Wave data assimilation using a hybrid approach in the Persian Gulf

**Mohammad Hadi Moeini · Amir Etemad-Shahidi ·
Vahid Chegini · Iraj Rahmani**

Abstract The main goal of this study is to develop an efficient approach for the assimilation of the hindcasted wave parameters in the Persian Gulf. Hence, the third generation SWAN model was employed for wave modeling forced by the 6-h ECMWF wind data with a resolution of 0.5°. In situ wave measurements at two stations were utilized to evaluate the assimilation approaches. It was found that since the model errors are not the same for wave height and period, adaptation of model parameter does not result in simultaneous and comprehensive improvement of them. Therefore, an approach based on the error prediction and updating of output variables was employed to modify wave height and period. In this approach, artificial neural networks (ANNs) were used to estimate the deviations between the simulated and measured wave parameters. The results showed that updating of output variables leads to significant improvement in a wide range of the predicted wave characteristics. It was revealed that the best input parameters for error prediction networks are mean wind speed, mean wind direction, wind duration, and the wave parameters. In addition, combination of the ANN estimated error with numerically modeled wave parameters leads to further improvement in the predicted wave parameters in contrast to direct estimation of the parameters by ANN.

Keywords Persian Gulf · Wave hindcasting · Neural network · SWAN · Error prediction

Responsible Editor: Chari Pattiaratchi

This article is part of the Topical Collection on *Physics of Estuaries and Coastal Seas 2010*

M. H. Moeini · A. Etemad-Shahidi (✉)
School of Civil Engineering,
Iran University of Science and Technology,
Tehran, Iran
e-mail: Etemad@iust.ac.ir

A. Etemad-Shahidi
e-mail: a.etemadshahidi@griffith.edu.au

M. H. Moeini
e-mail: mhmoeini@iust.ac.ir

A. Etemad-Shahidi
Griffith School of Engineering, Gold Coast Campus,
Griffith University,
Gold Coast, QLD 4222, Australia

V. Chegini
Iranian National institute for Oceanography,
Tehran, Iran
e-mail: v_chegini@inio.ac.ir

I. Rahmani
Transportation Research Institute,
Tehran, Iran
e-mail: rahmani@rahiran.ir

1 Introduction

Wind waves are the most important environmental factor in the design of coastal and offshore structures, sediment transport, coastal erosion, etc. Therefore, an accurate assessment of the wave climate is of great importance in the marine activities. Estimation of the wave climate can be performed using three different sources of long-term wave data, namely instrumentally measured, visually observed, or numerically modeled data. Due to the lack of long-term measurements and the recent development of state of the art wind-wave models, nowadays numerically simulated wave data are widely used for determining the wave climate. However, the results of numerical wind–wave model generally contain some inherent errors due to a number of factors mainly inaccuracies in the forcing terms (Komen et al. 1994). The wind forcing in the wind–wave models is generally

produced by the numerical weather prediction models such as those operated at the European Center of Medium-range Weather Forecast (ECMWF). These numerically simulated wind data contain some inevitable errors that lead to inaccurate prediction of wind waves (Cavaleri and Bertotti 2004; Signell et al. 2005; Ardhuin et al. 2007). Therefore, data assimilation is used to reduce the model errors (Kantha and Clayson 2000; Babovic et al. 2005; Sannasiraj et al. 2006).

Different data assimilation procedures can be categorized to four main groups. Updating of input parameters, updating of state variables such as wave spectrum, updating of model parameters such as whitecapping dissipation rate, and updating of output variables (error prediction) are the main data assimilation procedures (Refsgaard 1997; Babovic et al. 2005). The last one is an appropriate method to improve the accuracy of the numerical models (Babovic et al. 2001). Since the numerical model errors are not the same for different output variables (Moeini and Etemad-Shahidi 2007), the last assimilation procedure can be used to modify different wave characteristics separately.

Emmanouil et al. (2010) employed the Kalman filter in combination with the optimum interpolation data assimilation scheme to improve forecasted wave heights in an open ocean area (southwest US coast). In this study, the Kalman filters were implemented in the WAM model after the time integration of the two-dimensional frequency-direction wave spectra and before the data assimilation. The Kalman filter parameters were calculated based on the model predictions and the observations. Then, the correction effect of the Kalman filters was spread to a wider region by the use of a data assimilation system that follows the Kalman filters within the forecasting period. Their approach led to the extension of the assimilation impact to the whole forecasting period. A significant reduction of the magnitude and variability of the discrepancies between final forecasts and observations was achieved in this way (Emmanouil et al. 2010). Similar results are reported in Galanis et al. (2009) in that they employed some statistical tools to extend the impact of data assimilation on ocean wave prediction. In another study, Zhang et al. (2006) incorporated artificial neural networks and data assimilation techniques into the WAM model for wave forecasting. They assimilated altimeter significant wave height (H_s) data into the wave model by the statistical interpolation method to improve the accuracy of wave height prediction. Then, artificial neural networks were employed to mimic the effect of data assimilation and applied to situations where altimeter wave data were not available. They showed that by applying this approach for wind–wave simulation in the South China Sea, an apparent improvement in the accuracy of the forecasting can be obtained. Sannasiraj et al. (2005) also employed local model for wave data assimilation in Taiwan. The local model in their study was built over the time series embedding

theorem. They used the WAM model for wave forecasting and compared their results with buoy measurements. In their study, the numerical model outputs were updated using error prediction approach by the local model. According to their results, using a combination of numerical models and error prediction method reduce the errors of wave forecasting. Similar findings are reported by Babovic et al. (2005) and Sannasiraj et al. (2006).

Some other researchers have studied wave data assimilation, especially altimetry data, into the numerical models. Skandrani et al. (2004) evaluated the impact of using a multi-satellite of the altimeter data to improve wave model analyses and forecasts on a global scale. They employed the WAM model to generate significant wave height time series. They showed that the impact of data assimilation, when two or three sources of data are used instead of one, in terms of significant wave height, is larger in wave model analyses but smaller in wave model forecasts. They also showed that there is no improvement in terms of wave periods, both in the analysis and forecast periods. Greenslade (2001) assimilated satellite altimeter significant wave height data into the AUSWAM model (a version of WAM), using a statistical interpolation scheme. He validated the results of his study against a number of waverider buoys situated around the Australian coast. He found that assimilation of the ERS-2 SWH data decreased the systematic bias in analyzed SWH by approximately 10%, and the Scatter Index was reduced by 6%. The improvement in model skill was also retained throughout the forecast period. He showed that the assimilation has the greatest impact at locations on the west coast of Australia, where the sea state is dominated by swell. Greenslade and Young (2005) developed models for the structure of the background errors in the wave data assimilation systems based on the results of previous studies. They tested their model in a global wave data assimilation system, and the resulting wave forecasts were validated against observations from buoys. They showed that forecasts of significant wave height substantially improved while forecasts of peak period were not similarly improved. Hasselmann et al. (1997) presented an optimal interpolation scheme for assimilating two–dimensional wave spectra. They showed that the correction of wind data needs to be combined with other data in an atmospheric data assimilation scheme. They emphasized the need for the development of combined wind and wave data assimilation schemes for the optimal use of satellite wind and wave data. Bender and Glowacki (1996) evaluated two strategies for the assimilation of wave heights into a spectral wave model. They assimilated satellite altimetry data from GEOSAT into the modeled wave data and showed a significant improvement in the wave prediction. Breivik and Reistad (1994) assimilated the observed significant wave height from ERS-1 in an operational numerical wave model. They compared their

results with an independent buoy measurement in the North Sea. They showed that the assimilation correction has an impact on 12–24-h prognosis of the wave field. Dunlap et al. (1998), Young and Glowacki (1996), and Lionello et al. (1992) have also studied altimeter wave data assimilation into the wave models.

A few studies have been conducted on wave modeling in the Persian Gulf. Moeini et al. (2010) evaluated the quality of two sources of wind data, i.e., the ECMWF modeled winds and the measured data, for wave modeling in the Persian Gulf. They employed the third generation SWAN model for wave simulation and compared the results with one station recorded wave data. They found that the SWAN model overestimates the low wave heights and underestimates the higher ones because of the underestimation of high wind speeds by ECMWF. They also noted that the adaptation of the model parameters cannot lead to a comprehensive improvement of the model results. This implies the necessity of using more comprehensive methodologies, such as, error prediction, for a more accurate prediction of all ranges of wave data. Taebi et al. (2008) simulated wind waves in the Persian Gulf and the Gulf of Oman by the Mike21 SW model forced by ECMWF wind data. They found that the hindcasted wave heights have the Scatter Index of nearly 40% after calibration of the model. Al-Salem et al. (2005) used the third-generation WAM model to simulate the wind waves in the Persian Gulf with emphasis on Kuwait territorial waters. Wind data with a spatial resolution of 0.5° obtained from the ECMWF were used to force the model. The model was validated using measured waves at several locations in Kuwait waters. It was shown that the WAM model successfully predicts the wave conditions except for some storms where the value of H_s is underestimated. This underestimation was due to the underestimation of the storms by ECMWF. Similar results can be found in Rakha et al. (2007).

The main goal of this study is to improve the simulation of wind waves in the Persian Gulf using an error prediction approach. Hence, the SWAN model (Booij et al. 1999) was employed for wave hindcasting and reproducing time series of wave data. The model outputs were compared with field measurements at two stations to verify the results. After the wave simulation by the SWAN model, an artificial neural network was used to improve the results of wave modeling based on the modelled and measured data. The use of error prediction approach and updating of output variables has several advantages compared to other assimilation methods. At first, it can be used to improve different output variables such as wave height and period separately. Secondly, a wider range of output data can be improved using this method (Moeini et al. 2010). Finally, this assimilation approach will cover almost all of the error sources including forcing terms and wave model evolution.

It is worth noticing that using neural networks without combination with numerical model for wave prediction has some limitations. Since this method is generally appropriate for interpolation rather than extrapolation, the available data for learning process should cover all range of the probable events in the study area. Therefore, it may not be suitable for the prediction of extreme events. In addition, when this method is used for wave forecasting, the potential forecast horizon is typically quite short (Agrawal and Deo 2002).

This paper is organized as follows. "Section 2" introduces the study area and the field data. "Section 3" gives a brief description of the SWAN model and error prediction technique, and the results and discussions are described in "Section 4." Finally, "Section 5" covers the summary and conclusions.

2 Study area and data

The Persian Gulf, located in the southwest of the Asian continent is a shallow, semi-enclosed basin in a typical arid zone and is an arm of the Indian Ocean. It is located between the longitude of 48–57° E and the latitude of 24–30° N (Fig. 1). This Gulf is connected to the deep Gulf of Oman through the narrow Strait of Hormuz. The Persian Gulf covers an area of approximately 226,000 km^2 with a length of 990 km. Its width varies from 56 to 338 km, separating Iran from the Arabian Peninsula with the shortest distance of about 56 km in the Strait of Hormuz. This basin has an average depth of about 35 m, and the deepest water depth is approximately 107 m (Purser and Seibold 1973; Emery 1956).

In the present study, two sets of recorded wave data were used for verification of the results. The first recorded wave data set was collected by a surface buoy located in 50.7° E and 28.9° N near Bushehr port (hereinafter called Bushehr station). The water depth in the location of the buoy was nearly 27 m (Fig. 1). The wave data were recorded from the first of February, 1995 until the 31st of December 1996. The second measured station was located at 48.167° E and 29.167° N in the Kuwait territorial waters (Fig. 1). The wave data were recorded every 3 h from the first of January, until the 30st of August 1994. This measurement station is called Kuwait station hereinafter. Each set of wave data was divided randomly to three parts, i.e., training, testing, and validation data.

The wind data used were the 6-h 10-m operational ECMWF data with a spatial resolution of 0.5°. It should be noted that ECMWF operation was selected because of its higher spatial resolution than that of ECMWF re-analysis (1.125° spatial resolution).

Fig. 1 The Persian Gulf, modeled area and wave measurement stations
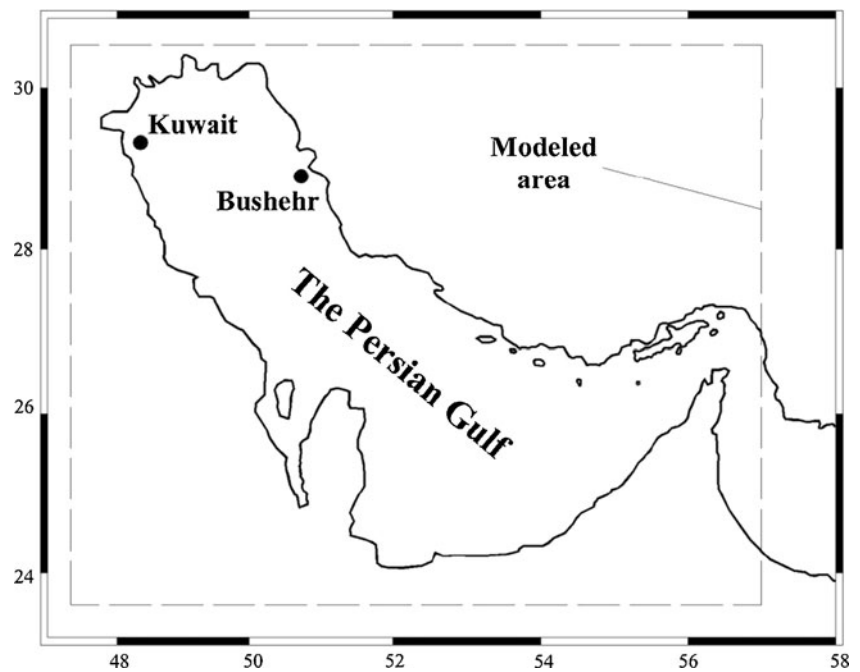
3 Numerical model and error prediction technique

In this study, the third generation spectral SWAN model (Booij et al. 1999) was employed to simulate wind waves over the Persian Gulf. The SWAN model is a spectral wind wave model, developed to obtain reliable estimates of wave parameters in coastal waters. In this model, action density spectrum is considered rather than energy density spectrum because it is conserved in the presence of currents unlike energy density. The action density is the energy density divided by the relative frequency:

$$N(\sigma,\theta) = E(\sigma,\theta)/\sigma \tag{1}$$

The relative frequency σ and the wave direction θ are the independent variables. In this model, the wave spectrum evolution in the space (x,y) and time (t) is described by the spectral action balance equation (Booij et al. 1999). In the present study, SWAN cycle III version 40.72 (The SWAN Team 2007) was used for wave simulation. The SWAN model was executed in a non-stationary mode and spherical coordinates. Since the Komen's formulation for wind input parameterization (Komen et al. 1984) leads to better results (Moeini and Etemad-Shahidi 2007), this formulation was used for exponential growths of wind input. Additionally, quadruplet wave interaction was activated for nonlinear interaction. Dissipation due to bottom friction, whitecapping, and depth-induced wave breaking were considered in

the simulation. The geographical space was discretized into a 200×160 cell grid over the Persian Gulf (Fig. 1) with a 0.05×0.05 degree resolution in x and y directions. The spectral space was divided into 20 logarithmically spaced frequencies, from 0.06 to 1 Hz and 24 equal directions in the rose ($\Delta\theta = 360/24 = 15$). This depicts that the lowest period of the simulated waves is 1 s, and the highest was nearly 17 s covering typical wind waves in the Persian Gulf. The computational time step was set as 10 min as well.

In the recent decades, artificial neural networks (ANNs) have been widely used in the ocean and coastal engineering applications especially wind and wave modeling (Agrawal and Deo 2002; Makarynskyy et al. 2005). The deviations between the simulations and the measured values, i.e., the model errors, are usually serially correlated. Therefore, it is possible to predict the values of these errors by means of time series models such as autoregressive moving average model or neural networks. The simulations can then be improved by adding the error predictions. This method, which is often referred to as error prediction, has been used in the real-time forecasting (Refsgaard 1997; Babovic et al. 2001, 2005). In the present study, a two-layer feed-forward network with the sigmoid hidden neurons and linear output neurons was employed for error prediction of the wave modeling. The network was trained with Levenberg-Marquardt back-propagation algorithm. The possible inputs of the ANN for error prediction of wave modeling at the current time step (analysis

Fig. 2 Comparison of the modeled wave heights and periods with default and updated parameter against the measurements at Bushehr station
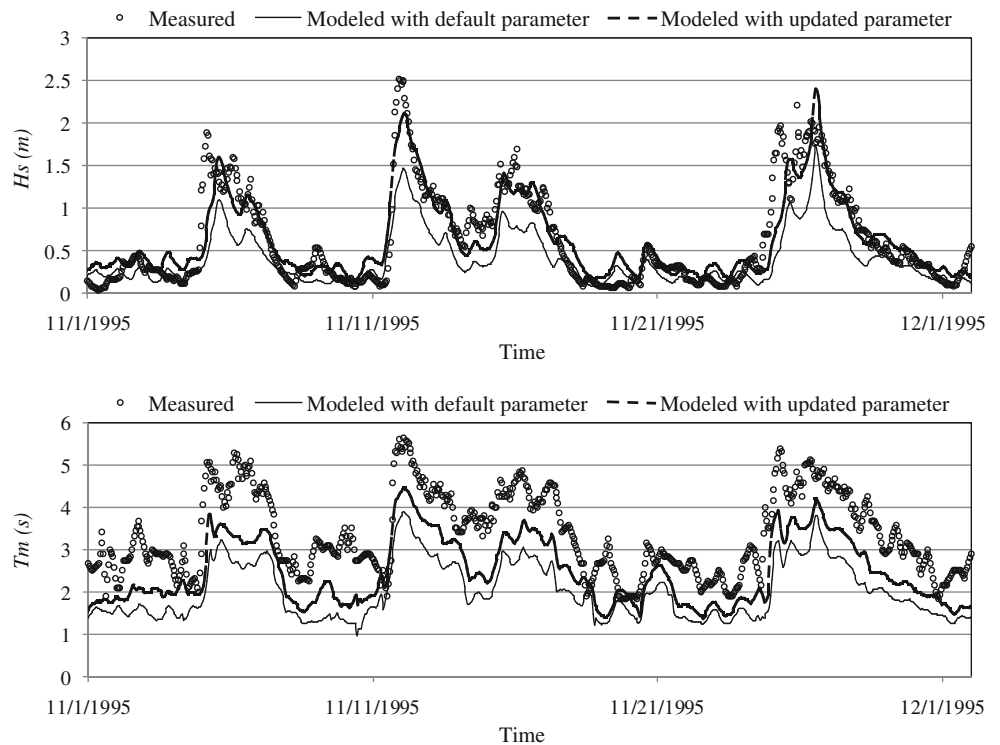
time) are the mean wind speed, mean wind direction, wind duration, fetch length, the modeled H_s, T_m and mean wave direction. Separated networks were employed for error estimation at each wave measurement station. To determine the wind duration, the definition of constant wind presented in the CEM (US army 2006) was used. In this way, a constant wind is defined when the difference between the

current wind speed and mean wind speed at previous time steps does not exceed a certain value (i.e., 2.5 ms^{-1}). In addition, wind direction should not vary more than 45 degrees. The presented data assimilation algorithm can be also used for future time steps. More details about neural network modeling can be found in Agrawal and Deo (2002).

Fig. 3 Comparison of the modeled wave heights and periods with default and updated parameter against the measurements at Kuwait station
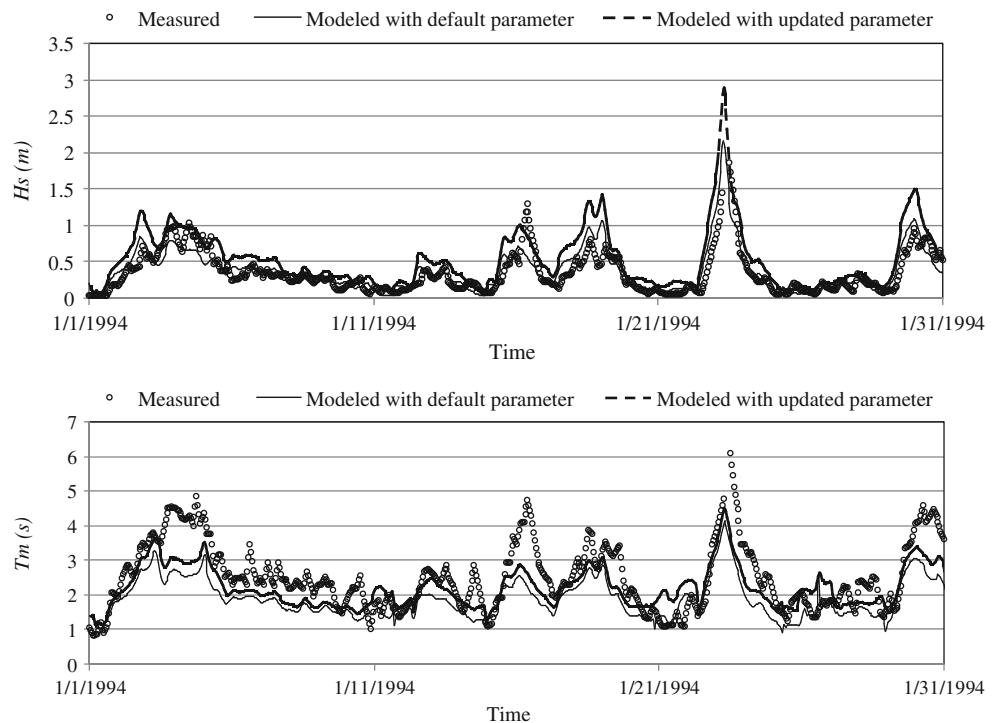
Table 1 The summary of statistical analysis of wave prediction in the analyzed scenarios for Bushehr station

Wave characteristics	Statistical Parameter	Modeled by default parameter	Modeled by updated parameter	Modified by ANN
Wave height	Bias (m)	−0.06	0.16	0.00
	RMSE (m)	0.29	0.35	0.09
	Scatter index (%)	65.19	78.7	20.20
	R^2	0.42	0.45	0.94
Wave period	Bias (s)	−0.98	−0.59	0.00
	RMSE (s)	1.13	0.80	0.22
	Scatter index (%)	37.84	26.91	7.40
	R^2	0.42	0.47	0.91

4 Results and discussion

4.1 Simulation

To investigate the performance of assimilation techniques, firstly all model parameters were set as default. Figure 2 shows time series of measured and modeled significant wave heights and mean wave periods during November 1995 at Bushehr station. Figure 3 shows the mentioned parameters at Kuwait station during January 1994. As seen in Fig. 2, high waves are generally underestimated by the SWAN model at Bushehr station mainly due to the underestimation of driving wind field (Moeini et al. 2010). The wave periods are significantly underestimated at this station. A comparison of the results at Kuwait station (Fig. 3) shows that H_s is well predicted by the model except for some storms where they are underestimated. The wave periods are underestimated generally at this station too. As seen in Tables 1 and 2, the RMSE of the prediction of significant wave height at Bushehr station is nearly two times of that for Kuwait station. The scatter index of the prediction of significant wave height at Bushehr station is almost 11% more than that of Kuwait station. The wave period is underestimated about 1 second at Bushehr station and half a second at Kuwait station. These findings are in agreement with those of Rakha et al. (2007) showing the underestimation of some

modeled H_s by the WAM model in Kuwait waters. These results show that the model errors are not the same for wave height and period because of the wave model evolution (Lin et al. 2002; Moeini and Etemad-Shahidi 2007).

4.2 Data assimilation

After assessment of the wave model, the model parameters were updated to evaluate their effects on the improvement of the results. In this step, the whitecapping dissipation rate was used as the tunable parameter. Other physical parameters such as bottom friction and wave breaking had no significant effect on the model results because of the deep water location of the buoys. The focus of this step was on the improvement of the wave height at Iranian waters, i.e., Bushehr station. The whitecapping dissipation rate was finally reduced to about 70%. Figures 2 and 3 also depict the qualitative comparisons of the hourly time series of the modeled H_s and mean wave period (T_m) using the updated parameter.

As seen in Fig. 2, updating of the model parameter leads to greater consistency between modeled and measured H_s especially in case of high waves at Bushehr station. However, some inconsistencies are observed between the measured and modeled wave heights because of inaccuracies in the forcing terms. In this case, low waves are generally

Table 2 The summary of statistical analysis of wave prediction in the analyzed scenarios for Kuwait station

Wave characteristics	Statistical Parameter	Modeled by default parameter	Modeled by updated parameter	Modified by ANN
Wave height	Bias (m)	0.051	0.21	0.00
	RMSE (m)	0.16	0.28	0.05
	Scatter index (%)	54.42	94.73	17.09
	R^2	0.67	0.70	0.96
Wave period	Bias (s)	−0.535	−0.23	0.00
	RMSE (s)	0.79	0.62	0.19
	Scatter index (%)	33.04	26.0	8.11
	R^2	0.66	0.64	0.96
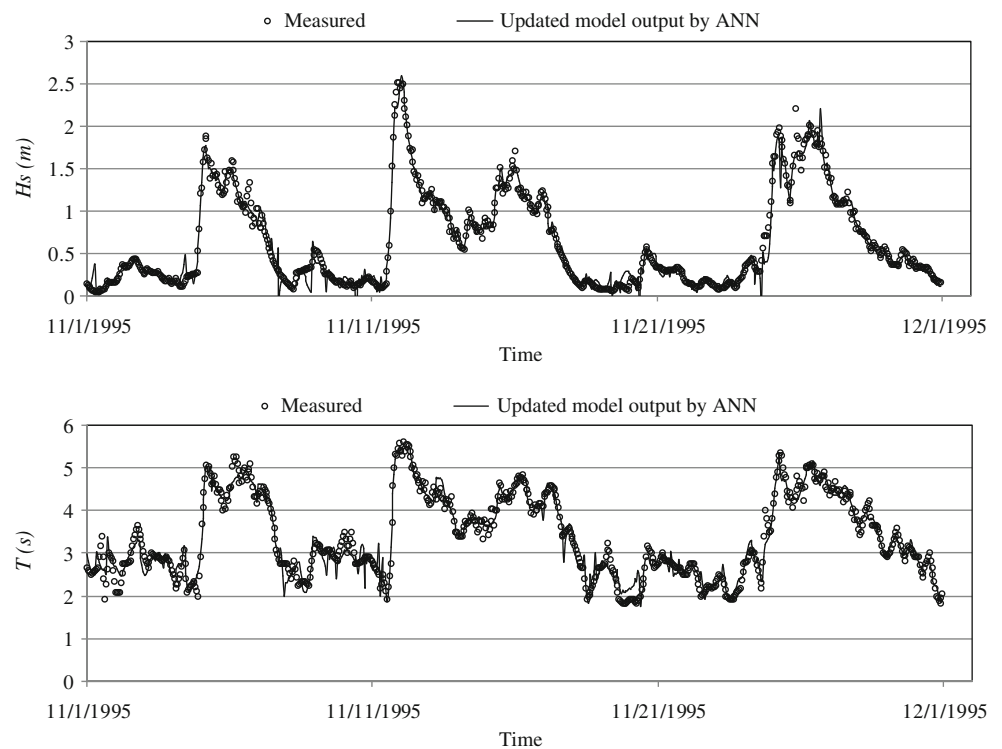
overestimated. Similar results can be found in Moeini et al.
(2010) for another area in the Persian Gulf. It is worth
noticing that the wave periods are still underestimated at
this station after adaptation of the model parameter based on
the modeled wave height. The effect of the updating of the
model parameter on the modeled wave data at Kuwait station

is shown in Fig. 3. This figure shows that updating of the
model parameter based on the modeled wave height at Bush-
ehr station results in the overestimation of wave heights at
Kuwait station. This fact reveals that spatial error of the output
variables is not uniformly distributed probably due to the
spatially nonuniform error covariance of the wind input, since
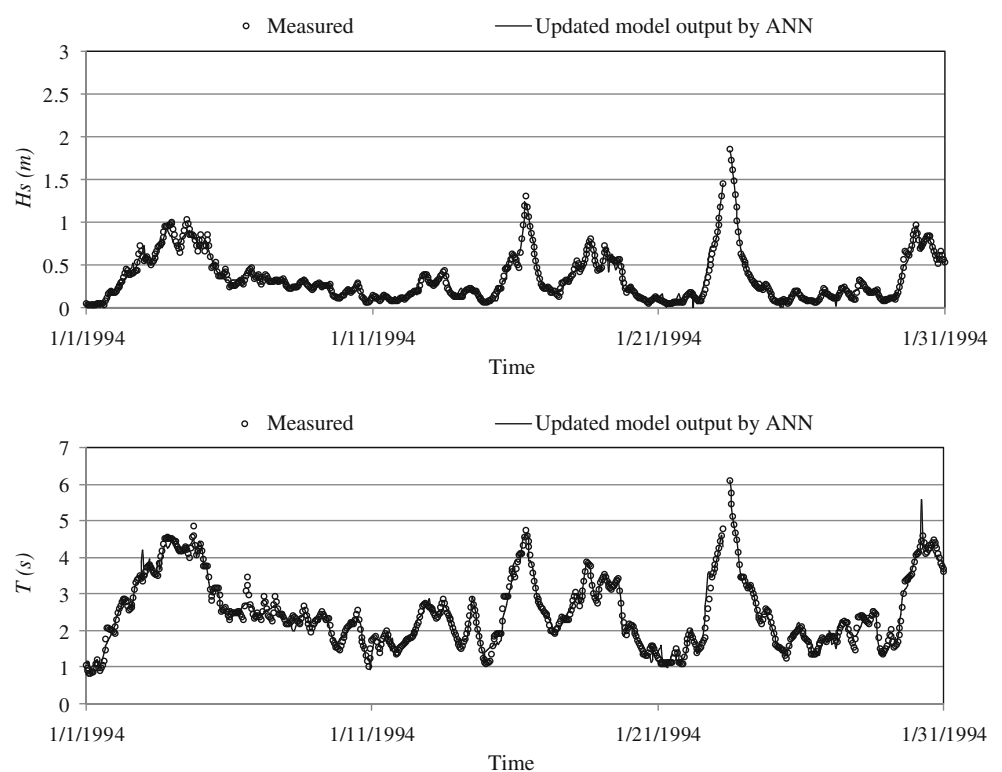
the error in the modeled waves is mainly due to the wind forcing error (Komen et al. 1994). In spite of the general overestimation of the modeled H_s at Kuwait station, Fig. 3 also shows that the wave period is underestimated in this condition. These findings show that owing to the difference between accuracy of the model for prediction of wave height and period, adaptation of the model parameter leads only to improvement in one variable. In other words, by updating the model parameter, different output variables cannot be improved simultaneously. On the other hand, adaptation of the model parameters can only enhance the results of one wave characteristic up to a certain level, depending on the accuracy of the wind input. These shortcomings in this data assimilation technique justify the necessity of employing more comprehensive methodologies, such as, error prediction, for a more accurate prediction of all ranges of different wave parameters. Therefore, updating of output variables (H_s and T_m) was considered in the next step for further improvement in the results.

In this step, ANN was employed to improve the model results. Different approaches were employed to predict the correct values of the wave parameters. In the first approach, the error predictions were combined with the modeled values. The errors of the modeled parameters, i.e., H_s and T_m, were estimated by ANN. In the second approach, the wave parameters were predicted directly by ANN. It will be shown that a combination of the ANN estimated error with modeled wave parameter leads to more accurate results than the prediction of the wave parameters by ANN. Let X be the wave parameter, i.e., the H_s or T_m. The error of the wave prediction can be defined as

$$E_{modeled} = X_{measured} - X_{modeled}, \qquad (2)$$

where $X_{modeled}$ is the modeled wave parameter (by the SWAN) and $X_{measured}$ is the corresponding measured value. If $E_{modeled}$ can be estimated by ANN, then the modified wave parameter can be obtained by

$$X_{modified} = X_{modeled} + E_{predited}, \qquad (3)$$

where $E_{predicted}$ is the estimated error of the wave parameter modelled by ANN. To determine the effective input parameters and the best topology, different networks with various combinations of input parameters were considered. The wind speed, wind direction, wind duration, wind shear velocity, modeled wave height, period, and direction were tested as the network inputs. It was revealed that the best inputs for the error prediction network of the H_s are the mean wind speed, mean wind direction, wind duration, and the modeled H_s. For the error prediction network of the T_m, the modeled wave height was replaced with the modeled wave period. The assessment of the best inputs was performed based on the sensitivity analysis and error statistics

of the results. The numerically modeled wave parameters have an important role in the error prediction approach that will be discussed later.

Figures 4 and 5 illustrate the comparison of time series of the updated wave parameters using ANN against the measurements at Bushehr and Kuwait stations, respectively. In contrast to Figs. 2 and 3, it can be seen that combination of the SWAN numerical model and ANN leads to better improvements in the wave simulation rather than updating of the model parameter. As seen, both the updated wave heights and periods follow the measurements very well. In addition, there is a great consistency between the measured and updated wave data for both low and high values. In other words, the employed assimilation approach successfully covers the error sources of the simulation and results in the improvement in a wider range of output variables. To have an overall assessment of the employed approaches, the scatter diagrams of H_s and T_m are depicted in Figs. 6 and 7. In addition, Tables 1 and 2 show the summary of error measures (bias, RMSE, scatter index, and R^2) at Bushehr and Kuwait stations for the whole time period that wave data were available. According to Fig. 6 and Table 1, the wave heights at Bushehr station modeled by SWAN with default parameter are underestimated (bias, −0.06 m) and have a large scatter (scatter index, 65.19%). After the updating of the model parameter, the modeled wave heights become overestimated (bias, 0.16 m), but the scatter index increases. These statistics were computed based on the all range of wave data. Because of the importance of high waves, updating of the model parameter was conducted based on the simulation of high waves. Since the ECMWF wind data are underestimated especially for high values, the calibration of high values has led to the overestimation of low wave heights (Moeini et al. 2010). Therefore, the statistics for all range of wave data after updating of the model parameter are worse than those before that. As an example, the SI (for the measured wave height larger than 1 m) was about 48% before updating of the model parameter, and it was reduced to 32% after updating.

The best results are obtained in the third approach that is the combination of error prediction by ANN and SWAN outputs. In this case, the simulated wave heights have no bias and the scatter index is significantly decreased (scatter index, 20.2%). In the case of T_m, the scatter index decreases about 80% after using an error prediction approach. Similar results are observed for the Kuwait station (Fig. 7; Table 2). In this station, the scatter index decreases about 69% for wave height and 75% for wave period after a combination of

Fig. 6 Scatter diagrams of the significant wave height (*left panel*) and wave period (*right panel*) at Bushehr station (**a**) SWAN model with default parameter (**b**) SWAN model with updated parameter (**c**) modified wave characteristics by ANN
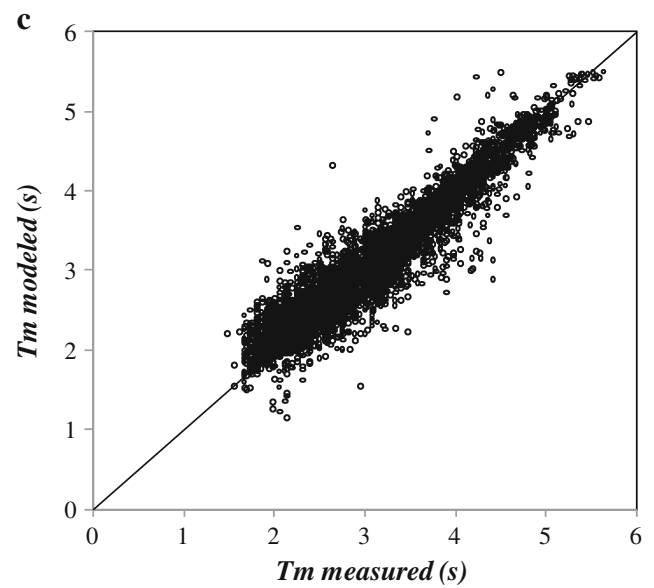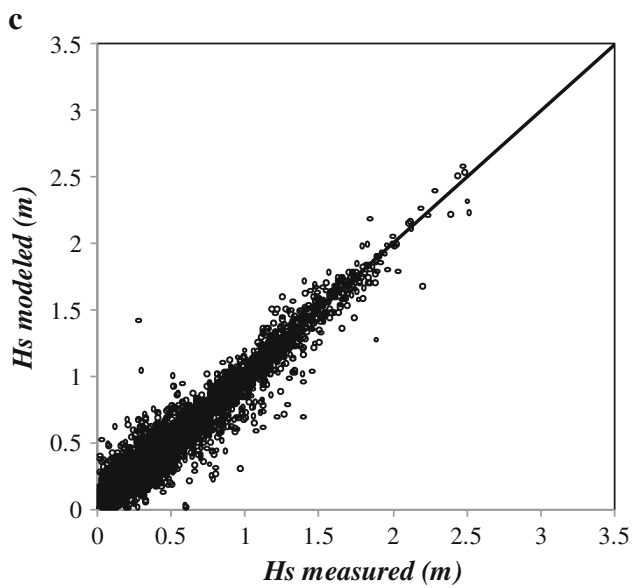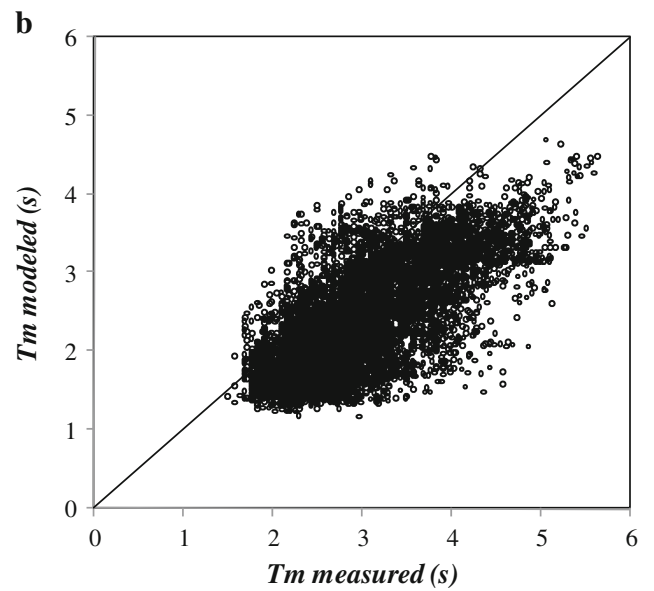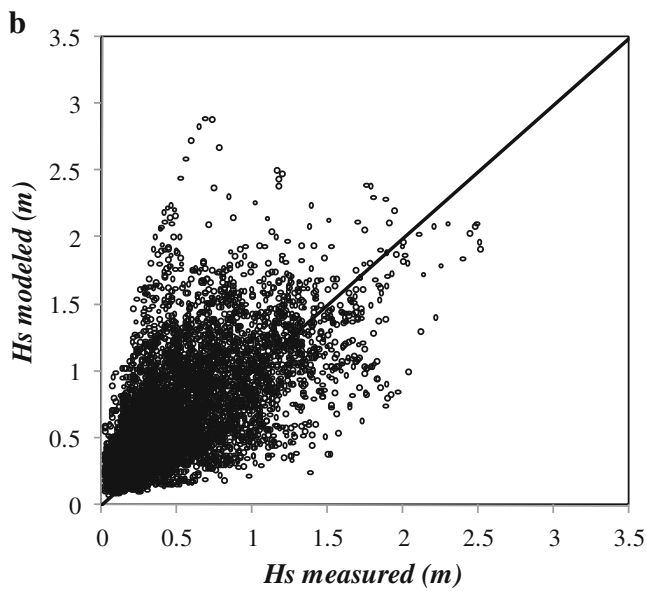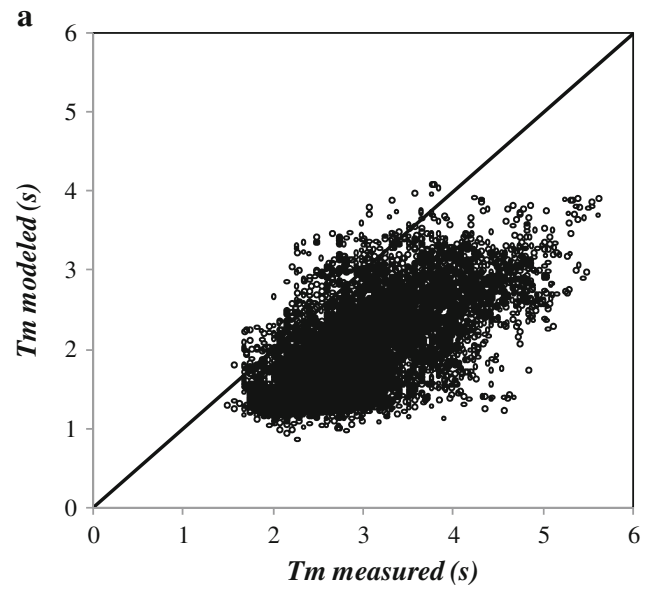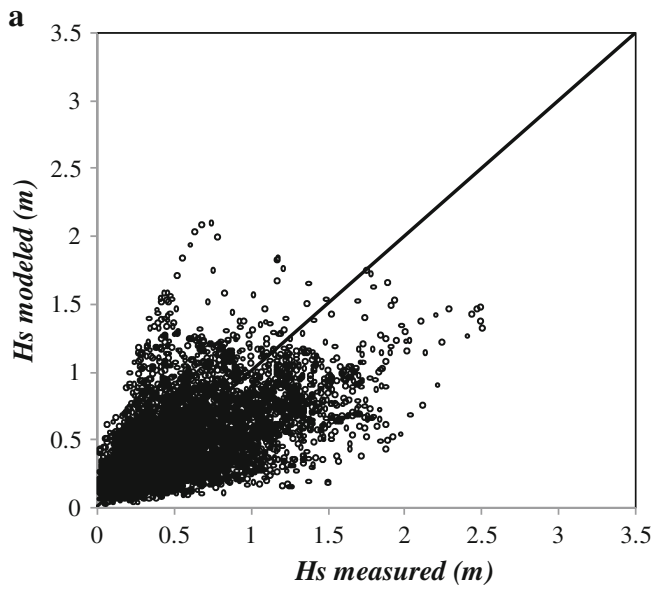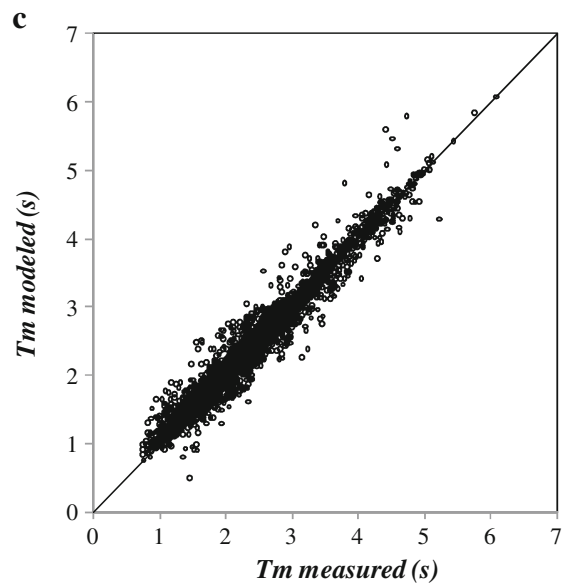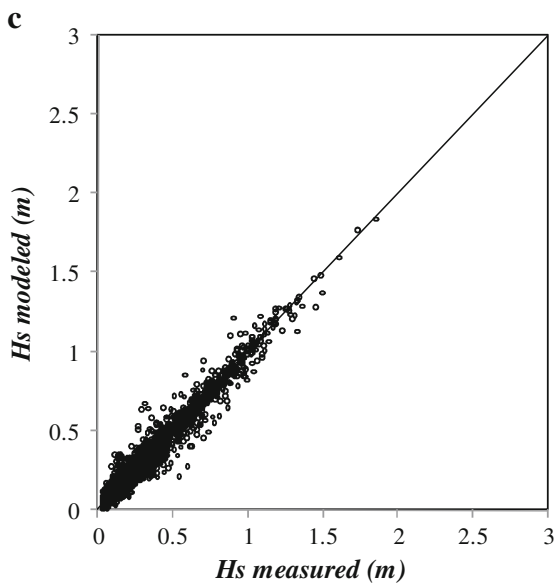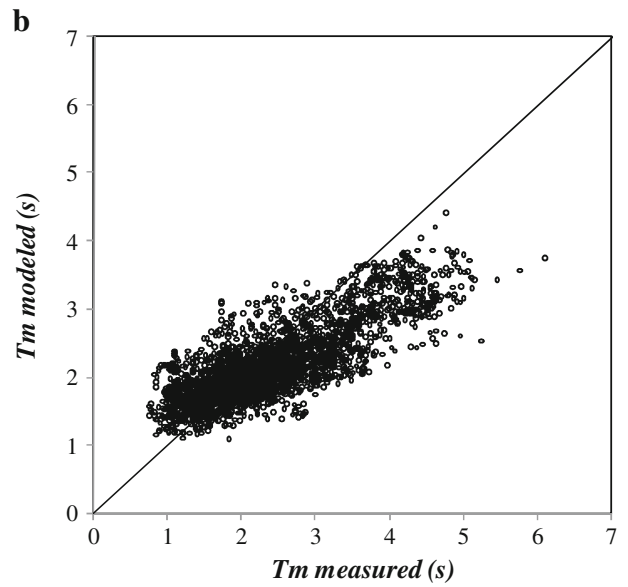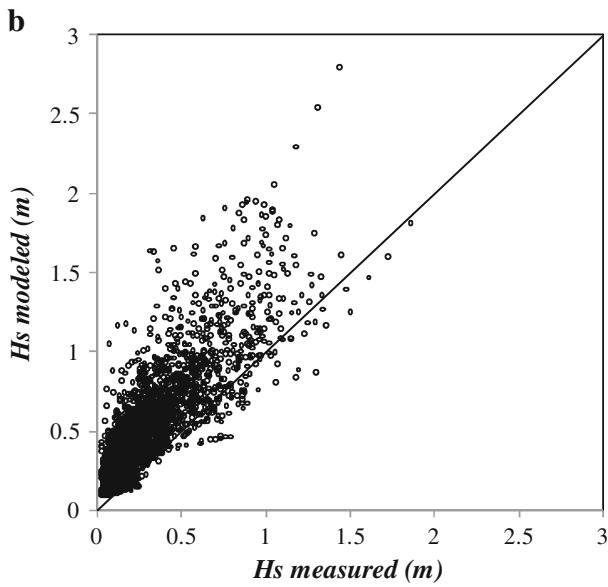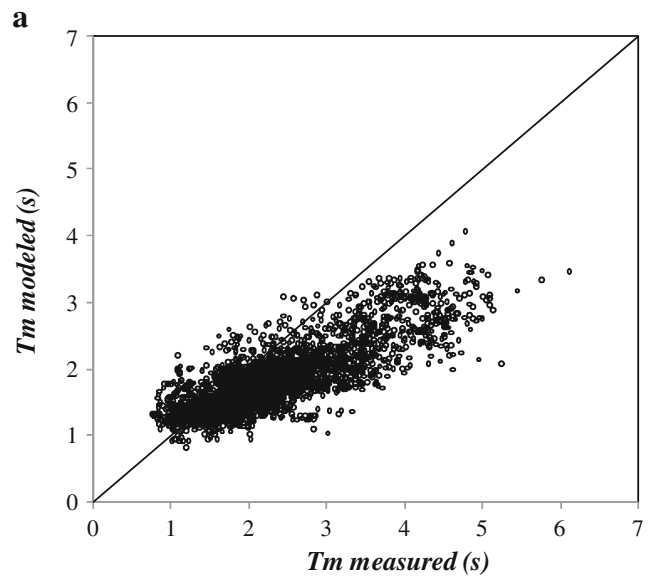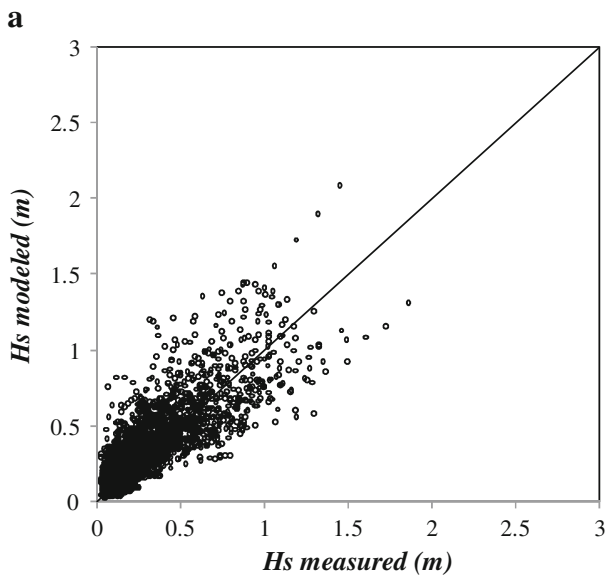
◀ **Fig. 7** Scatter diagrams of the significant wave height (*left panel*) and wave period (*right panel*) at Kuwait station (**a**) SWAN model with default parameter (**b**) SWAN model with updated parameter (**c**) modified wave characteristics by ANN

ANN and numerical model. These findings show that the output updating approach outperforms updating of the model parameters, which is in line with the results of Babovic et al. (2005) and Sannasiraj et al. (2005).

As mentioned before, different approaches were employed to predict the correct values of the wave parameters. The measured and modeled wave parameters at Bushehr station in January were analyzed. At first, ANN was employed to estimate measured wave parameters. Secondly, an error prediction approach was considered. Table 3 represents the statistical analysis of the predicted wave parameters in this period. As seen, the scatter index of predicted wave height decreases about 11% using the error prediction approach compared with that of prediction of the correct H_s using ANN. In the case of T_m, error prediction method results in a decrease of about 29% in the scatter index compared with that of T_m prediction by ANN.

Finally, it is worth noticing that the numerical wave modeling has an important role in the employed approach. The modeled wave parameters contain valuable information that is very important in the estimation of errors. In addition to wind forcing parameters, the error of the predictions depends on the modeled parameter as well. To reveal this fact, two different networks were considered. The inputs of the first network were the mean wind speed, mean wind direction, and wind duration. For the second networks, the modeled H_s and T_m were added to the networks inputs for error prediction of the wave height and period, respectively. Table 3 shows the results of this assessment at Bushehr station in January. As seen, a significant reduction in the scatter index of the predicted wave height is observed after adding the modeled H_s to the networks input. In this case, the scatter index shows an improvement of about 43%. In the case of T_m, a reduction of about 31% in the scatter index of predicted T_m is seen where the scatter index decreases

from 9.21% to 6.34%. This is mainly due to the fact that the wind field error depends also on the dynamics of the wind–wave system and the modeling of drag in the surface layer. In addition, waves are an integrated effect of the driving wind fields in space and time (Cavaleri and Bertotti 1997). Thus, it is suggested to use the employed approach for determining long-term time series of wave characteristics in the Persian Gulf.

The low number of wave measurement stations and the relatively short recording period may be the drawbacks of the present study. In addition, discovering the error distribution of the ECMWF winds and the modeled waves over the domain is also a matter of interest. Thus, it is recommended that more measured data such as satellite observations be employed for spatial assessment of the modeled wind and wave data.

5 Summary and conclusions

In this study, an efficient approach for the assimilation of the hindcasted wave parameters in the Persian Gulf was developed. To do so, the third generation SWAN model was employed for wave simulation forced by the 6-hourly ECMWF wind data with a resolution of 0.5°. In situ wave measurements in two stations were utilized for the evaluation of the studied approaches. Two main data assimilation procedures, i.e., updating of model parameters and updating of output variables, were analyzed for nudging the modeled data to the real one. Updating of output variables was carried out by means of ANN for estimation of the deviations between the simulated and measured wave characteristics. Different networks with various combinations of input parameters were considered to determine the effective input parameters and the best combination. The most important results are summarized as follows.

The model errors are not the same for wave height and period mainly due to the wave model evolution. Therefore, adaptation of the model parameter results in the improvement

Table 3 Comparison of the wave parameters estimated by ANN and error prediction approach

This table also shows the evaluation of the effect of the modeled wave parameters on the accuracy of the error prediction. ANN1 are the networks with modeled wave parameters and ANN2 are the networks without modeled wave parameters

Wave characteristics	Statistical Parameter	Direct prediction by ANN	Error prediction by ANN1	Error prediction by ANN2
Wave height	Bias (m)	0.00	0.00	0.00
	RMSE (m)	0.10	0.09	0.16
	Scatter index (%)	15.26	13.58	23.90
	R^2	0.97	0.98	0.92
Wave period	Bias (s)	0.00	0.00	−0.01
	RMSE (s)	0.30	0.21	0.31
	Scatter index (%)	8.91	6.34	9.21
	R^2	0.90	0.95	0.89

in only one output variable. In other words, by updating the model parameter, different output variables cannot be improved simultaneously. Additionally, even for one variable, adaptation of the model parameters can enhance the results up to a certain level depending on the accuracy of the wind input. Large scatter between modeled and measured data may still exist using this approach.

Updating of output variables outperforms updating of the model parameters for data assimilation. In this approach, both updated wave height and period follow the measurements very well. In addition, this assimilation approach covers almost all of the error sources of the wave modeling and results in the improvement of low and high values of output variables.

Combination of the ANN estimated error with numerically modeled wave parameters leads to better improvement in the predicted wave data in contrast to direct estimation of the desired parameters by ANN. In addition, the best inputs are the mean wind speed, mean wind direction, wind duration, and the modeled wave parameters. The modeled wave parameters, as the network inputs, increase the accuracy of the error estimation significantly.

References

Agrawal JD, Deo MC (2002) On-line wave prediction. Mar Struct 15 (1):57–74. doi:10.1016/S0951-8339(01)00014-4

Al-Salem K, Rakha KA, Sulisz W, Al-Nassar W (2005) Verification of a WAM model for the Arabian Gulf. In: Proceedings of the Arabian Coast 2005 Conference, Dubai

Ardhuin F, Bertotti L, Bidlot JR, Cavaleri L, Filipetto V, Lefevre JM, Wittmann P (2007) Comparison of wind and wave measurements and models in the Western Mediterranean Sea. Ocean Eng 34(3–4):526–541. doi:10.1016/j.oceaneng.2006.02.008

Babovic V, Canizares R, Jensen HR, Klinting A (2001) Neural networks as routine for error updating of numerical models. J Hydraul Eng 127(3):181–193. doi:10.1061/(ASCE)0733-9429 (2001)127:3(181)

Babovic V, Sannasiraj SA, Chan ES (2005) Error correction of a predictive ocean wave model using local model approximation. J Mar Syst 53(1–4):1–17. doi:10.1016/j.jmarsys.2004.05.028

Bender LC, Glowacki T (1996) The assimilation of altimeter data into the Australian wave model. Aust Meteorol Mag 45(1):41–48

Booij N, Ris RC, Holthuijsen LH (1999) A third-generation wave model for coastal regions 1. Model description and validation. J Geophys Res 104(C4):7649–7666

Breivik LA, Reistad M (1994) Assimilation of ERS-1 altimeter wave heights in an operational numerical wave model. Weather Forecast 9(3):440–451. doi:10.1175/1520-0434(1994)009<0440: aoawhi>2.0.co;2

Cavaleri L, Bertotti L (1997) In search of the correct wind and wave fields in a minor basin. Mon Weather Rev 125(8):1964–1975. doi:10.1175/1520-0493(1997)125<1964:isotcw>2.0.co;2

Cavaleri L, Bertotti L (2004) Accuracy of the modelled wind and wave fields in enclosed seas. Tellus, Series A: Dyn Meteorol Oceanogr 56(2):167–175. doi:10.1111/j.1600-0870.2004.00042.x

Dunlap EM, Olsen RB, Wilson L, De Margerie S, Lalbeharry R (1998) The effect of assimilating ERS-1 fast delivery wave data into the North Atlantic WAM model. J Geophys Res 103(C4):7901–7915. doi:10.1029/97jc02570

Emery KO (1956) Sediments and water of the Persian Gulf. Bull Am Assoc Petrol Geol 40(10):2354–2383

Emmanouil G, Galanis G, Kallos G (2010) A new methodology for using buoy measurements in sea wave data assimilation. Ocean Dyn 60(5):1205–1218. doi:10.1007/s10236-010-0328-9

Galanis G, Emmanouil G, Chu PC, Kallos G (2009) A new methodology for the extension of the impact of data assimilation on ocean wave prediction. Ocean Dynam 59(3):523–535. doi:10.1007/s10236-009-0191-8

Greenslade DJM (2001) The assimilation of ERS-2 significant wave height data in the Australian region. J Mar Syst 28(1–2):141–160. doi:10.1016/S0924-7963(01)00005-7

Greenslade DJM, Young IR (2005) The impact of inhomogenous background errors on a global wave data assimilation system. J Atmos Ocean Sci 10(2):61–93. doi:10.1080/17417530500089666

Hasselmann S, Lionello P, Hasselmann K (1997) An optimal interpolation scheme for the assimilation of spectral wave data. J Geophys Res 102(C7):15823–15836. doi:10.1029/96jc03453

Kantha LH, Clayson CA (2000) Numerical models of oceans and oceanic processes. Academic Press, California

Komen GJ, Hasselmann S, Hasselmann K (1984) On the existence of a fully developed windsea spectrum. J Phys Oceanogr 14(8):1271–1285. doi:10.1175/1520-0485(1984)014<1271:oteoaf>2.0.co;2

Komen GJ, Cavaleri L, Donelan M, Hasselmann K, Hasselmann S, Janssen PAEM (1994) Dynamics and modelling of ocen waves. Cambridge University Press, Cambridge

Lin W, Sanford LP, Suttles SE (2002) Wave measurement and modeling in Chesapeake Bay. Continent Shelf Res 22(18–19):2673–2686. doi:10.1016/S0278-4343(02)00120-6

Lionello P, Günther H, Janssen PAEM (1992) Assimilation of altimeter data in a global third-generation wave model. J Geophys Res 97 (C9):14453–14474

Makarynskyy O, Pires-Silva AA, Makarynska D, Ventura-Soares C (2005) Artificial neural networks in wave predictions at the west coast of Portugal. Comput Geosci 31(4):415–424. doi:10.1016/j.cageo.2004.10.005

Moeini MH, Etemad-Shahidi A (2007) Application of two numerical models for wave hindcasting in Lake Erie. Appl Ocean Res 29 (3):137–145. doi:10.1016/j.apor.2007.10.001

Moeini MH, Etemad-Shahidi A, Chegini V (2010) Wave modeling and extreme value analysis off the northern coast of the Persian Gulf. Appl Ocean Res 32(2):209–218. doi:10.1016/j.apor.2009.10.005

Purser BH, Seibold E (1973) The principal environmental factors influencing Holocene sedimentation and diagenesis in the Persian Gulf. The Persian Gulf, Springer-Verlag, Berlin

Rakha KA, Al-salem K, Neelamani S (2007) Hydrodynamic atlas for Kuwaiti territorial waters. Kuwait J Sci Eng 34:143–156

Refsgaard JC (1997) Validation and intercomparison of different updating procedures for real-time forecasting. Nord Hydrol 28 (2):65–84. doi:10.2166/nh.1997.005

Sannasiraj SA, Babovic V, Chan ES (2005) Local model approximation in the real time wave forecasting. Coast Eng 52(3):221–236. doi:10.1016/j.coastaleng.2004.12.004

Sannasiraj SA, Babovic V, Chan ES (2006) Wave data assimilation using ensemble error covariances for operational wave forecast. Ocean Model 14(1–2):102–121. doi:10.1016/j.ocemod.2006.04.001

Signell RP, Carniel S, Cavaleri L, Chiggiato J, Doyle JD, Pullen J, Sclavo M (2005) Assessment of wind quality for oceanographic modelling in semi-enclosed basins. J Mar Syst 53:217–233. doi:10.1016/j.jmarsys.2004.03.006

Skandrani C, Lefevre J-M, Queffeulou P (2004) Impact of multisatellite altimeter data assimilation on wave analysis and forecast. Mar Geodes 27(3–4):511–533. doi:10.1080/01490410490883496

Taebi S, Golshani A, Chegini V (2008) An approach towards wave climate study in The Persian Gulf and the Gulf of Oman: simulation and validation. J Mar Eng 4(7):11–26

The SWAN team (2007) SWAN user manual (Cycle III version 40.51). Delft University of Technology, Delft

US Army (2006) Coastal engineering manual, Chapter II-2, Meteorology and wave climate. Engineer manual 1110-2-1100. US Army Corps of Engineers, Washington, DC

Young IR, Glowacki TJ (1996) Assimilation of altimeter wave height data into a spectral wave model using statistical interpolation. Ocean Eng 23(8):667–686. doi:10.1016/0029-8018 (95)00066-6

Zhang Z, Li CW, Qi Y, Li YS (2006) Incorporation of artificial neural networks and data asssimilation techniques into a third-generation wind-wave model for wave forecasting. J Hydroinformatics 8 (1):65–76. doi:10.2166/jh.2006.005