Environ Earth Sci (2016)75:590
DOI 10.1007/s12665-015-5221-x

ORIGINAL ARTICLE

# A comparative study for estimation of wave height using traditional and hybrid soft-computing methods

**Chandrabhushan Roy[1] · Shervin Motamedi[1,2] · Roslan Hashim[1,2] · Shahaboddin Shamshirband[3] · Dalibor Petković[4]**

Received: 24 March 2015 / Accepted: 29 December 2015


**Abstract** The present study developed a wave height prediction model by the recorded climatic data. We used 1-year buoy data for training and testing the developed soft-computing model. Models were developed using a novel method based on the Support Vector Machine (SVM) coupled with the Firefly Algorithm (FFA). This research work used the FFA for estimating the optimum parameters. In addition, this work compared the predicted results of SVM-FFA model to the artificial neural networks (ANNs) and genetic programming (GP). The results indicate that the SVM-FFA approach attains an improvement in capability of generalization and predictive accuracy in comparison to the GP and ANN. A thorough statistical analysis was conducted to compare the predictions of three models i.e., among the SVM-FFA, ANN, and GP. A high $R^2$ value of 0.979 was obtained for the SVM-FFA predictions. Further, the ANN and GP results showed $R^2$ values of 0.524 and 0.525, respectively. Moreover, achieved results indicate that the developed SVM-FFA model can be used with confidence for future research works on formulating novel models for predictive strategy on wave height. The results also show that the new algorithm can learn thousands of times faster than the former popular learning algorithms. This study finds that the application of SVM-FFA is the likely alternative method for estimating the wave height.

**Keywords** Buoy data · Wave prediction · Hybrid methods · Support vector machine · Firefly algorithm

✉ Roslan Hashim
roslan@um.edu.my

✉ Shahaboddin Shamshirband
shamshirband1396@gmail.com

[1] Department of Civil Engineering, Faculty of Engineering, University of Malaya, 50603 Kuala Lumpur, Malaysia

[2] Institute of Ocean and Earth Sciences, University of Malaya, 50603 Kuala Lumpur, Malaysia

[3] Department of Computer System and Information Technology, Faculty of Computer Science and Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia

[4] Department for Mechatronics and Control, Faculty of Mechanical Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Nis, Serbia

## Introduction

Understanding the wave characteristics is vital for planning the offshore activities, such as maritime traffic (Maresca et al. 2014), fishing (Singhal et al. 2013), construction of marine structures (Jay 2010), extraction of marine renewable energy (Azzellino et al. 2013), and coastal protection (Kaliraj et al. 2014). Primarily investigators obtain the wave data for planning the offshore activities from the measurements of radars, buoys, and satellites. Researchers apply the empirical and numerical methods to the field measurements to study the wave characteristics. Offshore buoys are considered the most reliable measuring tools for obtaining wind and wave data (Alexandre et al. 2015; Battjes 1974). Buoy data are used in the empirical, numerical, and soft-computing methods for estimating and understanding wave characteristics (Abed-Elmdoust and Kerachian 2012; Alexandre et al. 2015; Jain et al. 2009; Thirumalaiah and Deo 2000).

Soft-computing techniques help in solving problems, which show non-linearity such as wind-driven waves. In recent years, many researchers have employed various soft-computing techniques for estimating the wave height. For

Published online: 28 March 2016

example, Altunkaynak and Wang (2012) used the Genetic Algorithm (GA) along with the Kalman Filtering (KF) for optimizing and predicting significant wave height in the Lake Okeechobee, Florida. Similarly, Mahjoobi et al. (2008) employed wave and wind data from a buoy in the Lake Ontario to train and test an alternative hindcast model based on the Fuzzy Inference System (FIS) and Artificial Neural Networks (ANNs). In a similar study, Makarynskyy (2004) used the ANN for improving short-term wave forecast. The developed model was tested through hourly observations of significant wave height and wave period from a buoy in the Atlantic Ocean. It was reported that the neural network improves the accuracy of forecast. In another study, Özger (2010) proposed a forecast scheme based on the wavelet and fuzzy logic (WFL) approach. The model was tested using buoy data from the west coast of the USA. The results showed high coefficient of efficiency (0.753), which indicates satisfactory predictions.

Recently, the SVM has gained reputation as a strong soft-computing approach (Burges 1998; Lee and Verri 2003; Ornella and Tapia 2010; Xu et al. 2014). The SVM has been applied in various engineering and science disciplines, such as mining (Jiang et al. 2014; Li et al. 2014), geo-technical (Ren et al. 2015; Wu et al. 2014), and coastal protection (Singh et al. 2014). Nevertheless, limited studies have implemented SVM for predicting significant wave height. In a recent study, Mahjoobi and Adeli Mosabbeb (2009) employed the SVM for estimating the significant wave height for a buoy station in the Lake Michigan. In their study, they compared the SVM with ANN, multi-layer perceptron (MLP) and radial basis function (RBF). It was observed that SVM requires less computational resource, and the results are marginally better than the ANN.

The accuracy of an SVM model relies on selection of the model parameters (Chung et al. 2003; Friedrichs and Igel 2005; Kisi et al. 2015; Lorena and De Carvalho 2008). Researcher uses smart strategies for selecting these model parameters. In addition, one needs to focus on aligning the model parameters. Researchers normally apply the conventional optimization algorithms, such as the grid search and gradient decent algorithms, but with limited success. Computational complexity is the downside of the grid search algorithm, which restricts their applicability in simple cases. Multiple local solutions exist for most of the optimization problems. New evolutionary algorithms, such as the Firefly algorithm (FFA) and genetic programming (GP), seem to be the best approaches, because they are capable of providing global solution to such optimization problems (Gocić et al. 2015; Kisi et al. 2015; Shamshirband et al. 2015; Yang 2010).

This study aimed to implement an innovative soft-computing method to predict significant wave height by combining the support vector machine (SVM) with the Firefly Algorithm (FFA). No study was found that implemented the SVM techniques with the FFA for predicting significant wave height. Therefore, this study developed a wave forecasting model using the SVM-FFA model and compared the prediction with other traditional methods, such as the ANN and GP. In this respect, 1-year buoy data were used for training and testing the models. The measurement data were obtained from the National Data Buoy Centre (NDBC) portal. Four meteorological parameters were used for estimating wave height, namely the wind speed, wind direction, seawater temperature and air temperature. Results of the SVM-FFA were compared to experimental measurements using statistical indicators, such as the root means square error (RMSE), Coefficient of determination ($R^2$) and Pearson coefficient ($r$). Further, the results of SVM-FFA were compared to other soft-computing techniques, such as ANN and GP.

## Methodology

The structure of "Methodology" is as following: "Description of study area" presents study area and measurement data for training and testing the predictive capability of SVM-FFA; "Measurement data" explains the selected input parameters, which have direct and indirect influence on wave height formation and prediction; and "Soft-computing simulation setup" presents the SVM-FFA, ANN, and GP models used for predicting the wave height. Figure 1 shows a schematic diagram of this research work.

### Description of study area

This study used recorded meteorological and climatological data for training and testing of the SVM-FFA. The measurement data were obtained from buoy station 44150 of the NDBC portal (Fig. 2). The observation stations were located in and around the Georges Bank about 170 nautical miles East of Hyannis, MA, USA. Figure 3 shows the bathymetry (sea depth) around the station. The bathymetry of the region was obtained by visualizing satellite data in *MATLAB* software. Bathymetry around the buoy varies from 50 to 500 m. We used the buoy data of 1 year, starting from 1st January to 31st December, 2014 on an hourly basis.

### Measurement data

The measurement data comprise significant wave height, wind speed, wind direction, air temperature, and sea water temperature. The climatological and meteorological data were recorded at an hourly interval. This study used 8505 data points for each input variable (42,525 data points).

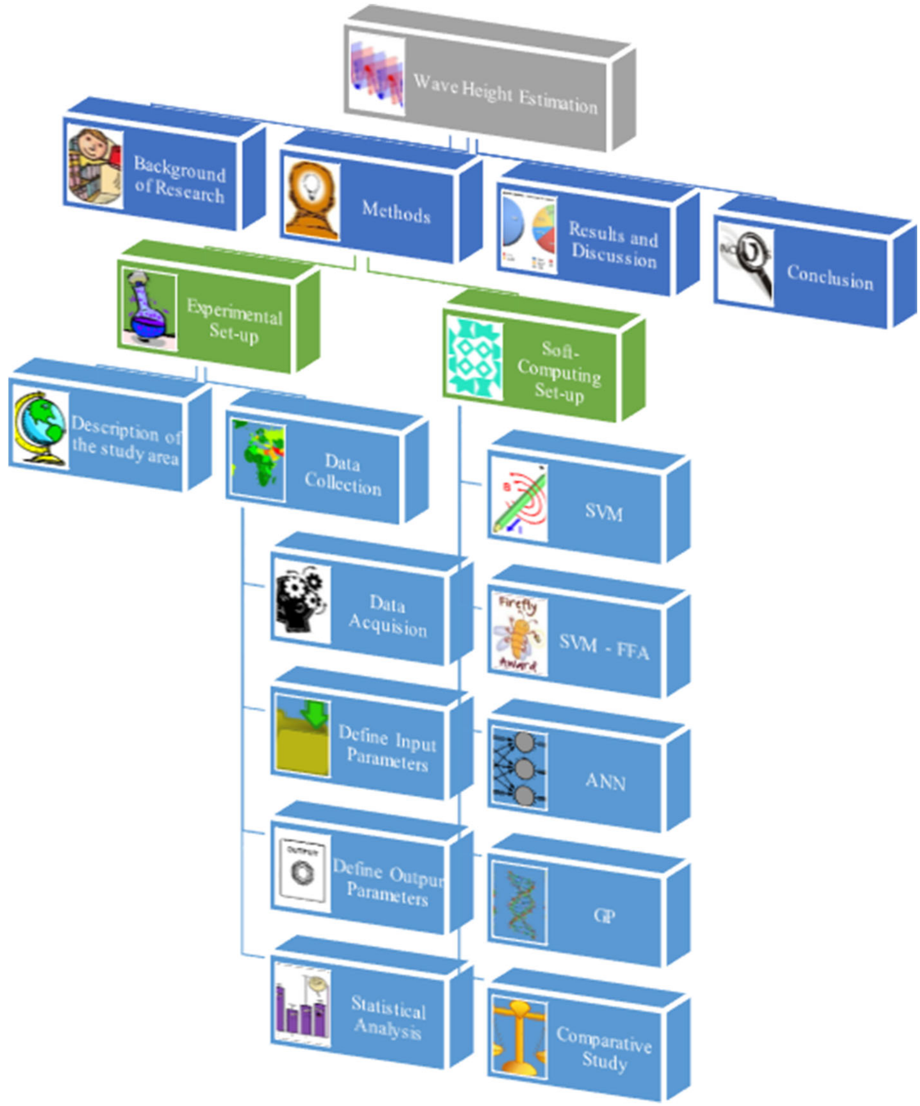


**Fig. 1** Flowchart showing the wave height prediction

Table 1 shows the statistical summary of measured data used in the study. Station 44150 recorded maximum and minimum wave heights of 10 and 0.4 m, respectively. Wind speed at the station shows a maximum reading of 24.5 m/s. South-west blowing winds are the most prevalent in the region. In addition, this region shows a wide variation in the air and water temperature. Figures 4 and 5 show the sea surface temperature and sea surface salinity around the stations on 1st January, 2014, which were obtained using the satellite and were visualized in *MATLAB*. Sea surface temperature varied from 2 to 6 °C, while the sea surface salinity varies from 31 to 32.6 ppm.

## Soft-computing simulation setup

This section describes the SVM-FFA model that was used for predicting significant wave height. In addition to the AVM-FFA, other soft-computing models, such as the ANN and GP, are also presented. "Model input and output variables" addresses the models' input and output parameters.

### *Model input and output variables*

Tables 2 and 3 illustrate input and output variables in terms of the definition and obtained values. These input parameters have direct or indirect influence on the wave height. For example, surface wind speed and direction have direct influence, whereas sea surface temperature or water salinity has indirect influence. Figure 6 shows schematic diagram of four inputs and one output used for training and testing the developed wave height prediction model.

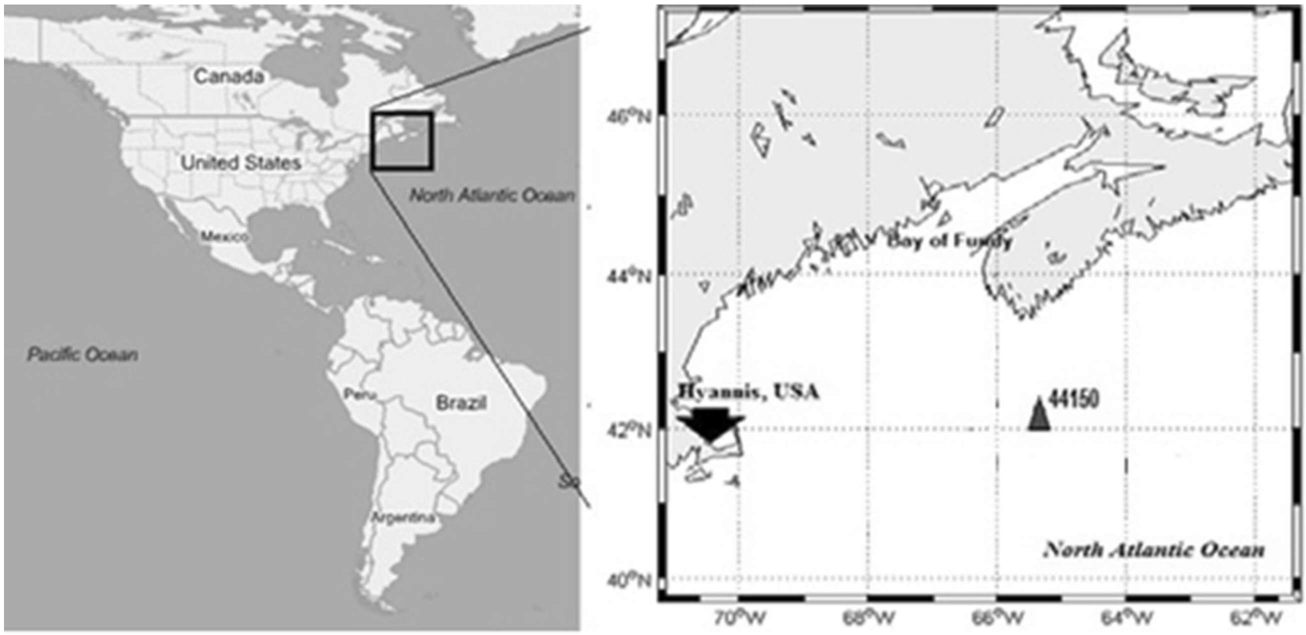


**Fig. 2** Location of the measurement station 44150 in the North Atlantic Ocean

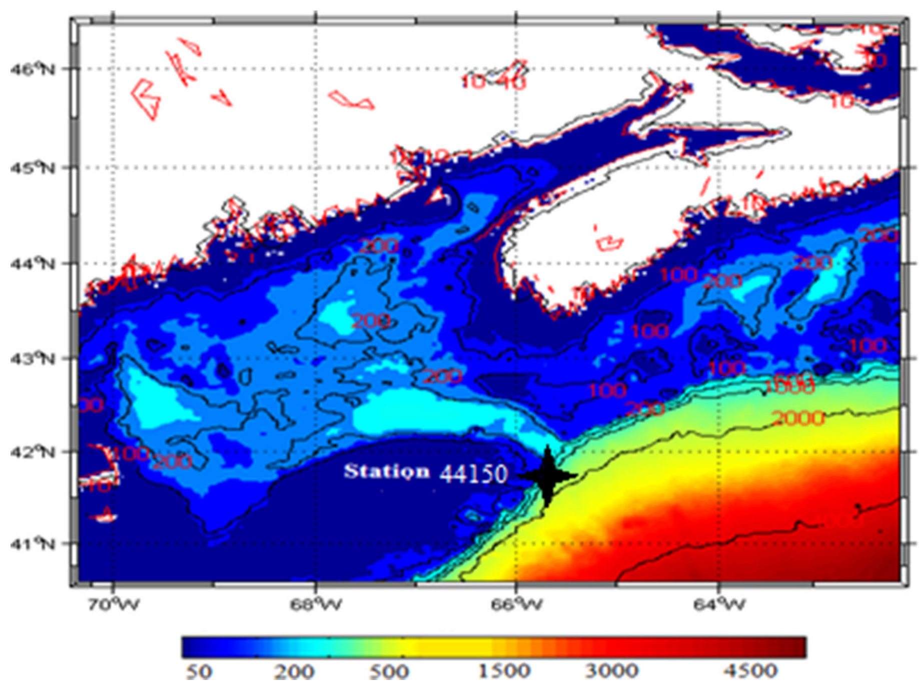


**Fig. 3** Satellite data showing bathymetry of the *North Atlantic Ocean* around station number 44150 registered with NDBC

**Table 1** Statistical summary of the measurement data

| | Maximum | Minimum | Average | Standard deviation |
|---|---|---|---|---|
| $H_s$ (m) | 10 | 0.40 | 1.89 | 1.12 |
| $W_s$ (m/s) | 24.50 | 0.00 | 6.76 | 3.58 |
| $W_\theta$ (°) | 360 | 0.00 | 204.28 | 97.69 |
| $T_a$ (°C) | 24.5 | −12.1 | 9.31 | 8.99 |
| $T_w$ (°C) | 22.3 | 2.1 | 13.92 | 4.62 |

*Support vector machine*

Fields of environment, computing, and hydrology widely use the Support vector machines (SVMs) as a novel soft-computing learning algorithm (Asefa et al. 2006; Ji and Sun 2013; Lee and Verri 2003; Lu and Wang 2005; Sun 2013). The SVMs solve problems related to pattern recognition, classification and regression analysis. The

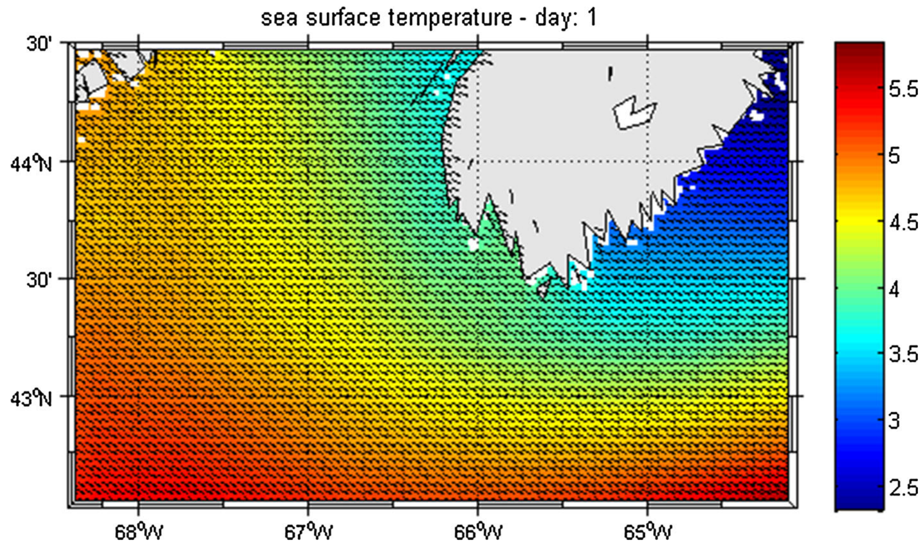


**Fig. 4** Sea surface temperature of *North Atlantic Ocean* around station number 44150 on 1st January, 2014 (Satellite data visualized in *MATLAB*)

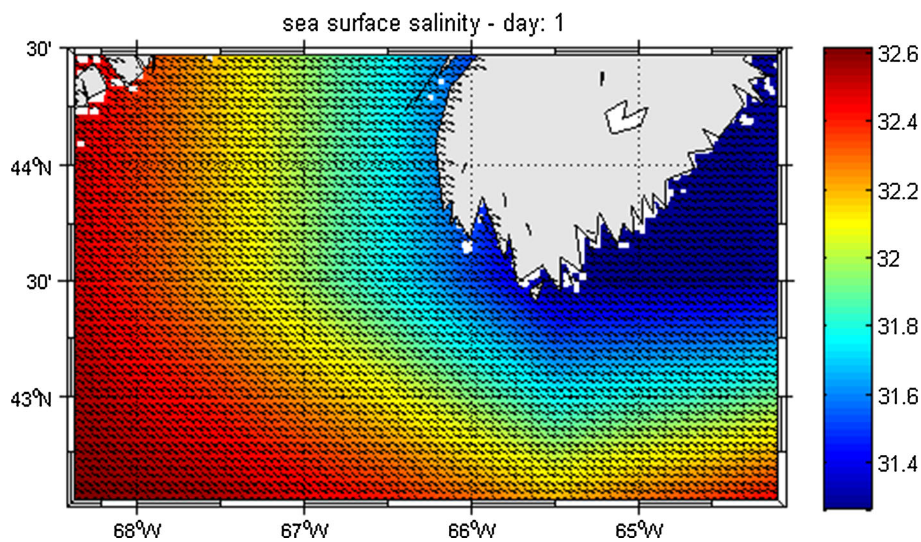


**Fig. 5** Sea surface salinity of *North Atlantic Ocean* around station number 44150 on 1st January, 2014 (Satellite data visualized in *MATLAB*)

**Table 2** Models' input parameters for wave height prediction study

| Inputs | Description of parameters | Characterization of parameters |
|---|---|---|
| Input 1 | Sea surface wind speed | $W_s$ (m/s) |
| Input 2 | Sea surface wind direction | $W_\theta$ (°) |
| Input 3 | Air temperature | $T_a$ (°C) |
| Input 4 | Sea surface temperature | $T_w$ (°C) |

**Table 3** Models' output parameter for wave height prediction study

| Output | Description of parameter | Characterization of parameter |
|---|---|---|
| Output | Wave height | $H_s$ (m) |

SVMs perform better than similar methodologies, such as classical statistical and neural network models (Collobert and Bengio 2000; Huang et al. 2002; Mukkamala et al. 2002; Sung and Mukkamala 2003; Cortes and vapnik 1995).

Vapnik and Vapnik (1998) and Vapnik (2000) recently proposed the theory behind the evolution of SVMs. Traditional methods minimize local training error, whereas the SVMs focus on minimizing upper bound for the generalization errors (Vapnik and Vapnik 1998). SVMs have upper hand over the traditional soft-computing methods because it is based on the Structural Risk Minimization principle (Vapnik and Vapnik 1998). Besides, because of the convex nature of the problem, unique solution is provided. This method maps the input data on a high-dimension space kernel function, which discreetly contains non-linear transformation.

According to Vapnik's theory, the SVM equations can be represented as given in Eqs. (1–4). Consider a set of $n$

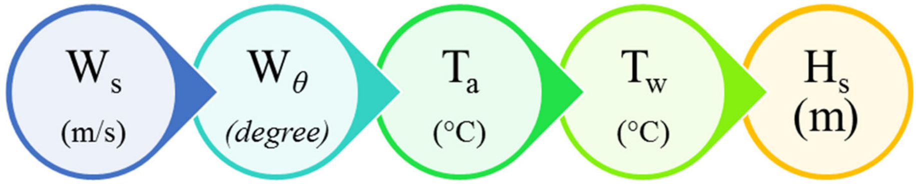


**Fig. 6** Model inputs and output used for training and testing the developed wave height prediction model

data points by $\{x_i, d_i\}_i^n$, then the SVMs approximate the functions given in Eqs. (1) and (2).

$$f(x) = w\varphi(x) + b \tag{1}$$

$$R_{SVMs}(C) = \frac{1}{2}\|w\|^2 + C\frac{1}{n}\sum_{i=1}^{n} L(x_i, d_i), \tag{2}$$

where $x_i$ is the input space vector, $d_i$ is the target value. $\varphi(x)$ is the high-dimension feature space for mapping the input $x$, $b$ is a scalar, $w$ is a normal vector, $C\frac{1}{n}\sum_{i=1}^{n} L(x_i, d_i)$ represents the empirical error, and $w$ and $b$ are the constants of the linear function $f(x)$.

Parameters $w$ and $b$ are computed using the Eq. (2). It introduces $\xi_i$ and $\xi_i^*$ which are the positive slack variables representing the lower and upper excess deviation (Vapnik and Vapnik 1998).

$$\text{Minimize } R_{\text{SVMs}}(w, \xi^{(*)}) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) \tag{3}$$

$$\text{Subject to } \begin{cases} d_i - w\varphi(x_i) + b_i \leq \varepsilon + \xi_i \\ w\varphi(x_i) + b_i - d_i \leq \varepsilon + \xi_i^* \\ \xi_i\xi_i^* \geq 0, \; i = 1, \ldots\ldots\ldots.l, \end{cases}$$

where the regularization term is $\frac{1}{2}\|w\|^2$, $\varepsilon$ is the loss function that is related to approximation accuracy of the training data point, $C$ represents the error penalty factor, and $l$ is the size of training data set.

By solving Eq. (1) a generic function is obtained given by Eq. (4).

$$f(x, a_i, a_i^*) = \sum_{i=1}^{n}(a_i - a_i^*)K(x, x_i) + b, \tag{4}$$

where $K(x, x_i) = \varphi(x_i)\varphi(x_j)$ and the kernel function is $K(x_i, x_j)$, which is obtained from the two inner vectors $x_i$ and $x_j$ in the feature spaces $\varphi(x_i)$ and $\varphi(x_j)$, respectively.

SVMs perform data correlation by non-linear mapping. One can build a direct computation method for the kernel function, if a means exits for direct calculation of the inner product of feature space as a function of original input variable.

The SVM provides four basic kernel functions. The four basic kernel functions are sigmoid, lineal, radial and polynomial. Researchers have found that the Radial basis function (RBF) is the most favorable kernel function. It is simple, efficient, adaptive and reliable. In addition, it can handle complex parameters (Rajasekaran et al. 2008; Wu and Wang 2009; Yang et al. 2009). The RBF kernel works with the linear equations for training purpose, instead of quadratic programming problem that is complex and computationally demanding (Shamshirband et al. 2014). Hence, this study adopted RBF with parameter $\sigma$. Further, Eq. 5 defines the non-linear radial basis kernel function as:

$$K(x_i, x_j) = \exp\left(-\gamma\left\|x_i - x_j\right\|^2\right), \tag{5}$$

where variables $x_i$ and $x_j$ are input space vectors (vectors computed from the training or testing data set).

The choice of three parameters ($\gamma, \varepsilon$ and $C$) determines the RBF kernel functions' predictive accuracy. The traditional method selects the parameters by the grid search method, which is an exhaustive tool that searches manually through the specified subset of the hyper parameter space. A grid search algorithm needs to be guided by certain performance metric. Therefore, the present study used the Firefly optimization algorithm for optimizing the parameters.

### *SVM parameters selection using firefly optimization algorithm*

Researchers have developed many optimization algorithms, inspired by nature for solving classical problems (Assareh et al. 2010; Whitley et al. 1990; Dorigo and Stützle 2003; Yang and Deb 2014). Some of such algorithms are: particle swarm optimization (PSO), cuckoo search (CS), ant colony optimization (ACO) and genetic algorithm (GA). Survival and selection of the fittest specie in nature form the basis of these algorithms. The firefly algorithm developed by Yang (2009) is the most recent biologically inspired optimization algorithm.

The firefly algorithm implements certain behavioral pattern of fireflies (insect), such as their flashing characteristic. This insect (Firefly) attracts mates and prey using its quality of bioluminescence. Other fireflies trail its path through the luminance produced. Research works suggest that the Fireflies algorithm (FA) is robust and efficient in comparison to other biological inspired algorithms (Fister et al. 2013; Pal et al. 2012; Yang 2013).

Fundamental rules of FA are as given below:

(1) It assumes that all the fireflies are of same gender, and can attract each other irrespective of their gender. (2)

Attraction between the fireflies is directly proportional to the luminance intensity. Intensity decreases with distance, and the firefly with lower intensity is attracted (move) to the one with higher intensity. (3) The brightness is proportional to the fitness (objective function) (Yang 2009).

Variation in intensity of light and formulation of objective functions are the major issues in developing the Firefly. Equation (6) represents the relationship between the intensity of light and distance.

$$I(r) = I_o \exp(-\gamma r^2), \tag{6}$$

where $I$ represents the light intensity, $r$ is distance from a firefly, $I_o$ represents initial intensity ($r = 0$), and $\gamma$ is the light absorption coefficient. $\gamma$ has a constant value between 0.1 and 10 (Yang 2009). It represents the attractiveness $\beta$ of a firefly at a distance $r$ as:

$$\beta(r) = \beta_o \exp(-\gamma r^2), \tag{7}$$

where $\beta_o$ is the attractiveness at $r = 0$. $T = 1\sqrt{\gamma}$ as the characteristic distance defines Eq. (7) that leads to obvious changes from $\beta_o$ to $\beta_o^{e-1}$.

Equation 8 represents distance between $i$ and $j$ fireflies.

$$r_{ij} = \|x_i + x_j\| = \sqrt{\sum\nolimits_{k=1}^{d} (x_{i,k} - x_{j,k})}. \tag{8}$$

Equation (9) represents the distance traveled by a firefly $i$ when attracted toward $j$ (brighter firefly).

$$\Delta x_i = \beta_o^{e^{-\gamma r^2}} (x_j - x_i) + \alpha \varepsilon_i, \tag{9}$$

where attraction is represented by the first term, while the randomization with $\alpha$ as randomization coefficient is represented by the second term. Value of the randomization coefficient varies between 0 and 1 (Ch et al. 2014), and $\varepsilon_i$ is the random number vector derived from a Gaussian distribution.

The $i$ movement is revised using Eq (10):

$$x_i^{i+1} = x_i + \Delta x_i. \tag{10}$$

Figure 7 summarizes the basic steps of the FA with the pseudo-code.

*Artificial neural networks*

The neural network architecture has multiple layers of feedforward network and uses a back propagation-learning algorithm. Three layers of the neural network are

(i) Input layer, with the input vectors $D \in R^n$ and $D = (X_1, X_2, \ldots, X_n)^T$;

(ii) Hidden or intermediate layer, with $Z = (Z_1, Z_2, \ldots, Z_n)^T$ as the outputs of $q$ neurons in the hidden layer ; and

**\Firefly Algorithm**

```
start
  Define the objective function, f(x), x = (x_i,,,,,,,,,,, x_d)T
  Generate initial population of fireflies x_i (i = 1,2, ... ... n)
  Determine light intensity I_i at x_i from f(x_i)
  Define light absorption coefficient γ
    while t < Max Generation
        Make a copy of population for movement function
      for i = 1:n all n fireflies
        for j = 1:i all n fireflies
            if (I_j > I_i)
            Move fireflies i and j in d-dimension;
            end if
            Attractiveness varies with distance r via exp [−γr]
            Evaluate new solution and update light intensity
        end
      end
      Rank the fireflies and find the current best
    end
    Post process results and visualization
end
```

**Fig. 7** Pseudo-code for Firefly Algorithm

(iii) an output layer, $Y = (Y_1, Y_2, \ldots, Y_n)^T$ as the outputs of the output layer.

Suppose $w_{ij}$ represents the weights and $y_j$ represents the thresholds between the input layer and hidden layer; $w_{jk}$ represents the weights and $y_k$ represents the thresholds between the hidden layer and output layer. Then, output from hidden and output layer can be represented as given below:

$$Z_j = f\left(\sum_{i=1}^{n} w_{ij} X_i - \theta_j\right) \tag{11}$$

$$Y_k = f\left(\sum_{j=1}^{q} w_{kj} Z_j - \theta_k\right), \tag{12}$$

where $f$ is the transfer function. Transfer function regulates mapping the sum of all inputs to their outputs. One can introduce non-linearity in the design of network by choosing suitable transfer function. Sigmoid is one such transfer function that increases monotonically (ranging from zero to one). One can refer to Govindaraju (2000) and Karunanithi et al. (1994) for further details on the ANNs.

*Genetic programming*

Genetic algorithm is based on the Darwinian theory of natural selection. It estimates relation between input and output variables in the form of equation. First, it populates initial set of equation with a random combination of input variables, functions and numbers that include logical, arithmetic and trigonometric operators. These set of random combinations are selected based on initial understanding of the process. Fitness of this initial set is then evaluated using an evolutionary process. Those elements that best describe relationship between the input and out variables are selected and the rest of elements are rejected.

In the next step, the selected equations are made to exchange information among them to produce a better equation using mutation and crossover techniques. Crossover is defined as the process in which the best part of the program is exchanged, whereas mutation is random exchange of parts of equation. This is similar to the natural reproduction process. This process repeats over many generation until an equation is found that best describes the relationship between the input and output variables. More in depth discussion on the GP can be found in the following works of Babovic and Keijzer (2005), Khu et al. (2001) and Koza (1992).

### Evaluating accuracy of proposed models

Predictive performances of proposed models are presented as root means square error (RMSE), Coefficient of determination ($R^2$) and Pearson coefficient ($r$). These statistics are defined as follows:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(P_i - O_i)^2}{n}}, \tag{13}$$

$$r = \frac{n\left(\sum_{i=1}^{n} O_i \cdot P_i\right) - \left(\sum_{i=1}^{n} O_i\right)\cdot\left(\sum_{i=1}^{n} P_i\right)}{\sqrt{\left(n\sum_{i=1}^{n} O_i^2 - \left(\sum_{i=1}^{n} O_i\right)^2\right)\cdot\left(n\sum_{i=1}^{n} P_i^2 - \left(\sum_{i=1}^{n} P_i\right)^2\right)}} \tag{14}$$

$$R^2 = \frac{\left[\sum_{i=1}^{n}(O_i - \overline{O_i})\cdot(P_i - \overline{P_i})\right]^2}{\sum_{i=1}^{n} O_i - \overline{O_i}\cdot\sum_{i=1}^{n} P_i - \overline{P_i}} \tag{15}$$

where $P_i$ and $O_i$ are the experimental and forecast values of wave height, respectively, and $n$ is the total number of test data.

## Results and discussions

This study predicted the significant wave height using the SVM-FFA. A thorough analysis of experimental and simulation results is presented and discussed in this section. The simulation results were compared to the field measurements using statistical indicators. In addition, this section provides the results from the simulation model from the SVM-FFA and compares its prediction to the ANN and GP soft-computing methods. First, the experimental buoy data that were used for training and testing purposes are presented. In the end, results from the SVM-FFA, ANN, and GP methods are compared using statistical parameters of root means square error (RMSE), Coefficient of determination ($R^2$) and Pearson coefficient ($r$).

### Experimental results

This study employed four meteorological parameters for estimating significant wave height, namely $W_s$, $W_\theta$, $T_a$, and $T_w$. This section presents the results of the buoy measurement that was used for training and testing the predictions of the soft-computing methods.

According to the measurement data from the buoy 44150, this site received maximum significant wave height of 10 m in February 2014. Figure 8 shows the annual trend of measured wave height at the station. Figure 9 shows the annual variation in wind speed. It receives an average wave height of 1.89 m. Similarly, this region received maximum wind speed of 24.50 m/s, while it recorded an average wind speed of 6.76 m/s. For most part of the year, wind blows in the south-west direction toward the coastline. It can be concluded from the Figs. 8 and 9 that the magnitude of wave height and wind speed have similar trends.

Trends of the air temperature and sea water temperature are shown in the Figs. 10 and 11. It can also be concluded that air temperature in the region shows wider variation than the water temperature. For example, buoy recorded maximum and minimum air temperature of 24.5 and −12.10 °C, respectively. On the other hand, water temperature shows maximum and minimum of 22.3 and 2.1 °C, respectively. These experimental field measurements were used for validating the predictions of soft-computing methods (see "Soft-computing results and discussions").

### Soft-computing results and discussions

The simulation results from the SVM-FFA model were validated with the field measurements presented in the previous section. Further, this study used experimental results for obtaining predictions from the ANN and GP soft-computing techniques, which were compared to the results of the SVM-FFA.

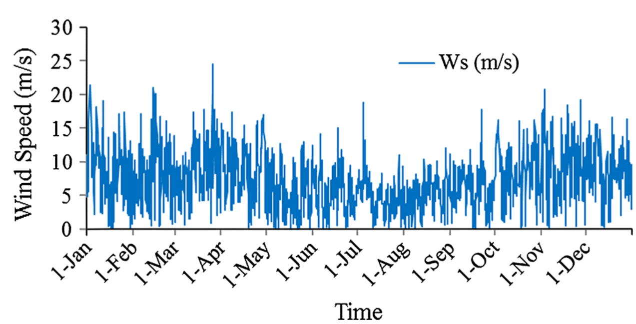


**Fig. 8** Variation in wave height at station 44150; from 1st January to 31st December, 2007

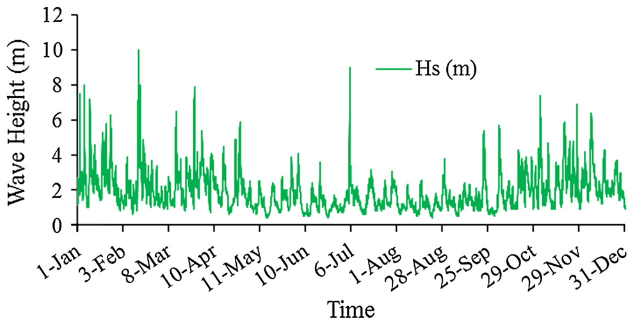


**Fig. 9** Variation of wind speed at station 44150; from 1st January to 31st December, 2014

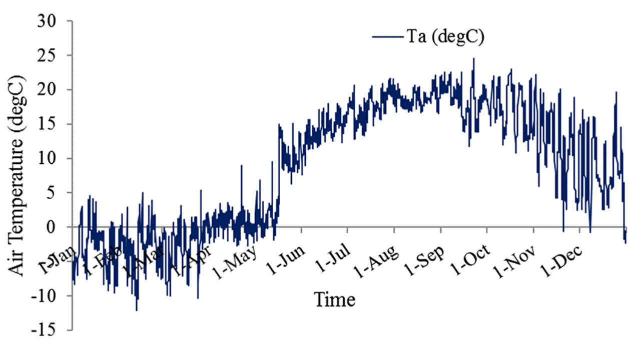


**Fig. 10** Variation of air temperature at station 44150; from 1st January to 31st December, 2014

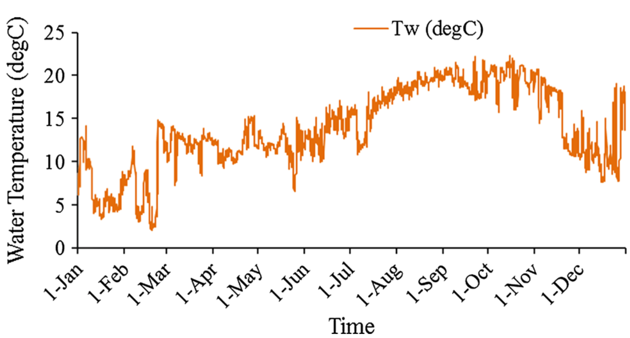


**Fig. 11** Variation of water temperature at station 44150; from 1st January to 31st December, 2014

*Evaluation of performance for the developed SVM-FFA model*

In this section, performance results of the SVM-FFA wave height predictive model are reported. Figure 12a shows the accuracy of the developed SVM-FFA wave height predictive model. Figure 12a compares the actual measurement with the predictions of the AVM-FFA using $R^2$. Figure 12b, c shows the accuracy of the developed GP and ANN wave height predictive models, respectively. Results show that wave height predictions of the SVM-FFA compare well with the field measurements. A strong $R^2$ value of 0.979 was obtained for the SVM-FFA predictions. Further, the ANN and GP results show $R^2$ values of 0.524 and 0.525, respectively, which are much lower than the prediction accuracy of the AVM-FFA.

It should be noted that even the wave height predicted from the ANN and GP models is within acceptable limits. Therefore, these observations confirm a very high value for the coefficient of determination for the SVM-FFA and acceptable values for the ANN and GP. Thus, prediction results are in highly consistent with the measured values for the SVM-FFA method. In addition, the number of produced overestimated and underestimated values is inadequate. Consequently, it is clear that the predicted values have high level of precision. This study further compared the results of the three models using multiple statistical indicators to demonstrate advantages and disadvantages of each model.

## Comparative study of the proposed SVM-FFA model with other soft-computing methods

To establish the virtues of the proposed SVM-FFA approach on a more positive and concrete basis, the prediction accuracy was compared to that of the GP and ANN methods, which were used as benchmark. RMSE, $r$ and $R^2$ statistical indicators were employed for comparison. Table 4 presents the results of prediction accuracy for test data sets, because the training error is not a credible indicator for prediction potential of a particular model.

The SVM-FFA model outperforms the GP and ANN models based on the results in Table 4. The SVM-FFA model provides significantly better results than the benchmark models. On the basis of RMSE analysis with comparison to the ANN and GP, it may be concluded that the proposed SVM-FFA outperformed the results obtained with the benchmark models.

Results of this study indicate that for predicting wave height, the SVM-FFA technique has advantage over the ANN and GP techniques, because its results are more consistent with the field measurements. This can be attributed to smart and efficient combination of the AVM with the Firefly algorithm, which is an innovative and smart strategy for selecting the model parameters of SVM. In addition, the SVM is computationally more efficient than the ANN and GP. Therefore, it can be concluded that the SVM-FFA technique can be used in prediction of wave height to improve prediction accuracy, coastal planning, and management of coastal activities, such as fishing, oil exploration, off shore construction, and coastal protection. The results also show that the new algorithm can learn thousands of times faster than former popular learning algorithms. This study applied the trained models for wave height prediction in the deeper ocean. Further research is

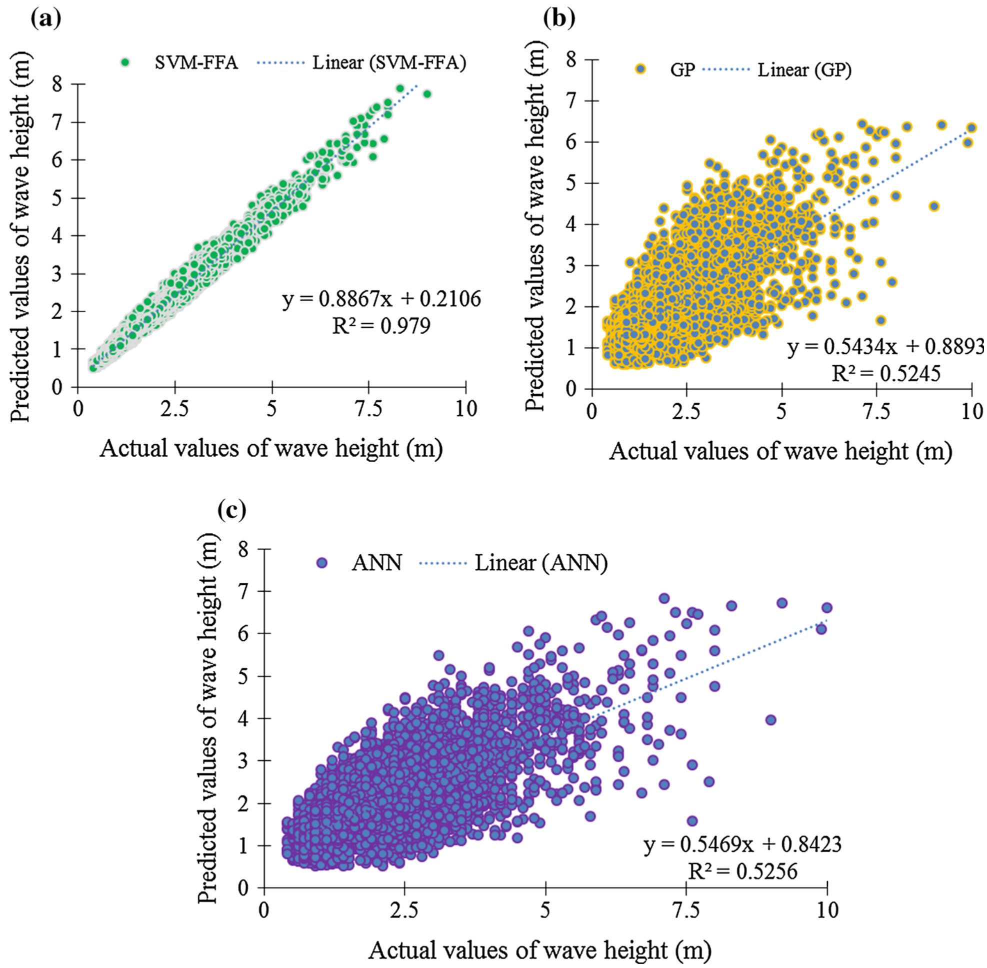


**Fig. 12** Scatter plots of actual and predicted values of wave height using **a** SVM-FFA, **b** GP and **c** ANN method

**Table 4** Comparative performance statistics of the SVM-FFA, ANN and GP wave height predictive models

| SVM-FFA | | | ANN | | | GP | | |
|---|---|---|---|---|---|---|---|---|
| RMSE | $R^2$ | $r$ | RMSE | $R^2$ | $r$ | RMSE | $R^2$ | $r$ |
| 0.175098 | 0.979 | 0.989441 | 0.700392 | 0.5256 | 0.725011 | 0.701226 | 0.5245 | 0.724229 |

required to test this model under various topographic conditions, such as shallow coastal water, estuaries, and narrow straits, where wave height is affected by local condition.

## Conclusion

Results from this study indicate that the hybrid soft-computing model i.e., the SVM-FFA better estimates the significant wave height than traditional methods, such as the ANN and GP. This was reflected in the statistical analysis, where the SVM-FFA yielded very high $R^2$ value of 0.979; while, ANN and GP results showed $R^2$ values of 0.524 and 0.525, respectively. This study considered four parameters for predicting wave height and developed a wave height prediction model using the SVM-FFA. One-year buoy measurement data were employed for testing and training the developed model. Parameters were selected carefully and comprised those with direct or indirect effect on wave height. The data used in the study show satisfactory performance of the SVM-FFA model in predicting wave height for deeper ocean.

First, this study collected actual measured data from the buoy station under the NDBC for each input variable. The study carried out a systematic approach by developing a SVM-FFA wave height predictive model. Later, measurement data were used as input for validating the SVM-FFA and other soft-computing methods. A comparison of the SVM-FFA method with ANN and GP technique was performed to evaluate the performance of prediction. The accuracy of the prediction results was measured in terms of

RMSE, $r$ and $R^2$. The results indicate that the SVM-FFA predictions are superior to the GP and ANN. Additionally, results revealed robustness of the method.

**Acknowledgments** The authors are thankful to the support received from the HIR-MOHE office, University of Malaya under Grant No. UM.C/HIR/MOHE/ENG/34.

## References

Abed-Elmdoust A, Kerachian R (2012) Wave height prediction using the rough set theory. Ocean Eng 54:244–250

Alexandre E, Cuadra L, Nieto-Borge J, Candil-García G, del Pino M, Salcedo-Sanz S (2015) A hybrid genetic algorithm—extreme learning machine approach for accurate significant wave height reconstruction. Ocean Model 92:115–123

Altunkaynak A, Wang K-H (2012) Estimation of significant wave height in shallow lakes using the expert system techniques. Expert Syst Appl 39:2549–2559. doi:10.1016/j.eswa.2011.08.106

Asefa T, Kemblowski M, McKee M, Khalil A (2006) Multi-time scale stream flow predictions: the support vector machines approach. J Hydrol 318:7–16

Assareh E, Behrang M, Assari M, Ghanbarzadeh A (2010) Application of PSO (particle swarm optimization) and GA (genetic algorithm) techniques on demand estimation of oil in Iran. Energy 35:5223–5229

Azzellino A, Ferrante V, Kofoed JP, Lanfredi C, Vicinanza D (2013) Optimal siting of offshore wind-power combined with wave energy through a marine spatial planning approach. Int J Mar Energy 3:e11–e25

Babovic V, Keijzer M (2002) Rainfall Runoff Modelling Based on Genetic Programming. Hydrol Res 33(5):331–346. http://hr.iwaponline.com/content/33/5/331

Battjes JA (1974) Computation of set-up, longshore currents, run-up and overtopping due to wind-generated waves. TU Delft, Delft University of Technology

Burges CJC (1998) A Tutorial on Support Vector Machines for Pattern Recognition. Data Min Knowl Discov 2(2):121–167. doi:10.1023/A:1009715923555

Ch S et al (2014) A support vector machine-firefly algorithm based forecasting model to determine malaria transmission. Neurocomputing 129:279–288

Chung K-M, Kao W-C, Sun C-L, Wang L-L, Lin C-J (2003) Radius margin bounds for support vector machines with the RBF kernel. Neural Comput 15:2643–2681

Collobert R, Bengio S (2000) Support vector machines for large-scale regression problems. Institut Dalle Molle d'Intelligence Artificelle Perceptive (IDIAP), Martigny, Switzerland, Tech Rep IDIAP-RR-00-17

Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20(3):273–297. doi:10.1007/BF00994018

Dorigo M, Stützle T (2003) The ant colony optimization metaheuristic: Algorithms, applications, and advances. In: Handbook of metaheuristics. Springer, pp 250–285

Fister I, Yang X-S, Brest J (2013) A comprehensive review of firefly algorithms. Swarm Evol Comput 13:34–46

Friedrichs F, Igel C (2005) Evolutionary tuning of multiple SVM parameters. Neurocomputing 64:107–117

Gocić M, Motamedi S, Shamshirband S, Petković D, Ch S, Hashim R, Arif M (2015) Soft computing approaches for forecasting reference evapotranspiration. Comput Electron Agric 113:164–173

Govindaraju RS (2000) Artificial neural networks in hydrology. I: preliminary concepts. J Hydrol Eng 5:115–123

Huang C, Davis L, Townshend J (2002) An assessment of support vector machines for land cover classification. Int J Remote Sens 23:725–749

Jain P, Garibaldi JM, Hirst JD (2009) Supervised machine learning algorithms for protein structure classification. Comput Biol Chem 33:216–223

Jay S (2010) Built at sea: Marine management and the construction of marine spatial planning. Town Plann Rev 81:173–192

Ji Y, Sun S (2013) Multitask multiclass support vector machines: model and experiments. Pattern Recognit 46:914–924

Jiang X, Lu W, Zhao H, Yang Q, Chen M (2014) Quantitative evaluation of mining geo-environmental quality in Northeast China: comprehensive index method and support vector machine models. Environ Earth Sci 73(12):7945–7955. doi:10.1007/s12665-014-3953-7

Kaliraj S, Chandrasekar N, Magesh N (2014) Impacts of wave energy and littoral currents on shoreline erosion/accretion along the south-west coast of Kanyakumari, Tamil Nadu using DSAS and geospatial technology. Environ Earth Sci 71:4523–4542

Karunanithi N, Grenney WJ, Whitley D, Bovee K (1994) Neural networks for river flow prediction. J Comput Civ Eng 8:201–220

Khu ST, Liong SY, Babovic V, Madsen H, Muttil N (2001) Genetic programming and its application in real-time runoff forecasting 1 Jawra. J Am Water Resour Assoc 37:439–451

Kisi O, Shiri J, Karimi S, Shamshirband S, Motamedi S, Petković D, Hashim R (2015) A survey of water level fluctuation predicting in Urmia Lake using support vector machine with firefly algorithm. Appl Math Comput 270:731–743. doi:10.1016/j.amc.2015.08.085

Koza JR (1992) Genetic programming: on the programming of computers by means of natural selection vol 1. MIT press

Lee S-W, Verri A (2003) Support vector machines for computer vision and pattern recognition. World Scientific

Li L, Wu K, Zhou D-W (2014) Extraction algorithm of mining subsidence information on water area based on support vector machine. Environ Earth Sci 72:3991–4000

Lorena AC, De Carvalho AC (2008) Evolutionary tuning of SVM parameter values in multiclass problems. Neurocomputing 71:3326–3334

Lu W-Z, Wang W-J (2005) Potential assessment of the "support vector machine" method in forecasting ambient air pollutant trends. Chemosphere 59:693–701

Mahjoobi J, Adeli Mosabbeb E (2009) Prediction of significant wave height using regressive support vector machines. Ocean Eng 36:339–347. doi:10.1016/j.oceaneng.2009.01.001

Mahjoobi J, Etemad-Shahidi A, Kazeminezhad MH (2008) Hindcasting of wave parameters using different soft computing methods. Appl Ocean Res 30:28–36. doi:10.1016/j.apor.2008.03.002

Makarynskyy O (2004) Improving wave predictions with artificial neural networks. Ocean Eng 31:709–724. doi:10.1016/j.oceaneng.2003.05.003

Maresca S, Braca P, Horstmann J, Grasso R (2014) Maritime surveillance using multiple high-frequency surface-wave radars. IEEE Trans Geosci Remote Sens 52:5056–5071

Mukkamala S, Janoski G, Sung A (2002) Intrusion detection using neural networks and support vector machines. In: Proceedings of the 2002 International Joint Conference on Neural Networks, IEEE 2002. IJCNN'02, 2002, pp 1702–1707

Ornella L, Tapia E (2010) Supervised machine learning and heterotic classification of maize (*Zea mays* L.) using molecular marker data. Comput Electron Agric 74:250–257. doi:10.1016/j.compag.2010.08.013

Özger M (2010) Significant wave height forecasting using wavelet fuzzy logic approach. Ocean Eng 37:1443–1451. doi:10.1016/j.oceaneng.2010.07.009

Pal SK, Rai C, Singh AP (2012) Comparative study of firefly algorithm and particle swarm optimization for noisy non-linear optimization problems. Int J Intell Syst Appl (IJISA) 4:50

Rajasekaran S, Gayathri S, Lee T-L (2008) Support vector regression methodology for storm surge predictions. Ocean Eng 35:1578–1587

Ren F, Wu X, Zhang K, Niu R (2015) Application of wavelet analysis and a particle swarm-optimized support vector machine to predict the displacement of the Shuping landslide in the Three Gorges. China Environ Earth Sci 73:4791–4804

Shamshirband S et al (2014) Wind turbine power coefficient estimation by soft computing methodologies: comparative study. Energy Convers Manag 81:520–526

Shamshirband S, Mohammadi K, Tong CW, Zamani M, Motamedi S, Ch S (2015) A hybrid SVM-FFA method for prediction of monthly mean global solar radiation. Theor Appl Climatol 1–13. doi:10.1007/s00704-015-1482-2

Singh SK, Srivastava PK, Gupta M, Thakur JK, Mukherjee S (2014) Appraisal of land use/land cover of mangrove forest ecosystem using support vector machine. Environ Earth Sci 71:2245–2255

Singhal G, Panchang VG, Nelson JA (2013) Sensitivity assessment of wave heights to surface forcing in Cook Inlet. Alaska Cont Shelf Res 63:S50–S62

Sun S (2013) A survey of multi-view machine learning. Neural Comput Appl 23:2031–2038

Sung AH, Mukkamala S (2003) Identifying important features for intrusion detection using support vector machines and neural networks. In: Proceedings. 2003 Symposium on, IEEE Applications and the Internet, 2003, pp 209–216

Thirumalaiah K, Deo MC (2000) Hydrological forecasting using neural networks. J Hydrol Eng 5:180–189

Vapnik V (2000) The nature of statistical learning theory. Springer

Vapnik VN, Vapnik V (1998) Statistical learning theory, vol 2. Wiley, New York

Whitley D, Starkweather T, Bogart C (1990) Genetic algorithms and neural networks: optimizing connections and connectivity. Parallel Comput 14(3):347–361. doi:10.1016/0167-8191(90)90086-O

Wu K-P, Wang S-D (2009) Choosing the kernel parameters for support vector machines by the inter-cluster distance in the feature space. Pattern Recognit 42:710–717

Wu X, Ren F, Niu R (2014) Landslide susceptibility assessment using object mapping units, decision tree, and support vector machine models in the Three Gorges of China. Environ Earth Sci 71:4725–4738

Xu Y, Cheng C, Zhang Y, Zhang D (2014) Identification of algal blooms based on support vector machine classification in Haizhou Bay. East China Sea Environ Earth Sci 71:475–482

Yang XS (2009) Stochastic Algorithms: Foundations and Applications: In: Watanabe O, Zeugmann T (eds) 5th International Symposium, Sapporo, Japan, October 2009. Proceedings, Springer Berlin Heidelberg, pp 169–178. http://doi.org/10.1007/978-3-642-04944-6_14

Yang XS (2010) Research and Development in Intelligent Systems XXVI. In: Bramer M, Ellis R, Petridis M (eds) Incorporating Applications and Innovations in Intelligent Systems XVII, Springer London, pp 209–218 doi:10.1007/978-1-84882-983-1_15

Yang X-S (2013) Multiobjective firefly algorithm for continuous optimization. Eng Comput 29:175–184

Yang X-S, Deb S (2014) Cuckoo search: recent advances and applications. Neural Comput Appl 24:169–174

Yang H, Huang K, King I, Lyu MR (2009) Localized support vector regression for time series prediction. Neurocomputing 72:2659–2669