www.cafetinnova.org

ISSN 0974-5904, Volume 10, No. 03

DOI:10.21276/ijee.2017.10.0322

International Journal of Earth Sciences and Engineering

June 2017, P.P. 633-638

Application of Support Vector Machine Technique for Damage Level Prediction of Tandem Breakwater

GEETHA KUNTOJI[1], SUBBA RAO[1], MANU[1] AND S MANDAL[2]

[1]*Department of Applied Mechanics & Hydraulics, National Institute of Technology Karnataka (NITK), Surathkal, Srinivasnagar Post, Karnataka, INDIA*

[2]*Department of Civil Engineering, PES University, Bengaluru, INDIA*

Email: *geeta.kuntoji@gmail.com, surakrec@gmail.com, manunitk77@gmail.com, smandal12341@rediffmail.com*

Abstract: *For decades, breakwaters played a vital role in the development of the port and which in turn assist in improving the economy of the country. Due to the development of industries and urbanization the coastal region is facing threat towards its normal and safe functioning. It challenges for the coastal engineers protect the environment and also support the urbanization in a protective way. Therefore, a need of protective structure along the coast arises. The tandem breakwater is one of the innovative types of structure, which consists of a conventional breakwater along with the submerged reef. It has been proved that submerged reef acts as a protective structure, when placed in-front of the conventional breakwater. The final geometry and the layout of tandem breakwater have been achieved through physical model studies. Support Vector Machine (SVM) which accounts for structural risk minimization compared to neural networks is used to model various problems of real time scenarios where mathematical modeling is difficult. By assigning appropriate weights, biases and e-insensitive loss function of the problem, SVM proves to be robust in addressing non-linear and non-stationary problems. In the present study SVM is applied to predict the damage level of the conventional breakwater of tandem breakwater. The experimental data available in the department of Applied Mechanics & Hydraulics are used for the analysis. Finally, comparison of the predicted damage level is made with the observed data of the experimental work using the statistical measures such as Root Mean Square Error (RMSE), Correlation Coefficient (CC), Scatter Index (SI) and Nash Sutcliffe Efficiency (NSE). It is observed that SVM technique using Radial basis kernel function (RBF) performed better with 0.9357 CC, 0.0765 RMSE, 0.5293 SI and 0.8327 NSE.*

Keywords: *Support Vector Machine, Conventional rubble mound breakwater of Tandem Breakwater, Damage level and Radial Basis function, Kernel functions.*

1. Introduction

1.1. Breakwater

Due to increased activity along the coasts there is a need to protect the coast, which is vulnerable to wave attacks. One of the protective measures that can be adopted is breakwater, which can also be used as a protective structure for ports and harbors. The main purpose of the breakwater is to dissipate wave energy before it reaches the coast. The effect of wave on the main breakwater can be decreased by putting a submerged reef in front of the main breakwater. This combination of submerged reef with main breakwater is called as tandem breakwater.

The Main challenge for coastal engineers is to dissipate the wave energy before it reaches the coast. The Breakwater is one of the solutions to dissipate the wave energy. On the action of waves a stable rubble mound breakwater slowly changes its shape to berm breakwater and then into a tandem breakwater Shirlal, K. G., and Subba, Rao.,[9] The wave action on the rubble mound breakwater will slowly affect its stability and it will transform into an S shaped profile bench with an extension at toe.

After that the extended toe was separated from the breakwater as a submerged reef to the front of main breakwater and the resulting structure is a tandem breakwater (see figure 1.) Cox, J. C., and Clark, G. R., [1].

Since from the beginning, breakwaters have been designed by conducting physical model studies Van der Meer., [10], Hall, and Kao., [3], Shirlal, K. G. et. al., [9], which are uneconomic and laborious.

Soft computing techniques are those methods which are capable to model real world problems, have got flexibility in information processing and it can handle difficult situations. These have got tolerance in imprecision, uncertainty, reasoning so that low cost solution can be achieved with robustness and tractability. These techniques are different from the conventional techniques such that it will provide us with an acceptably low cost solution for an imprecisely or precisely framed problem. Soft computing techniques consider human mind as a role model.

Some of the soft computing techniques are Fuzzy Logic (FL), Neural Computing (NC), Evolutionary

Received: January 05, 2017; Accepted: May 22, 2017; Published: June 30, 2017

Computation (EC) Machine Learning (ML) and Probabilistic Reasoning (PR).

In order to minimize the cost and time factor soft computing tools such as Artificial Neural Network (ANN), Support Vector Machine (SVM), Adaptive Neuro Fuzzy Inference System (ANFIS) etc. can be adopted to make predictions. These techniques are recent developments which can be applied in the area of structural design and computations. It will save a large amount of time in planning and design, which in turn reduce the overall cost.

2. Literature and purpose of paper

Mandal, S., et.al, [6] predicted the damage level of non-reshaped berm breakwater using ANN, ANFIS and SVM technique and from the results obtained from statistical measures such as MSE= 3.361 and 11.106, RMSE=1.833 and 3.33, SI= 0.153 and 0.250 for training and testing, respectively, they justified that SVM is one of the reliable technique which can be used as an alternate tool to predict the damage level of non- reshaped berm breakwater efficiently. It is noticed that only few literatures are available which is not sufficient to understand the efficiency of SVM. Hence, further research is required to confirm the reliability and efficiency of SVM technique to predict the damage level of different type of breakwaters. The present paper involves the development of Support Vector Machine (SVM) model for predicting the damage level of the main breakwater of tandem breakwater. The accuracy/reliability of the developed SVM model is checked by using statistical measures such as correlation coefficient (CC), root-mean-square error (RMSE), Nash Sutcliff Efficiency (NSE) and scatter index (SI) which are defined as:

Correlation coefficient:

$$CC = \frac{\sum_{i=1}^{n}(O_i - \bar{O}_i)(P_i - \bar{P}_i)}{\sqrt{\sum_{i=1}^{n}(O_i - \bar{O}_i)^2 (P_i - \bar{P}_i)^2}}$$

Root Mean Square Error:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(O_i - P_i)^2}$$

Scatter Index:

$$SI = \frac{RMSE}{\bar{O}_i}$$

Nash Sutcliff Efficiency:

$$NSE = 1 - \frac{\sum_{i}^{n}(O_i - \bar{P}_i)^2}{\sum_{i}^{n}(O_i - \bar{O}_i)^2}$$

Where, O_i and P_i are observed and predicted damage level respectively, n is the number of data set used. $\bar{O}_i$ and $\bar{P}_i$ are average observed and predicted damage level respectively.

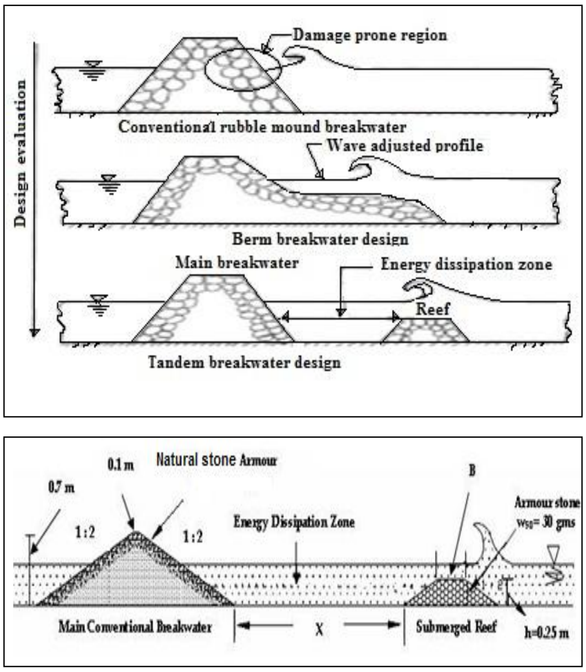


Figure 1. *Tandem breakwater evolution and experimental setup*

3. Methodology

Step.1: Experimental data collection
The data used in developing the SVM model for estimating the damage level of tandem breakwater are collected from the department of Applied Mechanics and Hydraulics, NITK Surathkal, Mangaluru, Shirlal, K. G. et. al., [9].

Eight input parameters, namely Incident wave steepness (H_i/gT^2), Relative reef spacing (X/d), Hudson's stability number ($H_i/\Delta D_{n50}$), Relative reef crest width (B/d), (B/Lo), Relative reef submergence (F/H_i), Relative reef crest height (h/d), Relative depth (d/gT^2) and output parameter damage level (S) are considered. The network is trained using the training data set containing 201 data samples (70%) and using the developed SVM model with weights and biases obtained, the network is then used for predicting with the testing data set of 87 data samples (30%).

Step.2: Normalization of data set
Normalization of the data set is required before presenting it to the network for its learning, so that it satisfies the activation function range. Normalization is also necessary if there is a wide difference in the ranges of input and output samples. Normalization process enhances the learning speed of the network and it avoids the possibility of early network saturation. Normalization of each variable is done using the equation below, to make them in the range (Min value, Max value).

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

Where,
X_{norm} : The data points normalized between 0 to 1
X : Each data point

X_{max} : maximum among all the data points
X_{min} : minimum among all the data points

Step.3: Separation of data
Once the data are collected, sorted the data points using the time series plots and then it is normalized. From the knowledge obtained from literatures 70% of the data is fixed for training and the rest 30% was used for testing based on the trial and error method.
Step.4: Development of SVM models
The consistency of the data was checked by plotting data time series plot. Using the data SVM models is developed by taking all the eight input variables using suitable kernel functions for predicting the damage level of the conventional main breakwater of the tandem breakwater. The results obtained is compared with the experimental data to see how much accurate is the prediction.

Step.5: Performance Evaluation
In the present study, the performance of the model is checked by using statistical measures viz correlation coefficient (CC), root-mean-square error (RMSE), Nash Sutcliff Efficiency (NSE) and scatter index (SI).

3.1. Principle and governing equations of SVM

In the present study, SVM technique is used for the prediction because the previous literature says that it is one of the reliable soft computing techniques compared to other conventional networks based on the following unique features. Many attractive features and promising empirical performance also helps to choose SVM as a tool for prediction. The formulation embodies the Structural Risk Minimization (SRM) principle by Gunn., [2], which has been shown as superior to traditional Empirical Risk Minimization (ERM) principle, employed by conventional neural networks by Mandal, S. et. al., [7] and Rao, S., et. al., [8] SRM minimizes an upper bound on the expected risk, as opposed to ERM that minimizes the error on the training data. It is this difference which equips SVM with a greater ability to generalize, which is the goal in statistical learning. SVMs are developed to solve the classification problem, but recently they have been extended to the domain of regression problems Vapnik et al., [11].

It is an exclusively data based nonlinear modeling paradigm. SVM based models help in greater extent to generalize, which works on the principle of structural risk minimization. Solving a quadratic optimization problem with the parameters of SVM model can be obtained. Quadratic form possesses a single minimum which is an objective function of SVM, thus avoiding the heuristic procedure involved in locating the global or the deepest local minimum on the error surface. The inputs are first nonlinearly mapped into a high dimensional feature space which is further correlated linearly with the output.

Let a training set is $\{ (x_i, y_i),\ i = 1,\ldots\ldots,n\}$, where x_i and y_i are i^{th} input training pattern & i^{th} target output and has *n* data sets for training. For nonlinear case SVR has the form:

$$f(x, \alpha) = (w \cdot \varphi (x)) + b \qquad (1)$$

Where, *w* is the weight vector and *b* is the bias. φ (x) is a mapping function to a higher dimensional feature space from input features. The regression problem will be similar as minimizing the regularized risk function Lee, et. al., [5]. Minimize,

$$R (f) = \frac{1}{n} \sum_{i=1}^{n} L (yi, f(xi, w)) \qquad (2)$$

Where,

$$L (yi , f(xi, w)) = \begin{cases} \varepsilon, \text{ if } | yi - f(xi, w) | \leq \varepsilon \\ | yi - f(xi, w) | - \varepsilon, \text{ otherwise} \end{cases}$$

Where, ε is insensitive loss function figure 2, on substitution in Eq. (2), optimization object will be Minimized,

$$\frac{1}{2} w \cdot w + C.(\sum_{i=1}^{n} \xi i^* + \sum_{i=1}^{n} \xi i) \qquad (3)$$

$$\text{subject to } \begin{cases} yi - w.xi - b \leq \varepsilon + \xi i^* \\ w \cdot xi + b - yi \leq \varepsilon + \xi i, \quad i = 1,\ldots\ldots,n \\ \xi i, \xi i^* \geq 0 \end{cases}$$

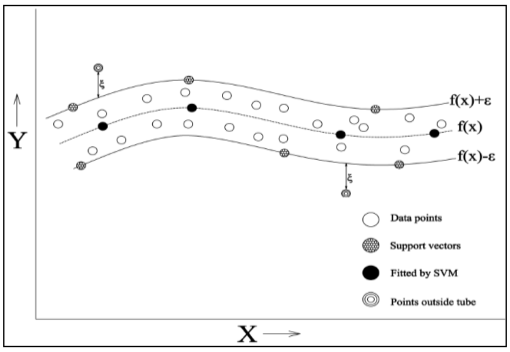


Figure 2. *A schematic diagram of support vector regression using e-insensitive loss function*

Where, the constant C > 0 means the penalty degree of the sample with error exceeds epsilon, ξ_i, ξ_i^* are slack variables. By using this optimization a dual problem can be attained by maximizing the function, maximize,

$$\sum_{i=1}^{n} yi (\alpha i^* - \alpha i^*) - \varepsilon \sum_{i=1}^{n} (\alpha i^* - \alpha i^*) -$$
$$-- \frac{1}{2} \sum_{i,j=1}^{n} (\alpha i^* - \alpha i) (\alpha i^* - \alpha i) \{ \varphi(xi) \cdot \varphi(xj) \}$$

$$\text{subject to } \begin{cases} \sum_{i=1}^{n} (\alpha i^* - \alpha i) = 0 \\ 0 \leq \alpha i^*, \ \alpha i \leq C \end{cases} \qquad (4)$$

Where α_i^* and α_i are Lagrange multiplier and $(\varphi (x_i)$. $\varphi (x_j) = K (x_i , x_j))$ is kernel function. By using the above maximization function the non-linear regression function is obtained,

$$f(x, \alpha^*, \alpha) = \sum_{i=1}^{n} (\alpha i^* - \alpha i) K(xi, x) + b \qquad (5)$$

Where,

$$w \cdot \varphi(x) = \sum_{SVs} (\alpha i^* - \alpha i) K(xi, x) \qquad (6)$$

$$b = -\frac{1}{2} \sum_{SVs} (\alpha i^* - \alpha i) [K(xr, xi) + K(xs, xi)] \qquad (7)$$

Where, x_r and x_s are support vectors, SVs number of support vectors and $0 \leq \alpha_i^*, \alpha_i \leq C$.

There are various kernels like linear, polynomial, radial basis function, sigmoid kernel, etc. Here we have obtained precise results of radial basis kernel with 3 number of kernel parameters, 0.00007 insensitive loss value obtained is less compared to other kernel functions. A form of RBF kernel is used here:

$$K(xi, xj) = ((\gamma * x * y) + C) \quad (8)$$

Where, γ is a kernel function parameter, C- upper bound of the SVM model. The performance of the SVM model relies on the selection of the upper bound C and γ kernel parameters. If the above parameters are improperly selected, then this leads to under fitting or over fitting problems Hsu et al., [4].

4. Results and discussions

The experimental data for developing the SVM models are taken from the physical model studies on the tandem breakwater by Shirlal, K. G., et. al., [9] in Marine Structures wave flume laboratory in Applied Mechanics & Hydraulics Department, NITK Surathkal, India.

The collected data set is checked for its consistency using time series plots, then normalized and further separated into two sets for training and testing. There are 288 data sets from which the data has been sorted at random such that 70% for training which includes all the range of data points for training the model and 30% for testing is taken after many trials with different combinations. Normalization of data points is done after checking consistency of data points using time series plots. Then statistical parameters are calculated and measured values versus predicted values are plotted in figure 3 and tabulated in table 2.

The scatter plots of SVM models with different kernel functions for the training and testing data, for damage level are shown in figure 3 in which the linear kernel function with SVM model shows poor correlation between observed and predicted data points compared to an SVM model with RBF kernel. The performance of SVM is assessed with different statistical measures, but most commonly used are the validation and estimation of leave-one-out error. The available data are divided into 70% for training and 30% for testing, which is chosen randomly to assess the performance of SVM.

Only if the system can perform well on test data which is not trained before, then the learning of SVM is considered to be successful. The good setting of meta-parameters C, ε , γ and kernel parameters such as kernel type and loss function type are responsible for better performance of SVM. The selection of optimal parameter is a complicated problem by the fact that the SVM model complexity depends on parameters such as C, ε,γ.

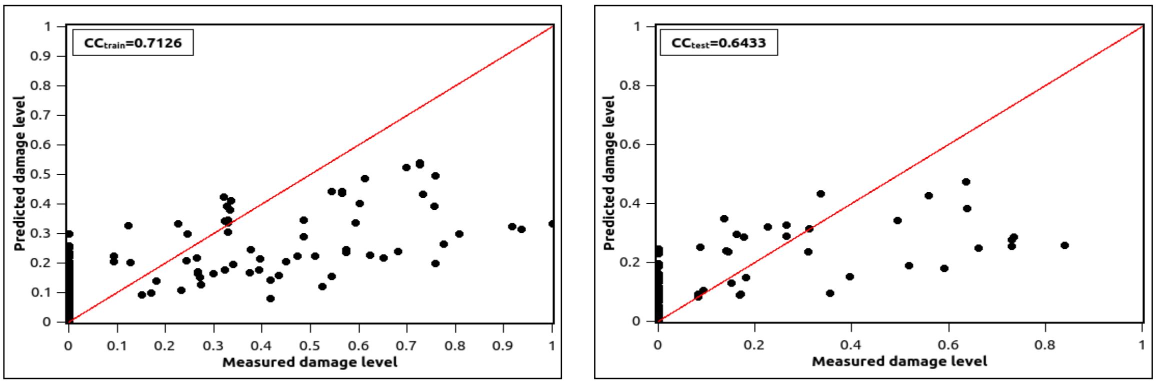
Table 1: Details of results obtained from the SVM model with different kernel functions

Type	Linear		Polynomial								RBF		Sigmoid	
			Quadratic		Cubic		Quartic		Quintic					
	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test
RMSE	0.175	0.170	0.094	0.091	0.102	0.073	0.106	0.071	0.122	0.072	**0.091**	**0.076**	0.194	0.180
NSE	0.488	0.412	0.854	0.832	0.912	0.787	0.914	0.771	0.911	0.697	**0.903**	**0.832**	0.376	0.342
CC	0.726	0.655	0.926	0.918	0.956	0.922	0.956	0.905	0.956	0.877	**0.953**	**0.935**	0.691	0.624
SI	1.293	1.215	0.690	0.648	0.776	0.504	0.806	0.496	0.927	0.504	**0.690**	**0.529**	1.367	1.341

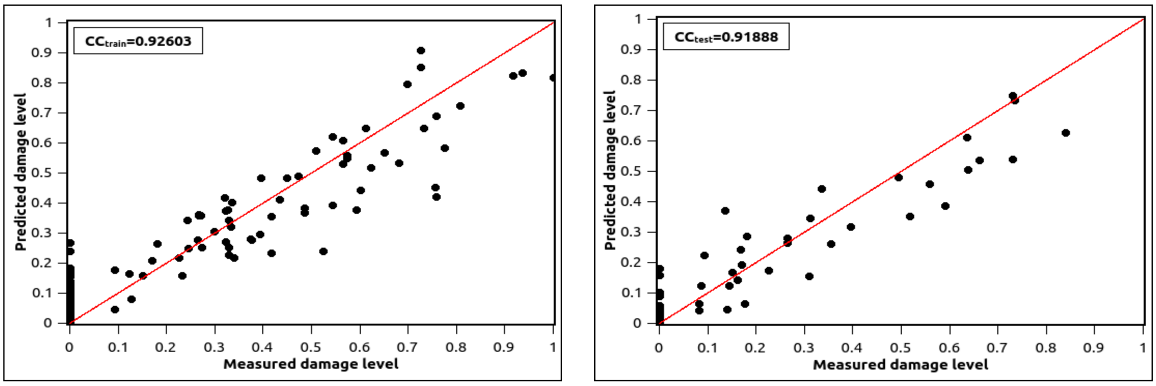
Table 2: Optimal parameters for SVM models with different kernel functions for S

Kernel Type	Linear	Polynomial	RBF	Sigmoid
C	2418	364	328	258
γ	-	-	3	4
ε	0.00169	0.00075	0.00007	0.00396
d	-	3	-	-
nsv	82	82	82	82

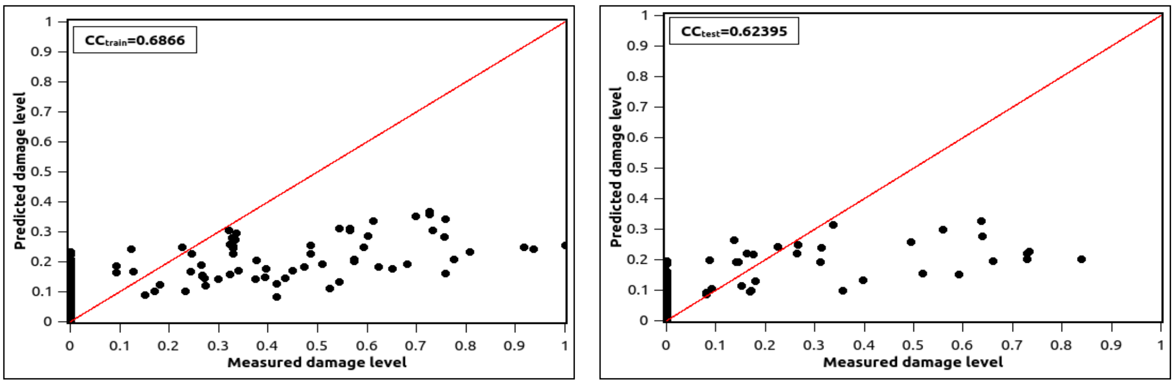
In the case of SVM model, the results obtained during training and testing processes for input parameters are evaluated using statistical measures like CC, RMSE and SI and NSE value as shown in table 1 for damage level. The performance of SVM depends on the good setting of kernel parameters. In developing SVM models, initially parameters are randomly selected by rough search (i.e. for C=100, 200, 300….2000; ε = 0.5, 1…2; = 1,2,…6 and γ = 1,2,...6) and d = 1,2,...6) to identify the near optimal values, and then a fine search (i.e. for C=1,10,20, 30….2000; ε = 0.000001,…2; γ = 0.01,0.02,…6 and d =1,2....6) is done to identify the final optimal values. In the case of RBF kernel, the optimal width γ obtained by the manual search is found to be 3. The number of support vectors for the prediction of the damage level of the main breakwater of the tandem breakwater is same for all the kernel functions. Cross validation, search is used for finding the optimal values of C, γ, ε as listed in table 2 for damage level. The SVM model with RBF kernel function shows better performance in prediction of S with CC=0.9537, RMSE=0.0910, SI=0.6900, NSE=0.9030 for training and CC=0.9357, RMSE=0.0765, SI=0.5293, NSE=0.8327 for testing.



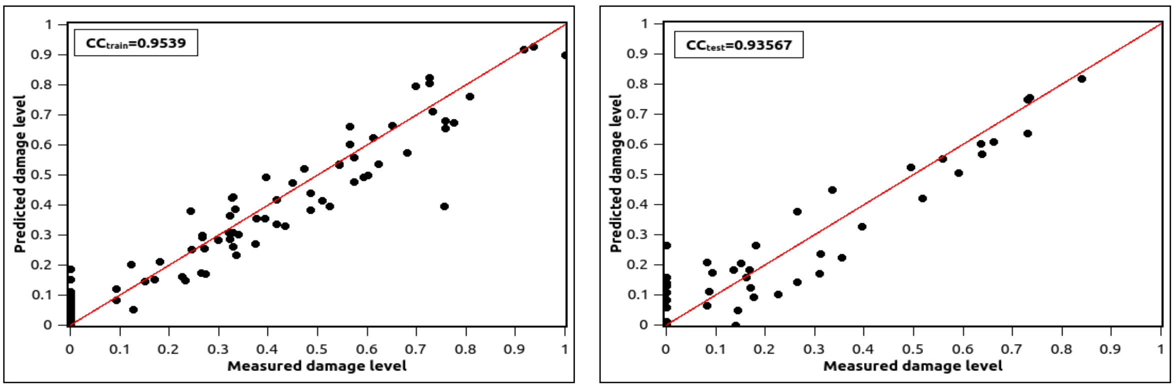
(a) Linear Kernel function train and test



(b) Polynomial Kernel function train and test



(c) Sigmoid Kernel function train and test



(d) RBF Kernel function train and test

Figure 3. *Scatter plots of predicted and observed damage level for an SVM model with different kernel function for training and testing*

In the figure 4 the performance variation of SVM (RBF kernel function) predictions with observed data is shown. The damage level predictions by SVM with RBF kernel function are in good agreement with observed values and are in good correlation with observed data points. The zero values in the graph indicate the zero damage at those places while conducting the experiments in a wave flume as well as in the predictions.

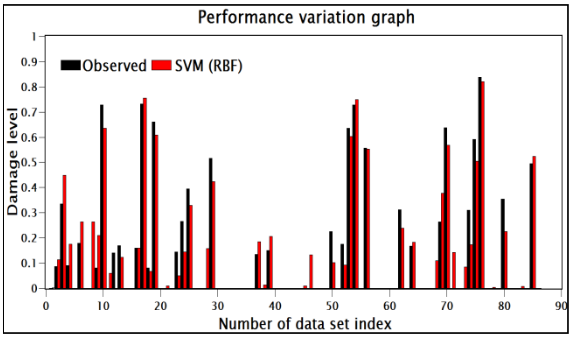


Figure 4. *Performance graph of SVM (RBF) in prediction of damage level with observed data*

5. Summary and conclusions

In this paper, we applied SVM to predict damage level. It is important and critical governing parameter on which the stability of the tandem breakwater is measured. The experimental data generated by wave flume at Marine Structure laboratory, NITK, Surathkal, India are used in this study. From 288 data sets randomly selected 201 (70%) data points are used for training the network and 87 (30%) data sets are used for testing the network.

The SVM models with different kernel functions are studied. Among these models, SVM model with RBF kernel function with upper bound C= 328 and γ= 3 kernel parameters yields better results compared to other kernel functions.

Based on the present study the following conclusions are drawn:

a. Among all the kernel functions in case of damage level prediction, SVM model with RBF (radial basis function) kernel function gives the better performance compared to other kernel functions with CC 0.953 and 0.935, RMSE 0.091 and 0.076, SI 0.690 and 0.529, NSE 0.903 and 0.832 for training and testing for damage level respectively.

b. From the present study it can be concluded that SVM with RBF kernel function model is an effective tool in predicting the damage level of the Conventional rubble mound breakwater of tandem breakwater and fast in predicting the damage values. Therefore, it can be used as an alternative tool to determine the damage level of tandem breakwater.

Acknowledgement

The authors are thankful to the Director of National Institute of Technology Karnataka, Surathkal and Head of the Department of Applied Mechanics and Hydraulics for the facilities provided and permission granted to publish the results. Also the authors would like to thank INCH, Ministry of Water Resources, GOI for providing financial support to carry out the work.

References

[1] Cox, J. C., and Clark, G. R., "Design Development of a Tandem Breakwater System for Hammond, Indiana", In Proceedings, Coastal Structures and Breakwaters Conference, Thomas Telford, London pp.111-121, 1992

[2] Gunn, S., "Support Vector Machines for Classification and Regression", ISIS Technical Report, 1997

[3] Hall, K. R, and Kao, J.S., "The influence of armor stone gradation on dynamically stable breakwaters", Journal of Coastal Engineering, 15(4), pp.333-346, 1991

[4] Hsu, C. W., Chang, C. C., and Lin, C. J., "A practical guide to support vector classification", (2003-06-15) (2012-03-01), 2004, http://www. csie. ntu. edu. Tw.

[5] Lee, J. J., Kim, D. K., Chang, S. K., and Lee, J. H., "Application of support vector regression for the prediction of concrete strength". Journal of Computers and Concrete, Vol. 4, No. 4, pp. 299-316, 2007

[6] Mandal, S., Rao, S., and Harish, N., "Damage level prediction of non-reshaped berm breakwater using ANN, SVM and ANFIS models", International Journal of Naval Architecture and Ocean Engineering, 4(2), pp.112-122, 2012

[7] Mandal, S., Rao, S., and Raju, D. H., "Ocean wave parameters estimation using back propagation neural networks", Marine structures, 18(3), pp.301-318, 2005

[8] Rao, S., and Mandal, S., "Hindcasting of storm waves using neural networks", Ocean Engineering, 32(5), pp.667-684, 2005

[9] Shirlal, K. G., and Rao, S., "Laboratory studies on the stability of tandem breakwater". ISH Journal of Hydraulic Engineering, 9(1), pp.36-45, 2003

[10] Van der Meer, J.W., "Stability of the seaward slope of berm breakwaters". J. Coastal Engineering, 16(2), pp.205-234, 1992

[11] Vapnik, V., "The Nature of Statistical Learning Theory", Springer Verlag, New York, 1995