Ocean Waves Measurement and Analysis, Fifth International Symposium WAVES 2005, 3rd-7th, July, 2005. Madrid, Spain
Paper number: 207

ON THE IMPORTANCE OF SPECTRAL WAVE OBSERVATIONS IN THE CONTINUED DEVELOPMENT OF GLOBAL WAVE MODELS

Bidlot J.-R.[1], Janssen P.A.E.M.[1], Abdalla S.[1]

Abstract: It is shown how comparisons between model and buoy spectra have nicely contributed to the recent improvement of the ECMWF wave model. It is argued that due to the present much improved quality of both atmospheric and wave models, detailed global spectral information is required to help steer the direction of further improvement.

INTRODUCTION

The European Centre for Medium range Forecasts (ECMWF) runs state of the art forecast models for the prediction of the evolution of the atmosphere, the oceans and the waves. The wave model products include the usual integrated wave parameters (wave height, period, mean wave direction, directional spread) as well as higher moment integrals connected to the prediction of enhanced probability of extreme waves (freak waves). Wave model spectra are also available for a full description of the sea state.

The quality of the different components of the forecasting system is routinely verified (Lalaurette et al., 2003). The common practice is to use the model analysis as the verifying basis for the different forecast steps. By construct, the atmospheric model analysis is the best estimate of the "truth" since it is the result of an advanced 4-dimensional variational analysis procedure (4dvar) with numerous sources of observations (Klinker et al., 2000). On the other hand, operational global wave model analyses have been limited by the relative small number of wave data sources on a global scale and by the crudeness of the analysis methods employed to combine observations and model estimates. Currently, ECMWF relies on altimeter wave height observations from ENVISAT and SAR wave spectra from ERS-2 to produce its wave model analysis using a simple sequential interpolation scheme. However, altimeters are solely providing information about the wave spectrum to the lowest order possible, namely the zero moment of the frequency spectrum in term of wave heights. The ENVISAT ASAR wave spectra will soon replace ERS-2 SAR data. But the data coverage is still limited and only a relative small portion of the total wave spectrum is actually properly observed. Because of these limitations, there is a limited amount of information available to properly update the wave model spectra. The wave model

1 European Centre for Medium range Weather Forecasts, Shinfield Park, Reading, RG2 9AX,U.K. Jean.Bidlot@ecmwf.int

analysis is therefore not "optimal". The use of the analysis as the basis of the truth for the validation of the different forecasts can be cast into doubt. There is therefore a need to validate the analysis itself.

For years, buoy data have been used to validate wave model hindcasts. These data are usually not used by the model data assimilation and hence constitute an ideal independent data set (Bidlot et al., 2002). This procedure is routinely used at ECMWF to assess the quality of the wave model (Janssen, 2004). Until recently however, only wave heights and possibly some measures of the wave periods were used in this routine exercise. Because the overall quality of the wind forcing has steadily improved and because of recent progress in wave modelling, a more detailed look at the wave model quality is needed. Global wave models are formulated in term of wave spectra. Therefore the actual spectra should be validated. For some years now, buoy wave spectra have been freely available for buoys along the North American Coasts. Extending a tool that was introduced by Voorrips et al. (2001) in a study to validate SAR and wave model spectra against buoy spectra, ECMWF now routinely uses 1-D spectra from those buoys to gain insight on the quality of its wave model. This has led to several developments some of which with marked positive impact on the system. A few examples are presented here.

ECMWF wave model

The ECMWF wave model (ECWAM) has evolved from the original WAM Cycle4. A summary of the key modifications is described in Janssen (2004). In operations, it runs coupled to the different forecasting configurations of the atmospheric model. It is also available as a stand-alone model. The results presented here were obtained using the current deterministic forecast resolution, namely a grid spacing of the order of 55 km, 24 directions and 30 frequencies (in use since November 21, 2000). Altimeter wave height data and SAR wave spectra are assimilated unless specified otherwise.

Buoy spectral data

We have processed buoy frequency spectra obtained from the National Oceanographic Data Center (NODC[2]) and from the Canadian Marine Environment Data Service (MEDS[3]). These buoys are deployed some distance offshore on both the Atlantic and Pacific sides of the North American continent, including Hawaii and Alaska (a list of all buoys used in this paper can be found in the appendix). Following a basic quality check that rejects unrealistic values, hourly buoy spectra were averaged in 4-hourly time windows around the main synoptic times for which model spectra are also produced. This produces more stable estimates of the spectral shape, more inline with the scale of the model spectra. Simple integration will yield traditional quantities such as significant wave height and mean wave periods. However, even more insight can be gained if model and observations are compared in terms of frequency bands. The information contained in the 1-D spectra is smoothed by integrating over frequency intervals corresponding to three consecutive wave model frequency bins and by converting the average energy density to 'equivalent' wave heights. This integration window is run across all frequency bins. The binned equivalent wave heights for the model and the buoys are compared for different frequencies or wave periods.

2 http://www.nodc.noaa.gov/

3 http://www.meds-sdmm.dfo-mpo.gc.ca/meds/Databases/WAVE/WAVE_e.htm

The resulting period-dependent bias, is plotted as a function of time in Fig. 1 for the comparison between model analysis and all selected US and Canadian buoys over a period from December 2000 to April 2005.

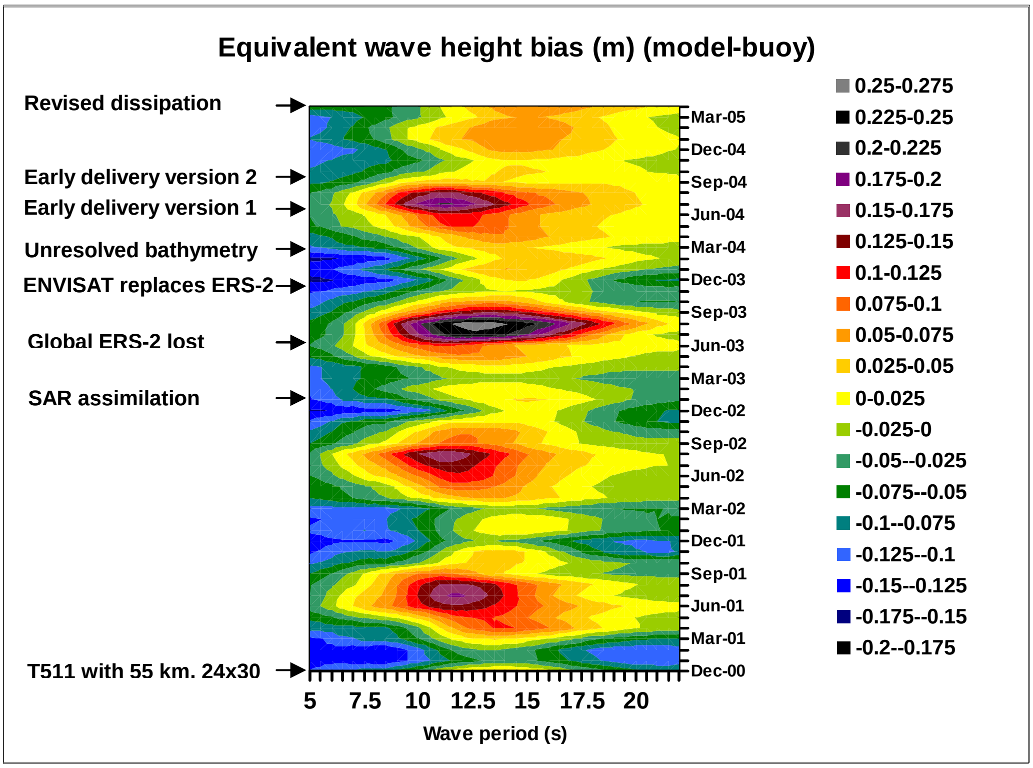


Figure 1: Period spectral bias (model-buoy) at all US and Canadian buoy locations. The 1-D spectra were smoothed out by averaging over 3 consecutive frequency bins and by converting each average value into an equivalent wave height (m). The operational model is used.

In the range of 10-15 seconds there is a clear seasonal dependence of the equivalent wave height bias, it being large in the summer time and much reduced to overall negative in the winter. Also visible are the changes in wave data usage over that period as well as some of the changes to the model or the forecasting system as will be described below.

Data assimilation validation

The operational wave model analysis at ECMWF has used altimeter wave heights since August 1993, first from ERS-1 and then from ERS-2. Currently, other satellites (ENVISAT and Jason) also provide near real time altimeter wave heights, resulting in a better coverage of the world oceans. ENVISAT altimeter data replaced ERS-2 data in late October 2003. Jason data have not yet been assimilated. Unfortunately, the altimeter only yields wave heights and wind speeds over a narrow swath. A more accurate description of the sea state requires the full two-dimensional wave energy spectrum. Such observations, albeit neither fully comprehensive nor independent, are available from the ERS synthetic aperture radar (SAR) and from the advanced SAR (ASAR) on board ENVISAT. Since January 2003, the operational models have also assimilated ERS-2 SAR spectra.

Spectra as derived from the ERS-2 SAR wave mode imagette spectra are processed operationally to retrieve ocean wave spectra using an inversion scheme based on the work done by Hasselmann et al. (1996). The inversion scheme relies on a model first guess to resolve the directional ambiguity, to provide first guess information on the low frequency part of the spectrum and to fill the gap at the high frequency part of the wave spectrum. Note however, that due to the motion of the scattering elements induced by long waves, the SAR only images part of the total wave spectrum. Waves with wavelength shorter than an observation dependent cut-off wavelength are not detected or are heavily distorted. A method based on the assimilation of wave systems as derived from a spectral partitioning scheme, which works on the principle of the inverted catchments area, is used (Hasselmann et al., 1997). The different wave systems are characterised by means of their mean energy, frequency and direction. The mean parameters are assimilated using an optimal interpolation scheme following a cross assignment procedure that correlates the observed and modelled wave systems. The analysed spectra are reconstructed by resizing and reshaping the model spectra based on the mean parameters obtained from the OI scheme.

Before the operational implementation of the SAR data assimilation, it was essential to investigate the spectral distribution of the wave energy by comparing the different runs with buoy spectra and other data sources (Abdalla et al., 2003). As an example, the impact of using satellite data on the wave energy distribution is displayed in Fig. 2. From these graphs, it is clear that altimeter data assimilation results in the largest decrease in errors at all periods, nonetheless, as expected, SAR data assimilation has an added positive impact on longer waves. This is particularly the case for May 2001, but less so for December 2000. Following encouraging results such as those displayed in Fig. 2, SAR data assimilation became operational on January 14, 2003.

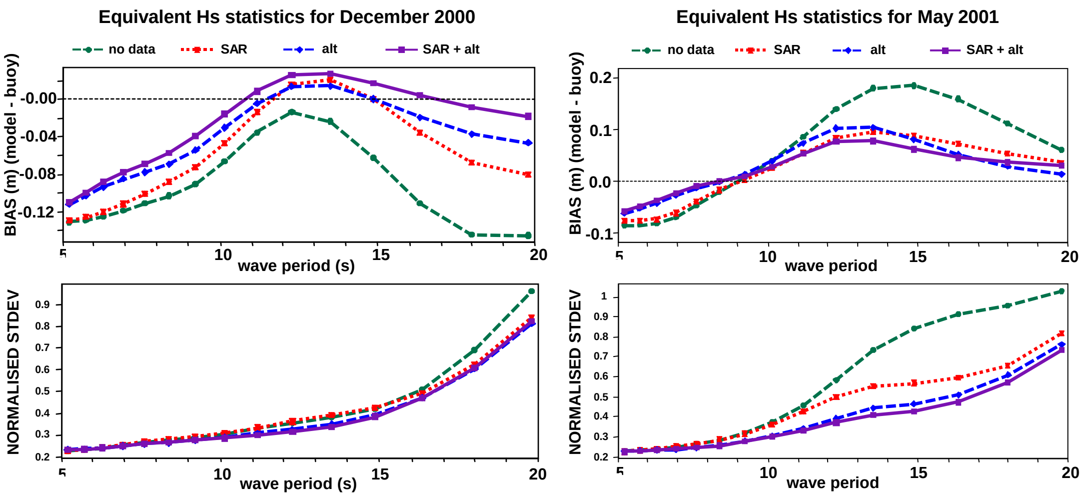


Figure 2: Comparison between stand-alone WAM hindcasts and 1-D wave spectra from US and Canadian buoys to access the impact of wave data assimilation from altimeter and SAR. Equivalent wave heights are compared (see text). The statistics are computed for each WAM discretised frequency and presented in term of equivalent period. The normalised STDEV is defined as the standard deviation of the difference normalised by standard deviation of the observations. Runs without any data assimilation (no data) are compared to runs with SAR data (SAR), with altimeter wave height data (alt) and with both (SAR+alt).

Change in wave data coverage

With the introduction of SAR data, the overestimation in the 10-15 second range was expected to be reduced during the summer period (see Fig. 1). Unfortunately, towards the end of June 2003, the ERS-2 global coverage was lost due to the failure of both tape recorders on-board the spacecraft. Since then, the European Space Agency has only been able to provide data for areas where the ERS-2 satellite is in view of a ground station, namely the North Atlantic and the west coast of North America.

This failure had devastating consequences for the wave model analysis because ERS-2 was then the only data source for the wave model analysis (altimeter and SAR data). For example, Fig. 1 indicates that the spectral bias for the 2003 summer was the worst for years. We have recently identified two of the reasons for the model overestimation of swell energy for spectral components with periods around 12 seconds as will be described below. Nevertheless, as was illustrated in Fig. 2, assimilating ERS-2 data had already a positive impact in reducing this overestimation.

Fortunately, ESA had successfully launched ENVISAT in February 2002. Following an extensive monitoring of the data, it was found that the Ku-band altimeter data are of very good quality apart from a small overestimation of the order of 4% (Janssen et al., 2003). The assimilation of these data became operational on October 22, 2003 (the remaining ERS-2 altimeter data were removed from the analysis). In preparation to the introduction of ENVISAT wave height data, a comparison with buoy data was made for experiments with and without ENVISAT data. As shown in Fig.3, buoy spectra are very useful in demonstrating the beneficial impact of ENVISAT assimilation. We can relate the reduction in bias in Fig. 3 for a possible similar impact in Fig. 1 if ENVISAT had been used.

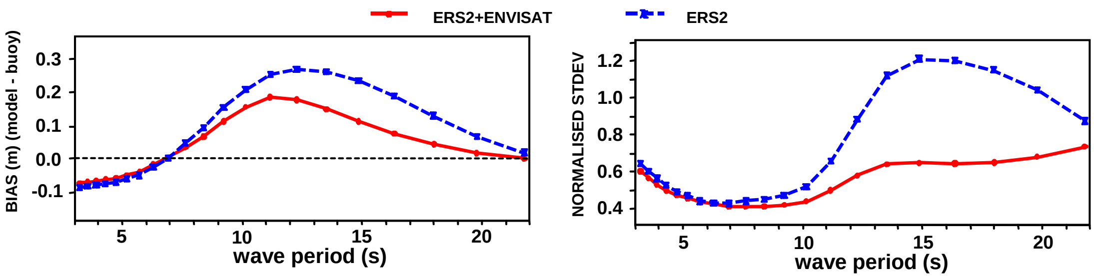


Figure 3: Comparison between stand-alone WAM hindcasts and 1-D wave spectra as in Fig. 2 for the period from July 23 to August 19, 2003 to assess the impact of the use of ENVISAT altimeter data for assimilation. The experiment in which ENVISAT altimeter data have been used (red solid curves) is compared to a reference (blue dash lines). Note that ERS-2 had a reduced coverage during that period (see text).

Unresolved bathymetry

Further analysis of the comparison between model and buoy spectra revealed that these large positive biases are partly related to swell events generated by storms in the southern hemisphere. An obvious candidate is the dissipation source function, because this source term is the least understood. However, a closer inspection of the results indicated that the main problem occurs in the Pacific Ocean and not in the Atlantic (not shown). Increasing the dissipation would therefore have a detrimental impact on the results in the Atlantic. As it turned out a revision of the model dissipation also yielded substantial gain (see below). Nevertheless, we also realised that an important difference between the Pacific and the Atlantic Oceans is the presence in the equatorial region of the Pacific of a vast number of small islands and atolls that were not resolved by the model at the time. Although these islands are small, they nonetheless block considerable amounts of low-frequency wave energy. Consequently, using a high resolution bathymetry data set[4], we determined a wavenumber dependent attenuation factor to the wave propagation based on similar ideas as those presented by Hardy et al. (2001) and Tolman (2003). Using only wave height information would have limited the analysis of this problem. With observed wave spectra, an even more detailed investigation of how

4: The 2 minute ETOP02 data set from the US National Geophysical data center http://www.ngdc.noaa.gov/mgg/global/global.html the new model affects the distribution of wave energy in term of frequency was obtained as displayed in Fig.4. As expected, the impact of the inclusion of unresolved bathymetry in the wave model advection is the largest for the runs without any data assimilation. Using ERS-2 altimeter data has a similar effect on the wave energy distribution around 10-12 seconds as was obtained when sub-grid effects were included. However, there is a small degradation for periods larger than 16 seconds (loss of correlation and increase in standard deviation). The assimilation of altimeter data seems the have a small detrimental effect on the low frequency wave energy distribution (a proper solution for this problem is still under investigation).

The treatment of the unresolved bathymetry was implemented on March 9, 2004 in operations (Bidlot and Janssen, 2003).

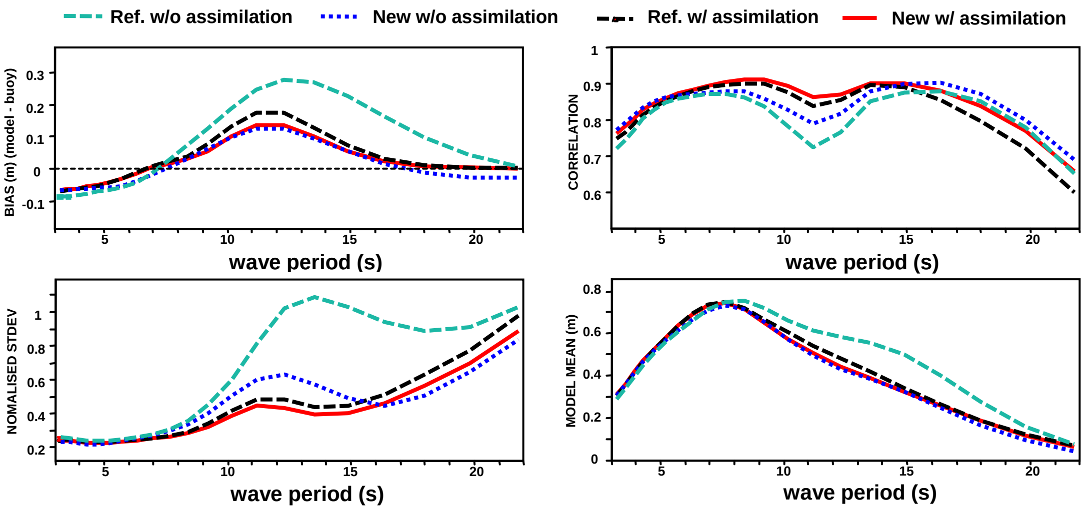


Figure 4: Comparison between 1-D buoy spectra and model hindcasts as in Fig. 2 for July 2001 to access the impact of unresolved bathymetry. Reference runs (Ref.) were obtained with the operational mean bathymetry of the time whereas the new runs (new) employed a mean bathymetry and attenuation coefficients for the treatment of unresolved bathymetry derived from the ETOP02 data set[4]. ERS-2 altimeter data were used for the runs with assimilation.

Change in the operational forecasting system

In order to improve on a timely delivery of ECMWF's forecast products, the early delivery suite was introduced on June 29, 2004. This operational suite as its name implies delivers products earlier than before by producing forecasts from a shorter 6-hour analysis that has not benefited from all observations that would be available it had waited a few extra hours. However, it was found that the impact of the late arrival data could be restored if another 12-hour analysis was performed later with all the data. The short term forecast from this delayed analysis is used to initialise the next analyses. In the context of the 12-hour 4dvar used by the atmospheric model, this 12-hour cycling was originally done with a 9-hour forecast since by nature of the 4dvar method a shorter 3-hour forecast from the analysis 6 hours later is statistically similar. However, the wave model data assimilation is still based on an instantaneous OI analysis, cycled by 6-hour forecasts. As a consequence, when cycling the coupled atmosphere-wave system by means of a 9-hour forecast from 0 and 12 UTC, the impact from 6 and 18 UTC wave analyses gets lost. Basically, it is as if the data coverage for the wave data was reduced by half. We have seen that reducing the data coverage could have a negative impact on the quality of the wave analysis (Fig.3).

The global data coverage was restored by shifting the cycling in 4dvar to a 3-hour forecast. Using spectral buoy data, it was indeed demonstrated that the impact of the 6 and 18 UTC analyses was restored (not shown). This new configuration of the forecasting system became operational on September 28, 2004.

New model around a revised whitecapping dissipation

The wave model dissipation source function was recently reformulated in terms of a mean wave steepness parameter and a mean frequency that gives more emphasis on the high-frequency part of the spectrum. The resulting model wave growth is therefore less sensitive to the presence of low frequency swell. Moreover, with this revised parameterisation, it was also possible to relax the prognostic frequency range over which the model equations are integrated. Prior to this change, ECWAM was only integrated up to a maximum discretised frequency, proportional to the mean frequency of the total sea. For frequencies above that threshold, a diagnostic f^{-5} spectral shape is appended, replacing any wave systems that might otherwise be there. In the presence of low-frequency swell, under light wind conditions, it is likely that the f^{-5} tail would prevent or delay the growth of newly generated windsea. Using buoy 1-D spectra, especially for those locations near the coast but still exposed to ocean swells, it was evident that such shortcoming existed in the model. Although, this is a relatively minor problem in terms of global wave height statistics, it adversely affects verification statistics for the mean period. American and Canadian buoys only report peak period on the GTS. However, with 1-D spectra from those buoys, it was possible to show that in terms of mean periods (T_z) the model performance was not as good as it would be thought if wave heights (H_s) are used for verification (Table 1). This was particularly true for the Pacific buoys during the summer when local weather conditions are quite gentle but swell from afar is still present. A few minor adjustments were also necessary to take advantage of the increased dynamic range of the model. The new model was extensively tested (Bidlot et al., 2005) as an example Fig. 5 shows the comparison with buoy spectra in terms of equivalent wave heights. The old formulation underestimates wave energy at high frequency (roughly speaking the windsea part of the spectrum) during the northern hemisphere summer. For winter months, the underestimation is also present at all frequencies. For summer months, the model largely overestimates swell around 12 second period. The new formulation does improve the prediction of windsea and also reduces the swell overestimation of the summer months at around 12 seconds. Note however, that it does not fit to the data for very long period swell (above 18 sec.). Further analysis indicates that there are still cases when low frequency swell is present in the model but not in the data. Moreover, by comparing the arrival time of those swell events between model and a given buoy indicates that the model swell usually arrives sooner and lasts longer than observed. This is indicative of too much diffusion in the model. The advection scheme in WAM is known to be quite diffusive. More work is needed to test other suitable advective schemes to see if they can be used to improve the prediction of low frequency swell.

Table 1. Impact of the New Model Dissipation Function on Wave Height and Mean Period Statistics (see Fig. 5)

R.M.S.E	Winter old	Winter new	Summer old	Summer new
H_s all buoys (m)	0.64	0.58	0.27	0.26
T_z all buoys (s)	0.88	0.68	0.69	0.62
H_s US West Coast (m)	0.75	0.70	0.34	0.28
T_z US West Coast (s)	1.14	0.85	1.30	0.72

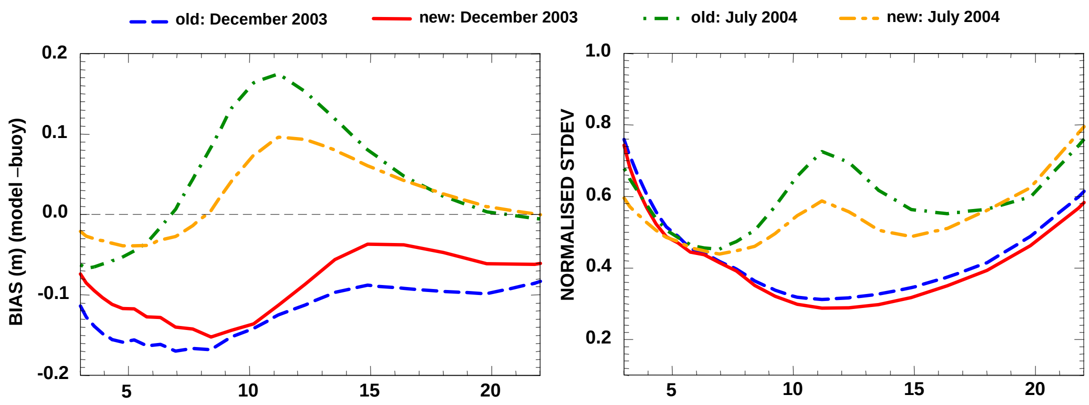


Figure 5: Comparison between 1-D frequency spectra from US and Canadian buoys and model hindcasts as in Fig. 2 for December 2003 and July 2004 to access the impact of the new model version. The model with the revised whitecapping dissipation and extended prognostic frequency range (new) is compared to the model with the old formulation (old). All runs used the stand-alone 55 km model without any data assimilation.

Following a long pre-operational testing, this new formulation was implemented into the operational system (along with changes to the atmospheric models) on April 9, 2005. Table 2 and Fig.6 confirm the improved quality of the new operational system (e-suite) compared to the old one (o-suite) in term of wave height (H_s), mean period (T_z) and the Benjamin-Feir Index (BFI, described below) as computed from the 1-D buoy spectra.

Freak waves modeling

Janssen (2003) showed that it is possible to make probabilistic statements regarding the occurrence of freak waves by making use of the model spectrum. New operational integral parameters are now produced that characterize extreme sea states such occur in the presence of freak waves (Janssen and Bidlot 2003). These parameters are Goda's peakedness parameter, which provides a measure for the width of the wave spectrum, the Benjamin Feir Index, which is the ratio of the integral steepness and the relative width of the spectrum, and the kurtosis of the sea surface.

It is of interest to validate modelled extreme statistics parameters against observed counterparts as computed from the buoy 1-D spectra. In Fig.6, we have plotted the new Benjamin Feir Index for the 3 first months of 2005. The relative poorer quality (as compared to H_s or T_z) of the BFI can be linked to the determination of the model peak wave number that enters in the definition of the integral steepness. Peak wave number is a parameter that is from experience difficult to predict. Nevertheless, the wave model seems to produce a useful, nearly unbiased estimate of the Benjamin Feir Index.

Table 2. Impact of the New Model Version on Wave Height, Mean Period and Benjamin-Feir Index Statistics (see Fig. 6)

Scatter Index (%) [5] [Bias (model-buoy)]	o-suite	e-suite
H_s all buoys (m)	16.8 [0.12]	16.2 [0.09]
T_z all buoys (s)	14.1 [0.57]	9.9 [0.15]
BFI all buoys (-)	45.7 [0.01]	44.5 [-0.02]

5 The scatter index is defined as the standard deviation of the difference normalised by the mean of the observations.

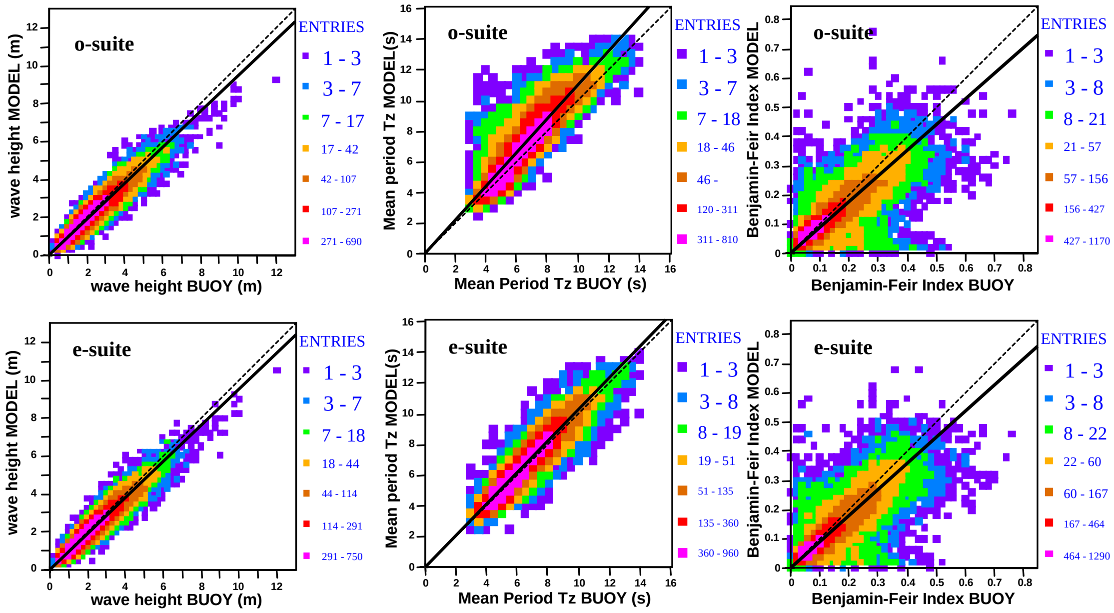


Figure 6: Comparison between 1-D frequency spectra from US and Canadian buoys and model operational analyses from January to March 2005. The model with the revised whitecapping dissipation and extended prognostic frequency range was part of the pre-operational suite (e-suite). It is compared to the model with the old formulation in the operational suite of the time (o-suite). See also table 2. The mean wave period is based on the integral of the second moment of the frequency spectrum

CONCLUSIONS

As demonstrated by the examples above, wave modellers should look at spectral wave observations to assess the different aspects of their model. It was shown that direction for possible improvements could be inferred. The quality of present wave models is such that one can no longer look at the total wave height alone. Comparison with wave model spectra should also be part of the validation. The buoy data used in this analysis are freely available. It is therefore very easy to set up a systematic verification procedure. More locations are however needed. We are urging data provider to facilitate the access to their data. It will greatly contribute to the continued improvement of wave model.

ACKNOWLEDGEMENTS

We would like to acknowledge the very useful contributions to the further development of wave model quality from the National Oceanographic Data Center and from the Canadian Marine Environment Data Service in freely providing the buoy data..

REFERENCES

Abdalla, S., P. Janssen, J.-R. Bidlot, 2003: Use of satellite data and enhanced physics to improve wave prediction. 28th International Conference on Coastal Engineering, Cardiff July 2002, *Word Scientific Publishing*.

Bidlot, J-R. , D. J. Holmes, P. A. Wittmann, R. Lalbeharry, H. S. Chen. 2002: Intercomparison of the performance of operational ocean wave forecasting systems with buoy data. *Wea. Forecasting*. 17. 287-310.

Bidlot, J-R., and Janssen P. 2003: Unresolved bathymetry, neutral wind, and new stress tables in WAM. *ECMWF Memo*. Research Department R60.9/JB/0400. ECMWF, Reading, United Kingdom, 17 pp. Available on request.

Bidlot, J-R., P. Janssen and S. Abdalla, 2005: A revised formulation for ocean dissipation in CY29R1. *ECMWF Memo*. Research Department R60.9/JB/0519. ECMWF, Reading, United Kingdom, 35 pp. Available on request.

Hardy, T.A., Mason, J.B., and McConochie, J.D., 2001: A wave generation model for the Great Barrier Reef, *Ocean Engineering*, 29, 45-70.

Hasselmann, S., C. Brüning, K. Hasselmann, and P. Heimbach, 1996: An improved algorithm for the retrieval of ocean wave spectra from SAR image spectra. *J. Geophys. Res.*, 101, C7, 16615- 16629.

Hasselmann, S., P. Lionello, and K. Hasselmann, 1997: An optimal interpolation scheme for the assimilation of spectral data. *J. Geophys. Res*., 102, C7, 15823-15836.

Klinker, E., Rabier F., Kelly F., Mahfouf J.-F., 2000: The ECMWF operational implementation of four-dimensional variational assimilation III: Experimental results and diagnostics with operational configuration, *Quart. J. Roy. Meteor. Soc*., 126, 1191-1215.

Lalaurette, F., L Ferranti, A. Ghelli, G. van der Grijn, 2003: Verification statistics and evaluation of ECMWF forecasts in 2002-2003. *ECMWF Tech. Memo*. 432. ECMWF, Reading, United Kingdom, 14 pp. Available via the ECMWF web pages [6].

Janssen, P.A.E.M., 2003: Nonlinear four-wave interaction and freak waves. J. Phys. Oceanogr. 33, 863-884.

Janssen, P., 2004: *The interaction of ocean waves and wind*, 300pp. Cambridge Univ. Press, Cambridge, UK.

Janssen, P.A.E.M., Abdalla, S. and Hersbach, H., 2003: Error estimation of buoy, satellite and model wave height data. *ECMWF Tech. Memo.* 402. Available via the ECMWF web pages [6].

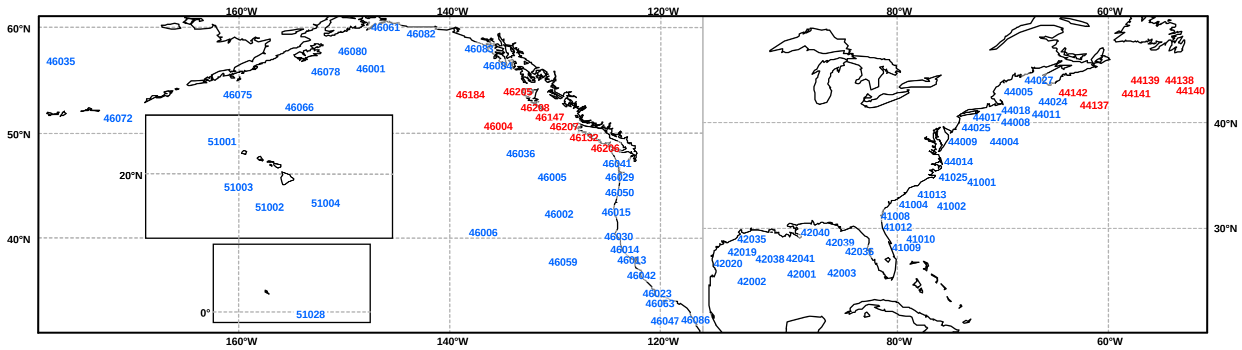
Janssen, P. and J. Bidlot, 2003: New wave parameters to characterize Freak Wave Conditions. *ECMWF Memo. Research Department* R60.9/PJ/0387. ECMWF, Reading, United Kingdom, 15 pp. Available on request.

Tolman, H. L., 2003: Treatment of unresolved islands and ice in wind wave models. *Ocean Modelling*, 5, 219-231

Vorripps, A. C., C. Mastenbroek, and B. Hansen, 2001. Validation of two algorithms to retrieve ocean wave spectra from ERS synthetic aperture radar. *J. Geophys. Res*. 106 (C8) 16825-16840.

Appendix: buoy locations

The following buoys were used. The NODC buoys are plotted in blue and the MEDS buoys in red.



6 http://www.ecmwf.int/publications/library/do/references/list/14