Inverse modeling to derive wind parameters from wave measurements

S.B. Charhate, M.C. Deo, S.N. Londhe
Department of Civil Engineering, Indian Institute of Technology Bombay, Powai, Mumbai 400076, India

Abstract

The problem of deriving wind parameters from measured waves is discussed in this paper. Such a need reportedly arises in the field when the wind sensor attached to a wave rider buoy at high elevation from the sea level gets disconnected during rough weather, or otherwise needs repairs. This task is viewed as an inverse modeling approach as against the direct and common one of evaluating the wind-wave relationship. Two purely nonlinear approaches of soft computing, namely genetic programming (GP) and artificial neural network (ANN) have been used. The study is oriented towards measurements made at five different offshore locations in the Arabian Sea and around the western Indian coastline. It is found that although the results of both soft approaches rival each other, GP has a tendency to produce more accurate results than the adopted ANN. It was also noticed that the equation-based GP model could be equally useful as the one based on computer programs, and hence for the sake of simplicity in implementation, the former can be adopted. In case the entire wave rider buoy does not function for some period, a common regional GP model prescribed in this work can still produce the desired wind parameters with the help of wave observations available from anywhere in the region. A graphical user interface is developed that puts the derived models to their actual use in the field.

1. Introduction

A large amount of ocean data is routinely collected with the help of floating wave rider buoys. Such wave buoys are fitted with both wave and wind sensors. Sometimes during rough weather, the wind anemometer usually located at the height of 3 m above mean sea level (MSL) gets disconnected, and the wind data collection stops, but the floating wave sensor continues to work. This problem has been experienced vividly by India’s National Institute of Ocean Technology which had deployed a large number of data collection buoys around the country’s coastline. The information of short interval wind speed and direction in a continuous and uninterrupted manner and also in a real time mode is needed for a variety of operation related work in the ocean, such as planning for coastal aircraft movements, recreational works, installation and construction activities, and also for output power prediction in case of wind turbines. Further, collection of historical wind data facilitates analysis and design of structures at a given location. In such cases it becomes necessary to carry out an inverse modeling to retrieve wind speed and direction from the measured wave height and period. It is therefore desirable to prescribe a method to derive the wind speed, along with its direction from the measured values of significant wave heights and periods. This work could be seen as inverse modeling, since it needs to be handled in an opposite manner to normal wind-wave modeling. Currently, there are no known methods to do so. This study therefore addresses this issue. Considering the complexity and non-linearity of the wind-wave relationship, it was decided to tackle this problem through two soft computing tools that are purely non-linear in nature, namely, genetic programming (GP) and artificial neural networks (ANN).

The attempts to derive wind speed using soft computing are limited. Oztopal [17], and Mehmet et al. [15] employed an artificial neural network (ANN) to carry out regional wind speed estimations at certain locations in Turkey, based on observations at nearby locations. These studies involved speed estimation only. Attempts to derive the direction of wind using ANN are sparse, a few noticeable among them are, Thiria et al. [18], who applied the ANN approach to obtain wind direction using simulated data as well as a spatial input context, Cornford et al. [3], who used ANN based techniques to estimate wind vectors form scatterometer data. Studies dealing wind speed and direction evaluation based on GP have however not been reported so far.

Applications of GP in civil engineering related to water flow started around 5 years ago. The tool of GP has been used for a variety of purposes, like pattern recognition, classification and regression. Unlike the other soft computing tool of ANN, the GP applications are restricted to relatively fewer areas in hydraulics and water resources, e.g., [4,22,16]. Some of these authors have presented a comparison with other models. Drecourt [4] reported that GP handles peak flow better, while ANN takes care of the noise efficiently. Muttil and Liong [16] found the performance of
GP marginally better than ANN. The applications of GP in ocean engineering are very sparse, and these include evaluation of ocean component concentration from sunlight reflectance or luminance values ([5]), coastal current prediction [2] and in-filling of missing wave data [20,8].

Some of the studies in recent past had worked out wave parameters and sea levels in an inverse manner, by establishing spatial correlations (unlike the present causal relationship) in between similar variables and based on ANN. These include Makarynskyy [12], who obtained wave conditions at a given location, based on the same at a distant site; Makasrynskyy et al. [14] who retrieved and predicted hourly tidal levels at one station given their values at another one, and Makarynska and Makarynskyy [13] who did a similar work in case of wave parameters. The use of fuzzy systems and adaptive neuro-fuzzy inference system (ANFIS) in addition to ANN was recently made by Mahjoobi et al. [11] to hindcast waves from wind through causal mapping.

2. The database used

In this study, the time series measurements of various ocean parameters made at different locations, namely: DS1, SW2, SW3, OB3 and DS7 (Fig. 1), in the Arabian Sea were considered. The observations were made under the national data buoy program of National Institute of Ocean Technology (NIOT) at Chennai, India. The parameters were the significant wave height, $H_s$, average zero cross wave period, $T_z$, average wave period, $T_m$, wind speed, $W_s$ and wind direction, $\theta$. The time interval of sampling was of 3 h. The location SW3 is in shallow water and close to the shore whereas all other sites are in deep water, and far away from the coast. Station DS7 is a fairly open location, exposed both to the Arabian Sea and the Indian Ocean. The period of data collection ranged from March 2004 to Dec. 2006, in general. The observations at all the stations had a few gaps, and in the present study these were filled by spatial correlations with measurements of adjacent stations.

3. Problem formulation and methodology

3.1. Artificial neural network

An ANN consists of an interconnection of computational elements or neurons (Fig. 2), each of which combines the input, determines its strength by comparing the combination with a bias or alternatively passing it through a non-linear transfer function, and fires out the result in proportion to such strength. Mathematically,

$$O = \frac{1}{1 + e^{-S}}$$

where,

$$S = (x_1w_1 + x_2w_2 + x_3w_3 + \cdots) + \theta$$

in which, $O =$ output from a neuron; $x_1, x_2, \ldots =$ input values; $w_1, w_2, \ldots =$ weights along the linkages connecting two neurons and they indicate the strength of connections; $\theta =$ bias value. Many applications of ANN in ocean engineering have so far used the feed forward type of the network as against the feedback or recurrent one. A feed forward multi-layer network would consist of an input layer, one or more hidden layers and an output layer of neurons as shown in the Fig. 2 referred to earlier. Before it is actually applied to solve a problem, the network is calibrated or trained using mathematical algorithms. Details of such training methods can be seen in text books like Kosko [9], Wasserman [21] and Wu [23]. A review of important ANN applications in ocean engineering recently reported, can be seen in [6].

3.2. Genetic programming

Genetic programming mimics the process of evolution occurring in nature, according to which the species continue as per the principle of ‘survival of the fittest’. Conceptually, it is similar to the genetic algorithm (GA), but unlike GA, its solution is a computer program or an equation as against a set of numbers in the GA. Koza [10] explains various concepts related to GP. In GP, a random population of individuals (equations or computer programs) is created, the fitness of individuals is evaluated, and then the ‘parents’ are selected out of these individuals. The parents are then made to yield ‘offspring’ by following the process of reproduction, mutation and cross-over. The creation of offspring continues (in an iterative manner) till a specified number of offspring in a generation are produced and further until another specified number of generations is created. The resulting offspring at the end of all this process (an equation or a computer program) is the solution of the problem. The GP thus transforms one population of individuals into another one in an iterative manner by following the natural genetic operations like reproduction, mutation and cross-over.

3.3. The wave parameters

Samples of sea surface elevation were collected at the rate of 1 Hz, upto 17 min at a time, once in every three hours. For every such short term wave record the wave spectrum, indicating the variation of spectral density function $S(f)$ for each wave frequency,

Fig. 1. Data observation stations in Arabian Sea, India.

Fig. 2. The ANN used.
$f$, was derived. Values of $H_s$, $T_z$ and $T_m$ were thereafter obtained using equations:

$$H_s = 4\sqrt{m_0}$$

$$T_z = \sqrt{\frac{m_0}{m_2}}$$

$$T_m = \frac{m_0}{m_1}$$

where

$$m_n = \int_0^\infty f^n S(f).df \quad \text{where} \ n = 0, 1, 2.$$  

In which $m_n = \text{moments of the spectrum about the origin and } S(f) = \text{spectral density function of the sea surface elevation for wave frequency } f$.

In order to know the exact input parameters to be used in the models, a sensitivity analysis was carried out in which the input variables were added one by one to check their effectiveness in producing the outcome. This also enabled us to know the effect of unavailability of the remaining parameters. Over the course of trials it was observed that both GP and ANN perform better with three input parameters, namely, $H_s$, $T_z$ and $T_m$. In the case of ANN the correlation among input variables was also studied through a Hinton diagram. Inclusion of the parameter $T_m$ in addition to $T_z$ was found to be useful, although like $T_z$, it is also a measure of the average value of wave period. This could be due to the additional flexibility in modeling it brings in. Further, it is to be noted that lower spectral moments such as $m_1$ are more stable than higher ones such as $m_2$ [1], and hence the use of $T_m$ might have resulted in reducing the effect of uncertainty in establishing average wave periods.

4. Estimation of wind speed

4.1. Station DS1

Station DS1 is in deep water, and it is located far away from the shoreline. The period of data collection varied from 09-04-2004 to 26-07-2005 and from 01-05-2006 to 30-09-2006; out of which observations for the first 12 months were used for training and those for the remaining 4 months (in the first set) were used for testing the network. Data for the year 2006 were additionally used to carry out the testing. The type of network used was feed forward. (Fig. 2). This was trained using a variety of learning schemes namely, back propagation, cascade correlation and variants of conjugate gradient. Out of these, the best training method turned out to be the ordinary back propagation. Figs. 3 and 4 show the test results along with the performance measures of correlation coefficient, $R$, root mean square error, RMSE and mean absolute error, MAE. The ANN had three input nodes, namely significant wave height ($H_s$), zero cross wave period ($T_z$) and average wave period ($T_m$), while the output node was one, and this belonged to the estimated wind speed. Using the same variables for input and output as in the ANN and a similar division of training and testing data, a GP (program based) model was also built. Figs. 5 and 6 show the testing performance of this GP model. These time series and scattered plots, together with the performance measures indicated on them, show that both GP and ANN were able to learn the underlying complex inverse relationship satisfactorily, with GP showing a tendency to produce slightly better estimates of wind speed. This indicates an efficient handling of the underlying non-linearity by GP. The algorithm of GP involves less mathematical rigidity and control than ANN and hence more flexibility in fitting and a reduced possibility of encountering problems, such as over-fitting and local minima. The testing exercise as above pertained to a duration of 4 months in the year 2005. Another testing exercise
was done for measurements of the year 2006 by GP (equation based), as opposed to the earlier GP (program based) model and this further confirmed the above findings. The outcome of GP in the form of a mathematical equation was as given below:

$$W_s = \left(\exp^{H_s - T_m} - \exp^{0.054}\right) \times \left(0.17 + T_z + H_s\right) + \log(H_s + H_s)$$  \hspace{1cm} \text{(7)}$$

Figs. 7 and 8 along with the performance measures indicated on these figures, showed that the equation based GP model performs almost identically to the program based GP model and hence the user may select either of them.

The high level of accuracy attained in this modeling is really noteworthy, considering the highly varying and arbitrary nature of the wind.

Innumerable studies in the past (e.g., [19,7]) have shown that statistical regression based models do not work as effectively as the ANN’s and hence in this study the comparison of performance with regression schemes was not again attempted.

4.2. Station: SW2

The Station SW2 is located in deep water (depth 2250 m), and at a distance of 53 Nm from the coast. The fluctuations in input values recorded here were found to be moderate compared to those over the locations SW3 (shallow water) and DS7 (open). The observation period was from Jan. 1, 2005 to July 31, 2006. The ANN and GP models were developed from the observations ranging from Jan. 1, 2005 to Jan. 14, 2006 while remaining measurements from Jan. 15, 2006 to July 31, 2006 were used for model testing. The test results in terms of time series and scattered plots along with performance statistics are shown in Table 1. The GP (program based) model can be seen to have an edge over the ANN. Eq. (8) is the result of the calibrated GP.

$$W_s = \frac{\left(\{\left(\{H_s / 0.054\}^2 + (1.91 + H_s)\}/T_m\} + H_s\right) / T_m}{7H_s + 3T_z + 3T_m}$$  \hspace{1cm} \text{(8)}$$

This model also showed good performance measures of $R = 0.90$, RMSE = 1.26 m/s, and MAE = 0.92 m/s that were similar to the program based GP model.

4.3. Station SW3

The Station SW3 is in shallow water, and it is located near the coastline. The observation period varied from March 2004 to June 2005. The models were trained on the observations made from March 2004 to Jan. 2005, while remaining measurements from February 2005 to June 2005 were used for validation purposes. The test results for ANN and GP (program based) models and their performances are shown in Table 1. The calibrated equation based GP model is given in Eq. (9):

$$W_s = \left\{\exp\left(\frac{\log(T_m) + \log(T_m)}{T_m}\right)\right\} + \left\{\log(H_s + T_z) / \log(T_m) \left[\left(\exp^{7.62} / (T_m/H_s) + (T_m^2/H_s)\right)\right](7.62 + T_z / \log(T_m))\right\}$$ \hspace{1cm} \text{(9)}$$

This equation produced results that were more or less of similar accuracy to that of the program based GP model. The corresponding error measures (testing) are: $R = 0.84$, RMSE =
1.62 m/s and MAE = 1.23 m/s. As expected, the accuracy level achieved while estimating the wind speed at this coastal location was lower than the one in earlier open and deep location DS1 and SW2.

4.4. Station OB3

The station OB3 is located away (200 Nm) from the shore line, and it is in moderate water depth (1750 m). The observation period varied from April 2004 to Sept. 2005. The ANN and GP models were built on the observations ranging from April 2004 to April 2005 while remaining measurements from May 2005 to Sept. 2005 were used for model testing. An excellent prediction performance by both ANN and GP was seen, with an edge of GP over the ANN. The performance measures indicated in Table 1 further confirmed the usefulness of these approaches. The resulting equation output is as given below

$$W_s = \left\{ \left[ 5H_s + 0.32 \right] + \exp\left( \frac{T_z}{T_m} \right) \right\} \left( \frac{T_m}{\exp\left( \frac{T_z}{T_m} \right) + 0.42} \right)^{H_s}$$

The error statistics for this model during the testing were: $R = 0.94$, RMSE = 1.30 m/s, MAE = 1.01 m/s.

4.5. Station DS7

The Station DS7 is in deep sea located towards south, and open to the Arabian Sea as well as the Bay of Bengal. The models were developed, based on the measurements of Jan. 2005 to March 2006 and validated for the observations of April 2006 to Dec. 2006. The test results are shown in Table 1. Like the previous cases here, both ANN and GP also performed well, with GP producing slightly better predictions. The equation as an outcome of GP is given in Box I.

This equation produced the error statistics during testing as $R = 0.86$, RMSE = 1.21 m/s, MAE = 0.96 m/s. The accuracy levels realized at this location were relatively low compared to other deep sea locations, and this could be due to the fact that it is open to both Arabian Sea as well as to Bay of Bengal, and as a result, very large variations in the input, degrading the model’s performance can only be expected.

5. Prediction of direction in addition to speed

The earlier section dealt with prediction of the magnitude of wind. Knowledge of the wind speed obtained in this way can be utilized for work such as derivation of statistical distributions, and estimation of design wind speed parameters. However, there are many applications where information of both speed and direction is required, such as drawing wind rose diagrams and deriving wave height and period values.

As reported in the preceding section, the GP technique was found to be very satisfactory for wind speed predictions at all the locations considered, namely, deep, shallow, open location and near shore. Hence to avoid repetition, the analysis reported below caters for the use of GP only.

In the current work, GP models were developed at each location of the Arabian Sea under consideration, based on the same datasets used in previous wind speed modeling. The input parameters were kept the same as $H_s$, $T_z$ and $T_m$. The outcome of wind speed and direction was obtained in terms of the speed component along the North–South direction ($u$) and the same along the East–West direction ($v$). Separate GP models were developed to obtain the $u$- and $v$-wind components respectively. The values of $u$–$v$ components so obtained, were used later on to predict the wind and direction as:

$$U = \sqrt{u^2 + v^2}$$

$$\theta = \tan^{-1} \frac{v}{u}.$$
The performance of models at each location is shown in Table 2. The second main column of Table 2 shows results of the study carried out in the previous section, using GP models while the third column indicates the performance of GP models in the study being discussed in this section, i.e., after combining the $u$-$v$ components. The resolution of wind speed into components may add to modeling complexity, due to the requirement of evaluating two instead of one parameter. This may have reduced the performance of the speed prediction model. The observed and predicted directions were from 0° to 360° with respect to the North. The error criterion of correlation coefficient ($R$) was seen to be within 0.72–0.88 while that of mean absolute errors in direction ranged from ±12 to ±28 degrees for deep and shallow water. It may be noted that, compared to the wind speed, the direction prediction is very difficult, due to its highly random rotation from 0° to 360°.

Figs. 9 and 10 show as an example the time series and scattered plots for direction predictions at DS1. It was observed that though the wind direction was very much more difficult to predict, the GP has produced a satisfactory forecast. The performance of models at the deep location DS7 is understandably less compared to that at location SW3 that is in shallow water, and near to the shore, where land proximity induced effect on wind direction could be predominant. Also DS7 is a highly open location, where frequent changes in wind pattern are expected.

6. Development of common regional GP model

The earlier sections dealt with the prediction of wind speed and its direction, based on the measured parameters of $H_s$, $T_z$ and $T_m$. The ANN and GP models were built at locations DS1, SW2, SW3, OB3 and DS7 individually and thereafter two separate models of GP were also developed in order to determine the wind speed and its direction. Calibrated GP models in the form of equations were specified at each of the sites as in Section 5.

An attempt was thereafter made to see if a common regional GP model applicable for the entire region encompassing all above

\[
W_s = \exp \left( \left[ \left( \log (H_s + e^{\log(T_m) + T_m}) + \log (3H_s) \right) + \left( \frac{\left[\log(H_s/T_m) + \log (T_z) \right]}{e^{(0.57)}} \right) + H_s \right] \right)
\]
Table 3
Performance comparison of the GP-evolved equations

<table>
<thead>
<tr>
<th>Station ID</th>
<th>Performance criteria</th>
<th>GP-E output from individual equation</th>
<th>Using GP-E of DS1</th>
<th>Using GP-E of SW2</th>
<th>Using GP-E of SW3</th>
<th>Using GP-E of OB3</th>
<th>Using GP-E of DS7</th>
</tr>
</thead>
<tbody>
<tr>
<td>DS1</td>
<td>R</td>
<td>0.94</td>
<td>0.92</td>
<td>0.94</td>
<td>0.94</td>
<td>0.92</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE (m/s)</td>
<td>1.26</td>
<td>2.08</td>
<td>1.10</td>
<td>1.35</td>
<td>4.70</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MAE (m/s)</td>
<td>0.96</td>
<td>1.78</td>
<td>0.82</td>
<td>1.02</td>
<td>4.28</td>
<td></td>
</tr>
<tr>
<td>SW2</td>
<td>R</td>
<td>0.90</td>
<td>0.89</td>
<td>0.88</td>
<td>0.89</td>
<td>0.85</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>1.26</td>
<td>1.30</td>
<td>1.46</td>
<td>1.37</td>
<td>2.96</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MAE</td>
<td>0.92</td>
<td>1.03</td>
<td>1.01</td>
<td>1.01</td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td>SW3</td>
<td>R</td>
<td>0.84</td>
<td>0.83</td>
<td>0.83</td>
<td>0.83</td>
<td>0.79</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>1.62</td>
<td>1.70</td>
<td>1.75</td>
<td>1.82</td>
<td>3.01</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MAE</td>
<td>1.23</td>
<td>1.31</td>
<td>1.37</td>
<td>1.39</td>
<td>2.45</td>
<td></td>
</tr>
<tr>
<td>OB3</td>
<td>R</td>
<td>0.94</td>
<td>0.94</td>
<td>0.92</td>
<td>0.94</td>
<td>0.93</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>1.3</td>
<td>1.18</td>
<td>1.80</td>
<td>1.10</td>
<td>3.59</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MAE</td>
<td>1.01</td>
<td>0.90</td>
<td>1.49</td>
<td>0.82</td>
<td>3.03</td>
<td></td>
</tr>
<tr>
<td>DS7</td>
<td>R</td>
<td>0.86</td>
<td>0.88</td>
<td>0.88</td>
<td>0.87</td>
<td>0.88</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>1.23</td>
<td>1.28</td>
<td>1.37</td>
<td>1.41</td>
<td>1.39</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MAE</td>
<td>0.96</td>
<td>1.06</td>
<td>1.11</td>
<td>1.11</td>
<td>1.08</td>
<td></td>
</tr>
</tbody>
</table>

(Note: GP-E Equation based GP model).

Fig. 11. The initial screen showing the buoy locations.

buoys was possible. Its advantage would be that even if both wind and wave sensors are out of order, it should still be possible to obtain the wind speed at a particular station from the wave observations at any other wave buoy. For this purpose, two approaches were followed. One of them consisted of using one of the 5 GP models to predict the wind speed to test its workability (in respect of individual testing sets) at other sites. The other approach involved combining the training data sets of all stations and developing a common GP model. Initial studies were carried out for wind speed predictions, and thereafter for both wind speed and direction predictions. The developments of such GP models are discussed in the following sections.

The individual equations developed using the GP approach is discussed in the preceding section. The resulting equations (7)–(10) and Box I formed at locations: DS1, SW2, SW3, OB3 and DS7 were tested simultaneously at all of the locations. The comparative performance of such equations is shown in Table 3. It can be observed from this table that the Eq. (7) developed at DS1 produced excellent results overall in terms of a high correlation coefficient, and low errors as compared to equations of other sites (see Table 3) (although the equation developed at site OB3 worked fairly similarly in this case). This could be due to its favorable location for a predominant wind direction from the south-west, and proximity to other locations except DS7. The largest database involved in training and testing at this location could also be another reason for this observation.

The results obtained using the above equation are given in Table 3, from which it can be seen that this equation yielded satisfactory results for the entire range of wind speeds.
7. A regional GP model based on combined data

After noticing good performance through a common GP model, as in the previous section, attempts were made to develop another equation trained on the entire dataset collected at all the five locations. Similar input–output schemes and methodology (used in the case of individual database) were adopted for development of such a common regional equation. The dataset used in training this model also included the cyclonic period at location DS1 and SW2. The development of such a model trained on entire data sets could help users to obtain wind information at any desired station where the recording had stopped, due to, say, maintenance activity.

The evolved equation was as follows:

\[
W_s = \frac{(H_s + T_z) + \sqrt{(T_z + \sqrt{T_m}) / (T_m / (H_s/0.26))}}{\left(\frac{T_m / (H_s/0.26)}{\sqrt{T_m / (H_s/0.26)}}\right) / (T_m / \sqrt{T_m/0.26})}.
\]

(13)
Eq. (13) was tested again at each site. The test results at each location are shown in Table 4. The second column of Table 4 shows the results reproduced from Table 3 and the third column shows the performance of the evolved equation (13).

It may be seen that both these exercises yielded comparable results, with the case shown in column (2) producing more acceptable error statistics. The slightly less accurate results seen in column (3) are mainly due to the large variations in the input used for training, including storm events.

8. A regional GP model based on the entire data and the resolved components

In another alternative for combined – speed and direction – two GP models were calibrated separately based on the u–v components of the speed. The evolved equations for u–v components are respectively as follows:

\[
W_s(u) = \left( \frac{T_m}{\exp(H_s)} - \sqrt{T_z} \right) - \left( \frac{H_s}{\log \left( \exp \left( \log \sqrt{T_m} \right) \right)} \right), \quad (14)
\]

\[
W_s(v) = \frac{\exp \sqrt{T_m} - (H_s + \log \sqrt{T_m})}{\log \left( T_m \sqrt{H_s} \right) \log \left( (H_s + \exp \sqrt{T_z}) + \sqrt{T_m} \right)}. \quad (15)
\]

The evolved equations were further tested at each location. A similar pattern of results as in the previous sections was obtained. It was accordingly noted that the model performance with the u–v combination was associated with somewhat inferior results, since the models were first trained separately for obtaining speed and direction and thereafter combined, and this procedure could have caused additional uncertainties. The combined prediction of wind speed and direction based on the u- and v-components was slightly less accurate in predicting speeds than separate speed and direction predictions, and hence when the speed prediction alone is desired, it is more appropriate to use separate models (developed with wind speed as a single component), while the wind direction estimate could be made from the u- and v-component models developed at the same time step. A common regional GP evolved equation was developed for all the five sites in the Arabian Sea.

9. Putting the developed model into practice

The studies described so far for the prediction of wind speed and direction, based on an inverse cause-effect modeling can be put into practice through an integrating platform of the graphical user interface (GUI). It was seen that although results produced by both ANN and GP were close to each other, the GP outcome was marginally, but consistently better, than that of the ANN. Hence the present GUI was developed on the basis of the GP models. The 3-hourly measurements of wind speed and direction

![Fig. 14. The screen showing the estimate of wind speed and direction.](image-url)
at a number of buoy locations in the Arabian Sea (Fig. 11) are currently telemetered to a server at NIOT, Chennai, India. Currently such observations are made available to registered clients of NIOT through a web based service. The GUI under consideration is intended to connect the models developed in the study to this web server.

The user has to click on the station where he wants to have the predictions (Fig. 11). This will generate the screen as shown in Fig. 12. Clicking on the ‘Load’ button will bring appropriate input files into the picture (Fig. 13), which in turn will be linked to the executable wind prediction program developed in this study. The estimations of speed and direction will be made after clicking on the button ‘Show Estimate’ and will be displayed as shown in Fig. 14.

The above software is of an intelligent type, in that if data of Hs, Tz, Tm do not get observed (missing values), then it will automatically generate the wind data from the regional equations (14) and (15) and report the estimated wind speed and direction.

10. Conclusions

The preceding sections described how the unobserved values of wind speed and direction at a wave buoy location can be calculated from the observed wave parameters through an inverse modeling exercise, based on genetic programming and artificial neural networks.

The GP models were able to learn the underlying complex inverses relationship well. Although the results of both soft approaches rivaled each other, the GP showed a tendency to produce more accurate values than the ANN adopted.

It was also noticed that the equation based GP model could be equally useful as the one based on a computer program, and hence for the sake of simplicity in implementation, the former can be adopted.

In case the entire wave rider buoy does not function for some period, a common regional GP model prescribed in this work can still produce desired wind parameters with the help of wave observations available from anywhere in the region.

A graphical user interface was developed that transfers the developed models for their actual use in the field by the data collection organization.

Acknowledgements

Authors thank Dr. K. Premkumar, Director, NDBP and Dr. G. Latha, Scientist (F) of NIOT, Chennai, India for sparing data used in this project and also for useful technical input.

References