α_i^*	= Lagrange multipliers
C	= Capacity factor
S	= Sensitivity
w/c	= water-cement ratio