Proceedings of the Ninth International Conference on the Mediterranean Coastal Environment,
MEDCOAST 09, E. Özhan (Editor), 10-14 November 2009, Sochi, Russia

Spectral Wave Data Assimilation in SWAN Wave Model

Yang-Ming Fan(1), Beng-Chun Lee(2) and Chia Chuen Kao(3)

(1) ***Coastal Ocean Monitoring Centre, National Cheng Kung University,***
701 Tainan, Taiwan, R.O.C.
Tel: + 886-6-2744058 ext.15 Fax: + 886-6-2098853
E-mail: yangming.fan@gmail.com

(2) ***Department of Environmental and Hazards-Resistant Design, Huafan University, 22301 Taipei, Taiwan, R.O.C.***
Tel: + 886-2-26632102 ext. 4566 Fax: + 886-2-26639003
E-mail: beng@huafan.hfu.edu.tw

(3) ***Department of Hydraulic and Ocean Engineering, National Cheng Kung University, 701 Tainan, Taiwan, R.O.C.***
Tel: + 886-6-2757575 ext.63265 Fax: + 886-6-2098853
E-mail: kaoshih@mail.ncku.edu.tw

Abstract

The purpose of this study is to enhance the accuracy of numerical wave forecast with data assimilation. The present paper is to investigate the potential use of the spectral observations from the pitch-and-roll buoys which supply measurements in near-real-time for assimilation in an operational forecast system. And how to do optimal interpolation when we only have one buoy on deep ocean will be discussed in this study. And then, the impact of the assimilation of those measurements on the wave analysis and forecast is studied over several typhoon periods in 2006, by comparing runs with and without assimilation.

Introduction

Application of data assimilation to operational wave modelling is a quickly developing subject. It was only a decade ago that first attempts were reported to improve the wave forecast by correcting the wave field with observation. Since then, the number of near-real time available wave and wind observations has grown drastically because of the launch of earth-observing satellites, such as the ERS-1 and ERS-2. With the advent of the ERS earth-observing satellites, however, global wave observations have become available in near-real time. This situation has inspired many researchers to investigate the possibility of including data assimilation methods in their operational wave forecasting. There are two advantages of assimilating wave observations in a forecast model. First, assimilation may improve the estimate of the present sea state. In particular in the case of swell, this will have a beneficial impact on the wave forecast as well. All presently operational data assimilation schemes are aiming at this goal. The other advantage of data assimilation stems from the high sensitivity of the wave field to the forcing winds. In some case studies, promising results have been obtained (de Valk, 1994; de las Heras *et al.*, 1994; Bauer *et al.*, 1996), but as far as the authors are aware, no systematic improvement of past wind fields has yet been demonstrated under operational conditions (e.g., Voorrips and de Valk, 1997).

In past decade, the most frequently used operational assimilation schemes are single-time-level schemes such as optimal interpolation(OI) (e.g., Janssen *et al.*, 1989; Lionello *et al.*, 1995; Hasselmann *et al.*, 1997; Voorrips *et al.*, 1997). These methods are computationally fast and therefore easily applicable to on-line wave analysis/ forecasting conditions, but they suffer from some drawbacks. Unfortunately, forecast errors are often inhomogeneously distributed over the wave spectrum, which limits the improvement obtained by wave height assimilation alone (Mastenbroek *et al.*, 1994). Therefore, some groups have taken up the challenge of using the SAR data (Hasselmann *et al.*, 1997). Although this may turn out to be useful for wave models of the world ocean, for regional coastal, the density of SAR observations is simply too low to have a serious impact on the wave analysis. Also, the spatial resolution of the SAR about 100m is a larger problem for (partly) sheltered seas, where the average wavelengths are substantially shorter than on the ocean. However, for the regional coastal there is a good alternative to the SAR data. It is rather densely covered with pitch-and-roll buoys, which also measure spectral information. Moreover, their spectral characteristics are more suited to the typical wavelengths encountered, and, since they measure continuously at a fixed position, they supply more data than satellites for the region.

The aim of the present paper is to investigate the potential use of the spectral observations from the pitch-and-roll buoys which supply measurements in near-real-time for assimilation in an operational forecast system. And how to do optimal interpolation when we only have one buoy on deep ocean will be discussed in this study. And then, the impact of the assimilation of those measurements on the wave analysis and forecast is studied over several typhoon periods in 2006, by comparing runs with and without assimilation.

Spectral Observations from Buoys

The spectral observations from Gagua Ridge buoy is used in doing testing and spectral data assimilation. The data is from July 1, 2006 to September 30, 2006. Hualien buoy is used in verifying the results of simulation.

Observed parameters

The directional spectra can be obtained by analyzing the observed heave, pitch-and-roll motion of the buoy hull. From these time series, one can determine the one-dimensional energy-density spectrum $F(f)$ and some information about the directional distribution of the energy. If one writes the two-dimensional energy density spectrum:

$$F(f,\theta)=F(f)D_f(\theta) \tag{1}$$

$$D_f(\theta)=\frac{1}{\pi}\left\{\frac{1}{2}+\sum_{n=1}^{\infty}\left[a_n(f)\cos(n\theta)+b_n(f)\sin(n\theta)\right]\right\} \tag{2}$$

Only the Fourier components for n=1, 2 of the directional distribution can be determined from the auto-spectra and cross spectra. For convenience, we will drop the frequency dependence of the Fourier components a_n and b_n in the notation below.

One can try to reconstruct the full directional distribution as well as possible based on only these first four Fourier parameters (Long and Hasselmann, 1979; Lygre and Krogstad, 1986) or fit the data to an assumed shape (e.g., Longuet-Higgins *et al.*, 1963). Subsequently, one could assimilate the "retrieved" two-dimensional spectra, analogous to the way the SAR spectra are assimilated. The disadvantage of this effort is that, although the obtained spectrum may be "best" according to some criterion, it suggests much more knowledge about the spectrum than what is actually measured. Errors in the various components of the "reconstructed" spectrum will necessarily be strongly correlated, which obscures the comparison with, e.g., a spectrum obtained from a wave forecast model at the same place and time.

The same objections could be raised to the SAR assimilation procedure. The nature of the SAR data, however, is quite different from the buoy data. First the SAR observations themselves are strongly nonlinearly distorted images of the wave spectrum, so direct assimilation in an optimal interpolation scheme seems impossible here. Second, the long wave part of the inverted SAR spectrum seems to be only weakly dependent on external (model first guess) information (Brüning and Hasselmann, 1994). Pitch-and-roll buoys measure the directionality of the wave spectrum relatively crudely, so more external information has to be added for the low frequencies in an "inversion" procedure.

Instead of constructing the whole directional spectrum, we use the one-dimensional energy density spectrum $F(f)$, which is uniquely determined by the first few Fourier components. Here we distribute wave direction and frequency to several parts.

Partition of buoy data

For the assimilation of pitch-and-roll buoy data, one option is to apply the original partition scheme to a "reconstructed" full wave spectrum. We, however, prefer to use only the truly observed parameters and devise a new partition scheme based on these data. Of course, the new scheme should stay as close as possible to the partitioning scheme for the full spectrum. The scheme used the energy-density spectrum $F(f)$ and the direction $\theta(f)$. Consider the computer time; we do experiment of assimilation procedure to get the optimal choice of direction and frequency in the following.

The Data Assimilation Scheme

Assimilation of wave observations in operational wave forecast models is a relatively new subject. For a long time, the lack of near-real-time available wave observations impeded the development of assimilation systems. How can assimilation of wave observations improve the model performance? Basically, there are two ways. The first method is to use an extensive data set to optimize the model parameters. In this way, we improve the wave model itself. Since this is an off-line task, it can be done with advanced, time-consuming data assimilation techniques.

The second method is to use observations on-line, to draw the modelled sea state to the observations. In this way, the model analysis and short-term forecasts can improve. Especially swell forecasts are expected to be improved, since they are not very sensitive to the quality of forecast wind fields. Since the data assimilation now has to be done during the operational forecast cycle itself, the method must be not too time consuming. Simpler assimilation methods are therefore need.

At present, for wave assimilation, as for assimilation in circulation models, there are two main classes of assimilations: sequential methods and variational methods. The advantage of sequential method is the relative simplicity and the relatively low requirement for computer resources. The minimization of the cost function typically requires a series of runs if the wave model. Consequently, the computer resources needed are larger than those needed to execute the model itself, while the resources that are requested by a sequential method are, in comparison, negligible.

Assimilation of wave spectra from pitch-and-roll buoy

The OI-P scheme is described extensively in Hasselmann *et al.* (1996) and in Voorrips *et al.* (1997). OI-P is a so-called sequential assimilation method. This means that it is called during the wave model forecast at every time at which new observations are available; it combines the model state at that time (the first guess) with the new observations to calculate an analyzed model state; and this analyzed state serves as the initial condition for a new model run, until a new set of observations is processed, etc.

Two major approximations are made in formulating the OI-P method. The first is to define pre-calculated (and constant) forecast and observation error covariances, which are used at every assimilation time step. This distinguishes OI schemes from the far more costly Kalman filter method, in which the error covariances are explicitly propagated by the model dynamics. The second approximation is to apply a technique called spectral partitioning, with which the number of free parameters in a wave spectrum is effectively reduced by more than an order of magnitude. With these two approximations, a scheme can be constructed which is very cost-effective: an assimilation step takes considerably less time than the propagation of the wave model between two consecutive assimilation times.

The concept of spectral partition was introduced by Gerling (1992). It is a method to describe the essential features of a two-dimensional wave variance spectrum $F(f, \theta)$ (f frequency, θ direction) with only a few parameters. This is done by separating the spectrum into a small number of distinct segments, so-called partitions, which correspond to the various "peaks" in the spectrum. The partitioning is a purely formal procedure; however, the partitions can be interpreted physically as representing independent wave systems. Details of the formalism used to calculate the partitions can be found in Hasselmann *et al.* (1996).

The original partitioning scheme as was devised by Hasselmann *et al.* (1994, 1996) can only be applied to a full two-dimensional wave spectrum, such as a model spectrum or an inverted SAR spectrum. Pitch-and-roll buoy data, however, contain only the one-dimensional wave spectrum $E(f)$, plus limited information about the directional distribution. To assimilate these data as well, an adapted version of the partitioning scheme was developed (Voorrips *et al.*, 1997) which needs only $E(f)$ and the wave propagation direction $\theta(f)$ as a function of frequency. Tests with synthetic buoy spectra which were extracted from full spectra showed very good agreement between the two partitioning schemes.

The analyzed partition parameters from the optimal interpolation are now combined with the first-guess spectra to obtain analyzed spectra. Every first-guess partition is multiplied by a scale factor and shifted in the (f, θ) plane such that its mean parameters are equal to the parameters obtained by the optimal interpolation. Small gaps in the spectrum which arise by the different shifts for different partitions are filled by two-dimensional parabolic interpolation.

Therefore, a new assimilation method was developed to assimilate detailed spectral wave observations:

$$E_A^i(f,\theta) = E_P^i(f,\theta) + \sigma_A^i \sum_{j=1}^{N_{obs}} W_{ij} \frac{E_O^j(f,\theta) - E_P^j(f,\theta)}{\sigma_P^j} \tag{3}$$

Description of the simulation region

In order to simulation in the model, basic data for the period Nov. 1 to Dec. 15, 2003 (winter monsoon) was used to drive the model. In this case, the domain of the SWAN wave model covers longitude 100°E to 145°E and latitude 0°N to 45°N, with a 0.5° grid resolution in longitude and latitude (Fig. 1, inset map). The domain of the wind field data also covers longitude 100°E to 145°E and latitude 0°N to 45°N, with a 0.5° grid resolution in longitude and latitude at 1-h resolution.

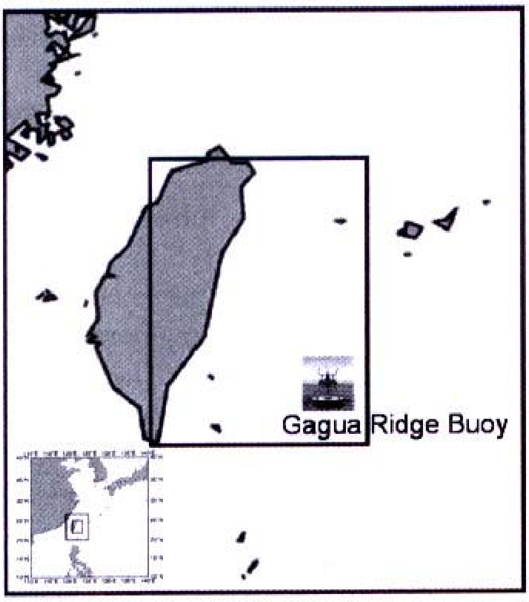

Fig. 1: Domain of model and buoy stations

In order to simulation wave effectively, we used the nesting scheme on the simulation region. It means the simulation region is larger, the grid resolution is rougher. Table 1 shows the domain, grid resolution and time step of model nesting. Because we focused on the eastern Taiwanese water in this study, we only need the detail wave information in this region. If the larger region used fine grid resolution, it should need lots of computer time. For the larger region, its mission is to offer bounder value for the next layer use.

Adjust the Optimum Parameter of OI-P

Optimal the frequency and direction

The next step in the assimilation procedure is to merge the model first-guess and observed partition parameters into an analyzed field of parameters. We have assumed that different partitions within a spectrum are uncorrelated since they are created by different typhoon events. So, we want to treat these partitions separately from each other in the assimilation. On the other hand, partitions in different spectra (e.g., model and observed spectra, or two model spectra at different locations) are correlated if they are created by the same event. Therefore, we have to define a cross-assignment criterion between the partitions of two different spectra.

The criterion which is used is based on the distance in spectral space between the parameters of two partitions. The ones which are closest to each other are cross-assigned. In case the number of partitions in the observed and model spectra do not match additional assumptions are needed.

Table 1: The domain, grid resolution and time step of model nesting

nesting	range	grid resolution	time step
1st layer	110.0°E-140.0°E 10.0°N-40.0°N	Δx=0.250° Δy=0.250°	60 min
2nd layer	119.0°E-125.0°E 20.0°N -27.0°N	Δx=0.067° Δy=0.067°	30 min
3nd layer	121.0°E-123.0°E 21.0°N -25.0°N	Δx=0.020° Δy=0.020°	12 min

When the cross-assignment is done, the parameters of the model and observed partitions can be combined to obtain an analyzed field of partition parameters. An important input for the OI-P procedure is the error covariances of the errors in the observed and the model parameters. The covariances were obtained by calculating long-term statistics of differences between observations and SWAN model hind casts. The observation errors were assumed to be spatially independent.

Although we only have one data buoy in the deep ocean, we use the first-guess spectra of neighbouring grid points of Gagua Ridge buoy as fictitious buoys data. Compare the wave spectral of assumed stations with field station; we could get the weight between assumed stations and field station. 3 months buoy data are used in doing statistic analysis during the process, and then we do OI-P by these wave spectra. The computer time will be influenced due to the number of direction and frequency. Therefore, 2 days assimilation in these experiment run is done to get the optimal choice (Table 2). The results show that the most accurate of model assimilation if we distribute wave direction to 32 and frequency to 41. But it spends a lot of computer time. In order to spend the less computer time and get higher accurate, we use 16 directions and 20 frequency for latter assimilation of typhoon events.

Optimal the number of assumed stations

Due to it needs two measured stations at least to do optimal interpolation, but we only have one Gagua Ridge buoy in this study. Therefore, we have to assumed stations for optimal interpolation use. Here we set 3 assumed stations, 5 assumed stations, and 7 assumed stations to do numerical test. The average errors of significant wave height and mean wave period for different assumed stations show in Table 3. It is obvious that the more assumed stations, the better accuracy in the model results. But considering the effective computer work, we choice 5 assumed stations for the following research.

Table 2: SWH RMSE statistics of the various numerical experiments against the data of Gagua Ridge buoy

Direction SWH RMSE(m) Frequency	8	16	32
10	0.85	0.73	0.71
20	0.77	0.47	0.43
41	0.58	0.39	0.34

Table 3: The average errors of significant wave height and mean wave period for different assumed stations

Average error	3 assumed stations	5 assumed stations	7 assumed stations
H_s (cm)	18.7	12.8	10.1
T_m (sec)	1.1	0.7	0.4

Verification of the Results from the Assimilation Run of Buoy Data against Buoy Observation

The influence of the assimilation on the wave analyses and forecasts was assessed by running the wave model with assimilation for a period of typhoon events. Runs were performed with the operational CWB wind field. In case the CWB wind fields were missing no run was done. These warm-up periods have been removed from the evaluation.

The main conclusion from that comparison is that the quality of these runs is comparable so the effect of OI-P assimilation in the SWAN model can be expected to be of the same order of magnitude as found in the results below which shows in Fig. 2 to Fig. 5. Direction wave spectra, one-dimensional spectra, and significant wave height are shows that assimilation run is closed to measurements. The reference run is lower than measurements. Fig. 2 shows the direction wave spectra at Hualien buoy for July 26, 2006, 05UT. The wave direction and wave energy in Fig. 3a is close to Fig. 3b. Fig. 4 shows the spectra at Hualien buoy for July 26, 2006, 05UT, and September 15, 2006, 15UT. We merge the spectra during assimilation procedure, so the wave spectra are similar after model assimilation. From the SWH time series and MWP time series (show in Fig. 4 to Fig. 5), the results show that model has great improve with data assimilation. The hind cast results always can't calculate the peak value or can't calculate the peak value at the right time if without data assimilation. Therefore, the data assimilation has good performance in SWAN wave model simulate.

Statistical Results over the Typhoon Periods

Compare OI-P scheme with OI-I scheme

We did a comparison of OI-P and OI-I analyzed wave spectra and significant wave height with measurements from the buoy. For the comparison the region is defined.

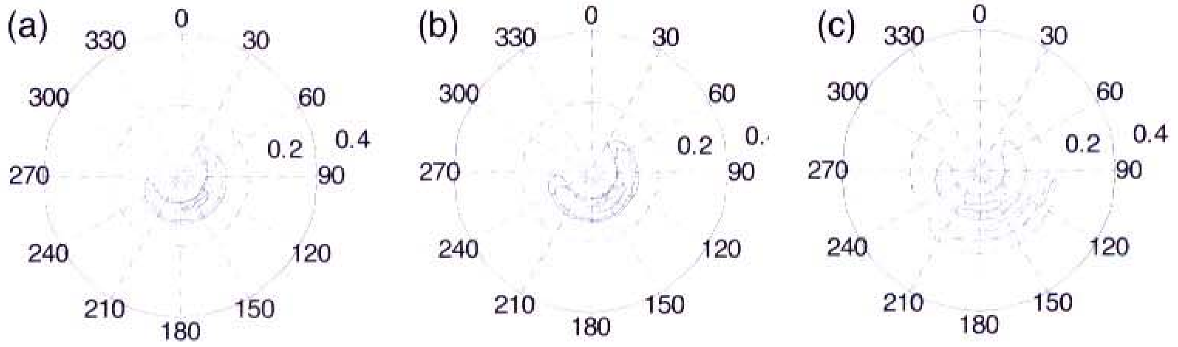

Fig. 2: Direction wave spectra at Hualien buoy for July 26, 2006, 05UT. (a) buoy observation, (b) assimilation run, and (c) reference run.

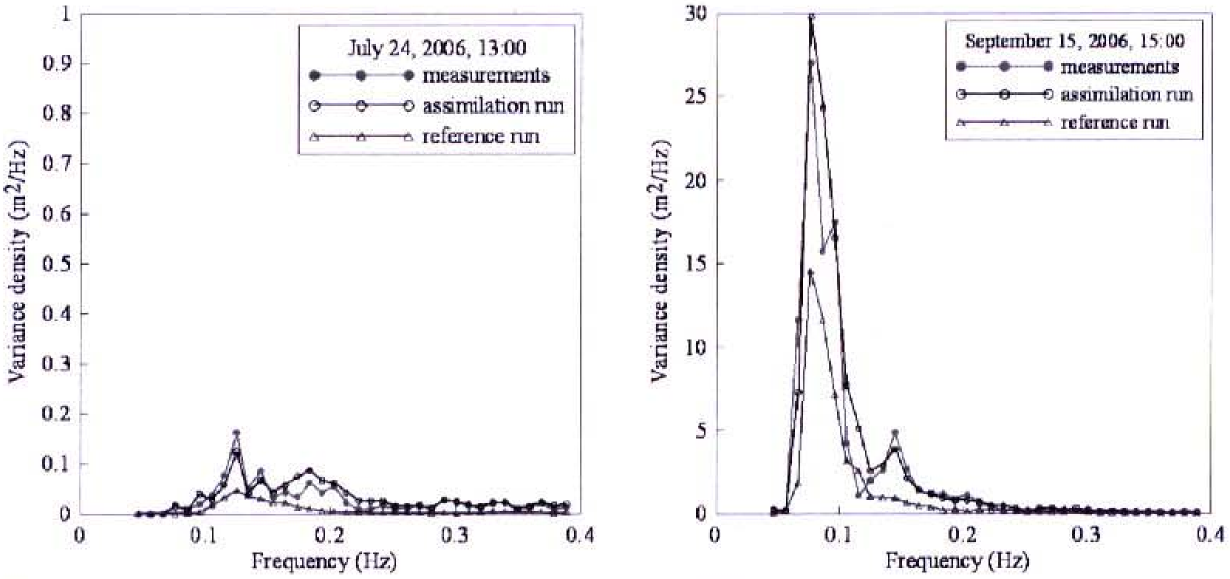

Fig. 3: Spectra at Hualien buoy for July 24, 2006, 05UT, and September 15, 2006, 15UT.

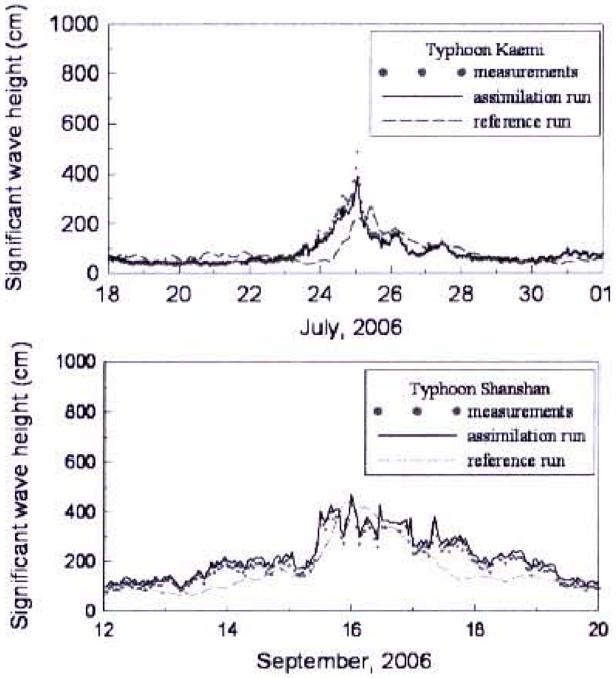

Fig. 4: SWH time series at Hualien buoy (a) Typhoon Kaemi; (b) Typhoon Shanshan.

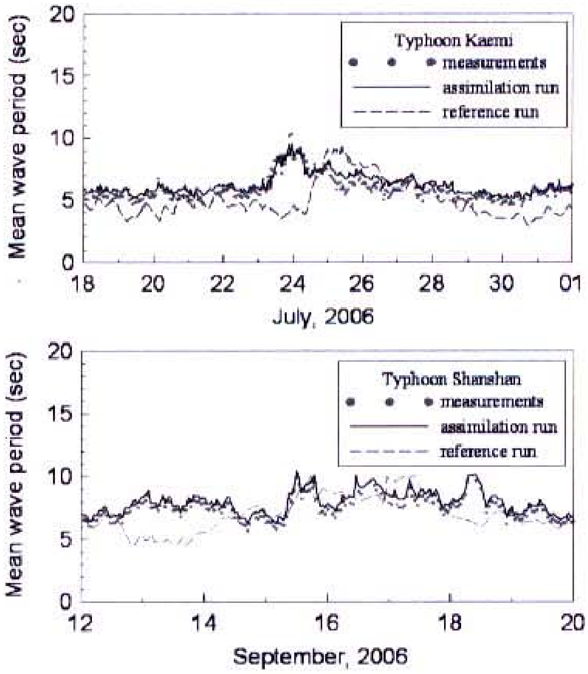

Fig. 5: MWP time series at Hualien buoy. (a) Typhoon Kaemi; (b) Typhoon Shanshan.

Comparison with both buoys shows OI-I scheme is overestimation than OI-P scheme. Differences between the results of OI-P are smaller than those OI-I scheme. And compare with buoy measurements, the OI-P scheme are more close to buoy measurements. This is not surprising, since the calibration of the ENVISAT winds has been obtained by a comparison with measurements from the same buoy. One way to explain the apparent discrepancy is the assumption that the model wind speed error is not homogeneous over the analysis areas in this relatively short period. The altimeter data is not continuous. But the wind speed change very fast during typhoon period. There is not enough data are available to make a reliable validation of the altimeter wind speed algorithms.

Statistical of the different assimilation scheme

The results of the analysis/forecast runs over the typhoon events have been validated against measurements of data buoy. Table 4 compare the OI-I scheme with the OI-P scheme, in terms of the root mean square error (RMSE).

Five parameters are compared: significant wave height (H_s), low-frequency significant wave height (H_{10}), mean wave period (T_m), mean wave direction (θ_w), and wind speed U_{10}. For all the times the ratio of the RMSE of OI-P scheme and the RMSE of OI-I scheme are given, expressing the relative impact of the assimilation on the analysis and forecast.

For different locations and wave parameters, the performance of the OI-I and OI-P is comparable. The main difference is the larger RMSE in wave height and period between OI-P scheme and OI-I scheme. The RMSE of the wind speed compared to observations is not too much difference due to wind speed have corrected before assimilation.

In the period considered, the impact of the assimilation on the forecast is seen up to around 12 hours. This is a shorter period than reported both for OI-I and for OI-P. Probably, the impact is relatively small because in this period, wind sea was in general the dominating wave system. In periods in which swell is of more importance, the impact period is longer.

Both schemes also correct the wind speed during the assimilation. The correction does not lead to a significant improvement or deterioration of the model wind compared to the buoy observations. The impact of ENVISAT assimilation in addition to the assimilation of buoy observations turned out to be negligible in the comparison with the buoy observations. This is to be expected, since especially in the neighbourhood of the buoys, the influence of the buoy measurements is much larger than the ENVISAT observations, which are sparse and often far away from the buoy location.

Conclusions and Outlooks

A spectral wave data assimilation scheme is presented which based on the wave spectrum into separate wave systems and subsequent optimal interpolation of wave partitions. The assimilation experiments at the eastern Taiwan show an improvement the sea state analysis.

In order to optimal the number of frequency and direction, numerical results show that 16 directions and 20 frequency is the optimal choice for data assimilation. In order to do optimal interpolation of wave spectral, we need to assumed stations. When the number of assumed stations is greater than 5 stations, the error tends to stability.

Table 4: RMSE statistics of the various numerical experiments against the data of Gagua Ridge buoy during typhoon Kaemi and typhoon Shanshan

Parameter	OI -P scheme	OI -I scheme	OI -P scheme	OI -I scheme
H_s	0.11 m	0.41 m	0.07 m	0.76 m
H_{10}	0.09 m	0.33 m	0.06 m	0.67 m
T_m	0.4 s	2.31 s	0.4 s	4.8 s
θ_w	7.3 deg	30.2 deg	6.7 deg	45.5 deg
U_{10}	1.92 m/s	3.52 m/s	1.28 m/s	3.96 m/s

Using altimeter data to do data assimilation, the results can improve the boundary values. It means the results can offer better boundary values for nesting using.

Compare OI-P scheme with OI-I scheme; OI-P scheme has better performance in the wave hindcast.

In the near future, the present scheme will be tested further in an operational forecasting setting. Various improvements on the scheme are under investigation. Better treatment of non-assigned partitions may prevent discarding valuable observations. Furthermore, it should be relatively straightforward to merge the OI-I and OI-P approaches in order to handle both spectral and integral observations simultaneously. Finally, research will concentrate on the incorporation of the model dynamics in the assimilation procedure.

Acknowledgement

The authors would like to thank the Central Weather Bureau for providing wind-field data and the observed wave data. The present study was supported by National Science Council grant NSC Project No. is NSC 95-2221-E-211-025.

References

Bauer, E., Hasselmann, K., Young, I.R. and Hasselmann, S. (1996), "Assimilation of wave data into the wave model WAM using an impulse response function method" *Journal of Geophysical Research*, 101, C2, 3801-3816.

De Las Heras, M.M., Burgers, G. and Janssen, P.A.E.M. (1994), "Variational wave data assimilation in a third generation wave model", *Journal of Atmospheric and Oceanic Technology*, 11, 1305-1369.

De Valk, C.F. (1994), "A wind and wave data assimilation scheme based on the adjoint technique", *Dynamics and Modelling of Ocean Waves*, Cambridge University Press, Cambridge, UK, 460-468.

Gerling, T.W. (1992), "Partitioning sequences and arrays of directional wave spectra into component wave systems", *Journal of Atmospheric and Oceanic Technology*, 9, 444-458.

Hasselmann, S., Brüning, C. and Lionello, P. (1994), "Towards a generalized optimal interpolation method for the assimilation of ERS-1 SAR retrieved wave spectra in a wave model", *Proceeding of Second ERS-1 Symposium, Hamburg*, Germany, ESA SP-361, 21-25.

Hasselmann, S., Brüning, C., Hasselmann, K., and Heimbach, P. (1996), "An improved algorithm for the retrieval of ocean wave spectra from synthetic aperture radar image spectra", *Journal of Geophysical Research*, 101, C7, 16615-16629.

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

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

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

Long, R.B. and Hasselmann, K. (1979), "A variational technique for extracting directional spectra from multi-component wave data", *Journal of Physical Oceanography*, 9, 373-381.

Longuet-Higgins, M.S., Cartwright, D.E. and Smith, N.D. (1963), "Observations of the directional spectrum of sea waves using the motions of a floating buoy", *Ocean Wave Spectra*, Prentice-Hall, Englewood Cliffs, New Jersey, 111-136.

Lygre, A. and Krogstad, H.E. (1986), "Maximum entropy estimation of the directional distribution in ocean wave spectra", *Journal of Physical Oceanography*, 16, 2052-2060

Mastenbroek, C., Makin, V.K., Voorrips, A.C. and Komen, G.J. (1994), "Validation 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., Makin, V.K. and Hasselmann, S. (1997), "Assimilation of wave spectra from pitch-and-roll buoys in a North Sea wave model", *Journal of Geophysical Research*, 102, C3, 5829-5849.

Voorrips, A.C. and De Valk, C. (1997), "A comparison of two operational wave assimilation methods", Royal Netherlands Meteorological Institute (KNMI), The Global Atmosphere and Ocean System, 46 pp.