Buoy data assimilation to improve wave height assessment in Bay of Bengal during monsoon seasons

Kalyani M1,*, Latha G1, Sannasiraj S A2 & Venkatesan R1

1National Institute of Ocean Technology, Pallikaranai, Chennai-600 100, India
2Department of Ocean Engineering, IIT Madras, Adyar, Chennai-600 036, India

*[E-mail: kalyani.niot00@gmail.com]

Received 23 June 2015; revised 04 August 2015

Significant wave height assessment in Bay of Bengal derived from third generation Wave Model (WAM cycle 4) has been improved by assimilating with continuous three hourly data from five moored buoys of National Institute of Ocean Technology using optimum interpolation technique. Covariance based weight function (gain) used in the study captures the model physics behind wind-wave process while distributing the errors at any buoy location. Gain composition at different buoy locations, due to assimilation of each buoy has been reported for both southwest (July) and northeast (November) monsoons, for year 2004. Qualitative and quantitative improvement in significant wave height due to buoy data assimilation during these two seasons are presented with reference to changing wind and wave characteristics. The order of improvement is found to be proportional to the magnitude of deviations found with respect to the measurements. Considerable model area has been influenced by assimilation.

[Key words: Significant wave height, WAM, Bay of Bengal, Moored buoy, Optimum interpolation, Assimilation]

Introduction

The wave climate in Indian Ocean is influenced by both summer (Southwest) and winter (Northeast) monsoons1. A methodology for the spatial calibration of wave hind cast data sources was developed based on empirical orthogonal function and a non-linear transformation of the spatial-time modes2. This method was applied to monthly long-term distribution of significant wave heights in the Western Mediterranean. This approach would not be applicable for the correction of hourly hind cast databases. In addition, since a prior distribution function of the data was assumed in the study area, this method could not be generalized.

An assimilation model based on optimal interpolation scheme was presented by using a gain function that had been formulated according to the process by which waves were generated by winds3. The efficiency of this method was evaluated by using buoy observations in the Arabian Sea. This assimilation approach led to a 30–50 percent reduction of root mean square error of wave height at the validation stations. But, in this method the gain function was formulated without considering the difference between wave climate in each season e.g. southwest monsoon, northeast monsoon and non-monsoon.

In the present study, the significant wave height characteristics in Bay of Bengal (BOB), derived from Wave Model (WAM) have been improved by using an assimilation model based on optimal interpolation scheme and by using a spatial error distribution function (gain function) which considers the underlying physical process by which wind generates the waves. This gain function has been formulated by considering the difference in wave climate during southwest (SW) and northeast (NE) monsoons for a hind cast study and the salient features have been discussed.

Materials and Methods

The third generation wave model WAM cycle 4 (The WAMDI Group)4 had been utilized to
simulate the wave characteristics in Indian Ocean. The model domain was extended from 30\(^{\circ}\) E to 120\(^{\circ}\) E in longitudinal and 50\(^{\circ}\)S to 30\(^{\circ}\) N in the latitudinal directions and covers both Arabian Sea and BOB in North Indian Ocean as well as the South Indian Ocean. The wind forcing was obtained from National Oceanic and Atmospheric Administration (NOAA) National Centres for Environmental Prediction (NCEP) winds of 1.00\(^{\circ}\) x 1.25\(^{\circ}\) degree resolution. The bathymetry was obtained from Earth Topography 2 Minute (ETOPO2) data sets. Grid spacing with 0.5 degree resolution was chosen for the model domain in both latitudinal and longitudinal directions. The significant wave height measurements were obtained from five buoys viz., MB12, MB11, DS3, DS4 and DS5, in Bay of Bengal during the year 2004 (Fig.1), as detailed in Table.1.

**Wave model**

Wave model WAM estimates the evolution of the energy spectrum for ocean waves by solving the wave transport equation explicitly without any presumption on the shape of the wave spectrum\(^4\).

\[
\frac{\partial F(f, \theta; x, t)}{\partial t} + \mathbf{v} \nabla_x F(f, \theta; x, t) = S
\]  

(1)

where, \(F(f, \theta; x, t)\) is the wave energy spectrum in terms of frequency \(f\) and propagation direction \(\theta\) at the position vector \(x\) and at time \(t\); \(\mathbf{v}\) is the group velocity. The second term on the left-hand side is the divergence of the convective energy flux \((\mathbf{v} \nabla_x F)\). The net source function, \(S\), takes into account all physical processes which contribute to the evolution of the wave spectrum. The source function is represented as superposition of source terms due to wind input, nonlinear wave–wave interaction, dissipation due to wave breaking, and bottom friction and is represented in the following form.

\[
S = S_{in} + S_{nl} + S_{ds} + S_{bot}
\]  

(2)

The wind input source function, \(S_{in}\), was adopted from Koman et al.\(^5\). The nonlinear wave–wave interaction term, \(S_{nl}\), represents the nonlinear conservative energy exchanges between all possible quadruplet wave components. The discretization of the governing equation in geographical and spectral space was performed using rectangular grid of 0.5 deg x 0.5 deg in both longitudinal and latitudinal directions. The details of the model equations and methods of solution were explained in WAM\(^1,5\).

**Optimum interpolation scheme:**

Optimum Interpolation (OI) is one of the efficient methods of assimilating measured data into a numerical model and is based on a statistical interpolation technique. The most widely adopted data assimilation schemes were sequential procedures like OI and successive correction\(^5,6\). Because of its relative adaptability, this method had been widely used for wave data assimilation. OI scheme was inbuilt into European Centre for Medium Range Weather Forecast (ECMRWF) system\(^2\) in which ERS-2 altimeter wave data was assimilated into a wave model, WAM. The weighting (gain) coefficients can be formulated with or without the dynamical model constraints. Improving the initial conditions for the wave model by updating input parameters.
using OI method\(^3\) revealed that the updated initial conditions could not spread over forecast period of more than twelve hours during severe wave climate with changing wind fields, as initial conditions would have no bearing on the wave field once it was developed. In the Kalman filter, the gain was computed and updated internally\(^5\).

OI-UOV method can be operated independent of time in the hind cast mode for any output parameter of the wave model. OI-UOV assimilation scheme could be executed independent of the wave model\(^3,9,10\) as all the corrections were carried out to the output variables of the model.

In the present study, the OI-UOV method had been applied to improve (update) the output variable (significant wave height, Hs) derived from the first guess WAM model output. This scheme distributes the error from discrete grid points to the entire model domain based on a weight matrix called gain matrix (G). The model improvement due to buoy data assimilation and its characteristics were discussed.

The state variable of the wave model (output) considered in the study was ‘significant wave height which was the ensemble parameter of the spectrum, \(F^s(f,\theta)\). Assimilation scheme could be written as

\[
F^a(f,\theta) = AF^f(Bf,\theta)
\]

where, \(F^a(f,\theta)\) was the analysed (corrected) wave spectrum computed from the first guess or the predicted wave spectrum, \(F^f(f,\theta)\) by rescaling the spectrum with two scale parameters \(A\) and \(B\). \(f\) and \(\theta\) were the frequency and directional components.

**Gain matrix:**

The wave height error evaluated at any location was the result of the cumulative effect of the interaction with neighbouring grids, and was connected in space and time, based on the wind-wave characteristics that vary with season. The H, at a location was also governed by site-specific characteristics. Therefore, error correction to H, evaluated at any single observation station should be meaningfully distributed to other model grids based on G that includes all these effects in it.

In a conventional OI technique, the coefficients of the gain matrix were mainly evaluated as a function of the distance between the observation station and the station of interest and its relative direction. The gain matrix would become more meaningful if the gains were estimated either from the model physics or from the knowledge of propagation of the state variables in some sense. In this study, the gain contours were formulated based on the physics of the modeling system and the observational error by adapting a covariance based weight function (G). This approach was well established and had been widely used by meteorologists in the climate studies\(^11,12\).

At each time level, all observations available in a time window centred at the time level were used to construct the ‘most likely’ estimate of the field at that time level. The basic idea was to compute an analysed or corrected value \(H^a\) of the true state (by assuming negligible measurement errors), \(H^f\). For \(i = 1,\ldots, n\);

\[
H_i^a = H_i^f + \sum_{j=1}^{m} G_{ij} \left( \alpha_j^f - \alpha_j^o \right)
\]

where, ‘\(H^f\)’ denotes the first guess field predicted from wave model, ‘\(\alpha^o\)’ denotes the observations from m stations, and ‘\(\alpha^f\)’ denotes their first guess counterparts. It was assumed that there exists a relation between the measured variable \(\alpha\) and the analysed field variable \(H^a\), and the transformation between these variables was accommodated in the gain matrix, \(G\). In the wave assimilation, the wave height was one of the measured variable, and also, the ensemble spectral parameter in the predictive model. The number of grid points which need to be updated were denoted with index ‘\(n\)’ and the number of observations (number of buoys in the present case) were denoted by the index ‘\(m\)’. Let \(H^o (= \alpha^o)\) and \(H^f (= \alpha^f)\) be written as a mean value (\(\langle H^o \rangle\) and \(\langle H^f \rangle\)where ‘\(\langle \cdot \rangle\)’ indicates temporal mean value) and a deviation (\(\delta\)),

\[
H^o = \langle H^o \rangle + \delta^o; \quad H^f = \langle H^f \rangle + \delta^f
\]

For obtaining best estimate of the variable, \(H^o\), the mean square error \((E)\) of analysed significant wave height \(H^a\) had to be minimized. The minimization of mean square error solves the interpolation weights \(G_{ij}\)\(^3\). The covariance between the observations and background errors could be neglected by the assumption of weak interaction. The optimal distribution of observed errors to the other grid locations was carried out by using Eq. 4 which was suitable for wave data such as buoy and radar altimeter derived wave heights.
Results and Discussion

The capability of the OI-UOV scheme had been demonstrated by assimilating significant wave height ($H_s$) data from five buoys in BOB during SW (July) and NE (November) monsoon period for the year 2004 with WAM output. $H_s$ was calculated at 3-hour interval throughout the simulation period.

The hind cast study was performed in offline mode but OI scheme can also be used in real time mode. Gain matrix for each season varies as per wind direction. Hence separate G was formulated for SW and NE monsoons. Monthly gain matrix had been generated with 3-hourly significant wave height date sets while care was taken to see that the length of time record clearly depicts a meaningful phenomenon and would not disrupt the physics in the averaging process (e.g. without splitting a peak during a monsoon). Minimum one month data were required to deduce seasonal variation in wave characteristics and to draw meaningful conclusions though more data if available would be good within each season. The calibration of G was performed by formulating G for one period and calibrating for another period. Calibration of MB12 buoy for SW and NE monsoons was shown in Fig. 2(a-b) respectively. For both SW and NE monsoons, the gain was 1.0 (100%) at self- buoy location (MB12 on MB12) while at neighbour buoys (DS5 or DS4 shown as an example) the gain was less than 1.0 and depends on the wind direction, season and the placement of buoys.

Gain contours for each buoy:

Single buoy assimilation had been carried out for each of the five buoys. Gain contours for SW and NE monsoon months in BOB were compared in Fig. 3(a-j) for all the five buoy stations.
The gain coefficient at buoy location was 1.0 and gradually decreased away from its location based on the wind direction. These plots indicate the weights given to each grid in correcting the model error in $H_S$ at a buoy location and were given in terms of the model error of the buoy under consideration. The entire model domain in BOB had been influenced by the error corrections due to these five buoys.

Assimilation of DS4 (Fig. 3(a)) seem to benefit only MB12 (which was in close proximity to the buoy assimilated, DS4) during SW month (July) when wind direction was from SW and away from neighbouring buoys. During NE monsoon month (November) when the wind direction was from NE (towards neighbouring buoys) all the buoys in the upwind direction were benefited with 90% to 40%, depending on their proximity to DS4 buoy (Fig. 3(b)). The farthest being DS3 with more than 40% gain. Assimilation of MB12 (Fig. 3(d)) and MB11 (Fig. 3(f)) also benefited all the neighbour buoys during NE monsoon with varying degree. MB12 assimilation brought significant gain (70%) at DS5 (on the upwind side) as the wind direction was from NE when compared to SW month (30%) with wind acting in opposite direction. MB11 was situated approximately at a centre location out of the 5 buoys in BOB and was in close proximity to all the neighbour buoys. Around 60% and more gain was observed for all the buoys during both the seasons (Fig. 3(e-f)) while at DS4 which was slightly off the wind direction indicated minimum gain of <20 % during SW monsoon.

Assimilation of DS5 (Fig. 3(g-h)) also showed more gain during NE monsoon at neighbour buoy locations except at DS3, which showed 25% increase in gain during SW monsoon. Assimilation of DS3 (Fig. 3(i-j)), which was on the southern-most side of all buoys, showed either equal or more gain during SW monsoon when the gain contours were bent in WSW-ENE direction when compared to that of NE monsoon. The exception being at DS4 which showed around 50% gain in NE month and 20% during SW month owing to its placement with respect to the wind direction during the two seasons. The seasonal variation in gain coefficients at neighbouring buoy stations was clearly evident from these figures.

**Error redistribution:**

After assimilation, at each model grid, the gain coefficients due to all the five buoys were available. Cumulative gain at each grid has been evaluated by redistributing these gains based on their respective covariances with the $H_S$ at grid under consideration, to avoid over correction$^5$. Using these redistributed gains the cumulative improvement in $H_S$ was found as per Eq. 4.

**Improvement in $H_S$ due to neighbor buoys:**

Maximum wave heights during SW monsoon for the year 2004 were approximately twice that of during NE monsoon. The correlation coefficients (CC) at each buoy station between the measured $H_S$ and the WAM derived $H_S$ (before and after the assimilation with neighbour buoys) had been plotted in Fig. 4(a-b) for SW and NE monsoons respectively.
It was to be noted that, the CC due to self-buoy assimilation alone was 1.0, which indicated full correction at its own site (e.g: MB12 on MB12 in Fig. 2) It was seen that neighbour buoy assimilation also considerably improved the CC (Fig. 4). Before assimilation the CC value varied between 0.74 at DS5 to 0.88 at MB11 and MB12 buoy sites during SW monsoon. After neighbour buoy assimilation CC was improved to 0.90 to 0.96 respectively. Before assimilation the CC value varied between 0.73 at DS4 to 0.85 at MB11 and DS3 buoy sites during NE monsoon. After neighbour buoy assimilation CC was improved to 0.89 to 0.95 respectively.

The reduction in root mean square error (RMSE) at buoy sites due to neighbour buoy assimilation had been quantitatively shown in Fig. 5(a-b) for SW and NE monsoons respectively. Model deviations were high during SW monsoon when compared to that of during NE monsoon. RMSE after assimilation with any buoy at its own site would be zero, while due to the assimilation of all neighbour buoys the reduction in RMSE was found to be considerable (Fig. 5).

During SW monsoon, the RMSE before assimilation was in the range between 0.25 m to 0.66 m at DS5 and DS4 respectively. With multiple neighbour buoy assimilation it was brought down to 0.11 m and 0.20 m respectively. The reduction in RMSE due to multiple neighbour buoys was found to be 57 to 74 % (ΔH_s achieved/ΔH_s required) at different buoy sites during SW monsoon. During NE monsoon, the RMSE before assimilation was in the range between 0.09 m to 0.36 m at MB11 and DS3 respectively. With multiple neighbour buoy assimilation it was brought down to 0.06 m and 0.26 m respectively during NE monsoon. The reduction in RMSE due to multiple neighbour buoys was found to be 30 to 50 % (ΔH_s achieved/ΔH_s required) at different buoy sites during NE monsoon. Neighbouring buoy assimilation effect could be used to improve model assessment during data gaps at validation stations.

**Overall H_s improvement in BOB:**
At any time instant, the cumulative improvement in H_s due to all five buoys at each model grid was evaluated using redistributed gains as discussed before. The gain matrix used
would be a steady state gain matrix for the entire period (one month). Hence, the improvement would be proportional to the model error at that instant. During peak events (high values) the model deviations would be generally high due to smoothening of wind vectors. Hence the model correction ($H_S$ improvement) also would be proportional to it. As an example, the $H_S$ improvement (in meters) in BOB on 27th July 2004 at 0:00 hours had been shown in Fig. 6(a), where the model errors at all the five buoy stations were considerable ($0.70$ m, $0.68$ m, $0.77$ m, $0.79$ m & $0.70$ m at DS4, MB12, MB11, DS5 & DS3 respectively before assimilation).

![Fig. 6(a)](image)

**Fig. 6(a)**—Improvement in $H_S$ due to assimilation of all five buoys in Bay of Bengal on 27th July 2004 0:00 hrs. (At any grid point all 5 buoy’s effect is considered. Effect of any buoy in its own location ($i=j$) is removed)

The effect of self-buoy at its own site (where $i = j$) had been removed to depict the neighbor buoy effect. Based on the wind direction during SW monsoon, the buoys on the upwind direction (DS4, MB12 & MB11) showed more correction in the order $0.4$ m to $0.5$ m when compared to that of DS5 and DS3, where, the improvement was between $0.3$ m to $0.4$ m. The improvement at any buoy site was found to be proportional to its model deviation.

The overall improvement in significant wave height at each model grid as a ratio of WAM derived model $H_S$ [$\Delta H_S/(H_S)_\text{Model}$] (as measurements are not available at each grid to know the actual model error at each grid) due to multiple buoy (five) assimilation for the month (averaged over a month) of July (SW monsoon) was shown in Fig. 6(b).

![Fig. 6(b)](image)

**Fig. 6(b)**—Improvement in $H_S$ due to assimilation of all five buoys in Bay of Bengal for SW monsoon (July). (At any grid point all 5 buoy’s effect is considered. Effect of any buoy in its own location ($i=j$) is removed)

The effect of each buoy at its own site (where $i = j$) had been removed and only neighbor buoy effect was shown. The similarity of the distribution could be seen from the two plots (Fig. 6(a) & (b)). Nearly 30 to 60 % improvement [$\Delta H_S/(H_S)_\text{Model}$] was observed in the neighborhood of the buoys in the BOB. From these plots, it was clear that buoy data at five locations could efficiently cover the entire BOB in improving the wave height assessment.

**Conclusions**

Buoy data assimilation using five buoys in Bay of Bengal, could correct the model error in $H_S$ by 100% at the respective buoy locations and also influenced the neighboring model domain considerably. Covariance based spatial correlation could efficiently capture seasonal variation in wave characteristics during SW and NE monsoons while evaluating the buoy influence in the neighboring domain. Excluding the self-buoy's effect on its own site, the overall improvement in $H_S$ in the model domain with respect to model $H_S$ at each grid [$\Delta H_S/(H_S)_\text{Model}$] due to five buoys was found to be up to 20% in southern BOB (south of 7.5°N), 20% to 40% in the central BOB (between 7.5°N to approximately 15°N) and was between 40%-60% in the northern BOB (north of 15°N) for SW monsoon season. The gain was found to be more on the upwind direction and its magnitude depended on the wind direction, proximity to the buoy assimilated and was proportional to that of the model error.
Acknowledgements
Authors would like to express their sincere thanks to Dr. M.A. Atmanand, The Director, National Institute of Ocean Technology, Chennai, for encouraging and supporting this work. Authors thank the Ministry of Earth Sciences (MoES), New Delhi and Indian National Centre for Ocean Information Services (INCOIS), Hyderabad for providing the moored buoy data.

References