Journal of Marine Science and Technology, Vol. 23, No. 6, pp. 943-950 (2015)
DOI: 10.6119/JMST-015-0610-13

IMPROVED BOUNDARY VALUES OF OCEAN WAVE FIELDS USING A DATA ASSIMILATION SCHEME

Yang-Ming Fan[1], Heinz Günther[2], Dong-Jiing Doong[3], and Chia Chuen Kao[3]

Key words: data assimilation, numerical wave model, altimeter data.

ABSTRACT

The objective of this study was to assess the accuracy of numerical wave forecasts through data assimilation of boundary values. A sequential data assimilation scheme was adopted to utilise altimeter data. The performance of the system in terms of operational applications, specifically for typhoon waves, was investigated. Two typhoons that occurred in 2006 around Taiwan (Kaemi and Shanshan) were used as case studies. The proposed data assimilation increased the forecast accuracy of the boundary values in terms of the wave parameters, such as the wave heights and periods. The results showed that after the assimilation, the assimilation model was significantly improved, especially the peak value of the wave and time of occurrence of the peak value.

I. INTRODUCTION

Data assimilation in operational wave modelling has been developed quickly over the past 20 years. The near real-time availability of wave and wind observations drastically increased after the launch of earth-observing satellites, such as the European remote sensing satellite-1 (ERS-1) and European remote sensing satellite-2 (ERS-2). Thus, many researchers have investigated the possibility of including data assimilation methods in operational wave forecasting systems to improve estimates of wave information.

Assimilation techniques for wave forecasting are commonly divided into sequential techniques (e.g., Lionello et al., 1992; Komen et al., 1994) and variation methods. Sequential methods are computationally cheap, and have successfully improved wave forecasts (e.g., Günther et al., 1993; Fan et al., 2014). Thus, these methods have been implemented into the operational wave analysis/forecast cycle at the European Centre for Medium-Range Weather Forecasts (ECMWF).

In sequential techniques, wave and wind observations are used to correct the winds and waves at each time step of the model, regardless of the previous model states. Because the space-time structure of the modelled wave field is not considered, the results are not completely consistent with the dynamics of the wave model. In the first attempt of wave data assimilation (Komen, 1985), this technique was used to improve swell forecasts in the southern North Sea with observed wave heights in the central North Sea. The model-predicted waves were replaced in the wave model with observations when and where such observations were available. However, the effect was relatively short-lived because the corrections were quickly lost due to the uncorrected winds and waves elsewhere in the wave model domain. Hasselmann et al. (1988) and Janssen et al. (1989) improved the effects by distributing the corrections over a larger area and by including wind corrections.

Over the past decade, the most frequently used operational assimilation schemes are single-time-step schemes, such as optimal interpolation (OI) (e.g., Janssen et al., 1989; Lionello et al., 1995; Hasselmann et al., 1997; Voorrips et al., 1997). This method is computationally fast; therefore, it is easily applicable to the online wave analysis/forecasting conditions. However, the method has some drawbacks. Unfortunately, forecast errors are often inhomogeneously distributed over the wave spectrum, limiting improvements obtained by the wave height assimilation alone (Mastenbroek et al., 1994). Thus, some groups were tried to use wave spectra from synthetic aperture radar (SAR) image spectra (Breivik et al., 1996; Hasselmann et al., 1997), which were useful for wave models of the world oceans.

The aim of the present paper was to investigate the potential use of altimeter data for assimilation in an operational forecast system and to provide reasonable boundary values for the following nesting layer. The impact of assimilation on the wave analysis was quantified by comparing runs with and without assimilation for several typhoons in 2006.

Paper submitted 11/17/14; revised 02/02/15; accepted 06/10/15. Author for correspondence: Yang-Ming Fan (e-mail: ymfan@mail.ncku.edu.tw).
[1] *Coastal Ocean Monitoring Center, National Cheng Kung University, Tainan, Taiwan, R.O.C.*
[2] *Institute of Coastal Research, Helmholtz-Zentrum Geesthacht Centre, Geesthacht, Germany.*
[3] *Department of Hydraulic and Ocean Engineering, National Cheng Kung University, Tainan, Taiwan, R.O.C.*

Table 1. The domain, grid resolution and time step of model nestings.

Nesting	Range	Grid resolution	Time step
1st layer	110°E~140°E/ 10°N~40°N	Δx = 0.250° Δy = 0.250°	60 min
2nd layer	119°E~125°E/ 20°N~27°N	Δx = 0. 067° Δy = 0.067°	30 min
3rd layer	121°E~124°E/ 21°N~25°N	Δx = 0.020° Δy = 0.020°	12 min

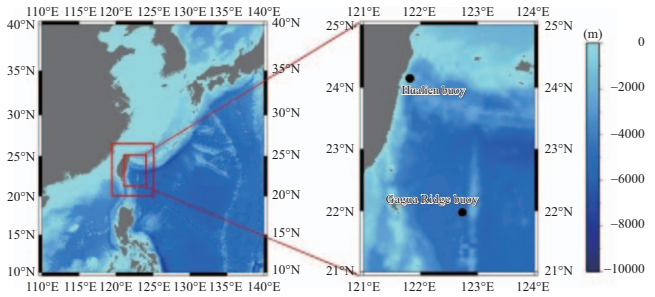

Fig. 1. Domain of the model and data buoy stations.

II. DESCRIPTIONS OF THE SIMULATION REGION

The study focused on the waters off eastern Taiwan. A three-level nesting scheme was applied to obtain detailed wave information in this region and to effectively simulate the wave field (see Fig. 1). The simulation domains, the grid resolutions and the time steps of the model nests are listed in Table 1. The larger region provided boundary values to the next finer layer. For this study, we concentrated on the fine-resolution of the layer 3 grid only. The SWAN wave model (Booij et al., 1999) was used for all layers.

All SWAN model runs were forced by the operational 1-hourly wind fields with a 0.5° resolution (longitude and latitude) provided by the Central Weather Bureau (CWB) of Taiwan. The fields were linearly interpolated over space and time.

The observed wave data from the Gagua Ridge buoy (122.78°E, 22.01°N) (see Fig. 1) were used for verification purposes. The Gagua Ridge buoy was located approximately 220 km east of Taiwan, where the water depth was approximately 6000 m. Pitch-and-roll buoys were developed, manufactured, and operated by the Coastal Ocean Monitoring Center (COMC) of National Cheng Kung University and were commissioned and supported by the CWB. The buoys report directional wave spectra every hour. Fast Fourier transform (FFT) was used to obtain the full two-dimensional wave spectrum (Brigham, 1988).

III. AN INTRODUCTION TO THE DATA ASSIMILATION SCHEME

Optimal interpolation (Hollingsworth, 1986) is a methodology used to construct the analysed significant wave height field. The optimal interpolation formulas (Lionello et al., 1992) are as follows.

The analysed wave height at each point x_i, denoted as H_A^i below, is expressed as a linear combination of H_P^i, which indicate the first-guess results produced by the model and H_O^k ($k = 1, \ldots, M_{obs}$), i.e., the observation:

$$H_A^i = H_P^i + \sigma_P^i \sum_{k=1}^{M_{obs}} W_{ik} \frac{H_O^k - H_P^k}{\sigma_P^k} \tag{1}$$

Here, σ_P^k is the root mean square error in the model prediction:

$$\sigma_P^k = \left\langle \left(H_P^k - H_T^k \right)^2 \right\rangle^{1/2} \tag{2}$$

where H_T^k represents the idealised true value of the wave directional spectra. The weights W_{ik} are chosen to minimise the root mean square error in the analysis of σ_A^k:

$$\sigma_A^k = \left\langle \left(H_A^k - H_T^k \right)^2 \right\rangle^{1/2}$$

The angle brackets indicate an average over a large number of realisations. Assuming that the errors in the model are unrelated to the errors in the measurements, the solution is:

$$W_{ik} = \sum_{m=1}^{N_{obs}} P_{im} M_{mk}^{-1} \tag{3}$$

where the element of matrix M is of the form:

$$M_{mk} = P_{mk} + O_{mk} \tag{4}$$

where P and O represent the error correlation matrices of the prediction and observation, respectively (both are actually scaled with σ_P^i):

$$P_{mk} = \left\langle \frac{\left(H_P^m - H_T^m \right)\left(H_P^k - H_T^k \right)}{\sigma_P^m \sigma_P^k} \right\rangle \tag{5}$$

$$O_{mk} = \left\langle \frac{\left(H_O^m - H_T^m \right)\left(H_O^k - H_T^k \right)}{\sigma_P^m \sigma_P^k} \right\rangle \tag{6}$$

Therefore, the prediction error correlation matrix P and the observation error correlation matrix O must be clearly speci-

fied. In practice, the statistics for both predictions and observations must be determined but are presently unavailable. If the idealised true value is known, then the RMSE between the observations and first guess results can be obtained. However, error exists in any observation technique; thus, we were unable to obtain the real observation. In this study, the prediction error correlation matrix is:

$$P_{mk} = \exp\left(-\frac{\left|\overline{x}_m - \overline{x}_k\right|}{L_{\max}}\right) \tag{7}$$

$$O_{mk} = \delta_{mk}\left(\sigma_O^m / \sigma_P^m\right) = \delta_{mk} R_m \tag{8}$$

where $\left|\overline{x}_m - \overline{x}_k\right|$ is the distance between grid point m and k. $L_{\max}$ is the correlation length. The effect of the variations in $L_{\max}$ and of the ratio between σ_O^m and σ_P^m on the results of the assimilation is discussed and verified below.

IV. ADJUSTMENTS OF THE OPTIMUM PARAMETER OF OI

We investigated the effect of the assimilation of real altimeter data on the SWAN model results. The data used were produced by ENVISAT from 18 July to 31 July 2006 and from 12 September to 19 September 2006. The wind fields were provided by the CWB.

A series of experiments were conducted. These experiments were given the same first step, during which the model was spun up for 14 days from 18 July to 31 July 2006 and for 7 days from 12 September to 19 September 2006. The spectra field at the end of the spin-up period was used as the initial condition for the assimilation experiments. The actual assimilation experiments began at the end of the spin-up period when the 12-hour data were assimilated at two time steps every hour (analysis period). At each time step, a field of analysed SWH was produced by OI and was subsequently used to construct the wave height. The analysed friction velocity was used to drive the wave model until a new stress field was provided by the atmospheric model. In the following 1.5 days (forecast period), the model was compared with the unused altimeter data to estimate the impact and advantage of the assimilation. The agreement between the model and the observations during the forecast period provided an estimate of the effectiveness of the assimilation. A comparison with a reference run, in which no assimilation was conducted, made it possible to study the decay of the effect of the assimilation. The length of the assimilation period was limited to half a day to analyse the spectrum only once at most of the grid points while providing coverage that was sufficient to analyse waves over most of the ocean.

This short experiment in which the actual length of the model run was only 2 days can be performed quite inexpensively, making it possible to produce a series of experiments

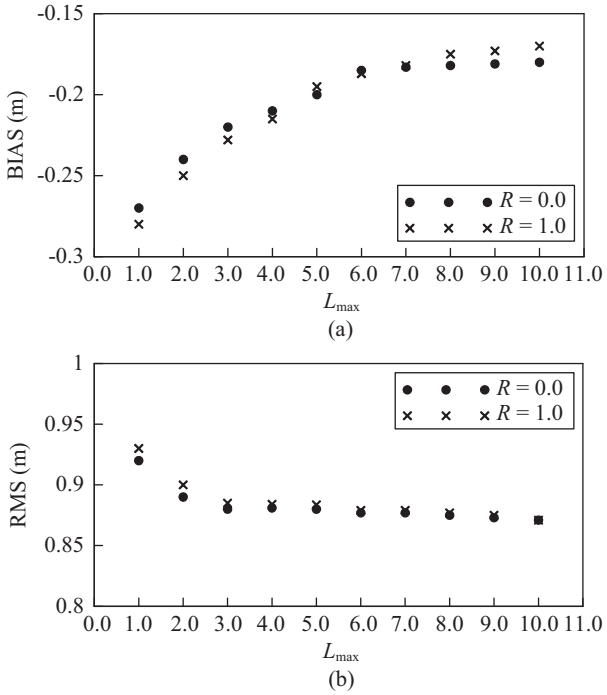

Fig. 2. Statistics of the model results compared with the altimeter observations for the entire forecast duration: (a) bias and (b) standard deviation. The reference run is represented by the solid circles (R = 0.0) and crosses (R = 1.0), where $R = \sigma_O/\sigma_P$.

and to obtain an evaluation of the effect of variations in the correlation length $L_{\max}$ and in the root mean square error of the observation σ_O^i. Because the method distinguishes between wind-induced waves and swells, the impacts of the two distinct contributions to the assimilation could be examined. Our discussion considered the dependence of the results on $L_{\max}$. We computed the statistics of the model results against the altimeter measurements. Specifically, we discussed the effect on the bias and root mean square error (rms):

$$\text{bias} = \frac{1}{N_{obs}}\sum_{j=1}^{N_{obs}} H_P^j - H_O^j$$

$$\text{rms} = \left[\frac{1}{N_{obs}}\sum_{j=1}^{N_{obs}}\left(H_P^j - H_O^j\right)^2\right]^{1/2} \tag{9}$$

Figs. 2(a) and 2(b) show the bias and standard deviation, respectively, for the entire forecast duration as functions of $L_{\max}$, which is expressed in grid units. All of the assimilation experiments significantly improved with respect to the reference run (whose bias and rms are represented by the solid line in both of the figures). The advantages increased as $L_{\max}$ increased. The differences are relevant for low values of $L_{\max}$, but the values eventually become more independent as $L_{\max}$ increases. In our opinion, this change was the consequence of separating adjacent satellite tracks (seven or eight grid points).

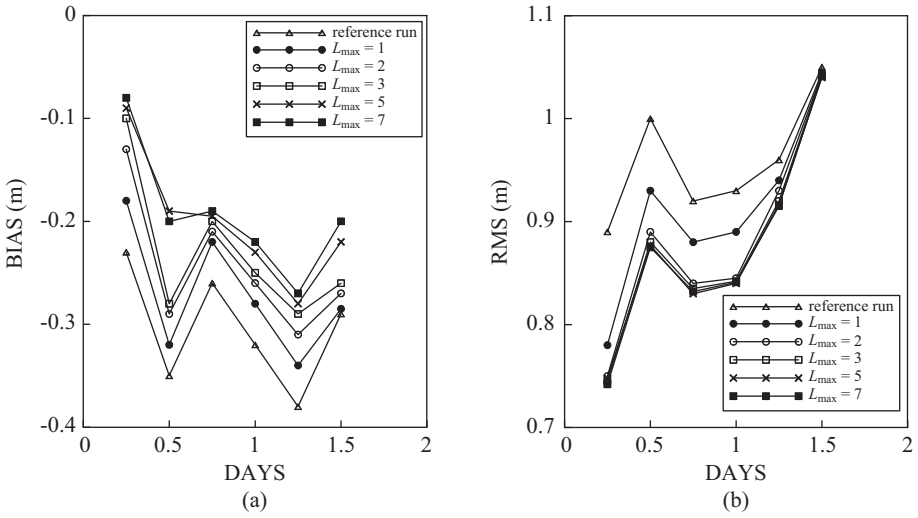

Fig. 3. Statistics of the model results compared with the altimeter observations computed every 6 hours during the forecast duration: (a) bias and (b) standard deviation. The triangles represent the reference run; the solid circles are L_{max} = 1; the open squares are L_{max} = 3; the plus signs are L_{max} = 5; and the solid squares are L_{max} = 7.

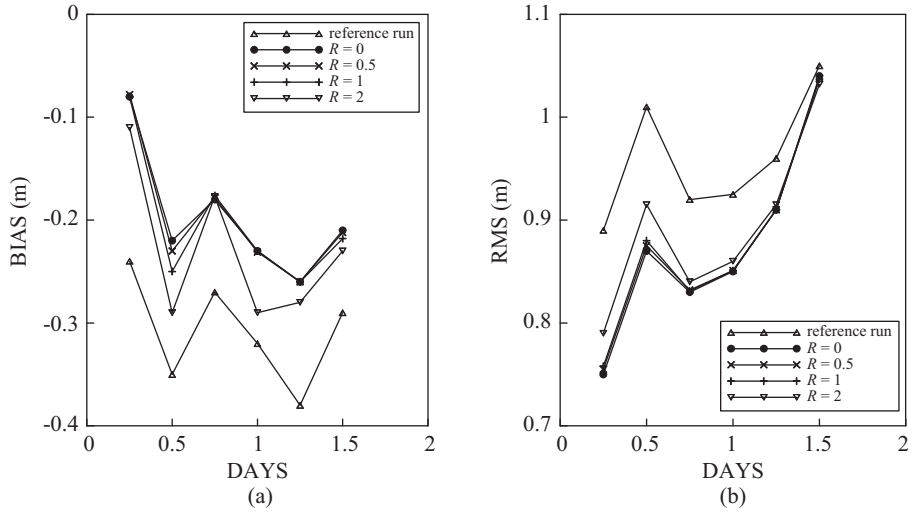

Fig. 4. Statistics of the model results compared with the altimeter observations computed every 6 hours during the forecast duration: (a) bias and (b) standard deviation. In all of the experiments, L_{max} = 5. The triangles represent the reference run; the solid circles represent R = 0; the crosses represent R = 0.5; the plus signs represent R = 1; and the nablas represent R = 2.

A satellite completes nearly seven orbits over 12 hours; therefore, almost any grid point was updated when L_{max} was larger than 4. A further increase impacted the use of the satellite measurements because it compensated for interruptions in the series along the tracks and locally changed the analysed SWH because more observations contributed to the interpolation; however, the resulting values were not substantially modified.

To investigate the persistence of the benefits of the assimilation, the statistics of the measurements were computed every 6 hours during the forecast period. The results are shown in Fig. 3 for the various L_{max} options. The absolute value decrease in the bias with respect to the reference run persisted for a long time. A reduction of 25% was still present 1.5 days after the end of the assimilation. The reduction of the standard deviation was less prolonged, but it was still approximately 10% 1 day after the end of the assimilation.

These numerical results indicated that the correlation length L_{max} was larger than 5. This value represents multiple situations. In fact, the waves had a spatial correlation similar to the spatial extent of the generating storm but smaller than the spatial correlation of the swell.

The results shown in Fig. 2 indicate that there was not a sensitive dependence on the value of the ratio $R = \sigma_O/\sigma_P$; in fact, the differences could not be considered statistically significant. Fig. 4 shows the temporal behaviour of the statistics for L_{max} = 5 and R = 0, 0.5, 1, 2. The choice R = 2 produced the

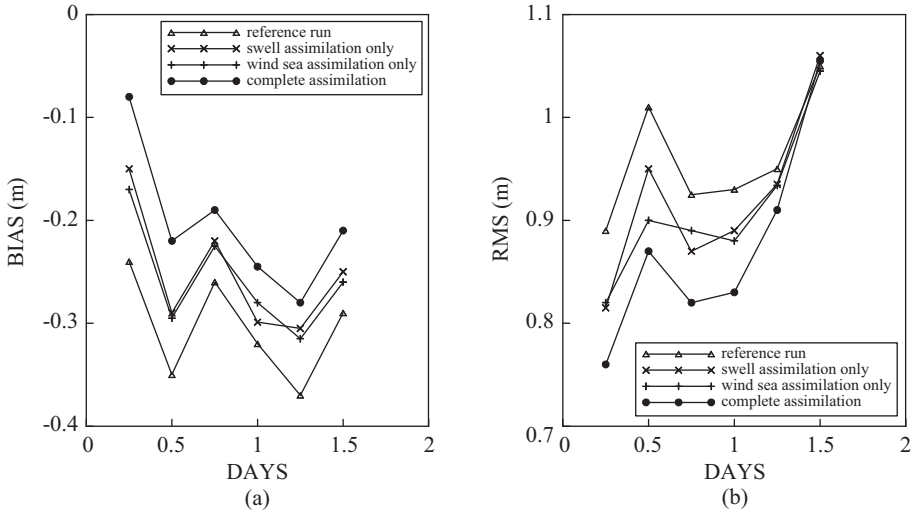

Fig. 5. Global statistics of the model results compared with the altimeter observations computed every 6 hours during the forecast duration (L_{max} = 5, R = 1): (a) bias and (b) standard deviation. The triangles represent the reference run; the crosses are the swell assimilation only; the plus signs are the assimilation only; and the solid circles are the complete assimilation.

worst results, whereas $R = 0$ and $R = 1$ produced quite similar results. Thus, the errors in the altimeter were comparable to or smaller than the errors in the SWAN model.

We briefly examined the impact of the assimilation by considering the case with $L_{max} = 5$ and $R = 1$. We simulated the differences between the assimilation and reference at the end of the assimilation period on 31 July and 19 September. Most of the assimilation effect was along the track of major typhoons. The general tendency of the model was to over-evaluate the measurements, with the exception of a typhoon that passed through our test area, where the SWH was under-evaluated. Because the over-evaluation was clearly distinguishable inside the generation areas, it was related to an over-evaluation of the wind speed. Consequently, the swell was generally over-evaluated as well. Notably, this situation was unlikely to greatly benefit from assimilation because at the end of the assimilation, the wind produced wave growth; thus, the energy decrease was ineffective. In contrast, when the energy of a wave system increased during the analysis, the advantages were evident; at the end of the assimilation, the analysed waves propagated over a field of less energetic waves, losing little energy and presenting a persistent pattern in the SWH. The individual features tended to move eastward because of the predominant direction of the waves generated by the typhoon in this area.

Two experiments were performed analysing only wave or swell spectra to investigate whether the wave or swell renewal was more effective. In the first case, the assimilation was only conducted at the points where most of the spectrum was waves (Eq. (10)). In the second case, only points where most of the spectrum was swell were renewed using Eq. (11). The waves were relatively localised on a global scale, which would suggest that updating of a swell could have a much larger impact. Moreover, the swell should have a longer memory. Although our experiments supported these arguments, they indicated that the impact of wave renewal was quite comparable to the impact of swell renewal. The difference was introduced by assimilating the swell and waves with respect to the assimilation run. The swell assimilation determined a much more widespread pattern, whereas the effects of the wave assimilation were very important in this region. One day after the end of the assimilation, the difference introduced by the swell assimilation was slightly larger than that introduced by the wave assimilation. However, the impacts of the two experiments were quite comparable due to the large differences introduced by the wave assimilation:

$$A = \frac{E_A}{E_p} B \quad B = \frac{f_{mP}}{f_{mA}} \tag{10}$$

$$A = \frac{E_A}{E_p} B \quad B = \Delta B \left(\frac{E_A}{E_P}\right)^{1/4} \tag{11}$$

The statistics supported the same conclusions. The statistics (Fig. 5) indicated that renewing the waves and swell had the same importance for a successive forecast. In fact, the two features were important in different regions. Fig. 6 shows the statistics that are limited to a region around the equator, where the swell was dominant and was responsible for almost the entire improvement obtained by the assimilation. The sea around the equator mostly contained swells radiated by mid-latitude storms. If the statistics were restricted to the southern part of the globe, then the effect of the wave analysis was more important because the storm influence on the waves was over-evaluated (see Fig. 7). These results indicated that a data assimilation approach should not be limited to the analysis of waves but that methods should also analyse the swell for successful global-scale results.

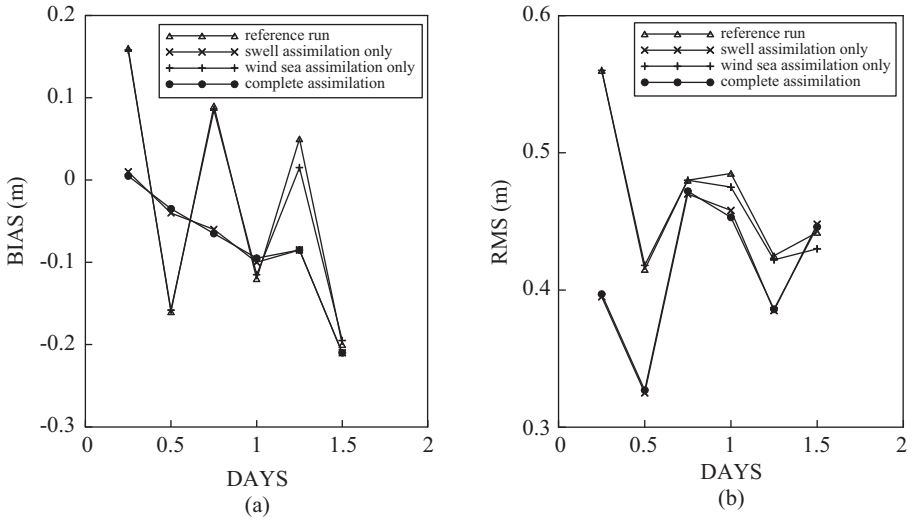

Fig. 6. Statistics of the model results compared with the altimeter observations computed every 6 hours during the forecast duration for the equatorial region (L_{max} = 5, R = 1): (a) bias and (b) standard deviation. The triangles represent the reference run; the crosses represent the swell assimilation only; the plus signs represent the wave assimilation only; and the solid circles represent the complete assimilation.

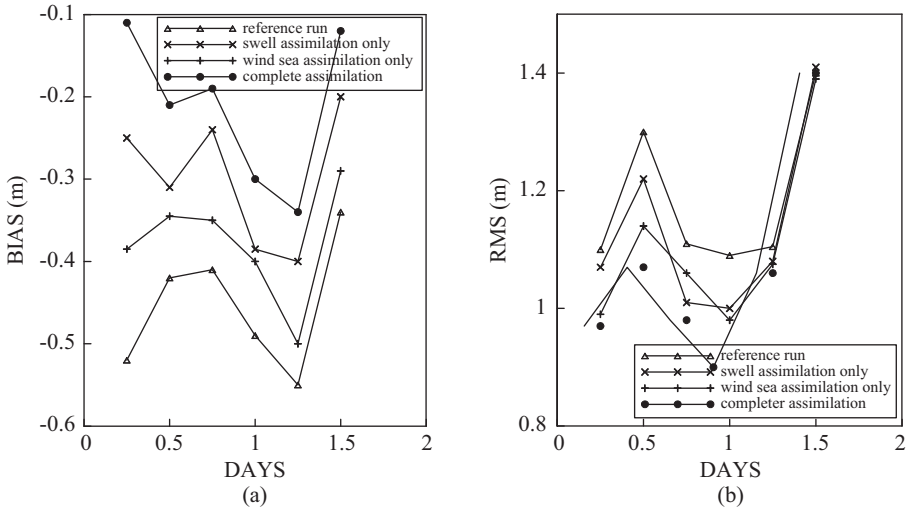

Fig. 7. Statistics of the model results compared with the altimeter observations computed every 6 hours during the forecast duration for the southern part of the oceans (L_{max} = 5, R = 1): (a) bias and (b) standard deviation. The triangles represent the reference run; the crosses represent the swell assimilation only; the plus signs represent the wave assimilation only; and the solid circles represent the complete assimilation.

V. VERIFICATIONS OF THE RESULTS FROM THE ASSIMILATION RUNS COMPARED WITH THE ALTIMETER DATA

We compared the results of the altimeter-data analysis with the buoy measurements, which were not used in the assimilation. The data were provided by CWB buoys for verifying the model results. To obtain an extensive dataset for the comparison, the assimilation of wave observations was conducted every hour during the typhoon periods. During these periods, the SWAN wave model was driven by the analysed wind fields produced by the CWB. The analysed wave data were provided by the altimeter mounted on the ENVISAT. A parallel reference run, without any assimilation, was conducted for comparison. The discussion involved some inspection of specific cases to explain the success or failure of the assimilation. Two relevant and representative situations are discussed by considering buoy measurements in Taiwanese water bodies.

The oceanic area around Taiwan was interesting because of the swell that was radiated by the typhoons. Fig. 8 shows the time series of the Gagua Ridge buoy and the model runs for the 4 previously mentioned cases. The arrows indicate the time when the spectra were compared in Fig. 9. The buoy-determined significant wave height was produced every hour, but the measurements were averaged over a 1-hour window for a more adequate comparison with the hourly model results.

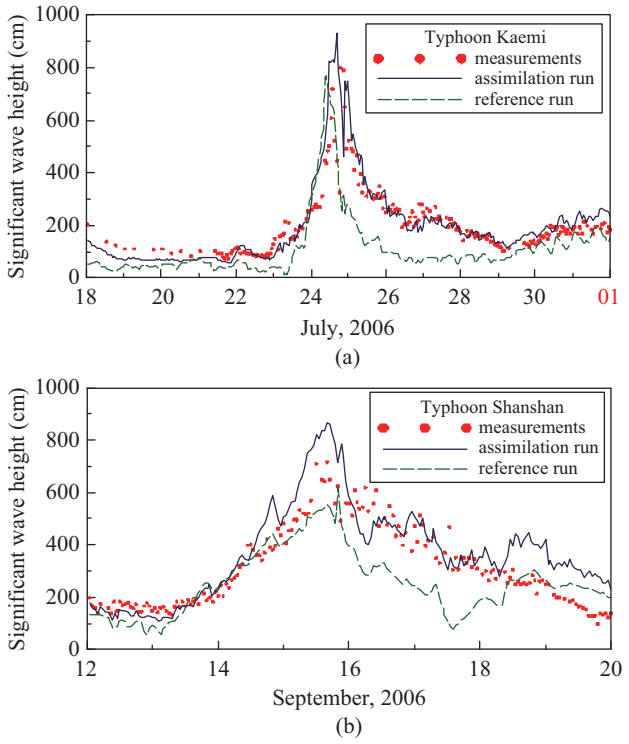

Fig. 8. SWH time series at Gagua Ridge buoy. (a) Typhoon Kaemi; and (b) Typhoon Shanshan.

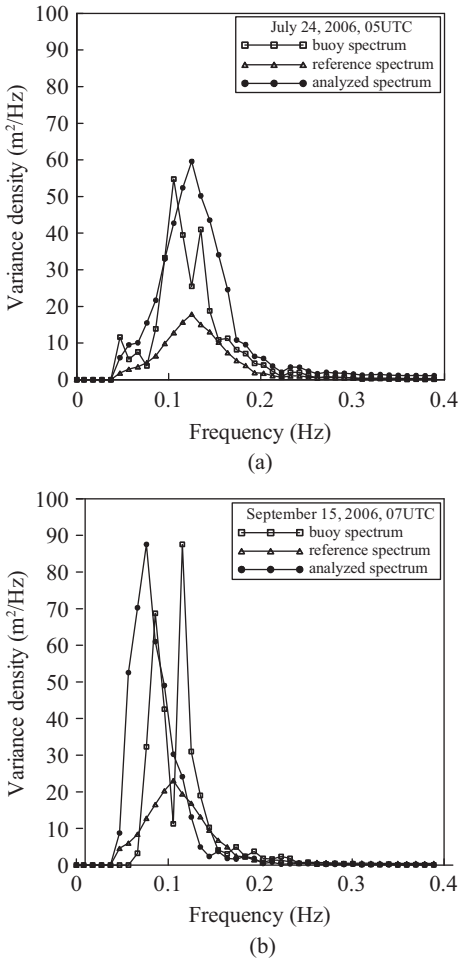

Fig. 9. Spectra at Gagua Ridge buoy for 24 July 2006 at 05:00 UTC, and 15 September 2006 at 07:00 UTC.

In Fig. 9, the buoy measurements are denoted with squares, the results of the reference run are denoted with triangles, and the results of the assimilation experiment are denoted with solid circles. The nearest grid point of the model was used for comparison. The impact of the assimilation was clearly positive: a series of relevant wave systems that was missing in the reference run but present in the buoy record was detected by the satellite. The improvements that correspond to the satellite passages were clearly distinguishable in the time series in Fig. 8.

VI. CONCLUSION

A sequential data assimilation scheme was adopted to utilise altimeter data. The operational performance of the system, specifically for typhoon waves, was investigated. The use of altimeter data to conduct data assimilation resulted in improved boundary values. Thus, the results could offer better boundary values for nesting.

The effect of the assimilation was irrelevant for some periods because of the presence of unreliable spikes in the data close to the buoy location, which prevented the use of data in a portion of the satellite tracks.

To evaluate the effect of variations in the correlation length $L_{\max}$ and R_m, the numerical results showed that the optimal value of the correlation length $L_{\max}$ was 5. The simulation analysis yielded a ratio between the observation and first guess with a standard deviation $R_m = 1$.

ACKNOWLEDGMENTS

This paper is part of the outcomes of the EU-FP7 PEARL project (THEME ENV.2013.6.4-3, Grant No: 603663) and funded supported by the Ministry of Science and Technology of Taiwan under grant no. MOST 104-2923-I-006-001-MY3. The field data used in this study are provided by Central Weather Bureau of Taiwan. The authors would like to express their sincere thanks.

REFERENCES

Booij, N., R. C. Ris and L. H. Holthuijsen (1999). A third-generation wave model for coastal regions, Part I, Model description and validation. Journal of Geophysical Research, 104, C4, 7649-7666.

Breivik, L. A., M. Reistad and H. Schyberg (1996). Assimilation of ERS SAR wave spectra in a numerical wave prediction model. DNMI Research Report 31, ISSN 0332-9879.

Brigham, E. O. (1988). The fast fourier transform and its applications. Prentice-Hall.

Fan, Y. M., H. Günther, C. C. Kao and B. C. Lee (2014). Assimilation of decomposed in situ directional wave spectra into a numerical wave model of typhoon waves. Natural Hazards and Earth System Sciences 14, 73-80.

Günther, H., P. Lionello and B. Hanssen (1993). The impact of the ERS-1 altimeter on the wave analysis and forecast. Report GKSS 93/E/44, GKSS Foreschungszentrum Geesthacht, Geesthacht, Germany.

Hasselmann, K., S. Hasselmann, E. Bauer, C. Bruening, S. Lehner, H. Graber and P. Lionello (1988). Development of a satellite SAR image spectra and altimeter wave height data assimilation system for ERS-1. Report 19, Max-Planck Institut für Meteorologie, Hamburg, Germany.

Hasselmann, S., P. Lionello and K. Hasselmann (1997). An optimal interpolation scheme for the assimilation of spectral wave data. Journal of Geophysical Research 102(C7), 15823-15836.

Hollingsworth, A. (1986). Objective analysis for numerical weather prediction. Journal of the Meteorological Society of Japan, WMO/IUGG/MWP Symposium Specification Issue, 11-60.

Janssen, P. A. E. M., P. Lionello, M. Reistad and A. Hollongsworth (1989). Hindcasts and data assimilation studies with the WAM model during the Seasat period. Journal of Geophysical Research 94(C1), 973-993.

Komen, G. J. (1985). Introduction to wave models and assimilation of satellite data in wave model in the use of satellites in climate models. ESA Scientific Publication SP-221, 21-25.

Komen, G. J., L. Cavaleri, M. Donelan, K. Hasselmann, S. Hasselmann and P. A. E. M., Janssen (1994). Dynamics and modeling of ocean waves. Cambridge University Press, Cambridge.

Lionello, P., H. Günther and P. A. E. M. Janssen (1992). Assimilation of altimeter data in a global third generation wave model. Journal of Geophysical Research 97(C9), 14453-14474.

Lionello, P., H. Günther and B. Hansen (1995). A sequential assimilation scheme applied to global wave analysis and prediction. Journal of Marine Systems 6, 87-107.

Mastenbroek, C., V. K. Makin, A. C. Voorrips and G. J. Komen (1994). Validaton of ERS-1 altimeter wave height measurements and assimilation in a North Sea wave model. Global Atmosphere and Ocean System 2, 143-161.

Voorrips, A. C., V. K. Makin and S. Hasselmann (1997). Assimilation of wave spectra from pitch-and-roll buoys in a North Sea wave model. Journal of Geophysical Research 102(C3), 5829-5849.