U. S. Department of Commerce
National Oceanic and Atmospheric Administration
National Weather Service
National Centers for Environmental Prediction
4700 Silver Hill Road, Mail Stop 9910
Washington, DC 20233-9910

Technical Note

Validation of WAVEWATCH III version 1.15 for a global domain.

Hendrik L. Tolman [‡]
SAIC-GSO at
Environmental Modeling Center
Ocean Modeling Branch

April 2002

[†] OMB Contribution No. 213.

[‡] e-mail: Hendrik.Tolman@NOAA.gov

Preface

The purpose of this report is to document the details regarding the original tuning and validation of the WAVEWATCH III model at NCEP for future reference. Most of this report was written prior to the operational acceptance of the global WAVEWATCH III model at NCEP in 1998. Because this report is mainly intended to fill a void in the documentation of the model development, no attempt has been made to update it with recent publications, particularly in the field of modelling source terms. More recent validation results can be found in Tolman et al. (2002). Note that the wind speed corrections suggested in Tolman (1998b) and used here no longer are needed as discussed in Tolman et al. (2002).

This report is available online as a full color pdf file from

http://polar.wwb.noaa.gov/omb/omb_notes.html

Abstract

Results of a study dealing with the validation of NCEP's new third-generation spectral ocean wave model (WAVEWATCH III v. 1.15) are presented. In this study, the model was used to provide global wave hindcasts in an operational environment. Data from buoys and the ERS1 altimeter are used for validation. The model results are also compared to those generated by the well-known WAM model. After some necessary modifications to the new model, including a systematic retuning and a modification of the interactions between the atmosphere and swell fields, WAVEWATCH shows excellent results for the entire global domain, with rms errors of typically 15% of the local mean observed wave height (based on altimeter and buoy data). WAVEWATCH generally outperforms WAM, particularly in the tropics and in the forecasts of extreme wave heights, but shows poorer results in selected high-latitude regions. This study furthermore shows that buoy data should not be used as the only data source in a global model validation, and that high-resolution validation with altimeter data identifies areas in which high-resolution wave modeling is required. Finally, shortcomings of present wave models and potential areas of improvement are discussed.

1 Introduction

This paper presents the second part of a validation study of a new ocean wind-wave forecast system at the US National Centers for Environmental Prediction (NCEP). This wave forecast system uses winds from NCEP's operational Global Data Assimilation System (GDAS, Kanamitsu, 1989; Derber et al., 1991), and from the operational Medium Range Forecast system (MRF Kanamitsu, 1989; Kanamitsu et al., 1991). In the first part of this study (Tolman, 1998b, henceforth denoted as part I), the quality of these wind fields has been assessed with buoy and satellite data for a three month period of December 1994 through February 1995. It has been shown that the open-ocean winds have a moderate bias, for which a statistical correction has been determined. With the corrected GDAS wind fields, the wave model is validated in the present study[1]. The validation is performed with buoy and

[1] In all results presented the model results for the first 8 days have been discarded to allow for model spin-up.

satellite data as described in section 2.

The present wave model is a recent version of WAVEWATCH (Tolman, 1991), as will be described in section 3. This model has been tuned for idealized fetch-limited conditions (Tolman and Chalikov, 1996, henceforth denoted as TC96). TC96 does not address (i) how to deal with the uncertainty of up to 40% in fetch-limited wave heights that can be attributed to atmospheric stability (Kahma and Calkoen, 1992, 1994, henceforth denoted as KC), (ii) the effects of feedback of momentum from swell to the atmosphere, and (iii) the 'saturation' level of wave growth. Consequently, there is room for fine-tuning the wave model in a practical application. Preliminary tests indicated that both the tuning of the model and the negative input for swell require additional attention, and that the effects of stability could be modelled explicitly. The necessary modifications to the wave model are discussed in section 3.

As a comparison with the state-of-the-art, results of WAVEWATCH are compared with results of a recent version of the well-known WAM model (WAMDIG, 1988; Komen et al., 1994). Computations with this model are described in section 4. Results of both models are presented and discussed in section 5. A final discussion and conclusions are presented in sections 6 and 7.

The validation of the wave model is partly performed with regression and bin-averaged (BA) analyses. Such methods are sensitive to observation errors. The present study uses error-corrected regression and BA techniques as suggested by Tolman (1998a), and as described in some more detail in part I. These techniques require an estimate of the observation error, which is discussed in Appendix A. The error-corrected analysis techniques will henceforth be denoted simply as the regression and BA analysis.

2 Data

Global wave model validations are conventionally performed with open-ocean buoy data. Hourly data for the 29 buoys presented in Table 1 (see also Fig. 1 of part I) have been obtained from the operational archive at NCEP. NCEP obtains these data in near real-time from GOES and GTS transmissions. Most buoys report wave spectra. From these spectra, wave heights can be computed up to the measurement accuracy. For the remaining buoys, only the wave height is transmitted, rounded off to the nearest 0.5m.

Table 1:
Buoys used in this study.

	WMO ID number	location		data type	group
1	46035	57.0°N	177.7°W	spectrum	
2	46003	51.9°N	150.9°W	spectrum	
3	46001	56.3°N	148.3°W	spectrum	
4	46184	53.9°N	138.8°W	wave height	A
5	46006	40.8°N	137.8°W	spectrum	
6	46005	46.1°N	131.0°W	spectrum	
7	46002	42.5°N	130.4°W	spectrum	
8	46059	38.0°N	130.0°W	spectrum	
9	51001	23.4°N	162.3°W	spectrum	
10	51003	19.2°N	160.8°W	spectrum	B
11	51002	17.2°N	157.8°W	spectrum	
12	51004	17.5°N	152.6°W	spectrum	
13	32302	18.0°S	85.1°W	spectrum	
14	41018	15.0°N	75.0°W	spectrum	C
15	42002	25.9°N	93.6°W	spectrum	
16	42001	25.9°N	89.7°W	spectrum	
17	41006	29.3°N	77.4°W	spectrum	
18	41002	32.3°N	75.2°W	spectrum	
19	41001	34.7°N	72.7°W	spectrum	
20	44008	40.5°N	69.4°W	spectrum	D
21	44011	41.1°N	66.6°W	spectrum	
22	44142	42.5°N	64.2°W	wave height	
23	44141	42.1°N	56.1°W	wave height	
24	44138	44.2°N	53.6°W	wave height	
25	62029	48.7°N	12.4°W	wave height	
26	62081	51.0°N	13.3°W	wave height	
27	62108	53.6°N	15.5°W	wave height	E
28	62105	55.9°N	14.2°W	wave height	
29	62106	57.0°N	9.9°W	wave height	

Unfortunately, buoy data cover only a small part of the oceans, and, therefore, cannot be the basis for a truly global model validation. A global validation can only be performed with satellite data. The present study therefore uses altimeter data. An altimeter renders a direct estimate of the variance of the water level within its footprint, and therefore represents a direct measurement of the wave height (unlike, for instance, altimeter wind observations, which are inferred from small-scale surface roughness). Here, fast delivery altimeter data from ERS-1 are used[2]. These data are not the most accurate altimeter data available, but systematic errors in the wave height retrievals are easily corrected (e.g., Cotton and Carter, 1994). The altimeter has a footprint with a diameter of approximately 6.5km. To get observation on a scale more similar to that of the wave model, and to apply additional quality control, these data have been averaged over 10s intervals. This results in an effective footprint of 6.5×65 km^2 (see also part I).

As mentioned above, the systematic error of altimeter wave heights should be corrected, and the random error needs to be estimated for the error-corrected validation techniques described in the introduction and in part I. Therefore, the averaged altimeter data have been collocated with buoy data with time and space lags of less than 30min and 100km, respectively. This resulted in 340 data pairs. As is discussed in Appendix A, the random observation error of the buoy data is much larger than the random altimeter error. Hence, the systematic error can be estimated with an 'inverse' regression of the buoy data on the altimeter results (e.g., Tolman, 1998a). From this regression, the error correction for the altimeter data can be estimated as

$$H_{a,c} = 0.10 + 1.17 H_a \,, \tag{1}$$

where $H_{a,c}$ and H_a are the corrected and original altimeter wave heights, respectively. The altimeter cannot observe extremely small wave heights, because an absence of waves corresponds to a discontinuous altimeter signal. The discretized altimeter signal can obviously never capture such conditions. It is therefore not surprising that (1) does not result in any observations below 0.55m. Such a low wave height cut-off might be expected to influence swell validation, and should therefore be removed. Unfortunately, the set of collocated data does not contain sufficient low-wave-height conditions to

[2] NCEP almost exclusively uses (reprocessed) fast delivery data due to operational constraints.

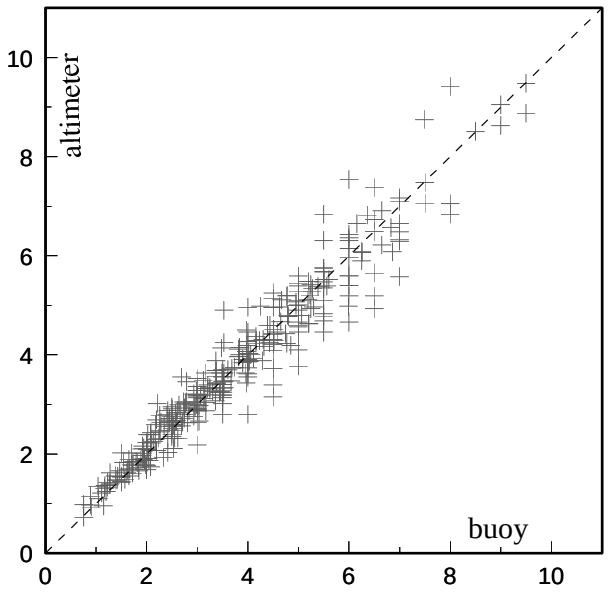

Fig. 1 : Collocated corrected altimeter wave heights (1) and buoy observations in meters. 340 observations.

perform a purely objective bias removal. Instead, (1) is somewhat arbitrarily replaced by a quadratic form for $H_a < 1.5$m with constant functional and derivative at the connection point and with $H_{a,c} = 0$ for $H_a = 0.5$m. The resulting collocated and corrected altimeter data are presented in Fig. 1. The averaged, quality controlled and error-corrected altimeter wave data will henceforth simply be denoted as the altimeter data.

3 The wave model

a General model description

The wave model used in NCEP's new forecast system is the most recent version of the WAVEWATCH model (WAVEWATCH III version 1.15), as described in detail in Tolman (1997). This model solves the action balance equation

$$\frac{D\,N(k,\theta;\mathbf{x},t)}{Dt} = \frac{S(k,\theta;\mathbf{x},t)}{\sigma}\,, \tag{2}$$

where N is the action density, defined as the variance density F divided by the intrinsic frequency σ, k and θ are the spectral wavenumber and direction, respectively, and $\mathbf{x}$ and t are space and time coordinates. Finally, S represents source terms for F. This model describes the spectral evolution of the wave field for slowly varying depths and currents. To avoid the loss of resolution in shallow water inherent to numerical modeling of the wavenumber spectrum, Eq. (2) is solved on a spatially varying wavenumber grid which corresponds to a spatially invariant σ-grid (Tolman and Chalikov, 1994; Tolman and Booij, 1998).

In the global application considered here, mean currents and water level variations are ignored. Shallow water effects are still included, although they are fairly irrelevant for a global grid. In such conditions, the governing equations of WAVEWATCH effectively reduce to the conventional balance equation for the variance spectrum $F(f,\theta)$. Converted to a longitude-latitude (λ-ϕ) grid, this equation becomes

$$\frac{\partial F}{\partial t} + \frac{1}{\cos\phi}\frac{\partial \dot{\phi} F \cos\phi}{\partial \phi} + \frac{\partial \dot{\lambda} F}{\partial \lambda} + \frac{\partial \dot{\theta} F}{\partial \theta} = S_i + S_n + S_d + S_b \,, \qquad (3)$$

where the source terms on the right side represent wind input, nonlinear interactions, dissipation ('whitecapping') and bottom friction, respectively. Expressions for the propagation velocities $\dot{\phi}$, $\dot{\lambda}$ and $\dot{\theta}$ can be found, for instance, in WAMDIG (1988).

WAVEWATCH has several options for the parameterization of the source terms and for the numerical propagation scheme. In the present application, the source terms of TC96 are used. These include an input (S_i) based on (Chalikov and Belevich, 1993, henceforth denoted as CB93), nonlinear interactions (S_n) based on Hasselmann et al. (1985), and the dissipation (S_d) of TC96. The bottom friction (S_b) is described with the conventional JONSWAP formulation (e.g. WAMDIG, 1988). A detailed description of these source terms can be found in the corresponding references, or in Tolman (1997).

In its present application, WAVEWATCH uses the third-order ULTIMATE QUICKEST propagation scheme (Leonard, 1979, 1991). As is shown in Tolman (1995), this scheme is sufficiently free of numerical diffusion to result in the disintegration of a continuous swell field into discrete swell fields (the 'garden sprinkler effect' Booij and Holthuijsen, 1987). This disintegration can be largely avoided by using appropriate propagation equations for

the discrete spectrum as derived by Booij and Holthuijsen. These equations include a diffusion tensor, the strength of which depends (amongst others) on the physical 'age' of the swell fields. A simplified version of these equations (as suggested by Booij and Holthuijsen), assumes a uniform and constant swell age, which is set at 4 days in the present model. Details of this propagation scheme can be found in Tolman (1995). Finally, the source term integration is performed with a dynamically adjusted time step algorithm (Tolman, 1992, 1997).

WAVEWATCH with the above described options is implemented on a longitude-latitude grid with a resolution of $1.25^\circ \times 1^\circ$ ranging from 78°N to 78°S. The minimum water depth allowed is 25m. The discrete spectrum consists of 24 directions and 25 frequencies, ranging from 0.042Hz to 0.41Hz (with the conventional exponential distribution with $f_{i+1} = 1.1 f_i$). Wind fields at 10m height are obtained from GDAS, and are corrected as described in Eqs (6) through (8) of part I. Grid points with a sea-ice concentration of more than 33% are treated as land. Ice concentrations are obtained from NCEP's automated passive microwave sea ice concentration analysis (Grumbine, 1996) and are updated daily. Finally, WAVEWATCH uses four different time steps. The first time step is the time increment at which the entire solution is propagated, and at which the input wind field is interpolated (linear in magnitude and direction). This time step also defines the maximum time step in the source term integration. The second time step is the minimum time step for source term integration. The third time step is the maximum propagation time step for the longest wave components in the spectrum. Maximum time steps for shorter components are set within the model, and scale linearly with the frequency. The final time step is the refraction time step. The four time steps are set to 3600s, 300s, 1300s and 3600s, respectively.

b Model modifications

Conventionally, model validation and tuning are performed using buoy data. However, as discussed in the pervious section, both the global coverage and quality of altimeter wave data makes these data more appropriate to use. It proved most practical to concentrate on the global distribution of systematic model errors (biases). The global distribution of random and total errors proved to be fairly close to minimal if biases were minimized. Based on the global minimization of biases relative to the altimeter data, three modifica-

tions to the wave model were deemed necessary or useful; (a) the model was retuned, (b) the swell momentum feedback mechanism was modified, and (c) effects of stability are modeled explicitly.

Observed growth curves show a large scatter, which for a significant part appears to be related to effects of atmospheric stratification. As shown by KC, the effects of stratification cannot be explained by the conventional effects of buoyancy on vertical momentum transport, and can therefore not be addressed by accounting for buoyancy effects in the surface stresses. The difference of up to 40% in wave heights for neutral versus unstable conditions, raises the question as to which growth curves the model should be tuned. Because on average atmospheric stratification is close to neutral for most of the oceans, it appears logical to tune a model for such conditions, and the corresponding constants in the TC96 source terms have therefore been used as a starting point in the model validation (TC96 Table 2, top row).

Perhaps somewhat surprisingly, such a model results in a systematic underestimation of wave heights in storm tracks of approximately 30% (figures not presented here). This suggest that the model should be tuned more closely to fetch-limited growth for unstable conditions. This apparent deficiency of the model is shared by the WAM model, which has been documented to give good results in the open ocean, yet reproduces fetch-limited growth curves for unstable conditions (e.g., TC96 Fig. 7). Possible reasons for this behavior will be discussed in section 6.

To remove these systematic biases the model was retuned. Intuitively, the dissipation source term could be retuned objectively, as outlined in TC96. This is however, a cumbersome process, which is further complicated because it is not apriori clear as to which growth curve the model should be tuned. Furthermore, retuning the model in this way compounds underlying problems with parameterization of the physics as will be discussed in section 6. An alternative method is to define an 'effective' wind speed internal to the model. This is a simple way to retune the model, and is fairly elegant because it keeps the present balance of source terms intact. It furthermore allows for a simple way to include explicit effects of stability at a later stage (see below). Considering that the wave height scales roughly with the square of the wind speed, the effective wind speed u_e is calculated from the actual wind speed u as

$$u_e = \sqrt{c_0}\, u \,, \tag{4}$$

where c_0 is the tunable parameter.

With $c_0 = 1.3$ WAVEWATCH gives reasonable results in closed basins and in storm tracks at higher latitudes, but results in a severe underestimation of the wave heights at lower latitudes. Particularly, in the tropical Pacific Ocean, negative biases exceed 1.5m or 50% of the mean wave height in large areas (figures not presented here). This model deficiency was traced to the effect of the input source term S_i on swell. This source term has the general form

$$S_i = \beta F \,, \tag{5}$$

where β is the wind-wave interaction parameter. A feature of the input term of CB93 is that β becomes negative for spectral components that travel faster than the wind or in directions opposite to it, representing a feedback of momentum from the waves to the atmosphere. Removal of the 'negative input' resulted in a reversal of the bias patterns, having little impact in storm tracks but resulting in positive biases larger than 1.5m in the tropics (figures not presented here). This implies that the momentum feedback mechanism is necessary, but that it is too strong in CB93. Considering this, a modified input source term $S_{i,m}$ is introduced

$$S_{i,m} = \begin{cases} S_i & \text{for} \quad \beta \geq 0 \quad \text{or} \quad f > 0.8 f_p \\ X_s S_i & \text{for} \quad \beta < 0 \quad \text{and} \quad f < 0.6 f_p \\ \mathcal{X}_s S_i & \text{for} \quad \beta < 0 \quad \text{and} \quad 0.6 f_p < f < 0.8 f_p \end{cases} \,, \tag{6}$$

where f is the frequency, f_p is the peak frequency of the wind sea, calculated from S_i using Eqs. (36) and (38) of TC96, and $0 < X_s < 1$ is a reduction factor for S_i, which is applied to swell with negative β only. $\mathcal{X}_s$ represents a linear reduction of X_s with f_p providing a smooth transition between the original and reduced input. Because this modification considers swell only, it does not influence growth behavior of the model as addressed by TC96 (figures not presented here). Biases in the open ocean are virtually eliminated by choosing $X_s = 0.125$ (figures not presented here).

The above correction of S_i is purely pragmatic, but can be justified on other grounds. First, the calculations on which the adjusted part of the CB93 input term is based were performed at coarse model resolution and partially interpolated and extrapolated. Chalikov (personal communication) recently has repeated the calculations of CB93 with a more efficient and ac-

curate model, and found that the negative input to swell was overestimated significantly in CB93. Furthermore, this input model implicitly assumes a correlation between resolved spectral wave components and unresolved high-frequency roughness. Whereas such a correlation might be expected for dominant spectral components of a wind sea, it is expected to be less pronounced for low-steepness swells. Reduced correlations result in a reduced momentum feedback to the atmosphere.

The last issue to address is the impact of stability on wave growth. Most of the atmospheric boundary layer will be close to neutral stability, but systematically unstable conditions occur over the Gulf Stream and the Kuroshio in the northern hemisphere winter. To test if stability has a systematic influence on model errors, such errors are assessed as a function of the stability parameter

$$\mathcal{S} = \frac{hg}{u_h^2} \frac{T_a - T_s}{T_0} , \tag{7}$$

where u_h and T_a are the wind speed and air temperature at height h, and T_s and T_0 are the sea and reference temperature, respectively. Because only the buoy data contain all necessary observations to assess $\mathcal{S}$, the analysis will consider buoy data only. To asses effects of stability on wave growth, wind sea events have to be isolated. Furthermore, dependent observations can be removed and sampling errors in the observations can be reduced by considering one observation per event only. The consolidation of data into events is discussed in Appendix B.

Figure 2 shows normalized errors H_m/H_o of the retuned wave model with $X_s = 0.125$ as a function of the stability parameter $\mathcal{S}$, where the suffices m and o identify the model and observations, respectively. To isolate wind seas, only modeled wave ages $u/c_p < 1.35$ are considered, where c_p is the phase velocity corresponding to the peak frequency of the one-dimensional frequency spectrum. This figure suggests that the model error is a (fairly weak) function of the atmospheric stability. The apparent underestimation of wave heights in unstable conditions is consistent with the observed higher growth rates in such conditions (KC). Ideally, the effects of unstable stratification are incorporated in the input source term using physical considerations. Developing such a parameterization is well outside the scope of the present study. Instead, an empirical stratification correction can be defined based on the effective wind speed as defined above in the retuning of the model. Using the dashed line in Fig. 2 as an estimate of the wave model behavior, and

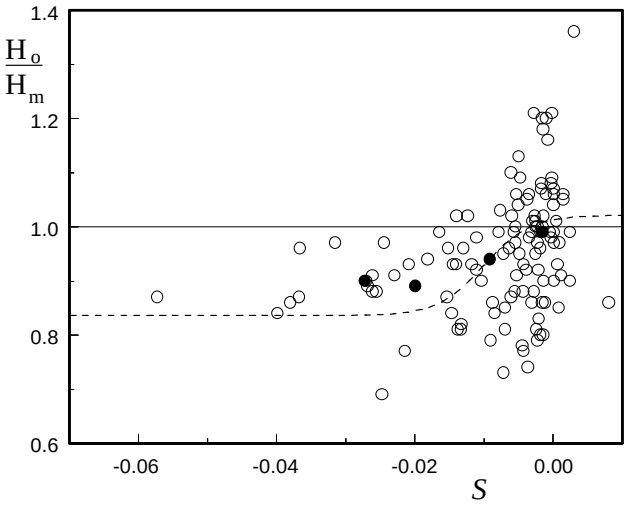

Fig. 2 : Normalized wave model errors H_m/H_o as a function of the stability parameter $\mathcal{S}$ for events at buoy locations for wind sea events (see text). ∘ : data. •: averaged data for ranges $\Delta\mathcal{S} = 0.01$. dashed line: assumed model behavior on which Eq. (8) is based.

assuming that the wave height scales with the square of the wind speed, the following effective wind speed is used.

$$u_e = u \left(\frac{c_0}{1 \pm c_1 \tanh[\pm c_2(\mathcal{S} - \mathcal{S}_o)]} \right)^{1/2} , \tag{8}$$

where c_0, c_1, c_2 and $\mathcal{S}_o$ are set to 1.4, 0.1, 150 and -0.01, respectively, and where the sign equals the sign of $(\mathcal{S} - \mathcal{S}_o)$. Compared to the effective wind speed in Eq. (4), this correction is expected to decrease wave height for stable conditions by about 3%, and increase wave height for unstable conditions by up to 17%.

Evaluation of $\mathcal{S}$ requires global estimates of the air and sea surface temperatures T_a and T_s. These temperature are obtained along with the wind speeds from the lowest level of GDAS. These sea surface temperatures are highly smoothed, and therefore do not represent the Gulf Stream and Labrador currents particularly well. This is not a serious drawback for the present study as the stability changes are mostly due to variability of the air temperature.

In the following, WAVEWATCH with the effective wind speed (8) and the

corrected input source term (6) will simply be denoted as WAVEWATCH.

4 The WAM model

The performance of the new model is evaluated by comparing its wave forecasts and those produced by the WAM model with buoy and altimeter observations. WAM cycle 4 with the with available updates from MPI Hamburg has been used for this purpose and it has been implemented to run with the same spatial and spectral grids and wind input as used for WAVEWATCH. The overall time step in WAM is set to 1200s. Ice coverage was not included in this model, but its effects proved highly localized in NCEP's operational implementation of WAM. By considering results north of 60^o S in the following only, effects of omitting ice coverage in the WAM runs is expected to be negligible in the model comparison.

5 Model validation and comparison

As discussed in the previous sections, the validation of WAVEWATCH and the comparison with WAM should concentrate on altimeter data. To collocate modeled wave heights with altimeter data hourly wave height fields were generated for both models. The model results at the altimeter track (and observation time) were obtained by tri-linear interpolation from these hourly fields. From such collocated results scatter plots etc. can be produced directly. To obtain an impression of the spatial distribution of model errors, collocated altimeter data and wave model results in turn are collocated with the closest model grid point, requiring a minimum of 7 collocations per grid point to assess error parameters. To remove signatures of individual satellite tracks from such results, they are smoothed in longitudinal direction with a filter width of approximately 5°.

Figure 3 shows the resulting average altimeter wave heights for the three month period considered. As expected, the highest wave heights occur in the storm tracks at high latitudes, in particular in the northern hemisphere. Low wave heights are found in the tropics, with a local maximum in the Pacific ITCZ (0-15^o N).

Figure 4 shows the global distribution of biases and random errors of WAVEWATCH as estimated from the altimeter data. The biases (panel a)

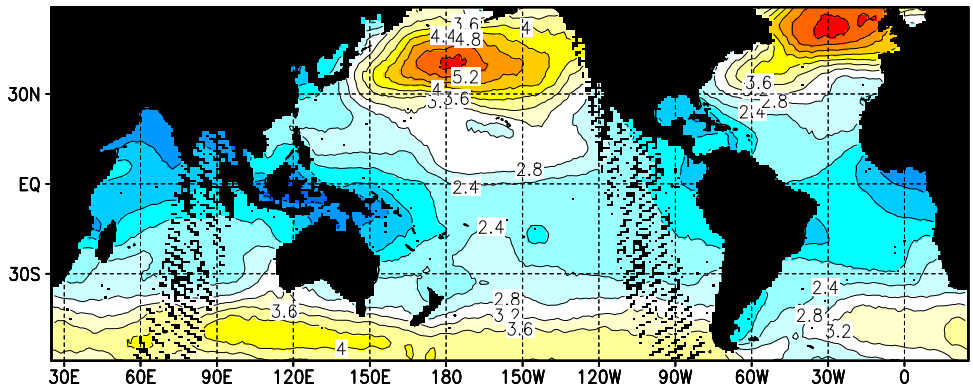

Fig. 3 : Mean altimeter wave height in meters for December 9 1994 through February 28 1995 as obtained by collocation with the wave model grid. Black indicates missing data (less than 7 observation per grid point). Grid lines at 30^o intervals. Latitudinal smoothing with a filter width of approximately 5^o applied.

are generally small, i.e., of the order of $\pm$0.3m or less, and are more or less randomly distributed. Larger biases occur in several areas. Large positive biases exceeding 0.75m occur in the Bering Sea north of the Aleutian Islands. These biases can be attributed to the fact that the Aleutian Islands are not resolved by the model grid. Hence, in the model swell is traveling unhampered into the Bering sea from the south, whereas such swell in reality is dissipated at the islands. This will lead to positive biases around unresolved Islands and other topographical features. Similar positive biases appear around the Bismarck and Solomon Islands, and to a lesser extent around French Polynesia, the Caribbean Islands, the Shetland Islands and Hawaii. Random model errors (Fig. 4b) range from typically 0.3m in the tropics to 1m in the northern Pacific storm tracks. Random errors seem to be somewhat increased by unresolved Island chains (particularly the Aleutian Islands). The effects of such unresolved features on random errors is much less pronounced than the corresponding impact on biases.

Figure 5 shows the corresponding biases and random errors of WAM. The biases of WAM are generally somewhat larger than those of WAVEWATCH (compare Figs. 4a and 5a), but are generally well below 0.5m The biases of WAM show a systematic distribution with low biases in the storm tracks at high latitude, and high biases in the tropics. Because the latter wave fields

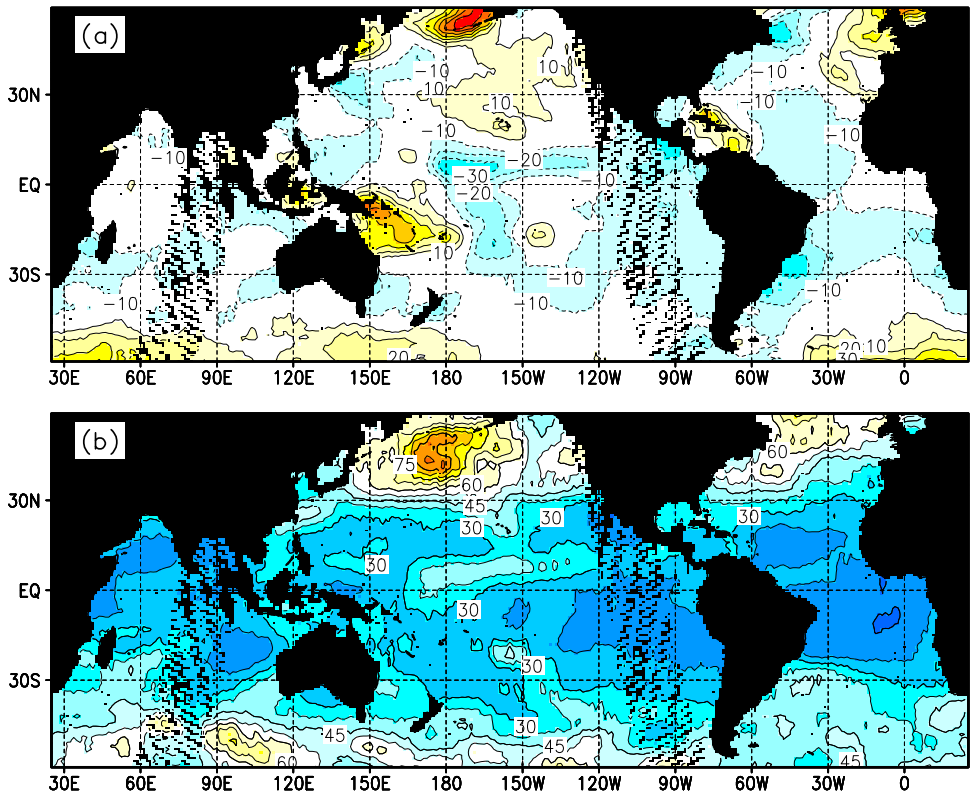

Fig. 4 : Global distribution of biases (panel a, in cm) and standard deviation errors (panel b, in cm) of WAVEWATCH relative to altimeter observations as in Fig. 3.

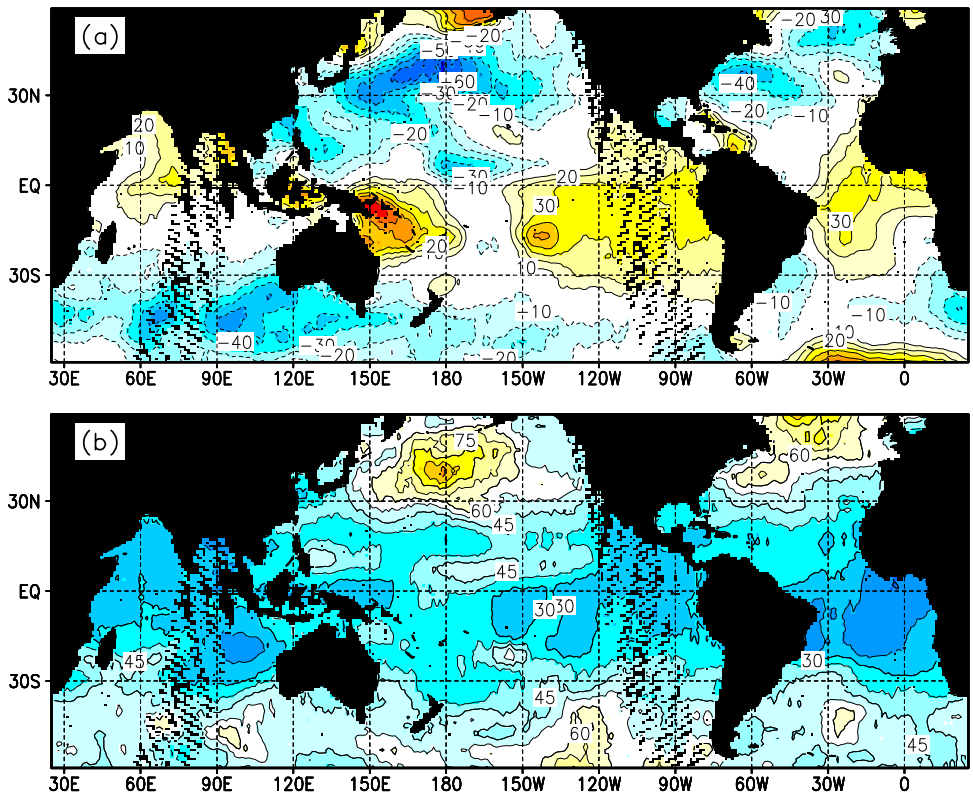

Fig. 5 : Like Fig. 4 for WAM.

consist mostly of swell, this suggests that swell attenuation is insufficient in WAM. This could furthermore be interpreted as supporting the need for a swell momentum feedback mechanism ('negative input') as is present in WAVEWATCH but absent in WAM. Random errors of WAM and WAVEWATCH are fairly similarly distributed (compare Figs. 4b and 5b), with somewhat smaller errors for WAM at high latitudes, and somewhat smaller errors for WAVEWATCH elsewhere.

A concise way to address the relative importance of model errors or improvements is by considering the normalized total model error. This parameter is also known as the 'scatter index' (S.I.) and is defined as the rms model error normalized with the mean observed value. For WAVEWATCH the scatter index S.I. < 0.15 for most of the deep ocean (Fig. 6a), i.e., the rms model error is generally less than 15%. Larger errors are found in the areas where unresolved islands negatively impact model results, and in areas where low wave heights make the S.I. sensitive, particularly in the Arabian Sea, at the East Coast of South America and around Central America. Scatter indices for WAM are generally higher (Fig. 6b), particularly in tropical waters where S.I. $>$ 0.20 for large areas. The improvements of WAVEWATCH over WAM are clear when the differences in S.I. between models are plotted (Fig. 6c). In large areas in the tropics the S.I. of WAVEWATCH is better that the S.I. of WAM by more than 0.1, in particular in the eastern tropical Pacific. Only in the Norwegian Sea and in the Indian Ocean south of 45° S WAM systematically outperforms WAVEWATCH with S.I.'s which are better by up to 0.05. WAM also gives better results in the Bering Sea, but both models are suspect in this area due to the unresolved Aleutian Islands.

The above analysis of model errors considers mean error parameters. Also of interest is the behavior of errors as a function of wave height. Therefore, biases (β), standard deviations (σ) and rms errors (ϵ) as a function of the observed wave height have been estimated using a bin-averaged (BA) analysis. The required estimates for observation errors are discussed in Appendix A. In this analysis data for areas with significant errors due to unresolved geographical features have been omitted. These areas are the Bering Sea (165°E to 160°W north of 50°N), the Indonesian, Solomon and Bismarck Islands (120°E to 170°E and 0 to 20°S as well as 105°E to 120°E and 0 to 8°S) and the Caribbean Islands (70°W to 90°W and 20°N to 30°N).

Figure 7 shows biases β as a function of the wave height H_s. WAVEWATCH (solid lines) typically shows moderate high biases for low wave heights and low biases for high wave heights. As discussed in part I for

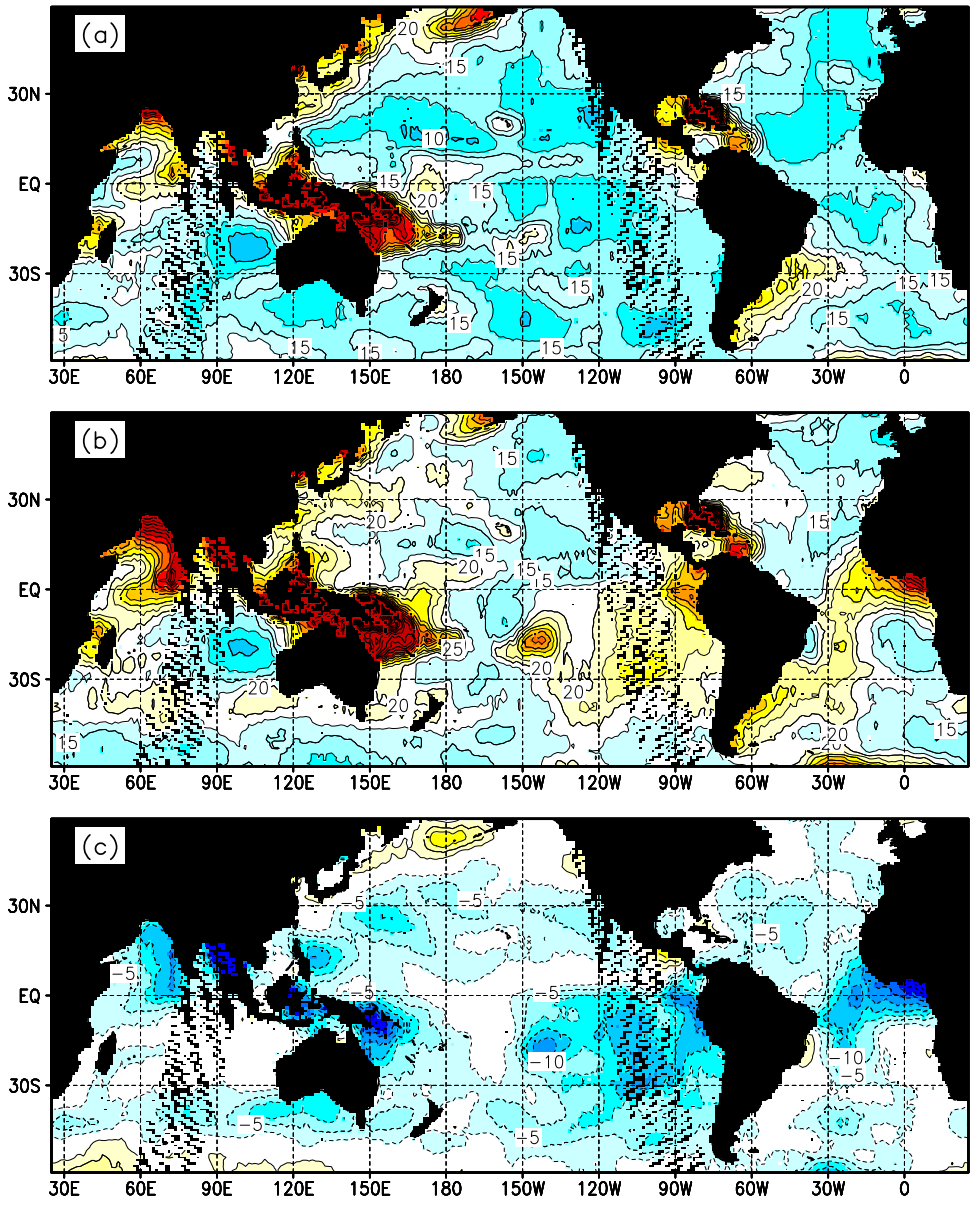

Fig. 6 : Global distribution of scatter indices (S.I.) in % for WAVEWATCH (panel a), WAM (panel b) and the differences (WAVEWATCH - WAM) relative to altimeter observations as in Fig. 3.

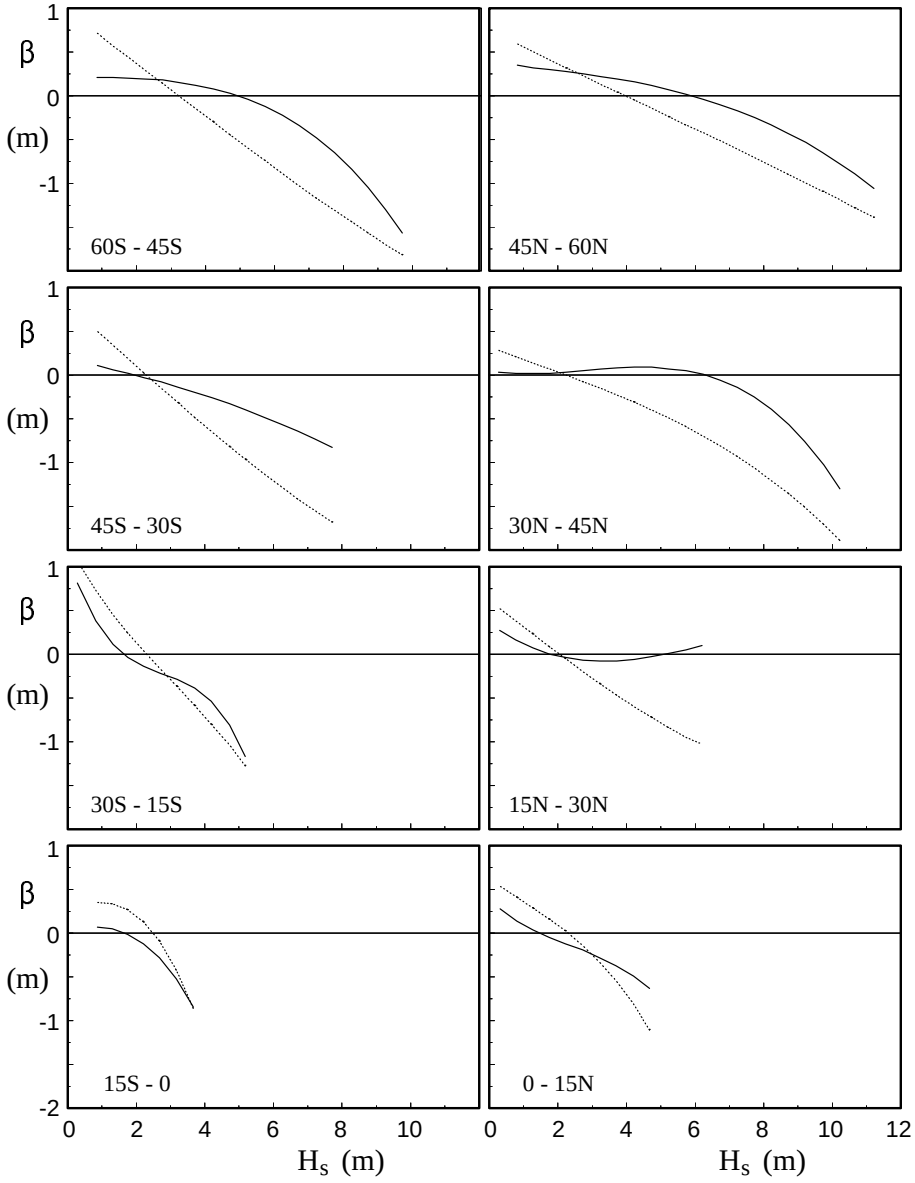

Fig. 7 : Model biases β relative to ERS1 data as a function of the observed significant wave height H_s for eight latitude bands as determined with a BA analysis. Bin width 0.5 m, minimum number of observations per bin is 20. Third-order polynomials fit to actual BA results. Solid lines: WAVEWATCH. Dotted lines: WAM.

wind speeds, such behavior can be due to either systematic or random errors in location and strength of wind and wave systems. To assess the origin of the biases, observed and modelled wave height distributions have been analyzed. Selected distributions are presented in Fig. 8.

The BA analysis for the latitude range of 15-30°N shows virtually no model biases for WAVEWATCH (solid line in corresponding panel of Fig. 7). The corresponding wave height distributions for the altimeter and WAVEWATCH ($\circ$ and solid line in Fig. 8a) are also virtually identical, even for large wave heights (see enlarged inset in Fig. 8a). The latter suggest that the model results indeed are virtually free of systematic errors for arbitrary wave heights in this latitude range. The BA analysis for the latitude range of 30-45°N shows virtually no model biases for WAVEWATCH for wave heights below 8m, but low biases for higher waves. The corresponding wave height distributions (Fig. 8b) in general show a good agreement between observed ($\circ$) and modelled (solid lines) distributions, but also that the model systematically underestimates the number of events with wave heights above 10m (inset in Fig. 8b). This suggests that the negative biases for high wave heights for this latitude range in Fig. 7, at least partially, represent a systematic model error. The BA analysis for the latitude range of 30-45°S shows a systematic evolution of the wave bias as a function of the wave height. The corresponding wave height distributions (Fig. 8c) again are fairly similar, but for high wave heights they show that the corresponding bias estimates again are at least partially due to systematic model behavior. For all other latitude ranges, wave height distributions (not presented here) show a similar consistency with bias estimates.

Figures 7 and 8 also show the corresponding results of WAM. WAM shows low biases for high wave heights and vise versa which are systematically larger than those of WAVEWATCH (compare dotted and dashed lines in Fig. 7). This results in wave height distributions which are systematically too narrow (compare dotted lines and $\circ$ in Fig. 8).

Figure 9 shows standard deviations σ and overall rms errors ϵ as a function of the wave height H_s for the eight latitude bands. WAVEWATCH typically shows random errors $\sigma \approx 0.3 - 0.4$m for low wave heights, and of $\sigma \approx 0.15 H_s$ for higher waves (lower solid lines). Due to the generally small biases rms errors ϵ (upper solid lines) are nearly identical to the random errors σ except for the highest waves in all distributions. WAM shows generally lower random errors σ (compare lower dotted and solid lines), particularly at high latitudes. If, however, the biases are also considered, the overall errors ϵ (upper solid

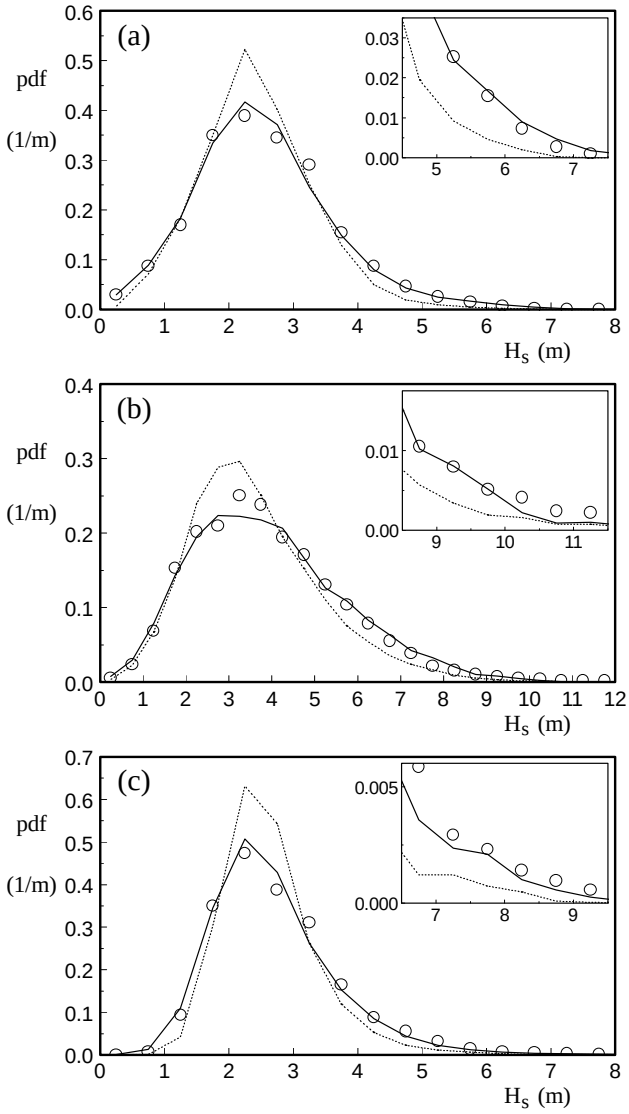

Fig. 8 : Probability density functions (pdf's) of significant wave heights for selected latitude bands from altimeter (∘), WAVEWATCH (solid lines) and WAM (dotted lines). Bin width 0.5 m. (a) 15°N - 30°N. (b) 30°N - 45°N. (c) 30°S - 45°S.

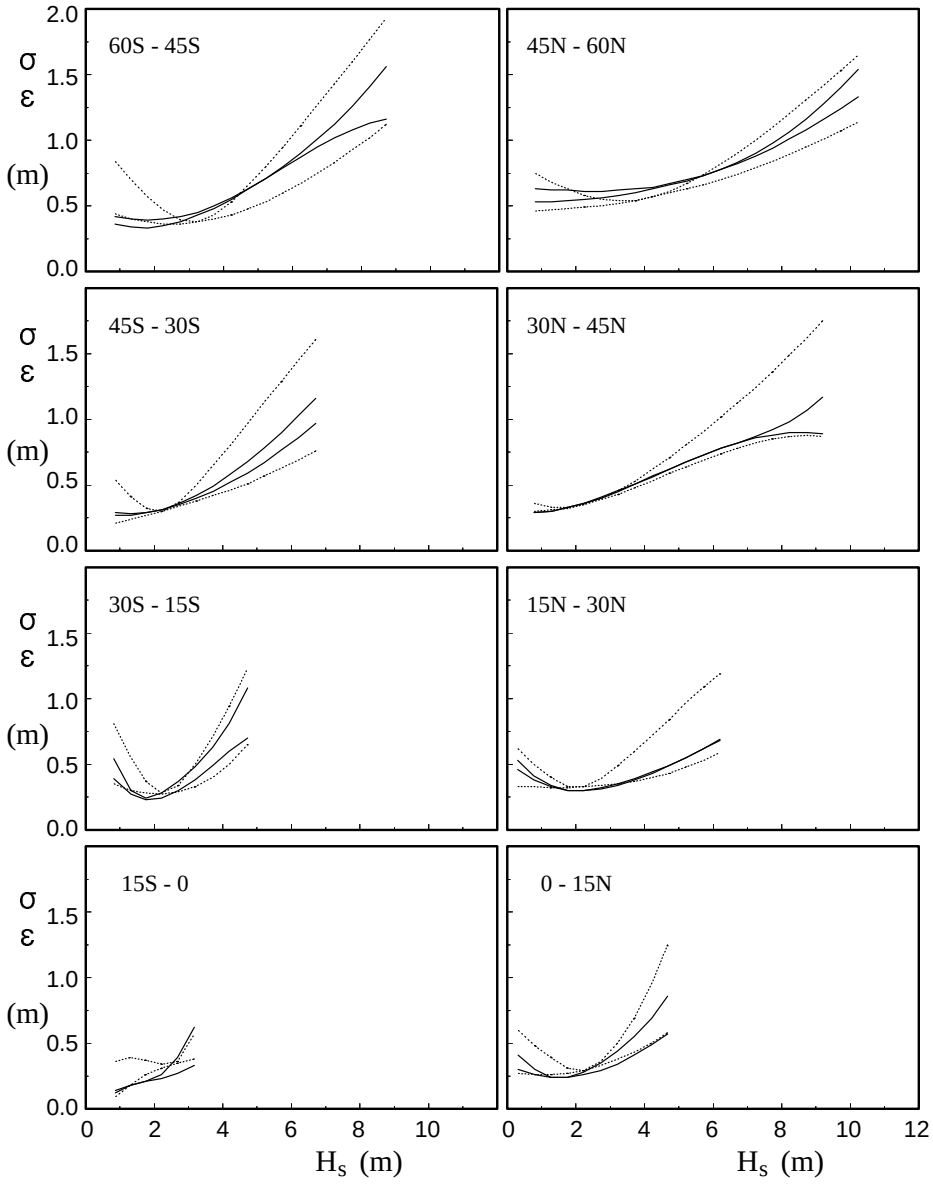

Fig. 9 : Like Fig. 7 standard deviation errors (σ, lower lines) and rms errors (ϵ, upper lines). Minimum number of observations per bin is 40.

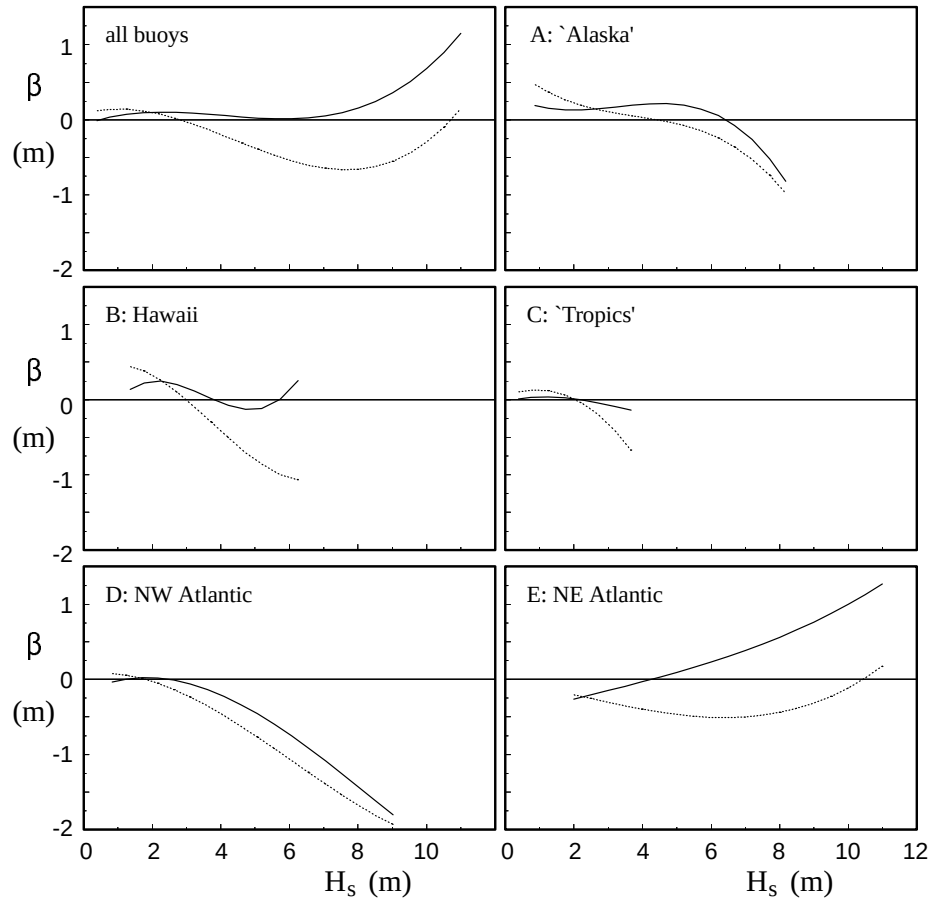

Fig. 10 : Like Fig. 7 using buoy data for regions as defined in Table 1.

and dotted lines) of WAM and WAVEWATCH are similar near the center of data distributions, but are systematically larger for WAM for extremely high or low wave conditions.

A model validation and comparison cannot be complete without also considering buoy data. Therefore biases, distributions, random and rms errors as determined from the buoy data are presented in Figs. 10 through 12. The amount of data available is sufficiently large to allow for a subdivision into geographical regions Table 1). In general the comparison to the buoy data gives qualitatively and quantitatively similar results as the above comparison to the altimeter data, i.e., (i) biases tend to be low for high wave heights

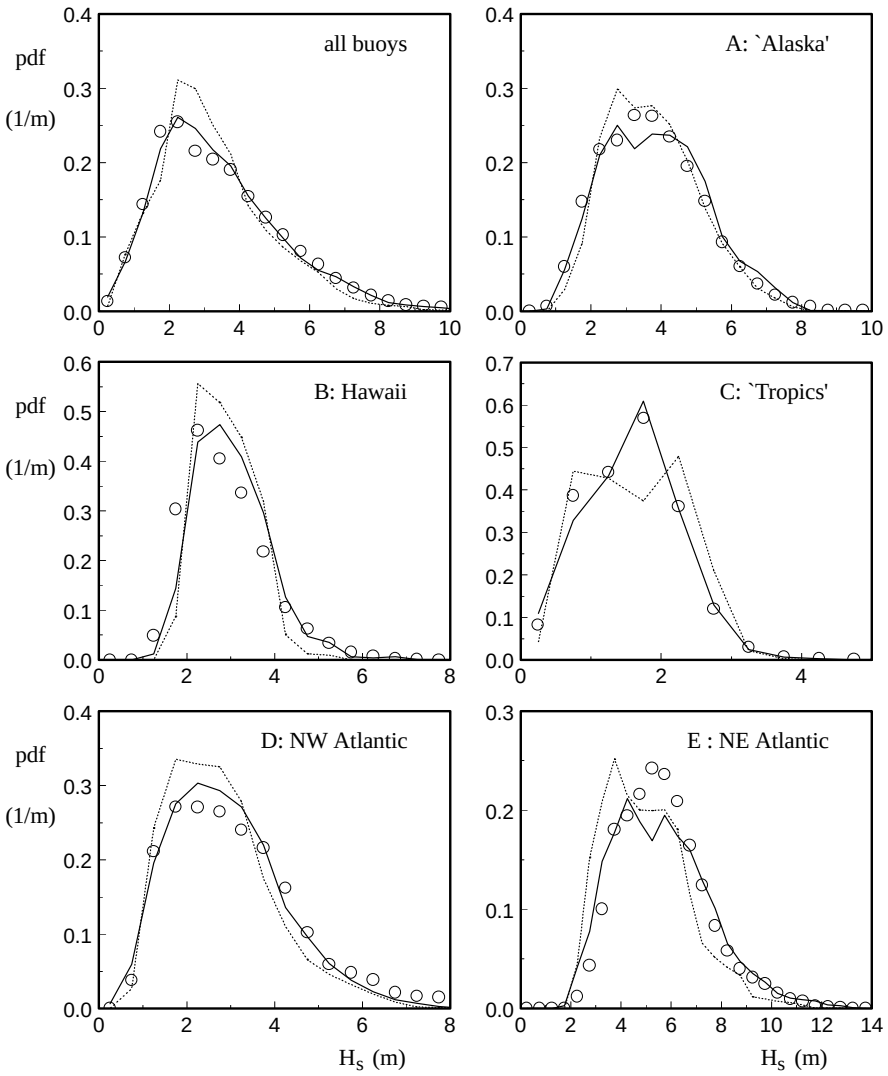

Fig. 11 : Like Fig. 8 using buoy data for regions as defined in Table 1.

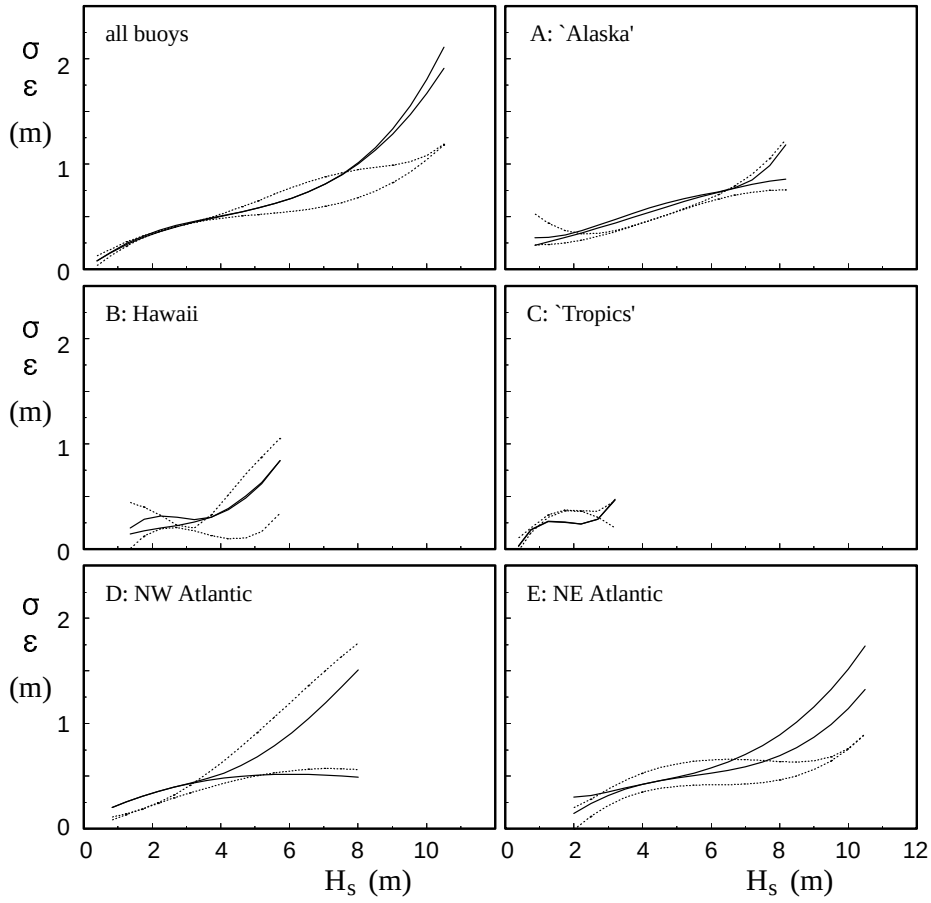

Fig. 12 : Like Fig. 9 using buoy data for regions as defined in Table 1.

and vise versa, with (ii) a stronger dependence of the bias β on the wave height H_s for WAM than for WAVEWATCH (Fig. 10). (iii) WAVEWATCH reproduces the observed wave height distributions better than WAM which produces distributions that are too narrow (Fig. 11). (iv) Random errors σ are generally somewhat lower for WAM, particularly at higher latitudes, but rms errors ϵ are generally lower for WAVEWATCH, in particular for extreme wave heights (Fig. 12). There are, however, several differences that should be addressed in some more detail.

First, the results for the buoy data in region D (NW Atlantic Ocean) show an anomalously negative bias for high wave heights, whereas the data in region E (NE Atlantic Ocean) result in an anomalously positive bias for high wave heights when compared to all other results (Figs. 10 and 7). Because this behavior is present in both models, it might well be, at least partially, due to peculiarities of the wind fields or of the model discretization (several of the buoys considered are rather close to the coast). It furthermore shows that localized buoy observations do not necessarily produce a globally valid model validation.

Secondly, WAM performs systematically better when compared to buoy data than when compared to altimeter data (e.g., rms errors in Figs. 9 and 12), whereas WAVEWATCH shows more consistent errors against both data sets (particularly for extreme wave heights).

Finally the wave height distribution of WAM for region C is anomalously broad (Fig. 11). This is due to the fact that the wave heights of WAM for buoy 32302 are systematically high, whereas the wave heights for the other three buoys in this group are low, resulting in the rather broad bimodal distribution.

6 Discussion

This study presents the results of the application of WAVEWATCH-III version 1.15 to generate operational global wave forecasts. This model includes third-order propagation schemes with a correction for the description of continuous dispersion in a discrete spectral model, and new input and dissipation source terms of TC96. The latter source terms had not been tested in realistic conditions, and required some modifications as described in section 3.b. With these modifications, WAVEWATCH shows excellent results; biases are generally small and are more or less randomly distributed over the global

domain (Fig. 4a), and overall rms model errors are for most of the global domain of the order of 15% (Fig. 6a). WAVEWATCH tends to systematically underestimate the highest waves, and to somewhat overestimate the lowest waves (Figs. 7, 8 and 10). In particular the former, at least partially, can be traced to numerical rather than physical errors; first, the wave model is driven by wind fields at roughly the nominal resolution of the GDAS. This implies that small and intense systems are subject to some smoothing, resulting in a systematic underestimation of peak wind speeds and hence peak wave heights. Secondly, data and model results are collocated using multi-dimensional interpolation, which is also tends to reduce extreme values.

WAVEWATCH compares favorably to the well established WAM model. Mean biases are generally smaller, and more randomly distributed over the global domain (Figs. 4a and 5a). Furthermore, biases are generally less dependent on the wave heights and are smaller for extreme wave heights (Figs. 7 and 10). The rms errors of WAVEWATCH are significantly smaller in the tropics (Fig. 6) and for extreme wave heights (Figs. 9 and 12). Finally, the results of WAVEWATCH more closely reproduce observed wave height distributions (Figs. 8 and 11). WAM, on the other hand, shows smaller rms errors in some regions at high latitudes (Fig. 6c) and smaller biases and rms errors for extreme wave heights when compared to the buoy data in the NE Atlantic Ocean.

It should be noticed that WAVEWATCH and WAM have been tuned and/or tested in different ways. WAM has been compared extensively to buoy data for almost a decade. The testing and tuning of WAVEWATCH (section 3), on the other hand, leans more heavily on altimeter data. In this context, several issues require additional discussion.

First, WAVEWATCH is partially tuned using the present validation data, so that the good model results might well be due to the tuning to this particular data set. However, the tuning only considered a very limited set of the validation parameters, i.e., mostly the global distribution of mean biases. Furthermore, the present validation results for WAVEWATCH appear to be closely reproduced by more recent computations and comparisons with altimeter and buoy data (to be presented elsewhere). It therefore does not appear that the dual use of the data set contaminates the validation results.

Second, considering the different tuning strategies, it is not surprising that the altimeter data more strongly favors WAVEWATCH than the buoy data. In this context it is encouraging to see how well WAVEWATCH verifies against the buoy data which have played only a minor role in its tuning.

Third, the present study illustrates the danger of using buoy data only for the validation of a global model. By nature, buoy data have been available in selected areas only, and therefore can only produce local error estimates. For instance, the bias behavior of the models relative to the NE and NW Atlantic buoy data (Fig. 10) appears somewhat anomalous when compared to the global data. Furthermore, the buoy data suggest that WAM produces scatter indices S.I. of typically 10 to 15%, whereas in large areas not covered by buoy data the S.I.'s are much larger. This, of course, does not mean that buoy data are not important. In fact, the buoy data are very helpful to identify local peculiarities of a wave forecast system, and are of paramount importance in the validation and tuning of the altimeter wave height retrieval algorithms.

Fourth, the detailed global analysis based on the altimeter data has the added benefit that it indicates the areas in which significant errors occur due to unresolved geographical features such as the Aleutian Islands. Such areas will benefit most from high-resolution nested models, or from a sub-grid approach to such features.

Wave models such as WAM and WAVEWATCH with scatter indices of the order of 15% can generally be considered to be excellent forecast tools, particularly because scatter indices of wind fields are generally larger (see part I). The latter may be somewhat surprising, because the wave height is usually assumed to scale with the square of the wind speed. This in principle implies that the wave model doubles the relative error of the driving wind field. The wave model, however, mostly reacts to synoptic scale variations in the wind fields. This implies that a systematic increase of the wind speed by say 2% will indeed increase the wave height by about 4%. However, the time scales of wave growth are typically too long for the wave model to respond to turbulence, gustiness and fast mesoscale wind variations. A significant part of the random wind speed error occurs at such scales and will therefore not be of importance for the wave model. Hence, larger relative wind speed errors are not in conflict with the quadratic scaling behavior of the wave model.

In spite of the fairly good performance of the wave models considered here, there is much room left for improvement. Particularly, WAVEWATCH and WAM share the deficiency that open-ocean wave growth can only be reproduced adequately with a model that over-estimates the growth in idealized fetch-limited conditions (section 3). This reduces the applicability of the present models and their physics parameterizations to arbitrary scales and conditions. Tentatively, the source of this deficiency might be found

in the shared parameterizations of the physics, i.e., the Discrete Interaction Approximation to the nonlinear interactions (DIA Hasselmann et al., 1985).

Whereas the development of the DIA was a giant step forward for numerical wave modeling, the DIA has two distinct shortcomings that influence wave growth behavior. First, the DIA severely overestimates the nonlinear interactions at high frequencies (Hasselmann et al., 1985, Fig. 7). In particular for initial growth, the spuriously large positive lobe at low frequencies will intersect the cut-off frequency in a wave model, resulting in a net dissipation for the resolved part of the spectrum. Numerical experiments have shown that this net dissipation can be up to 30% of the integral input over the resolved spectrum. Such a large artificial dissipation influences the source term balance and hampers tuning of the model through adaptation of the dissipation source term. Furthermore, compared to the 'exact' interaction calculations, the DIA significantly broadens the spectrum (Hasselmann et al., 1985, Figs. 8 and 10). This influences the relation between fetch- and time-limited growth. For idealized fetch- and time-limited growth, the growth rate for the total energy E can be estimated as

$$\frac{\partial E}{\partial x} \approx \frac{S_{int}}{\overline{c}_x} \quad , \quad \frac{\partial E}{\partial t} \approx S_{int} \quad , \tag{9}$$

respectively, where S_{int} is the source term integrated over the spectrum, and where $\overline{c}_x$ is the spectral-mean offshore propagation velocity. For an identical integral source term and hence time-limited growth, but a broader spectrum, $\overline{c}_x$ will decrease and hence the fetch-limited growth rates will increase. For spectral widths as in Hasselmann et al. (1985), this might result in an artificial increase of fetch-limited growth rates of roughly 10%.

The input source term CB93 is also open for further improvements. First, the wind swell interaction needs to be reconsidered as described in section 3. Second, the effects of atmospheric stability have a potentially significant impact on wave growth (KC). The most logical way to model this is by adding a physical parameterization of the effects of stability on wave growth. Such parameterizations have not yet been suggested. Finally, all present input source terms assume quasi-linear superposition of input for separate spectral components. This might well not be realistic, considering that the input is related to the interaction between the turbulent (nonlinear) atmospheric boundary layer and the instantaneous water surface.

Finally, the dissipation source term is still largely unknown, and generally used as a closure term in the tuning of models. Unfortunately, not much

progress has been made in the development of physical parameterizations of spectral wave energy dissipation in the last two decades.

7 Conclusions

NCEP's new third-generation spectral ocean wave model (WAVEWATCH III version 1.15) has been applied successfully to produce wave forecasts in an operational environment over a global domain. After some initial modifications, including a retuning and a modification of the interactions between swell and the atmosphere, this model is shown to perform well. The rms model errors against altimeter and buoy data are 15% of the mean observed wave heights for most of the global domain. In general, the results are better than those of the well known WAM model, in particular with respect to swell prediction in the tropics, and extreme wave height prediction and wave height distributions in general. WAM shows better results in selected high-latitude regions. In spite of the good results, wave models are still open for further improvements. In particular, WAVEWATCH and WAM share a discrepancy between wave growth for short fetches and for the deep ocean. The present study furthermore shows that buoy data alone should not form the basis of a global model validation or comparison. Finally, the high-resolution validation with altimeter data indicates where higher-resolution model applications are necessary.

Acknowledgements. The author likes to thank B. Balasubramaniyan and V. Gerald for their help in archiving and managing the data, and D.B. Rao, D. Chalikov and J. Sienkiewicz for comments on early drafts of this paper.

Appendix A: Observation errors.

Buoy wave height observation errors relative to the altimeter observations have been investigated in detail by Monaldo (1988) (henceforth denoted as M88). M88 expects the buoy instrument error to be negligible, in particular in comparison to the sizeable sampling error inherent to the relatively short observation period. Donelan and Pierson (1983) estimate this sampling error

as 8% (or even more for extremely sharp spectra). Steele and Earle (1979) find an rms difference of wave heights of buoys less than 100m apart of 7%, which corresponds to a variability per buoy of 5%. This smaller variability is not necessarily inconsistent with the results of Donelan and Pierson, because the close proximity of the buoys is expected to result in correlated instead of uncorrelated samples. The buoy and altimeter data are on average separated by 15min and 50km. By analyzing time series of wave height from a buoy, M88 estimates that the effect of the time separation is about 0.1m or 5%. This estimate, however, includes the above sampling variability. This implies that the effect of a 15min separation is negligible compared to the sampling variability of the buoy observation. From along-track variability of altimeter wave heights, M88 estimates the rms difference of wave heights separated by 50km to be approximately 10%. This estimate includes the sampling variability of the altimeter observation, which is expected to be much smaller than the sampling variability of the buoy (and hence negligible) due to the implicit area averaging. Thus, the observation error of the buoy relative to the altimeter is dominated by the sampling variability and the effects of spatial separation, and is estimated as 12.5%.

From the collocated data, the maximum random error supported by the data ($\overline{\sigma}_{\max}$) can be estimated from the error-corrected and 'inverse' regressions as (Tolman, 1998a, Eqs (20) and (33))

$$\overline{\sigma}_{\max} = \sqrt{s_{oo} - \frac{s_{om}^2}{s_{mm}}}\,, \tag{A1}$$

where s_{oo} and s_{mm} are the variances of the buoy and altimeter wave height, and s_{om} is their covariance. For the present data set, this results in a maximum supported observation error of 10.3%, which is slightly smaller than the above estimate. This implies that the random altimeter error is negligible compared to the buoy observation error, implying that their functional relation is well described with the 'inverse' regression of the buoy wave height on the altimeter wave height.

Buoy observation error relative to wave model consists of the above sampling (instrument) error, collocation errors and mismatch errors. As the wave model results are interpolated to the buoy location, the collocation error becomes part of the scale error. The temporal scale error is included in the sampling error; the spatial scale error can be estimated from the above separated data as approximately 10%. Hence, the buoy observation error relative

to the wave model is estimated as approximately 12%.

Altimeter observation errors relative to the wave model similarly incorporate a small instrument error, and small scale errors because the averaged wave model considers the evolution of the wave field at scales similar to those of the wave model. The actual magnitude of the error is difficult to estimate, but might be expected to be in the 3-5% range. Such small errors are practically negligible in the error-corrected analysis techniques.

Appendix B: Event analysis.

The event data is derived from hourly buoy data and model output. Individual events are centered on local wave height maxima in a moving window of $\pm$12h. Within this window, data and model results are averaged over a period of $\pm$6h, considering model data only for those hours for which buoy data is available. Ideally, the event selection should be based on observations. However, dropouts and sampling variability of the data make the use of model results for event selection more convenient. Visual observation of the model results and observations indicate that this choice has little impact on the selected events and the averaged data. For events to be accepted it is furthermore required that (i) the maximum wave height should be at least 1.15 times the minimum wave height before and after it within the event window to eliminate model noise, and (ii) a minimum of six observations is used in the average to reduce sampling variability in the observation.

References

Booij, N. and L. H. Holthuijsen, 1987: Propagation of ocean waves in discrete spectral wave models. *J. Comput. Physics*, **68**, 307–326.

Chalikov, D. V. and M. Y. Belevich, 1993: One-dimensional theory of the wave boundary layer. *Bound. Layer Meteor.*, **63**, 65–96.

Cotton, P. D. and D. J. T. Carter, 1994: Cross calibration of TOPEX, ERS-1 and Geosat wave heights. *J. Geophys. Res.*, **99**, 25,025–25,033.

Derber, J. C., D. F. Parish and S. Lord, 1991: The new global operational analysis system at the National Meteorological Center. *Wea. Forecasting*, **6**, 538–547.

Donelan, M. and W. J. Pierson, 1983:: The sampling variability of estimates of spectra of wind-generated gravity waves. *J. Geophys. Res.*, **88**, 4381–4392.

Grumbine, R. W., 1996: Automated passive microwave sea ice concentration analysis at NCEP. Tech. Note 120, NOAA/NWS/NCEP/OMB, 13 pp.

Hasselmann, S., K. Hasselmann, J. H. Allender and T. P. Barnett, 1985: Computations and parameterizations of the nonlinear energy transfer in a gravity-wave spectrum, Part II: parameterizations of the nonlinear energy transfer for application in wave models. *J. Phys. Oceanogr.*, **15**, 1378–1391.

Kahma, K. K. and C. J. Calkoen, 1992: Reconciling discrepancies in the observed growth rates of wind waves. *J. Phys. Oceanogr.*, **22**, 1389–1405.

Kahma, K. K. and C. J. Calkoen, 1994: *Growth curve observations.* In Komen et al (1994), 174–182.

Kanamitsu, M., 1989: Description of the NMC global data assimilation and forecast system. *Wea. Forecasting*, **4**, 335–243.

Kanamitsu, M., J. C. Alpert, K. A. Campana, M. P. Caplan, D. G. Deaven, M. Iredell, B. Katz, H.-L. Pan, J. Sela and G. H. White, 1991: Recent changes implemented into the global forecast system at NMC. *Wea. Forecasting*, **6**, 425–435.

Komen, G. J., L. Cavaleri, M. Donelan, K. Hasselmann, S. Hasselmann and P. E. A. M. Janssen, 1994: *Dynamics and modelling of ocean waves.* Cambridge University Press, 532 pp.

Leonard, B. P., 1979: A stable and accurate convective modelling procedure based on quadratic upstream interpolation. *Comput. Methods Appl. Mech. Engng.*, **18**, 59–98.

Leonard, B. P., 1991: The ULTIMATE conservative difference scheme applied to unsteady one-dimensional advection. *Comput. Methods Appl. Mech. Engng.*, **88**, 17–74.

Monaldo, F., 1988:: Expected differences between buoy and radar altimeter estimates of wind speed and significant wave height and their implications on buoy-altimeter comparisons. *J. Geophys. Res.*, **93**, 2285–2302.

Steele, K. E. and M. D. Earle, 1979: The status of the data produced by NDBC wave data analyzer (WDA) systems. *Oceans '79*, pp. 1–10.

Tolman, H. L., 1991: A third-generation model for wind waves on slowly varying, unsteady and inhomogeneous depths and currents. *J. Phys. Oceanogr.*, **21**, 782–797.

Tolman, H. L., 1992: Effects of numerics on the physics in a third-generation wind-wave model. *J. Phys. Oceanogr.*, **22**, 1095–1111.

Tolman, H. L., 1995: On the selection of propagation schemes for a spectral wind wave model. Office Note 411, NWS/NCEP, 30 pp + figures.

Tolman, H. L., 1997: User manual and system documentation of WAVEWATCH III version 1.15. Tech. Note 151, NOAA/NWS/NCEP/OMB, 97 pp.

Tolman, H. L., 1998a: Effects of observation errors in linear regression and bin-average analyses. *Quart. J. Roy. Meteor. Soc.*, **124**, 897–917.

Tolman, H. L., 1998b: Validation of NCEP's ocean winds for the use in wind wave models. *The Global Atmosphere and Ocean System*, **6**, 243–268.

Tolman, H. L., B. Balasubramaniyan, L. D. Burroughs, D. V. Chalikov, Y. Y. Chao, H. S. Chen and V. M. Gerald, 2002: Development and implementation of wind generated ocean surface wave models at NCEP. *Wea. Forecasting*, **17**, 311–333.

Tolman, H. L. and N. Booij, 1998: Modeling wind waves using wavenumber-direction spectra and a variable wavenumber grid. *The Global Atmosphere and Ocean System*, **6**, 295–309.

Tolman, H. L. and D. V. Chalikov, 1994: The third generation wave prediction model with improved physics. in *AGU Ocean science meeting. EOS supplement*, p. 100.

Tolman, H. L. and D. V. Chalikov, 1996: Source terms in a third-generation wind-wave model. *J. Phys. Oceanogr.*, **26**, 2497–2518.

WAMDIG, 1988: The WAM model – a third generation ocean wave prediction model. *J. Phys. Oceanogr.*, **18**, 1775–1809.