Significant Wave Height Forecasting using GMDH Model

Sajad Shahabi Ph. D Candidate civil Engineering Department Shahid Bahonar University of Kerman Mohammad-Javad Khanjani Professor civil Engineering Department Shahid Bahonar University of Kerman Masoud-reza Hessami Kermani Assistant Professor civil Engineering Department Shahid Bahonar University of Kerman

ABSTRACT

Forecasting of significant wave height (SWH) is necessary for most of ocean engineering activities. Different models have been applied to forecast SWH at various lead times. Here, group method of data handling as a data learning machine method is used to forecast the SWH for next 3, 6 and 12. The SWH data are collected from station 41036 located in the North Atlantic Ocean. The model performance was evaluated using three different index including root mean square error (RMSE), coefficient of correlation (R) and index of agreement (I_a). The results shows that in short lead times, the predicted significant wave height mostly correlated to the observed significant wave height but in larger lead times this correlation decreased.

Keywords

GMDH, lead time, significant wave height, time series.

1. INTRODUCTION

Chakrabarti (1987) reported that the concept of significant wave height (SWH) was first introduced by Sverdrup and Munk as the average of the highest one-third of all waves in a special location of sea. It shows by $H_{1/3}$ or H_s . It is to be close to the height reported by sea captains through visual observations.

Prediction of SWH is an important issue for human activities in coastal zone. For instance port design and operability, dispersion and diffusion of pollutants require an exact knowledge of predicted SWH. Accurate forecasting of wave properties is very important for most of coastal activities.

It seems that the time series of SWH can be modeled as a random process, but it has some correlations that can be exploited to extrapolate the future from its past values (Nitsure et al., 2012). Therefore, Many SWH forecasting models have been developed in recent years: see (Shahabi and Khanjani, 2015; Shahabi et. al., 2016).

In last decade, the soft computing methods are the most used methods. One of the most common of these methods is artificial neural network (ANN). ANN has been applied extensively in coastal and ocean engineering (Deo and Naida, 1999; Deo et al., 2001; Tsai et al., 2002; Londhe and Panchang, 2006).

In recent years, some other artificial intelligence techniques such as genetic algorithm (GA), genetic programming (GP) and fuzzy logic (FL) have been provided for modeling and forecasting the SWH (Ozger and Sen, 2007; Gaur and Deo, 2008; Altunkayank., 2015).

Several time series based model such as autoregressive (AR), autoregressive moving average (ARMA), autoregressive integrated moving average (ARIMA) and Kalman filter were used to SWH forecasting (Ozger, 2010; Altunkayank and Wang, 2012).

Some other models, used for wave height forecasting, such as regressive support vector machine, fuzzy interface system and genetic algorithm (GA) were described by Kazeminezhad et al. (2005), Mahjoobi and Mosabeb (2009) and Canellas et al. (2010).

The group method of data handling (GMDH) was used in recent decades, extensively. Zhang et al. (2013) were employed an improved GMDH model to predict of debris flow by integrating dynamic rainfall data and environmental factors. Their results showed that the GMDH for prediction of debris flow performed well using the training, validation, and testing sets. In other hand, some other researchers used a hybridized GMDH model (Zhu et al., 2012; Atashrouz et al., 2014).

Literature review on forecasting of SWH shows that, while prediction models give satisfactory results up to 6 hour lead time, the accuracy falls remarkably for higher lead times

Therefore, this study is intended to illustrate an evolutionary method for forecasting of SWH. This method can be used for every time series of SWH. For this purpose, a GMDH model was employed as a predictor method.

2. GMDH THEORY

The GMDH is a learning machine based on the principle of exploratory self- organizing (Ivakhnenko, 1968).

A schematic diagram of this model was shown in Fig. 1. The GMDH can be combined by other evolutionary or artificial intelligence models (Shahabi et al., 2016).



Figure 1. A schematic diagram of GMDH model

The output of the GMDH model (the predicted amount) is a vector such as y_{sim} based on some vectors such as $(x_1, x_2, ..., x_n)$. The input vectors can be written as:

 $y_i = f(x_{i1}, x_{i2}, \dots, x_{ij})$, $i = 1, 2, 3, \dots, k$ (1)

And Eq. (2) presents the forecasted output.

$$y_{sim} = f(x_{i1}, x_{i2}, \dots, x_{ij})$$
 , $i=1,2,3,\dots,k$ (2)

The GMDH model is explored a relationship between input and output vectors. It selects the best relationship based on minimizing the square of difference between forecasted and observed data. That means:

$$\sum_{i=1}^{k} (f_{sim}(x_{i1}, x_{i2}, \dots, x_{ij}) - y_i)^2 \Rightarrow Min$$
⁽³⁾

The relationship between input and output data in GMDH model can be described by Eq. (4).

$$y = w_0 + \sum_{i=1}^n w_i x_i + \sum_{i=1}^n \sum_{j=1}^n w_{ij} x_i x_j + \cdots$$
(4)

where w is the weighting term.

Here, the quadratic polynomial of GMDH is used such as predictors function and described by:

$$y_{simn} = H(x_i, y_j) = w_0 + w_1 x_i + w_2 x_j + w_3 x_i x_j + w_4 x_i^2 + x_5 x_j^2$$
(5)

In this study determination coefficient are used to assessment of the GMDH model in every step.

3. STUDY AREA AND DATA

The data used in this study are significant wave height (SWH) of the stations 41036 (Latitude 34°12′25″ and longitude 76°56′56″) located in North Atlantic Ocean as shown in Fig. 2 which were downloaded from NOAA's National Data Buoy Center (http://www.ndbc.noaa.gov/historical_data.shtml).



Figure 2. Location map of the study area

The SWH time series are divided into two parts, training and testing data. The statistical properties of all data set including full data time series, training and testing data set were presented in Table 1, separately.

Table 1. Statistical properties of all data

Station	Data	Significant wave height		
ID		Min. (m)	Max. (m)	Mean (m)
41036	Full series	0.27	5.18	1.24
	Training data	0.27	4.7	1.22
	Testing data	0.32	5.18	1.29

4. RESULTS AND DISCUSSION

An exploratory model called GMDH was utilized to predict the significant wave height for station 41036 located in the North Atlantic Ocean off the east coast of the USA as shown in Fig. 2.

Table 2. shows a summary of results for the employmentof the GMDH approach for station 41036.

Lead		RMSE	R	Ia
time				
3h	Training data	0.1937	0.9395	0.9680
	Testing data	0.2149	0.9429	0.9699
6h	Training data	0.3041	0.8433	0.9095
	Testing data	0.3426	0.8468	0.9123
12h	Training data	0.4425	0.6239	0.7377
	Testing data	0.4993	0.6306	0.7452

Assessment of models is made using scatter plots between the observed and predicted SWH. The figures 3-8 indicate satisfactory predictions in lower lead times. In addition, root mean square (RMSE), correlation coefficient (R) and index of agreement are employed to evaluate of the model. These indices are shown in Table 2.



Figure 3. Scatter plot of observed and predicted SWH by GMDH model at 3h lead time (training data)



Figure 4. Scatter plot of observed and predicted SWH by GMDH model at 6h lead time (training data)



Figure 5. Scatter plot of observed and predicted SWH by GMDH model at 12h lead time (training data)

Fig.6 and Fig. 7 indicate very good accuracy of SWH estimation by the GMDH model at 3 and 6h lead times. The correlation coefficient is 0.94 and 0.85, respectively. In other hand the index of agreement decreased from 0.97 to 0.91 for 3 and 6h lead times.



Figure 6. Scatter plot of observed and predicted SWH by GMDH model at 3h lead time (testing data)



Figure 7. Scatter plot of observed and predicted SWH by GMDH model at 6h lead time (testing data)

As seen in Fig. 8, the scatter plot shows that the predicted SWH at 12h lead time is not bad although the model is more accurate in 3 and 6h lead times.

The RMSE, R and I_a at 3, 6 and 12h lead times indicate acceptable accuracy of predicted SWH, especially in 3 and 6h lead times. Table 2 summarizes these three indices for all lead times at station 41036.



Figure 8. Scatter plot of observed and predicted SWH by GMDH model at 12h lead time (testing data)

In summary, it can be concluded that the GMDH model can be employed to forecast SWH at lower lead times, satisfactory. But in higher lead times it appears that a hybrid model can be presented a more acceptable results.

5. CONCLUSIONS

Here, a GMDH model have been developed and performed to forecast significant wave height (SWH) for 3, 6 and 12 h lead times at west of North Atlantic Ocean. The model performance was assessed using three different indices. The model predictions in 3 and 6 h lead times are better than 12 h lead time. The index of agreement values decreased from 0.970 to 0.745 for 3 to 12h lead times. In addition, the correlation coefficient of the model changed between 0.943 to 0.631 for mentioned lead times. The GMDH model obtained results in shorter lead times were more accurate than the larger lead times such as similar studies. Also, this method can be simply used for planning of similar problem in the field of water resources engineering.

6. REFERENCES

- Nitsure, S. P., Londhe, S. N., and Khare, K. C, "Wave forecasts using wind information and genetic programming", Ocean Engineering, 2012, 54, 61-69.
- [2] Chakrabarti, S. K. Hydrodynamics of offshore structures. Southampton, Boston: Computational Mechanics Publications, (Chapter 5). 1987.
- [3] Shahabi, S., Khanjani, M. J., and Kermani, M. R. H., "Hybrid wavelet-GMDH model to forecast significant wave height", Water Science and Technology: Water Supply, 2016, In Press.
- [4] Ivakhnenko, A. G. The group method of data handling (GMDH). Automation. 1968. 3, 57-83.
- [5] Shahabi S, Khanjani MJ. 2015a. Modeling of significant wave height Using wavelet and genetic programming. In

International Journal of Computer Applications (0975 – 8887) Volume 133 – No.16, January 2016

Proceedings of 36th IAHR World Congress, IAHR2015, 28th June-3rd July; Netherlands.

- [6] Altunkayank, A., "Prediction of significant wave height using spatial function", Ocean Engineering, 2015, 106, 220-226.
- [7] Deo, M. C., Jha, A., Chaphekar, A. S., and Ravikar. 2001. "Neural networks for wave forecasting", Ocean Engineerong, 2001, 28, 889-898.
- [8] Tsai, C. P., Lin, C., & Shen, J. N., "Neural network for wave forecasting among multi-stations", Ocean engineering, 2002, 29, 1683-1695.
- [9] Londhe, S., Panchang, V. "One-day wave forecasts using buoy data and artificial neural networks", Journal of Atmospheric and ocean Technology, 2006, 3, 2119-2123.
- [10] Ozger, M., & Sen, Z. 2007. Prediction of wave parameters by using fuzzy logic approach, Ocean Engineering, 34, 460-469.

- [11] Gaur S, Deo MC. Real-time wave forecasting using genetic programming. Ocean Engineering 2008; 35: 1166-1172.
- [12] Zhang, H., Liu, X., Cai, E., Huang, G., & Ding, C. "Integration of dynamic rainfall data with environmental factors to forecast debris flow using an improved GMDH model". Computers & Geosciences, 2013, 56, 23-31.
- [13] Zhu, B., He, C. Z., Liatsis, P., & Li, X. Y. A GMDHbased fuzzy modeling approach for constructing TS model. Fuzzy Sets and Systems, 2012, 189 (1), 19-29.
- [14] Amanifard, N., Nariman-Zadeh, N., Farahani, M-H., &Khalkhali, A. "Modeling of multiple short – lengthscale stall cells in an axial compressor using evolved GMDH neural networks", Energy Conversation and Management, 2008, 49 (10), 2588-2594.
- [15] Atashrouz, S., Pazuki, G. R., & Alimoradi, Y. "Estimation of the viscosity of nine nanofluids using a hybrid GMDH-type neural network system", Fluid Phase Equilibria, 2014, 372(25), 43-38.