

Error distribution and correction of the predicted wave characteristics over the Persian Gulf

Mohammad Hadi Moeini¹, Amir Etemad-Shahidi², Vahid Chegini³, Iraj Rahmani⁴ and Mona Moghaddam⁴

¹*School of Civil Engineering, Iran University of Science and Technology, Tehran, Iran.*

Tel.: +9821 77240399, Fax: +9821 77454053, Email: mhmoeni@iust.ac.ir

²*Griffith School of Engineering, Gold Coast Campus, Griffith University, QLD, 4222, Australia.*

Tel: +61 7 5552 9267, Fax: +61 7 5552 8065, Email: a.etemadshahidi@griffith.edu.au,

³*Iranian National Institute for Oceanography. Tehran, Iran.*

Email: v_chegini@inio.ac.ir

⁴*Transportation Research Institute. Tehran, Iran.*

Email: iraj.rahmani@gmail.com, mona.moghaddam@yahoo.com

Abstract

Wind-waves are the most important environmental parameter for the design of coastal and offshore structures, sediment transport, coastal erosion etc. Therefore, an accurate evaluation of the wave climate is of great importance. Due to the lack of long term measurements, nowadays numerically modeled wave data are widely used for determining the wave climate. The numerically simulated wave data are continuous in time and space, but generally inaccurate in enclosed and semi-enclosed basins mainly due to the inaccurate wind input data. The main goal of this study is to develop a new and efficient approach to improve the hindcasted wave parameters in the Persian Gulf. Hence, the third generation SWAN model was employed for the wave modeling forced by the 6-hourly ECMWF wind data with a resolution of 0.5°. A new methodology was introduced for the distribution of wave prediction errors from discrete observation points to the other points of interest. It was found that the proposed method which considers the wave generation physics, leads to a significant improvement in the predicted wave parameters. In addition, it was revealed that the improvements in prediction of waves with higher wave heights and longer periods are more than those of others. This was shown to be due to the higher correlation between high values of output parameters which contain larger errors. The influence radius in the error distribution procedure was found to be near 2 degrees (~200km).

Keywords: Persian Gulf; Wave hindcasting; SWAN; error distribution; Wave modeling

1. Introduction

Wind generated waves are the most important parameter in the design of coastal and offshore structures in the marine environment. The wave climate can be determined using long term wave data from instrumental measurement such as buoys, observation by satellites or numerical models (e.g. Goda, 2000). Each of these data sources has its advantages and disadvantages. The buoys present more accurate and frequent data but their locations are very sparse and mostly close to coasts. In addition, the long term buoy data may not be available in many regions. The altimetry data are continuous and accurate, but very intermittent and inaccurate near the coast (Krogstad and Barstow, 1999). The numerically modeled wave data are continuous in time and space, but generally inaccurate (mainly underestimated) in enclosed and semi-enclosed basins (Cavaleri and Sclavo, 2006). The most important reason for the underestimation of modeled data is inaccurate wind input data (Cavaleri and Sclavo, 2006; Komen et al., 1994). The wind data in the wave models are generally produced by numerical weather prediction models such as those operated at the ECMWF (European Center of Medium-range Weather Forecast). These simulated wind data are inevitably inaccurate and lead to inaccurate prediction of wind waves (Ardhuin et al., 2007; Cavaleri and Bertotti, 2004; Signell et al., 2005). Nevertheless, due to some irrecoverable drawbacks of buoys and altimetry data, nowadays numerically modeled data are widely used for determining the wave climate in different regions. To reduce the model errors, data assimilation techniques are employed to combine the modeled data with measurements at various points in the domain and produce the best possible estimation of the sea state over the entire domain (Kantha and Clayson, 2000).

Data assimilation techniques can be classified into four main groups namely updating of input parameters, updating of state variables, updating of model parameters and updating of output variables (UOV). (Babovic et al., 2005; Refsgaard, 1997). In these methods buoy and altimetry data are used to improve the accuracy of the simulated wave data in the hindcast or forecast databases (Babovic et al., 2005; Mínguez et al., 2011; Moeini et al., 2012). The UOV technique is an appropriate method to reduce the error of the numerical models (Babovic et al., 2001). This method can be employed to modify different wave characteristics noting that the numerical model errors are not the same for different outputs (e.g. Moeini and Etemad-Shahidi, 2007). Some researchers have studied the correction of the simulated wave data based on the information taken from the measurements. Caires and Sterl (2005) presented a nonparametric

approach to correct model outputs. They applied their approach to the significant wave height dataset of the 45 years European Centre for Medium-Range Weather Forecasts Re-Analysis (ERA-40). Their results showed clear improvements in bias, scatter, and quantiles of the corrected data in the whole range of significant wave height. Cavaleri and Sclavo (2006) used information taken from buoy, altimetry and model data to calibrate decadal time series at a large number of points, distributed at 0.5° intervals in the Mediterranean sea. They derived single calibration coefficients for wind speed and wave height at the various grid points. They showed that the model captures well the structure of the wave fields, i.e. wave direction; and applied no correction to the wave direction. In the above-mentioned studies, the calibration process was conducted on a point-to-point basis, without considering the spatial correlation between adjacent points or taking into account the effect of wave climate. Tomas et al. (2008) developed a methodology for the spatial calibration of wave hindcast data sources based on an empirical orthogonal function and a non-linear transformation of the spatial–time modes. They applied their method to monthly long-term distribution of significant wave height in the Western Mediterranean. This approach may not be applicable for the correction of hourly hindcast data bases. In addition, since they assume a prior distribution function of the data in their study area, their method may not be generalized. Sannasiraj and Goldstein (2009) presented an assimilation model based on an optimal interpolation scheme. They formulated the gain matrix (weighting matrix) according to the model physics (by which the waves are generated by winds). They evaluated the efficiency of their method by using buoy observations in the Arabian Sea. Their assimilation approach led to a 30–50% reduction of root mean square error of wave height at the validation stations. In their method, the gain matrix was formulated without considering the difference between wave climate in each season e.g. south-west monsoon, north-east monsoon, and non-monsoon. Mínguez et al. (2011) proposed a calibration procedure for wave hindcast databases based on a nonlinear regression method. In their method, the correction parameters were formulated based on the wave propagation direction and varied smoothly along the possible wave directions by means of cubic splines. They showed the performance of their method using buoy and satellite data from different locations around Spain.

Limited researches have been conducted on wave modeling and data assimilation in the Persian Gulf. Al-Salem et al. (2005) employed the third-generation spectral WAM model to simulate the wind waves in the Persian Gulf with focus on Kuwait territorial waters. They used wind data

with a spatial resolution of 0.5° obtained from the ECMWF to force the model. The model was validated using measurements at several locations near Kuwait waters. They showed that the WAM model successfully simulates the wave conditions except for some storm events where the wave heights were underestimated. This underestimation was due to the underestimation of the storms' intensity by ECMWF. Similar results were echoed by Rakha et al. (2007). Moeini et al. (2010) evaluated the quality of two wind data sources, i.e. numerically modeled winds and the measured data, for wave hindcasting in the Persian Gulf using the SWAN model. They found that the SWAN model overestimates the low wave heights and underestimates the higher ones because of the smoothing of wind field by ECMWF. They also noted that the calibration of the model parameters cannot lead to a comprehensive improvement of the model results. Moeini et al. (2012) developed an approach based on the point to point correction of the hindcasted wave parameters in the Persian Gulf. They showed that calibration of model parameters does not result in simultaneous and wide-ranging improvement of wave heights and periods. They proposed an error prediction approach using artificial neural networks (ANN) to correct wave heights and periods, separately. In addition, they showed that combination of the ANN estimated error with numerically modeled wave parameters leads to better results showing the importance of numerical wave modeling in the assimilation procedure.

The main goal of this study is to develop an efficient approach for improvement of the simulated wave parameters over the entire domain based on the error distribution. Two approaches are investigated for the error distribution and their results are compared with each other. In the first approach, the errors at the desired grid point are estimated based on the linear combination of the predicted errors at the observation points. In the second one, it is assumed that the error ratios at the different grid points are the same as the ratio of the wave heights at these points.

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 $48\text{--}57^\circ$ E and the latitude of $24\text{--}30^\circ$ 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 $226,000\text{ km}^2$ with a length of 990 km. Its width varies from 56 to 338 km, separating Iran from the Arabian Peninsula where the shortest distance is 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).

Generally, the input wind data for wind-wave simulation are generated by the meteorological models. Many different organizations run global atmospheric model producing 2-D gridded wind data all over the world. In this study, the wave model was forced by 6-hourly ECMWF operational data. These data are corresponding to 10 m level above the water surface and have spatial resolution of 0.5° . It should be noted that ECMWF operational data were selected because of their higher spatial resolution compared to those of ECMWF re-analysis (1.125° spatial resolution).

3. Wind-wave model

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, the action density spectrum is considered rather than the 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:

$$N(\sigma, \theta) = E(\sigma, \theta) / \sigma \quad (1)$$

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 nonstationary 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 activated in the simulation. This study focused on the assimilation method based on updating of output variables not on the updating of the model parameters. Therefore, all of the model's tunable coefficients such as bottom friction and whitecapping coefficient were set to their default values. For example, the used default value

of the coefficient for determining the rate of whitecapping dissipation (cds2 coefficient in the model set up) was 2.36e-5. Millar et al. (2007) have also used the default values for the model coefficients to analyze the change of the shoreline wave climate due to installation of a wave farm off the north coast of Cornwall, UK. Although the rate of whitecapping has a great influence on the output wave parameters, Moeini et al. (2012) showed that adaptation of model parameters is not a comprehensive methodology for accurate prediction of wide-range wave data.

The geographical space was discretized into 200×160 cell grid over the Persian Gulf with a 0.05×0.05 degree resolution in x and y directions. The origin of the computational grid was selected at 47 E, 23 N in the spherical coordinates (Fig. 1). 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 second and the highest was nearly 17 seconds covering typical wind waves in the Persian Gulf. The computational time step was set as 10 minutes as well. It should be noted that all of the model discretization parameters such as discretization of time, space and spectrum (frequency and direction) were selected by sensitivity analysis and then optimized by trial and error.

4. Error distribution approaches

Two different approaches were used to distribute the error of the predicted wave heights over the modeled domain. In the first approach, the errors at the desired grid points were estimated using the linear combination of the predicted errors at the observational points. If the errors are known at m observation points and need to be determined at n grid points (Fig. 2), the modified wave heights, based on the linear combination of the errors, can be written as:

$$H_i^a = H_i^f + \left(\sum_{j=1}^m g_{ij} E_j^o \right) / \left(\sum_{j=1}^m g_{ij} \right) \quad (2)$$

$$i = 1, 2, \dots, n$$

$$j = 1, 2, \dots, m$$

where, H_i^a is the updated (assimilated) wave height at grid point i , H_i^f is the simulated (first-guess) wave height at that grid point. E_j^o is the error of the predicted wave height at the observation point j :

$$E_j^o = H_j^o - H_j^f \quad (3)$$

where, g_{ij} in Eq. 2 presents the weight of the error of each observation points in the assimilation process. In the present study, the values of g_{ij} were defined as a function of the distance between observation and grid points (r_{ij}):

$$g_{ij} = \exp \left[- \left(\frac{r_{ij}}{d} \right)^2 \right] \quad (4)$$

where, d is the influence radius (Kantha and Clayson, 2000) determined by trial and error.

In the second approach, it was assumed that the error ratios of the different grid points are the same as the ratios of the wave heights at these points. Some researchers employed this approach without using any weight for different points (e.g. Kumar et al., 2009). Here we used the same weight used in the first approach (Eq. 4). In this way, the assimilated wave height at the desired grid point is estimated as follows:

$$H_i^a = \frac{H_i^f}{\left(\sum_{j=1}^m g_{ij} \right)} \sum_{j=1}^m \left(g_{ij} \frac{H_j^o}{H_j^f} \right) \quad (5)$$

The above equation can be rewritten based on the error at the observation point as:

$$H_i^a = H_i^f + \frac{H_i^f}{\sum_{j=1}^m g_{ij}} \sum_{j=1}^m g_{ij} \left(\frac{H_j^o - H_j^f}{H_j^f} \right) \quad (6)$$

Where, $(H^o - H^f)_j$ is the value of the error at the observation point. It is worth mentioning that by contrast to the linear combination method, the physics of the wave generation is considered in this data assimilation method. In this method, the value of the error at a desired grid point ($H_i^a - H_i^f$) depends on the value of the error at the observation point as well as the first-guess value of the wave parameter at that grid point (H_i^f). Since the waves are an integrated effect of the generating wind fields in time and space (Cavaleri and Bertotto, 2004), the higher waves are usually simulated with larger errors. Therefore, considering the simulated value of the wave parameter at a desired grid point in the error distribution methodology improves the efficiency of the data assimilation approach.

5. Results and discussion

As mentioned in section 4, two different approaches were investigated for updating of the simulated wave parameters over the modeled domain. In this section, the results of these approaches are presented. Generally, approaches can be assessed by comparison of their results with measured data. Since the main goal of this study is to develop an efficient approach for improvement of the simulated wave parameters over the entire domain, the results of these approaches should be compared with accurate measurement over the simulated area. Until now, in-situ simultaneous wave data (at several points) are not available over the Persian Gulf area. Therefore, altimeter wave measurements are the only available data over the model domain that can be used for evaluation of the error distribution approaches. However, this altimeter data contain some errors (Krogstad and Barstow, 1999; Greenslade, 2001) and are not accurate enough to be used as the data base for assessment of the assimilation approaches. For example, the error of the altimeter wave height data has been estimated about 0.5 m by Greenslade (2001). If these data are being used for assessment of the error distribution approaches, some of the outcome error will be due to the error in altimeter data. Therefore, another source of data was used as the basic data for evaluation of the error distribution approaches and improvement of the model outputs over the computational domain. In this regard, the SWAN model was run using available ECMWF data and the results were considered as the observations (H^o). The simulation period covered 1994-1995 and the model was forced by 0.5 degree 6-hourly operational ECMWF. Since the focus of this study was on the assimilation method based on updating of output variables not on the modification of the model parameters, all model parameters such as bottom friction and whitecapping coefficient were set to their default values. Then, the simulated wave characteristics containing some errors (H^f) were produced. To produce these erroneous first-guess wave heights and periods, the input values of wind speeds were multiplied randomly by a factor between 0.75 and 1.25 and the wave model was executed again. The erroneous model results are called first-guess model, hereinafter. Sannasiraj et al. (2006) have also employed this method for producing of basic and first-guess data. It should be noted that all of the wave model conditions such as discretization of time, space and spectrum and physical parameters remained unchanged during simulation of the first-guess wave data. The simulation period was also same as the simulation of observations, i.e. 1994-1995.

The assimilation approaches were then assessed based on the observation (H^o) and first-guess (H^f) values in the modeled domain. This assessment was conducted by two scenarios, using one observation point and six observation points. In the first scenario, the observation point was considered to be at the Bushehr station where buoy measurements were available (Moeini et al., 2012). Five grid points were also selected to control the assimilation results. Fig.1 shows the location of the observation and grid points in the first scenario.

Figures 3 to 6 depict the scatter diagrams of the wave heights at the grid points using two assimilation approaches. As seen in these figures, both approaches improve the results.

Generally, by increasing the distance between the grid point and observation point the accuracy decreases. Additionally, in all cases, the accuracy of the second approach (based on the wave height ratio) is more than that of using the linear combination. Tables 1 to 4 present error measures of the data assimilation procedure at four selected grid points. In these tables, Assimilation 1 refers to the assimilation approach based on the linear combination and Assimilation 2 refers to the second approach based on the wave height ratio. As seen in Tables 1-4, at the nearest point to the observation point, i.e. point B81; the scatter index reduced from 26 percent to about 11 percent when using the linear combination and 8 percent when using the wave height ratio (assimilation 2) method. Point 88 has a distance about 330 km from the observation point. At this point, the scatter index reduced from about 25 percent to nearly 18 percent using the linear combination and to about 11 percent using the second assimilation approach. In other words, using the wave height ratio for error distribution leads to a reduction of more than 50 percent in the scatter index at the most distant point. The reason for the skillfulness of the wave height ratio approach is that it considers the wave climate in the computational domain. Because of the nature of the error in the wind input data, the error of an output parameter depends on its simulated value and this is considered in the wave height ratio approach.

Similar results are found in the improvement of the wave period. For brevity, only the results for points B81 and 77 are presented. Figures 7 and 8 depict scatter diagrams of the wave period at the mentioned points using two assimilation procedures. As seen, by increasing the distance, the improvement in the results reduces. In other words, at point 77 which is located about 2.4 degrees from the observation point; significant improvement in the wave period is not seen especially for low values of wave period. On the other hand, the improvement in the wave height

was significant for this distant point. For the qualitative assessment of improvements, the relevant error measures are presented in tables 5 and 6. As seen, in both approaches the scatter indices reduce from 9 percent to less than 6 percent at point B81. This improvement is less at point 77. The reason of the less effectiveness of the assimilation procedure on the wave period (rather than on the wave height) may be the lower value of the produced error in the wave period. Comparison of the tables 1- 4 and 5-6 shows that the produced error in the wave height is about 26 percent and about 9 percent for the wave period. Therefore, since the correlations between the values of the errors at different points decrease when the error values decrease, the assimilation procedure is not very effective for the wave period. The reduction of the correlation of the errors between different points for lower values can be seen in the scatter diagrams of the wave height and period of all points. These diagrams depict that the scatter of the first guess data for high values of wave heights and periods is less than those of low values of wave heights and periods. After assimilation and improving the results, the scatter of the high values is significantly reduced while the scatter of the low values even increased in some cases. Next, the assimilation process was evaluated assuming 6 observational points. To control the results, four grid points (P1 to P4) were considered over the Persian Gulf (Fig. 9). Figures 10 to 13 depict the scatter diagram of the wave heights for first guess and updated data. Tables 7-10 show the statistical parameters of the assimilation process. As seen in the scatter diagrams and statistical parameters, when 6 observation points exist in the computational domain, updating of the model results led to a similar improvement of the predicted wave height at all grid points. The scatter indices decrease to about one fourth in the case of linear combination and to one fifth in the case of the wave height ratio (assimilation 2) approach. In addition, at all grid points, the assimilation approach based on the wave height ratio outperforms the linear combination one. This was observed and discussed for the case of one observational point. Another point is that the assimilation procedure yields the best results at point P2 and the worst ones at P4. This is due to the different values of the output parameters at these points. Based on the modeling results, point P2 is located in an area where the highest waves exist whereas the smallest waves are generated at P4 area. Since the error correlation between different grid points is less for lower values of the output parameters, assimilation procedure at point P2 which contains higher values leads to better results than other points. It should be noted that the best value for the influence radius was found to be nearly 2 degrees by trial and error.

6. Summary and conclusion

An efficient approach to distribute error of simulated wave heights over the domain was developed in this study. 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° . Two approaches for error distribution and improving the results were assessed. In the first approach, the value of the error at a desired grid point was estimated based on the linear combination of the error at the observation points. In the second approach, the errors of the output parameter of the model at a desired point were estimated based on the combination of the ratio of simulated values of the output parameters in the observational and grid points. The weight of each observational point was determined using an exponential function of the distance. The results of the study showed that the second approach leads to better results due to considering the physics of wind-wave generation.

The estimation of the weights based only on the distance between observation and grid points can influence the error distribution procedure. Thus, the selection of the gain values considering different wave regime is very important. In addition, optimum selection of the weight vectors to achieve minimum errors over the domain can enhance the quality of the error distribution process. Detecting highly correlated points over the simulation domain is also an interesting future research topic.

Acknowledgements

We acknowledge the Iranian National Institute for Oceanography, the Iranian Meteorological Office and the Iranian National Cartographic Center for providing the used data. We would also thank the SWAN group at the Delft University of Technology (Department of Fluid Mechanics) for providing the wave model. This study was supported by the Transportation Research Institute, Tehran, I.R. Iran, grant no. 89B8T8P09.

References

- Al-Salem, K., Rakha, K.A., 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, J.R., Cavaleri, L., Filipetto, V., Lefevre, J.M., Wittmann, P., 2007. Comparison of wind and wave measurements and models in the Western Mediterranean Sea. *Ocean Engineering* 34(3-4), 526-541.
- Babovic, V., Cañizares, R., Jensen, H.R., Klinting, A., 2001. Neural networks as routine for error updating of numerical models. *Journal of Hydraulic Engineering* 127(3), 181-193.
- Babovic, V., Sannasiraj, S.A., Chan, E.S., 2005. Error correction of a predictive ocean wave model using local model approximation. *Journal of Marine Systems* 53(1-4), 1-17.
- Booij, N., Ris, R.C. Holthuijsen, L.H., 1999. A third-generation wave model for coastal regions 1. Model description and validation. *J. Geophys. Res.* 104(C4), 7649-7666.
- Caires, S., Sterl, A., 2005. A new nonparametric method to correct model data: Application to significant wave height from the ERA-40 re-analysis. *Journal of Atmospheric and Oceanic Technology* 22(4), 443-459.
- Cavaleri, L., Bertotti, L., 2004. Accuracy of the modelled wind and wave fields in enclosed seas. *Tellus A* 56(2), 167-175.
- Cavaleri, L., Sclavo, M., 2006. The calibration of wind and wave model data in the Mediterranean Sea. *Coastal Engineering* 53(7), 613-627.
- Emery, K.O., 1956. Sediments and water of the Persian Gulf. *Bulletin of the American Association of Petroleum Geologists* 40(10), 2354-2383.
- Goda, Y., 2000. Random seas and design of maritime structures. Advanced series on ocean engineering, 15. World Scientific Publishing Co.
- Greenslade, D. J. M., 2001. The assimilation of ERS-2 significant wave height data in the Australian region. *Journal of Marine Systems* 28(1-2), 141-160.
- Kantha, L.H. Clayson, C.A., 2000. Numerical models of oceans and oceanic processes. International geophysics series. Academic press.
- Komen, G.J., Cavaleri, L., Donelan, M., Hasselmann, K., Hasselmann, S., Janssen, P., 1994. Dynamics and Modelling of Ocean Waves. Cambridge University Press.
- Komen, G.J., Hasselmann, S. Hasselmann, K., 1984. On the existence of a fully developed windsea spectrum. *Journal of Physical Oceanography*, 14(8), 1271-1285.

- Krogstad, H.E., Barstow, S.F., 1999. Satellite wave measurements for coastal engineering applications. *Coastal Engineering* 37(3-4), 283-307.
- Kumar, R., Bhowmick, S.A., Ray, S., Bhatt, V., Surendran, S., Basu, S., Sarkar, A., Agarwal, V.K., 2009. Improvement in predictability of waves over the Indian Ocean. *Natural Hazards* 49(2), 275-291.
- Millar, D.L., Smith, H.C.M., Reeve, D.E., 2007. Modelling analysis of the sensitivity of shoreline change to a wave farm. *Ocean Engineering* 34, 884–901.
- Mínguez, R., Espejo, A., Tomás, A., Méndez, F.J., Losada, I.J., 2011. Directional calibration of wave reanalysis databases using instrumental data. *Journal of Atmospheric and Oceanic Technology* 28(11), 1466-1485.
- Moeini, M.H., Etemad-Shahidi, A., 2007. Application of two numerical models for wave hindcasting in Lake Erie. *Applied Ocean Research* 29(3), 137-145.
- Moeini, M.H., Etemad-Shahidi, A., Chegini, V., 2010. Wave modeling and extreme value analysis off the northern coast of the Persian Gulf. *Applied Ocean Research* 32(2), 209-218.
- Moeini, M.H., Etemad-Shahidi, A., Chegini, V., Rahmani, I., 2012. Wave data assimilation using a hybrid approach in the Persian Gulf. *Ocean Dynamics*, 1-13.
- Purser, B.H., Seibold, E., 1973. The principal environmental factors influencing Holocene sedimentation and diagenesis in the Persian Gulf. *Persian Gulf*, Berlin, 1–9.
- Rakha, K.A., Al-salem, K., Neelamani, S., 2007. Hydrodynamic atlas for Kuwaiti territorial waters. *Kuwait Journal of Science and Engineering* 34, 143-156.
- Refsgaard, J.C., 1997. Validation and intercomparison of different updating procedures for real-time forecasting. *Nordic Hydrology* 28(2), 65-84.
- Sannasiraj, S. A., Babovic, V., Chan, E. S., 2006. Wave data assimilation using ensemble error covariances for operational wave forecast. *Ocean Modelling* 14(1-2), 102-121.
- Sannasiraj, S.A., Goldstein, M.G., 2009. Optimal interpolation of buoy data into a deterministic wind-wave model. *Natural Hazards* 49(2), 261-274.
- Signell, R.P., Carniel, S., Cavaleri, L., Chiggiato, J., Doyle, J.D., Pullen, J., Sclavo, M., 2005. Assessment of wind quality for oceanographic modelling in semi-enclosed basins. *Marine Systems* 53, 217– 233.
- The SWAN team, 2007. SWAN User Manual (Cycle III version 40.72). Delft University of Technology, Delft.

Tomas, A., Mendez, F.J., Losada, I.J., 2008. A method for spatial calibration of wave hindcast data bases. *Continental Shelf Research* 28(3), 391-398.

Figure Caption

Fig. 1. The Persian Gulf, modeled area and the location of the observation and grid points used for assessment of the error distribution results

Fig. 2. The used grid and observation points in the data assimilation process

Fig. 3. Scatter diagram of the wave height at point B81. Left: first guess data, middle: updated based on the linear combination, right: updated based on the wave height ratio.

Fig. 4. Same as Fig.3 but for Point 65

Fig. 5. Same as Fig.3 but for Point 77

Fig. 6. Same as Fig.3 but for Point 88

Fig. 7. Scatter diagrams of the wave period at point B81. Left: first guess data, middle: updated based on linear combination, right: updated based on wave height ratio.

Fig. 8. Same as Fig.7 but for Point 77

Fig. 9. The location of the observation and grid points

Fig. 10. Scatter diagram of the wave height at point P1. Left: first guess data, middle: updated based on linear combination, right: updated based on wave height ratio.

Fig. 11. Same as Fig.10 but for P2.

Fig. 12. Same as Fig.10 but for P3.

Fig. 13. Same as Fig.10 but for P4.

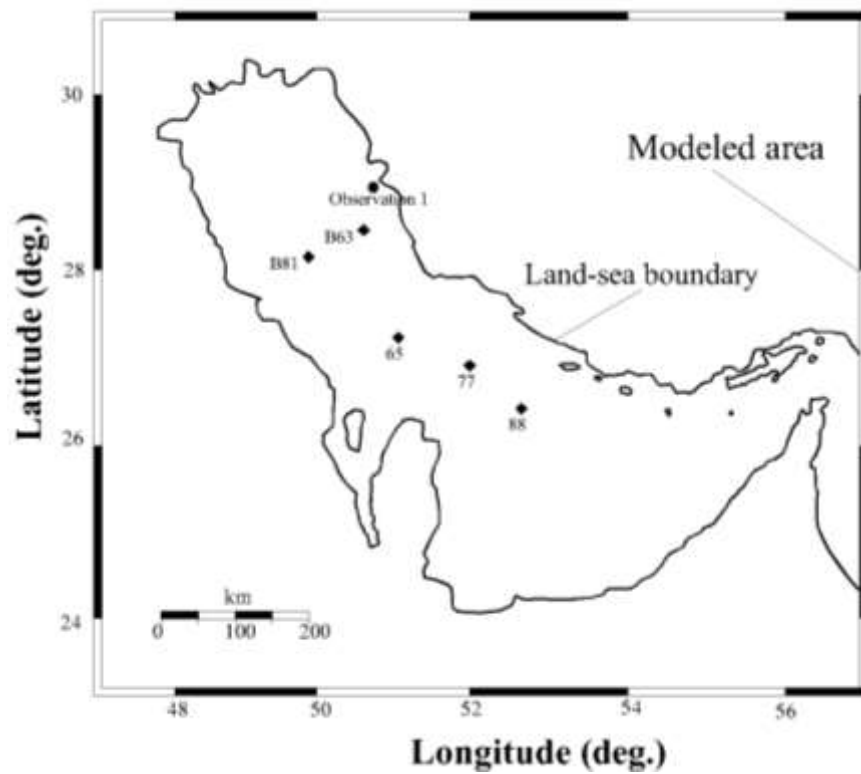


Fig 1. The Persian Gulf, modeled area and the location of the observation and grid points used for assessment of the error distribution results

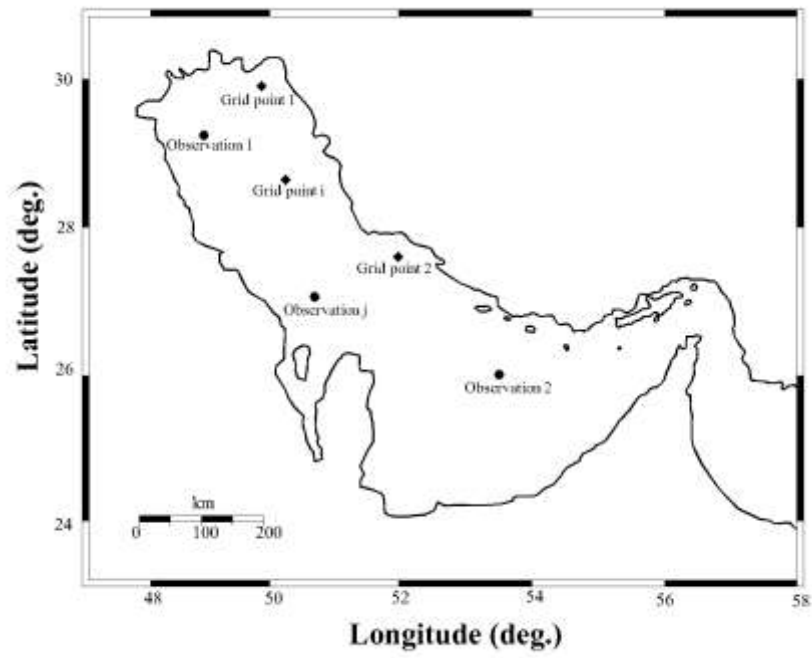


Fig. 2. The used grid and observation points in the data assimilation process

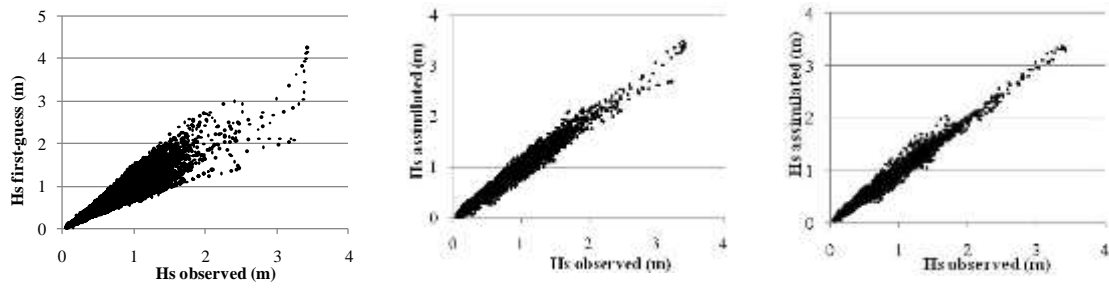


Fig. 3. Scatter diagram of the wave height at point B81. Left: first guess data, middle: updated based on the linear combination, right: updated based on the wave height ratio.

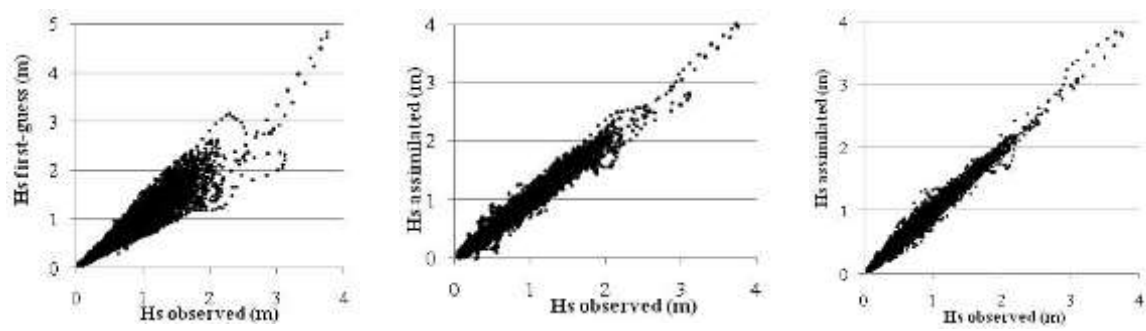


Fig. 4. Same as Fig.3 but for Point 65

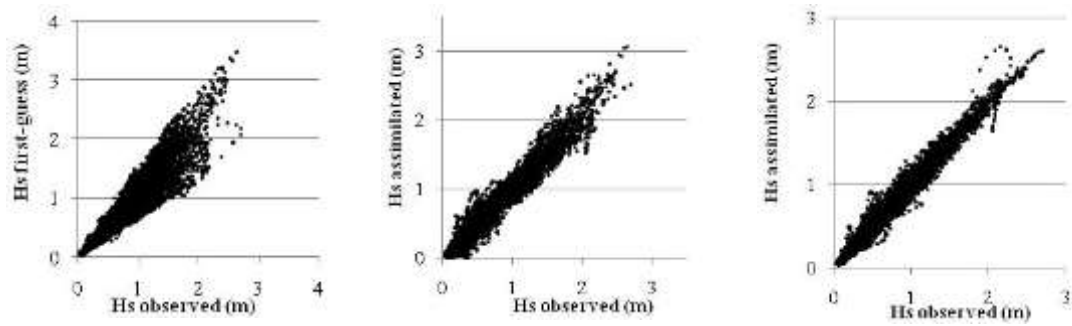


Fig. 5. Same as Fig.3 but for Point 77

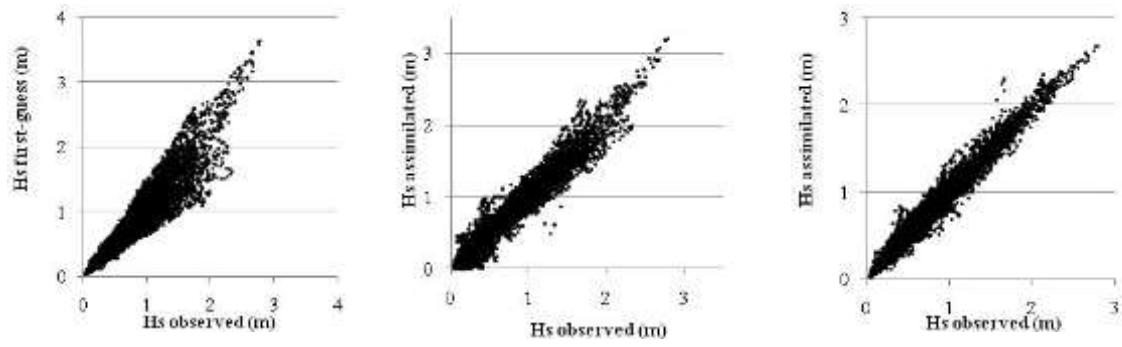


Fig. 6. Same as Fig.3 but for Point 88

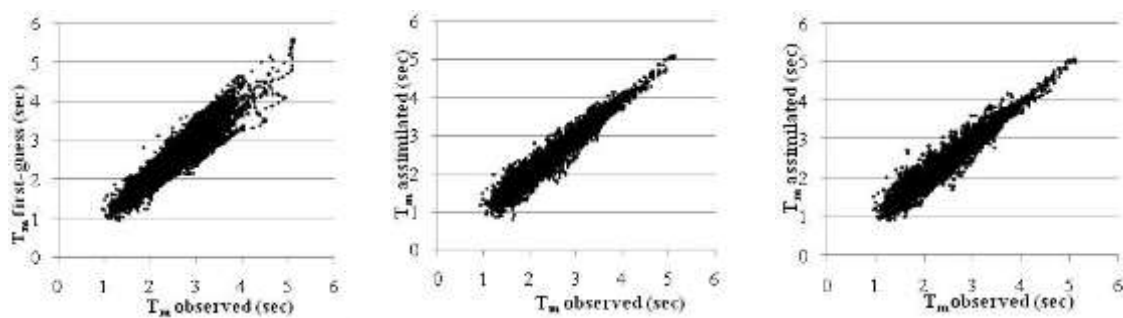


Fig. 7. Scatter diagrams of the wave period at point B81. Left: first guess data, middle: updated based on linear combination, right: updated based on wave height ratio.

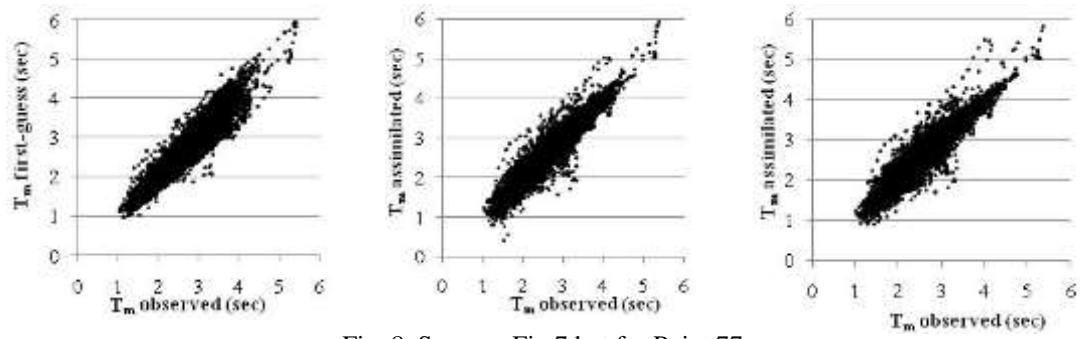


Fig. 8. Same as Fig.7 but for Point 77

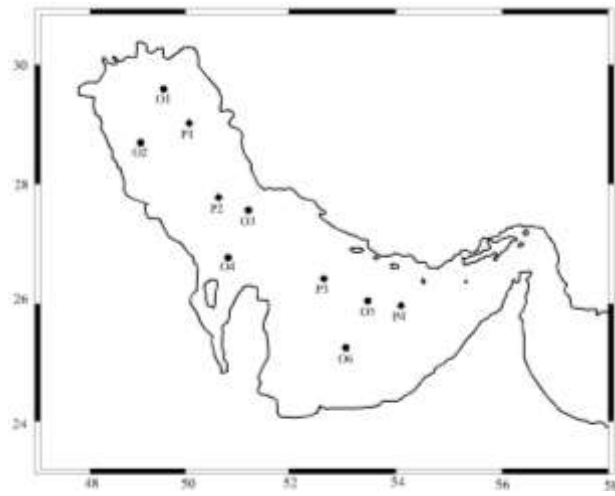


Fig. 9. The location of the observation and grid points

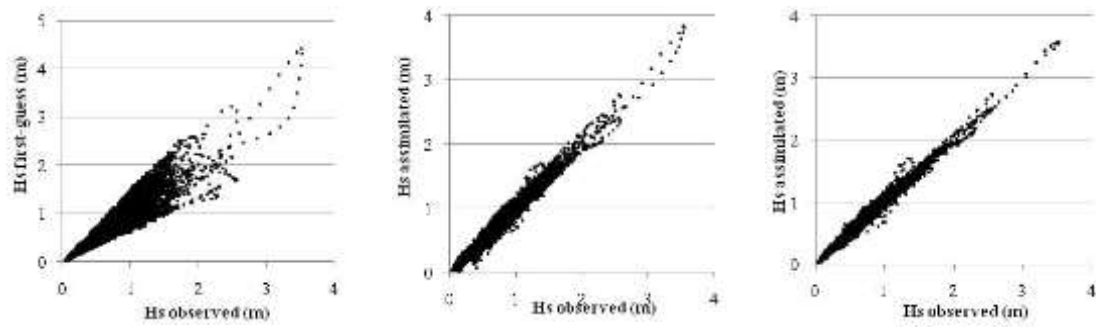


Fig. 10. Scatter diagram of the wave height at point P1. Left: first guess data, middle: updated based on linear combination, right: updated based on wave height ratio.

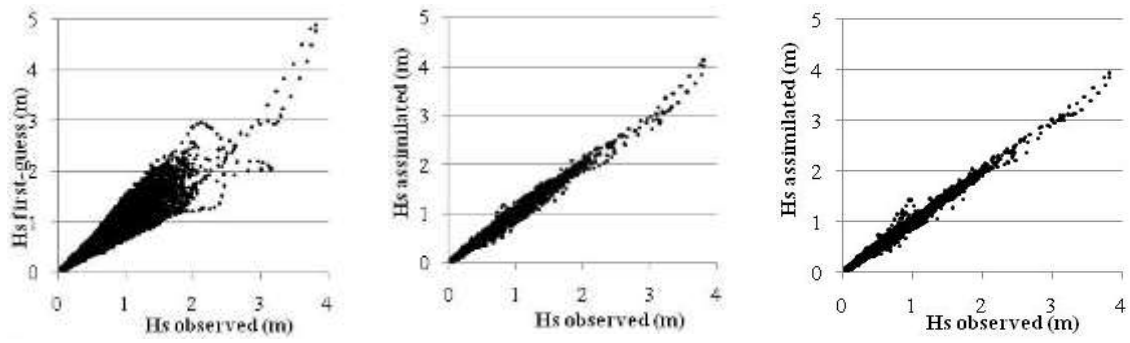


Fig. 11. Same as Fig. 10 but for P2.

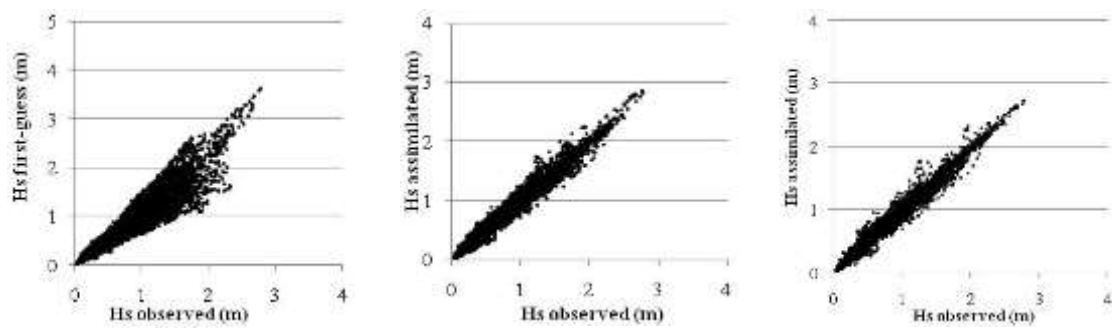


Fig. 12. Same as Fig. 10 but for P3.

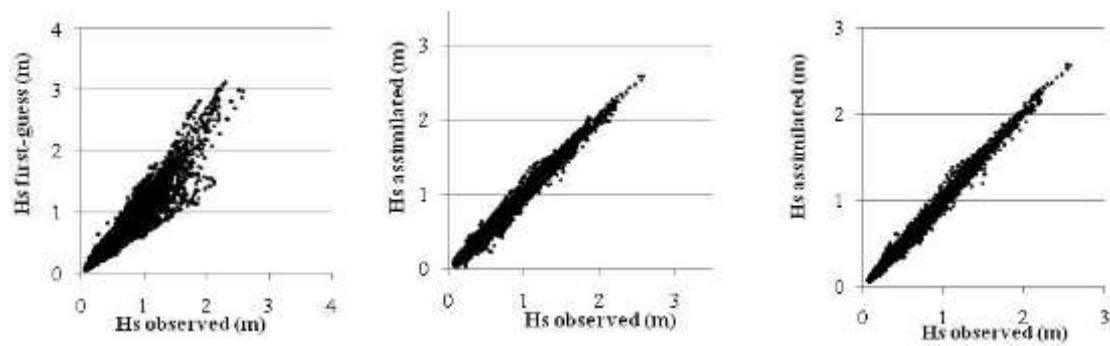


Fig. 13. Same as Fig. 10 but for P4.

Table 1. Statistical parameters of the data assimilation procedure at point B81

Point B81	First guess	Assimilation 1	Assimilation 2
Bias (m)	0.015	0.00	0.00
RMSE	0.143	0.06	0.04
SI (%)	26.009	10.82	8.02
Correlation Coef.	0.900	0.98	0.99

Table 2. Statistical parameters of the data assimilation procedure at point 65

Point 65	First guess	Assimilation 1	Assimilation 2
Bias (m)	0.019	0.00	0.00
RMSE	0.150	0.07	0.05
SI (%)	25.410	12.12	8.66
Correlation Coef.	0.905	0.98	0.99

Table 3. Statistical parameters of the data assimilation procedure at point 77

Point 77	First guess	Assimilation 1	Assimilation 2
Bias (m)	0.020	0.01	0.00
RMSE	0.145	0.09	0.06
SI (%)	25.37	15.60	10.28
Correlation Coef.	0.909	0.96	0.98

Table 4. Statistical parameters of the data assimilation procedure at point 88

Point 88	First guess	Assimilation 1	Assimilation 2
Bias (m)	0.022	0.01	0.01
RMSE	0.143	0.10	0.06
SI (%)	25.318	17.65	11.42
Correlation Coef.	0.915	0.95	0.98

Table 5. Statistical parameters of the assimilated wave periods, point B81

Point B81	First guess	Assimilation 1	Assimilation 2
Bias (m)	0.028	0.00	0.00
RMSE	0.206	0.13	0.13
SI (%)	9.055	5.66	5.61
Correlation Coef.	0.900	0.96	0.96

Table 6. Statistical parameters of the assimilated wave periods, point 77

Point 77	First guess	Assimilation 1	Assimilation 2
Bias (m)	0.035	0.01	0.01
RMSE	0.214	0.18	0.18
SI (%)	9.09	7.68	7.62
Correlation Coef.	0.909	0.93	0.93

Table 7. Statistical parameters for assimilation procedures at point P1

P1	First guess	Assimilation 1	Assimilation 2
Bias (m)	0.015	0.00	0.00
RMSE	0.143	0.04	0.03
SI (%)	26.227	7.23	5.52
Correlation Coef.	0.896	0.99	0.99

Table 8. Statistical parameters for assimilation procedures at point P2

P2	First guess	Assimilation 1	Assimilation 2
Bias (m)	0.017	0.00	0.00
RMSE	0.146	0.03	0.03
SI (%)	25.406	5.99	4.61
Correlation Coef.	0.904	0.99	1.00

Table 9. Statistical parameters for assimilation procedures at point P3

P3	First guess	Assimilation 1	Assimilation 2
Bias (m)	0.022	0.00	0.00
RMSE	0.143	0.04	0.03
SI (%)	25.318	7.52	6.00
Correlation Coef.	0.915	0.99	0.99

Table 10. Statistical parameters for assimilation procedures at point P4

P4	First guess	Assimilation 1	Assimilation 2
Bias (m)	0.022	0.00	0.00
RMSE	0.123	0.04	0.03
SI (%)	24.750	7.70	6.13
Correlation Coef.	0.915	0.99	0.99