Journal of Marine Science and Application (2019) 18:167–175
https://doi.org/10.1007/s11804-019-00088-4

RESEARCH ARTICLE

Below the Data Range Prediction of Soft Computing Wave Reflection of Semicircular Breakwater

Suman Kundapura [1] · Vittal Hegde Arkal [1] · Jose L. S. Pinho [2]

Received: 3 July 2017 / Accepted: 18 November 2018 / Published online: 8 April 2019

Abstract
Coastal defenses such as the breakwaters are important structures to maintain the navigation conditions in a harbor. The estimation of their hydrodynamic characteristics is conventionally done using physical models, subjecting to higher costs and prolonged procedures. Soft computing methods prove to be useful tools, in cases where the data availability from physical models is limited. The present paper employs adaptive neuro-fuzzy inference system (ANFIS) and artificial neural network (ANN) models to the data obtained from physical model studies to develop a novel methodology to predict the reflection coefficient (K_r) of seaside perforated semicircular breakwaters under low wave heights, for which no physical model data is available. The prediction was done using the input parameters viz., incident wave height (H_i), wave period (T), center-to-center spacing of perforations (S), diameter of perforations (D), radius of semicircular caisson (R), water depth (d), and semicircular breakwater structure height (h_s). The study shows the prediction below the available data range of wave heights is possible by ANFIS and ANN models. However, the ANFIS performed better with $R^2 = 0.9775$ and the error reduced in comparison with the ANN model with $R^2 = 0.9751$. Study includes conventional data segregation and prediction using ANN and ANFIS.

Keywords Semicircular breakwater · Wave reflection · Below the data range · Artificial neural network · Adaptive neuro-fuzzy inference system

1 Introduction

The semicircular breakwaters have been of interest to researchers in the past few decades owing to its advantages over the conventional breakwaters. The semicircular breakwaters are composite structures precasted with reinforced concrete with a caisson semicircular in shape and a bottom slab. These breakwaters are having better stability, cost-effectiveness, and are suitable for soft soil foundations (Dhinakaran et al. 2009). Although the physical model studies are reliable, they are expensive and the complexities of coastal processes pose as a limitation. The empirical formulae and numerical models can also be used to estimate the wave reflection coefficient and attempt to integrate the complexity of coastal processes. These require more computing facilities, and a need for alternate approach over conventional techniques is found necessary to predict the hydrodynamic characteristics of the semicircular breakwaters. The study by Kudumula and Mutukuru (2013) shows that the reflection coefficient of perforated semicircular breakwater decreased with the increase in percentage of perforations. The wave reflection coefficient being an important aspect in breakwater stability assessment directly impacts the scouring at toe, its study is important (Zanuttigh et al. 2013). The use of soft computing tools (individual and hybrid) for the prediction of breakwater parameters has

Article Highlights
• In cases where the data availability from physical models is limited to know the hydrodynamic characteristics of the semicircular breakwaters soft computing methods prove to be useful tools.
• ANFIS and ANN models were employed to the data obtained from physical model studies to develop a novel methodology to predict the reflection coefficient (K_r) of seaside perforated semicircular breakwaters under low wave heights, for which no physical model data is available.
• The possibility of prediction for below the ranges of trained data is checked by two models and compared with each other and with physical experiment data and the results were found to be good.

✉ Suman Kundapura
sumankundapura@yahoo.com

[1] Department of Applied Mechanics and Hydraulics, National Institute of Technology Karnataka, Surathkal, Mangaluru 575025, India

[2] Department of Civil Engineering, University of Minho, 4710-057 Braga, Portugal

been carried out in the past (Mandal et al. 2009; Jain and Deo 2008; Janardhan et al. 2015; Harish et al. 2015). Yagci et al. (2005) predicted damage ratio of the breakwater using artificial neural network (ANN) in the case of inadequate experimental data sets. The prediction of maximum wave run-up on rubble mound structures was done by using ANN (Erdik et al. 2009). ANN was found better than the regression analysis as the latter has uncertainties in the computation of hydrodynamic characteristics (Deo 2010). Mandal et al. (2009) was successful in using ANNs in the wave transmission coefficient prediction of multilayer floating pipe breakwater. Further, the results were compared with the ANFIS wave transmission coefficient predictions of multilayer floating pipe breakwater. The ANFIS outperformed the ANN in the predictions of wave transmission coefficient predictions of multilayer floating pipe breakwater (Patil et al. 2011). Jabbari and Talebi (2011) used ANN for estimation of scour at the head of vertical wall breakwater. The ANN model performed well when trained by dimensional parameters in comparison with dimensionless parameters. Harish et al. (2015) developed the ANFIS model to predict berm breakwater damage level and compared it with the ANN model, and found that ANFIS is an efficient tool in predicting the damage level of non-reshaped berm breakwater. Reflection coefficient estimation for various coastal structures using ANN was found to be more effective than empirical formulae of the past (Zanuttigh et al. 2013). The wave transmission coefficient of quarter circular breakwater has been predicted by Goyal et al. (2014) using ANN and found the prediction efficient with respect to multiple regression methods. Kim et al. (2014) efficiently predicted the breakwater damage using ANN by introducing the tidal variation as a parameter of prime importance along with the deepwater wave heights. The rock armor stability number was better estimated using ANN with principal component analysis (PCA) (Lee et al. 2015). Raju et al. (2015) attempted to predict the wave reflection coefficient for emerged quarter circular breakwater for beyond the ranges of data (for wave period) using of ANN and ANFIS. For a set of dimensional and dimensionless parameters, ANFIS performed reasonably well with dimensional parameters.

A need was felt to develop the soft computing methods by using “below the data range” of wave height. Hence, the paper objective is to develop the ANN and ANFIS models for estimation of wave reflection coefficient of semicircular breakwaters below the data range of wave heights, used in the experiments. This below the data range of wave heights is not used during training and the robustness of the model is assessed. Also, a comparison between the conventional method of prediction and below the data range prediction is done. The ANN, ANFIS prediction performance is compared with the experimental observations, and also the conventional method of prediction with ANN and ANFIS.

2 Materials and Methods

2.1 Experimental Setup

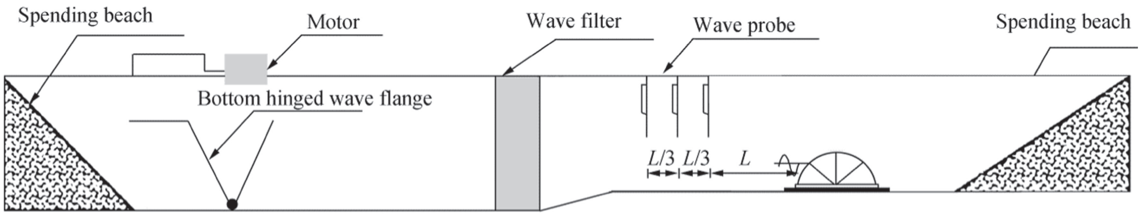
The longitudinal section of the wave flume is presented in Fig. 1a for a regular wave flume as in Marine Structures Laboratory, National Institute of Technology Karnataka, Surathkal, India. The dimensions of the flume are 50 m length, 0.71 m width, and 1.1 m depth. At a distance of 28 m from the wave flap the semicircular breakwater is placed on a rubble mound foundation. The model scale is adopted as 1:30 since it corresponds to the conditions of the Arabian Sea. The incident and reflected wave heights are measured by three capacitance-type wave probes as per the three-probe method proposed by Issacson (1991). In the flume setup, the wave probe is placed at a distance of L from the offshore face of the semicircular breakwater model and the spacing between the successive probes are kept equal to $L/3$, where L is the wavelength. To avoid successive reflection and re-reflection of waves, the waves were forced to attack the semicircular breakwaters in bursts of five waves at each time. The probe measured surface elevation was recorded by the wave recorder, and thereby the measured voltage signals get converted into wave heights and wave periods using the lab wave recorder software provided by EMCON (Environmental Measurements and Controls), Kochi, India. Figure 1b shows the detail of perforations in the semicircular breakwater.

2.2 Data Used

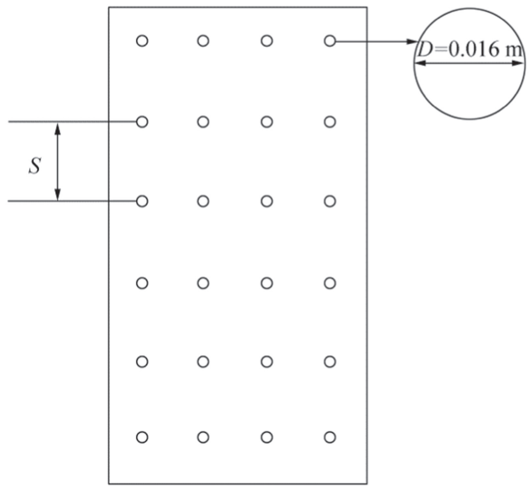
The data obtained from physical model studies of emerged seaside perforated semicircular breakwater (SBW) (Nishanth 2008; Sooraj 2009; Vishal 2010; Sreejith 2015) were fed as inputs to the ANN and ANFIS. The downscaled input parameters used in the experiments to represent the conditions of the Arabian Sea the typical Mangaluru coast are presented in Table 1. Dattatri (1993) found that the mixed-type tides are prevailing in Mangaluru with semi-diurnal components dominating. The tidal variation with respect to mean sea level ± 1.0 m approximately has been reported, with the largest single wave recorded off Mangalore coast was 5.4 m. The predominant wave period during the monsoon is 9–10 s, while longer period waves are experienced in the fair-weather season.

2.3 Data Segregation

The present study is an attempt to predict the wave reflection coefficient for emerged seaside perforated semicircular breakwater for below the data range of wave heights (H_i) which is not trained into ANN and ANFIS. The entire dataset (consisting of 1020 input-output data points) is called global data (GD) and is sorted in the increasing order of wave heights



(a) Experimental Setup of wave flume (Kundapura et al. 2019)



(b) Typical detailing of semicircular breakwater perforations

Fig. 1 Experimental setup of wave flume (Kundapura et al. 2019) and typical detailing of semicircular breakwater perforations

(H_i). The line diagram in Fig. 2 shows the data segregation. The lower ranges of wave height (15%, 0.06 m, 153 data points) and higher ranges of wave height (15%, above 0.15–0.18 m, 153 data points), i.e., a total of 30% is segregated. The inner 70% of the data (consisting 714 data points) is randomized and is called curtailed data (CD) of which 80% (571 data points) is used for training and the remaining 20% (143 data points) is used for testing. This 20% test data is replaced by the available 15% of lower ranges of data and it is named as "below the range." The network is expected to predict the reflection coefficient for those ranges not involved in the training. A prediction for conventional data segregation method is also done (80:20) for the same dataset, i.e., including below the data range values but randomized. Similarly, a prediction for purely higher ranges of wave height (15%, above 0.15 m) not involved in training the network has been carried out by the authors (Kundapura et al. 2019) and it is named as "beyond the range" predictions.

Table 1 Experimental parameter ranges (Kundapura et al. 2019)

Input wave parameters	Data range
Incident wave height, H_i/m	0.06–0.18
Wave period, T/s	1.2–2.6
Depth of water, d/m	0.35, 0.40, 0.45, 0.50
Radius of the semicircular caisson, R/m	0.45, 0.60
Perforation spacing, S/m	0.032, 0.048, 0.064, 0.096, 0.128,
Perforation diameter, D/m	0.012, 0.016
Structure-specific parameters	
SBW structure height h_s/m	0.502, 0.652, 0.730

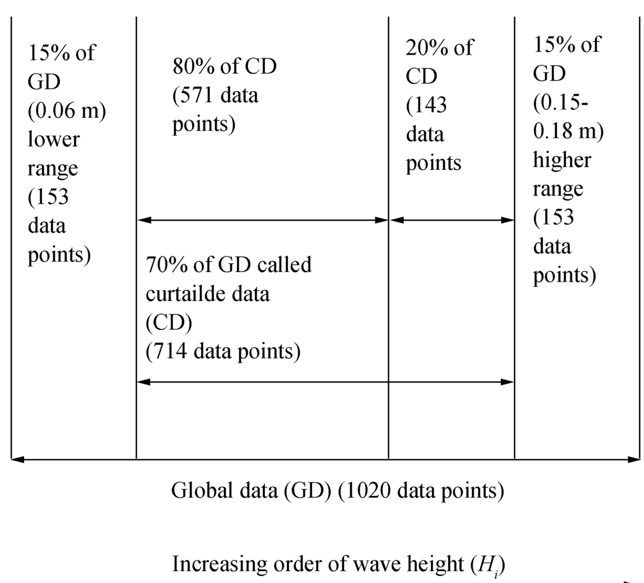


Fig. 2 Typical data segregation procedure for performance prediction of SBW for incident wave height (H_i)

2.4 Development of ANN Model

The basic structure of ANN has an input, a hidden and an output layer. ANN learns from the training of the input-output pairs and regulates the connection weight values in the hidden layer and bias (Azmathulla and Ghani Ab 2011). In the present study, the prediction of K_r for emerged seaside perforated semicircular breakwaters for below the data range of wave heights is done. A feed-forward back-propagation neural network (FFBPNN) as seen in Fig. 3 is used for training the input-output data sets using Levenberg-Marquardt algorithm with transfer functions like “tansig” (hidden layer) and “purelin” (output layer). In the FFBPNN, the error is propagated back in a direction opposite to the way activities propagate in a network.

The FFBPNN is mathematically expressed as follows:

$$Z_k(x) = \sum_{j=i}^{m} W_{kj} \times T_r(y) + b_{ko} \tag{1}$$

$$y_j = \sum_{i=1}^{n} W_{ji} \times x_i + b_{ji} \tag{2}$$

where x is input values from 1 to n, hidden layer neurons is y_j. W_{ji} and W_{kj} are the weights between input and hidden layer and weights between hidden and output layer respectively. Also, b_{ji} and b_{ko} are bias at hidden and output layer respectively. Number of hidden layer nodes is m and $T_r(y)$ is transfer function.

A nonlinear conversion of summed inputs is possible with this transfer function $T_r(y)$. This is done by tansig when employed between the input and hidden nodes and is expressed as follows:

$$T_r(y) = \left[\frac{2}{1+\exp(-2\times y)} - 1\right] \tag{3}$$

where y is summation of input values with weights and biases.

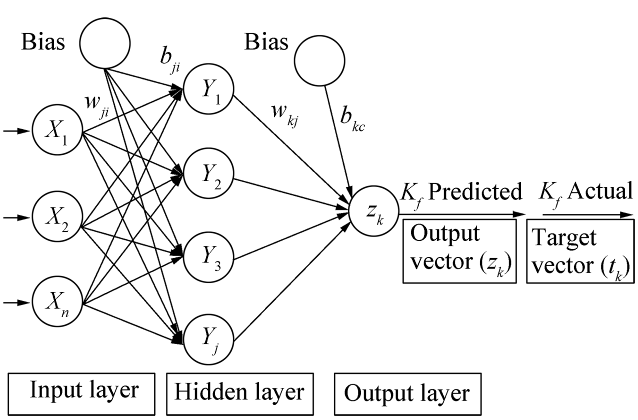


Fig. 3 ANN structure of feedforward backpropagation network (Kundapura et al. 2019)

The transfer function increases the generalization capability of the network and expedites the learning process convergence. In every iteration, the bias adjustments for both hidden and output layer happens. The updated Levenberg-Marquardt algorithm calculates the weights between the hidden and output layer. Purelin a linear transfer function employed between hidden and output layer and is expressed as follows:

$$\text{Purelin}\ (n) = n \tag{4}$$

The objective of the algorithm is to reduce the global error, E, is defined as follows:

$$E = \frac{1}{p}\sum_{p=1}^{P}\left[\sum_{k=1}^{K}\left(d_{kp} - o_{kp}\right)^2\right] \tag{5}$$

wherein, a total number of training patterns are p, desired values of the kth output and the pth pattern are d_{kp} and o_{kp}, the actual value of the kth output and pth pattern.

Literature supports the fact that a single hidden layer is sufficient to solve most of the nonlinear problems and the number of neurons in each layer is determined by a trial and error method (Erdik et al. 2009; Karsoliya 2012; Panchal and Panchal 2014). In the current study, several ANNs were trained altering the number of neurons in the hidden layer dimensional parameters. The network with highest correlation coefficient with the experimental data and least error is chosen. The ANNs predict well if properly trained with the datasets; however, over-training of the network should be avoided. The parameters influencing the reflection coefficient (K_r) are taken as input into the ANN models in dimensional form.

2.5 Development of an Adaptive Neuro-Fuzzy Inference Systems (ANFIS) Model

Jang 1993 originally presented the ANFIS technique. ANFIS constructs fuzzy if-then rules with the required membership functions to generate a functional mapping of input-output pairs (Al-hmouz et al. 2012). ANFIS is a hybrid method combining the neural network learning to tune the parameters of a fuzzy inference system. In ANFIS, the inputs are first converted into fuzzy membership functions and then rule-based learning happens to obtain the output membership functions, by defuzzification, we get the required output (Azmathulla 2011). There are a couple of methods to develop the initial fuzzy membership functions in the literature here that the subtractive clustering method has been employed. The subtractive clustering algorithm is used to automatically generate Gaussian membership functions. A set of fuzzy if-then rules are generated by the algorithm which is equal to the number of cluster centers, each representing the characteristic of the cluster. The algorithm is explained in the subsequent section (Chiu 1994).

2.5.1 Subtractive Clustering

Subtractive clustering evaluates the potential of each data point with their respective dimensions assuming all the data points to have fallen inside a hyperbox (Ratrout 2011). This is worked out on the basis of the density of each point by the formula as follows:

$$\rho_i = \sum_{j=1}^{n} \exp\left(-\frac{\left\|x_i - x_j\right\|^2}{\left(\frac{r_a}{2}\right)^2}\right) \tag{6}$$

where x_i, x_j are the data points with radius of influence r_a and x_{c1} is the first cluster center with highest density. A data point with highest density value is chosen as the first cluster center x_{c1}. A point with highest density occurs when there lie large number of data points in its vicinity (Mohan et al. 2015). Further, the next iteration density measure of each data point x_i is obtained from the Eq. (7) below:

$$\rho_i = \rho_i - \rho_{ci} \sum_{j=1}^{n} \exp\left(-\frac{\left\|x_i - x_j\right\|^2}{\left(\frac{r_a}{2}\right)^2}\right) \tag{7}$$

The process of changing the density of data point and finding the new cluster center is continued until all of the data points are within the range of cluster radius. Thus, the FIS model is obtained and the obtained FIS is based on the first-order Sugeno model which is used to initialize ANFIS model. To map the inputs to the outputs, most commonly adopted FIS are Mamdani inference system and Sugeno inference system, it is found that the Sugeno model is more efficient and compatible (Tiwari et al. 2018).

In a first-order Takegi Sugeno Kang model, if a model has only two inputs the X_1 and X_2 and one output Y and the rule base has only two fuzzy if-then rules then the rules can be represented as follows:

$$\begin{aligned} &\text{Rule 1 : If } X_1 \textit{ is } A_1 \text{ and } X_2 \textit{ is } B_1 \text{ then } F_1 = P_1X_1 + Q_1X_2 + R_1 \\ &\text{Rule 2 : If } X_1 \textit{ is } A_2 \text{ and } X_2 \textit{ is } B_2 \text{ then } F_2 = P_2X_1 + Q_2X_2 + R_2 \end{aligned} \tag{8}$$

where P_1, P_2, Q_1, Q_2 and R_1, R_2 are the linear parameters in the consequent part of the Sugeno fuzzy inference system.

2.5.2 ANFIS Architecture

The architecture of ANFIS has five layers and each layer has several nodes defined by functions.

Layer 1 is a fuzzification layer with each node representing the membership grades of inputs (X_1, X_2) and each nodes output denoted as O_i^j , i.e., the output of ith node in jth layer.

$$\begin{aligned} O_i^1 &= \mu_{A_i}(X_1) \quad i = 1,2 \\ O_i^1 &= \mu_{B_{i-2}}(X_1) \quad i = 3,4 \end{aligned} \tag{9}$$

A_i and B_i are the linguistic labels (like small, medium…) for particular node characterized by membership functions $\mu(X_1)$ and $\mu(X_2)$ respectively. O_i^1 is the membership grade of the fuzzy set. The Gaussian membership function is given by the following:

$$\begin{aligned} \mu_{A_i}(X_1) &= e^{-\left(\frac{(x_1 - b_i)^2}{2a_i^2}\right)} \\ \mu_{A_i}(X_2) &= e^{-\left(\frac{(x_2 - b_i)^2}{2a_i^2}\right)} \end{aligned} \tag{10}$$

where a_i and b_i are the premise parameters of the membership function in the premise part of fuzzy if-then rules that modify the shapes of membership functions.

Layer 2 is the rule layer where each node calculates the rule weight, i.e., the firing strength of the associated rule as in Eq. (11).

$$O_i^2 = w_i = \mu_{A_i}(X_1)\mu_{B_i}(X_2) \tag{11}$$

Layer 3 being the normalization layer represents the ratio of ith rules firing weight to the summation of all rules' firing weight as in Eq. (12).

$$O_i^3 = \overline{w_l} = \frac{w_i}{w_1 + w_2} \tag{12}$$

Layer 4 being the defuzzification layer, Eq. (13) shows the contribution of the ith rule to the total output, where $\overline{w_i}$ is the Layer 3 output and f_i shows the fuzzy if-then rule of the Takagi sugeno type as in Eq. (8):

$$O_i^4 = \overline{w_l} f_i = \overline{w_l}(P_iX_i + Q_iX_i + R_1) \tag{13}$$

Layer 5 is the total output layer, as in Eq. (14) here, the single node computes output (O_i^5 - single output) by summation of all the rules from the previous layer.

$$O_i^5 = \sum_i \overline{w_l} f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \tag{14}$$

Basically, the ANFIS fine tunes the model parameters with gradient descent backpropagation and mean least squares optimization algorithms, which is a hybrid of two techniques. This hybrid technique has a forward pass and backward pass. The mean least square algorithm identifies the consequent parameters in the forward pass (in Layer 4). In every epoch, the sum of the squared difference error (SSE) is propagated backward to update the premise parameters by gradient descent. Once the optimal premise parameters are learned for the generated model, the overall output K_r is a linear combination of consequent parameters (Tiwari et al. 2018; Zhou et al. 2016).

In this study, an ANFIS to generate a Takagi Sugeno Kang-type FIS was done using "genfis2" the subtractive clustering. The genfis2 can handle high dimension of input data unlike "genfis1" which has the curse of dimensionality as a limitation (Hiremath and Patra 2010) and uses Gaussian membership function as the input membership function, whereas output membership function is linear. The parameter influencing the output is cluster radius which could be varied from 0 to 1 for all the variables or the default value of 0.5 can be set (Chiu 1996). The other algorithm parameters are set to default, i.e., squash factor as 1.25, accept ratio as 0.5 and reject ratio as 0.15. In this study, the subtractive clusters construct an FIS with 7 input parameters and 1 output parameter. The subtractive clustering finds the number of membership functions for input and output being equal to the number of clusters (Mohammady 2016). Figure 4 gives the general working of ANFIS model.

The performance of both the ANN and ANFIS models during training and testing is validated using the error metrics like coefficient of determination (R^2), which shows the degree of association between model prediction and actual experiment values. Mean absolute error (MAE) is the average of absolute error between the predicted and actual values, in this case, all the individual differences have equal weight. Root mean square error (RMSE) is the square root of the average of squared differences between the predicted and actual values. The RMSE gives high weight to large errors as the error is squared before taking the average. Both MAE and RMSE should be close to zero, although they vary from zero to infinity.

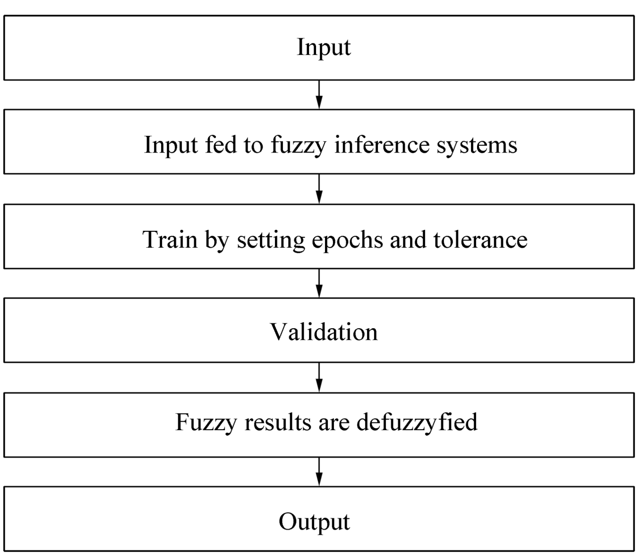


Fig. 4 ANFIS structured flowchart

3 Results and Discussion

The data segregation corresponding to "below the data range" is described in Section 2.3. The conventional method of prediction is done for curtailed data including the data below the range but randomized. The main objective of the paper is to check whether the ANN and ANFIS models can predict for ranges which it is not trained for.

3.1 Training and Testing of ANN Model for Prediction Below the Data Range

The training and testing of ANN model for prediction below the data range were carried out. The best ANN architecture for the available data sets was determined using a trial and error basis with respect to the error metrics. The network was fed with seven inputs and the prediction of one output was done by varying the number of neurons in the one set of the hidden layer. Among the different model architectures tested, the model with one hidden layer consisting of 12 neurons predicted the wave reflection coefficient (K_r) with the least error. The values of the measure of error is presented in Table 2. The coefficient of determination (R^2) was found to be 0.9751; however, the error is highest of all the four cases. Upon comparison of MAE and RMSE during testing, RMSE was found to be slightly higher. This indicates that there is a variance of the frequency distribution of error magnitudes. Figure 5 shows the comparison of model prediction and actual values of the reflection coefficient for various data points for below the data range testing using ANN.

3.2 Training and Testing of ANN Model for Conventional Method of Data Segregation

The curtailed data including below the data range values is randomized and divided into two sets (80% for training and 20% for testing). This follows the typical conventional method of data segregation and prediction. As mentioned earlier, seven inputs were fed into the model and with the expected output being reflection coefficient (K_r). For this case, the best architecture was found to be 12 neurons in the hidden layer. Table 3 shows the error measure for training and testing by the

Table 2 ANN and ANFIS model statistics obtained for K_r prediction below the data range

Error measure	ANN		ANFIS	
	Training	Testing	Training	Testing
R^2	0.9799	0.9751	0.9848	0.9775
MAE	0.0208	0.0428	0.0161	0.0235
RMSE	0.0273	0.0517	0.0228	0.0303

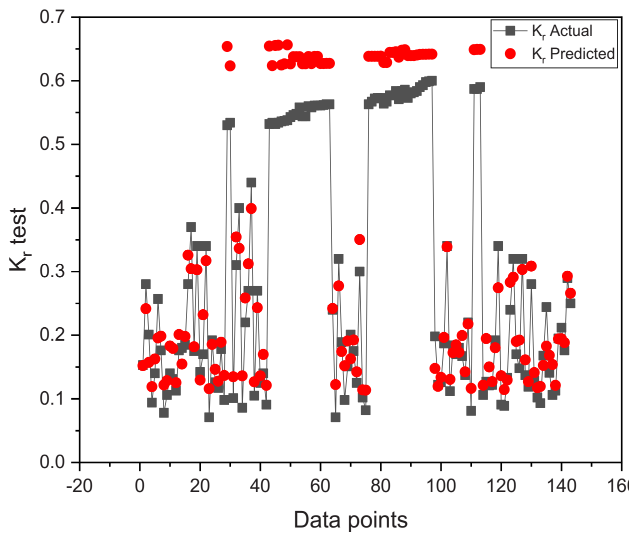


Fig. 5 ANN model prediction and actual values of K_r vs. data points for testing of below the data range

conventional method of prediction. $R^2 = 0.9808$, in the case of testing, is found closer to 1, indicating better fit of the model compared to ANN testing of below the data range. ANN model comparison of MAE and RMSE of conventional method of data segregation to testing of below the data range shows error reduction in the conventional method of data segregation. Figure 6 shows the comparison of model prediction and actual values of the reflection coefficient for testing by the conventional method of data segregation using ANN.

3.3 Training and Testing Results of ANFIS Model for Prediction Below the Data Range

Training data loaded to generate input membership function consisted of seven inputs and one output data. In this case, the "genfis2" is used for training the data with a step size of 0.1 and number of epochs 20. The radius of influence was varied from 0 to 1 and found that the RMSE to be least when radius was 0.3 for all 7 inputs, and the genfis2, makes Gaussian membership functions for each input. Training and testing performance of ANFIS model for below the data range is validated by error measure as shown in Table 2 and also a quantitative comparison between ANN and ANFIS model was done. The ANFIS model gave slightly better results with

Table 3 ANN and ANFIS model statistics obtained for K_r prediction for conventional method of data segregation

Error measure	ANN		ANFIS	
	Training	Testing	Training	Testing
R^2	0.9826	0.9808	0.9868	0.9844
MAE	0.0189	0.0195	0.0153	0.0170
RMSE	0.0252	0.0265	0.0213	0.0233

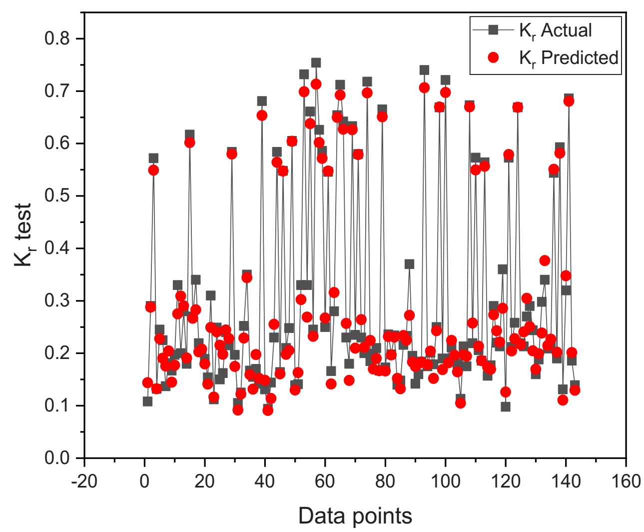


Fig. 6 ANN model prediction and actual values of K_r vs. data points for testing of conventional method of data segregation

higher $R^2 = 0.9775$, as seen in Table 2 with respect to ANN testing. The MAE and RMSE value was found to be lower than MAE and RMSE in the case of ANN model. The actual reflection coefficient and the ANFIS model prediction are plotted for various data points in Fig. 7. The input-output network of the ANFIS model for below the data range prediction is represented in the Fig. 8.

3.4 Training and Testing Results of ANFIS Model for the Conventional Method of Data Segregation

Training data with seven inputs and one output data was loaded to generate input membership function. In the study, genfis2 is used for training the data with a step size of 0.1 and number of epochs 20 with RMSE least when radius was set to 0.9. Training and testing performance of ANFIS model

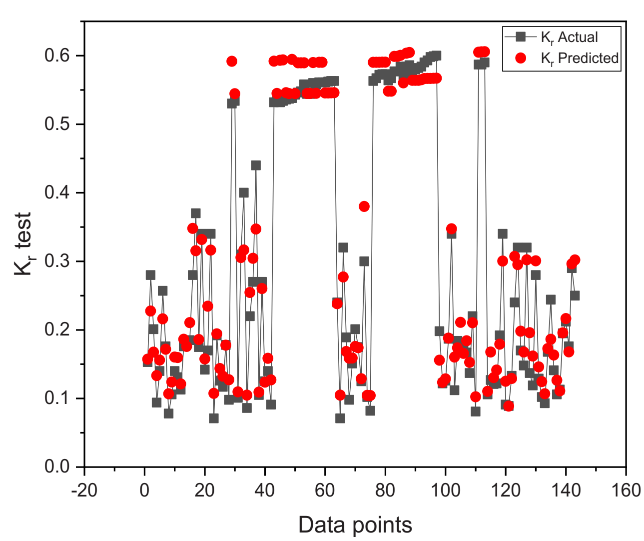


Fig. 7 ANFIS model prediction and actual values of K_r vs. data points for testing of below the data range

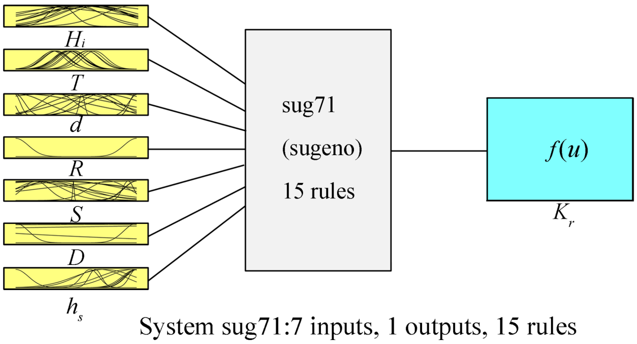


Fig. 8 ANFIS model input and output for below the data range prediction

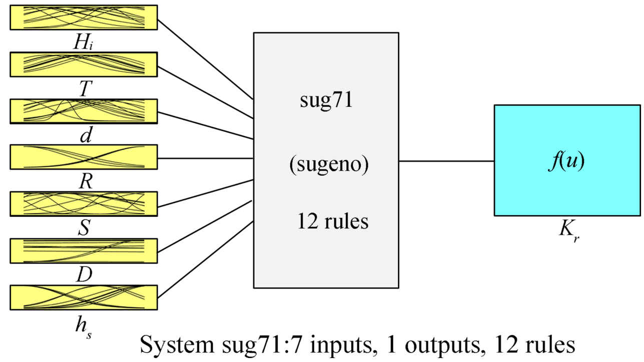


Fig. 10 ANFIS model input and output for conventional method of data segregation

for the conventional method of prediction is validated by error measure as shown in Table 3 and found that the predictions made by ANFIS were reasonably good. Also, a quantitative comparison between ANN and ANFIS model can be seen in Table 3 which shows that ANFIS results slightly improved over the ANN result in terms of coefficient of determination $R^2 = 0.9844$. The MAE and RMSE values was found to be the lowest in this particular case of ANFIS model testing compared to that of ANN testing errors indicating the reduction of errors in the conventional method of prediction using ANFIS. In all the four cases, RMSE was found to be slightly higher than MAE and indicates that there is variance in the frequency distribution of error magnitudes. The actual reflection coefficient and the ANFIS model prediction are plotted for various data points in Fig. 9. The input-output network of the ANFIS model for prediction of conventional method of data segregation is represented in the Fig. 10.

The study aims to predict the reflection coefficient (K_r) when the wave height was less than the minimum value of wave height used in the experiments. For instance, total wave heights available from the experiments ranged from 0.06 to 0.18 m. To check the possible prediction without conducting, the physical model study for some lower ranges of wave height was the priority. The validation of such below the data range prediction was done for $H_i = 0.06$ m lesser than the minimum value of H_i in the data range used for training. The prediction was compared with the available experimental values and the error metrics was checked. The predictions made for below the data range and conventional method of data segregation are close in terms of percentage of coefficient of determination, i.e., 97% and 98% respectively and hence acceptable. The study shows that below the data range prediction up to 25% is possible. Similarly, a study for the higher ranges of wave height prediction called beyond the range predictions have been conducted and the results are encouraging with the ANFIS model outperforming the ANN model with percentage of coefficient of determination 96% (Kundapura et al. 2019).

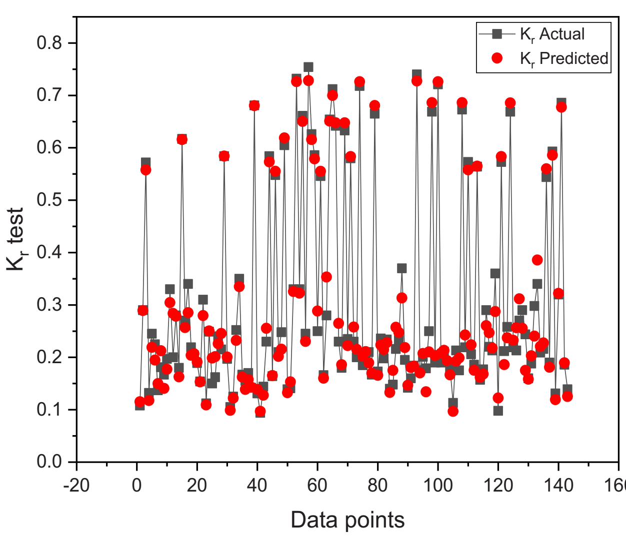


Fig. 9 ANFIS model prediction and actual values of K_r vs. data points for testing of conventional method of data segregation

4 Conclusions

The current approach of data segregation, i.e., below the data range approach is first of its kind wherein certain ranges of data is excluded from training the ANN and ANFIS models. The developed models are expected to emulate the wave reflection coefficient of emerged seaside perforated semicircular breakwater.

1) The predictions for "below the data range" approach is possible up to 25% and found to be close to the predictions by conventional method of data segregation approach using ANN and ANFIS models.
2) The study found that below the data range predictions made by the ANFIS model outperformed the ANN model.
3) This study could be useful to the coastal engineers in the prediction of the reflection coefficient of emerged seaside perforated semicircular breakwaters, for lower wave heights for which no physical model data is available.

References

Al-hmouz A, Shen J, Al-Hmouz R, Yan J (2012) Modeling and simulation of an adaptive neuro-fuzzy inference system (ANFIS) for mobile learning. IEEE Trans Learn Technol 5(3):226–237. https://doi.org/10.1109/TLT.2011.36

Azmathulla HMD, Ghani Ab A (2011) ANFIS-based approach for predicting the scour depth at culvert outlets. J Pipeline Syst Eng Pract 2(1):35–40. https://doi.org/10.1061/(ASCE)PS.1949-1204.0000066

Chiu S (1994) Fuzzy model identification based on cluster estimation. J Intell Fuzzy Syst 2(3):267–278

Chiu S (1996) Method and software for extracting fuzzy classification rules by subtractive clustering. *North American Fuzzy Information Processing Society Conf.* (NAFIPS'96), Berkeley, 19–22. https://doi.org/10.1109/NAFIPS.1996.534778

Dattatri J (1993) Waves off Mangalore harbour-west coast of India. J Waterw Port Coast Ocean Eng, ASCE 99(2):39–57

Deo MC (2010) Artificial neural networks in coastal and ocean engineering. Indian J Mar Sci 39(4):589–596

Dhinakaran G, Sundar V, Sundaravadivelu R, Graw KU (2009) Effect of perforations and rubble mound height on wave transformation characteristics of surface piercing semicircular breakwaters. Ocean Eng 36:1182–1198. https://doi.org/10.1016/j.oceaneng.2009.08.005

Erdik T, Savci ME, Sen Z (2009) Artificial neural network for predicting maximum wave runup on rubble mound structures. Expert Syst Appl 36(3):6403–6408. https://doi.org/10.1016/j.eswa.2008.07.049

Goyal R, Singh K, Hegde AV (2014) Quarter circular breakwater : prediction and artificial neural network. Mar Technol Soc J 48:1–7. https://doi.org/10.4031/MTSJ.48.1.7

Harish N, Mandal S, Rao S, Patil SG (2015) Particle Swarm Optimization based support vector machine for damage level prediction of non-reshaped berm breakwater. Appl Soft Comput 27:313–321. https://doi.org/10.1016/j.asoc.2014.10.041

Hiremath S, Patra SK (2010) Transmission rate prediction for cognitive radio using adaptive neural fuzzy inference system. Proc IEEE Ind Inf Syst Conf Mangalore, India, p 92–97

Issacson M (1991) Measurement of regular wave reflection. J Waterway Port Coastal Ocean Eng ASCE 117(6):553–569

Jabbari E, Talebi O (2011) Using artificial neural networks for estimation of scour at the head of vertical wall breakwater. J Coast Res SI 64(ICS2011):521–526

Jain P, Deo MC (2008) Artificial neural networks for coastal and ocean studies. 12th Int Conf Int Assoc Comput Methods Adv Geomech, Goa, India, p 1655–1663

Janardhan P, Harish N, Rao S, Shirlal KG (2015) Performance of variable selection method for the damage level prediction of reshaped berm breakwater. Aquat Procedia 4:302–307. https://doi.org/10.1016/j.aqpro.2015.02.041

Jang JR (1993) *ANFIS :* adaptive-network-based fuzzy inference system. IEEE Trans Syst Man Cybern 23(3):665–685

Karsoliya S (2012) Approximating number of hidden layer neurons in multiple hidden layer BPNN. Architecture Int J Eng Trends Technol 6:714–717

Kim DH, Kim YJ, Hur DS (2014) Artificial neural network based breakwater damage estimation considering tidal level variation. Ocean Eng 87:185–190. https://doi.org/10.1016/j.oceaneng.2014.06.001

Kudumula SR, Mutukuru MRG (2013) Experimental studies on low crested rubble mound, semicircular breakwaters and vertical wall system. Int J Ocean Climate Syst 4(3):213–226 http://journals.sagepub.com/doi/abs/10.1260/1759-3131.4.3.213

Kundapura S, Hegde AV, Wazerkar AV (2019) Beyond the data range approach to soft compute the reflection coefficient for emerged perforated semicircular breakwater, In: Murali K, Sriram V, Samad A, Saha N (eds) Proceedings of the Fourth International Conference in Ocean Engineering (ICOE2018). Lecture Notes in Civil Engineering, vol 23. Springer, p 281–292

Lee A, Kim SE, Suh KD (2015) Estimation of stability number of rock armor using artificial neural network combined with principal component analysis. Procedia Eng 116:149–154. https://doi.org/10.1016/j.proeng.2015.08.276

Mandal S, Patil S, Hegde AV (2009) Wave transmission prediction of multilayer floating breakwater using neural network. Proc 3rd International Conference in Ocean Engineering, ICOE-2009, IIT Madras, Chennai, p 574–585

Mohammady S (2016) Optimization of adaptive neuro-fuzzy inference system based urban growth model. City Territ Archit 3(10):1–15. https://doi.org/10.1186/s40410-016-0039-8

Mohan RU, Sood YR, Jarial RK (2015) Subtractive clustering fuzzy expert system for engineering applications. Procedia Comput Sci 48(C):77–83. https://doi.org/10.1016/j.procs.2015.04.153

Nishanth N (2008) Hydrodynamic performance characteristics of semicircular breakwaters. Master Thesis, National Insitutute of Technology Karnataka, Surathkal, Mangaluru, India, Appendix 5–13

Panchal FS, Panchal M (2014) Review on methods of selecting number of hidden nodes in artificial neural network. Int J Comput Sci Mob Comput 3(11):455–464

Patil SG, Mandal S, Hegde AV, Alavandar S (2011) Neuro-fuzzy based approach for wave transmission prediction of horizontally interlaced multilayer moored floating pipe breakwater. Ocean Eng 38:186–196. https://doi.org/10.1016/j.oceaneng.2010.10.009

Raju B, Hegde AV, Chandrashekar O (2015) Computational intelligence on hydrodynamic performance characteristics of emerged perforated quarter circle breakwater. Procedia Eng 116:118–124. https://doi.org/10.1016/j.proeng.2015.08.272

Ratrout NT (2011) Subtractive clustering-based K -means technique for determining optimum time-of-day breakpoints. J Comput Civ Eng 25(5):380–387. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000099

Sooraj M (2009) Sliding stability and hydrodynamic performance of emerged semicircular breakwater, Master Thesis, NITK, Surathkal, Mangaluru, India, AppendixII, 96–102

Sreejith KU (2015) Sliding stability and hydrodynamic performance of emerged semicircular breakwater, Master Thesis, NITK, Surathkal, Mangaluru, India, Appendix 95–104

Tiwari S, Babbar R, Kaur G (2018) Performance evaluation of two ANFIS models for predicting water quality index of river Satluj (India). Adv Civil Eng 2018:8971079. https://doi.org/10.1155/2018/8971079

Vishal K (2010) Hydrodynamic performance characteristics of one side and two side perforated semicircular breakwater. Master Thesis, NITK, Surathkal, Mangaluru, India, Appendix I, 79–83

Yagci O, Mercan DE, Cigizoglu HK, Kabdasli MS (2005) Artificial intelligence methods in breakwater damage ratio estimation. Ocean Eng 32:2088–2106. https://doi.org/10.1016/j.oceaneng.2005.03.004

Zanuttigh B, Mizar S, Briganti R (2013) A neural network for the prediction of wave reflection from coastal and harbor structures. Coast Eng 80:49–67. https://doi.org/10.1016/j.coastaleng.2013.05.004

Zhou Q, Wang F, Zhu F (2016) Estimation of compressive strength of hollow concrete masonry prisms using artificial neural networks and adaptive neuro-fuzzy inference systems. Constr Build Mater 125:417–426. https://doi.org/10.1016/j.conbuildmat.2016.08.064