International Journal of Current Engineering and Technology
E-ISSN 2277 – 4106, P-ISSN 2347 – 5161
Available at http://inpressco.com/category/ijcet

Research Article

Significant Deep Wave Height Prediction by using Support Vector Machine Approach (Alexandria as case of study)

Tamer Elgohary[#*] and Amir Mubasher[!] and Hassan Salah^

[#]Civil Engineering Dept., Heliopolis University, Egypt
[!]Civil Engineering Dept., El Azhar University, Egypt
^Offshore Dept., ENPPI petroleum Co, Egypt

Accepted 02 Feb 2017, Available online 07 Feb 2017, **Vol.7, No.1 (Feb 2017)**

Abstract

The numerous increase in offshore operational activities demands improved wave forecasting techniques. If the accurate wave data is available, it is possible to carry out the marine activities easily and safely (i.e. offshore drilling, offshore platforms and pipelines installation, naval operations, near shore construction activities, etc.). This paper focuses on the prediction of significant wave heights (H_s) by support vector machine (SVM), using various kernel functions. This study aims to evaluate the influence of fetch and meteorological data over SVM approach and also perform a comparison between SVM kernel methods. Measured sea waves off Alexandria, west coast of Egypt, and meteorological data are used in this study. Six SVM (linear kernel) models comprising of various input combinations for wind speed, fetch, sea level pressure and air temperature have been performed to evaluate the wave height prediction performance of fetch and meteorological parameter. The results indicated that the SVM model (linear kernel function) gave the satisfactory results with all parameters (wind speed, fetch, sea level pressure and air temperature). Furthermore, the analysis showed that wind speed is the most important parameters for wave prediction. The results showed also that the fetch could also be useful for the wave height estimations, especially when used in combined with wind speed. Furthermore, the SVM kernels named sigmoid and a radial basis function (RBF) comprising all parameters were investigated. The results indicated that the SVM (linear kernel) gave the same results extracted from the SVM (sigmoid kernel). However, SVM (RBF kernel) gave the best prediction performance over other SVM kernels. Their results indicated that the error statistics of SVM models are generally within an acceptable range. Therefore, SVM can be used successfully for prediction of Hs.

Keywords: *Deep waves height, wave prediction, support vector machines, linear kernel function, radial basis function, Alexandria waves heights.*

1. Introduction

The waves are considering the most effective factor in many activities related to the ocean environment such as the offshore structures installation, port and terminal, maritime transportation and navigation, shoreline protection etc.

For this propose, there are several empirical and numerical methods described in literature, such as; SMB (Bretschneider, 1970), Wilson (Wilson, 1965), JONSWAP (Hasselmann *et al*., 1973), (Donelan, 1980), Shore Protection Manuel (SPM, 1984), Coastal Engineering Manuel (CEM, 2003), Kinsman (1965), World Meteorological Organization (WMO, 1988) and Goda (2003).

These methods aim to provide us the most suitable probability distributions and predicting the most appropriate wave characteristics, such as; wave height and wave period from fetch length and meteorological data. Usually; the empirical methods are developed based on the dimensionless parameters which are affecting on the wave generation. These simplified methods are particularly preferred for solving of the practical engineering problems. Numerical models are generally based on a form of the spectral energy or action balance equation. However, due to their complexity of implementation, high amount of processor time is required, and the need for accurate local bathymetric surveys, their implementation is not an easy task (Browne *et al*., 2007).

When the enormous amount of information is not available and the computational resources and expertise are limited, data mining and machine learning approaches would be very good choices (Mahjoobi and Mosabbeb, 2009). Empirical model is meant to learn and infer the behavior of the problem. The performance of this empirical model depends on quantity and quality of the data used. In traditional

*Corresponding author: **Tamer Elgohary**

statistical methods for problem solving, one needs to make some prior model assumptions, such as normality, linearity, .etc. On the other hand, the machine learning approaches these assumptions are not required. This is also expressed in the statistical learning theory (Vapnik, 1995) in other words, which tries to minimize the empirical risk of the model built over the data that minimizes the error on the training data. SVMs applications were developed to solve the classification problems, but recently they have been extended to the domain of regression problems (Vapnik, 1998). Support vector machines have been applied in many applications in the field of water engineering. Mohandes *et al*. (2004) used SVM for wind speed prediction and compared the results with multi-layer perceptron (MLP) neural network. They indicated that SVM outperforms MLP for their purpose. Another work on soil classification (Bhattacharya and Solomatine, 2006) used three different machine learning models namely ANN, SVM and decision trees. This work acquired nearly the same results for the tree models. Asefa *et al*. (2006) utilized SVM for multi-time scale stream flow predictions. They have achieved better results compared to those of physical models. Yu *et al*. (2006) also used SVM but for real-time flood stage forecasting and obtained satisfactory results. One important point that they have noted is that the SVM model is not easily understood and interpreted. This is one of the shortcomings of SVM in comparisons to traditional ANN. Also, SVMs are used for estimation of discharge and end depth in trapezoidal channel in Pal and Goel (2007). They noted that in comparisons to back propagation neural network, both radial basis function (RBF) and polynomial kernel-based approaches work better for different datasets. In addition, they have introduced a smaller computational time for SVM comparing with ANN.

A comprehensive review of support vector machine applications in ocean engineering was performed by J. Mahjoobi and Ehsan Adeli Mosabbeb (2009) and M. S. Elbisy (2013). They indicated in their paper that SVM can provide a good alternative to traditional statistical regression, numerical methods and approaches of this kind. The advantages are due to the improved accuracy, less complexity, smaller computational efforts and in several cases reduced data requirements. Mahjoobi and Mosabbeb (2009) used a support vector machine (SVM) to predict significant wave height. These authors compared the SVM results with those of NNs (Multilayer Perceptron and Radial Basis Function, or RBF) and reported that the SVM model was superior to the NN model. Moreover, this previous study concluded that using optimization algorithms, such as genetic algorithms (GAs), to select SVM parameters and choosing alternate kernel functions improves SVM model performance.

Elbisy, M.S. (2013) predicted the sea wave parameters by using SVM for different kernel functions. A genetic algorithms (GAs) was used in this study to determine the optimal values of the parameters for the different kernel functions of the SVMs and compared these results with those obtained from the field data and a Back-Propagation Neural Network (BPNN) and a Cascade-Correlation Neural Network (CCNN) models. The results showed that the SVM (RBF kernel) model out-performed the other methods. Furthermore, the SVM (RBF kernel) model has the highest accuracy and better generalization performance than the CCNN and BPNN models for all wave height and period ranges. The results obtained in this investigation demonstrated that the SVM (RBF kernel) model is a promising alternative to NN for wave parameter forecasting.

The effects of the meteorological factors such as; sea level pressure and air temperature were implicitly included in wind data measurement. Consequently, the evaluation of the effect of the meteorological data explicitly may be useful through the machine learning approach.

In this paper, the 3 hourly significant wave heights (H_s) were predicted from fetch data (F) and meteorological data such as wind speed (u), sea level pressure (p), and air temperature (C_a) based on hourly observations data by using the SVM approach (software package for predictive model) with different kernel functions.

This paper is organized as follows: the next section introduces method used in this study. Section 3 describes the studied area and data used. Section 4 presents the results of the SVM methods. Finally, conclusions are reported in the last section.

2. SVM Method

SVMs are methods of supervised learning, which are commonly used for classification and regression purposes. A Support Vector Machine (SVM) is a relatively modern approach that has shown great promise at generating accurate models for a variety of engineering problems. The original Support vector machine algorithm was created in 1963 by Vladimir N. Vapnik and Alexey Ya. Chervonenkis. In 1992, Bernhard E. Boser, Isabelle M. Guyon and Vladimir N. Vapnik suggested a technique to make nonlinear classifiers by applying the kernel trick to maximum-margin hyper planes (separating lines of data sets). The modern SVM depending on the soft margin was developed by Corinna Cortes and Vapnik in 1993 and published in 1995.

SVM models are built around a kernel function (Linear, Radial Basis Function (RBF), Sigmoid (S-shaped) and Polynomial) that transforms the input data into an n-dimensional space where a hyper plane can be constructed to partition the data.

Using SVMs requires an understanding of how they work. When training an SVM we need to make a number of decisions; how to preprocess the data, what kernel to use, and finally, selecting the parameters of the SVM and the kernel. In some cases, uninformed choices may be reduced the results performance.

A SVM constructs a separating hyper plane between the classes in the n-dimensional space of the inputs. This hyper plane maximizes the margin between the two data sets of the two input classes. The margin is defined as the distance between the two parallel hyper planes, on each side of the separating one, pushed against each of the two datasets. Simply, the largest margin gives the minimum error of the datasets.

Figure1 shows the distance between the dashed lines which is called the margin. The vectors (points) that constrain the width of the margin are the support vectors. An SVM analysis finds the hyper plane (i.e. a line) that is oriented so that the margin between the support vectors is maximized. In the figure below, the line in the right side is superior to the line in the left side. For the case of regression the only difference is that the SVM attempts to fit a curve, according to the kernel function applied, on the data points such that the points lie between the two marginal hyper planes to minimize the regression error.

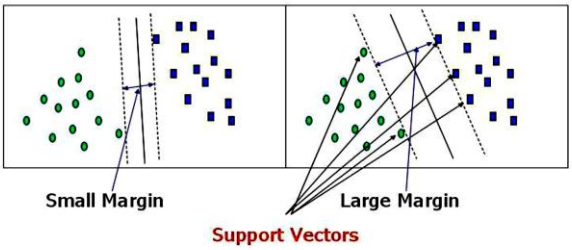

Fig.1 Illustrates Two-Dimensional Example

3. SVM Concepts

However, for simplification SVM concepts, we can describe SVM for two-class classification problems only. Assume that we have some data points, that each is a member of a class. The target is determining which class a new data point is belonging to. Hence the objective is maximizing the separation margin. It is expected that the data points will give several separating lines as shown in figure 2, but the one gives the largest margin of these data (see figure 3). This separating line (say hyper plane) in figure 3 is considered the best one relative to the separating lines in figure 2.

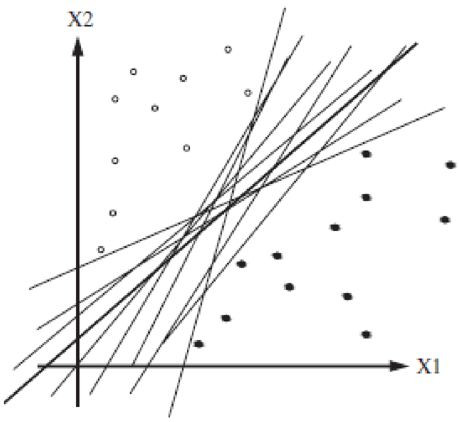

Fig.2 several possible separating lines (hyper planes)

So, the idea of support vector machine is to create a hyper plane in between data sets to indicate which class it belongs to.

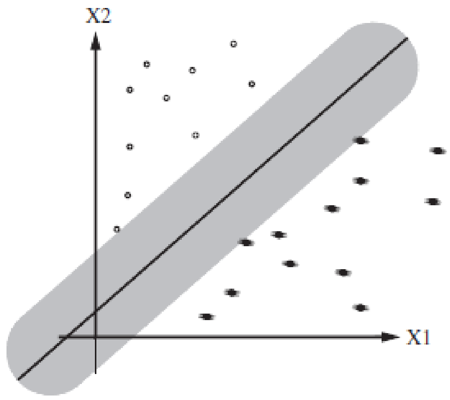

Fig.3 The maximum separating line (hyper plane)

For example, If we have a dataset of n points as(x_i, y_i), where (i= 1, 2, 3, ... n), $x_i \in R$ and $y_i \in \{+1,-1\}$. Any hyperplane can be written as the set of points x satisfying:

$$w.x_i + b = 0 \quad (1)$$

Where w is known as the weight vector and b is called the bias. The bias b translates the hyper plane away from the origin. The parameter $\frac{\mathbf{b}}{\|\mathbf{w}\|}$determines the offset of the hyper plane from the origin which is called "hard margin". If the data is found in the linearly separable case, with the hard margin, such as in figure 4, these hyper plane can be described by the equations:

$$w.x_i + b \geq +1 \; for \; y_i = +1 \quad (2)$$

$$w.x_i + b \leq -1 \; for \; y_i = -1 \quad (3)$$

Geometrically, the distance between these two hyper planes is $\frac{2}{\|w\|}$. Therefore to maximize the distance $\frac{1}{\|w\|}$ (between the origin line and one hyper plane), we want to minimize$\|w\|$ (for one hyper plane) or minimize the value $\|w\|^2$ (between two hyper planes). By combining of equations 2 and 3 in one set of inequalities, we can be rewritten as:

$$y_i(w.x_i + b) \geq +1, \qquad \text{for } i = 1, \dots. n \quad (4)$$

The SVM gets an optimal separating hyper plane with the maximum margin using the following optimization problem:

$$\text{Minimize } \frac{1}{2}\|w\|^2 \quad \text{subjected to: } y_i(w.x_i + b) \geq +1 \quad (5)$$

In figure 4, the data points lying on the marginal hyper planes are called the support vectors.

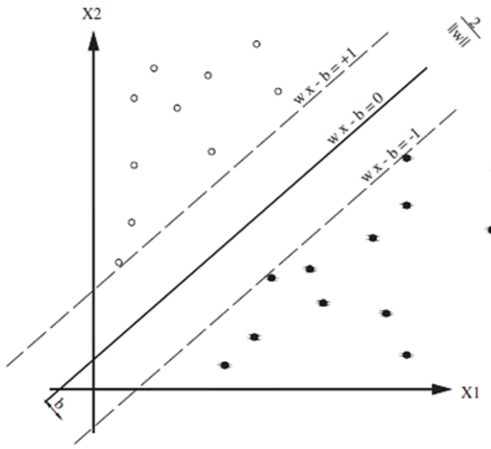

Fig.4 The linearly separable case (hard margin) and the support vectors points

To extend SVM to covers also the data in non-linearly separable cases in an attempt to build a hyper plane with the smallest number of errors. Therefore the non-negative slack variables ($\xi_i \geq 0,\ i = 1, \ldots, n$) were introduced by (Haykin, 1999) to obtain the following formal setting of this problem:

$$w.x_i + b \geq +1 - \xi_i \quad \text{for } y_i = +1 \qquad (6)$$

$$w.x_i + b \leq -1 + \xi_i \quad \text{for } y_i = -1 \qquad (7)$$

where $\xi_i \geq 0$is slack variables that allow an example to be in the margin ($0 \leq \xi_i \leq 1$, also called a margin error) or to be misclassified ($\xi_i > 1$). Since an example is misclassified if the value of its slack variable is greater than 1, $\sum_i \xi_i$ is a bound on the number of misclassified examples. Our objective of maximizing the margin, i.e. minimizing $\frac{1}{2}\|\mathrm{w}\|^2$ will be amplified with a term $\mathrm{C}\sum_i \xi_i$ to penalize misclassification and margin errors (where C and ξ are SVM parameters).

The optimization problem now is:

$$\text{Minimize } \frac{1}{2}\|\mathrm{w}\|^2 + \mathrm{C}\sum_{i=1}^{n} \xi_i \qquad (8)$$

subjected to: $y_i(w.x_i + b) \geq 1 - \xi_i\ ,\xi_i \geq 0$ and $\mathrm{C} > 0$

This formulation (equation 8) is called the soft-margin SVM, and was introduced by Boser *et al.*, 1992 and Cortes and Vapnik, 1995.

Where C is the parameter determines the tradeoff between increasing in the margin-size and the ensuring that the x_i lie on the correct side of the margin. Therefore, for sufficiently small values of C, the soft-margin SVM will behave identically to the hard-margin SVM if the input data are linearly classifiable, but will still learn a feasible classification rule if not.

By using the Lagrange multipliers, we can find the dual formulation which is expressed in terms of variables α_i (also known as the dual representation of the decision boundary):

$$\text{Minimize } \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\sum_{i=1}^{n}\ \sum_{j=1}^{n} \alpha_i\alpha_j y_i y_j \phi(x_i)^T\phi(x_j) \qquad (9)$$

subjected to: $\sum_{i=1}^{n} \alpha_i y_i = 0,\ 0 \leq\ \alpha_i \leq \mathrm{C}$

where $\phi(.)$ denotes a set of nonlinear transformation between the input space and the feature apace. So we can define the kernel function as $\mathrm{K}(\mathrm{x_i}, \mathrm{x_j}) = \phi(\mathrm{x_i})^{\mathrm{T}}\phi(\mathrm{x_j})$, which is induced by Mercer's theorem (Haykin, 1999). Accordingly, Kernel functions are used to compute a non-linearly separable function and then transform into a higher dimension linearly separable function (see figure 5).

So the problem is changed to be:

$$\text{Minimize } \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\sum_{i=1}^{n}\ \sum_{j=1}^{n} \alpha_i\alpha_j y_i y_j \mathrm{K}(\mathrm{x_i}, \mathrm{x_j}) \quad (10)$$

Kernel functions are used to change the dimensionality of the input space, in order to perform the classification (or regression) task with more confidence. Some common kernel functions are used:

- Linear (homogeneous) : $K(x_i, x_j) = x_i x_j$
- Polynomial(inhomogeneous: $K(x_i, x_j) = (x_i x_j + 1)^p$
- Radial basis function (Gaussian): $K(x_i, x_j) = \exp\left(-\gamma\ \|\mathrm{x_i} - \mathrm{x_j}\|^2\right)$ for $\gamma > 0$
- Sigmoid (Hyperbolic tangent): $K(x_i, x_j) = \tanh(\gamma\ x_i x_j + \nu)$ for $\gamma > 0$ and $\nu \leq 0$

Where γ, p and ν are kernel parameters

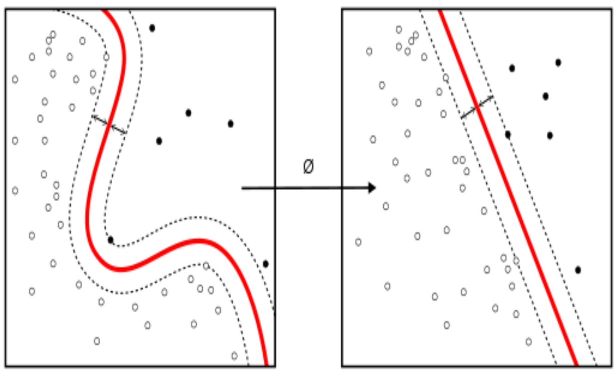

Fig.5 The kernel functions transform nonlinear into linear by ϕ

3. Parameters Selection

It is well-known that SVM accuracy depends on a selection of SVM parameters C and ε and kernel parameters γ, p and ν. The choices of C and ξ control the prediction (regression) model complexity. The problem of optimal parameter selection is further complicated by the fact that SVM model complexity depends on above parameters (Smola and Scholkopf, 1998). The penalty constant C is a positive constant that can be close to infinity; however, C =1000 and slack variable (or relaxation factor) ξ = 0.0001 are adequate for many operations (Mao *et al*., 2005 and Elbisy, M.S. 2015). In 2009, Mahjoobi and Mosabbeb used SVM parameters C =100, ξ = 0.001 and kernel parameters p =1.0, γ = 0.01 in significant wave height prediction.

The assumption of SVM parameters will be used in this study are C =500, ξ = 0.0001 and for kernel parameters γ =5, p =20 and ν=0.

4. Study area and data

The meteorological, fetch and wave data were gathered from deep water location in Ras El-Teen coastal zone, open sea area, located at the west of Alexandria on the North West Egyptian Nile delta coast (Figure 6).

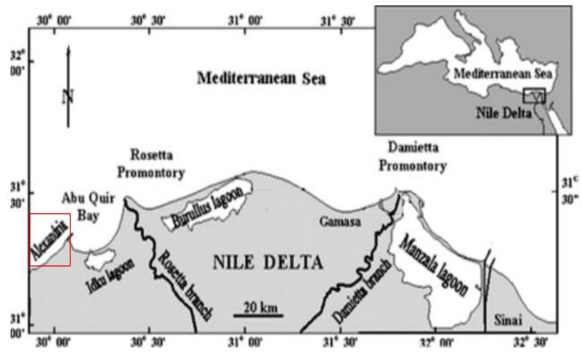

Fig.6 The Study Area Location

The data used in this study was measured from 1 January, 2010 to 31 December, 2012. The offshore data set was collected from S4DWI buoy (31° 10' 60" N and 29° 50' 00" E), where the buoy is owned and maintained by Egyptian Navy Forces, Meteorological and Oceanographic Division, through the specialized people in this area.

Since the water depth and wave direction does not affect the form of waves in deep water zone, so they are not used in this study. The 3 hourly significant wave heights (H_s) accompanied with wind speed (u), fetch data (F), sea level pressure (p), and air temperature (C_a) used in this study based on hourly observations data were provided by the buoy mentioned above.

The table represents the values of the minimum, maximum, and mean values of parameters u, F, p, C_a and H_s of data sets for 3 years (from 2010 to 2012).

Table 1 Min, Max. and Mean values for datasets

Parameter	Min. value	Max. value	Mean value
u (m/s)	0.00	23.66	5.90
F (m)	0.00	5720	2826.4
p (Kg/m2)	10012	10286	10128.3
C_a (°C)	9.60	40.20	21.90
H_s (m)	0.00	15.02	1.15

6. Results and discussions

6.1 Models accuracy

For comparison of models accuracy correlation coefficient (R), mean square error (MSE), mean absolute error (MAE) and scatter index (SI) are used. These statistical measures are defined as follows:-

Correlation coefficient (R)

$$R = \frac{\sum_{i=1}^{n} (P_i - \bar{P})(O_i - \bar{O})}{\sqrt{\sum_{i=1}^{n}(P_i - \bar{P})^2 \sum_{i=1}^{n}(O_i - \bar{O})^2}} \tag{1}$$

Mean Square Error (MSE)

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(P_i - O_i)^2 \tag{2}$$

Mean Absolute Error (MAE)

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|(P_i - O_i)| \tag{3}$$

Scatter Index

$$SI = \frac{RMSE}{\bar{O}} \tag{4}$$

In all formulas, the O_i's represent the observation value, the P_i's represent the predicted value, n is the total number of observations, $\bar{O}$ is the mean of O_i and $\bar{P}$ is the mean of *Pi*.

6.2 Results of SVM (Linear)

Six combinations of the u, F, p and C_a parameters were used as inputs to predict H_s using SVM approach, linear kernel function, in an attempt to evaluate the parameters importance, impact on the results accuracy and models performance. The parameters combinations are illustrated in Table 2.

Table 2 Parameters combinations used in the prediction of H_s

Model No.	Wind speed u (m/s)	Fetch F (m)	Sea level pressure p (Kg/m2)	Air temperature C_a (°C)
1	Yes	-	-	-
2	-	Yes	-	-
3	-	-	Yes	-
4	-	-	-	Yes
5	Yes	Yes	-	-
6	Yes	Yes	Yes	Yes

The R, MSE (m), MAE (m) and SI (%) values of the models results for data used in this study (3 years) are shown in Table 3.

The first four models, in table 3 above, were performed to evaluate the prediction performance of H_s relative to each parameter (u, F, p and C_a) separately.

Table 3 Results values of statistical measures for each model

Model No.	Inputs	R	MSE (m)	MAE (m)	SI (%)	SVM
1	u	0.940	0.152	0.214	12.6	Linear
2	F	0.866	0.300	0.274	24.9	Linear
3	p	0.143	1.181	0.731	98.09	Linear
4	C_a	0.165	1.184	0.754	98.3	Linear
5	u, F	0.955	0.121	0.171	10.05	Linear
6	u, F, p, C_a	0.957	0.111	0.180	9.20	Linear

For these models; the results of R, MSE, MAE and SI% revealed that the performance of model (1), with the input of wind speed (u) only, gave the best performance comparing with the other models (2), (3) and (4). This indicates that the wind action is the dominant factor in the wind-waves growth as a fact.

Also, the performance of the fetch data used in model (2), when used as a single input, was fairly close to model (1) (e.g. the correlation value of model (2) equal 0.866 relative to correlation value of model (1) equal 0.940). So, the reasonable results of fetch model (2) indicated that the fetch could also be useful in the wave height prediction. It is obvious that the performance of model (5) did not significantly different the performance of model (1) (R equal to 0.955 and 0.940 for models (5) and (1) respectively) where the effect of the fetch located in deep water area (open sea) is not valuable relative to, which is located in shallow water area, as well as the confined water areas such as the lakes and bays.

The results from model (3), with sea level pressure (p) only, and also model (4), with Air temperature (C_a) only, showed that p and C_a are not effective on the H_s prediction using SVM method when used as individual input. As shown in table 3, model (3) and (4) give smallest correlation values R (0.143 and 0.165 for model (3) and (4) respectively) and highest values of MSE, MAE and SI%. Accordingly, these results showed that the models (3) and (4) have poorer performance than models (1) and (2). Although the sea level pressure and the Air temperature did not affect significantly the performance of models (3) and (4) where it is implicitly included in the measurements of the waves and wind speed, but they increase slightly the accuracy of model when used in participation with the fetch and wind speed as will be shown later in the results of model (5).

Based on the satisfactory results from models (1) and (2), wind speed (u) and fetch (F) will be used as inputs to build model (5). The value of correlation for model (5) is equal (0.955) which is higher than models from (1) and (2). Also, the values of MSE, MARE and SI% values gave smaller values than the first two models, which means that the model (5) has slightly better performance than model (1), with wind speed only, while it increased the prediction performance comparing with model (2), with fetch only.

Model (6) was performed to evaluate the performance of all parameters together. The results revealed that the model (6) gave the highest value in the correlation factor (0.957) and the smallest values for MSE, MAE and SI% comparing with the first five models stated in table 3.From the above results, it can be concluded that using the combinations of u, F, p and C_a increased the prediction performance of model. The model (5) used u and F in the input data provided results values nearly identical for model (1). Sea level pressure (p) and air temperature (C_a) have a minor effective on the wave height prediction when used as single input such as in models (3) and (4) respectively, however if they used in combination such as in model (6) they increased the model accuracy. Model (6) whose input data consists of u, F, p and C_a provided the best prediction performance. Figures 7 (a ~ d) present the correlation between the actual (observed) and the predicted values of Hs for data used in models (1), (2), (5) and (6) respectively (also called a Residual Chart). In models (3) and (4) respectively the graphs showed that both sea level pressure (p) and air temperature (Ca) give the poor prediction performance for significant wave height (Hs) when used, each one individually, as a single input. Accordingly; they were evaluated as out of preferable results.

In Figures 7 (a ~ d) generally, the X coordinate of a point is the actual target (observed) value and Y coordinate of the point is the corresponding predicted target value. The points (in blue) are offset from the diagonal line show the different for some of the predicted values against the actual values, and the vertical distance (in red) from the line to the point corresponds to the error (residual).

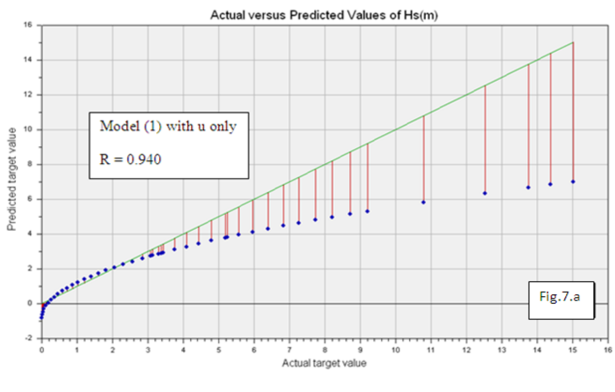

Fig.7.a for Actual and Predicted values of Hs (Model 1)

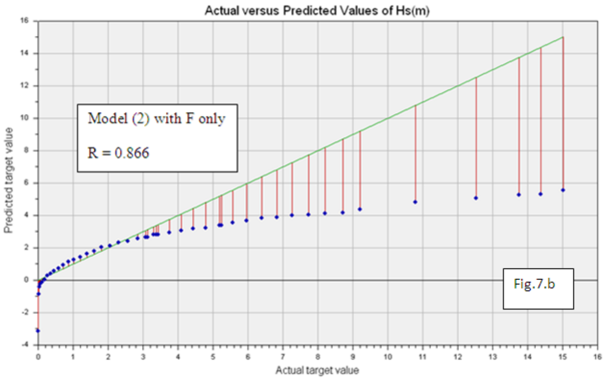

Fig.7.b for Actual and Predicted values of Hs (Model 2)

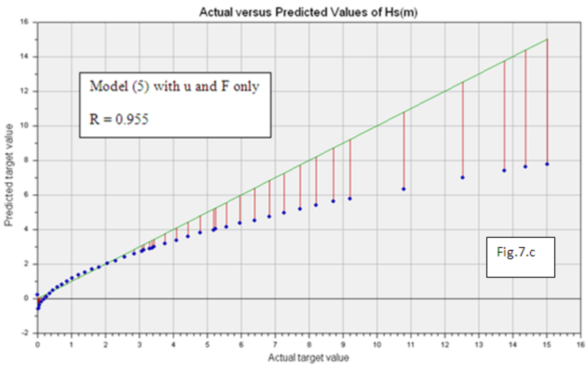

Fig.7.c for Actual and Predicted values of Hs (Model 5)

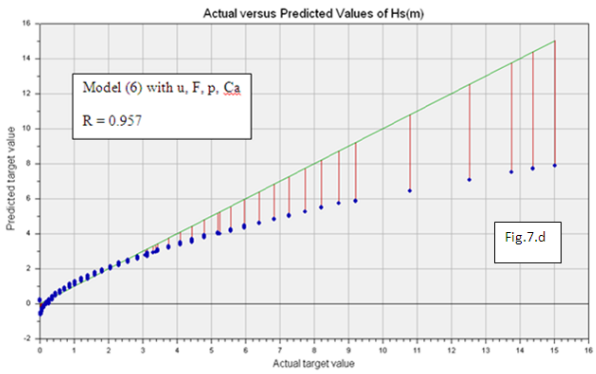

Fig.7.d for Actual and Predicted values of Hs (Model 6)

From Figures 7 (a ~ d), it is obvious that the correlation between the observed and the predicted values of Hs for all represented models decreased while the significant wave height (Hs) increased, especially for Hs values more than 5.0 m. Correlations in figures (7.a), (7.c) and (7.d) gave the nearly models results, however model (1) in figure (7.a) was slightly less performance than models (5) and (6) in figures (7.c) and (7.d) respectively.

Figure 8 presents the overall importance of parameters for model (6) relative to Hs using the linear SVM method. The results revealed that the weight of fetch relative to prediction of Hs is about 8% while the sea level pressure and air temperature are equal 0.23% and 0.01 % respectively. Also, the graph confirmed that the wind speed is the important factor can be affected in the prediction of the significant wave height (Hs)

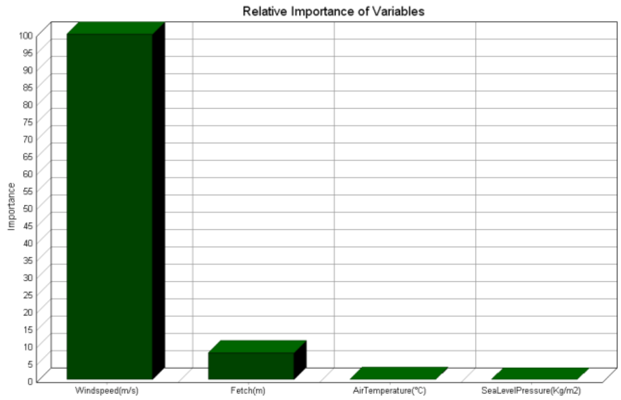

Fig.8 parameters importance in linear SVM model (6)

The results of R, MSE, MAE and SI%, stated in table 4, showed that the performance of model (7) using sigmoid kernel function gave the same performance for model (6) with the linear kernel function, while the results of model (8) using RBF kernel function gave the better performance than models (6) and (7).

Table 4 Results values of statistical measures for Sigmoid and RBF

Model No.	Inputs	R	MSE (m)	MAE (m)	SI (%)	SVM Kernel Function
7	u, F, p, C_a	0.957	0.111	0.180	9.20	Sigmoid
8	u, F, p, C_a	0.999	0.000003	0.0013	0.0002	RBF

Correlation factor "R" in model (8) gave the highest value of "R" relative to all previous models, where it closed to 1.00 and also gave the smallest values of MSE, MAE and SI%, where they closed to zero. Accordingly, model (8) indicated that the SVM with RBF gave the superior performance which is more near from the measured data.

Figures (8.a) and (8.b) present the actual (observed) versus predicted Hs values for data used in models (7) and (8) respectively.

Figure (8.a) gave the same results of correlation value such as value in figure (7.d) stated in section 4.2, which means that the SVM with sigmoid kernel function does not improve the model accuracy over SVM with linear kernel function.

Figure (8.b) showed that the correlation between the observed and the predicted values of Hs close to 1.00 when using SVM with RBF. Therefore, model (8) is considered as the perfect model where almost of all points located on the diagonal line (in green).

Furthermore, a comparison of the correlation (R) values between the three models (6), (7) and (8) illustrated in figure (9) hereafter. A comparison shows that the highest R in the three models (linear, sigmoid and RBF) is observed with the RBF in model (8).

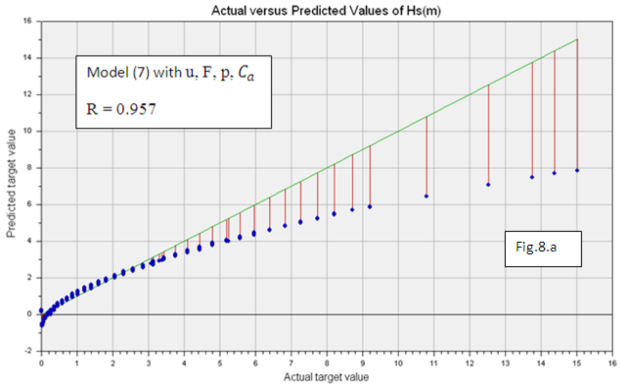

Fig.8.a for Actual and Predicted values of Hs (Model 7)

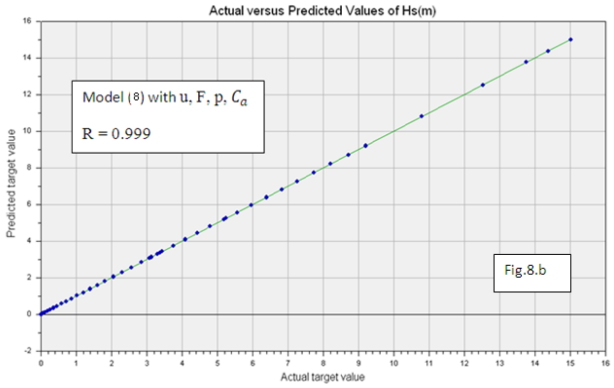

Fig.8.b for Actual and Predicted values of Hs (Model 8)

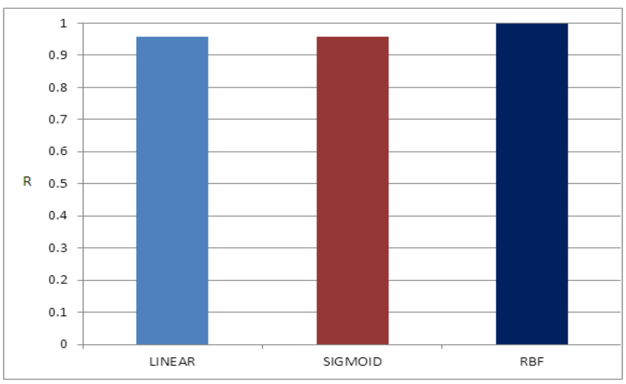

Fig.9 comparison of the correlation (R) of the linear, sigmoid and RBF models

Conclusions

Significant wave height prediction is an essential step to study of many projects whether in the offshore or in the coastal area. In this study, the prediction of 3 hourly significant wave heights (Hs) based on fetch data and the meteorological data named as u, p and Ca was investigated by using SVM with linear, sigmoid and radial basis function methods. It was concluded from the six models of SVM (linear kernel function) that:

1) The best prediction performance was obtained from the SVM model (6) including fetch and all meteorological data as the inputs.
2) The results of model (1) confirmed that the wind speed is the most effective factor on wave height prediction which gave the satisfactory results.
3) On the other hand, models (3) and (4) revealed that the mean sea level pressure (p) and air temperature (Ca) were not effective on wave height prediction in case of used each one as single input in the model. Nevertheless, when add p and Ca into input combination with fetch and wind speed data, the performance of SVM model increased.
4) The fetch data showed good prediction performance for both single and two input cases in models (5) and (6) and could be useful for wave height prediction.

   Based on the results from SVM (linear kernel), all parameters were used during the new SVM kernel functions named sigmoid and RBF. The comparison between the three models concluded that:
5) The results from models (6) and (7) showed that there is no difference in the results between SVM linear method and SVM sigmoid method.
6) The results from model (8) for SVM RBF method gave the best results for all model combinations used in this study.

Finally, this study has confirmed the following:

7) SVM linear and sigmoid approaches could not gave good results with the high wave ranges and it works well in the small and medium wave height ranges.
8) SVM linear approach could not gave good results with the high wave ranges, but it works well with the small and medium wave height ranges. On the contrary, SVM RBF kernel approach gives a very good results with any range of wave height.
9) The SVM (RBF kernel) in the wave height prediction gave a superior performance comparing with the SVM (linear and sigmoid kernel). SVM (RBF kernel) has the highest accuracy as well as gave a lower predicting error.

Recommendations

1) Theoretical analysis study for deep wave prediction in Alexandria region should be achieved to compare it with support vector machine approach methods.
2) Both Theoretical analysis and support vector machine approach methods for shallow wave

prediction in Alexandria region should be achieved.

3) Study the effect of sea water salinity on deep wave prediction should be achieved.
4) Study the effect of sea water level rise due to climate changes on deep wave prediction should be achieved.

References

Bretschneider, C.L., 1970. Wave forecasting relations for wave generation. Look Lab/Hawaii, 1(3), 31–34.

Wilson, B.W., 1965. Numerical prediction of ocean waves in the North Atlantic for December 1959. Deutsche Hydrographische Zeitschrift 18 (3), 114–130.

Hasselmann, K., Barnett, T.P., Bouws, E., Carlson, H., Cartwright, D.E., Enke, K., Weing, J.A., Gienapp, H., Hasselmann, D.E., Kruseman, P., Meerburg, A., Muller, P., Olbers, K.J., Richter, K., Sell, W., Walden, W.H., 1973. Measurements of wind wave growth and swell decay during the Joint North Sea Wave Project (JONSWAP). Deutsche Hydrograph, Zeit, Erganzung-self Reihe, A 8 (12).

Donelan, M.A., 1980. Similarity theory applied to the forecasting of wave heights, periods and directions. Proceedings of the Canadian Coastal Conference (National Research Council of Canada), pp. 47– 61.

SPM, 1984. Shore Protection Manual. Coastal Engineering Research Center, Waterway Experiment Station, Corps of Engineers, Department of the Army, Vicksburg, MS.

CEM, 2003. Coastal Engineering Manual. US Army Engineer Research and Development Center, Coastal and Hydraulics Laboratory, Vicksburg, MS.

Kinsman, B., 1965. Wind Waves. Prentice-Hall, Englewood Cliffs, NJ, USA. Lee, .L., 2006. Neural network prediction of a storm surge. Ocean Engineering 33, 483–494.

WMO, 1988. Guide to wave analysis and forecasting. World Meteorological Organization WMO no. 702, Geneva, Switzerland.

Goda, Y., 2003. Revisiting Wilson's formulas for simplified wind-wave prediction. Journal of Waterway, Port, Coastal and Ocean Engineering 129 (2), 93–95.

Browne, M., Castelle, B., Strauss, D., Tomilnson, R., Blumenstein, M., Lane, C., 2007. Near-shore swell estimation from a global wind–wave model: spectral process, linear and artificial neural network models. Coastal Eng. 54, 445–460.

J. Mahjoobi and EhsanAdeliMosabbeb, 2009. Prediction of significant wave height using regressive support vector machines. Ocean Engineering 36 (2009) 339–347. Journal ELSEVIER.

Vapnik, V., 1995. The Nature of Statistical Learning Theory. Springer, NY.

Vapnik, V., 1998. Statistical Learning Theory. Springer, NY.

Mohandes, M.A., Halawani, T.O., Rehman, S., Hussain, Ahmed A., 2004. Support vector machines for wind speed prediction. Renewable Energy 29, 939–947.

Agrawal, J.D., Deo, M.C., 2002. On-line wave prediction. Mar. Struct. 15, 57–74. Bhattacharya, B., Solomatine, D.P., 2006. Machine learning in soil classification. Neural Networks 19, 186–195.

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.

Yu, P.S., Chen, S.T., Chang, I.F., 2006. Support vector regression for real-time flood stage forecasting. J. Hydrology 328, 704–716.

Pal, M., Goel, A., 2007. Estimation of discharge and end depth in trapezoidal channel by support vector machines. Water Resour. Manage. 21, 1763–1780.

Moussa S. Elbisy, 2015. Sea Wave Parameters Prediction by Support Vector Machine Using a Genetic Algorithm. Journal of Coastal Research: Volume 31, Issue 4: pp. 892–899.

Boser, B. E.; Guyon, I. M.; Vapnik, V. N. (1992). "A training algorithm for optimal margin classifiers". Proceedings of the fifth annual workshop on Computational learning theory – COLT '92. p. 144.

Cortes, C.; Vapnik, V. (1995). "Support-vector networks". Machine Learning. 20 (3): 273–297.

Smola, A.J., Scholkopf, B.,1998. A Tutorial on Support Vector Regression. Royal Holloway College,London,UK,NeuroCOLTTech.Rep.TR1998-030.

Mao, M.; Zhou, X.B.; Pi, D.Y.; Sun, Y.X., and Wong, S.T., 2005. Parameters selection in gene selection using Gaussian kernel support vector machines by genetic algorithm. Journal of Zhejiang University Science, 6(10), 961–973.

Shai Shalev-Shwartz and Shai Ben-David, 2014. Understanding Machine Learning From Theory to Algorithms. Cambridge University.

Vladimir Cherkassky and Yunqian Ma, 2003. Practical Selection of SVM Parameters and Noise Estimation for SVM Regression. University of Minnesota, USA.

BS 6349-1:2000. Maritime structures Part 1: Code of practice for general criteria

Chih-Wei Hsu, Chih-Chung Chang, and Chih-Jen Lin, 2010. A Practical Guide to Support Vector Classification. National Taiwan University, Taiwan.

Ben-Hur and Weston, 2007. A User's Guide to Support Vector Machines. Colorado State University and NEC Labs America Princeton, USA.

Kemal Gunaydın, 2008. The estimation of monthly mean significant wave heights by using artificialneural network and regression methods. Ocean Engineering 35 (2008) 1406–1415. Journal Elsevier.