Hybrid genetic algorithm tuned support vector machine regression for wave transmission prediction of horizontally interlaced multilayer moored floating pipe breakwater

S.G. Patil  
Department of Built and Natural Environment, Caledonian College of Engineering, PO Box: 2322, CPO Seeb, PC 111, Sultanate of Oman

S. Mandal  
Ocean Engineering Division, National Institute of Oceanography, Goa. 403 004, India

A.V. Hegde  
Department of Applied Mechanics and Hydraulics, National Institute of Technology Karnataka, Surathkal, Srinivasanagar, Mangalore 575025, India

A. Muruganandam  
Department of Mechanical and Industrial Engineering, Caledonian College of Engineering, PO Box: 2322, CPO Seeb, PC 111, Sultanate of Oman

Abstract
Support Vector Machine (SVM) works on structural risk minimization principle that has greater generalization ability and is superior to the empirical risk minimization principle as adopted in conventional neural network models. However, it is noticed that one particular model in isolation cannot capture all data patterns easily. In the present paper, a hybrid genetic algorithm tuned support vector machine regression (HGASVMR) model was developed to predict wave transmission of horizontally interlaced multilayer moored floating pipe breakwater (HIMMFPB). Furthermore, parameters of both linear and nonlinear SVM models are determined by Genetic Algorithm. HGASVMR model was trained on the dataset obtained from experimental wave transmission of HIMMFPB using regular wave flume at Marine Structure Laboratory, National Institute of Technology, Surathkal, India. The results are compared with artificial neural network (ANN) model in terms of Correlation Coefficient, Root Mean Square Error and Scatter Index. Performance of HGASVMR is found to be reliably superior.

1 INTRODUCTION

There is a great volume of published work dealing with floating breakwaters (Bishop, 1982; Harms, 1979; Harris and Webber, 1968; Homma et al., 1964; Kennedy and Marsalek, 1968; Leach, McDougal and Solitt, 1985), but it is noticed that there is a lack of a simple mathematical model to predict breakwater performance characteristics such as the transmission coefficient. A number of studies have been carried out considering a floating breakwater in basic form with some assumptions common in hydrodynamics which shows less improvement. In the last two decades, floating breakwaters (Mani, 1991; McCartney, 1985; Murali and Mani, 1997; Sannasiraj et al., 1998; Sundar et al., 2003) have generated a great interest in the field of coastal engineering, as they are less expensive compared to conventional type breakwaters. In addition, they have several desirable characteristics such as, comparatively small capital cost, adoption to varying harbour shapes and sizes, short construction time and freedom from silting and scouring. For effective design of floating pipe breakwater, it is necessary to study the hydrodynamic performance characteristics of this structure. Hence, a study on wave transmission of the floating pipe breakwater would provide a proper configuration to the structure. In recent years, the research interest in Artificial Neural Networks (ANN) has increased and many efforts have been made on applications of neural networks to various coastal engineering problems. ANN in coastal/ocean engineering is commonly used by the researchers to predict ocean wave parameters like wave height, wave period, impact wave force etc. (Deo et al., 2001; Deo and Jagdale, 2003; Gunaydin, 2008; Londhe and Deo, 2003). Apart from this, it has provided promising results in prediction of tidal levels (Chang and Lin, 2006), damages to coastal structures (Mandal et al., 2007), depth of eroded caves in
needs to find network parameters such as number of hidden layers and neurons by trial and error, which is time consuming. Jeng et al. (2004) adopt the concept of genetic algorithm based training of ANN models in an effort to overcome the problems inherent in ANN training procedures while providing accurate results for determining maximum liquefaction depth in a real world application. It is also noticed that apart from improving the performance of ANN, computational effort and time needed for training and testing the model is significantly reduced compared to traditional methods (Mandal et al., 2007). This fact leads us to expect neural networks to be an excellent tool for solving the motion characteristics of the floating pipe breakwater while overcoming complexity and non-linearity associated with wave-structure interaction. When the performance of ANN alone is poor in mapping input-output relation many researchers developed a hybrid model by combining the ANN with fuzzy systems, ANN with numerical wave modeling (Kazeminezhad et al., 2005), (Sylaios et al., 2009), (Malekmohamadi et al., 2008). Apart from ANNs, many authors have used a new approach to solve coastal engineering problems like genetic programming by Gaur and Deo (2008) for real time wave forecasting, Guven et al. (2009) for prediction of circular pile scour. Adaptive neuro-fuzzy inference system by Sylaios (2009) for wave wave modeling, Model trees by Shahidi and Mahjoobi (2009) for prediction of significant wave height. Support vector machines by Han et al. (2007) for flood forecasting, Radhika and Shashi (2009) for prediction of atmospheric temperature, Msiza et al. (2008) used ANN and SVR for water demand prediction, Rajasekaran et al. (2008) develop a support vector machine regression (SVMR) model for forecasting storm surges. He compared these results with numerical methods and ANN which indicates that storm surges and surge deviations are efficiently predicted using SVMR. According to Mahjoobi and Mosabbeb (2009) SVM creates a more reliable model with better generalization error, in comparison to ANN, He also reveals that SVMs do not over fit, while ANNs may face such problem and need to deal with it.

SVM and ANN were used for other applications as discussed above. However, it is observed that there are hardly any applications of soft computing tools on wave transmission of floating breakwater. This fact leads us to use ANN and SVM model in this work.

In the present paper, the performance of HGASVMR models for predicting wave transmission coefficient of Horizontally Interlaced Multilayer Moored Floating Pipe Breakwater (HIMMFPB) is investigated. GAs is used to optimize the SVMR and kernel parameters. Results of HGASVMR models are compared with that of ANN model. The paper is organized as follows—section 1 starts with literature associated with floating breakwaters and applications of soft computing techniques in coastal engineering. Section 2 details experimental HIMMFPB. Section 3 describes ANN model for estimation of $K_i$. Fundamentals of SVMR, GAs for parameter selection and proposed HGASVMR are detailed in section 4. Results and discussions are described in section 5. Conclusions and acknowledgments are presented in section 6 and 7 respectively.

2 ABOUT EXPERIMENTAL HIMMFPB MODEL AND DATA USED

The detail of the HIMMFPB has been shown in Figure 1 (Deepak, 2006; Hegde et al., 2007; Jagadisha, 2007; Karmat, 2010). The breakwater comprises of the rigid Poly Vinyl Chloride (PVC) pipes. The pipes are placed parallel to each other with certain spacing between them in each layer and the adjacent layers are oriented at right angles to each other so as to form an interlacing. Longitudinal pipes are placed along the direction of propagation of waves and transverse pipes are placed and tied perpendicular to longitudinal pipes. The length of the longitudinal pipes defines the width of the breakwater. It is felt that with proper number of layers, spacing of pipes and relative breakwater width, it is possible to achieve a considerable and effective attenuation of waves. Figure 1 shows a pictorial representation of the HIMMFPB model in plan and section. The wave specific parameters and structure specific parameters considered in the experiments are shown in Table 1. The experimental study carried out by Kamat (2010) shows hydrodynamic characteristics of horizontally interleaved three and five layer floating breakwater systems, in which wave transmission is less for five layer systems. These experimental data are divided
into two sets for training and testing the ANNs and SVMR models.

The input parameters that influence the wave transmission ($K_t$) of floating breakwater such as spacing of pipes relative to pipe diameter ($S/D$), breakwater width relative to wave length ($W/L$), incident wave relative to water depth ($H_i/d$), incident wave relative to wave length ($H_i/L$) are used to train ANN and SVMR models as shown in Table 2.

### 3 ESTIMATION OF $K_t$ BY ANN

Artificial neural network (ANN) is an information processing paradigm that is inspired by the way of biological nervous system like, the human brain process information. The ANN structure constructed consists of four input nodes, five hidden nodes and one output node. After training the network model, weights and biases of the network are fixed. Here each input value gets multiplied with the weight and adds with bias value, total sum is then the input at each hidden node and this is passed through a tansig transfer function. Further the outputs from hidden node get multiplied with the weight and add with the bias value and then the total sum was passed through purelin to get $K_t$.

### 4 HYBRID GENETIC ALGORITHM TUNED SUPPORT VECTOR MACHINE REGRESSION (HGASVMR)

#### 4.1 Fundamentals of Support Vector Machine Regression (SVMR)

The support vector machines (SVMs) were proposed by Vapnik (1998) and is based on statistical learning theory. The basic idea of support vector machines is to map the original data $X$ into a feature space with high dimensionality through a nonlinear mapping function and construct an optical hyper-plane in new space. Hence, given a set of data $S = \{(x_i,d_i)\}_{i=1}^{N}$, where $x_i$ is the input vector; $d_i$ is the desired result, and $N$ corresponds to the size of the data set. The SVM regression function (Smola and Scholkopf, 1998) is:

$$y = f(x) = w \phi(x) + b$$  \hspace{1cm} (1)

where $\phi(x)$ is the nonlinear function in feature of inputs $x$, and both $w$ and $b$ are coefficients which are estimated by minimizing the regularized risk function:

$$\text{Minimize} \ : \ R(C) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{N} L_{\epsilon}(d_i, y_i)$$ \hspace{1cm} (2)

where,

$$L_{\epsilon}(d_i, y_i) = \begin{cases} |d_i - y_i| - \epsilon, & \text{if } |d_i - y_i| \geq \epsilon, \\ 0, & \text{otherwise,} \end{cases}$$ \hspace{1cm} (3)

Equation 3 defines a range where the loss will be zero if the forecasted value is within the $\epsilon$-tube (Fig. 2). However, if the value is out of the $\epsilon$-tube then the loss is the absolute value which is the difference between forecasted value and $\epsilon$. Introducing two positive slack variables $\zeta_i$ and $\zeta_i^*$ in Equation 2 it is possible to transform it into a primal objective function. The detail explanation is given in Smola and Sholkopf (1998).

In the present paper, we have experimented with the six kernels. To optimize these parameters for better generalization of SVM model, SVM model is hybridized with GAs. Section 4.2 details interface of genetic algorithim with support vector machine regression to obtain the best hybrid model.

#### 4.2 Genetic Algorithm for Selecting Parameters in the SVMR Model

Genetic algorithms are search methods based on principles of natural selection and genetics (Holland, 1975). The algorithm is based on the principle of the survival of the fittest which tries to retain genetic information from generation to generation. In this
paper GAs is used to search for better combination of C, ε and kernel parameters to maximize the generalization performance of SVMR model. Figure 3 shows the proposed HGASVMR model. The codes are written in MATLAB 7 Release 14. The steps involved in GA for selecting SVMs and kernel parameters are presented as follows:

**Step 1. (Initialization):** Initial population of chromosomes is generated. In the present paper population size was set to 50. The chromosomes are real coded string consists of SVMs parameters C, ε and kernel parameters.

**Step 2. (Evaluating fitness):** evaluate the fitness of each chromosome. In the present paper a negative normalized mean square error is used as the fitness function, which is defined as

\[
\text{Fitness Function} = \frac{1}{\sigma^2 N} \sum_{i=1}^{N} (d_i - y_i)^2
\]  

(4)

where \( \sigma^2 = \frac{1}{N} \sum_{i=1}^{N} (d_i - \overline{d})^2 \)

N is the total number of data in the test set, \( \overline{d} \) is the mean of the actual value, \( d_i \) is the actual value and \( y_i \) is the predicted value.

**Step 3. (New population):** new population is created by repeating following steps until the new population is complete

i) [Selection]: In the present study two parent chromosomes from a population are selected according to fitness function Equation 4. The roulette wheel selection principle (Holland, 1975) was used to select chromosomes for reproduction.

ii) [Crossover]: with crossover probability cross over the parents to form new offspring’s (children). In cross over, chromosomes are paired randomly.

iii) [Mutation]: After cross over operation is performed the string is subjected to mutation operation. The variable in the string to be mutated was selected randomly, where incremental operator is used. The rate of crossover and mutation was determined by probabilities. In the present paper the probabilities of crossover and mutation was set to 0.8 and 0.05 respectively.

iv) [Accepting]: place new offspring in the new population.

**Step 4. (Replace):** use new generated population for a further run of the algorithm

**Step 5. (Stop condition):** if the end condition is satisfied, stop, and return the best solution in current population. Otherwise go to step 2.

### 4.3 The Proposed HGASVMR Model

In the present study MATLAB support vector machine toolbox (Gunn 1998) was interface with genetic algorithm to optimize the SVMs and kernel parameters simultaneously for better generalization of the proposed HGASVMR model. The main goal is to find the optimal or near optimal parameters that produces the most accurate prediction. If the calculated fitness value satisfy the terminal condition in GAs, then the optimal parameters are selected, the new generation of the population is produced by applying genetic operators such as selection, crossover and mutation. These optimize SVMs and kernel parameters are shown in Table 3.

Optimized parameters obtained by GA are tested with the test data. The final decision about the optimum models is not based on the training data, but on the testing data, as illustrated in Figure 3. Once the testing is over the six models with different kernels
are compared based on statistical measures (Table 4) to get the best model.

5 RESULTS AND DISCUSSION

In the present study, updated LM algorithm (Wilamoski, et al, 2001) is used to train ANN Model. A high correlation coefficient (CC Test = 0.9419) is obtained at epoch equal to 200 with hidden nodes equal to five for ANN model, with SI 0.090 for trained data and 0.096 in case of test data (Table 4).

Figure 4 shows the comparison of predicted and measured $K_t$ for ANN model. The performance of ANN model is better compared to HGASVMR (linear) model, whereas the HGASVMR (bspline) outperformed ANN model with CC Test = 0.9741. Figure 5 shows the comparison of predicted and measured $K_t$ for HGASVMR (bspline) model. In comparison to all the models linear kernel function has shown poor generalization performance in prediction of $K_t$ of HIMMFPB. Increasing the will disturbed the solution, but it can be helpful for other kernels as shown in Table 3. In case of bspline kernel function is 15, whereas, in case of spline kernel it is 40000, both the kernel function has better generalization performance with RMSE 0.036. The optimal width of the function $\gamma$ for rbf and erbf are 0.3 and 0.4 respectively. The optimal value of $d$ in case of polynomial

<table>
<thead>
<tr>
<th>Kernel</th>
<th>nsv</th>
<th>C</th>
<th>$\varepsilon$</th>
<th>$\gamma$</th>
<th>$d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>2131</td>
<td>100</td>
<td>0.001</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Polynomial</td>
<td>2131</td>
<td>100</td>
<td>0.00001</td>
<td>–</td>
<td>6</td>
</tr>
<tr>
<td>Rbf</td>
<td>2131</td>
<td>150</td>
<td>0.001</td>
<td>0.3</td>
<td>–</td>
</tr>
<tr>
<td>Erbf</td>
<td>2131</td>
<td>100</td>
<td>0.001</td>
<td>0.4</td>
<td>–</td>
</tr>
<tr>
<td>Spline</td>
<td>2131</td>
<td>40000</td>
<td>0.05</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Bspline</td>
<td>2131</td>
<td>15</td>
<td>0.05</td>
<td>–</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 3 Values of optimal parameters for HGASVMR models with different kernels

Table 4 ANN and HGASVMR models with statistical measures for trained and test data

<table>
<thead>
<tr>
<th>Model</th>
<th>CC test</th>
<th>Train data</th>
<th>Test data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSE</td>
<td>SI</td>
<td>RMSE</td>
</tr>
<tr>
<td>HGASVMR (Linear)</td>
<td>0.8924</td>
<td>0.070</td>
<td>0.123</td>
</tr>
<tr>
<td>HGASVMR (Polynomial)</td>
<td>0.9513</td>
<td>0.046</td>
<td>0.081</td>
</tr>
<tr>
<td>HGASVMR (Rbf)</td>
<td>0.9478</td>
<td>0.046</td>
<td>0.081</td>
</tr>
<tr>
<td>HGASVMR (Erbf)</td>
<td>0.9478</td>
<td>0.005</td>
<td>0.009</td>
</tr>
<tr>
<td>HGASVMR (Spline)</td>
<td>0.9735</td>
<td>0.028</td>
<td>0.050</td>
</tr>
<tr>
<td>HGASVMR (Bspline)</td>
<td>0.9741</td>
<td>0.022</td>
<td>0.039</td>
</tr>
<tr>
<td>ANN</td>
<td>0.9419</td>
<td>0.051</td>
<td>0.090</td>
</tr>
</tbody>
</table>

Figure 4 Comparison of predicted and measured $K_t$ for HGASVMR (bspline) model.

6 CONCLUSIONS

An application of hybrid genetic algorithm tuned support vector machine regression for prediction of wave transmission for HIMMFPB is presented in this paper. Our proposed model optimizes SVMs and kernel parameters simultaneously. The performance of HGASVMR was compared with ANN. The results obtained show that HGASVMR with polynomial, rbf, erbf, spline and Bspline kernel functions performs better than ANN trained with back propagation LM updated algorithm. The forecasting performance of SVMR appears to be highly influenced by the choice of its kernel function. The Bspline kernel function performed superior than other kernel. It was also observed that parameter selection in
the case of SVMR has a significant effect on the performance of the model. HGASVSMR can replace the neural network based models for wave transmission prediction of HIMMFPB. HGASVSMR can be utilized to provide a fast and reliable solution in prediction of the wave transmission for HIMMFPB, thereby making HGASVSMR as an alternate approach to map the wave structure interactions of HIMMFPB.

7 ACKNOWLEDGEMENTS

The authors are grateful to the Director, and Head, Department of Applied Mechanics and Hydraulics, NITK, Surathkal, India for support and encouragement provided to them and for permission to publish the paper. Thanks are also due to Ministry of Earth Sciences, GOI for sponsoring the project on HIMMFPB at NITK, Surathkal, India.

REFERENCES


Holland, J. H., 1975. Adaption in natural and artificial system, the MIT Press, Massachusetts.


