Ocean Modelling 59–60 (2012) 11–23

Contents lists available at SciVerse ScienceDirect

Ocean Modelling

journal homepage: www.elsevier.com/locate/ocemod

Review

Combination of statistical Kalman filters and data assimilation for improving ocean waves analysis and forecasting

George Emmanouil [a,b,*], George Galanis [a,c], George Kallos [a]

[a] University of Athens, School of Physics, Division of Applied Physics, Atmospheric Modeling and Weather Forecasting Group, University Campus, Bldg. PHYS-V, Athens 15784, Greece
[b] Hellenic National Meteorological Service, E. Venizelou 14, Helliniko 16777, Greece
[c] Hellenic Naval Academy, Section of Mathematics, Hatzikyriakion, Piraeus 18539, Greece

ARTICLE INFO

Article history:
Received 19 February 2012
Received in revised form 7 September 2012
Accepted 9 September 2012
Available online 5 October 2012

Keywords:
Data assimilation
Kalman filters
Optimum interpolation
Ocean wave modeling

ABSTRACT

The issue of ocean wave analysis and forecasting is today of increasing importance for a variety of scientific and social-economic purposes. In this framework, a number of state-of-the-art research and operational tools have been developed, mainly based on numerical modeling and advanced statistical techniques. The performance of the latter is essentially dependant on the utilization of external information (remote sensing and in situ measurements). In this work, a combination of ocean wave numerical models, statistical Kalman filters and data assimilation techniques is used for improvement of simulation-accuracy. More precisely, the systematic deviations of the wave model results are minimized by the use of Kalman filtering algorithms in areas with continuous flow of observations. Then, the improved outputs are assimilated by an optimum interpolation scheme, into the forecasting period of the wave model, in order to extend the assimilation impact in time and space. The case studied concerns four one-monthly intervals in the North Atlantic Ocean.

Contents

1. Introduction

The knowledge of sea state conditions is nowadays crucial for a number of different applications, since water surfaces are covering a major part of our planet. Among the issues requiring an accurate description of ocean conditions are climate change, renewable energy, marine pollution, searouting, ship safety and commercial transportation.

* Corresponding author at: University of Athens, School of Physics, Division of Applied Physics, Atmospheric Modeling and Weather Forecasting Group, University Campus, Bldg. PHYS-V, Athens 15784, Greece. Tel.: +30 2106011124; fax: +30 2107276765.

E-mail addresses: gemman@mg.uoa.gr, george.emmanouil@yahoo.gr (G. Emmanouil).

1463-5003/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.ocemod.2012.09.004

Today, wave analysis and forecasting are mainly based on numerical atmospheric and ocean wave models, as well as on data assimilation (DA) systems (see Siddons, 2007; Greenslade and Young, 2004; Breivik and Reistad, 1994; Greenslade, 2001; Voorrips et al., 1997). The latter use sea state information, in order to produce analysis and more accurate initial conditions, therefore resulting in the improvement of the final wave forecasts. The most common technique in ocean-waves DA systems is the optimum interpolation (OI) (Abdalla et al., 2005; Greenslade and Young, 2004; Breivik and Reistad, 1994; Lionello et al., 1992). In recent years, an efficient low-rank approximation to the Kalman filter, presented by Voorrips et al. (1999), has been also considered as a DA methodology.

Despite the progress achieved in this field, a serious disadvantage of the assimilation systems remains their limited temporal and spatial impact to the improvement of the final predictions, especially that of long forecasting horizons (Emmanouil et al., 2007). There are two main issues that have to be considered about ocean waves. The first one is their dependence to atmospheric winds and to the atmospheric model output quality (something not controlled by wave models). The second one is the memory of the system, meaning that any "wrong" information (for any reason e.g. bad quality of wind input) may propagate and cause low quality of wave forecasts at a long distance. The corrections caused by the observations and DA disperse quickly because the (potentially biased) forecasted wind forcing is bringing the model back to its "old' state rapidly. As a result, any discrepancies coming from wind prediction, as well as those related to the wave model evolution, will force the subsequent wave forecasted fields to diverge. Furthermore, the limited number of available and quality-controlled wave observations (compared to the surface covered by oceans and compared to the relevant atmospheric data) contributes to the above-mentioned problem. In this way, only a short-time part of the model forecasting results is improved.

For the improvement of atmospheric model results, Kalman filters – KF – (Kalman, 1960; Kalman and Bucy, 1961; Kalnay, 2002) have been also employed in combination with observations, as post processes for the elimination of the systematic bias in atmospheric and wave modeling in several previous works (Evensen, 2003, 2004, Galanis and Anadranistakis, 2002; Galanis et al., 2006; van der Grij, 2009; Persson, 1990).

The methodology presented by Emmanouil et al. (2010) is a technique to keep the observation information longer in the simulation procedure by applying a bias correction within the forecasting horizon and is based on the incorporation of such filters into the wave model integration, instead of using them as an external post procedure. More precisely, KF are employed in a different way than the one presented by Voorrips et al. (1999) in order to provide improved model forecasts, which are utilized as additional information ("forecasted observations") for DA inside the forecasting period and are spatially propagated by the subsequent use of the OI.

This methodology has been presented in Emmanouil et al. (2010), in a "testing"-preliminary level project and simple cases were studied in order to examine how the system performs when only buoy observations are used.

In this work, we focus on the operational use of the above system in real time applications. The basic developments, as proposed for the KF algorithm are tested under various conditions and by using different inputs. All available observations (satellite and in situ) are exploited by the DA scheme towards a detailed sensitivity study of the system, under complicated weather and wave conditions, and taking into account the local characteristics over a wider region: deep and shallow water cases, wind wave and swell dominated areas, etc. On the other hand, different time periods and weather conditions are studied, providing information on potential seasonal dependence of the system performance.

The paper is organized as follows: a short description of the wave model, the data assimilation scheme, as well as of the Kalman filter algorithms and the necessary modifications for its introduction into the wave model are presented in Section 2. The model configuration and the application of different techniques are discussed in Section 3. In Section 4 may be found the presentation of the results, while the main conclusions are summarized in Section 5.

2. Numerical models and statistical tools

2.1. Ocean wave model and data assimilation scheme

In this study, the WAM model version of ECMWF (European Centre for Medium-Range Weather Forecasts) is the one corresponding to the physics as described in Bidlot et al. (2007). This is a third generation wave model which solves the wave transport equation explicitly without any assumptions on the shape of the wave spectrum. It represents the physics of the wave evolution in accordance with our knowledge today for the full set of degrees of freedom of a 2d wave spectrum. The description of the model is presented in detail in WAMDIG (1988), Komen et al. (1994), Jansen (2000, 2004), and Bidlot et al. (2007). Details of the model configuration used in the present work are described in Section 3.

The data assimilation scheme for the WAM model was developed at ECMWF (Lionello et al. (1992)) and it is based on an Optimal Interpolation method, as outlined in Lorenc (1981). In the first step of the DA procedure, optimum interpolation creates an analyzed field of significant wave heights. From this field, the full two-dimensional wave spectrum is retrieved from a first-guess spectrum, in order to transform the information of a single wave height measurement into different corrections for the wind sea and swell parts of the spectrum. The identification of the sea state as wind sea or swell is performed without further discretization of the spectral components. Therefore, the two-dimensional spectrum is corrected by the introduction of appropriate rescaling factors to the energy and frequency scales of the wind sea and swell. Moreover, the local wind speed forcing is updated. The rescaling factors are derived from duration limited growth relations for the wind sea; for the swell, it is assumed that the wave steepness is conserved. A detailed description of the method can be found in Komen et al. (1994).

2.2. Statistical Kalman filter algorithm

In this paragraph, a short description of the Kalman filter is presented using the unified notation proposed by Ide et al. (1997). Such type of filters simulate the evolution in time of an unknown process (state vector x), whose observational value at time t_i is denoted by $x^t(t_i)$ and it is combined with a corresponding record y_i^O. The change of x in time and the relation between the observation and the unknown vectors are described by the following (observation and system, respectively) equations:

$$x^t(t_i) = M_{i-1}[x^t(t_{i-1})] + \boldsymbol{\eta}(t_{i-1}), \quad y_i^O = H_i[x^t(t_i)] + \boldsymbol{\varepsilon}_i \qquad (1)$$

The system operator M_{i-1}, the observational one H_i as well as the covariance matrices of the Gaussian (non-systematic) errors $\boldsymbol{\eta}(t_i)$ and $\boldsymbol{\varepsilon}_i$ respectively, have to be determined before the application of the filter.

The above described Kalman algorithm has been mostly utilized until now in post processing mode for the elimination of

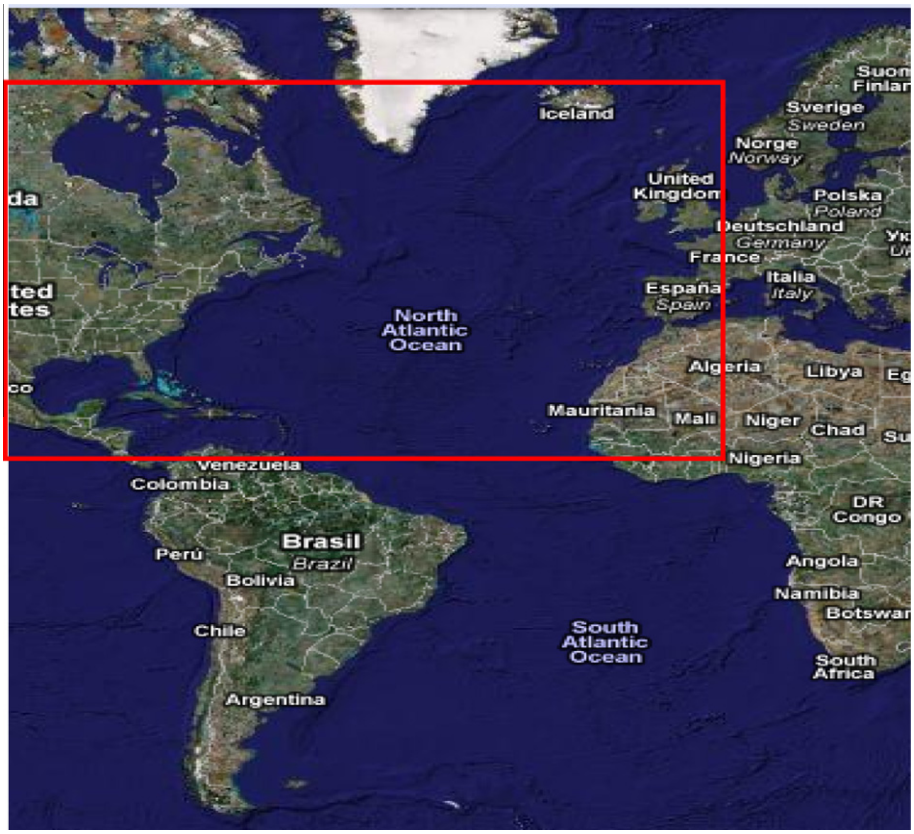

Fig. 2.2.1. The study area (red rectangle-map from Google Earth). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this paper.)

systematic errors (Galanis et al., 2009). In Emmanouil et al. (2010), new developments on the use of KF in combination with the ocean wave DA scheme were proposed: ways of using Kalman filters with the available observational timeseries in real time, different Kalman filters in different areas and optimization of the calculation of the KF covariance matrices Q and R in an operational environment, where Q denotes the covariance matrix of the system equation and R that of the observation equation. In the present paper, this new KF system is tested under real time-operational conditions in the North Atlantic Ocean (Fig. 2.2.1), as part of an integrated ocean wave prediction system. The issue addressed is the exploitation of the corresponding results (KF-"forecasted" values), in combination with an OI assimilation scheme so as to improve wave prediction by maximizing the benefit of both techniques.

The Kalman filter was applied to a single forecasted parameter: the significant wave height (swh). As proposed by Galanis et al. (2009), the filter was adapted such that the corresponding swh bias, now denoted by y_i^O, was estimated as a function of the forecasted direct model output swh_i:

$$y_i^O(v) = a_{0,i}(v) + a_{1,i}(v)swh_i(v) + a_{2,i}(v)swh_i^2(v) + \varepsilon_i(v), \tag{2}$$

where the coefficients $\{\alpha_{0,i}(v)\alpha_{1,i}(v)\alpha_{2,i}(v)\}$ have to be estimated by the filter. Parameter ε_i stands for the Gaussian (non systematic) error of the previous procedure. In this way the state vector of the filter becomes $x(t_i, v) = [\alpha_{0,i}(v)\ \alpha_{1,i}(v)\ \alpha_{2,i}(v)]^{\mathrm{T}}$, the bias $y_i^O(v)$ is used as the known parameter, the observation matrix takes the form $H_i(v) = [1\ \ swh_i(v)\ \ swh_i^2(v)]$ and the system matrix $M_i(v)$ is the three-dimensional identity. The system and observation equations (1) take the following initial values:

$$x = 0, \quad y_0^O((v)) = \varepsilon_0, \quad P(t_0, v) = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix},$$
$$Q(t_0, v) = I_3, \quad R(t_0, v) = 6, \tag{3}$$

where P denotes the error covariance matrix of the state vector x. No correlations between different coordinates of the state vector x (swh) are assumed. The high values for R and the diagonal elements of P indicate low credibility of the first guess and ensure fast adjustment to new conditions. The obtained KF-estimated bias y_i^O is then used for the correction of the forecasted significant wave height.

The covariance matrices of the filter mentioned above are calculated from a seven timestep window (as described in detail in Emmanouil et al. (2010)). The system was trained by taking an observation every 3 h. If any data was missing, the seven previous available observations were used for the necessary calculations of KF.

It is worth noting that the selection of a third order non-linear function in KF (Eq. (2)) in the present study compensates, at least partially, the disadvantage of the application of such linear filters in (non-linear) wave models.

3. Model configuration and applications

The wave model WAM was integrated in this work over a global domain, with horizontal resolution of 0.5° × 0.5°. The main target was to test the proposed methodology in a time-efficient configuration. Analogous studies performed for atmospheric models, showed that the successful use of KF do not really depend on the horizontal resolution of the main model (Galanis et al., 2006). The wave spectrum was discretized with 30 frequencies and 24

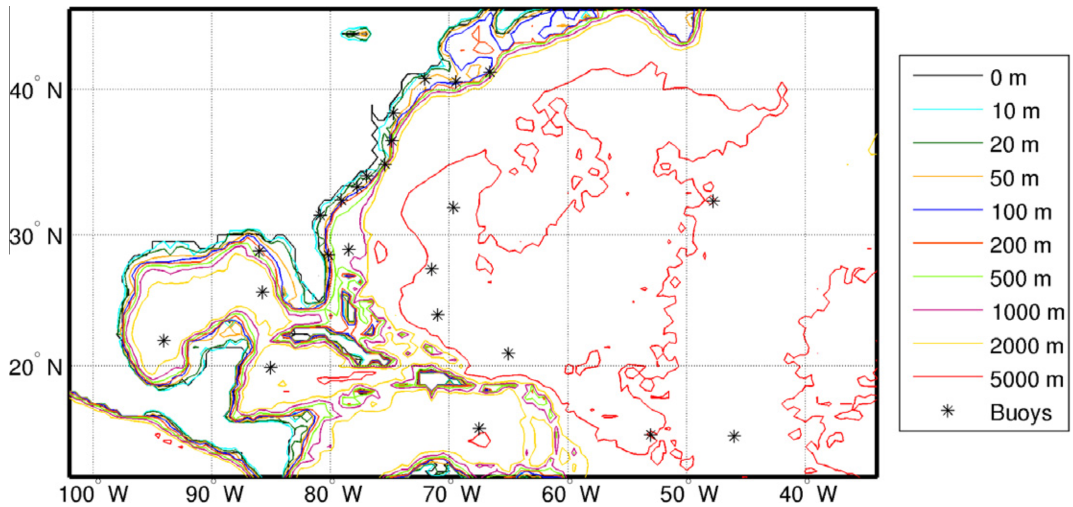

Fig. 3.1. Buoys used (with black stars) and bathymetry (contour lines).

directions. The lowest frequency was defined to 0.0417 Hz, while the propagation and the integration time step were set to 300 s. The timestep of DA was set to 6 h. The model ran in deep water mode with no refraction. It is important to underline here that all the versions were ran with the same parameters. The unresolved bathymetry scheme was switched off while the sea ice mask was taken from the GFS database on a monthly base. The necessary atmospheric input (10 m wind speed and direction forecasts) was obtained from NCEP/GFS global model (horizontal grid resolution $0.5^\circ \times 0.5^\circ$) at a time step of 6 h. It worth to be noted that NCEP/GFS forecasts have been reported in some cases, to overestimate surface winds (e.g. Wave Forecast Verification Project produced for WMO-IOC Joint Technical Commission for Oceanography and Marine Meteorology). On the other hand, NCEP is one of the most reliable environmental operational centers globally, providing free real time weather forecasts.

Even though the area of interest was over the North Atlantic (Fig. 2.2.1), an open sea area, the model was set to cover the whole North and South Atlantic; this choice ensures the right representation of long waves (swell) and high availability of observational data. Moreover, weather and sea conditions in the chosen domain cover a wide range of cases and, by this way, a better evaluation may be performed. Finally, the chosen area lies between two of the most developed continents of the world (Europe and North America), a fact that increases the interest on sea conditions for transport, scientific studies and other applications.

The sources of the observational data used were:

1. The RA2 instrument onboard satellite Envisat (ESA, 2007) from where the data were captured in near real time (with a 3–5 h of delay).
2. The altimeter instrument of the NASA/CNES satellite Jason-1 (Picot et al., 2003) and the data were also captured in near real time.
3. Buoys of the National Data Buoy Center (NDBC) network of NOAA, in real time, which are indicated in Fig. 3.1.

The above data were not quality controlled, nor calibrated (altimeter data) since they were used in real time. Then, they were inserted to the DA scheme as individual along track observations. It has been shown (Durrant et al., 2009; Abdalla et al., 2011) that Jason-1 data are noisier than other altimeter data set such as ENVISAT. Operational centers, like ECMWF, have accounted for that by increasing the observation error associated to Jason-1 (Abdalla et al., 2005). In our experiments, the use of Jason-1 data has not led to elevated observation error.

Eight different experimental versions of WAM were employed, as shown in Table 3.1, which differ in the DA and KF schemes used, as well as in the corresponding wave data input:

1. The first one (referred from now on as WAM0) does not use any assimilation system.
2. In the second version (WAM1), two different observation types were assimilated:
   i. The buoy observations and Envisat RA2 records within the assimilation window.
   ii. Improved-filtered forecasts of WAM, obtained by Kalman filters, which are used as "forecasted values", assimilated inside the forecasting period by the standard OI assimilation system, as shown in Fig. 3.2. In these Kalman-filtered values, an important part of the systematic error has been removed, while their impact

Table 3.1
Experimental versions of WAM. Symbol • means that use of the relevant scheme-data was made.

	Data assimilation	Kalman filters	Calculation of matrices Q and R of the Kalman filters		Input for data assimilation scheme		
			Continuous-dynamic	Semi-continuous	Buoys	Envisat RA2 data	Jason1 data
WAM0							
WAM1	•	•	•		•	•	
WAM2	•	•		•	•	•	
WAM3	•	•	•		•	•	•
WAM4	•	•		•	•	•	•
WAM5	•	•	•		•		•
WAM6	•	•		•	•		•
WAM7	•				•	•	•

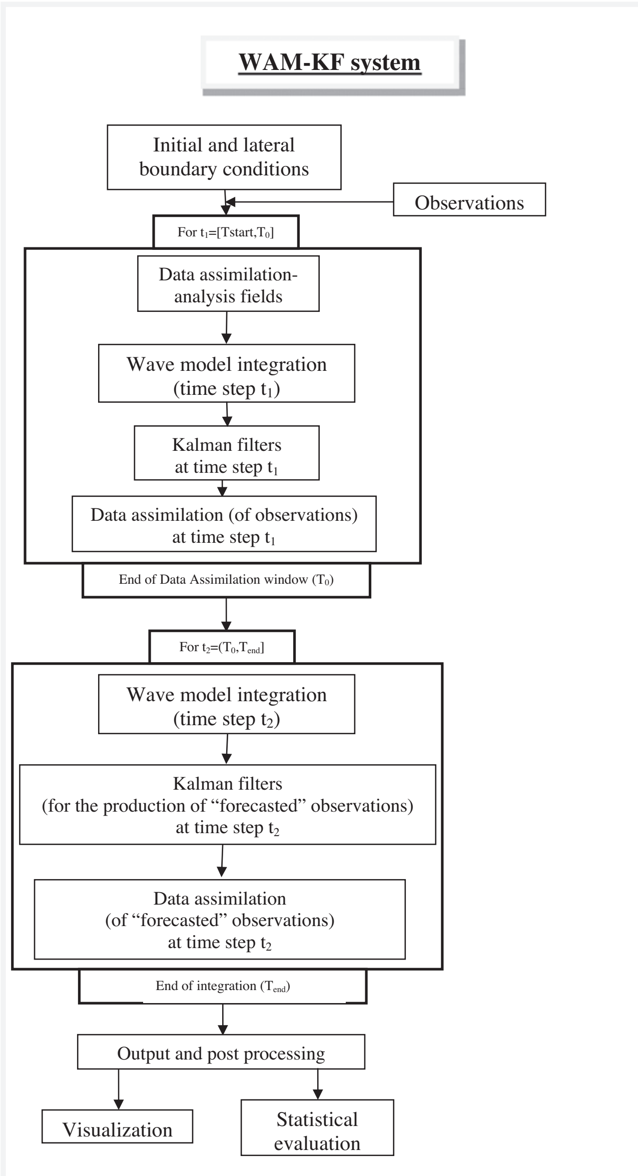

Fig. 3.2. Flowchart of the new wave analysis and forecasting platform.

is extended to a greater area and to the entire forecasting period. However, it should be noted that Kalman filters can be used only in cases with a continuous flow of data, like the buoy data in this study.

The Kalman filter covariance matrices Q and R, of the system and the observation equation respectively are calculated in a continuous way by the last seven values of n and ε at Eq. (1) (which are available at 3-h intervals in the present study). This 7-step period resulted as the optimum one after a series of relevant tests (Galanis et al., 2006). On the other hand, this choice allows fast adjustment to possible changes. The values used for this estimation are either observations (into the assimilation window), or "forecasted values" from the Kalman filter (into the forecasting period). The purpose is to explore the advantages obtained by the continuous-dynamical calculation of the covariance matrices and, therefore, to study the adjustment of the filter to the new conditions appearing in the forecasting period.

3. The third test (WAM2) is similar to WAM1, using again "forecasted values" (produced by KF) as input to the OI – assimilation scheme, though having the Kalman covariance matrices (Q and R) updated in a semi-constant way: only when observations are available (inside the assimilation window). During the forecasting period, mean values of Q and R are used, which is calculated by the last seven observations. This value is changing at each forecasting cycle. In this way, the most recent observations are employed, increasing the accuracy of the filter.The data used by DA in this test are the same with WAM1 (Envisat RA2 and buoy measurements).
4. WAM3 also uses KF-forecasted values as in WAM1. The difference in this case concerns the data used by DA inside the assimilation window, which comes from Envisat RA2, Jason-1 and buoys. The KF covariance matrices are estimated following the "continuous" way.
5. WAM4 is similar to WAM2, but the data used by DA inside the assimilation window comes from Envisat RA2, Jason-1 and buoys, as in WAM3. The KF covariance matrices are estimated following the "semi-continuous" way.
6. WAM5 has the same characteristics with WAM1, with the assimilated data coming only from Jason-1 and buoys. The KF covariance matrices are estimated following the "continuous" way.
7. WAM6 is similar to WAM2, using observations from Jason-1 and buoys, as in WAM5. The KF covariance matrices are estimated following the "semi-continuous" way.
8. Finally, WAM7 employs only the DA system described in Section 2 (Lionello et al., 1992), without KF, serving by this way as a reference platform. This scheme is widely used from operational centers and meteorological services worldwide and is based on an OI technique. In this case, the DA scheme assimilates Envisat RA2, Jason-1 and buoy observations, available before the beginning of the forecasting period (analysis time) in order to study the differences with the proposed system.

The testing period for WAM0-4 consists of four one-month intervals, one in each season (November 2007, February 2008, April 2008 and June 2008). In this way, the performance of the proposed system is evaluated under different atmospheric-synoptic conditions. WAM5 and 6 were executed for June 2008, in order to study the influence of assimilating data from satellites with different characteristics (e.g. repetition time, resolution) in the accuracy of the corresponding sea state forecasts. Finally, WAM7 was executed for November 2007 so to compare the classical DA methodology (OI) with the new system (WAM1-4), and not for a detailed study of an already tested ocean wave DA system (e.g. Abdalla et al., 2005; Breivik and Reistad, 1994).

Each monthly run was initialized (i.e. cold start, no swell) ten days earlier in order to better reproduce all the information from long traveling swell waves which might affect our study area. The above mentioned experiments were performed in real-time conditions. In each forecasting cycle, the first 24 h were used as the DA window for providing the necessary analysis fields and then a 48-h forecasting period was started. The forecasted results were evaluated against buoy and satellite measurements which were not used by DA. The comparison to the observations was done with the nearest model grid point, at the closer time (maximum: 1.5 h of time difference) for each along track observation.

This verification procedure concerns the forecast accuracy and the assimilation impact in time and space. The statistical analysis was based on the following parameters:

1. Bias of forecasted values:

$$Bias = \frac{1}{k} \cdot \sum_{i=1}^{k} (for(i) - obs(i)) \tag{4}$$

Here $obs(i)$ denotes the recorded (observed) value at time i, $for(i)$ the respected forecast and k the size of the sample.

2. Normalized Bias (N.Bias):

$$\text{N.Bias} = \frac{1}{k} \sum_{i=1}^{k} \left| \frac{for(i) - obs(i)}{obs(i)} \right| \tag{5}$$

where || stands for the absolute value, revealing the normalized divergence of the forecasts as a proportion of the observations.

3. Root Mean Square Error (RMSE) and standard deviation (St_Dev) of the error, two classical variation and divergence measures respectively:

$$RMSE = \sqrt{\frac{1}{k} \cdot \sum_{i=1}^{k} (for(i) - obs(i))^2}, St\_Dev = \sqrt{\frac{1}{k} \cdot \sum_{i=1}^{k} ((for(i) - obs(i)) - Bias)^2} \tag{6}$$

4. Results and discussion

As it has been mentioned in previous sections, the main purpose of this study is to propose a new technique for the extension of the ocean-wave assimilation benefits in time and space, addressing, in this way, one of the main problems of the classical DA schemes: the limited influence in the wave forecasts (Emmanouil et al., 2007). In Emmanouil et al. (2010), simple cases were studied to check the system (KF and DA) performance with only buoy observations. Here, we use operationally the new system in real time applications, with all available observations (satellite and in situ) under complicated weather and wave conditions.

A main advantage of the proposed methodology is the correction of the bias between direct model forecasts and buoy measurements even in cases where such discrepancies do not have a constant behavior, since Kalman filters can be dynamically involved in the wave integration (Galanis et al., 2009; Emmanouil et al., 2010).

The study performed concerns the following main issues:

4.1. Comparison with the existing wave simulation systems

As a first step in the evaluation procedure of this new integrated system, we compare the proposed methodology as applied in experiments WAM1-4 with the classical DA scheme (OI) of WAM (WAM7) and WAM without DA (WAM0). The results show that KF combined with classical DA scheme in the forecasting period improves the accuracy of the wave model in all the experimental versions (Table 4.1). More precisely, the systematic bias is significantly decreased thanks to KF at the observation areas. This bias is mainly a result of the deviations of the relevant forecasted atmospheric input, limitations of the physical parameterizations and the numerical scheme used. After this, the DA scheme spreads the improved forecasts to neighboring areas by using OI and, by this way, contributes to the reduction of the deviation of the model results in comparison with the relevant measurements. In contrast, the classical DA scheme (WAM7), although it contributes to the bias reduction, does not reduce the scatter and the deviation indexes (RMSE and standard deviation). Such a behavior may be attributed to the presence of a systematic bias in the model or to the use of Jason satellite data which are associated with the problems mentioned in the previous section.

The improvement of the model results by the use of Kalman filters affected a greater area due to the combination with DA. The wave model forecasted fields were modified in comparison with the standard version of WAM (WAM0), as well as with WAM7, as shown in Figs. 4.1–4.3, where a short testing period, concerning March 2008, is presented (Emmanouil, 2010). Here, we can see that in experiment WAM7 (classical DA with data from two satellites and buoys), after 30 h of forecast, there is very minor difference with WAM0 (no DA). On the other hand, by the use of KF, the ocean waves forecasted fields are still affected. The comparison that follows proves that the new results are more accurate.

The above conclusions are clearer if we look at the absolute differences between the tests performed, as shown in Figs. 4.4–4.8. The differences between WAM0 and WAM7 after 30 h of forecasting, are trivial (Fig. 4.4). In the proposed WAM versions which used data from one satellite and buoys (WAM1 and 2) combined with KF, the differences with WAM0 reach 0.5 m and are spread to the whole study area (Figs. 4.5 and 4.6). The same magnitude of differences hold for the comparison with WAM7. Finally, in the tests, where two satellites were used (WAM3-4), the differences with WAM0 and WAM7 are higher reaching even the level of 1 m (Figs. 4.7 and 4.8).

The problems of positive anomaly near the equator are mainly due to unexpected integration/simulation values near the boundaries of the area of interest at the dates presented, reminding about the "dangers" of boundary conditions. A detailed examination of other simulation dates, in the same area, showed much smaller differences (0.1–0.25 m). Taking into account that long swell travels at about 20 m/s, after 30 h of integration, these anomalies cannot affect the results, in any way, in the study area.

The improvement of the statistics concerns all the proposed versions of WAM. It is worth noting that better results compared with the classical DA scheme (WAM7) have been achieved even with the assimilation of data from only one satellite (WAM1-2). However, the optimum performance came from WAM4, where the calculation of the Kalman covariance matrices is performed only from real measurements and a mean value – changing in each forecasting cycle – is used during the forecasting period (semi-con-

Table 4.1
Comparison with satellite data from Envisat-RA2 (110,000 measurements) and Jason-1 (26,000 measurements) for the autumn period and the first 24 h of forecasting horizon.

Automn (11/2007)	0–24 Forecasting hours	WAM0	WAM1	WAM2	WAM3	WAM4	WAM7
Bias	Envisat-RA2	0.214	0.038	0.047	−0.012	−0.018	0.082
	Jason-1	0.401	0.124	0.126	0.077	0.003	0.163
RMSE	Envisat-RA2	0.623	0.565	0.584	0.562	0.502	0.632
	Jason-1	0.838	0.801	0.791	0.766	0.477	0.831
N.Bias	Envisat-RA2	0.200	0.171	0.165	0.156	0.142	0.190
	Jason-1	0.182	0.160	0.158	0.150	0.092	0.170
Standard deviation	Envisat-RA2	0.585	0.564	0.582	0.562	0.502	0.627
	Jason-1	0.735	0.791	0.781	0.762	0.477	0.815

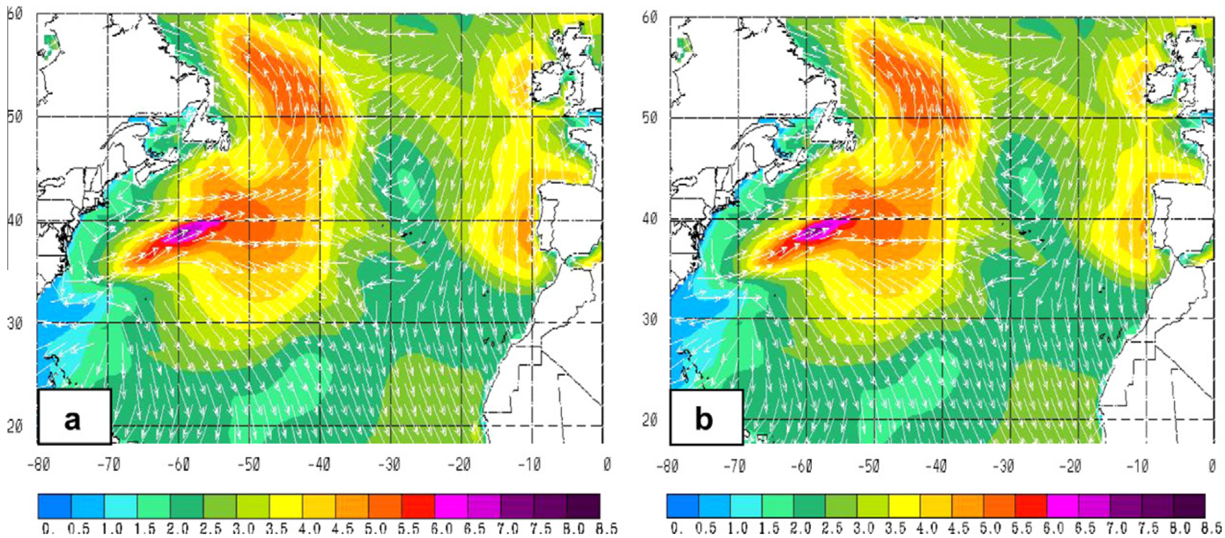

Fig. 4.1. Significant wave height and direction fields at 06UTC of 23/03/2008, from WAM0 (a) and WAM7 (b).

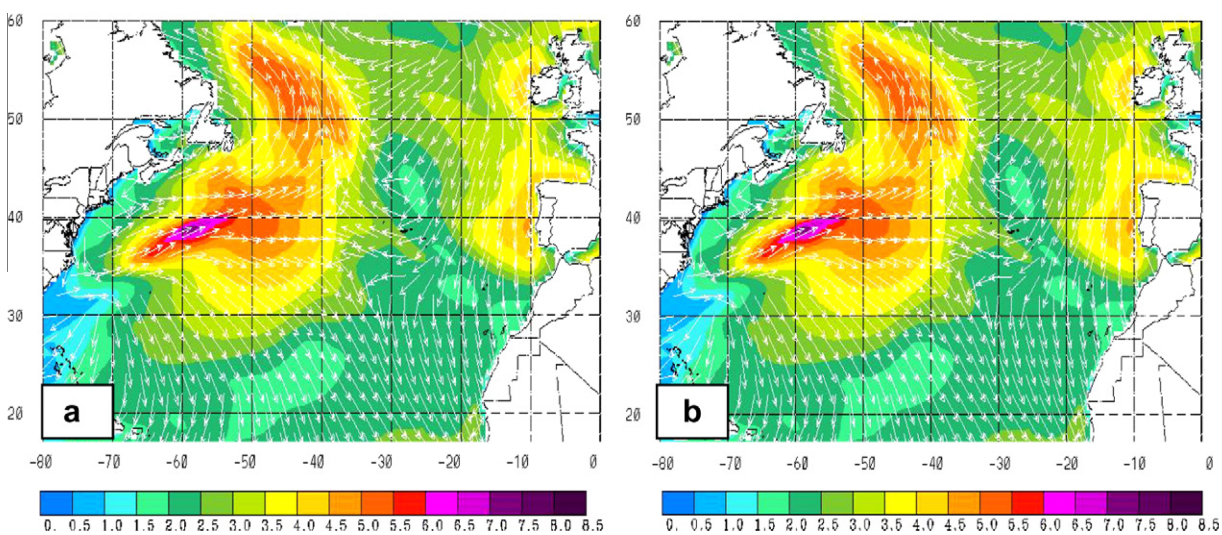

Fig. 4.2. Significant wave height and direction fields at 06UTC of 23/03/2008, from WAM1 (a) and WAM2 (b).

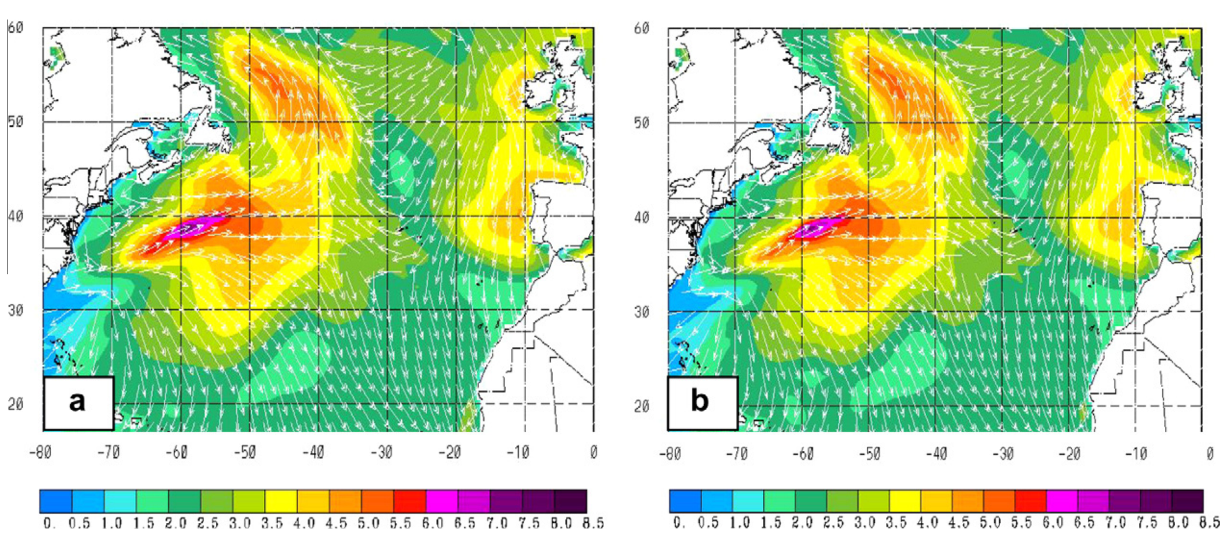

Fig. 4.3. Significant wave height and direction fields at 06UTC of 23/03/2008, from WAM3 (a) and WAM4 (b).

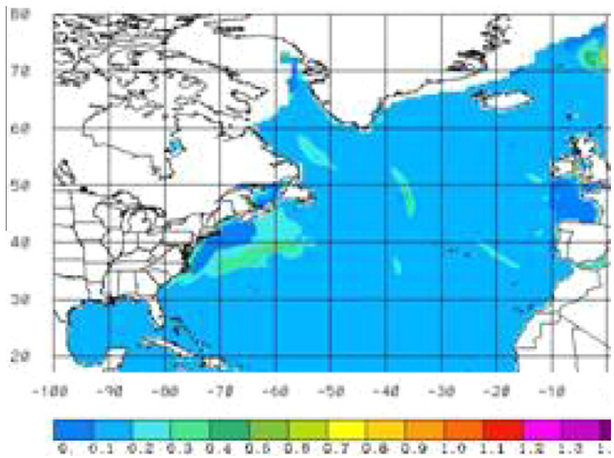

Fig. 4.4. Absolute difference in significant wave height between WAM0 and WAM7 after 30 h of forecast (23/03/2008, 06UTC, from the forecasting cycle of 22/03/2008, 00UTC).

stant way). Moreover, the use of additional observations further improves the performance of the system. In conclusion, going from continuous to semi-continuous update did not seem to result in large improvements from WAM2 with respect to WAM1. Neither did adding Jason-1 when comparing WAM1 to WAM3. But the two changes together did have an important impact! This might mean that there is a threshold in the amount of data needed to remove the systematic error in the model before it pops back up and influences the dynamic of the continuous update of the filter.

4.2. Extention of the impact in time

The proposed methodology extended in time the effects of the DA scheme. This is due to the fact that the use of Kalman filters improved the model results at areas with available measurements by reducing the systematic bias. The subsequent use of DA into the forecasting period extended this "improvement" in the neighborhood where it is expected to have similar wave behavior. The obtained benefits become obvious in Tables 4.2–4.4 where the improvement of the results of the new system is lasting even more than 24 h of forecasting. By this way, an important issue of the ocean waves DA schemes is, at least partially, confronted.

4.3. Seasonal comparison between the proposed methodologies

The better results come from WAM4 (data from two satellites and semi-constant calculation of KF covariance matrices). When

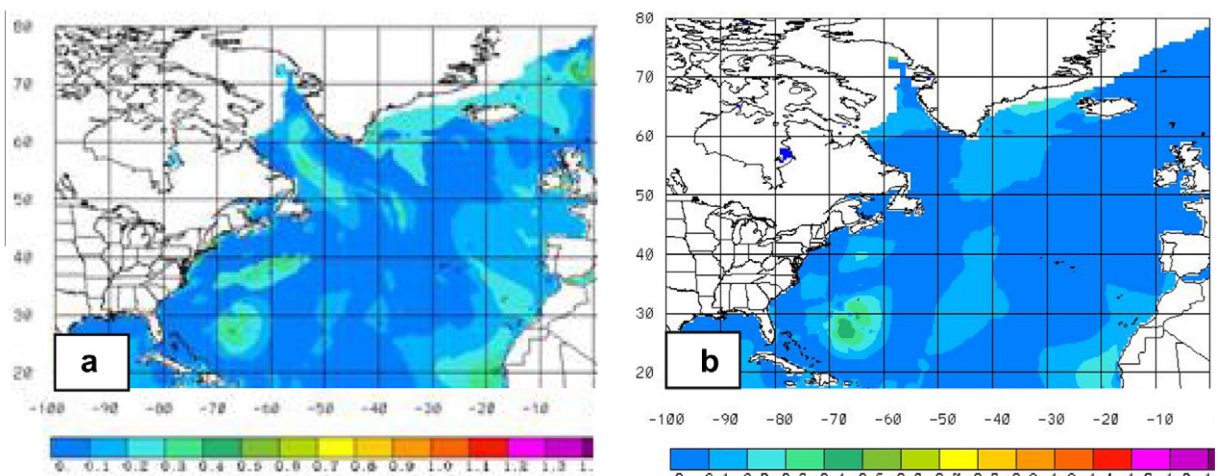

Fig. 4.5. Absolute difference in significant wave height between: (a) WAM0 and WAM1 and (b) WAM7 and WAM1 at the same time with Fig. 4.4.

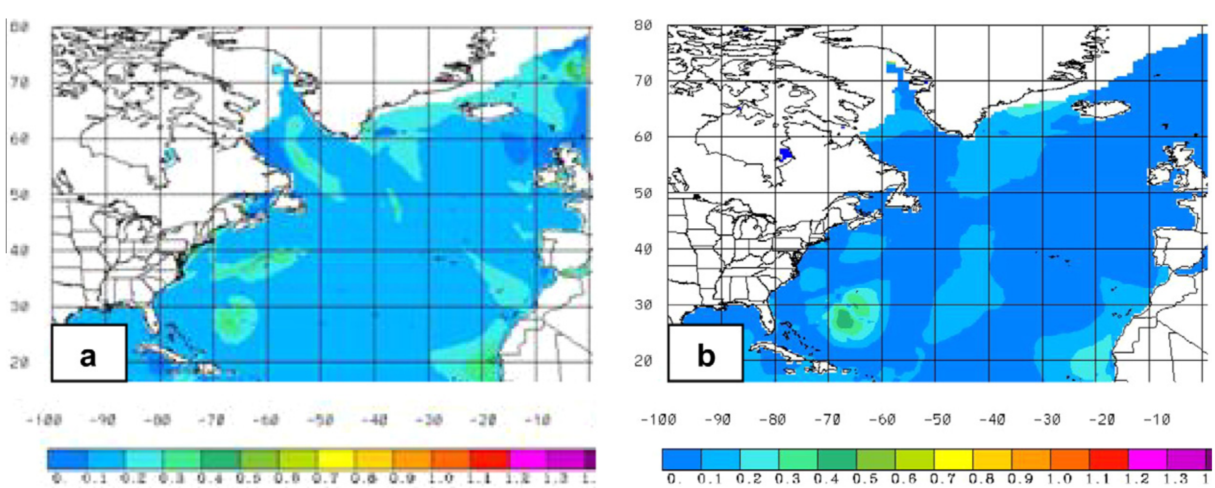

Fig. 4.6. Absolute difference in significant wave height between: (a) WAM0 and WAM2 and (b) WAM7 and WAM2 at the same time with Fig. 4.4.

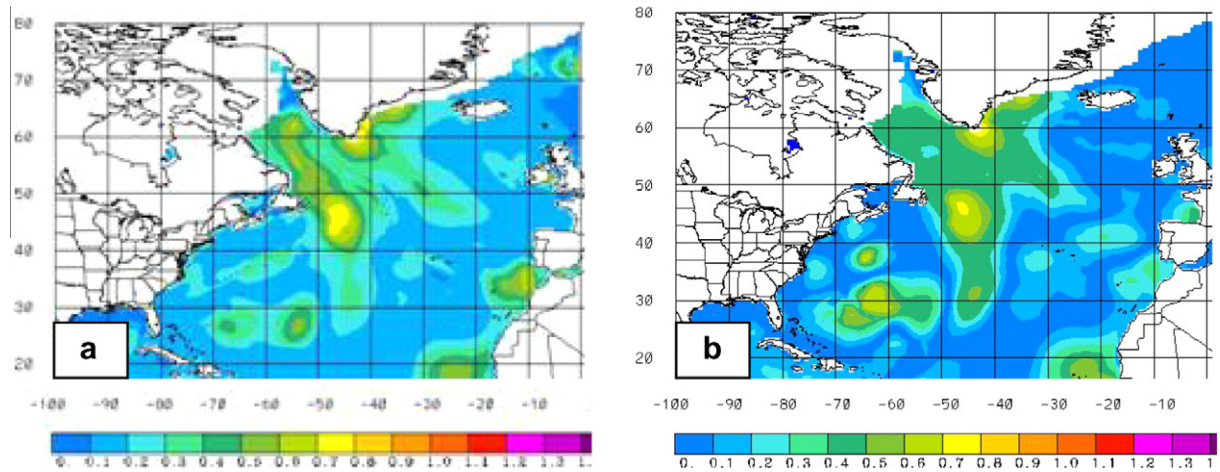

Fig. 4.7. Absolute difference in significant wave height between: (a) WAM0 and WAM3 and (b) WAM7 and WAM3 at the same time with Fig. 4.4.

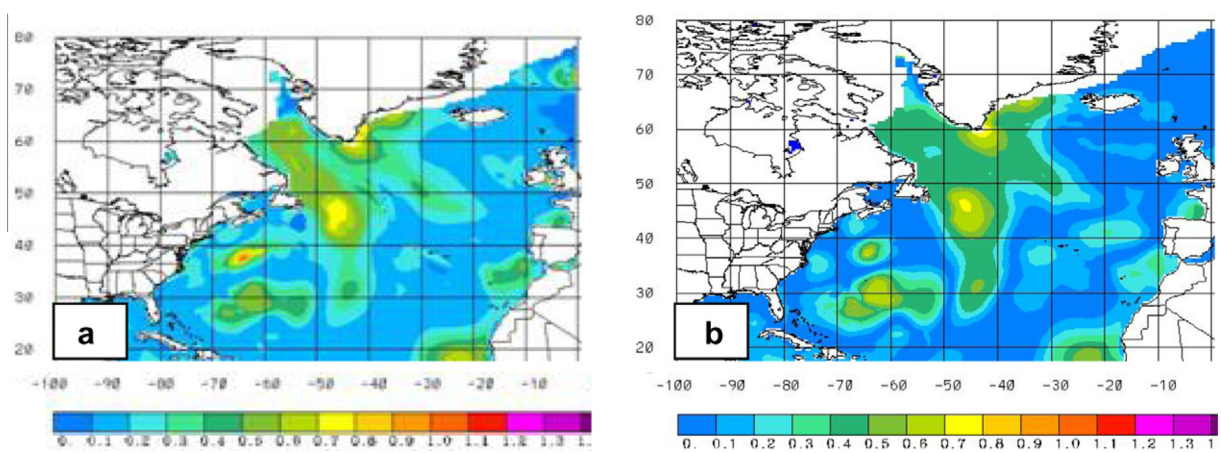

Fig. 4.8. Absolute difference in significant wave height between: (a) WAM0 and WAM4 and (b) WAM7 and WAM4 at the same time with Fig. 4.4.

Table 4.2
Comparison with satellite data from Envisat-RA2 and Jason-1 for the autumn period and the second 24 h of forecasting horizon.

Automn (11/2007)	24–48 Forecasting hours	WAM0	WAM1	WAM2	WAM3	WAM4
Bias	Envisat-RA2	0.219	0.068	0.069	0.029	−0.005
	Jason-1	0.366	0.095	0.095	0.065	−0.0005
RMSE	Envisat-RA2	0.736	0.700	0.700	0.690	0.580
	Jason-1	0.966	0.934	0.933	0.929	0.508
N.Bias	Envisat-RA2	0.217	0.192	0.192	0.185	0.155
	Jason-1	0.199	0.183	0.183	0.180	0.097
Standard deviation	Envisat-RA2	0.703	0.697	0.697	0.690	0.580
	Jason-1	0.872	0.929	0.928	0.926	0.507

using Envisat data as reference, WAM1 and 2 did not lead to improved standard deviation in winter and spring, when in the case of comparing with Jason-1, the standard deviation is improved only in WAM4, with the exception of the summer period. These exceptions from the general positive impact, and especially the increase of standard deviation, are a common fact in the use of DA schemes. It may be attributed to the fact that non-continuous information is incorporated in the homogenous smooth model fields, leading to higher variance values. Moreover, we have to remind that the assimilation scheme relies on reshaping the wave model spectrabased on very simplistic assumptions. When the sea state is complicated, this reshaping is far from optimal and has a tendency to generate undesired noise. The most significant differences between model and measurements were observed during the winter experiment and secondly in autumn, spring and summer. On the other hand, the best improvement in percentage was obtained in winter. In this period, the atmospheric conditions are characterized by the prevalence of well organized low pressure systems which are simulated sufficiently well by the numerical models and the divergences are mainly due to the wave model discrepancies. In these cases, which lead to systematic deviation, the use of KF improves significantly the results.

During the mid periods (spring, autumn), the atmospheric instability is more important, resulting in lower accuracy of the

Table 4.3
Comparison with satellite data from Envisat-RA2 for the winter, spring and summer period (total of 525,000 measurements).

Forecasting hours		WAM0		WAM1		WAM2		WAM3		WAM4	
		0–24h	24–48h	0–24h	24–48h	0–24h	24–48h	0–24h	24–48h	0–24h	24–48h
Bias	Winter (02/2008)	0.262	0.276	0.105	0.108	0.079	0.114	0.027	0.080	0.001	0.009
	Spring (04/2008)	0.189	0.203	0.028	0.067	0.038	0.067	−0.018	0.031	−0.034	−0.022
	Summer (06/2008)	0.175	0.178	0.002	0.009	0.002	0.009	−0.023	−0.010	−0.023	−0.009
RMSE	Winter (02/2008)	0.692	0.789	0.686	0.763	0.626	0.748	0.580	0.735	0.541	0.625
	Spring (04/2008)	0.600	0.706	0.580	0.668	0.578	0.668	0.543	0.657	0.511	0.567
	Summer (06/2008)	0.545	0.616	0.504	0.553	0.504	0.554	0.476	0.548	0.476	0.548
N.Bias	Winter (02/2008)	0.202	0.219	0.176	0.202	0.162	0.188	0.147	0.179	0.142	0.157
	Spring (04/2008)	0.190	0.207	0.187	0.191	0.165	0.191	0.151	0.182	0.145	0.160
	Summer (06/2008)	0.193	0.207	0.171	0.194	0.171	0.194	0.159	0.188	0.159	0.188
Standard deviation	Winter (02/2008)	0.640	0.739	0.677	0.806	0.621	0.739	0.580	0.730	0.541	0.625
	Spring (04/2008)	0.570	0.634	0.599	0.664	0.576	0.664	0.543	0.656	0.509	0.566
	Summer (06/2008)	0.516	0.590	0.504	0.573	0.504	0.573	0.476	0.558	0.476	0.558

Table 4.4
Comparison with satellite data from Jason-1 for the winter, spring and summer period (total of 82,000 measurements).

Forecasting hours		WAM0		WAM1		WAM2		WAM3		WAM4	
		0–24h	24–48h	0–24h	24–48h	0–24h	24–48h	0–24h	24–48h	0–24h	24–48h
Bias	Winter (02/2008)	0.506	0.448	0.217	0.238	0.214	0.238	0.142	0.203	0.028	0.025
	Spring (04/2008)	0.353	0.346	0.154	0.117	0.125	0.117	0.073	0.084	0.003	0.0005
	Summer (06/2008)	0.430	0.426	0.064	0.046	0.063	0.046	0.032	0.031	0.032	0.031
RMSE	Winter (02/2008)	0.971	1.063	0.883	1.041	0.898	1.040	0.842	1.031	0.494	0.531
	Spring (04/2008)	0.833	1.007	0.812	0.945	0.809	0.945	0.774	0.937	0.484	0.520
	Summer (06/2008)	0.859	0.990	0.769	0.940	0.769	0.940	0.738	0.934	0.738	0.934
N.Bias	Winter (02/2008)	0.201	0.208	0.162	0.192	0.163	0.192	0.149	0.188	0.091	0.096
	Spring (04/2008)	0.172	0.187	0.159	0.185	0.158	0.185	0.147	0.180	0.093	0.098
	Summer (06/2008)	0.183	0.203	0.161	0.193	0.161	0.193	0.153	0.190	0.153	0.190
Standard deviation	Winter(02/2008)	0.829	0.878	0.856	1.013	0.872	1.012	0.830	1.011	0.493	0.530
	Spring (04/2008)	0.732	0.838	0.818	0.938	0.799	0.938	0.770	0.933	0.484	0.520
	Summer (06/2008)	0.744	0.894	0.737	0.909	0.737	0.909	0.708	0.884	0.708	0.884

Table 4.5
Comparison with satellite data (Envisat, 290,000 measurements, and Jason, 32000 measurements) for the summer period, for the first and the second 24 h forecasting horizon.

Summer (06/2008)	Forecasting horizon	Comparison to	WAM5	WAM6
Bias (m)	0–24 h	Envisat	−0.026	−0.026
		Jason	0.026	0.026
	24–48 h	Envisat	−0.012	−0.012
		Jason	0.029	0.029
RMSE (m)	0–24 h	Envisat	0.485	0.485
		Jason	0.744	0.744
	24–48 h	Envisat	0.549	0.550
		Jason	0.936	0.936
N.Bias	0–24 h	Envisat	0.163	0.163
		Jason	0.155	0.155
	24–48 h	Envisat	0.189	0.189
		Jason	0.191	0.191
Standard deviation (m)	0–24 h	Envisat	0.484	0.484
		Jason	0.713	0.713
	24–48 h	Envisat	0.559	0.560
		Jason	0.896	0.896

atmospheric model results. The resulting deviations at the wave model results are often non-systematic and the impact of the KF application is restricted but still important. Finally, in the summer time, there are no significant changes in the atmospheric fields and the results of the models are close with each other and to the relevant measurements.

4.4. Satellite comparison

The characteristics of the satellite whose data are used for input in the DA scheme affect the results of the new wave system. The comparison between WAM1-2, where only Envisat data were used, and WAM5-6, where the data come from Jason-1, showed that better results were achieved in the second case (Tables 4.3–4.5). This leads to the conclusion that, in wave simulations at a relatively low model resolution, the most important factor is the number of measurements and the shorter repetition cycle. It should be mentioned that the observations were collocated with the nearest model grid point.

4.5. Evaluation against buoys

The comparison of the proposed system with buoy measurements (for all the testing period covered in each experimental version of WAM) reconfirms the already mentioned conclusions (Fig. 4.9):

- The new system gives better results in all cases when compared with WAM0 and WAM7 (standard DA scheme).
- The optimum results come when we use data from Jason-1. In contrast, when comparing with satellite data, the best results are achieved when we used data from both satellites.
- Between the two ways of calculating the covariance matrices, the semi-constant way leads to more accurate outputs.
- Finally, it is obvious that the new system achieves longer durations of improvement compared with the classical DA scheme.

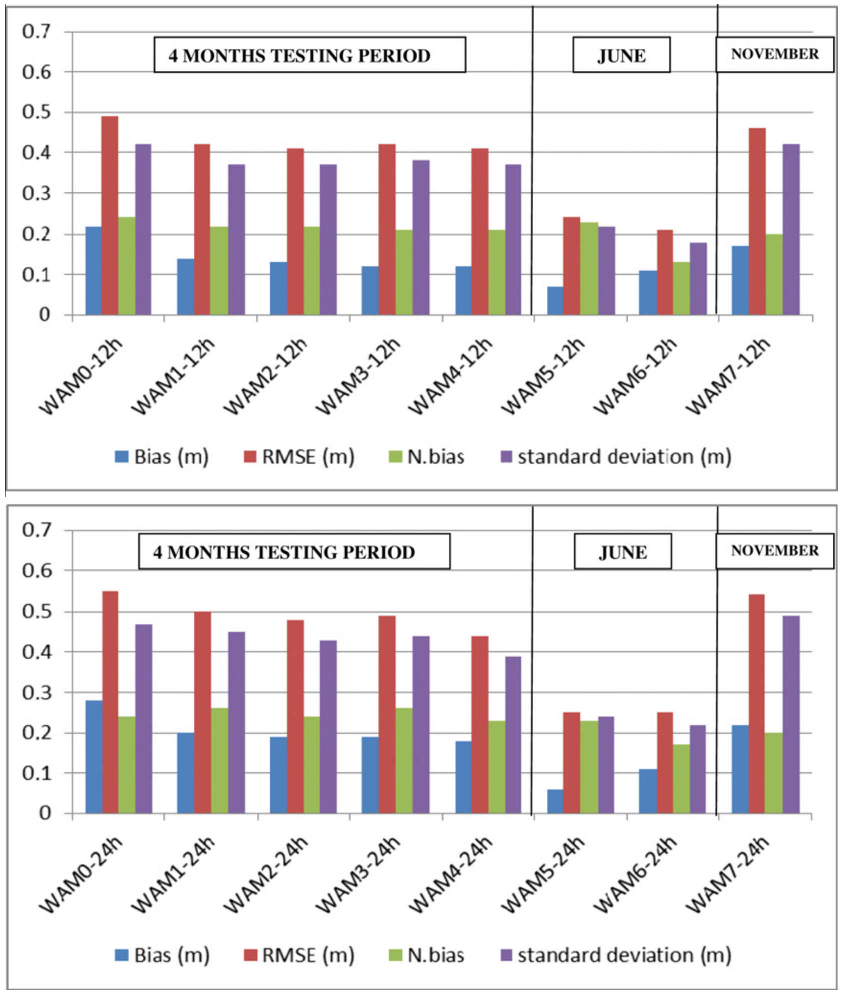

Fig. 4.9a. Comparison [bias (m), RMSE (m), N.Bias, standard deviation (m)] with buoy measurements (240 records for each buoy and each month of comparison) for 12 and 24 h of forecasting horizon for all the testing period available in each case.

We should mention here that the results are for different months on the same figure and they should not be compared. One could compare WAM0-4 which refer to the same period, as well as WAM5-6.

The above comparisons are focusing on swh since this is the parameter that the proposed methodology aims to correct. In general, one of the main problems in wave data assimilation based on wave height observations is to find the optimal way to distribute the update in the spectral domain. Therefore, the utilization of more integrated wave parameters would be enlightening but it is out of the scope of the present work.

5. Conclusions

The results of ocean wave forecasting systems are improved by data assimilation for only a limited time period. This is mainly due to the fact that biases from the atmospheric input (since forecast winds have forecast error due to the intrinsic nature of weather forecasting) or the dynamics of the wave models lead to the reappearance of the initially emerged discrepancies as soon as the external information is no longer available.

In this study, a new technique has been tested in an operational configuration in order to reduce the consequences of this drawback. The proposed approach produces improved model forecasts, obtained by the use of Kalman filters as part of the wave system. Afterwards, these "forecasted observations" are utilized by the DA scheme, inside the forecasting period. This technique leads to the reduction of the systematic error, spreading at the same time this positive impact over a greater area, compared to the one with only observations. On the other hand, the use of the Kalman filter as a part of the wave model guarantees the compatibility of the relevant outputs with the physics of the simulated wave system. The proposed integrated system also improves the forecasting accuracy in time since the obtained impact is extended even until the end of the forecasting period, while by using the classical DA scheme it is limited just to the first 24 h at maximum. Different alternatives for estimating the covariance matrices of the Kalman filter are also explored. The optimum approach proved to be the calculation of a mean value for each forecasting cycle, based on the most recent observations, in each area of Kalman filters application. However, better results than the classical DA were achieved, even when the calculation was made by a continuous-dynamic way, during the assimilation window and the forecasting period.

The statistics in the verification of the proposed system become better when the number of satellites providing data for assimilation is increased. On the other hand, when only one satellite is used, the role of the number of measurements and of the repetition time proves to be more important than the resolution of the supplied data. Finally, the most significant divergences between wave

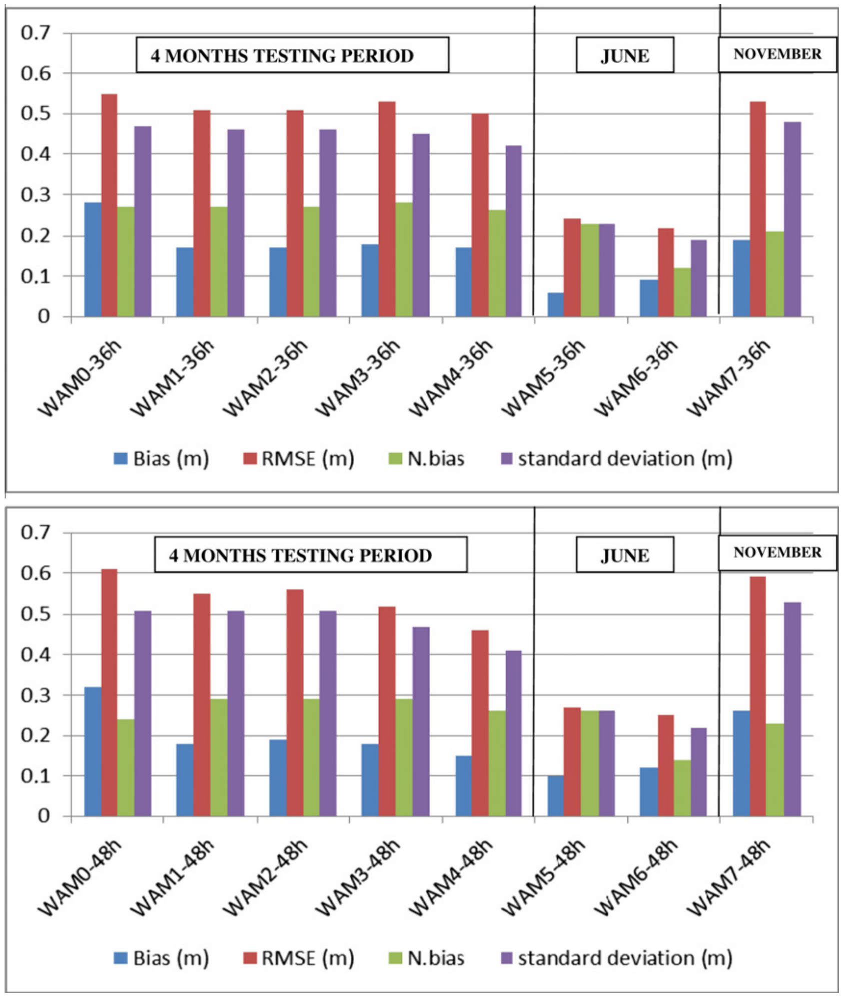

Fig. 4.9b. Comparison [bias (m), RMSE (m), N.Bias, standard deviation (m)] with buoy measurements (240 records for each buoy and each month of comparison) for 36 and 48 h of forecasting horizon for all the testing period available in each case.

model results and observations appear during winter time. However, the use of the proposed system is significantly improving the final results.

It is worth noting that the proposed methodology can be successfully applied only in the presence of continuous time series of observational data (e.g. buoys), which is not always the case, especially in the open ocean. However, such type of data is available for areas of increased interest, like big harbors, touristic coasts, commercial areas, etc. In these cases, the proposed technique can contribute in the improvement of the wave forecasts in a considerable way.

Acknowledgments

The support of the ocean waves team of the ECMWF and especially Dr. Jean Bidlot and Dr. Saleh Abdalla in the proper use of the DA scheme in WAM model is greatly appreciated, as well as the partial support from the FP7 project MARINA PLATFORM (Grant Agreement No. 241402).

References

Abdalla, S., Bidlot, J., Janssen, P., 2005. Jason altimeter wave height verification and assimilation. In: Ozhan, E. (Ed.), Proceedings of the Seventh International Conference on the Mediterranean Coastal Environment (MEDCOAST 05), Kusadasi, Turkey, 25–29 October, pp. 1179–1185.

Abdalla, S., Bidlot, J., Janssen, P., 2005. Assimilation of ERS and ENVISAT wave data at ECMWF. In: Proceedings of the 2004 ENVISAT & ERS Symposium, Salzburg, Austria, 6–10 September 2004 (ESA SP-572, April 2005).

Abdalla, S., Janssen, P., Bidlot, J., 2011. Altimeter near real time wind and wave products: random error estimation. Marine Geodesy 34 (3–4), 393–406.

Bidlot, J.-R., Janssen, P., Abdalla, S., Hersbach, H., 2007. A Revised Formulation of Ocean Wave Dissipation and its Model Impact. ECMWF Tech. Memo. 509. ECMWF, Reading, United Kingdom, 27pp. Available online at: <http://www.ecmwf.int/publications/>.

Breivik, L.A., Reistad, M., 1994. Assimilation of ERS-1 altimeter wave heights in an operational numerical wave model. Weather and Forecasting 9 (3).

Durrant, T., Greenslade, D., Simmonds, I., 2009. Validation of Jason-1 and Envisat remotely sensed wave heights. Journal of Atmospheric and Oceanic Technology 26, 123–134. http://dx.doi.org/10.1175/2008JTECHO598.1.

Emmanouil, G., 2010. New Methodologies for Analyzing and Forecasting Sea Waves with Numerical Models, Data Assimilation and Statistical Filters. Ph.D. Thesis, University of Athens, Department of Physics.

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

Emmanouil, G., Galanis, G., Kallos, G., 2007. Assimilation of radar altimeter data in numerical wave models: an impact study in two different wave climate regions. Annales Geophysicae 25, 1–15.

Evensen, G., 2003. The ensemble Kalman filter: theoretical formulation and practical implementation. Ocean Dynamics 53, 343–367.

Evensen, G., 2004. Sampling strategies and square root analysis schemes for the EnKF. Ocean Dynamics 54, 539–560. http://dx.doi.org/10.1007/s10236-004-0099-2.

ESA, ENVISAT RA2/MWR Product Handbook, 2007. Issue 2.2.

Galanis, G., Anadranistakis, M., 2002. A one dimensional Kalman filter for the correction of near surface temperature forecasts. Meteorological Applications 9, 437–441.

Galanis, G., Louka, P., Katsafados, P., Kallos, G., Pytharoulis, I., 2006. Applications of Kalman filters based on non-linear functions to numerical weather predictions. Annales Geophysicae 24, 2451–2460.

Galanis, G., Emmanouil, G., Chu, P.C., Kallos, G., 2009. A new methodology for the extension of the impact in sea wave assimilation systems. Ocean Dynamics. http://dx.doi.org/10.1007/s10236-009-0191-8.

Greenslade, D., 2001. The assimilation of ERS-2 significant wave height data in the Australian region. Journal of Marine Systems 28, 141–160.

Greenslade, D.J.M., Young, I.R., 2004. Background errors in a global wave model determined from altimeter data. Journal of Geophysical Research 109 (C09007), 1–24.

Ide, K., Courtier, P., Ghil, M., Lorenc, A.C., 1997. Unified notation for data assimilation: operational, sequential and variational. Journal of the Meteorological Society 75, 181–189.

Jansen, P.A.E.M., 2000. ECMWF wave modeling and satellite altimeter wave data. In: Halpern, D. (Ed.), Satellites, Oceanography and Society, pp. 35–36. Elsevier.

Janssen, P., 2004. The Interaction of Ocean Waves and Wind. Cambridge University Press, 300pp.

Kalman, R.E., 1960. A new approach to linear filtering and prediction problems. Transactions of ASME, Series D 82, 35–45.

Kalman, R.E., Bucy, R.S., 1961. New results in linear filtering and prediction problems. Transactions of ASME, Series D 83, 95–108.

Kalnay, E., 2002. Atmospheric Modeling, Data Assimilation and Predictability. Cambridge University Press, p. 341.

Komen, G.J., Cavaleri, L., Donelan, M., Hasselmann, K., Hasselmann, S., Janssen, P.A.E.M., 1994. Dynamics and Modelling of Ocean Waves. Cambridge University Press.

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

Lorenc, A.C., 1981. A global three-dimensional multivariate statistical interpolation scheme. Monthly Weather Review 109, 701–721.

Persson, A., 1990. Kalman filtering a new approach to adaptive statistical interpretation of numerical meteorological forecasts. ECMWF Newsletter.

Picot, N., Case, K., Desai, S., Vincent, P., 2003. AVISO and PODAAC User Handbook, IGDR and GDR Jason Products'', SMM-MU-M5-OP-13184-CN (AVISO), JPL D-21352 (PODAAC).

Siddons, Lee, 2007. Data Assimilation of HF Radar Data into Coastal Wave Models. Ph.D. Thesis of the Department of Applied Mathematics, School of Mathematics and Statistics, The University of Sheffield (U.K.), pp. 260.

van der Grijn, G., 2009. Use of ECMWF products at MeteoGroup. In: ECMWF Forecast Products Users Meeting, 10–12 June 2009.

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

Voorrips, A.C., Heemink, A.W., Komen, G.J., 1999. Wave data assimilation with the Kalman filter. Journal of Marine Systems.

WAMDIG, 1988. The WAM-Development and Implementation Group: Hasselmann, S., Hasselmann, K., Bauer, E., Bertotti, L., Cardone, C.V., Ewing, J.A., Greenwood, J.A., Guillaume, A., Janssen, P.A.E.M., Komen, G.J., Lionello, P., Reistad, M., Zambresky, L. The WAM Model – a third generation ocean wave prediction model. Journal of Physical Oceanography, 18 (12), 1775–1810.