NCEP NOTES

Development and Implementation of Wind-Generated Ocean Surface Wave Models at NCEP*

HENDRIK L. TOLMAN,+,# BHAVANI BALASUBRAMANIYAN,# LAWRENCE D. BURROUGHS, DMITRY V. CHALIKOV,+ YUNG Y. CHAO, HSUAN S. CHEN, AND VERA M. GERALD

NOAA/NCEP/Environmental Modeling Center, Camp Springs, Maryland

13 April 2001 and 30 October 2001

ABSTRACT

A brief historical overview of numerical wind wave forecast modeling efforts at the National Centers for Environmental Prediction (NCEP) is presented, followed by an in-depth discussion of the new operational National Oceanic and Atmospheric Administration (NOAA) ''WAVEWATCH III'' (NWW3) wave forecast system. This discussion mainly focuses on a parallel comparison of the new NWW3 system with the previously operational Wave Model (WAM) system, using extensive buoy and *European Remote Sensing Satellite-2* (*ERS-2*) altimeter data. The new system is shown to describe the variability of the wave height more realistically, with similar or smaller random errors and generally better correlation coefficients and regression slopes than WAM. NWW3 outperforms WAM in the Tropics and in the Southern Hemisphere, and they both show fairly similar behavior at northern high latitudes. Dissemination of NWW3 products, and plans for its further development, are briefly discussed.

1. Introduction

Wind waves generated and propagated on the ocean surface potentially represent a serious hazard to life and property in various maritime and coastal activities. Hence, it is necessary to develop the capability to forecast wave conditions over global and regional ocean domains to minimize loss of life and property.

The Ocean Modeling Branch (OMB) of the National Centers for Environmental Prediction (NCEP) and its predecessors have a long history of providing the marine forecasters of the National Weather Service (NWS) with operational numerical wave forecast guidance. This note will start with a brief history of numerical wave modeling in general, and at NCEP in particular. The remainder of the paper will concentrate on recent developments at NCEP.

The layout of the present paper is as follows. In section 2, a brief review is given of previous wave model development, leading up to the justification for developing a new model. In section 3 a brief description of the new model is given. Section 4 provides some details of the global and regional applications of this model that have been developed for use by NCEP for operational purposes. In section 5 validation results are presented. The model is compared with the operational model it replaces, the Wave Model (WAM) system, concentrating on the global model. Furthermore, ongoing validation results for the global and regional models are presented. Products produced by the new suite of operational wave models and their dissemination methods are briefly discussed in section 6. A summary and conclusions are presented in section 7.

2. Wave model development

The first operational wave forecasts were made in preparation for the Normandy invasion of World War II in 1944, culminating in the work of Sverdrup and Munk (1946, 1947). The first computer-generated wave forecasts for the NWS were made in July 1956 (Hubert 1957). These early models produced a single wave height and period at each grid point, using a direct relation between the local wind speed and the wave height and period. Originally, only wind seas generated by local and recent wind speeds were calculated. The wind sea was described by its significant wave height H_s,

* Ocean Modeling Branch Contribution Number 208.
+ UCAR Project Scientist.
# Current affiliation: SAIC/GSO, Beltsville, Maryland.

Corresponding author address: Dr. Hendrik L. Tolman, Ocean Modeling Branch, Environmental Modeling Center, NCEP, NOAA, 5200 Auth Rd., Rm. 209, Camp Springs, MD 20746.
E-mail: hendrik.tolman@noaa.gov

which is defined as the average height of the 33% of the waves that are the highest, and by the corresponding significant wave period. Later a single (dominant) significant wave height for swell and a significant wave height for the combined sea state were added (e.g., Pore and Richardson 1968). This prediction approach is generally identified as the ''representative wave'' approach. Such relatively simple models remained operational at the NWS until 1985.

At the time when these models were implemented at the NWS, it was understood that the representative wave approach does not do justice to the inherent complexity of the wave field on the ocean surface. For this reason, forecast skills of such models were limited. It had become clear that the sea state consists of a random superposition of waves of different wavelengths, propagating in different directions, and that the only logical way to describe it is through a statistical description of the inherent spectrum. Hence, a more complete description of the sea state, and the potential for much better forecasts, can be realized by predicting the so-called energy spectrum $F(f, \theta)$. This spectrum describes the distribution of wave energy over wave frequency f and wave propagation direction θ. From such a spectrum, the significant wave height can be calculated as

$$H_s = 4\left[\int\int F(f, \theta)\, df\, d\theta\right]^{1/2}. \qquad (1)$$

The development of the wave energy spectrum in space and time is governed by the basic transport or energy balance equation

$$\frac{DF}{Dt} = S = S_{\text{in}} + S_{\text{nl}} + S_{\text{ds}} + \cdots. \qquad (2)$$

Similar equations can be derived for spectra based on wavenumber and direction $F(k, \theta)$ or on the wavenumber vector $F(\mathbf{k})$. The left side of this equation describes changes in the local spectrum due to (conservative and linear) propagation of the wave energy of individual spectral components with their group velocity (and sometimes also due to the effects of externally prescribed mean currents). The right side represents a combination of nonconservative sources and sinks of wave energy, such as the wind input (S_{in}); dissipation due to wave breaking (S_{ds}); other (mostly shallow water) processes, denoted here with the ellipsis; and a term S_{nl} representing the transfer of energy due to nonlinear interactions between the spectral wave components. The latter term only exchanges energy between spectral components, but does not change the total wave energy.

Numerical wave models in which the spectrum F is discretized and Eq. (2) is solved are identified as spectral wave models. After the pioneering work of Gelci et al. (1956, 1957), many such models have been developed. A review and classification of spectral models can be found in The Sea Wave Modeling Project (SWAMP) Group (1985). The classification of different spectral models is largely based on the treatment of the nonlinear interaction term (S_{nl}) in Eq. (2). In the so-called first-generation models, S_{nl} is not modeled explicitly, so that all spectral components evolve independently. Dissipation for wind seas is generally modeled as an on–off mechanism, limiting the spectral evolution to some predescribed spectral shape. In second-generation models, simple approximations for nonlinear interactions are introduced, either treating the entire wind sea part of the spectrum using empirical growth relations and idealized spectral shapes (so-called hybrid models), or by modeling S_{nl} based on results for simplified spectral shapes (so-called discrete models).

The first spectral model used operationally at NCEP (then known as the National Meteorological Center, NMC) was based on the second-generation ''SAIL'' model (Cardone and Ross 1977). The NMC version of the SAIL model, named the National Oceanic and Atmospheric Administration (NOAA) Operational Wave Model (NOW), became operational in 1985 (Chin 1986). It represented a significant improvement over the previous operational models in terms of the general quality of the products, but also because it provided forecasters with much more detailed wave field information (i.e., spectra) at selected model output points (Chin and Burroughs 1988; Esteva and Kidwell 1990). The global NOW model was augmented with similar second-generation regional models for the Gulf of Mexico in 1988 (Chao 1991) and the Gulf of Alaska in 1994 (Chao 1995).

With the comprehensive intercomparison of first- and second-generation wave models in the SWAMP study, it became apparent that, ''All present second-generation models suffer from limitations in the parameterization of the nonlinear energy transfer, S_{nl}'' (SWAMP Group 1985, p. 136, item 4). Particularly, second-generation models such as NOW give poor results in rapidly changing wind and wave conditions. In NOW, such shortcomings resulted in generally suppressed initial wave growth, as well as large positive coastal biases.

Within the international community, the shortcomings of second-generation models led to the establishment of the WAM group, which later became the World Meteorological Organization's Scientific Committee on Ocean Research (WMO SCOR) Work Group 83 (see preface to Komen et al. 1994). The purpose of this group was to develop an economically feasible model that would integrate Eq. (2) based on first principles, that is, by directly parameterizing all sources S without a priori assumptions on spectral shapes. Such a model is denoted as a third-generation wave model. The WAM group succeeded in developing a simple procedure to estimate S_{nl} in an economical manner. With the development of this procedure, referred to as the discrete interaction approximation (DIA; Hasselmann et al. 1985), the WAM (WAMDI Group 1988; Komen et al.

1994) became the first operationally feasible third-generation wave model.

In 1994, an implementation of WAM cycle 4 replaced the global NOW model as the operational model at NCEP and resulted in a major improvement in the quality of the numerical forecasts of significant wave heights (Chen 1995). Subsequently, a regional WAM was developed for the U.S. east coast and the Gulf of Mexico (Chao 1997). This model replaced the operational second-generation Gulf of Mexico model. The second-generation Gulf of Alaska model at that time was left unchanged.

Although the WAM was a major step forward in wave modeling, it became clear from evaluations carried out at NCEP that this model also left room for further improvement. WAM uses first-order numerics in its propagation terms, and this adversely influences swell propagation. Source terms are integrated with large fixed time steps, which is expected to result in spectral shape errors in rapidly changing wave conditions. Furthermore, extreme wave conditions were systematically underestimated away from storm tracks in areas such as Hawaii (as will be illustrated below). Although this might be due to numerical issues, it is more likely an artifact of the physical parameterizations in WAM. Similarly, extremely low wave conditions appear to be systematically overestimated.

The underlying design of the WAM dates back to the early days of supercomputers and was tailored to run efficiently on early vector computers. This design now hampers further development of WAM, in such a way that some of the envisioned improvements simply could not be incorporated in the structure of WAM. For this reason NCEP decided to develop a new model (''WAVEWATCH III''). Development of this model started in 1993, and the model was tested and validated comprehensively. The new model, referred to as NOAA WAVEWATCH III, or NWW3, formally became operational at NCEP on 9 March 2000 (Chen et al. 1999) for global application. At the same time, two new regional models were implemented to replace previous East Coast and Alaska regional models. These new regional models are high-resolution versions of the global NWW3 model (Chao et al. 1999a,b).

WAVEWATCH III differs from WAM in several important ways. The basic model design is focused on model transparency and plug compatibility for both numerical and physical approaches. Unlike WAM, WAVEWATCH III is based on the fully unsteady spectral action density equation, in order to take into account large-scale wave–current interactions (see section 3). This was done with an eye on the future as presently we do not have sufficiently accurate current fields to consider wave–current interactions in operational wave forecasts. WAVEWATCH III furthermore uses new physics parameterizations for most source terms and more accurate numerical integration schemes. A brief description of this model is given in the following section.

It should be noted that the actual wave model only constitutes part of a wave forecast system. The second part is the atmospheric model providing wind fields used to force the wave model. The atmospheric model obviously is an integral and important part of the wave forecast system. Without good winds, even the best wave model will have no chance to provide good wave forecasts. Like the wave models discussed here, wind forecasts have also gone through major developments in the past decade. Such development have also had a major impact on the quality of the wave forecasts. A description of recent developments at NCEP can be found in Kanamitsu (1989), Derber et al. (1991), Kanamitsu et al. (1991), and Caplan et al. (1997). A detailed description of the atmospheric models is out of the scope of this paper.

3. WAVEWATCH III

In this section a brief description of the generic WAVEWATCH III model is given. For details, see Tolman (1999). The model solves the linear balance equation for the spectral wave action density A in terms of wavenumber k and wave direction θ, as a slowly varying function of space $\mathbf{x}$ and time t,

$$\frac{DA(k, \theta; \mathbf{x}, t)}{Dt} = S(k, \theta; \mathbf{x}, t), \tag{3}$$

which is closely related to Eq. (2). The implicit assumption in this equation is that the space scale and timescale of individual waves are much smaller than the corresponding scales of change of the spectrum and of the mean depth and current. The action density spectrum A is related to the energy density spectrum F as $A = F/\sigma$, where σ is the intrinsic wave frequency. Similarly, $S = S/\sigma$. The intrinsic frequency is related to the wavenumber through the dispersion relation

$$\sigma^2 = gk \tanh kd, \tag{4}$$

where d is the mean water depth. The intrinsic or relative frequency is related to the absolute frequency ω (as observed in a fixed frame of reference) through the Doppler equation:

$$\omega = \sigma + \mathbf{k} \cdot \mathbf{U}, \tag{5}$$

where $\mathbf{U}$ is the mean current velocity vector. In all applications discussed in the present paper, currents are ignored ($\mathbf{U} \equiv 0$). In this case $\omega = \sigma$, and Eq. (3) reduces to the form of Eq. (2).

Virtually every spectral wave model solves an equation similar to Eqs. (3) or (2) using a fractional step method, where parts of the equation are solved consecutively. WAVEWATCH III consecutively solves equations for spatial propagation, intraspectral propagation, and source terms. In the model k, θ, and $\mathbf{x}$ space are discretized, using a spatially varying discretization of k as suggested by Tolman and Booij (1998). This spatially varying wavenumber grid corresponds to a spatially in-

variant σ grid. For **x** space, a latitude–longitude grid is used (see Tolman 1999, p. 8). The equations are solved by marching forward in time with an overall global model time step Δt_g. In individual fractional steps reduced time steps are used as discussed below.

The spatial propagation equation is solved using the third-order-accurate ''ULTIMATE QUICKEST'' scheme of Leonard (1979, 1991). This scheme is sufficiently free of numerical diffusion to result in the so-called garden sprinkler effect. This implies that the discrete description of the spectrum results in a disintegration of a continuous swell field into individual discrete swell fields. To avoid such aphysical behavior, the modified spatial propagation equations of Booij and Holthuijsen (1987) have been used (see also Tolman 1995). For each discrete model frequency σ, a maximum propagation time step $\Delta t_{p,\max}$ is defined. To satisfy Courant–Friedrichs–Lewy (CFL) criteria in an economical way, $\Delta t_{p,\max}$ scales linearly with σ. If $\Delta t_{p,\max} < \Delta t_g$, spatial propagation is performed in several substeps.

For intraspectral propagation (i.e., local shifts of energy and action in k and θ spaces), the ULTIMATE QUICKEST scheme is also used. In this part of the model, numerical stability can become an issue for poorly resolved transitions from deep to shallow water, and for refraction in general in extremely shallow water. To avoid the need for extremely small time steps in this context, refraction velocities are filtered (Tolman 1999, p. 35). Furthermore, the model allows for an intraspectral propagation time step Δt_i to be smaller than the global time step Δt_g, as in spatial propagation.

The source terms considered in WAVEWATCH III are wind input, nonlinear interactions, whitecapping dissipation, and bottom friction. Input and dissipation are modeled following Tolman and Chalikov (1996), with several modifications. Swell attenuation is reduced (Tolman 1999, p. 15), and growth rates are retuned and corrected for effects of atmospheric stability (Tolman 1999, p. 18). Nonlinear interactions are modeled using the discrete interaction approximation of Hasselmann et al. (1985). Bottom friction is modeled using the Joint North Sea Wave Project (JONSWAP) parameterization (Hasselmann et al. 1973). The numerical integration of the source terms uses a modified version of the semi-implicit scheme of WAM (WAMDI Group 1988). In this scheme, the integration time step is dynamically adjusted for each spatial grid point depending on the rate of spectral change due to the source terms (Tolman 1992). This scheme allows for more accurate integration of source terms in conditions of rapid change, and more economical integration otherwise.

WAVEWATCH III uses as input wind fields and air–sea temperature differences, which can be provided at arbitrary (and irregular) intervals. Wind speed and direction are linearly interpolated at the time interval Δt_g. The model can furthermore ingest polar ice concentrations. If the ice concentration becomes larger than a cutoff value (typically 0.33 or 0.5), the corresponding spatial grid points are taken out of the calculations, as if covered by land. The model can finally ingest unsteady currents and water levels (not considered in present applications).

4. Global and regional models for NCEP operations

The initial tuning of the model (presented elsewhere) has been done by using a global model to provide wave hindcasts and validate them using buoy and altimeter data. This wave forecast system with its driving wind and temperature fields was originally named the NOAA Experimental Wave model (NEW), and has recently become the operational NOAA WAVEWATCH III model. For simplicity, this forecast system will be exclusively denoted as NWW3 here. The NWW3 wave model has a spatial resolution of $1^\circ \times 1.25^\circ$ in latitude and longitude, with the grid ranging from 78°N to 78°S. This spatial grid consists of 288×157 points, of which 30 030 are sea points. The spectrum is discretized with 25 frequencies, ranging from 0.041 to 0.42 Hz with a 10% increment, and 24 directions with a 15° increment. The global time step is $\Delta t_g = 3600$ s, the maximum propagation time step for the lowest frequency σ_1 is $\Delta t_{p,\max}(\sigma_1) = 1300$ s, and the minimum allowed time step in the source term integration is set to 300 s. Parameter settings in the physics parameterizations are generally the default WAVEWATCH III settings, and additional details can be found online at the NWW3 Web site (http://polar.ncep.noaa.gov/waves/implementations.html) and in Chen et al. (1999).

The NWW3 wave forecast system uses winds and temperatures 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; Caplan et al. 1997), which are available at 3-h intervals. Initially, the resolution of these models was T126 with 28 levels. On 24 January 1999 the resolution was increased to T170 with 42 levels. Ice concentrations are obtained from NCEP's automated passive microwave sea ice concentration analysis (Grumbine 1996) and are updated daily.

Initial validation of the wind fields for the 1994/95 Northern Hemisphere winter identified systematic errors (Tolman 1998b). To avoid the need for retuning the wave model when such errors change, a systematic error correction for the driving winds was introduced. This error correction obviously needs to be monitored continuously, and updated as needed. The correction suggested by Tolman (1998b) consisted of a coastal error correction up to 150 km offshore, a deep-ocean correction beyond 300 km offshore, and a smooth blending in between.

Error corrections as initially suggested by Tolman (1998b) were substantial, typically

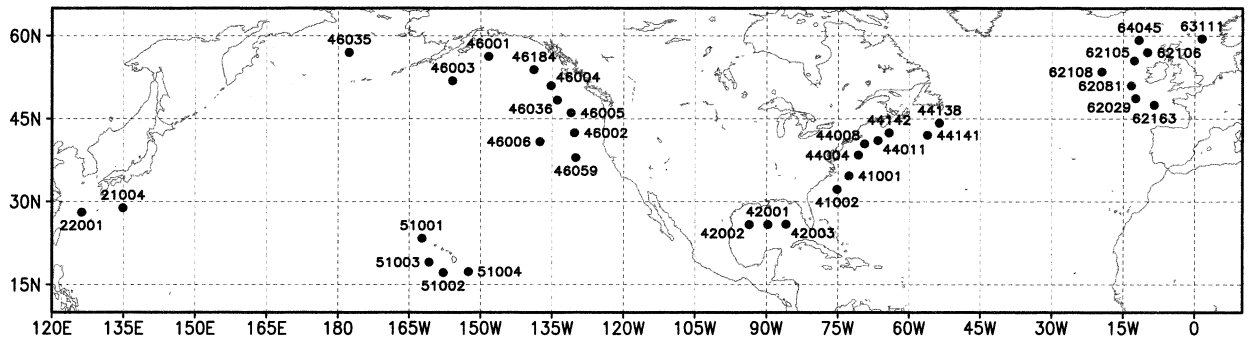


FIG. 1. Location of validation points used for global NWW3 model with corresponding WMO buoy identifiers.

$$U_c = -1.5 \text{ m s}^{-1} + 1.10U_o, \quad (6)$$

where U_c and U_o are the corrected and original wind speeds, respectively. By early 1997 wind speed biases were greatly reduced, and initially the following error corrections were used in NWW3:

$$U_c = -0.3 \text{ m s}^{-1} + U_o \quad \text{(coastal) and} \quad (7)$$

$$U_c = -1.0 \text{ m s}^{-1} + 1.05U_o \text{ (deep ocean).} \quad (8)$$

Subsequent improvements of GDAS and MRF in June 1998 again substantially reduced wind speed biases. On 15 June 1998, the wind speed bias correction was removed from NWW3.

Parallel runs of this wave forecast system to compare its performance with that of the then-operational WAM-based wave forecast system were initiated on 29 January 1997. The model was run twice daily at the 0000 and 1200 UTC model run cycles, and consisted of a 12-h hindcast and a 72-h forecast. The first data were made available to the public in January 1998. After extensive validations of the parallel tests as well as evaluations by the NWS forecasters in the field and at the Marine Prediction Center (MPC), the global and two regional versions of NWW3 (described below) were made operational at NCEP on 9 March 2000. On 31 May 2000, their forecasts were extended out to 126 h.

In addition to the global wave model, two regional wave models have been constructed; the Alaskan Waters (AKW) model using 0.25° × 0.50° latitude–longitude resolution, and the Western North Atlantic (WNA) model using 0.25° × 0.25° latitude–longitude resolution. Both models obtain hourly boundary data from the global NWW3 model and cover identical hindcast and forecast ranges. Wind and ice data are also obtained from the same sources. A more detailed description of these models can be found at the NWW3 Web site or in Chao et al. (1999a,b).

5. Validation

a. Data

The present model validations use fixed buoy observations and *European Remote Sensing Satellite-2* (*ERS-2*) altimeter data. Figure 1 shows the locations of the buoys that have been used to validate the global NWW3 model. All buoys are located in the Northern Hemisphere, and even there do not cover the deep ocean well. Buoy data are nevertheless important because they provide continuous time series, as well as for historical reasons. Where possible, the buoy data have been obtained from the National Data Buoy Center's online archive of quality controlled data (http://ndbc.noaa.gov); otherwise, they were obtained from NCEP's real-time operational data flow. All buoy data have furthermore been manually quality controlled at NCEP/OMB. Both wind and significant wave height data from buoys have been used. All wind data were converted to 10-m heights.

Fast-delivery *ERS-2* data are also retrieved from the operational data flow at NCEP. As in Tolman (1998b), these data are averaged along the track in 10-s intervals to result in observations with scales comparable to those of the wave models. Furthermore, fast-delivery wave height data are known to include systematic errors (e.g., Cotton and Carter 1994), which can be removed with a simple linear correction. Here, the corrected altimeter wave height $H_{a,c}$ is calculated from the fast-delivery wave height $H_{a,\text{FD}}$ as

$$H_{a,c} = 0.03 + 1.09H_{a,\text{FD}}. \quad (9)$$

For low wave heights a nonlinear correction is needed because the altimeter is unable to produce wave heights below approximately 0.5 m. An ad hoc quadratic correction has been constructed, which results in $H_{a,c} = 0$ for a retrieved wave height of $H_{a,\text{FD}} = 0.6$ m, and which fits with constant value and derivative to the above linear correction at a retrieved wave height of $H_{a,\text{FD}} = 2.0$ m. These corrections are based on 1725 collocations with buoy data for the buoys in Fig. 1 from March 1997 through February 1998. Resulting averaged and corrected collocations are presented in Fig. 2. Note that the resulting scatter of about 10% can be attributed to the combined sampling error in the buoy observations and the collocation error. Hence, the averaged and bias-corrected altimeter data are arguably of better quality than the buoy data. These data have been collocated with

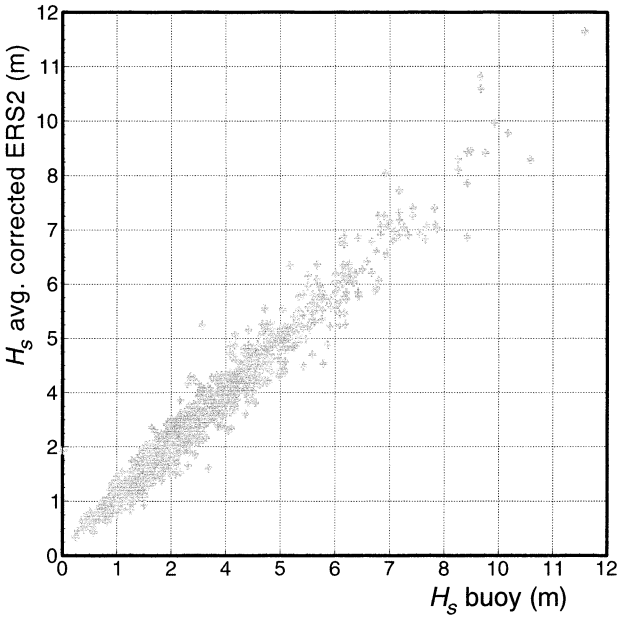


FIG. 2. Averaged and corrected *ERS-2* altimeter wave height retrievals as a function of buoy observations: Mar 1997–Feb 1998, 1725 data pairs.

model results using trilinear interpolation from hourly model wave height fields. Wind retrievals from *ERS-2* have not been considered, because these retrievals are systematically contaminated by the background wave fields (e.g., Tolman 1998b).

Note that the Ocean Topography Experiment (TOPEX) altimeter provides an additional source of high-quality global wave data. Because these data have only recently become available operationally at NCEP, they have not been used in the present study.

b. Parallel comparisons with previous models

The parallel comparison with the previous operational wave forecast system at NCEP mainly focused on the global model. The previous operational global model was a version of WAM cycle 4 (see section 1 and Chen 1995). This forecast system was not only using a different wave model, but it also had a significantly coarser resolution of 2.5° × 2.5° in latitude and longitude, and 12 spectral directions with a directional increment of 30°. Furthermore, the aerial extent of WAM was truncated at 67.5°S, and WAM did not take into account ice coverage for the period considered here. The operational WAM and the new NWW3 model used the same source for wind forcing (GDAS and MRF). Interpolation and grid discretization result in small but systematic differences. The main difference in forcing is the error correction given by Eqs. (7) and (8), which was not used in the old wave forecast system. Below we will for simplicity denote the old forecast system as WAM. Because we compare forecasts systems with different resolutions and forcing, differences between NWW3 and WAM should not be interpreted as differences between the underlying wave models alone (WAVEWATCH III and WAM).

The parallel comparison and validation period started 12 January 1998 and ended 30 June 1998. The comparison is somewhat complicated because, starting 9 February, all validation data[1] were assimilated into the WAM. For the WAM hindcast, the buoy and altimeter data are therefore no longer independent validation data starting at this date. For WAM forecasts, however, this is not the case. Extensive validation data, including time series for each buoy and each month, are available from the parallel comparison Web site (http://polar.ncep.noaa.gov/waves/NEW-WAM.html). Here, we obviously have to be much more concise.

To illustrate different behaviors of the systems, we will first consider selected time series for selected buoy locations in Fig. 3. Hourly hindcast time series are considered because they represent the optimal performance of the forecast systems. Furthermore, contiguous time series are available for the hindcast but not for the forecast. Only time series for January are considered, to assure that the data used for validation constitute a truly independent set for both models (as mentioned earlier, buoy and altimeter data have been used for assimilation in WAM hindcasts starting February).

The most pronounced differences between NWW3 and WAM were found near Hawaii. As an example, Fig. 3a shows a time series for buoy 51001. In this area, NWW3 (solid line) generally follows the observations (○) well and displays similar variability. WAM (dashed line), on the other hand, displays a much smoother wave height evolution, systemically missing the highest and lowest wave conditions. Very encouraging is the capability of NWW3 to realistically capture rapidly changing wave conditions early on 28 January.

Similar systematic differences were found in the Gulf of Mexico. As an example, Fig. 3b shows a time series for buoy 42001. Here, NWW3 (solid line) captures the minima, maxima, and variability in the observations (○) much better than WAM (dashed line). Note that particularly in this relatively small enclosed basin, the coarse spatial resolution of WAM is expected to be detrimental. Note furthermore that on 14 and 15 January both models completely miss two well-defined wave events. The timescale of these events indicates that the temporal scales of the generating wind events (and therefore probably also the spatial scales) cannot be resolved properly by the driving wind fields. Obviously, without these events occurring in the driving wind fields, both wave models fail to produce the corresponding wave fields.

In other regions differences between the two forecast systems are less systematic, with some buoy locations distinctly showing the above behavior, and other loca-

[1] Altimeter data averaged and error corrected as described above.

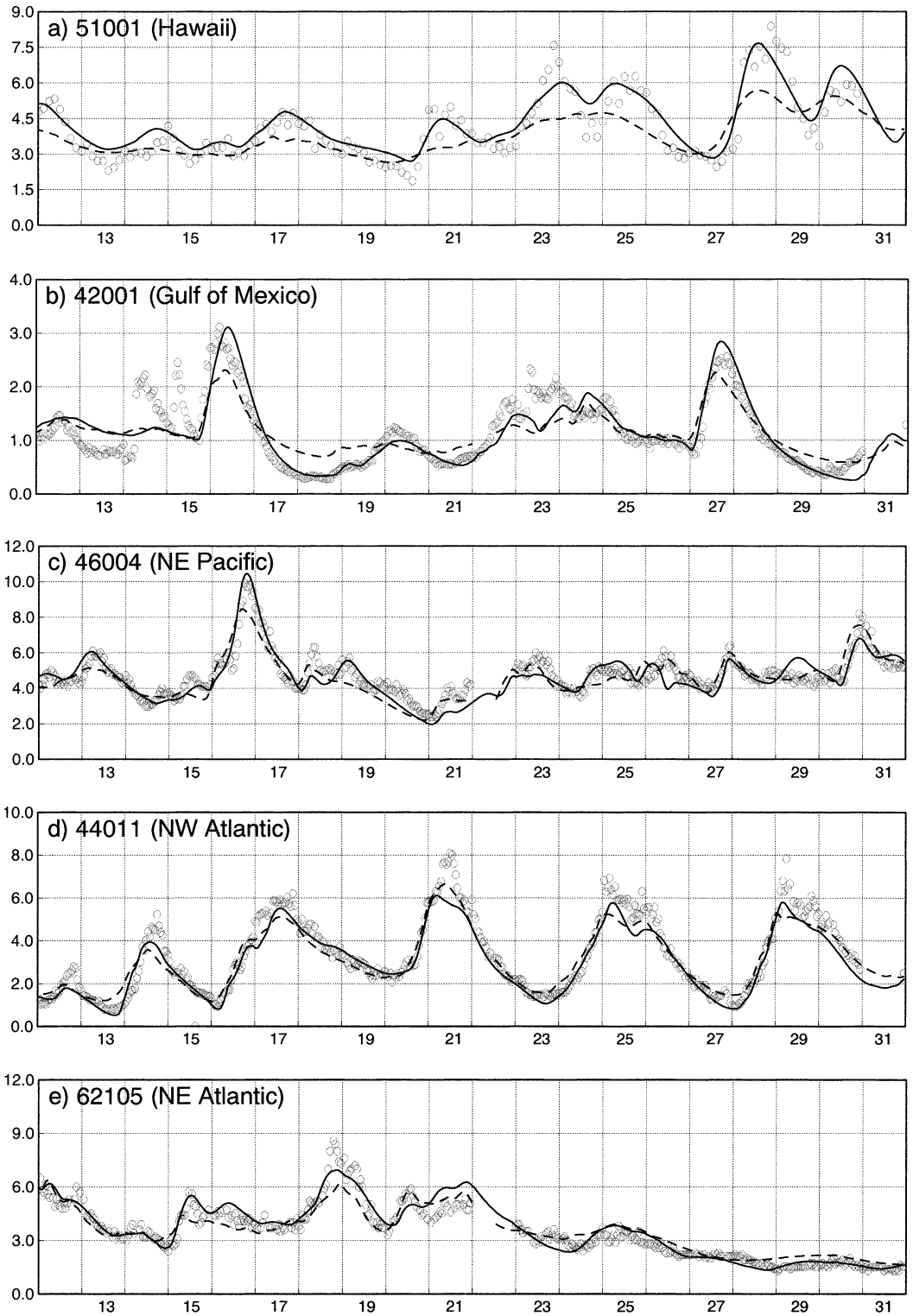


FIG. 3. Hindcast time series of the significant wave height H_s (m) for Jan 1998: solid line, NWW3; dashed line, WAM; ○, observations. Vertical grid lines at 0000 UTC each day. Buoy locations as identified in panels [(a) 51001, (b) 42001, (c) 46004, (d) 44011, (e) 62105].

tions showing much more comparable model behavior. Examples of more comparable behavior are presented in Figs. 3c–e for locations 46004 (NE Pacific), 44011 (NW Atlantic), and 62105 (NE Atlantic), respectively. Although NWW3 (solid lines) does not show clearly better behavior than WAM (dashed lines) for these locations, the systematically more realistic variability of the wave heights produced by NWW3 is clear at all locations.

Because of the large volume of data, time series cannot present a comprehensive analysis of the buoy validation dataset. For a more comprehensive validation, joint probability density functions (PDF) of model results and buoy observations and bulk model statistics are presented in Figs. 4–6. Bulk statistics considered are linear regression lines (in particular, their slope), biases, standard deviations (labeled std in figures), root-mean-square (rms) errors, scatter indices (SI), and correlation coefficients (c.c.; see the appendix for details).

Before one looks further at wave heights, wind speeds deserve some attention, to assure that systematic wave model behavior cannot be attributed to systematic properties of driving wind fields. The top panels of Fig. 4 show the PDFs and statistics for wind speeds as used by both hindcasts (i.e., GDAS winds) against buoy data for the entire parallel comparison period, for NWW3 (left panel) and WAM (right panel). These figures show that the winds for both systems are essentially unbiased, and they do not show systematic biases for lower or higher wind speeds. The regression slope of the WAM winds is nearly 4% lower than the slope for NWW3. This is a direct effect of the error correction [Eq. (8)] in NWW3. The difference in slope is expected to result in slope differences for wave height regressions of up to 7% due to the quadratic scaling of wave height with wind speed. Random wind errors (standard deviation) for WAM are slightly larger. This could be expected, as WAM data are interpolated from a coarser grid. Note that the error correction in Eq. (8) counteracts this effect, as it increases the random error by 10%.

The middle panels of Fig. 4 show the statistics for the hindcast significant wave height for the entire validation period. Here, WAM shows a significantly tighter fit to the data than NWW3. This was expected, however, because most of these validation data were assimilated into WAM, and hence it no longer constitutes an independent dataset. An independent validation can be performed only with the data obtained before 9 February 1998, that is, before assimilation started in the operational WAM. The corresponding wave height statistics are presented in the bottom panels of Fig. 4.

For NWW3, the statistics for the entire period (middle-left panel of Fig. 4) and the first part of the period (bottom-left panel) are very similar, indicating that the first period is representative for the entire validation period. For WAM (middle- and bottom-right panels), they are very different, indicating that the hindcast results for WAM including data assimilation (i.e., the WAM ''analysis'') indeed should not be used in the model comparison.

A detailed comparison of both models (bottom panels of Fig. 4) shows slightly smaller standard deviations and rms errors for NWW3, as well as a better correlation coefficient. NWW3, however, has a larger bias and appears in general too energetic against buoy data, with a regression slope that is about 11% too large. Conversely, WAM is not sufficiently energetic, with regression slopes about 7% too small. Note that the differences between regression slopes are much larger than expected based on the wind statistics in the upper panels of this figure, suggesting that the underlying wave models are responsible for a significant part of these differences.

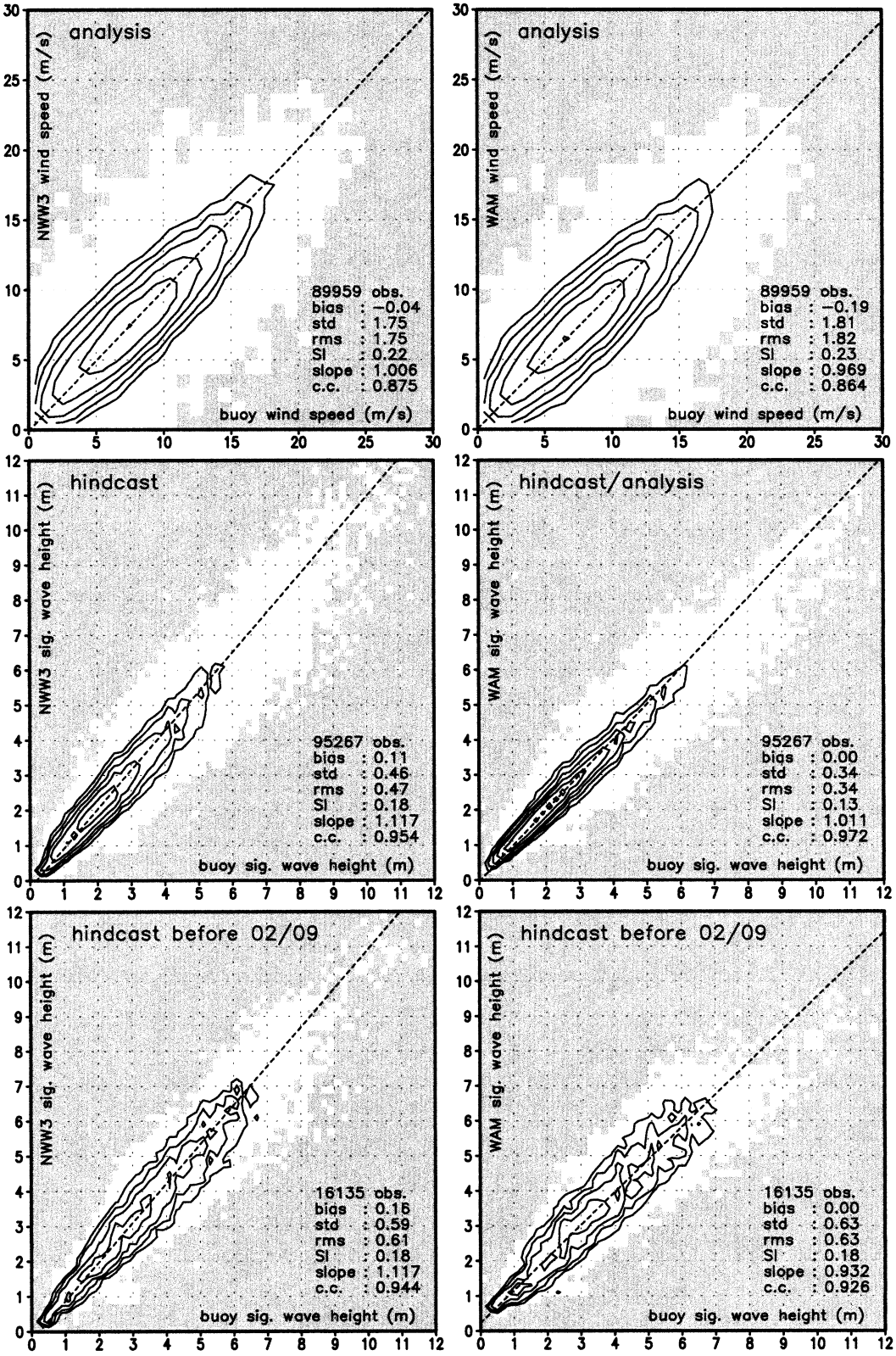
Figure 5 shows wave height statistics for NWW3 (left panels) and WAM (right panels) for the 24-, 48-, and 72-h forecast times (top, middle, and bottom panels, respectively). Unlike hindcast runs, in which validation data had been assimilated in WAM, forecasts by definition are free of any influence of validation data. Consequently, all data can be considered independent validation data. The statistics in this figure, therefore, cover the entire 6-month comparison period.

Both models show a systematic increase of the bias with forecast time. This is due to a corresponding increase in the wind speed bias (figures not presented here). Both models furthermore show a systematic increase of the standard deviations and rms errors with forecast time, as well as a systematic reduction of the correlation coefficient. This represents the effects of an increased random error in the wind fields on the wave forecasts. The regression slopes for both models show only minor changes with forecast time.

Note that the error growth of NWW3 [left panels of Fig. 5, approximately 0.13 m $(24\ \text{h})^{-1}$] is much larger than the error growth of WAM [right panels, approximately 0.08 m $(24\ \text{h})^{-1}$]. This is due to the fact that NWW3 is a more energetic model than WAM, and it hence reacts more vigorously to wind speed errors than

→

FIG. 4. Joint PDF of buoy observations and NWW3 (left panels) or WAM (right panels) for hindcast/analysis. Upper panels show 10-m wind speeds; $\Delta u = 1\ \text{m s}^{-1}$, with lowest contour level at $0.001\ \text{m}^{-2}\ \text{s}^{2}$. Contours increment by a factor of 2. Middle panels show corresponding wave height distributions for entire model comparison period (after 9 Feb, WAM used validation data in assimilation). Here $\Delta H_s = 0.2$ m, with lowest contour level at $0.02\ \text{m}^{-2}$. Bottom panels, as in middle panels, show data before start of assimilation in WAM only. Gray areas are devoid of data. See appendix for statistical parameters presented.


analysis
NWW3 wind speed (m/s)
89959 obs.
bias : −0.04
std : 1.75
rms : 1.75
SI : 0.22
slope : 1.006
c.c. : 0.875
buoy wind speed (m/s)
analysis
WAM wind speed (m/s)
89959 obs.
bias : −0.19
std : 1.81
rms : 1.82
SI : 0.23
slope : 0.969
c.c. : 0.864
buoy wind speed (m/s)
hindcast
NWW3 sig. wave height (m)
95267 obs.
bias : 0.11
std : 0.46
rms : 0.47
SI : 0.18
slope : 1.117
c.c. : 0.954
buoy sig. wave height (m)
hindcast/analysis
WAM sig. wave height (m)
95267 obs.
bias : 0.00
std : 0.34
rms : 0.34
SI : 0.13
slope : 1.011
c.c. : 0.972
buoy sig. wave height (m)
hindcast before 02/09
NWW3 sig. wave height (m)
16135 obs.
bias : 0.16
std : 0.59
rms : 0.61
SI : 0.18
slope : 1.117
c.c. : 0.944
buoy sig. wave height (m)
hindcast before 02/09
WAM sig. wave height (m)
16135 obs.
bias : 0.00
std : 0.63
rms : 0.63
SI : 0.18
slope : 0.932
c.c. : 0.926
buoy sig. wave height (m)

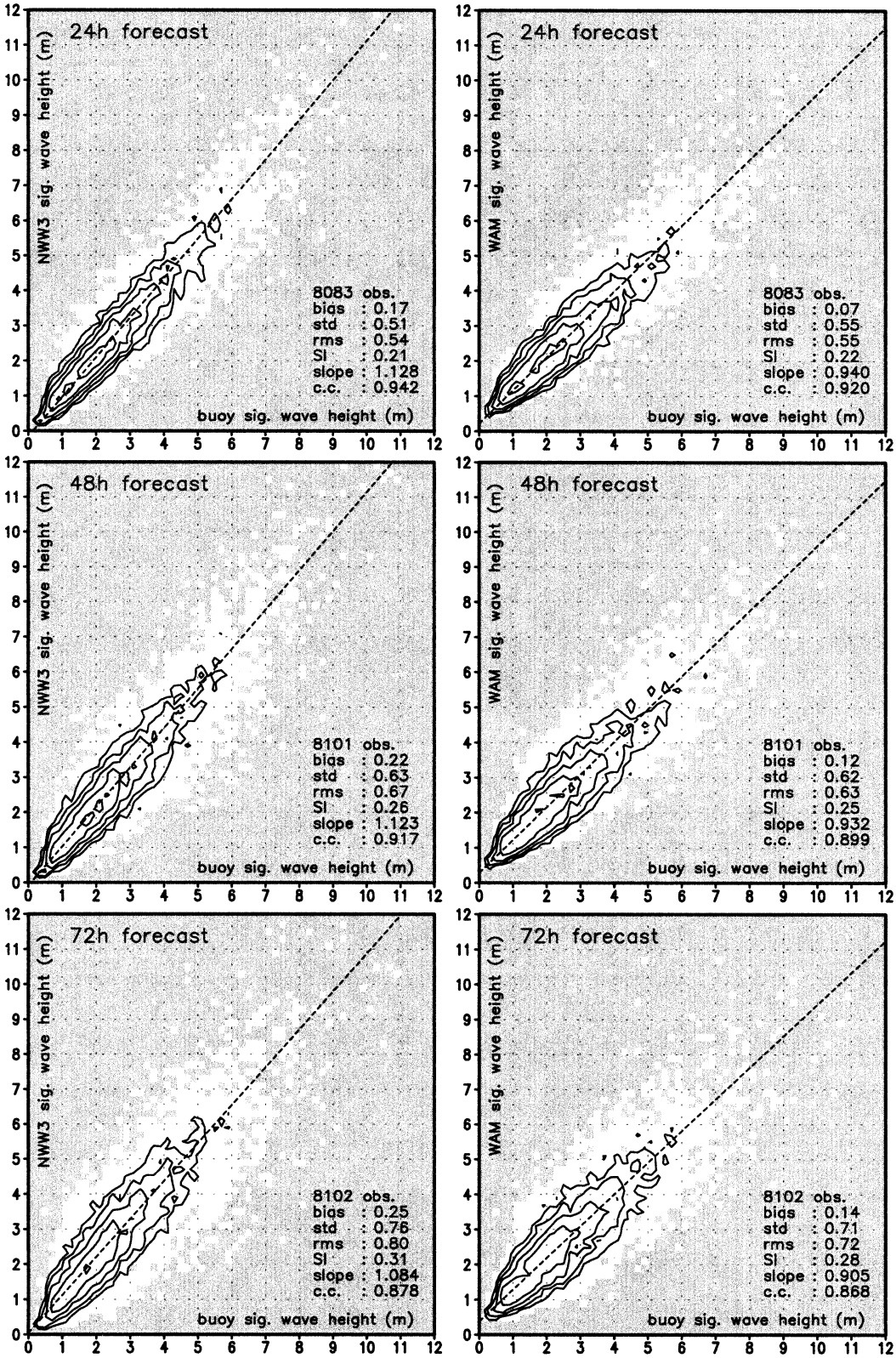


FIG. 5. Joint significant wave height PDF of buoy observations and NWW3 (left panels) or WAM (right panels) for 24-, 48-, and 72-h forecasts (upper, middle, and lower panels, respectively). Here $\Delta H_s = 0.2$ m, with lowest contour level at 0.02 m^{-2}; contours increment by a factor of 2. Gray areas are devoid of data. See appendix for statistical parameters presented.

the less energetic WAM does. Time series as presented in Fig. 3 suggest that NWW3 produces a more realistic wave height variability than WAM. This would suggest that NWW3 error growth rates are also more realistic than those of WAM, and that the faster growth rates should not be interpreted as poorer model behavior. This argumentation, however, is not fully supported by the bulk statistics presented in Figs. 4 and 5. The latter figures suggest that NWW3 is too energetic, which would translate into an error growth rate that is too large. Similarly, these figures do confirm that error growth rates for WAM are expected to be too low.

Last, statistics for the 24-h forecast for both models against buoy data for separate geographical regions are presented in Fig. 6. For most statistical parameters considered, the differences between systems are small and are similar to the differences for the overall dataset as presented in the upper panels of Fig. 5. The exception is the regression slope, where differences between models per region generally are much larger than differences for the composite dataset.

For Japan, the Gulf of Mexico, and the NW Atlantic, NWW3 appears to behave significantly better than WAM, primarily based on the regression slopes. For the NE Pacific and Atlantic, NWW3 has much better correlation coefficients and smaller standard deviations and rms errors than WAM. However, NWW3 overestimates the slope of the regression line by 16%, whereas WAM underestimates the slope by a much smaller margin.

Results for Hawaii deserve some additional attention. Here, NWW3 has smaller errors (standard deviation, rms error, SI) and a much better correlation coefficient, yet dramatically overestimates extreme events (regression slope 34% too high). The latter can be explained because extreme events in this area are associated with storms tracking north of Hawaii. For such conditions, all buoys except for 51001 are more or less sheltered behind the islands (see Fig. 1). Because neither model resolved the islands, systematic positive biases for extreme events are expected for all buoys except for 51001. To illustrate the impact of the lack of sheltering in the model, statistics for the individual Hawaiian buoys are presented in Table 1.[2] This table indeed shows good results for NWW3 at buoy 51001, with a regression slope comparable to those of the NE Pacific and Atlantic buoys. Buoys 51002 and 51003 show extremely high regression slopes, consistent with the sheltering behind the Hawaiian Islands that is not modeled. Note that the generally good regression line for WAM near Hawaii in Fig. 6 appears to be due to canceling of errors of systematic model behavior at 51001, and the absence of sheltering at 51002 and 51003.

Figure 6 clearly shows the potential pitfalls of validating with buoy data only. Validation statistics vary greatly from region to region, and there is no way to assess how representative the local observations are for global model behavior. This makes a truly global validation with altimeter data of paramount importance. Figures 7–10 show some examples of such a validation. To avoid dealing with dependent data in WAM, and to minimize wind errors in the model, these figures mostly deal with 24-h forecasts. Maps of error statistics against altimeter data have been constructed as in Tolman (1998b).

Figure 7 shows NWW3 and WAM 24-h forecast biases against the altimeter data for the entire validation period. Both models show alternate areas with positive and negative biases in the range of -0.5 to 0.5 m. The distribution of the biases, however, is very different. NWW3 (Fig. 7a) generally shows positive biases in the storm tracks north of 30°N and south of 30°S, and generally negative biases in the Tropics. WAM (Fig. 7b) shows almost exactly opposite behavior.

Figure 8a shows the corresponding scatter indices for NWW3. Areas with SI $>$ 20% are shaded light gray. Such areas with large scatter indices correspond to the following: (i) major unresolved island groups such as Solomon–Bismarck–Fiji (10°–20°S, 140°–150°W), French Polynesia (0°–20°S, 150°E–180°), the Aleutian Islands (50°N, 160°W–180°), and others; (ii) western sides of large basins and enclosed smaller basins such as the Gulf of Mexico, where conditions are dominated by local, potentially poorly resolved wind sea systems; and (iii) the Weddell Sea (50°–70°S, 10°–40°W), where partial ice coverage (not included in models) might be responsible for systematic high biases in both models (see also Fig. 7).

Figure 8b shows the corresponding differences in scatter indices between NWW3 and WAM. Areas where WAM has a smaller SI are shaded light gray. Note that areas with SI $>$ 20% in Fig. 8a do not correspond to areas with large differences between the models. Hence, NWW3 and WAM share these ''problem areas.'' At high latitudes (north of 30°N and south of 30°S), SI differences are small, generally $|\Delta\text{SI}| < 5\%$, with both systems alternately showing lower scatter indices. At lower latitudes (between 30°N and 30°S), NWW3 shows systematically lower scatter indices. Particularly in the eastern Pacific, differences are large. With SI $<$ 10% for NWW3 and ΔSI $>$ 10%, the SI for NWW3 is less than one-half the SI of WAM. Thus, generally, differences in scatter indices are largest when NWW3 performs better.

To assess global model error growth, Fig. 9 shows the corresponding scatter indices for the 72-h forecasts. A comparison of the 24- and 72-h forecast scatter indices of NWW3 (cf. Figs. 8a and 9a) indicates that scatter indices increase with forecast time mainly at higher latitudes. Hence, error growth appears (as expected) to be mostly related to wind forecast errors along dominant storm tracks. In these storm tracks, 72-h forecast scatter indices for WAM are systematically lower than those of NWW3, whereas the opposite remains the case in the Tropics (Fig. 9b). A comparison of Figs. 8b

[2] See also time series at http://polar.ncep.noaa.gov/NEW-WAM.

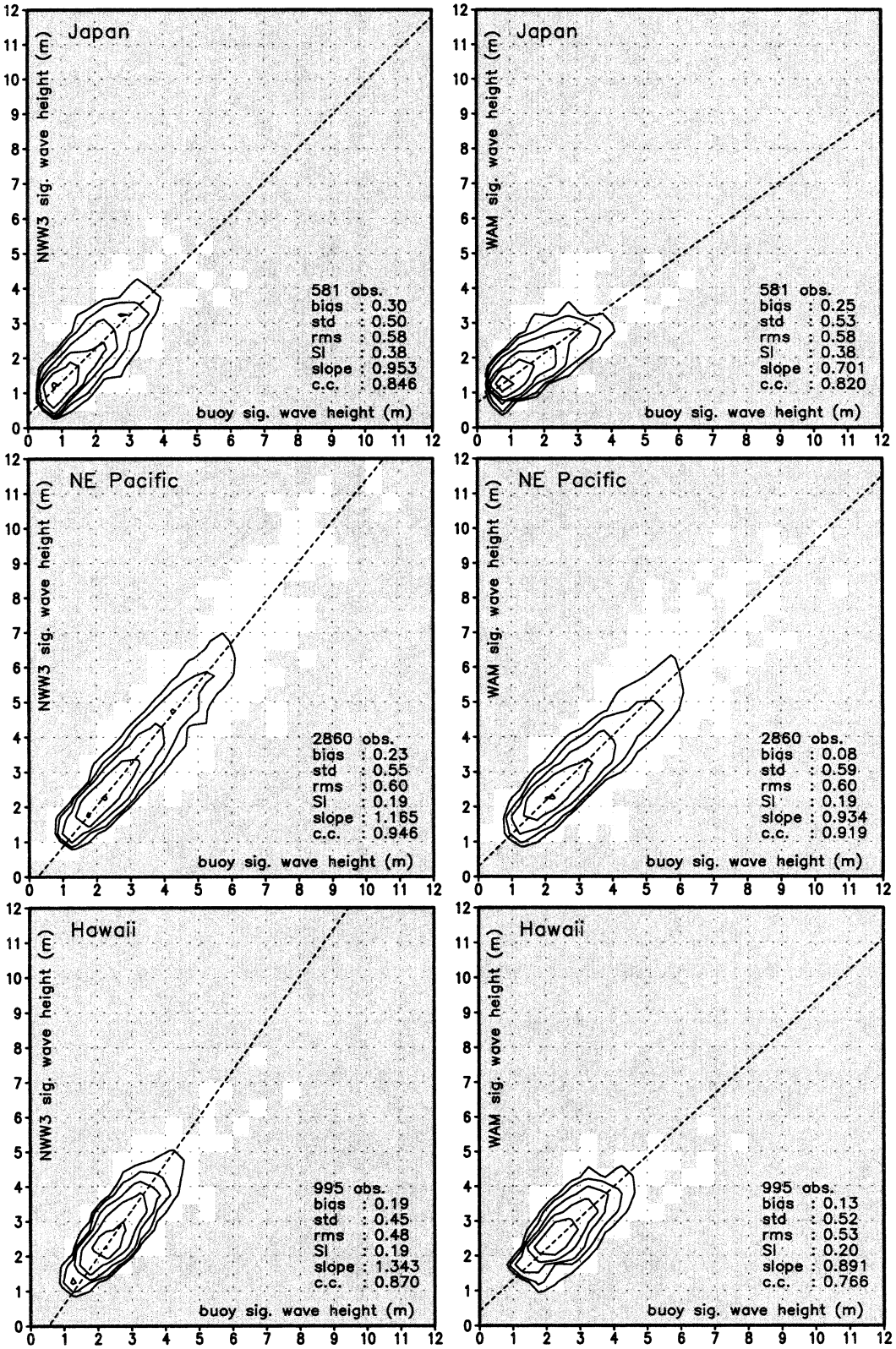


FIG. 6. Joint significant wave height PDF of buoy observations and NWW3 (left panels) or WAM (right panels) for 24-h forecasts for six areas labeled in top left of each panel. Here $\Delta H_s = 0.5$ m, with lowest contour level at 0.02 m^{-2}; contours increment by a factor of 2. Gray areas are devoid of data. See appendix for statistical parameters presented.

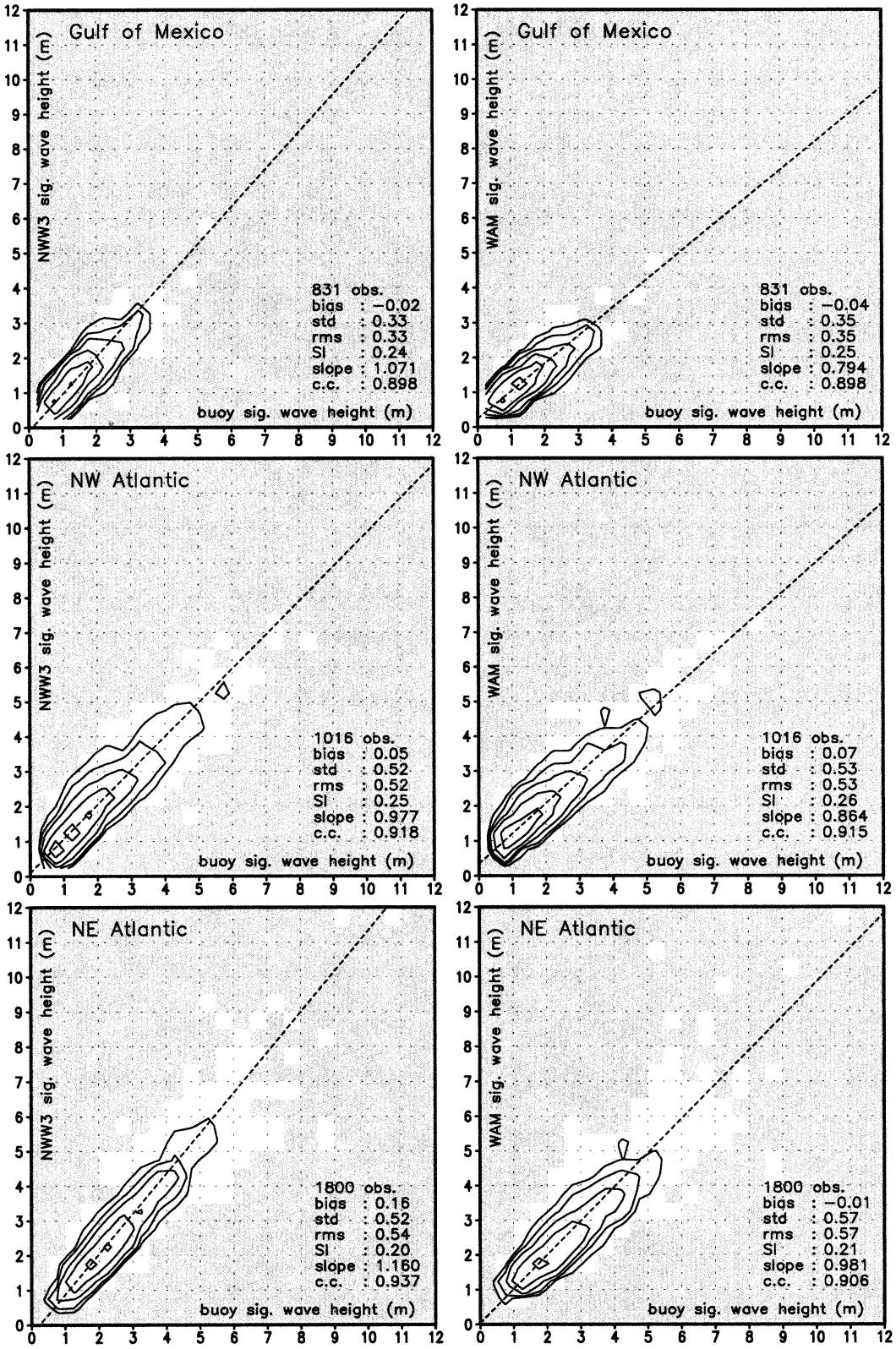


FIG. 6. (*Continued*)

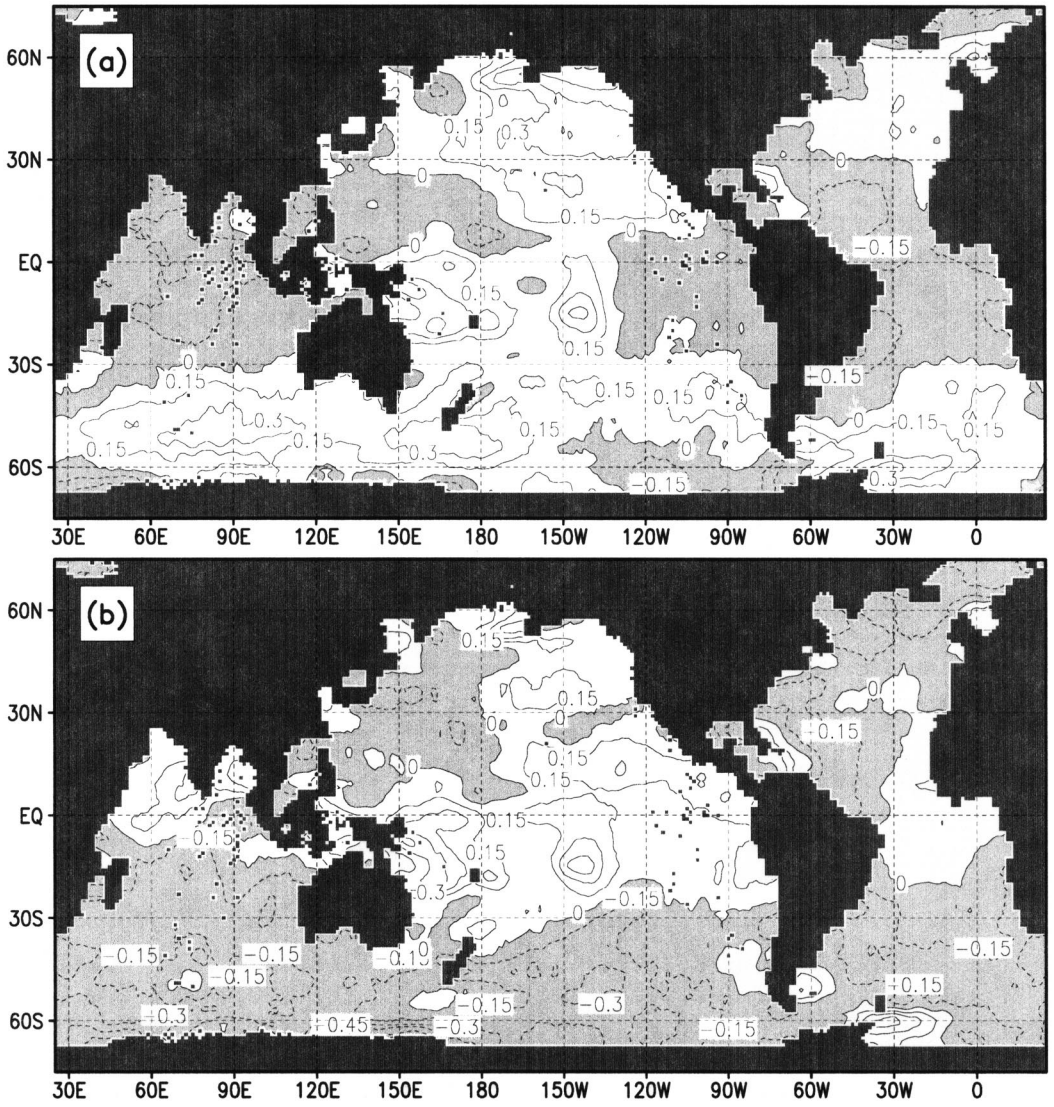


FIG. 7. Wave height bias of 24-h model forecasts against *ERS-2* altimeter data (m) for (a) NWW3 and (b) WAM. Dark gray areas have insufficient data (including land and areas not covered by models); light gray areas have negative biases.

and 9b shows that error growth with forecast time in storm tracks for NWW3 is significantly faster than for WAM, as was also observed and discussed in relation to buoy data before.

Figure 10 shows bulk statistics and PDF for 24-h model forecasts against altimeter data for high latitudes and the Tropics, separated by 30°N and 30°S. For both models, the northern high-latitude statistics (upper panels in Fig. 10) closely resemble the 24-h model statistics against buoy data (Fig. 5, top panels). The most distinct difference is that the regression slopes against the altimeter data for both models are about 3% lower than against buoy data, making the overshoot in slope for NWW3 about equal to the undershoot for WAM. This difference with validation statistics against buoy data is not unexpected, because most buoy data come from the eastern side of basins, where the altimeter indicates that NWW3 has more systematic positive biases.

For the Tropics (middle panels of Fig. 10), most statistics clearly favor NWW3, with the exception of the

TABLE 1. Statistics for NWW3 and WAM 24-h forecast against buoy data near Hawaii. Note that buoy 51004 produced insuffcient data for the parallel comparison period. Left number in each column corresponds to NWW3; right number to WAM.

Buoy	No. of obs	Bias (m)		Rmse (m)		Slope (−)		C.c. (−)	
51001	305	0.14	0.00	0.48	0.59	1.17	0.74	0.91	0.82
51002	314	0.24	0.22	0.52	0.53	1.62	1.15	0.81	0.72
51003	316	0.19	0.18	0.49	0.50	1.58	1.08	0.85	0.76

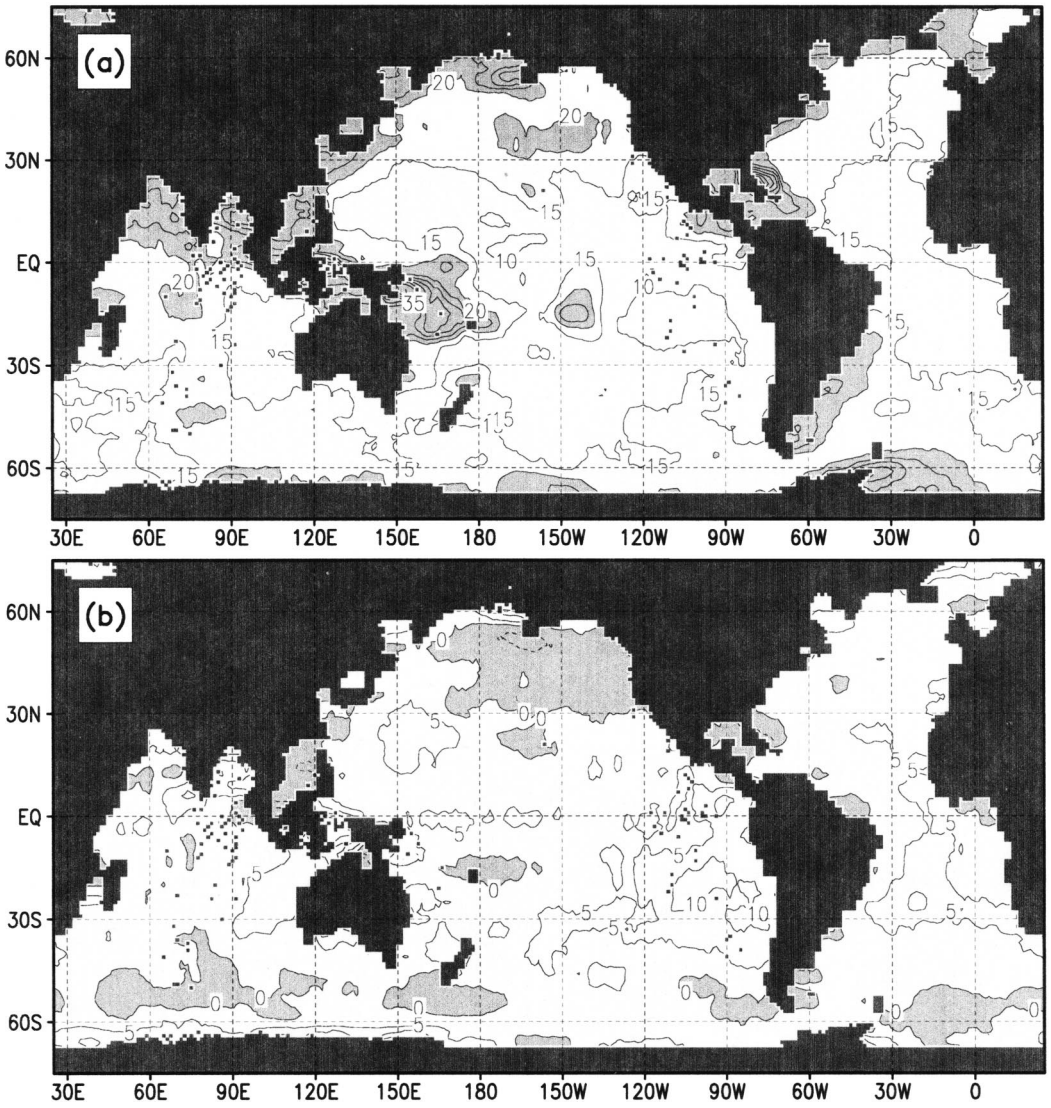


FIG. 8. Wave height SI of 24-h model forecasts against *ERS-2* altimeter data (%). (a) NWW3; light gray indicates SI > 20%. (b) The SI of WAM minus SI of NWW3; light gray indicates smaller SI for WAM. Dark gray areas have insufficient data (including land and areas not covered by models).

regression slope, which is about 23% too high. The latter behavior is similar to the model behavior against buoy data around Hawaii (Fig. 6), and probably could also be caused at least in part by the presence of many unresolved island groups in the models.

In the Southern Hemisphere (bottom panels of Fig. 10), all statistics clearly favor NWW3. Note that NWW3 shows fairly similar behavior for the Northern and Southern Hemispheres. WAM, on the contrary, shows significantly poorer regression slopes and correlation coefficients in the Southern Hemisphere than in the Northern Hemisphere (cf. upper and lower panels in Fig. 10).

Parallel comparisons for the regional models have been much less detailed for three primary reasons. First, for the regional models, NWW3 replaces WAM or even older wave modeling technology that was previously proven to be inferior to WAM. It is therefore sufficient to show that the regional and global models have similar characteristics. This was done in Chao et al. (1999a,b).

Second, the new regional models cover much larger domains than the models they replaced, to cover more areas for which local NWS Weather Forecast Offices (WFOs) have forecast responsibilities. For the AKW model, such areas include the Beaufort Sea. For the WNA model such areas include the tropical regions covered by the Tropical Prediction Center (TPC). The increased domains by themselves are a good reason to replace the previous regional models.

Third, replacing all models by the same generic wave model was part of the overall development plan at

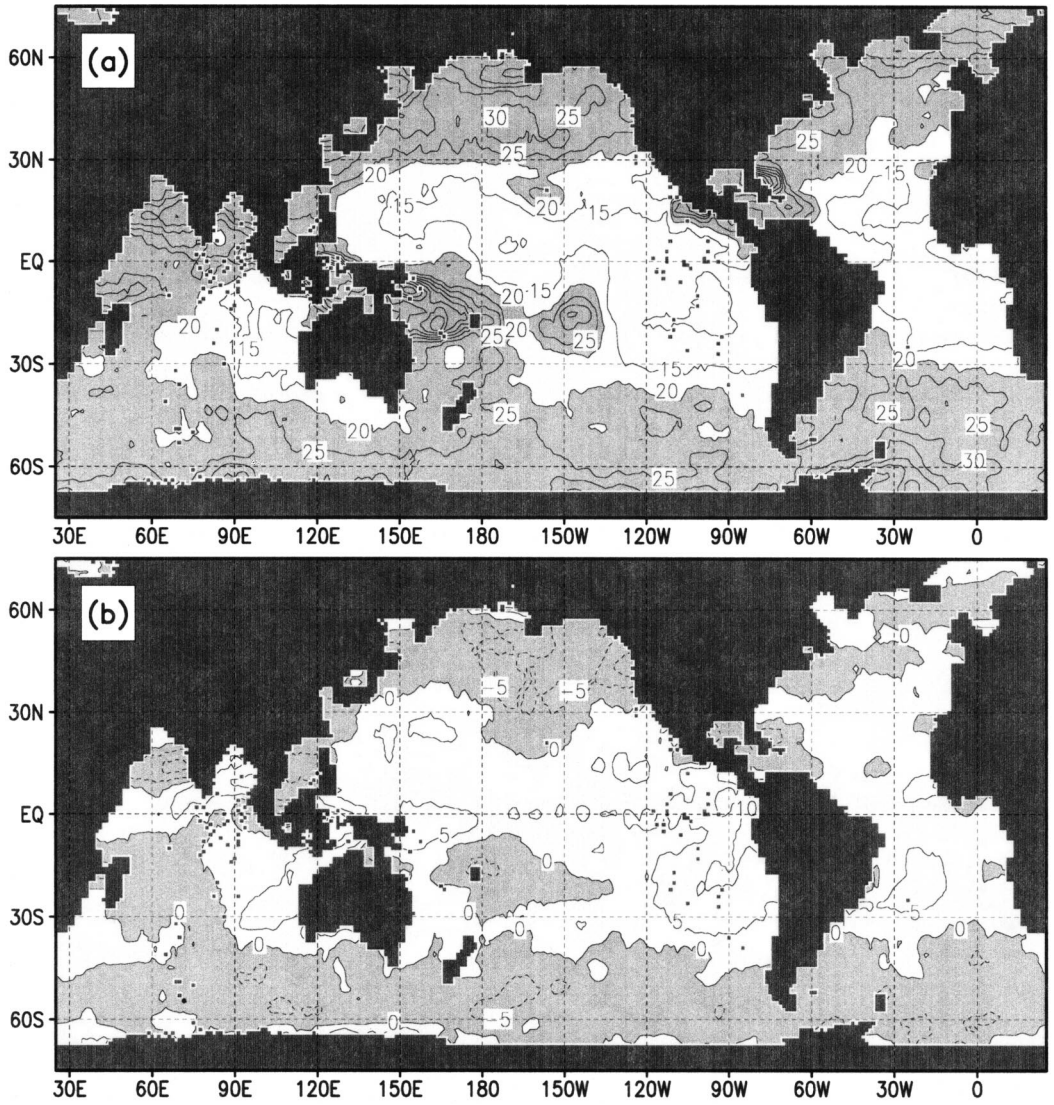


FIG. 9. As in Fig. 8 but for 72-h model forecasts.

NCEP/OMB. Having a single model applied to different regions greatly simplifies maintenance of the operational job suite at NCEP.

Considering the above, no further validation data are presented here for the regional models, and references are made to Chao et al. (1999a,b) for details on the limited parallel comparisons available.

c. Ongoing validation

The parallel comparison between NWW3 and WAM presented above was carried out to document the improvement of NWW3 over WAM, in order to justify replacing NCEP's operational wave forecast systems. As a matter of practice, NCEP continues to monitor the performance of all operational models to continue to identify areas where improvements are needed. The validation of NWW3 (global and regional) is therefore an ongoing effort. Monthly comparisons of the models against buoy data, and seasonal comparisons against buoy and altimeter data are presented at the NWW3 Web site (http://polar.ncep.noaa.gov/waves; see separate links for global and regional validation pages), and are upgraded regularly. For the regional models, buoys of Fig. 1 that fall within the domain are used for validation. Additionally, near-coastal buoys, whose locations were not properly resolved by the global model, are considered in the regional model validation. Figures 11 and 12 identify all validation buoy locations used for the two regional models.

Here, we will only present monthly bias, rms errors, and scatter indices for hindcasts for all three models against all buoy data from the start of the parallel model runs to the present (Fig. 13). For the global NWW3

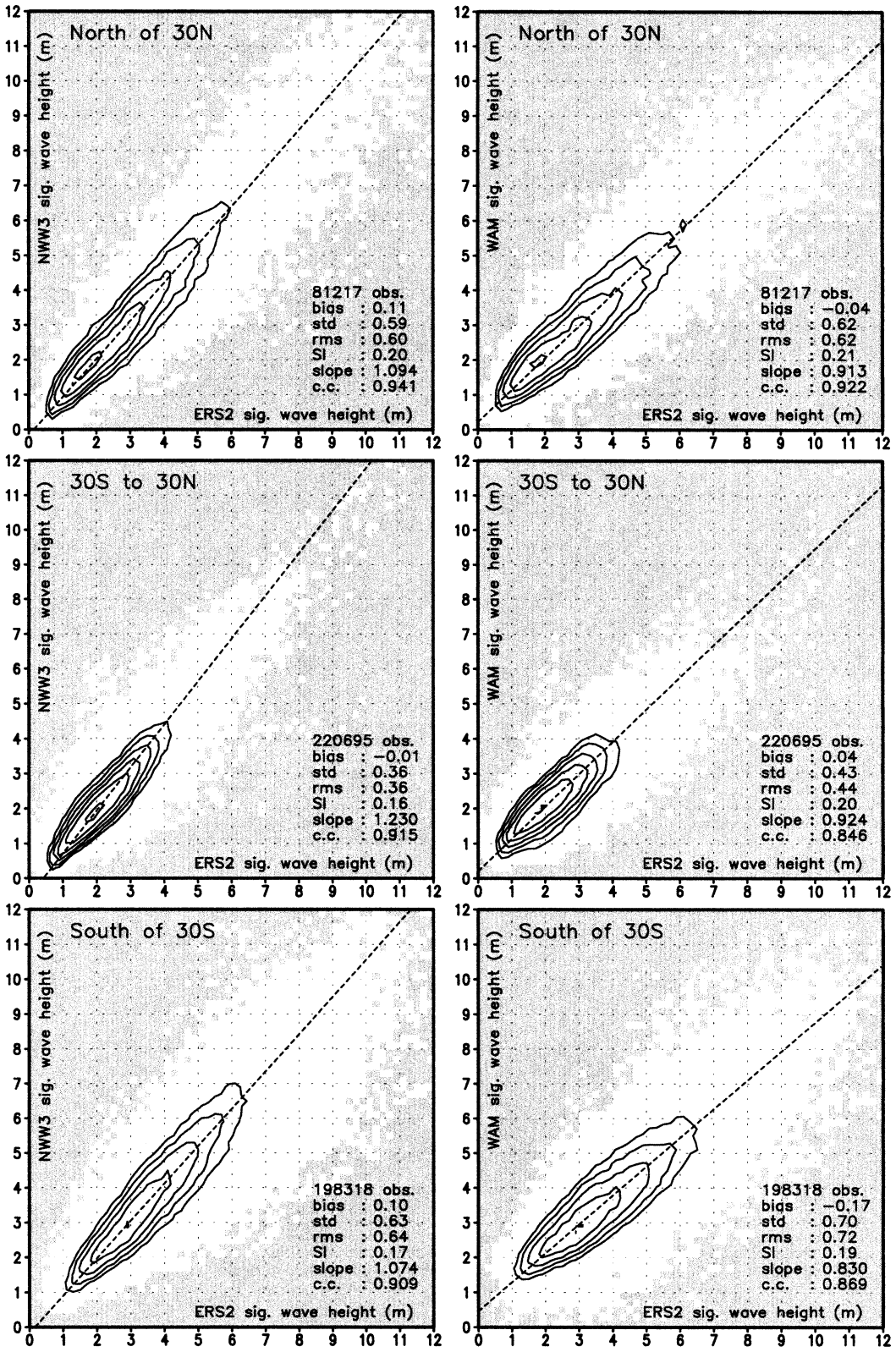


FIG. 10. Joint significant wave height PDF of *ERS-2* altimeter observations and NWW3 (left panels) or WAM (right panels) for 24-h forecasts for three latitude ranges. Here $\Delta H_s = 0.2$ m, with lowest contour level at 0.02 m^{-2}; contours increment by a factor of 2. Gray areas are devoid of data. See appendix for statistical parameters presented.

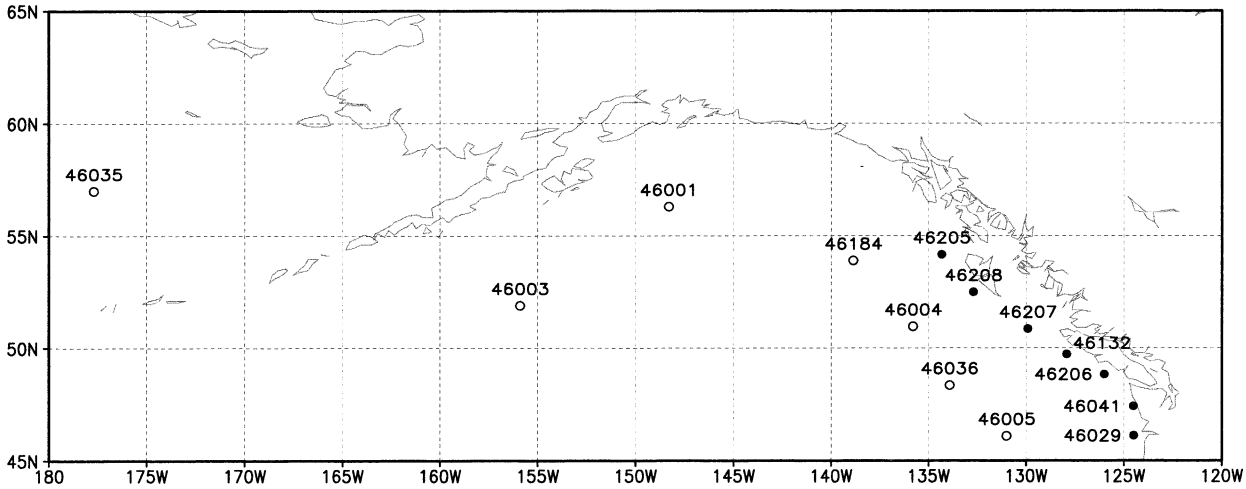


FIG. 11. Location of validation points used for the regional AKW model with corresponding WMO buoy identifiers: ○, output locations shared with global model; ●, output locations not shared with global model.

model this covers a 4-yr period; for the regional models just over 1 yr is covered. These time series of model errors are presented to illustrate seasonal behavior, consistency of model results, and effects of model changes and problems. Note that some of the month-to-month variability of the validation results can be attributed to random failure of instruments.

For the global NWW3 model, the time series of error statistics are sufficiently long to identify seasonal cycles. The bias and rms error (solid lines in Figs. 13a and 13b) show a clear seasonal cycle, with the largest errors in the Northern Hemisphere winter, that is, coinciding with the most extreme wave conditions. The scatter index, however (solid line in Fig. 13c), shows no discernible seasonable cycle. Hence the seasonal cycle in the bias and rms is a direct consequence of more active wave conditions in the winter. It is very encouraging to see that NWW3 has similar relative accuracy in benign (summer) and extreme (winter) wave condition.

Because of the absence of a seasonal cycle, the scatter index can be used to identify changes in model behavior, as well as periods with anomalous model behavior. The first such period, marked as A in Fig. 13c, appears to show the best model results of the entire 4-yr period (lowest scatter index). This is most likely a spurious observation, because it coincides with the failure of a large number of buoys in the NW Atlantic, where relative errors tend to be larger in spring and summer (see the NWW3 Web page). This period is directly followed by a period with anomalously large errors (marked B).

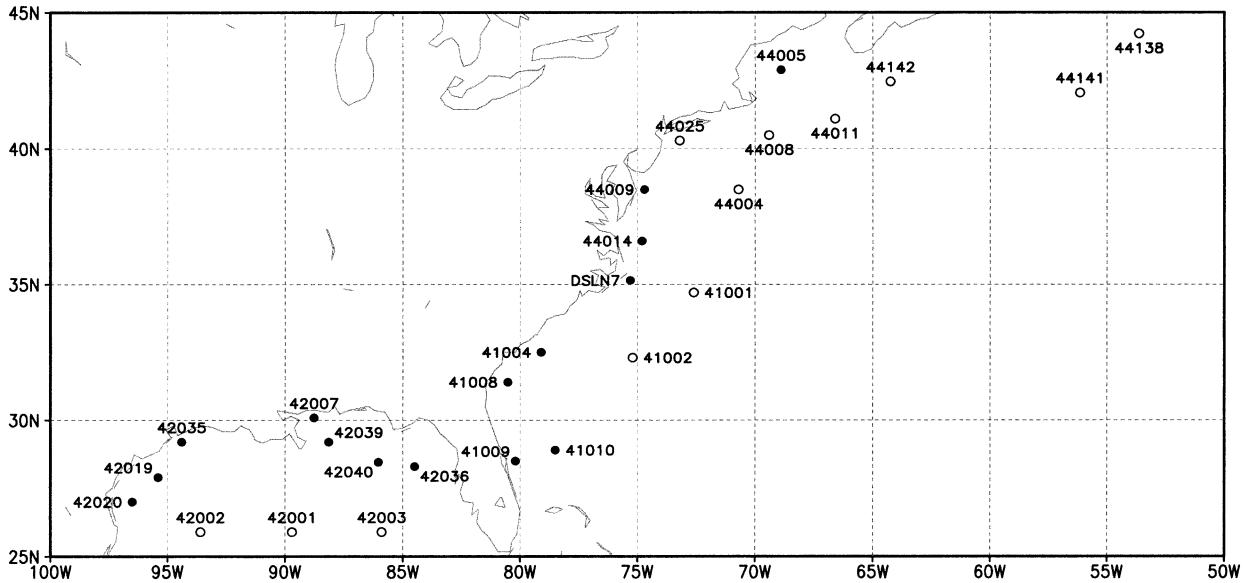


FIG. 12. As in Fig. 11 but for the regional WNA model.

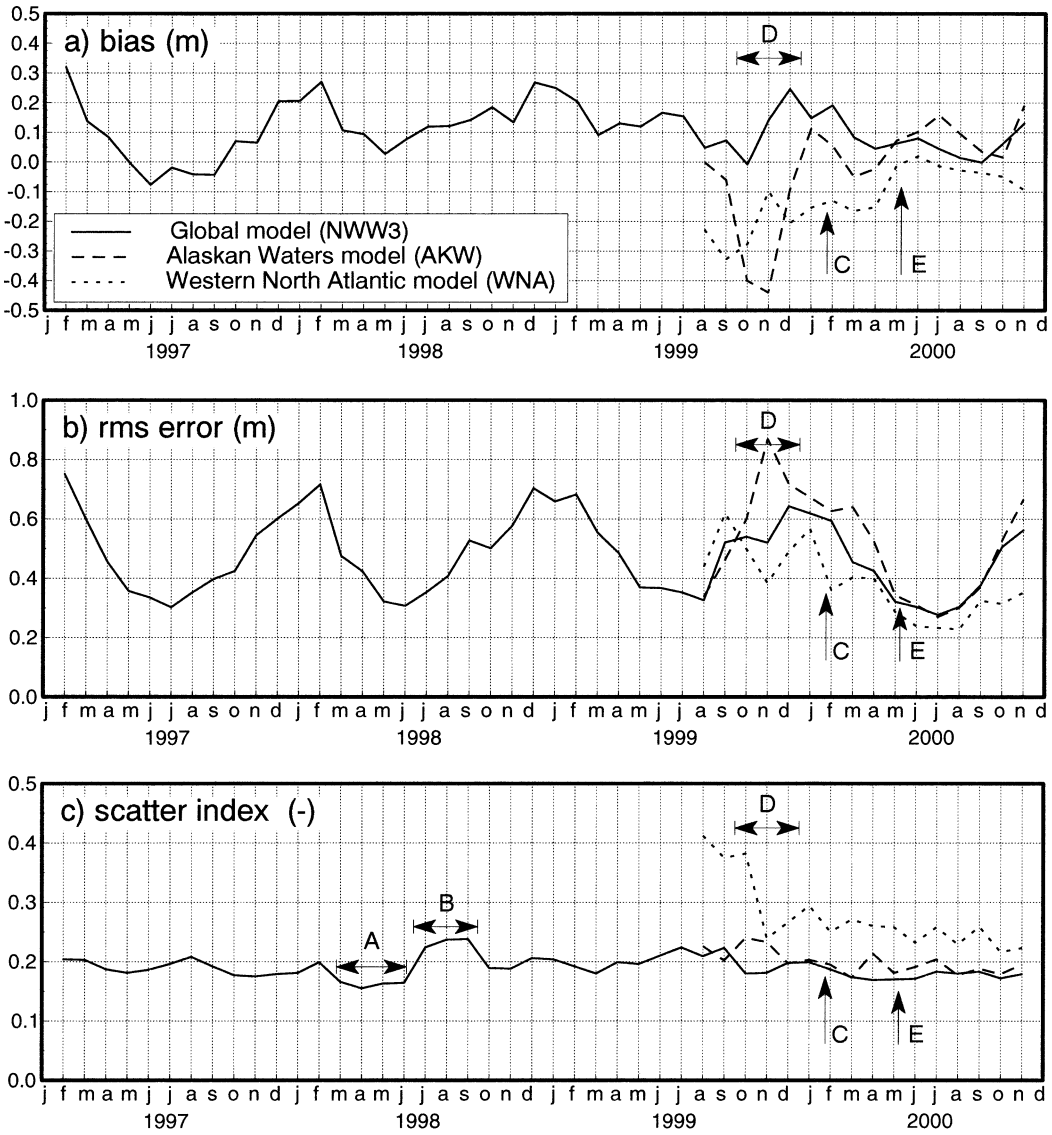


FIG. 13. Wave height hindcast validation statistic against buoy data for global and regional NWW3 models for 1997–2000: (a) bias (m), (b) rms error (m), and (c) SI (rms error normalized with mean observation). Here A–E are periods with anomalous model behavior as discussed in text. The model legend in (a) applies to (b) and (c).

During this time frame, an attempt was made to increase the resolution of the operational GDAS and Aviation Model (AVN) wind fields. Unexpected problems resulted in a noticeable reduction of the quality of the wind fields, which translated into larger wave model errors. At the end of this period, attempts to run the GDAS and AVN models operationally at higher resolution were abandoned and postponed until late January 2000 (marked C). From this point on, the NWW3 scatter index appears to be systematically reduced, suggesting that a better quality of the wind fields has translated into better wave model results.

The shorter time series of error statistics for the regional models makes it harder to assess their systematic behavior. Furthermore, during the period marked D, the regional models in effect did not get boundary data from the global model because of a coding error. This is particularly clear from the spuriously large negative bias for the AKW model (dashed line in Fig. 13a). For the WNA model this is less obvious, because most buoy locations in this model are far from the open boundaries. Also, at time E, a bug suppressing initial wave growth and growth after long periods of calm was removed, and bottom friction in the regional models was retuned. This is particularly important in enclosed basins such as the Gulf of Mexico and, therefore, has a clear impact on the bias of the WNA model (dotted line in Fig. 13a). The time series are too short to assess the systematic impact of these changes, particularly since seasonal behavior of both regional models has not yet been established.

The short time series do allow for a comparison between the global and regional models. The Alaskan Waters and global models (dashed and solid lines in Fig. 13, respectively) show similar biases, rms errors, and scatter indices, particularly after the bug fix ending period D. This could be expected because (i) the part of the Pacific in which all buoys are located is fairly representative for the global model and (ii) wave and weather patterns travel mostly eastward here. The wave climate for additional coastal buoys in the regional model (filled circles in Fig. 11) is therefore expected to be similar to the climate at the deep ocean buoys (open circles in Fig. 11) in the regional and global models. The same is expected to be true for error statistics.

The Western North Atlantic and global models (dotted and solid lines in Fig. 13, respectively) behave more differently. Absolute errors in the WNA model are smaller (Fig. 13b), but relative errors are much larger (Fig. 13c). This could be expected because additional coastal buoys in the regional model (filled circles in Fig. 12) are in the enclosed Gulf of Mexico, and/or close to the coast in areas where weather patterns mostly move offshore. In such areas, wave heights in general are much smaller than in the deep ocean. One might therefore hope that absolute errors (Fig. 13b) are also smaller. However, simultaneously spatial and temporal scales of waves fields are smaller here, which is expected to lead to larger relative errors (Fig. 13c).

6. Products

The most complete set of global and regional NWW3 products can be found on the NCEP/OMB waves Web site (http://polar.ncep.noaa.gov/waves), which also includes a full (and regularly updated) documentation of all available products (http://polar.ncep.noaa.gov/waves/products.html). Here, we can only give a cursory overview of what is available.

Three types of products are available and are updated every 12 h: (i) graphics products consisting of maps of wind speeds, wave heights, and peak period and direction (at selected times and animations); plots of full wave spectra at all output points for selected times; and plots of spectra and source terms at all output points for the nowcast (i.e., 0-h forecast); (ii) text products consisting of bulletins for model output points, describing individual wave fields making up the complete spectra; and (iii) binary products consisting of fields of wave heights and so on in WMO gridded binary (GRIB) format, as well as full spectral data for output points in compressed ASCII format. Also available online are validation results and historical archives for all three models.

The dissemination of NWW3 products on the Internet started as a method to quickly get experimental data out to other NWS organizations, in order to get feedback on model performance during parallel testing. The Web pages have become an important method of disseminating wave model results to the public in general, with approximately 20 000 graphics products and 1 Gb of binary data downloaded each day (July 2001 statistics). The Web pages, however, are not an official operational product of NCEP and may, therefore, occasionally suffer outages.

Within the NWS, operational dissemination of NWW3 products takes place through the Satellite Broadcast Network (SBN), and visualization is enabled through the Advanced Weather Information Processing System (AWIPS). Presently, the above-described GRIB fields are transmitted though the SBN for all three models, at a reduced time resolution as compared with the products available on the Web site. Furthermore, only the global NWW3 fields can be processed by AWIPS. The regional GRIB data have been scheduled for inclusion in upcoming releases of AWIPS. Also available on SBN and AWIPS are the text bulletins for output locations of all models. Graphical products for output locations are not available through SBN and AWIPS, and are presently not scheduled for such.

7. Summary and conclusions

In this paper, we presented a brief review of numerical wave modeling efforts at NCEP. We then presented the new operational NOAA WAVEWATCH III (NWW3) wave forecast system. An extensive parallel comparison between this system and the previously operational WAM-based system (denoted simply as WAM) is presented, using deep-ocean buoy data as well as *ERS-2* altimeter wave observations.

Hindcast time series of wave heights at selected buoy locations indicate that, particularly away from storm tracks, NWW3 represents maximum and minimum wave heights much more realistically than WAM (Fig. 3). Particularly for swells around Hawaii, and in the Gulf of Mexico, time series of wave heights show dramatically better results for NWW3 (Figs. 3a and 3b).

Bulk statistical comparisons against buoy data as presented in Figs. 4–6 suggest that NWW3 might be too energetic, systematically overestimating wave heights (regression slopes) by about 10%. Some of this behavior can be attributed to the lack of proper sheltering by unresolved islands (Table 1). WAM systematically underestimates regression slopes in spite of the lack of sheltering.

In spite of the larger wave height variability displayed by NWW3, its standard deviations and rms errors against buoys are generally similar to, or smaller than, those of WAM. Correlation coefficients are systematically better. Regional comparisons of both models to buoy data show much larger differences, in particular in regression slopes between the models, with generally better model behavior for NWW3. Excessively high regression slopes of NWW3 near Hawaii are clearly dominated by the lack of proper sheltering of several buoys by the islands in the model.

A side effect of the more responsive nature of NWW3 is that it is more sensitive to errors in the wind forecasts; that is, it shows a much more rapid error growth with forecast time than WAM. The fact that regression slopes for NWW3 are generally too high may suggest that error growth rates in NWW3 are too large. This depends on whether the overestimation of regression slopes is dominated by the responsive nature of the model or by the lack of sheltering in certain cases. In the former case the error growth rates are expected to be too high; in the latter case they are not. The present data cannot conclusively identify the source of the overestimation of the regression coefficient in NWW3. Because the regression coefficient in WAM is systematically too low, except in cases of extreme lack of sheltering, it is clear the error growth rates in WAM are too low.

A global validation of both systems against altimeter data (Figs. 7–10) augments the validation with buoy data. The validation results against altimeter data for the higher northern latitudes are very similar to those against all buoy data (also exclusively in the Northern Hemisphere), with fairly similar behavior for both systems. In the Tropics and in the Southern Hemisphere, NWW3 clearly outperforms WAM. A detailed global validation of both models points out some shortcomings of both models. In particular, unresolved island chains such as the Aleutian Islands and French Polynesia show up as areas of significantly increased wave model errors.

Long-term validation of NWW3 against buoy data (Fig. 13) shows consistent model behavior outside the time frame of the NWW3–WAM intercomparison. It also shows a clear seasonal cycle in the bias and rms error but no discernible seasonal cycle in the scatter index. The latter parameter can be used to identify problems in the wave forecast, as well as impact of improvements in the wind fields. Improvement of the wind fields in late January 2000 has clearly had a positive impact on NWW3. This again illustrates the well-known fact that better winds produce better wave forecasts, and that the quality of wind fields is a critical element of a wave forecast system.

The model validation results presented here solely focus on wave heights. This implicitly validates up to some level wave periods and wave spectra, because much of the data consist of swell. Accurate swell predictions are critically dependent on accurate wave periods and spectra in the dominant wave generation areas. Nevertheless, more detailed spectral data can be used for validation. This is presently done in a separate study, the first results of which are presented in Wingeart et al. (2001).

Presently, all wave model products are available on the Internet. Selected products are available through SBN and on AWIPS. Getting the full suite of wave model products on SBN and AWIPS is a high priority item at NCEP.

Recently a version of the regional WNA model driven with high-resolution hurricane winds has been developed. The first results of this experiment (Chao and Tolman 2000) were promising, and during the preparation of this manuscript, this model has also become operational. A high-resolution model for the U.S. west coast and Hawaii is presently under development, first using GDAS and MRF winds, and later to be augmented with a special hurricane version of this model.

Further development of the generic wave model is also an ongoing project. Present test versions of WAVEWATCH III include a new fully modular FORTRAN 90 setup, and further improved integration of the source terms. New propagation concepts that more elegantly deal with the garden sprinkler effect, and that explicitly model unresolved islands, are presently being tested (see Tolman 2001). Long-range development plans include systematic research into the further improvement of all source terms, data assimilation, and coupling to atmospheric and oceanic models.

Acknowledgments. The authors thank Bob Grumbine for his help and encouragement in setting up and maintaining the Web pages, the Marine Prediction Center staff for ongoing validation efforts, and the NCEP Central Operations (NCO) staff for their effort in making and keeping the models operational. We also thank D. B. Rao, Bob Grumbine, Bill Gemmill, Sajal Kar, Naomi Surgi, and three anonymous reviewers for their constructive remarks on early drafts of this manuscript. Development and testing of NWW3 has been supported through funding provided by the NOAA High Performance Computing and Communication (HPCC) office.

APPENDIX

Statistical Methods

If large numbers of data are considered, conventional scatterplots become misleading. The number of outliers in such plots scales with the number of data points N, spuriously suggesting increasingly poor model behavior with increasing N. Such behavior is not displayed by the joint probability density function (PDF) of the observations and the model results. Binning the data pairs in bins with size Δx, the joint probability density function p of the observation x_o and model x_m is estimated as

$$p(x_o, x_m) \approx \frac{n}{N}\Delta x^{-2}, \tag{A1}$$

where n is number of observation in a bin with size $\Delta x \times \Delta x$ around (x_o, x_m). If N is large enough to make n insensitive to sampling error for a given Δx, the PDF becomes independent of N. Furthermore, Δx governs the detail of the PDF that can be displayed. Thus there is a trade-off between resolution and sampling error regarding the choice of Δx, considering that N usually is given. For small N, scatterplots remain the only reasonable display method of such validation data.

Bias, standard deviation, root-mean-square error, cor-

relation coefficients, and regression lines are conventional bulk statistical validation parameters used. Their definition can be found in any textbook, and will not be reproduced here. In wave modeling the scatter index, defined here as the rms error normalized by the average observation, is also commonly used. With the exception of the bias, all these parameters are influenced by random observation errors (e.g., Draper and Smith 1981; Tolman 1998a). For most parameters, this is not important because only relative differences between models are relevant. For the slope of the regression line, however, absolute values are critical. Defining the regression line as $x_m = a + bx_o$, the conventional estimate for the slope b is

$$b = \frac{s_{\text{om}}}{s_{\text{oo}}}, \tag{A2}$$

where s_{om} is the covariance of model and observations and s_{oo} is the variance of the observation. Whereas s_{om} is independent of random observation errors, s_{oo} is systematically increased by such errors, which therefore results in a systematic underestimation of b (e.g., Draper and Smith 1981, section 2.14). If the average error variance of the observation s'_{oo} can be estimated, an error-corrected slope estimate can be made as (Tolman 1998a)

$$b = \frac{s_{\text{om}}}{s_{\text{oo}} - s'_{\text{oo}}}. \tag{A3}$$

This slope estimator requires estimates of observation errors. Here, we need an estimate for the buoy wind and wave errors, and for the altimeter wave height error. Following Tolman (1998b) and Monaldo (1998), these are estimated here as

$$\max(2 \text{ m s}^{-1}, 0.13u), \tag{A4}$$

$$\max(0.1 \text{ m}, 0.08H_s), \quad \text{and} \tag{A5}$$

$$\max(0.1 \text{ m}, 0.05H_s), \tag{A6}$$

respectively.

REFERENCES

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

Caplan, P., J. Derber, W. Gemmill, S.-Y. Hong, H.-L. Pan, and D. Parish, 1997: Changes to the 1995 NCEP operational medium-range forecast model analysis/forecast system. *Wea. Forecasting,* **12,** 581–594.

Cardone, V. J., and D. Ross, 1977: State of the art wave prediction methods and data requirements. *Ocean Wave Climate,* M. Earle and A. Malahoff, Eds., Plenum Press, 61–92.

Chao, Y. Y., 1991: The Gulf of Mexico spectral wave forecast model and products. NOAA/NWS Tech. Procedures Bull. 381, 4 pp.

——, 1995: The Gulf of Alaska regional wave model. NOAA/NWS Tech. Procedures Bull. 427, 10 pp.

——, cited 1997: The U.S. east coast–Gulf of Mexico wave forecasting model. NOAA/NWS Tech. Procedures Bull. 446. [Available online at http://polar.wwb.noaa.gov/omb/tpbs/tpb446/tpb97.html.]

——, and H. L. Tolman, 2000: Numerical experiments on predicting hurricane generated wind waves. Preprints, *Sixth Int. Workshop on Wave Hindcasting and Forecasting,* Monterey, CA, Environment Canada, 167–179.

——, L. D. Burroughs, and H. L. Tolman, cited 1999a: Wave forecasting for Alaskan waters. NOAA/NWS Tech. Procedures Bull. 456. [Available online at http://polar.wwb.noaa.gov/omb/tpbs/akwtpb/akwtpb.html.]

——, ——, and ——, cited 1999b: Wave forecasting for western North Atlantic and adjacent waters. NOAA/NWS Tech. Procedures Bull. 459. [Available online at http://polar.wwb.noaa.gov/omb/tpbs/wnatpb/wnatpb.html.]

Chen, H. S., 1995: Ocean surface waves. NOAA/NWS Tech. Procedures Bull. 426, 17 pp.

——, L. D. Burroughs, and H. L. Tolman, cited 1999: Ocean surface waves. NOAA/NWS Tech. Procedures Bull. 453. [Available online at http://polar.wwb.noaa.gov/omb/tpbs/nww3tpb/nww3tpb.html.]

Chin, H., 1986: Guidance in high seas ocean wave height forecasts. NOAA/NWS Tech. Procedures Bull. 364, 5 pp.

——, and L. D. Burroughs, 1988: Ocean surface wave directional-frequency spectra forecast data: Alphanumeric format. NOAA/NWS Tech. Procedures Bull. 374, 4 pp.

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. J. Lord, 1991: The new global operational analysis system at the National Meteorological Center. *Wea. Forecasting,* **6,** 538–547.

Draper, N. R., and H. Smith, 1981: *Applied Regresion Analysis.* John Wiley and Sons, 709 pp.

Esteva, D., and D. Kidwell, 1990: Ocean surface waves. NOAA/NWS Tech. Procedures Bull. 388, 4 pp.

Gelci, R., H. Cazalé, and J. Vasal, 1956: Utilization des diagrammes de propagation à la prévision énergétique de la houle. *Bull. Inf. Com. Central Oceanogr. Etud. Côtes,* **8,** 169–197.

——, ——, and ——, 1957: Prévision de la houle. La méthode des densités spectroangulaires. *Bull. Inf. Com. Central Oceanogr. Etud. Côtes,* **9,** 416–435.

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

Hasselmann, K., and Coauthors, 1973: Measurements of wind-wave growth and swell decay during the Joint North Sea Wave Project (JONSWAP). *Dtsch. Hydrogr. Z.,* **8A** (12), 1–95.

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.

Hubert, W. E., 1957: A preliminary report on numerical sea condition forecasts. *Mon. Wea. Rev.,* **85,** 200–204.

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

——, and Coauthors, 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. A. E. 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. Eng.,* **19,** 59–98.

——, 1991: The ULTIMATE conservative difference scheme applied to unsteady one-dimensional advection. *Comput. Methods Appl. Mech. Eng.,* **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.

Pore, N. A., and W. S. Richardson, 1968: Wind-wave, swell and

combined wave forecast. NWS Tech. Procedures Bull. 17, 8 pp. plus figures.

Sverdrup, H. U., and W. H. Munk, 1946: Empirical and theoretical relations between wind, sea and swell. *Trans. Amer. Geophys. Union,* **27,** 823–827.

——, and ——, 1947: Wind, sea and swell: Theory of relations for forecasting. U.S. Hydrographic Office (HO) Tech. Rep. 1, HO Publ. 601, 44 pp.

SWAMP Group, 1985: *Ocean Wave Modelling.* Plenum Press, 256 pp.

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

——, 1995: On the selection of propagation schemes for a spectral wind-wave model. NWS/NCEP Office Note 411, 30 pp. plus figures.

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

——, 1998b: Validation of NCEP's ocean winds for the use in wind wave models. *Global Atmos. Ocean Syst.,* **6,** 243–268.

——, 1999: User manual and system documentation of WAVEWATCH III version 1.18. NOAA/NWS/NCEP/OMB Tech. Note 166, 110 pp. [Available online at http://polar.ncep.noaa.gov/waves/wavewatch.]

——, 2001: Improving propagation in ocean wave models. *Proc. Fourth Int. Symp. on Ocean Wave Measurement and Analysis,* San Francisco, CA, ASCE, in press.

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

——, and N. Booij, 1998: Modeling wind waves using wavenumber-direction spectra and a variable wavenumber grid. *Global Atmos. Ocean Syst.,* **6,** 295–309.

WAMDI Group, 1988: The WAM model—A third generation ocean wave prediction model. *J. Phys. Oceanogr.,* **18,** 1775–1810.

Wingeart, K. M., W. C. O'Reilly, T. H. C. Herbers, P. A. Wittmann, R. E. Jenssen, and H. L. Tolman, 2001: Validation of operational global wave prediction models with spectral buoy data. *Proc. Fourth Int. Symp. on Ocean Wave Measurement and Analysis,* San Francisco, CA, ASCE, in press.