

# Prediction of Transmitted Wave Height of Tandem Breakwater using Support Vector Regression

Jithin J. S.<sup>1</sup>, Geetha Kuntoji<sup>2</sup>, Subba Rao<sup>3</sup>, Manu<sup>3</sup>, Mandal S.<sup>3</sup>

P.G. Student, Dept. of Applied Mechanics & Hydraulics, NITK, Surathkal, India<sup>1</sup>

Research Scholar, Dept. of Applied Mechanics & Hydraulics, NITK, Surathkal, India<sup>2</sup>

Professor, Dept. of Applied Mechanics & Hydraulics, NITK, Surathkal, India<sup>3</sup>

**ABSTRACT:** India has got a fast growing economy which has been possible due to the improvements in connectivity from coast to the interior parts of the country. These improvements has brought too much stress on coastal environment leading to its damage. Therefore a need of protective structures along the coast arise. Breakwaters are the most commonly used protective structure and it can also be used as a facility for loading passengers and cargo. Final layout and cross section will be selected after conducting physical model studies. Soft Computing can be used to model various problems of real case scenario where mathematical modelling is difficult. Support Vector Machine (SVM) is a machine learning tool where algorithms can be formulated for data analysis, to recognize patterns, to classify data, regression analysis, etc. For the regression analysis Support Vector Regression (SVR) is widely used.

**KEYWORDS:** Tandem Breakwater, Support Vector Regression, Kernel

## I. INTRODUCTION

### 1.1 Breakwater

Breakwaters are those structures which are built to protect ports and harbors from wave attack and also to maintain a calm condition, it's also used to protect river mouth from silting and to protect beaches. The breakwater can be used as a structure for both dissipating wave energy and as a facility for loading and unloading passengers and cargos. Effect of wave attack on breakwater can be reduced by placing a submerged breakwater in front of the main breakwater, this is called tandem breakwater. The submerged breakwater will dissipate the some of the wave energy of the incoming wave and reduce its impact on the main breakwater. This combination of a main breakwater with a submerged breakwater in the front is called tandem breakwater.

#### 1.1.1 Tandem Breakwater

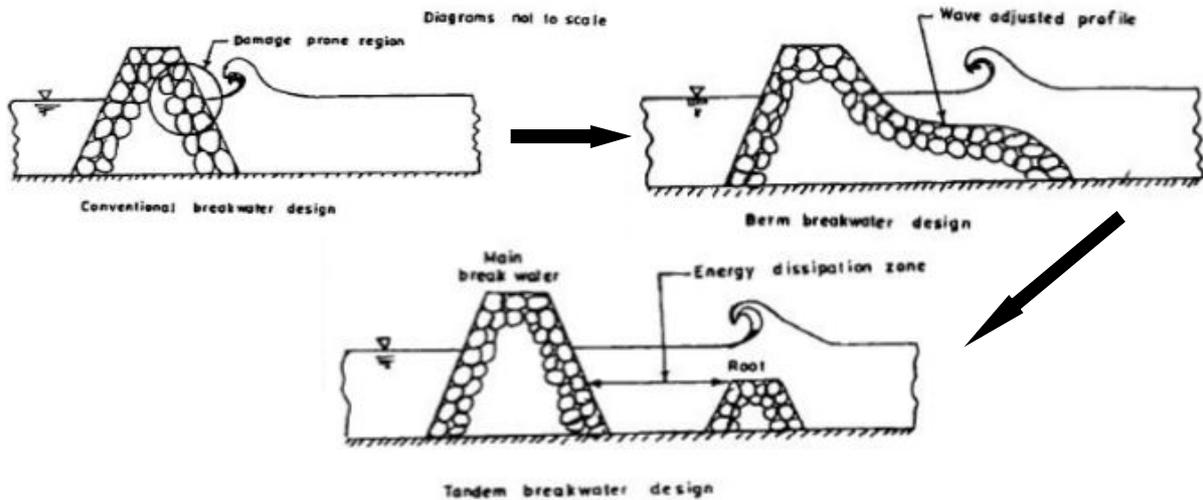
The tandem breakwater was derived from the most stable rubble mound breakwater to berm breakwater and then finally to tandem breakwater [5]. Rubble mound breakwater slowly stabilise by forming an S shaped profile bench with a toe extension. Later it was noted that we can shift the extended toe as a separate submerged breakwater to the front of main breakwater and the result is tandem breakwater (see Figure 1) [2].

From ancient days the breakwaters have been designed by conducting experiments on physical models (Van der Meer, 1988, 1992; Hall and Kao 1991; Subba Rao, Pramod and Rao, 2004; Subba Rao, Subrahmanya, Rao and Chandramohan, 2008) which are time consuming and costly. For minimizing both soft computing tools like Artificial Neural Network (ANN), SVM, Adaptive Neuro Fuzzy Inference System (ANFIS), etc can be utilized for predicting the experimental results. So these are new areas of interest which can be applied to the field of structural design and calculation there by reducing an enormous amount of time in planning and design which will in turn lead to the reduction in overall cost.

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Special Issue 9, May 2016



**Figure 1. Tandem Breakwater Evolution**

## 1.2 Soft Computing Techniques

Soft computing techniques is a group of methods which are capable of modelling real world problems with more flexibility in processing information and has got good capacity to handle vague situations. These techniques can handle for a tolerance in imprecision, uncertainty, reasoning so that it can provide with solutions of low cost and achieve robustness as well as tractability. The main idea of these techniques is to give us an acceptable low cost solution for an imprecisely or precisely framed problem. Thus soft computing techniques is entirely different from conventional techniques. In short human mind is the role model for soft computing techniques.

Some of the soft computing techniques are Fuzzy Logic (FL), Neural Computing (NC), Evolutionary Computation (EC) Machine Learning (ML) and Probabilistic Reasoning (PR).

### 1.2.1 Support Vector Regression

Vapnik 1995, developed SVM and is getting acceptance because of its striking features and its better performance. SVM uses Structural Risk Minimization (SRM) principle for the formulation [4] which is better than outdated Empirical Risk Minimization (ERM) principle which conventional neural networks uses. SVM was primarily used for classification problems, then its use was extended to regression problems [13].

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

$$f(x, a) = (w \cdot \phi(x)) + b \dots\dots\dots (1)$$

where,  $w$  is the weight vector and  $b$  is the bias.  $\phi(x)$  is a mapping function to a higher dimensional feature space from input features. The regression problem will be similar as minimising the regularized risk function [6].

$$\text{Minimize,} \quad R(f) = \frac{1}{n} \sum_{i=1}^n L(y_i, f(x_i, w)) \dots\dots\dots (2)$$

$$\text{where,} \quad L(y_i, f(x_i, w)) = \begin{cases} \varepsilon, & \text{if } |y_i - f(x_i, w)| \leq \varepsilon \\ |y_i - f(x_i, w)| - \varepsilon, & \text{otherwise} \end{cases} \dots\dots\dots (3)$$

where,  $\varepsilon$  is insensitive loss function, on substitution in Eq. (2), optimization object will be [5]:

$$\text{Minimize} \quad \frac{1}{2} w \cdot w + C \left( \sum_{i=1}^n \xi_i^* + \sum_{i=1}^n \xi_i \right) \dots\dots\dots (4)$$

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Special Issue 9, May 2016

$$\text{subject to } \begin{cases} y_i - w \cdot x_i - b \leq \epsilon + \xi_i^* \\ w \cdot x_i + b - y_i \leq \epsilon + \xi_i^*, \quad i = 1, \dots, n \\ \xi_i, \xi_i^* \geq 0 \end{cases}$$

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

$$\text{Maximize, } \sum_{i=1}^n y_i (\alpha_i^* - \alpha_i) - \epsilon \sum_{i=1}^n (\alpha_i^* - \alpha_i) - \frac{1}{2} \sum_{i,j=1}^n (\alpha_i^* - \alpha_i) (\alpha_j^* - \alpha_j) \{ \varphi(x_i) \cdot \varphi(x_j) \} \dots \dots \dots (5)$$

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

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

$$f(x, \alpha^*, \alpha) = \sum_{i=1}^n (\alpha_i^* - \alpha_i) K(x_i, x) + b \dots \dots \dots (6)$$

where,

$$w \cdot \varphi(x) = \sum_{i \in SV} (\alpha_i^* - \alpha_i) K(x_i, x) \dots \dots \dots (7)$$

$$b = -\frac{1}{2} \sum_{r,s \in SV} (\alpha_r^* - \alpha_r) [K(x_r, x_r) + K(x_s, x_s)] \dots \dots \dots (8)$$

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

Few kernel functions are linear, polynomial, radial basis function, sigmoid kernel, etc. For the present study the most accurate result was given by the polynomial kernel with fourth degree. The polynomial kernel used by the SVR for the best performance is:

$$K(x_i, x_j) = (\gamma * x * y + c)^d \dots \dots \dots (8)$$

where,  $\gamma$  is a kernel function parameter,  $c$  is the coefficient of the polynomial when  $c = 0$  polynomial is homogenous and  $d$  is the degree of the polynomial.

### 1.3 Statistical Parameters

For the current study various statistical parameters like Root Mean Square Error (RMSE), Nash-Sutcliffe Efficiency (NSE), Mean Absolute Error (MAE), Scatter Index (SI) and Correlation Coefficient (CC) were used to check the accuracy of predicted values with measured data sets. The model will be considered good when RMSE, MAE, SI are minimum while NSE and CC are maximum. Consider  $ob$  and  $pr$  as the observed and predicted values,  $\overline{ob}$  and  $\overline{pr}$  as the average of observed and predicted values number of observations is  $n$ , then,

RMSE is calculated using the formula:

$$\sqrt{\frac{\sum_{i=1}^n (ob - pr)^2}{n}} \dots \dots \dots (9)$$

NSE is calculated using the formula:

$$1 - \frac{\sum_{i=1}^n (ob - pr)^2}{\sum_{i=1}^n (ob - \overline{ob})^2} \dots \dots \dots (10)$$

MAE is calculated using the formula:

$$\frac{1}{n} \sum_{i=1}^n |pr - ob| \dots \dots \dots (11)$$

SI is calculated using the formula:

$$\frac{RMSE}{\overline{ob}} \dots \dots \dots (12)$$

CC is calculated using the formula:

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Special Issue 9, May 2016

$$\frac{\sum_i (a_i - \bar{a})(p_i - \bar{p})}{\sqrt{\sum_i (a_i - \bar{a})^2 \sum_i (p_i - \bar{p})^2}} \dots \dots \dots (13)$$

## II. OBJECTIVES AND METHODOLOGY

Objectives of the paper:

1. To check the applicability of a soft computing technique in the tandem breakwater design from the experimental data sets.
2. To find out which kernel in SVM is suited here.

Methodology adopted:

1. Collection of data
2. Checking consistency of the data
3. Dividing data for training and testing
4. Generating the model
5. Evaluation of various statistical parameters

## III. PRESENT STUDY

The present study focus on the prediction of the transmitted wave heights of submerged breakwater in front of the main breakwater. The experiment on the physical model of tandem breakwater has been conducted by Dr. Manu in Marine Structures flume lab in Applied Mechanics & Hydraulics Department, NITK Surathkal. The data sets for SVR is taken from the experimental data of 108 data sets. This 108 data set has been sorted at random and divided as 70% and 30% for training and testing respectively after various trials in different combinations. The data was normalised and then used in the model. The consistency of data set is checked by plotting data time series. The input for the epsilon-SVR model is  $F/H_i$ ,  $B/d$ ,  $X/d$ ,  $H_i/gT^2$  and the output was  $H_t/H_{t\max}$ . From the output the measured values vs predicted values is plotted.

## IV. RESULTS

Table 1 show the results obtained using SVR with different kernel functions. The various statistical parameters for comparison of models using different kernel functions were also calculated using the formulas as described in the previously.

**Table 1: Statistical values compared with observed and predicted values obtained from SVM**

Kernel Type	Linear	Polynomial				Radial Basis Function	Sigmoid
		Quadratic	Cubic	Quartic	Quintic		
<b>RMSE</b>	0.18493	0.09751	0.12254	0.15244	0.55091	0.10596	10.44114
<b>NSE</b>	0.30464	0.80666	0.69471	0.52749	-5.17097	0.77170	-2215.58078
<b>MAE</b>	0.14957	0.07229	0.09620	0.12565	0.31165	0.08455	7.44694
<b>SI</b>	0.39462	0.20808	0.26147	0.32529	1.17557	0.22611	22.27996
<b>CC</b>	0.55488	0.90033	0.83625	0.73808	0.28663	0.90697	-0.37782

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Special Issue 9, May 2016

Graphs were plotted for measured values vs predicted values with the results obtained using SVR for different kernel functions, for comparing all the results in similar manner a 45° line was drawn in all the plots as shown in graphs in figure 2. In the graph it's clearly visible how the points are spread around the 45° line, when points are closer towards the line much better is the correlation coefficient between measured and predicted values.

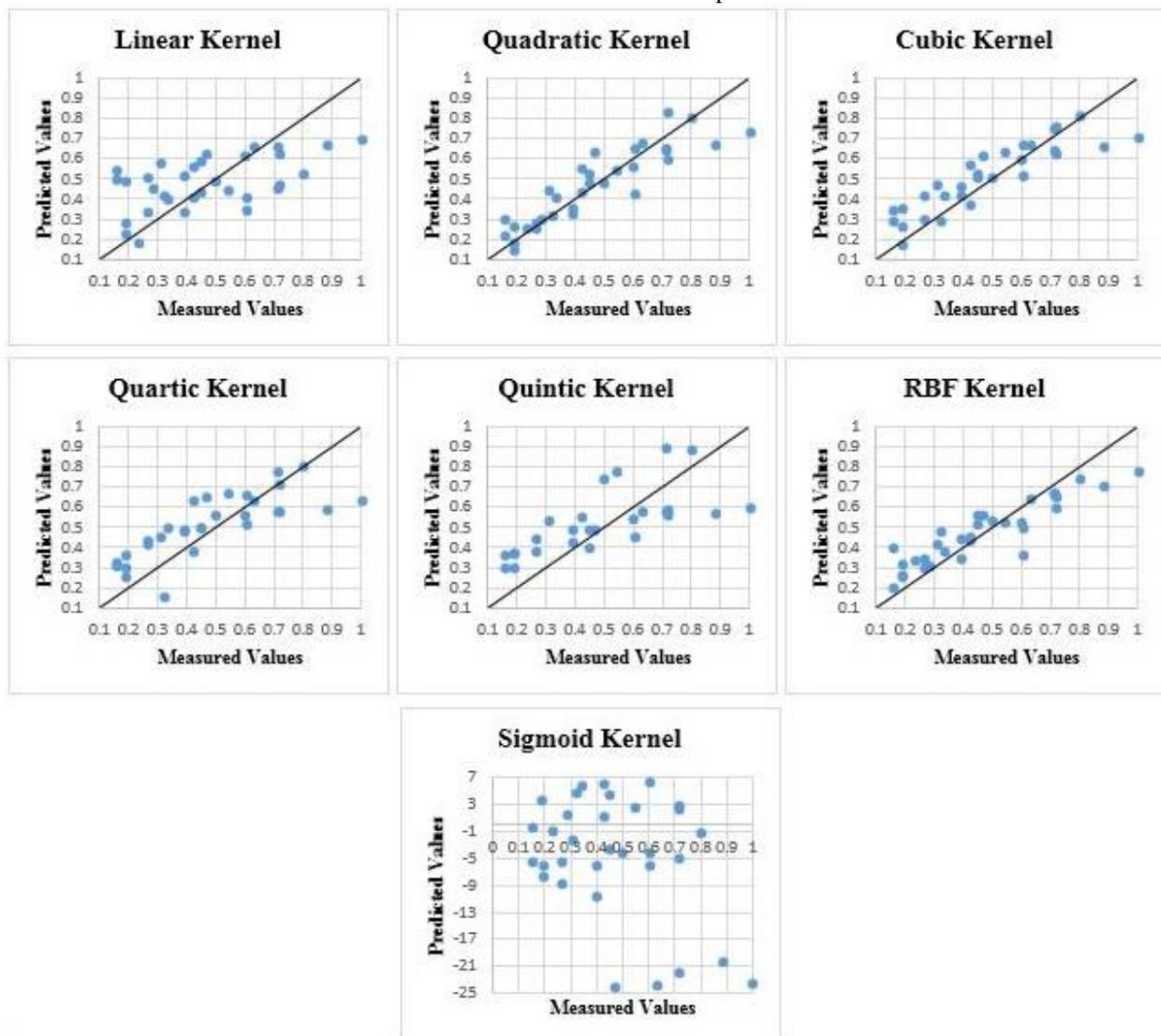


Figure 2. Plots of Measured Ht vs Predicted Ht obtained using SVR with different kernels

## V. DISCUSSIONS

When comparing the statistical values obtained from epsilon-SVR model using different kernel function we can see that the quadratic kernel and Radial Basis Function (RBF) kernel gives comparable results. The correlation coefficient obtained from the RBF kernel is slightly more than that obtained from the quadratic kernel. But quadratic kernel gives all other statistical parameters like RMSE, NSE, MAE and SI a slight better result than the RBF kernel. By looking on to the plots of measured values vs predicted values we can see that the above quoted discussion is true, because when the points are closer to the 45° line then the CC value is maximum and also when the points are closely

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Special Issue 9, May 2016

occurring on the plots it indicates that the measured values and the predicted values is not differing much which indirectly says that NSE is comparatively more whereas RMSE, MAE and SI will be comparatively less. But the sigmoid kernel is providing a result which is way out of the actual results. This may be due to the data used for this particular model. This can be concluded only with further studies.

## VI. CONCLUSIONS

1. Based on the above results we can say that soft computing techniques will yield good results as that of experimental results, because here we obtained a correlation coefficient around 0.90 which is good. This can be improved by using various optimization techniques like ant colony optimisation, particle swarm optimisation, cuckoo search etc.
2. For the present study the best kernel suited for SVR was quadratic kernel. Also the RBF kernel provide comparable results.
- 3.

## REFERENCES

- [1] C. Cortes and V. Vapnik, "Support vector networks," Machine Learning, vol. 20, pp. 273–297, 1995
- [2] Cox, J. C. and Clark (1992). Design Development of a Tandem Breakwater System for Hammond Indiana. Proc. of Conf. on Coastal Structures and Breakwaters, ICE, Thomas Telford Publishers London, pp. 111-121.
- [3] Hall, K.R. and Kao, J.S., 1991. The influence of armor stone gradation on dynamically stable breakwaters. J. Coastal Engineering, 15(4), pp.333-346.
- [4] Gunn,S. (1997). "Support Vector Machines for Classification and Regression", ISIS Technical Report.
- [5]K. G. Shirlal & Subba Rao M.ISH (2003) laboratory studies on the stability of tandem breakwater, ISH Journal of Hydraulic Engineering, 9:1, pp 36-45
- [6] Lee, J. J., Kim, D. K., Chang, S. K., and Lee, J. H. (2007). "Application of support vector regression for the prediction of concrete strength." Computers and Concrete, Vol. 4, No. 4, pp. 299-316.
- [7] Stitson, M. O., Weston, J. A. E., Gammerman, A., Vork, V., and Vapnik, V. (1996). Theory of support vector machines, Technical Report CSD-TR-96-17, Department of Computer Science, Royal Holloway College, University of London.
- [8] Subba Rao, Pramod, C.H. and Rao, B.K., 2004. Stability of berm breakwater with reduced armor stone weight. J. Ocean Engineering, 31(11-12), pp.1577–1589.
- [9] Subba Rao, Subrahmanya, K., Rao, B.K. and Chandramohan, V.R., 2008. Stability aspects of nonreshaped berm break-waters with reduced armour weight. Journal of Waterway, Port, Coastal and Ocean Engineering, 134(2), pp.81-87.
- [10] Tanimoto, K. and Goda, Y. (1992). Historical Development of Breakwater Structures in the World. Coastal Structures and Breakwaters, Thomas Telford, London, pp. 193-206.
- [11] Van der Meer, J.W., 1988. Deterministic and probabilistic design of breakwater armor layer. Journal of Waterway, Port, Coastal and Ocean Engineering, 114(1), pp.66-80.
- [12] Van der Meer, J.W., 1992. Stability of the seaward slope of berm breakwaters. J. Coastal Engineering, 16(2), pp.205-234.
- [13] Vapnik, V., Golowich, S. and Smola, A.(1996). "Support vector method for function approximation regression estimation and signal processing", Adv. in Neural Inform.Proces.Syst, 9 281-287.
- [14] Vapnik, V. (1995). "The Nature of Statistical Learning Theory", Springer Verlag, New York.