Intercomparison of Different Wind–Wave Reanalyses

S. CAIRES AND A. STERL

Royal Netherlands Meteorological Institute, De Bilt, Netherlands

J.-R. BIDLOT

European Centre for Medium-Range Weather Forecasts, Reading, United Kingdom

N. GRAHAM

Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California

V. SWAIL

Environment Canada, Toronto, Ontario, Canada

(Manuscript received 3 December 2002, in final form 17 November 2003)

ABSTRACT

This paper describes the comparison of wind speed and significant wave height data from several reanalyses. The data are assessed against time-averaged altimeter and buoy measurements. The comparisons between the datasets are made in terms of description of short-scale features, monthly means, and long-scale features—namely trends and variability. The results show that although the quality of the datasets in terms of their comparisons with observations differs, most of the long-scale features are equally present in all datasets. The differences between the several wave datasets are larger than those between the wind speed datasets; moreover, differences in wave datasets exist even when the forcing winds used to produce the different wave reanalyses are the same. Most of the discrepancies between the datasets occur in the Tropics, testifying that the physics in that region is still poorly known. The data before the mid-1980s show significant discrepancies also in the Southern Hemisphere, most of which is a consequence of the lack of measurements in those regions in the presatellite era.

1. Introduction

The knowledge of the ocean wind and wave climate, its variability, and possible trends is of great importance to the safety of lives at sea, the design of offshore structures, the protection of coastal areas, and the planning of operations at sea, among other things. The knowledge of the wind speed is particularly important because it is used to derive wave conditions and to compute structural loads.

Before the appearance of wave models the only sources of wave conditions, mainly significant wave height (H_s), were measurements. Visual observations from voluntary observing ships (VOSs) are available since the midnineteenth century, but their raw quality is sometimes poor (Gulev et al. 2003). Observations and measurements from Light Vessels are available since the 1950s, and buoy observations since the 1970s (see, e.g., Gilhousen 1999). All of these data are restricted to coastal locations or to ship routes, and all are mainly in the Northern Hemisphere. Since the advent of satellites, altimeter measurements of H_s have been available globally (see, e.g., Cotton and Carter 1994).

Since World War II numerical models have been used to forecast and mainly hindcast wave conditions. The models use as input the wind speed at a 10-m height (U_{10}). The physics and numerics used in the models have been improving with time, as has the quality of input wind fields. Along with the introduction of data assimilation in the mid-1990s (Komen et al. 1994), these improvements make the present wave predictions so reliable that they are provided operationally at the major weather institutes (see, e.g., Bidlot et al. 2002 and Tolman et al. 2002).

The study of wave climatology and climate variability requires good quality data with a reasonable time and space resolution and coverage. One way of obtaining such data would be to collect *analyzed* wave data pro-

Corresponding author address: Dr. S. Caires, Royal Netherlands Meteorological Institute, P.O. Box 201, NL-3730 AE De Bilt, Netherlands.
E-mail: caires@knmi.nl

duced by one or more of the meteorological institutes over the years. By analyzed data we mean, as usual, model output for a given date corrected through data assimilation using observations available at that date. The quality of the data thereby obtained would, however, be quite inhomogeneous over time, with the inhomogenities being due to two major sources:

- analysis technique—over the years, as mentioned, wave prediction techniques have evolved from hand-made wave maps, using empirical charts, to the present third-generation wave models, using modern assimilation techniques, each time with finer grid resolutions; and
- quality, coverage, and resolution of the observations used as input and in the assimilation—for example, the presently available altimeter measurements provide a lot of data in the South Hemisphere, where in the past there were hardly any observations.

Although little can be done to improve the past quality and coverage of observations, the past wave analyses can be redone by running the same numerical model throughout the period in question. This is the goal of *reanalysis* studies: to produce a dataset with no inhomogeneities as far as the technique of analysis is concerned, by using the same numerical model throughout.

Wind fields from four reanalysis efforts have been used online or offline to provide wave conditions (some of which produced more than one dataset):

1) The European Centre for Medium-Range Weather Forecasts (ECMWF) produced a 15-year world climatology reanalysis, ERA-15 (Gibson et al. 1997). There was no wave model coupled with the system used, but Sterl et al. (1998) computed the corresponding (offline) ERA-15 ocean wave field using the ERA-15 surface winds to force the wave model (WAM; WAMDI 1988), producing the first global ocean wave reanalysis.
2) The National Centers for Environmental Prediction (NCEP), in collaboration with the National Center for Atmospheric Research (NCAR), also produced a reanalysis of data from 1958 to 1997 (Kalnay et al. 1996), the NCEP–NCAR reanalysis, which was then extended and now covers the period of 1948–2003 (Kistler et al. 2001). Again, in this reanalysis there was no wave model coupled, and the ocean wave fields were obtained offline by Cox and Swail (2001), using the Ocean Data Gathering Program version 2 (ODGP2) wave model (see Cox and Swail 2001, and references therein), for the whole globe, and by Pacific Weather Analysis (PWA; see, e.g., Graham et al. 2002), using the Wavewatch III model (WWIII; Tolman 1999) for the Pacific Ocean.
3) Motivated by some deficiencies in the NCEP–NCAR reanalysis winds, Swail and Cox (2000) carried out an intensive kinematic reanalysis of the NCEP–NCAR surface wind fields in the North Atlantic. The resulting improved winds were used to force the Oceanweather, Inc., (OWI) third-generation (3-G) wave model (see Swail and Cox 2000, and references therein).
4) The success of the ERA-15 products led ECMWF to conduct a reanalysis for the longer period of 1957–2002, named ERA-40 (Uppala 2001). Contrary to the other reanalyses there was a wave model coupled with the atmospheric model. Further, the last decade of wave data benefited from the assimilation of European Remote Sensing Satellite (ERS) H_s altimeter measurements.

Thus, there are now several wave reanalysis datasets available that cover the last four decades. Because they were produced using different wave models and different reanalysis wind fields, it is interesting to know which dataset is more adequate for which purpose. Without any assessment of the data, one may expect, for example, the dataset of Swail and Cox (2000) (due to the high quality of the wind fields) to be the most appropriate for a study of the wind and wave conditions in the North Atlantic on a fine time scale, particularly for extremes, and the ERA-40 dataset to be the most adequate for a global study of the wave conditions in the late 1990s, because it is the only one benefiting from the assimilation of altimeter H_s observations. However, it is not clear whether the datasets differ in terms of climatology, for instance, in terms of monthly means, or in terms of large-time-scale features in the data, such as trends. The goal of this article is precisely to assess the quality of each H_s reanalysis dataset at different time scales. More precisely, we will assess the description of short-term features provided by each reanalysis by comparing them with in situ and global measurements to see which compares better with the observations and whether there is a time and/or spatial dependence in the quality of each dataset. Further, we will try to assess whether differences in the several datasets in terms of the description of features at *short time scales,* defined as that of 6-hourly fields, are also present in the monthly means, and, at a later stage, do the same in terms of the description of features at time scales longer than 1 yr, which for simplicity we shall call *long time scales.*

Although our main goal is to compare the different wave reanalyses, we will also assess the reanalysis wind speed fields, and in the same way. Our motivation is to assess how important the choice of wind speed fields is for the wave reanalysis and to see whether the differences–correspondences between the wind fields of the several reanalysis at longer time scales are equally present–absent in the corresponding significant wave height datasets.

The ERA-15 wind and offline wave fields will not be considered here. These fields have been compared to the ERA-40 fields (Caires and Sterl 2001) and the conclusions of those comparisons will be briefly described.

This article is divided in 7 sections. In section 2 we describe the reanalysis datasets used in this study and the buoy and altimeter observations that we have used in their validation. In section 3 we describe the strategy of the validation and intercomparison. In section 4 the different reanalyses of 6-hourly wind and wave fields are validated against buoy and TOPEX/Poseidon altimeter observations. In section 5 we compare the fields of monthly means of the different datasets. In section 6 we compare the datasets in terms of their description of features at long time scales. We finish in section 7 with the discussion of the results and recommendations.

2. Data description

a. *Reanalysis data*

1) ERA-40

Sterl et al. (1998) produced the first global wave reanalysis fields by forcing WAM on a 1.5° latitude × 1.5° longitude grid covering the whole globe with the ERA-15 winds from 1979 to 1993. In their study they analyzed the H_s climatology in terms of annual cycles and trends. Following ERA-15, the ECMWF conducted ERA-40 for the longer period of 1957–2002. This is a reanalysis of, among other things, global ocean wind and waves. It uses ECMWF's Integrated Forecasting System—a coupled atmosphere–wave model with variational data assimilation, which is a state-of-the-art model very similar to the one used operationally, but with a lower resolution. WAM is used, and it is coupled to the atmospheric model through the sea state–dependent Charnock parameter (see Janssen et al. 2002).

There were several sets of observations assimilated into ERA-40. We will briefly mention the ones that more directly affect the quality of the H_s and U_{10} fields. Shipborne wind observations contained in the Comprehensive Ocean–Atmosphere Data Set (COADS; Woodruff et al. 1998) were used. They played an important role especially for the period preceding the availability of satellite measurements, when VOS observations were the only oceanic source of information. Onboard satellite wind measurements of the Special Sensor Microwave Imager (SSM/I) from 1987 and of *ERS-1*/*-2* scatterometer from 1993 were assimilated. *ERS-1* and *ERS-2* H_s altimeter measurements were also assimilated into the model for the period in which they are available (1992–2001).

The wave model grid is the same as that used in the ERA-15 study. Although the results are not presented here, we have compared the ERA-40 and ERA-15 wind speed and wave data, and concluded that the ERA-40 data compare better with the observations than the corresponding ERA-15 data (Caires and Sterl 2001). The superiority of the ERA-40 data relative to the ERA-15 data can be attributed, among other things, to local improvements in the wind fields due to the correction of errors identified in ERA-15.

The ERA-40 significant wave heights from January 1992 until May 1993 are corrupted because during this period erroneous fast delivery product (FDP) *ERS-1* H_s measurements were assimilated into ERA-40.

It should be noted that there is more than one 10-m wind speed parameter available from ERA-40, namely, the 10-m atmospheric wind speed and the 10-m wave model wind speed—the one used in this study. The differences between these two U_{10} products have to do with the way the coupling of the wave model with the atmosphere is done and with the three-dimensional variational data assimilation (3DVAR) scheme used in ERA-40. Roughly speaking, the wave model is forced by hourly winds from the latest 6-h forecast instead of by the analyzed winds (see Janssen et al. 2002).

2) ERA-40/ODGP2

In an independent study the ERA-40 wind fields for 1988 were also used to force the ODGP2 spectral ocean wave model on a 1.25° latitude × 2.5° longitude grid covering the whole globe. We shall refer to this dataset as ERA-40/ODGP2.

3) NCEP–NCAR WINDS AND DERIVED WAVE REANALYSIS

The first long global reanalysis of the surface winds was produced by NCEP–NCAR; it initially covered data from 1958 to 1997 and continues to be extended. This reanalysis also benefited from the assimilation of data from COADS; no satellite wind measurements were used. These wind fields are available on a Gaussian global grid of 1.875°. The fields analyzed here appear as they were used in the Cox and Swail (2001) wave reanalysis, on a global 1.25° latitude × 2.5° longitude grid.

4) CS01

Cox and Swail (2001) used these winds to force the ODGP2 spectral ocean wave model and produced the first 40-yr wave reanalysis covering the whole globe, hereafter referred to as the CS01 dataset. These results were studied in terms of seasonal extremes of wave height by Wang and Swail (2001).

5) PWA-R

PWA produced a 50-yr wave reanalysis using NCEP–NCAR reanalysis winds to force the WWIII on a 1° latitude × 1.5° longitude grid covering the North Pacific Ocean during the winter months (December–March), some of the results of which were reported in Graham and Diaz (2001) and Graham et al. (2002). Building upon that effort, PWA produced a new wave reanalyses using again the NCEP–NCAR reanalysis winds to force WWIII on a 1.5° latitude × 2° longitude grid covering this time the whole Pacific Ocean for the whole year.

TABLE 1. Specific features of every reanalysis dataset used in this study.

Dataset acronym	Resolution (lat × lon)	Spatial coverage	Temporal coverage	Quantity	Wave model	Forcing wind
NCEP–NCAR reanalysis	1.25° × 2.5°	Global	1958 onward	U_{10}		
ERA-40	1.5° × 1.5°	Global	1957–2002	$U_{10,}$ H_s	WAM	ERA-40
ERA-40/ODGP2	1.25° × 2.5°	Global	1988	H_s	ODGP2	ERA-40
CS01	1.25° × 2.5°	Global	1958–1997	H_s	ODGP2	NCEP–NCAR reanalysis
PWA-R	1.5° × 2°	Pacific Ocean	1979–2002	H_s	WW III	NCEP–NCAR reanalysis
AES40	0.625° × 0.833°	North Atlantic	1958–2000	$U_{10,}$ H_s	OWI 3-G	AES40

The reanalysis is for data from 1979 to 2002. There was some tuning in the creation of the data. In order to achieve best results for major winter swell events from 1981 to 1998 at the National Oceanic and Atmospheric Administration (NOAA)/National Data Buoy Center (NDBC) buoy 46011 (35°N, 121°W) near Point Arguello, the U_{10} fields' height was assigned as 6 m. In this article we will only assess this later reanalysis effort and refer to is as PWA-R.

6) AES40

Swail and Cox (2000) carried out an intensive kinematic reanalysis of the NCEP–NCAR reanalysis surface wind fields from 1958 to 2000 in the North Atlantic. The kinematic reanalysis was done interactively by trained marine meteorologists correcting the initial fields with data from different sources. Datasets taken into account were COADS, other buoy, Coastal-Marine Automated Network (C-MAN), and ship observations, but this time attention was paid to the height at which the wind measurements were made, *ERS-1*/*-2* scatterometer winds, tropical cyclone boundary layer fields, and H_s altimeter observations.

The resulting improved winds were used (offline) to force the OWI 3-G wave model on a 0.625° latitude × 0.833° longitude grid. These wind and wave datasets are hereafter referred to as the AES40 datasets. These data were studied in terms of seasonal extremes of wave height by Wang and Swail (2002).

Table 1 synthesises the characteristics of the reanalysis datasets just described, which will be assessed in this study.

b. Measurements

Most of the existing reliable sets of climate observations have been used as input or in the assimilation schemes of the reanalyses. Because validation is possible only if reliable, established, and independent (not used in production of the reanalyses) measurements are used, the observations used in the reanalyses must be borne in mind. With the objective of validating the reanalyses against reliable and, if possible, independent observations, we have assessed the reanalyses U_{10} and H_s fields against buoy data from the NOAA/NDBC and TOPEX/Poseidon altimeter measurements. The buoy's H_s measurements and the TOPEX/Poseidon U_{10} and H_s altimeter measurements were not used in the production of any of the reanalysis data and allow an independent assessment of the quantities. This is not true for the buoy's U_{10} measurements because most of the NOAA/NDBC buoy wind measurements are available in COADS and possibly some were assimilated into the wind reanalyses. So far, buoy observations are considered the most reliable wave observations, but they are limited to some locations along the coast, mainly in the Northern Hemisphere, and are available only at a small number of locations before 1978. From 1978 onward NOAA/NDBC buoy observations off the coast of North America are available (information online at http://seaboard.ndbc.noaa.gov/). TOPEX/Poseidon altimeter observations are available since 1993 and have a global coverage.

In order to compare the reanalysis results with the observations, time and space scales must be brought as close to each other as possible. The reanalysis results are available at synoptic times (every 6 h) and each value is an estimate of the average condition in a grid cell; on the other hand, both the buoy and the altimeter measurements are local. Because the ERA-40 resolution is inbetween the resolution of the other reanalysis products, we will use the resolution of the ERA-40 data as a reference (the implications of which are discussed in section 7). In order to make the time and space scales of the data compatible, the reanalysis data will be compared with 3-h averages of buoy observations (which is the approximate time a long wave would take to cross the diagonal of a 1.5° latitude × 1.5° longitude grid cell at midlatitude) and with the average of the altimeter measurements within a 1.5° latitude × 1.5° longitude cell. The origin of the buoy and altimeter data and the treatment we have applied to its measurements are described below.

1) NOAA/NDBC BUOY MEASUREMENTS

From all of the NOAA/NDBC data buoy locations available during this period, we have selected a total of 19 locations for these comparisons: one off the coast of Peru (buoy 32302), four around the Hawaiian Islands (buoys 51001, 51002, 51003, and 51004), three in the Gulf of Mexico (buoys 42001, 42002, and 42003), four in the northwest Atlantic (buoys 41001, 41002, 41010, and 44004), three off the coast of Alaska (46001, 46003, and 46004), three in the northeast Pacific (46002, 46005, and 46006), and one off the coast of California (buoy 46059). The buoys are sometimes slightly relocated. The precise location of the buoys at different times along with the dates at which the data are available can be obtained from the NOAA/NDBC Web site. The selection of the locations took into account their distance from the coast and the water depth. Only deep-water locations can be taken into account because no shallow-water effects are accounted for in the wave models, and the buoy should not be too close to the coast in order for the corresponding grid point to be located at sea. The buoy H_s and U_{10} measurements are available hourly from 20- and 10-min-long records, respectively. These measurements have gone through some quality control; we do, however, still process the time series further, using a procedure similar to the one used at ECMWF (Bidlot at al. 2002) and described in Caires and Sterl (2003, 43.2–43.3). When the anemometers of the buoys are not at a height of 10 m, the wind speed measurements are adjusted to that height using a logarithmic profile under neutral stability (see, e.g., Bidlot et al. 2002, p. 291). The reanalysis data at the synoptic time around which the buoy measurements were averaged is interpolated bilinearly to the buoy location.

2) TOPEX/Poseidon altimeter measurements

The TOPEX/Poseidon along-track quality checked deep-water altimeter measurements of H_s and the normalized radar cross sections (σ_0) were obtained from the Southampton Oceanography Centre (SOC) Global Altimeter Processing Scheme (GAPS) interface (information available online at http://www.soc.soton.ac.uk/ALTIMETER/GAPS.php; Snaith 2000).

There are several corrections available to bring the altimeter H_s measurements closer to that of the buoys. The TOPEX/Poseidon wave height observations for 1997–99 (cycles 170–235) have drifted; the drift is corrected according to Challenor and Cotton (1999). Caires and Sterl (2003), using a functional relationship model, found that TOPEX/Poseidon data relate to the buoy data according to $H_{\text{sbuoy}} = 1.05H_{\text{stopex}} - 0.07$. We have made the TOPEX/Poseidon observations used here compatible with the buoy observations by applying this linear relationship.

Although altimeters do not measure wind speeds directly, the altimeter backscatter depends and correlates highly with the sea surface wind speed. There are several empirical algorithms available to compute the wind speeds up to 20 m s^{-1} from σ_0. The most recent algorithm is due to Gourrion et al. (2002) and is used here. For wind speeds above 20 m s^{-1} the relation of Young (1993) is used. The satellite measurements are performed about every second with a spacing of about 5.8 km. From these we form altimeter ''observations'' by grouping together the consecutive observations crossing a 1.5° latitude $\times$ 1.5° longitude region (observations at most 30 s or $1.5\sqrt{2}°$ apart). The altimeter observation is taken as the mean of these grouped data points after a quality control similar to the one applied to the buoy data. The reanalysis data at the synoptic times before and after the time of the altimeter observation are interpolated bilinearly to the mean observation location and these two data points are then linearly interpolated in time to the mean time of the observation.

3. Strategy of validation and intercomparison

When trying to assess so many datasets with so many years of data, a strategy must be defined. From the previous section it is obvious that validation against observations can be done only after 1978. For a detailed validation of the reanalysis data against observations we have selected data from 1978, 1988, 1994, and 1997: 1978 was chosen because (i) it is the first year that can be validated against observations, (ii) it is, out of the four, the year for which the quality of the reanalysis data is expected to be worst, and (iii) it also characterizes the quality of all the reanalyses data before the use of satellite observations; 1988 was chosen because it corresponds to a period with more observations relative to all those preceding, at least in the case of the ERA-40 dataset on which (since 1987) SSM/I winds were assimilated (Uppala 2001); and 1994 and 1997 were chosen because they possibly characterize periods of different error properties due to the use of *ERS-1* (1992–95) and *ERS-2* (1996–2000) altimeter H_s observations in ERA-40 and also scatterometer measurements in the AES40 reanalyses. Obviously, only data from these two last years will be assessed against TOPEX/Poseidon observations. The 6-hourly fields from these 4 yr were assessed by looking at different plots, such as scatterplots comparing the different reanalysis products with the buoy and altimeter observations, time series plots in the case of the comparison with buoy data, histograms, and quantile plots. Most of these plots are not shown here, but they are all available at the ERA-40 ocean wave product validation and analysis Web site (online at http://www.knmi.nl/onderzk/oceano/waves/era40/index.html). The differences between the reanalysis products and the observations were also quantified by computing some standard statistics such as the bias ($\overline{y} - \overline{x}$), the root-mean-square error [rmse $= \sqrt{n^{-1} \sum (y_i - x_i)^2}$], the scatter index (SI) [SI $= \sqrt{n^{-1} \sum [(y_i - \overline{y}) - (x_i - \overline{x})]^2}/\overline{x}$], and the correlation coefficient [$\rho = \sum (x_i - \overline{x})(y_i - \overline{y}) / \sqrt{\sum (x_i - \overline{x})^2 \sum (y_i - \overline{y})^2}$]. In all these formulas the x_i's represent the observations, the y_i's represent the reanalysis products, n is the number of observations, and a bar over a variable represents its average. An overview of these results is presented in Tables 2–11 and will be commented separately in terms of wind speed and significant wave height in section 4.

TABLE 2. Wind speed (m s^{-1}) statistics of different reanalysis products vs buoy measurements in different ocean basins. Data for 1978.

Region	n	$\bar{x}$	Reanalysis	Bias	Rmse	SI	ρ
Gulf of Mexico	2522	5.20	ERA-40	0.34	1.55	0.29	0.81
	2522	5.20	NCEP–NCAR reanalysis	1.06	2.06	0.34	0.74
Northwest Atlantic	2325	7.83	ERA-40	−0.91	2.10	0.24	0.88
	2325	7.83	NCEP–NCAR reanalysis	−0.30	2.15	0.27	0.84
	2325	7.83	AES40	−0.73	2.22	0.27	0.86
Alaska	3058	8.17	ERA-40	−0.35	1.80	0.22	0.88
	3058	8.17	NCEP–NCAR reanalysis	0.05	1.89	0.23	0.86
Northeast Pacific	2419	8.86	ERA-40	−1.51	2.83	0.27	0.84
	2419	8.86	NCEP–NCAR reanalysis	−1.17	2.78	0.29	0.81

In order to analyze the differences in monthly means between the several reanalysis sets we produced surface plots of the relative differences [$(\bar{x} - \bar{y})/\bar{x}$] between each pair of datasets and marked the locations at which the Mann–Whitney nonparametric test (see, e.g., Mood et al. 1974, 522–524) rejects the equality of means at the 5% significance level. This analysis was done again for 1978, 1988, 1994, and 1997 and the results are described in section 5.

In order to study differences in the datasets at long time scales, we analyzed the monthly mean fields of the different products for the initial decade of the datasets (1958–67) and for a period at the end (1990–97). We have computed correlations between the datasets and trends for the various months of the different periods. Trends are computed in order to check whether general tendencies of variability are comparable in the different datasets. The trend analysis was done in the same way as described in Wang and Swail (2001). The Mann–Kendall nonparametric test was used to identify the significant trends at a 5% level and the trend estimator is based on Kendall's rank correlation. Correlations were also computed on a monthly basis, and also for the whole period; in the latter case, the annual cycle was removed from the data. Correlations will help in assessing whether there are differences in the short-term variability of the datasets that cannot be identified using the trends. In the correlation map locations at which the Mann–Whitney nonparametric test rejects the equality of means at the 5% significance level are also marked. Obviously, for the PWA-R dataset we will analyze only data from 1979 onward.

4. Data validation

a. *Wind speed*

1) COMPARISON WITH NOAA/NDBC BUOY MEASUREMENTS

We start with the comparisons of reanalysis data against buoy measurements. Tables 2–5 provide the number of paired measurements, reanalysis values (n), average of the measurements ($\bar{x}$), bias (average of the reanalysis data minus the average of the measurements), rmse, SI, and correlation (ρ) of the three U_{10} datasets versus the buoy observations in different basins. The availability of the buoy observations changes with time, 1978 (Table 2) being the year with fewest observations available (see online at http://seaboard.ndbc.noaa.gov/). The analysis of the buoy assessments according to the basin is as follows:

- On the Peruvian coast there are only buoy measurements available for 1988 and 1994. Both the NCEP–NCAR reanalysis and ERA-40 dataset tend to underestimate (negative bias) the buoy measurements from December to May, and overestimate throughout the rest of the year. As can be observed in Tables 3 and 4, the comparative statistics of both reanalyses are very similar.

TABLE 3. The same as Table 2, but for 1988.

Region	n	$\bar{x}$	Reanalysis	Bias	Rmse	SI	ρ
Peruvian coast	1460	6.85	ERA-40	0.25	1.01	0.14	0.87
			NCEP–NCAR reanalysis	0.27	1.26	0.18	0.80
Hawaiian Islands	3396	7.07	ERA-40	−0.28	1.28	0.18	0.83
			NCEP–NCAR reanalysis	−0.74	1.64	0.21	0.77
Gulf of Mexico	3032	5.82	ERA-40	−0.17	1.44	0.25	0.87
			NCEP–NCAR reanalysis	0.58	1.75	0.28	0.83
Northwest Atlantic	4182	7.02	ERA-40	−0.23	1.71	0.24	0.85
			NCEP–NCAR reanalysis	0.44	1.98	0.27	0.81
			AES40	0.27	0.93	0.13	0.96
Alaska	4051	8.16	ERA-40	0.24	1.76	0.21	0.89
			NCEP–NCAR reanalysis	0.65	2.14	0.25	0.86
Northeast Pacific	1941	7.44	ERA-40	−0.03	1.47	0.20	0.91
			NCEP–NCAR reanalysis	0.38	1.74	0.23	0.88

TABLE 4. The same as Table 2, but for 1994.

Region	n	$\bar{x}$	Reanalysis	Bias	Rmse	SI	ρ
Peruvian coast	1400	6.77	ERA-40	−0.02	1.05	0.15	0.82
			NCEP–NCAR reanalysis	−0.01	1.15	0.17	0.78
Hawaiian Islands	4570	8.33	ERA-40	−0.68	1.51	0.16	0.85
			NCEP–NCAR reanalysis	−1.3	2.13	0.20	0.76
Gulf of Mexico	3661	6.04	ERA-40	−0.55	1.53	0.24	0.87
			NCEP–NCAR reanalysis	0.06	1.70	0.28	0.82
Northwest Atlantic	6522	6.94	ERA-40	−0.36	1.68	0.24	0.88
			NCEP–NCAR reanalysis	0.13	1.87	0.27	0.85
			AES40	0.25	0.79	0.11	0.98
Alaska	3329	8.07	ERA-40	−0.25	1.60	0.20	0.91
			NCEP–NCAR reanalysis	0.22	1.90	0.23	0.88
Northeast Pacific	2602	7.52	ERA-40	0.12	1.54	0.20	0.90
			NCEP–NCAR reanalysis	0.30	1.91	0.25	0.86
California	294	8.40	ERA-40	−0.39	1.52	0.17	0.91
			NCEP–NCAR reanalysis	0.29	1.70	0.20	0.88

- Buoy measurements around the Hawaiian Islands are available for 1988, 1994, and 1997. Both ERA-40 and NCEP–NCAR reanalysis fields underestimate the measurements all year round. The rms errors of the NCEP–NCAR reanalysis data are at least 0.31 m s^{-1} higher than those of the ERA-40 data. There is no indication that the quality of the reanalysis fields at these locations depends on the year.
- In the Gulf of Mexico the wind conditions tend to be underestimated by the ERA-40 and overestimated by the NCEP–NCAR reanalysis fields. The ERA-40 data compare better with the observations, but both reanalyses sets have a high scatter index—above 0.24. Again, there is no evidence of a dependence of the error characteristics on the year considered.
- The buoy observations located off of the East Coast of the United States can also be used to assess the AES40 U_{10} fields. As the results presented in Tables 3–5 testify, the AES40 wind speed at the buoy locations compare much better with the observations than the ERA-40 and the NCEP–NCAR reanalysis wind speeds, except for 1978. The scatter index shows values of about 13%, compared with more than 23% for the other products, and those of the rmses are about half of those of the NCEP–NCAR reanalysis and ERA-40 data. The AES40 data for 1978 (Table 2) compare equally well with the observations as the other reanalysis data. The error statistics for ERA-40 are better than those for the NCEP–NCAR reanalysis data.
- At the buoy locations off the Alaskan coast again the ERA-40 data compare much better with observations than do the NCEP–NCAR reanalysis data. ERA-40 underestimates most of the data while the NCEP–NCAR reanalysis overestimates it.
- In the northeast Pacific locations the reanalysis data for 1978 are clearly worse than for the following years. This may be an indication that these buoy's wind speeds are being used by the models. The ERA-40 data tend to overestimate the winter values and underestimate the summer ones. The NCEP–NCAR reanalysis winds underestimate the observations in 1978 and overestimate them in the following years.
- On the coast of California there are only observations available for the last 3 months of 1994 and for 1997. The ERA-40 data have a higher bias but less SI and rmse than the NCEP–NCAR reanalysis wind speeds.

2) COMPARISON WITH TOPEX/POSEIDON ALTIMETER MEASUREMENTS

A global view of the wind quality can be obtained by comparing the reanalysis data with the TOPEX/Poseidon altimeter observations. For the validation and comparison of the NCEP–NCAR reanalysis and the ERA-40 data, we have considered three latitude bands in our comparisons: north of 20°N, south of 20°S, and the region between 20°S and 20°N. For the comparisons with the AES40 dataset we considered two North Atlantic regions: one north of 20°N and the other between the equator and 20°N. The statistics of the different reanalysis comparisons with the TOPEX/Poseidon measurements are presented in Table 6. Because the TOPEX data were collocated to the resolution of the ERA-40 data, the number of measurements used in each reanalysis comparison varies.

As for the in situ comparisons, the ERA-40 U_{10} fields compare better with the observations than those from NCEP–NCAR reanalysis for all statistics but the bias, especially in the southern region where the rmses of the NCEP–NCAR reanalysis data are almost 0.5 m s^{-1} higher than those of the ERA-40 data. Both reanalyses compare better with observations in the northern region than in the southern, but in the case of ERA-40, only marginally.

Contrary to the results of the validation of the AES40 with buoy observations, the quality of the AES40 winds is comparable to those of the ERA-40 and NCEP–NCAR reanalysis fields when considering the whole North Atlantic basin. This is a clear indication that the AES40 wind reanalysis relies heavily on buoy mea-

TABLE 5. The same as Table 2, but for 1997.

Region	n	$\bar{x}$	Reanalysis	Bias	Rmse	SI	ρ
Hawaiian islands	5456	7.31	ERA-40	−0.33	1.26	0.17	0.87
			NCEP–NCAR reanalysis	−0.63	1.57	0.20	0.81
Gulf of Mexico	3663	6.03	ERA-40	−0.62	1.57	0.24	0.87
			NCEP–NCAR reanalysis	−0.06	1.69	0.28	0.83
Northwest Atlantic	4453	6.92	ERA-40	−0.24	1.77	0.25	0.86
			NCEP–NCAR reanalysis	0.05	1.98	0.29	0.82
			AES40	0.21	0.95	0.13	0.96
Alaska	3787	7.56	ERA-40	0.05	1.72	0.23	0.88
			NCEP–NCAR reanalysis	0.40	2.03	0.26	0.84
Northeast Pacific	1551	7.77	ERA-40	−0.01	1.54	0.20	0.89
			NCEP–NCAR reanalysis	0.28	1.86	0.24	0.85
California	1408	7.16	ERA-40	−0.21	1.35	0.19	0.90
			NCEP–NCAR reanalysis	0.06	1.45	0.20	0.88

surements and that these cannot be used for its independent assessment. Still, the error statistics show that the AES40 data have a higher quality, but only slightly. In the tropical Atlantic region the rmse of the ERA-40 U_{10} fields is at least 0.20 m s^{-1} lower than those of the AES40 fields.

The 1%–99% quantile–quantile (Q–Q) plots, comparing the reanalysis data with the TOPEX/Poseidon observations, help in visualising the differences and deficiencies of the datasets. Figure 1 shows the Q–Q plots of the TOPEX/Poseiden U_{10} observations and the corresponding reanalysis data for 1994 and 1997. All of the datasets tend to overestimate wind speeds below 4 m s^{-1} and underestimate wind speeds above 14 m s^{-1}. The underestimation is more severe by the ERA-40 dataset. The plots show that the AES40 dataset is an improvement on the NCEP–NCAR reanalysis dataset, most obviously for high wind speeds.

3) SUMMARY

To sum up: The quality of each dataset differs regionally. In terms of error statistics, the ERA-40 U_{10} fields compare better with the observations than those of the NCEP–NCAR, reanalysis; the kinematically improved wind speeds of Swail and Cox (2000) are clearly of superior quality. The observed distributions are described best by AES40 in the Atlantic Ocean and by

TABLE 6. Wind speed (m s^{-1}) 1994 and 1997 statistics of different reanalysis products versus TOPEX/Poseidon measurements in different ocean basins.

Region	n	$\bar{x}$	Reanalysis	Bias	Rmse	SI	ρ
			1994				
20°–80°N	94759	7.75	ERA-40	−0.19	1.51	0.19	0.91
	93563	7.75	NCEP–NCAR reanalysis	0.10	1.72	0.22	0.88
20°S–20°N	147101	6.28	ERA-40	−0.05	1.31	0.21	0.84
	143517	6.29	NCEP–NCAR reanalysis	−0.09	1.65	0.26	0.75
80°–20°S	226563	8.89	ERA-40	−0.12	1.65	0.18	0.90
	225898	8.91	NCEP–NCAR reanalysis	−0.14	2.20	0.25	0.81
20°–80°N, 70°W–20°E	34712	7.95	ERA-40	−0.23	1.51	0.19	0.92
	34586	7.94	NCEP–NCAR reanalysis	0.09	1.71	0.21	0.89
	34116	7.99	AES40	0.12	1.63	0.20	0.90
0°–20°N, 70°W–20°E	14766	6.58	ERA-40	−0.10	1.16	0.18	0.86
	14545	6.60	NCEP–NCAR reanalysis	0.02	1.54	0.23	0.75
	14832	6.58	AES40	0.00	1.49	0.23	0.77
			1997				
20°–80°N	91994	7.66	ERA-40	−0.17	1.49	0.19	0.91
	90712	7.66	NCEP–NCAR reanalysis	0.11	1.71	0.22	0.88
20°S–20°N	142563	6.27	ERA-40	−0.04	1.30	0.21	0.85
	139033	6.29	NCEP–NCAR reanalysis	−0.11	1.67	0.27	0.75
80°–20°S	220409	8.87	ERA-40	−0.10	1.58	0.18	0.90
	219611	8.89	NCEP–NCAR reanalysis	−0.02	2.15	0.24	0.82
20°–80°N, 70°W–20°E	33731	7.79	ERA-40	−0.22	1.51	0.19	0.92
	33557	7.78	NCEP–NCAR reanalysis	0.08	1.71	0.22	0.89
	33132	7.83	AES40	0.08	1.54	0.20	0.91
0°–20°N, 70°W–20°E	14323	6.44	ERA-40	−0.21	1.21	0.18	0.86
	14108	6.46	NCEP–NCAR reanalysis	−0.07	1.53	0.24	0.76
	14388	6.44	AES40	−0.11	1.42	0.22	0.80

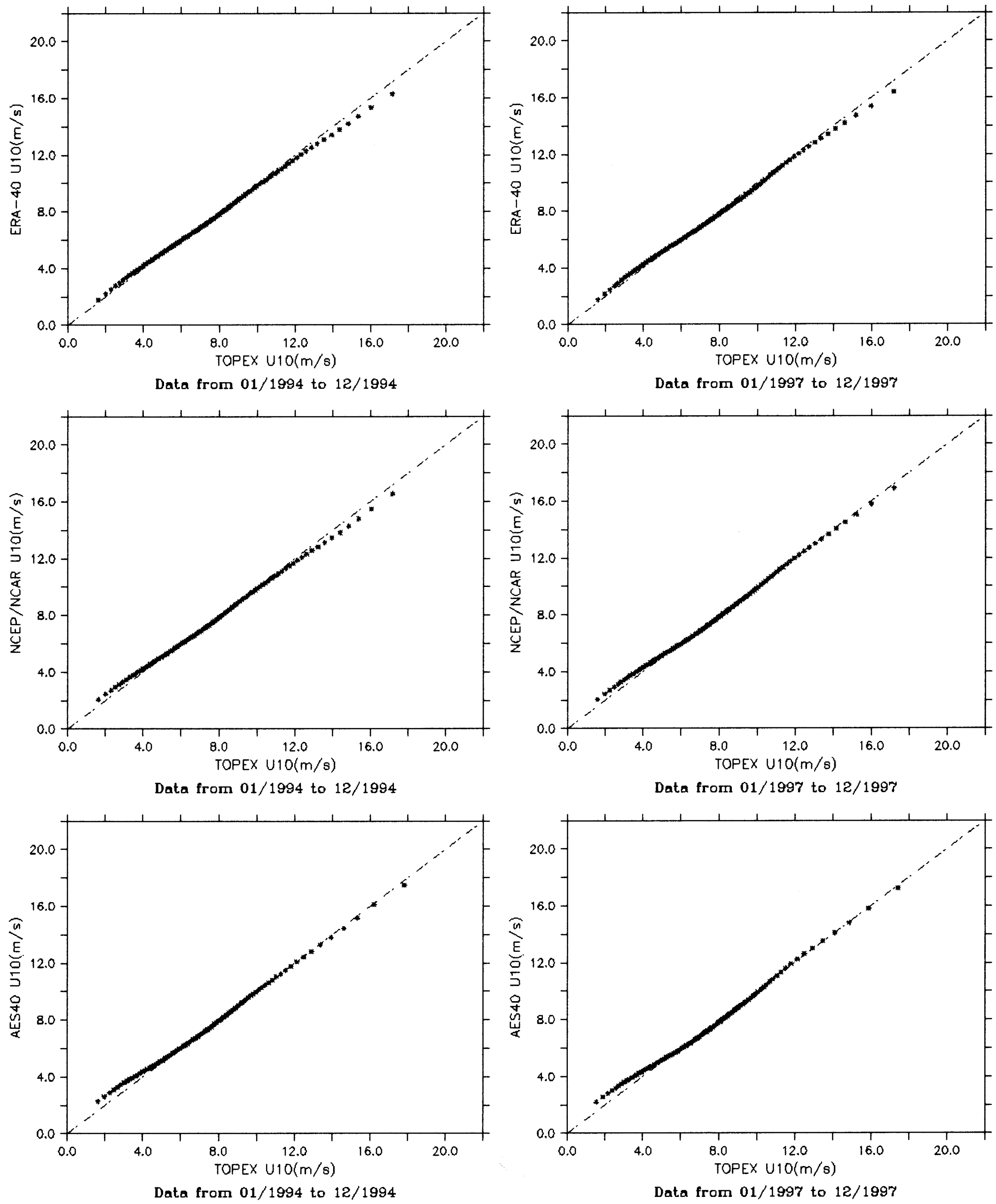


FIG. 1. Graphs comparing the 1%–99% quantiles of TOPEX/Poseidon observations against reanalysis U_{10} data. Data from (left) 1994 and (right) 1997: (top) ERA-40, (middle) NCEP–NCAR reanalysis, and (bottom) AES40. The ERA-40 and NCEP–NCAR reanalysis data are global, and the AES40 data are just for the North Atlantic Ocean.

the NCEP–NCAR reanalysis globally. There is some indication that the quality of the U_{10} fields is better in the years after 1978. This is particularly noticeable for the AES40 U_{10} fields. The rmse values for the buoy comparisons for 1978 are quite close to the values obtained for the NCEP–NCAR reanalysis wind, while in the following years they are only half as large. This is explained by the smaller number of observations available for the kinematical analyses during and prior to this period.

b. Wave height

In terms of H_s we are able not only to assess the five different reanalysis products (although one of which is just of 1 yr), but also the effect of forcing different second- and third-generation wave models using the same wind field, and forcing the same model using different wind fields (see Table 1). The NCEP–NCAR reanalysis winds fields were used to produce the CS01 dataset using a second-generation wave model, as well

TABLE 7. Significant wave height (m) statistics of different reanalysis products vs buoy measurements in different ocean basins. Data for 1978.

Region	n	$\bar{x}$	Reanalysis	Bias	Rmse	SI	ρ
Gulf of Mexico	2654	1.05	ERA-40	−0.24	0.38	0.29	0.90
			CS01	0.35	0.48	0.31	0.86
Northwest Atlantic	2499	1.78	ERA-40	−0.22	0.41	0.20	0.94
			CS01	0.23	0.60	0.31	0.89
			AES40	0.24	0.51	0.25	0.93
Alaska	3313	2.54	ERA-40	0.13	0.48	0.18	0.92
			CS01	0.62	0.85	0.23	0.92
Northeast Pacific	2666	2.31	ERA-40	0.19	0.49	0.19	0.91
			CS01	0.52	0.76	0.24	0.91

as the PWA-R dataset using a third-generation wave model. The ERA-40 winds were used to produce both the ERA-40 H_s and the ERA-40/ODPG2 fields using a third- and second-generation wave model, respectively. The ODPG2 model was used to create both the CS01 and the ERA-40/ODPG2 datasets using the NCEP–NCAR reanalysis and the ERA-40 wind fields, respectively.

1) COMPARISON WITH NOAA/NDBC BUOY MEASUREMENTS

The statistics of the comparisons of the reanalysis datasets with the buoys for 1978, 1988, 1994, and 1997 are presented in Tables 7–10, respectively. The information is the same as provided for wind comparisons in Tables 2–5. In the different ocean basins the reanalysis products compare with the buoy observations as follows.

- As for the wind speed assessment, comparisons in the Peruvian coast buoy location are available only for 1988 and 1994, and all four datasets cover this location. The datasets show comparable quality. The bias of ERA-40 data is closer to 0 than that of the other products, but the monthly biases oscillate between months of underestimation and overestimation.
- In the locations around the Hawaiian Islands the quality of the ERA-40 and CS01 data is quite similar in 1988, but the ERA-40 comparison is clearly better with observations for 1994 and 1997. The PWA-R dataset compares worse with the observations than the CS01 dataset. Because the forcing winds are the same in both cases, differences are due to the choice of the wave model and their calibration. In this region the ERA-40/ODGP2 data can also be compared with the 1998 observations. The error statistics obtained are close to the ones for the respective ERA-40 data.
- As for the assessment of wind speed, both reanalysis datasets have quite high scatter index statistics in the Gulf of Mexico, with the ERA-40 dataset comparing better with observations. The ERA-40/ODPG dataset does a better job depicting the conditions in that basin.

TABLE 8. The same as Table 7, but for 1988.

Region	n	$\bar{x}$	Reanalysis	Bias	Rmse	SI	ρ
Peruvian coast	1461	2.21	ERA-40	−0.03	0.33	0.15	0.84
			ERA-40/ODGP2	−0.30	0.41	0.13	0.87
			CS01	−0.24	0.40	0.15	0.83
			PWA-R	−0.14	0.38	0.16	0.84
Hawaiian Islands	3399	2.20	ERA-40	−0.23	0.42	0.16	0.87
			ERA-40/ODGP2	−0.31	0.45	0.15	0.87
			CS01	−0.16	0.40	0.17	0.83
			PWA-R	−0.45	0.62	0.19	0.81
Gulf of Mexico	3452	1.14	ERA-40	−0.28	0.45	0.31	0.93
			ERA-40/ODGP2	0.15	0.34	0.26	0.94
			CS01	0.33	0.49	0.32	0.90
Northwest Atlantic	4568	1.89	ERA-40	−0.43	0.60	0.22	0.92
			ERA-40/ODGP2	−0.21	0.45	0.21	0.90
			CS01	0.01	0.48	0.25	0.86
			AES40	0.03	0.37	0.19	0.92
Alaska	4054	3.18	ERA-40	−0.35	0.68	0.18	0.94
			ERA-40/ODGP2	−0.10	0.57	0.18	0.93
			CS01	0.29	0.71	0.20	0.92
			PWA-R	−0.14	0.80	0.25	0.91
Northeast Pacific	2179	2.83	ERA-40	−0.17	0.61	0.21	0.94
			ERA-40/ODGP2	−0.09	0.49	0.17	0.94
			CS01	0.14	0.56	0.19	0.93
			PWA-R	−0.02	0.61	0.23	0.94

Table 9. The same as Table 7, but for 1994.

Region	n	$\bar{x}$	Reanalysis	Bias	Rmse	SI	ρ
Peruvian coast	1457	2.18	ERA-40	−0.14	0.30	0.12	0.92
			CS01	−0.43	0.56	0.17	0.82
			PWA-R	−0.33	0.48	0.16	0.83
Hawaiian Islands	4570	2.55	ERA-40	−0.38	0.51	0.13	0.90
			CS01	−0.46	0.62	0.17	0.81
			PWA-R	−0.59	0.73	0.17	0.83
Gulf of Mexico	3884	1.09	ERA-40	−0.18	0.33	0.25	0.93
			CS01	0.32	0.45	0.30	0.89
Northwest Atlantic	6677	1.82	ERA-40	−0.29	0.50	0.22	0.95
			CS01	0.00	0.45	0.25	0.91
			AES40	0.07	0.34	0.18	0.95
Alaska	3783	2.91	ERA-40	−0.37	0.60	0.16	0.97
			CS01	0.20	0.59	0.19	0.94
			PWA-R	−0.17	0.70	0.24	0.93
Northeast Pacific	2910	2.88	ERA-40	−0.38	0.60	0.16	0.94
			CS01	0.02	0.54	0.19	0.93
			PWA-R	−0.14	0.66	0.23	0.93
California	294	3.57	ERA-40	−0.54	0.79	0.15	0.95
			CS01	−0.03	0.58	0.16	0.91
			PWA-R	0.10	0.63	0.17	0.92

- The ERA-40 data persistently underestimate the wave conditions in the Atlantic locations during the 4 yr, although less severely in 1997. The CS01 dataset compares poorly with observations from 1978 and has almost no bias in the following years. The AES40 dataset is the one that compares best with the observations. The 1988 ERA-40/ODGP2 data have an rmse that is 0.15 m smaller than that of the ERA-40 data.
- In the comparisons with the buoy observations off the Alaskan coast and off the coast of the northeast Pacific, the quality of the ERA-40 and the CS01 datasets is comparable for 1988 and 1994, with the ERA-40 data comparing better with observations from 1978 and 1997. This is probably due to the assimilation of the *ERS-2* altimeter wave heights in 1997 and to the fact that the ERA-40 winds compared generally better than the NCEP–NCAR reanalysis winds with observations from 1978. The PWA-R dataset again compares the worst with the observations.
- In the location off of the coast of California, the ERA-40 dataset compares slightly better with the observations than the CS01 and the PWA-R dataset for 1997. For 1994, the error statistics of ERA-40 are the worst. This is because the only observations available for 1994 are winter observations, so the mean of the data is higher, and because ERA-40 systematically underestimates high values the underestimation is more severe.

2) Comparison with TOPEX/Poseidon altimeter measurements

As was the case for the wind speed comparisons, the quality of the H_s datasets was assessed globally by comparing the reanalysis data with the TOPEX/Poseidon altimeter observations. For the validation and comparison of the NCEP–NCAR reanalysis and the ERA-40 data, we have considered three latitude bands in our comparisons: north of 20°N, south of 20°S, and the region between 20°S and 20°N. For the comparisons with the AES40 dataset we considered two North Atlantic regions: one north of 20°N and the other between the equator and 20°N. For the comparisons with the PWA-R dataset we considered three Pacific Ocean regions: north of 20°N, south of 20°S, and the region between 20°S and 20°N. The statistics of the different reanalysis comparisons with the TOPEX/Poseidon measurements are presented in Table 11. Analyzing the statistics presented in the table, we can see that the ERA-40 data compare better with observations than the CS01 dataset in terms of rmse and SI. The CS01 data, however, do not have so much underestimation of high waves (high observations means) as the ERA-40 data (see Fig. 2). In the North Atlantic the AES40 results generally compare better with the observations; however, the ERA-40 data show also a low scatter index and rms error at low latitudes, especially for 1997. The PWA-R data show the worst correspondence with the measurements.

The Q–Q plots comparing the reanalysis data with the TOPEX/Poseidon observations help in visualizing the differences and deficiencies of the datasets. Figure 2 shows the Q–Q plots of the TOPEX/Poseidon significant wave height observations and the corresponding reanalysis data for 1994 and 1997. The ERA-40 and CS01 data are global, the AES40 data are for the North Atlantic Ocean, and the PWA-R data are for the Pacific Ocean. The ERA-40 dataset shows the most severe differences in distributions relative to the observations; it underestimates most of the high peaks of significant wave height and shows good correspondence only with the observations at low sea states. This underestimation is, to some extent, due to the respective underestimation

TABLE 10. The same as Table 7, but for 1997.

Region	n	$\bar{x}$	Reanalysis	Bias	Rmse	SI	ρ
Hawaiian Islands	5569	2.37	ERA-40	−0.16	0.35	0.13	0.90
			CS01	−0.31	0.48	0.15	0.85
			PWA-R	−0.37	0.58	0.19	0.84
Gulf of Mexico	3671	1.09	ERA-40	−0.09	0.31	0.27	0.92
			CS01	0.27	0.41	0.29	0.90
Northwest Atlantic	4774	1.74	ERA-40	−0.15	0.42	0.23	0.94
			CS01	−0.01	0.47	0.27	0.88
			AES40	0.05	0.37	0.21	0.92
Alaska	3788	2.87	ERA-40	−0.21	0.50	0.16	0.96
			CS01	0.20	0.65	0.22	0.92
			PWA-R	−0.22	0.75	0.26	0.91
Northeast Pacific	1911	2.89	ERA-40	−0.14	0.45	0.15	0.95
			CS01	0.19	0.60	0.20	0.92
			PWA-R	−0.02	0.63	0.23	0.92
California	1460	2.59	ERA-40	−0.13	0.41	0.15	0.95
			CS01	0.01	0.45	0.17	0.92
			PWA-R	−0.08	0.51	0.20	0.92

of high wind speeds (Fig. 1) as was already noted in Caires and Sterl (2003). The behavior of CS01 data in terms of the H_s range is much harder to identify. There are ranges of underestimation (H_s values around 2 m and above 6 m) and of overestimation (H_s values around 4.5 m). The AES40 data show quite a good correspondence with the observations, apart from some underestimation of H_s values above 3 m. The data of PWA-R have a negative bias for sea states with H_s below 6 m and a positive bias above this value. The PWA-R and the CS01 data show the same error characteristics in 1994 and 1997. The ERA-40 H_s data between 2 and 3 m compare much better with observations in 1997 than in 1994, and the AES40 data show more underestimation of high H_s values in 1994 than in 1997.

3) SUMMARY

In general terms the AES40 dataset is the one that compares better with observations in the North Atlantic. The two wave datasets produced with the NCEP–NCAR reanalysis winds in the Pacific seem to be of a similar quality, but the time series compare quite differently: there are periods of overestimation, for instance, in the CS01 data, corresponding to periods of underestimation in the PWA-R dataset. The ERA-40 data quality is better for 1997 than for the previous years, which seems to be a direct result of the assimilation of the *ERS-2* H_s altimeter measurements in 1997. There are some indications that the CS01 and AES40 data compare worse with the H_s observations for 1978 than for the following years. In terms of error statistics, the ERA-40 dataset compares better than the CS01 dataset with observations, and worse than the AES40 dataset, except for 1997, when especially the ERA-40 observations at low latitudes compare quite well with observations. In terms of distribution, the ERA-40 dataset compares worse with observations at high values and the PWA-R dataset at low values. The data produced by forcing the ODGP2 spectral wave model with the ERA-40 wind compare generally better with the buoy observations than the ERA-40 data. A close look at the time series shows that the ERA-40/ODGP2 data capture high H_s peaks better than do the ERA-40 data.

5. Comparison of monthly means

We have produced and analyzed surface plots with the relative differences between the H_s and U_{10} data of different reanalyses for the years of 1978, 1988, 1994, and 1997, and have applied the Mann–Whitney test to find where the differences were significant at a 5% level.

Figure 3 presents surface plots of the relative differences between the various wind reanalyses for December 1997: there are minor differences in the plots for different months, but the chosen example is representative of the extent and location of the differences found. The ERA-40 and the NCEP–NCAR reanalysis data differ mainly in the Tropics, in the Southern Hemisphere (especially in 1978), and in coastal regions. There are

→

FIG. 2. Graphs comparing the 1%–99% quantiles of TOPEX/Poseidon observations against reanalysis H_s data. Data from (left) 1994 and (right) 1997: (top) ERA-40, (second) CS01, (third) AES40, and (bottom) PWA-R. The ERA-40 and CS01 data are global, the AES40 data are for the North Atlantic Ocean, and the PWA-R data are for the Pacific Ocean.

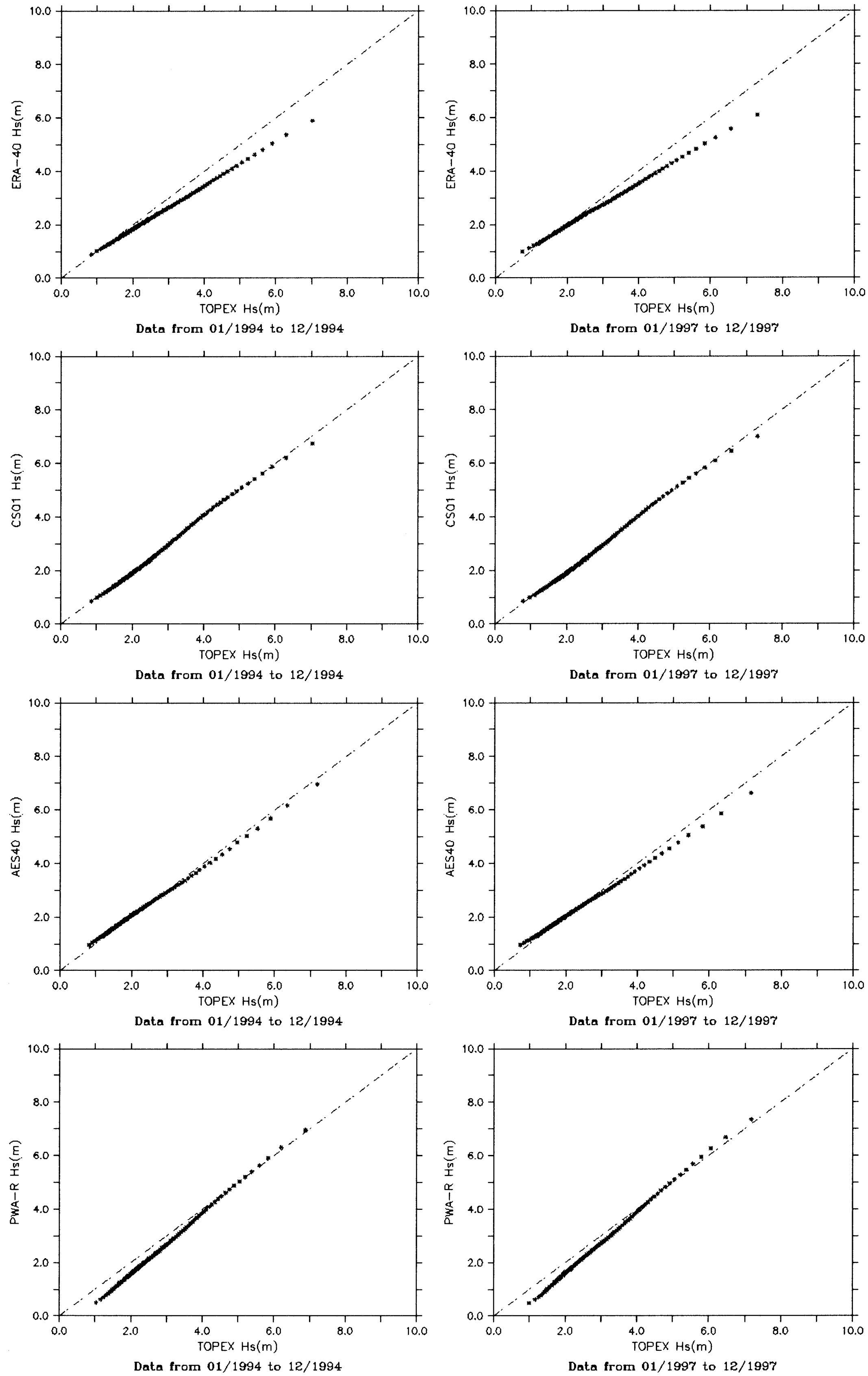
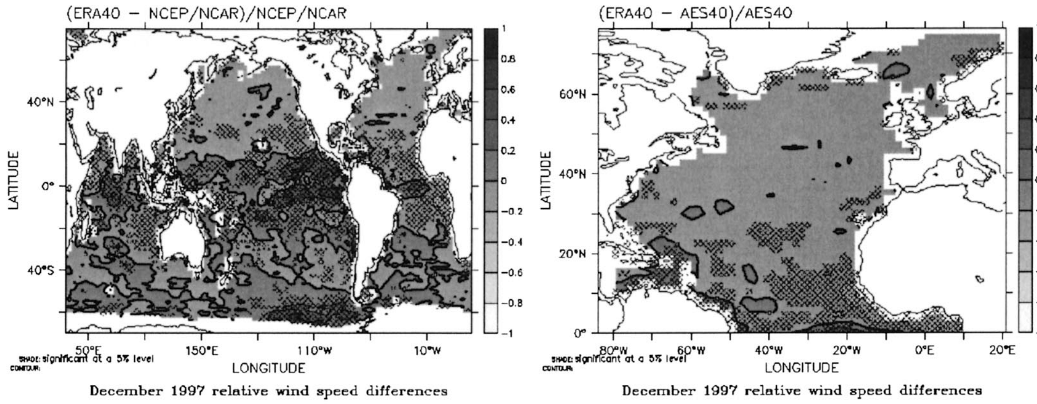

ERA-40 Hs(m)
TOPEX Hs(m)
Data from 01/1994 to 12/1994
ERA-40 Hs(m)
TOPEX Hs(m)
Data from 01/1997 to 12/1997
CS01 Hs(m)
TOPEX Hs(m)
Data from 01/1994 to 12/1994
CS01 Hs(m)
TOPEX Hs(m)
Data from 01/1997 to 12/1997
AES40 Hs(m)
TOPEX Hs(m)
Data from 01/1994 to 12/1994
AES40 Hs(m)
TOPEX Hs(m)
Data from 01/1997 to 12/1997
PWA-R Hs(m)
TOPEX Hs(m)
Data from 01/1994 to 12/1994
PWA-R Hs(m)
TOPEX Hs(m)
Data from 01/1997 to 12/1997

TABLE 11. Significant wave height (m) 1997 statistics of different reanalysis products vs TOPEX/Poseidon measurements in different latitude bands.

Region	n	$\bar{x}$	Reanalysis	Bias	Rmse	SI	ρ
			1994				
20°–80°N	94759	2.62	ERA-40	−0.34	0.55	0.16	0.96
	93563	2.62	CS01	0.03	0.55	0.21	0.94
20°S–20°N	147101	2.01	ERA-40	−0.17	0.32	0.13	0.92
	143517	2.02	CS01	−0.10	0.42	0.20	0.79
80°–20°S	226563	3.43	ERA-40	−0.41	0.61	0.13	0.95
	225898	3.43	CS01	−0.05	0.73	0.21	0.85
20°–80°N, 70°W–20°E	24712	2.67	ERA-40	−0.37	0.56	0.16	0.97
	34586	2.67	CS01	0.03	0.52	0.20	0.94
	34025	2.70	AES40	−0.02	0.47	0.17	0.95
0°–20°N, 70°W–20°E	14766	1.89	ERA-40	−0.21	0.30	0.11	0.91
	14545	1.89	CS01	0.08	0.32	0.16	0.79
	14832	1.89	AES40	0.13	0.30	0.14	0.85
20°–80°N, 115°E–90°W	56983	2.64	ERA-40	−0.32	0.54	0.16	0.95
	56212	2.65	CS01	0.02	0.57	0.22	0.91
	51597	2.67	PWA-R	−0.33	0.73	0.25	0.90
20°S–20°N, 115°E–70°W	8553	2.03	ERA-40	−0.17	0.31	0.13	0.91
	84002	2.04	CS01	−0.16	0.45	0.20	0.74
	81198	2.05	PWA-R	−0.36	0.53	0.19	0.81
40°–20°S, 115°E–70°W	41365	2.79	ERA-40	−0.34	0.54	0.15	0.93
	40415	2.79	CS01	−0.35	0.64	0.19	0.85
	40966	2.79	PWA-R	−0.36	0.69	0.21	0.85
			1997				
20°–80°N	91994	2.68	ERA-40	−0.25	0.54	0.18	0.96
	90712	2.69	CS01	−0.03	0.57	0.21	0.92
20°S–20°N	142563	2.03	ERA-40	−0.04	0.28	0.13	0.93
	139033	2.04	CS01	−0.09	0.41	0.20	0.81
80°–20°S	220409	3.51	ERA-40	−0.34	0.58	0.14	0.96
	219611	3.51	CS01	−0.05	0.75	0.21	0.86
20°–80°N, 70°W–20°E	33731	2.70	ERA-40	−0.28	0.56	0.18	0.96
	33557	2.70	CS01	−0.01	0.55	0.20	0.93
	33122	2.73	AES40	−0.09	0.49	0.18	0.95
0°–20°N, 70°W–20°E	14323	1.87	ERA-40	−0.08	0.25	0.13	0.91
	14108	1.87	CS01	0.03	0.34	0.18	0.79
	14388	1.86	AES40	0.11	0.30	0.15	0.86
20°–80°N, 115°E–90°W	55310	2.73	ERA-40	−0.23	0.53	0.18	0.96
	54488	2.74	CS01	−0.04	0.59	0.21	0.91
	49901	2.77	PWA-R	−0.35	0.76	0.24	0.91
20°S–20°N, 115°E–70°W	83004	2.09	ERA-40	−0.04	0.28	0.13	0.91
	81432	2.10	CS01	−0.16	0.44	0.20	0.76
	78696	2.16	PWA-R	−0.29	0.52	0.20	0.79
40°–20°S, 115°E–70°W	40031	2.80	ERA-40	−0.21	0.46	0.15	0.94
	39126	2.80	CS01	−0.32	0.61	0.18	0.86
	39622	2.80	PWA-R	−0.29	0.66	0.21	0.85

no significant differences in the northern storm tracks. In the comparisons between ERA-40 and AES40 data the differences are mainly south of 30°N and in the coastal regions. Between the NCEP–NCAR reanalysis and the AES40 data, no pattern can be identified in the differences; they are significant only in the coastal locations, and in some locations south of 30°N in the summer months, especially in 1997.

Figure 4 presents surface plots of the relative differences between the various wave reanalyses for December 1997, the same month for which the differences in wind fields are exemplified. In spite of the fact that the differences between the ERA-40 and the NCEP–NCAR reanalysis winds are only significant south of 20°N, the ERA-40 and CS01 H_s fields are significantly different almost everywhere. Even the CS01 and the PWA-R H_s fields, which were produced from the same wind fields, are significantly different everywhere. The only fields that show some agreement are CS01 and AES40, with most of the differences being south of 30°N. These differences are due, at least in part and in the summer months, to the explicit treatment of tropical storms in the creation of the AES40 data, whereby the tropical storm wind fields are rigorously reanalyzed, using National Hurricane Center high quality, high-resolution reconnaissance data, and incorporated into the kinematic



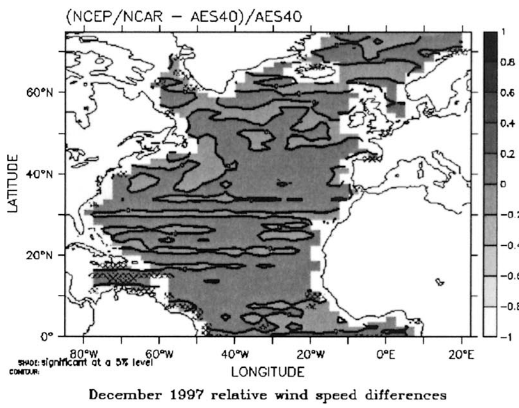


FIG. 3. Surface plots of the relative differences between the Dec 1997 monthly means of the wind speeds from the different reanalysis, with the regions where the differences are significant at a 5% level (shaded).

analysis. Another contributing factor is the use of scatterometer winds in the kinematic analysis that helps correct the initially poor NCEP–NCAR reanalysis winds in the Tropics.

6. Long time-scale features

We have used the monthly mean fields of the different reanalyses to analyze the trends in monthly means from 1958 to 1967 and from 1990 to 1997. This was done in the same way as described in Wang and Swail (2001): the Mann–Kendall nonparametric test was used to identify the significant results—namely the existence of trends—at a 5% level, and trend estimates were obtained from Kendall's rank correlation. Figure 5 describes the trends in the January means of U_{10} from the different reanalyses during the period of 1990–97. These results are representative of what happens in the other months: the trend patterns and the regions where trends were detected are essentially the same in the ERA-40, NCEP–NCAR reanalysis, and AES40 wind fields.

In order to synthesize the comparisons at long time scales, and to compare the monthly and annual variability of the reanalysis datasets, we present in Figs. 6 and 7 the correlations between the different U_{10} datasets for the periods of 1957–68 and 1990–97, and the locations at which the Mann–Whitney test rejects the equality of means at the 5% significance level. For these calculations the annual cycle was removed from the data. The correlations between the NCEP–NCAR reanalysis and the AES40 winds are in both periods above 0.8 everywhere, for which reason we have omitted the respective figures. The correlations between the ERA-40 and the NCEP–NCAR reanalysis data for the period of 1957–68 are below 0.8 in the Tropics and Southern Hemisphere; in the period of 1990–97 the regions with lower correlations are restricted to the Tropics. Nowhere does the Mann–Whitney test give significant results when comparing the ERA-40 and NCEP–NCAR reanalysis data. The correlations between the ERA-40 and the AES40 wind fields are above 0.8, except for some locations in the Tropics.

We now analyze the long-time-scale features in the H_s datasets: For all calendar months of the 1990–97 period, the trends in the PWA-R, CS01, and AES40 data have a similar pattern, but the trends of the latter are somewhat more pronounced. The trends of the ERA-40 data have a different pattern from that of the other reanalyses—there is a caveat in these trends because the ERA-40 significant wave height is corrupted from January 1992 to May 1993—but the areas with significant results tend to be the same. These conclusions are illustrated in Fig. 8, which presents the trends in the January means of H_s from the different reanalyses during the period of 1990–97. There are some discrepancies between the trends of the CS01 and ERA-40 datasets

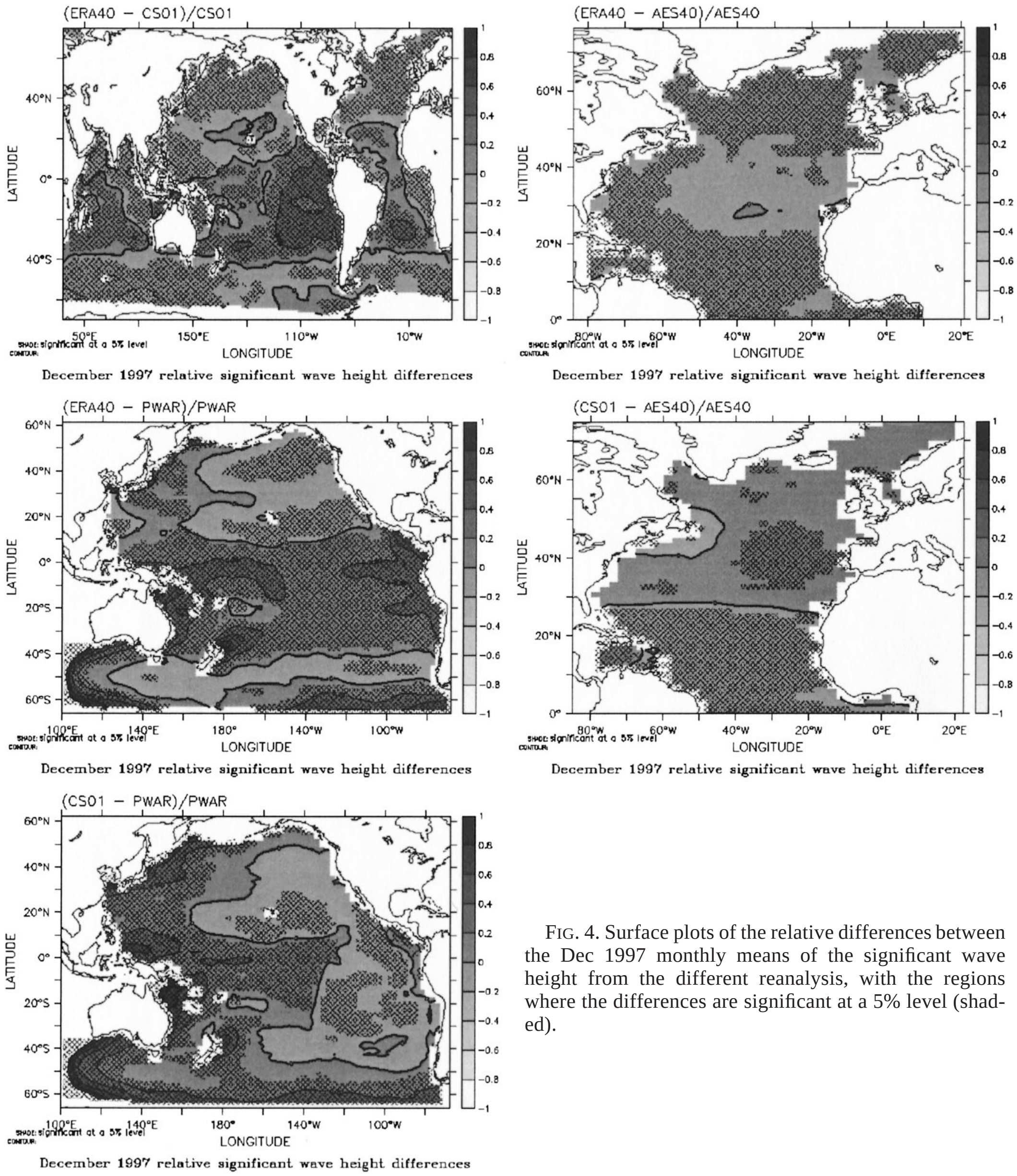


FIG. 4. Surface plots of the relative differences between the Dec 1997 monthly means of the significant wave height from the different reanalysis, with the regions where the differences are significant at a 5% level (shaded).

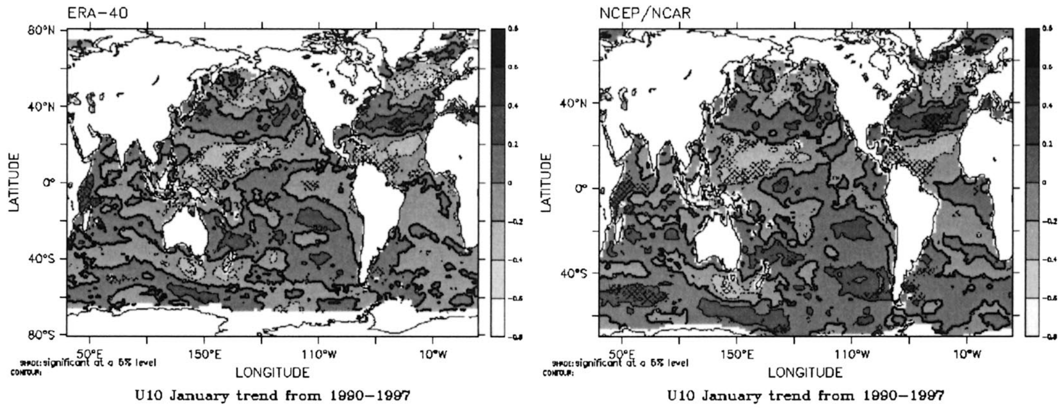
for the Southern Hemisphere during the 1958–67 period (not shown).

Figures 9 and 10 present the correlations between the different reanalysis datasets of H_s monthly anomalies for the periods of 1957–68 and 1990–97, respectively. The correlations between the ERA-40 and CS01 data in the first period are below 0.8 in most of the Tropics and Southern Hemisphere. In the second period, correlations below 0.8 are restricted mainly to the Tropics; in contrast with the first period, the correlations are high also in the Southern Hemisphere storm track region. The correlations between the AES40 and the ERA-40 and CS01 data are below 0.8 only in some boundary locations and in the Tropics. The correlations between the PWA-R dataset and the other datasets are available only for the 1990–97 period; high correlations with the ERA-40 are found only in the North Pacific and Southern Hemisphere storm tracks. Correlations between the PWA-R and the CS01 datasets below 0.8 are present only in the Tropics.

7. Discussion and recommendations

We have collected, assessed, and compared the wind speed and significant wave height data from several reanalyses.

The data were assessed against time-averaged altimeter and buoy measurements. As a rough approximation,



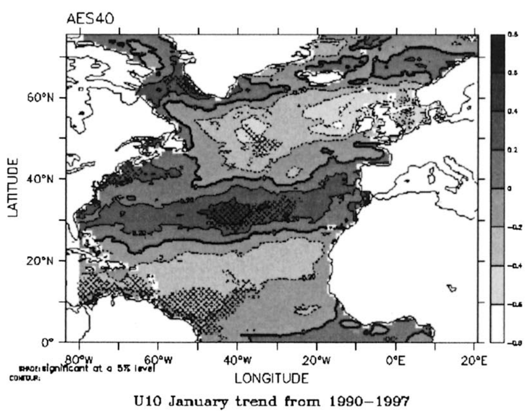


FIG. 5. Surface plots of the trend (m s^{-1} yr^{-1}) in the Jan monthly mean wind speed data of the different reanalyses from 1990 to 1997, with the areas where the trend is significant at a 5% level (shaded).

the ERA-40 grid resolution was used as a reference for the averaging of the measurements. The use of this approximation will not seriously affect the assessment of the other reanalysis data; at most, one might expect occasional and small underestimations of high peaks by the PWA-R and CS01 wave data and the NCEP–NCAR reanalysis winds, and some small overestimations of high peaks by the AES40 waves and winds.

Our assessment indicates that the AES40 data, which are restricted to the North Atlantic, best represent the measurements within that basin. The ERA-40 data also compare well with the observations, and generally better than the other reanalyses in terms of standard statistical measures. This is especially true for the H_s data for 1997, a period where *ERS-2* altimeter measurements are assimilated and in which the statistics are comparable with those obtained for the AES40 data, especially in the low latitudes. The results of forcing the ODGP2 spectral wave model with the ERA-40 winds show that the use of another wave model forced with the ERA-40 winds can produce waves that compare better with observations. The same conclusion is drawn from the comparisons between the CS01 and PWA-R datasets, which were produced using the same wind fields to force

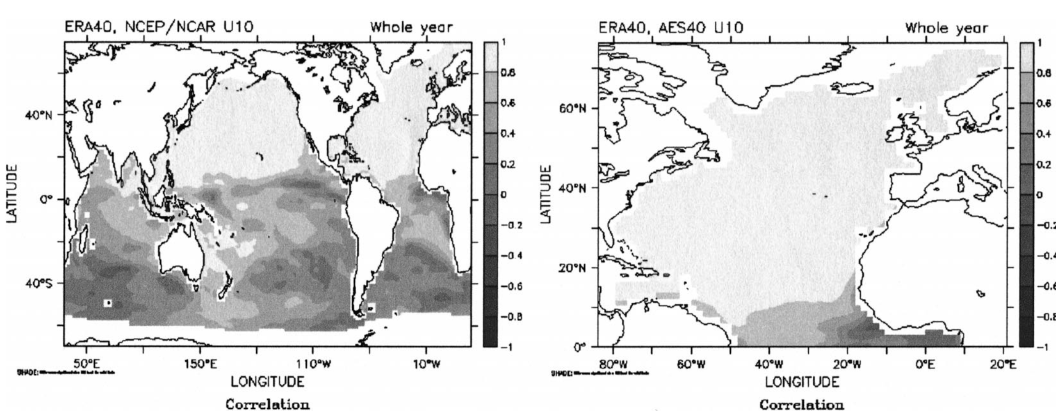


FIG. 6. Correlation between the different U_{10} reanalysis monthly means anomalies from 1958 to 1967.

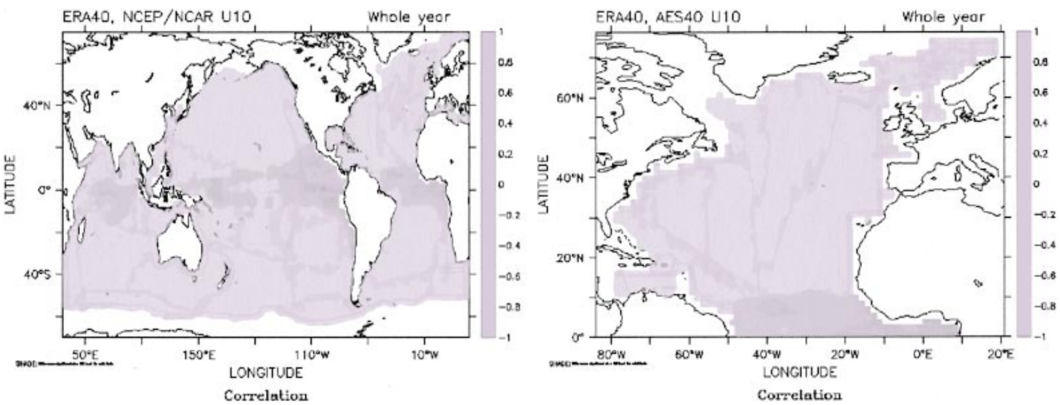


FIG. 7. The same as Fig. 6, but for data from 1990 to 1997.

different wave models, and whose results are quite distinct—not only at short time scales but also in terms of monthly means. It should be emphasized that the dataset of PWA-R was tuned in order to give the best results for major winter swell events from 1981 to 1998; this tuning improved the agreement with the observations for high waves, but made it worse for low waves.

An interesting feature of the comparison of wind speed data is the existence of large differences between the reanalyses in the Tropics; another is the fact that the differences are usually larger in the Southern than in the Northern Hemisphere, testifying to the present limitations of modeling those regions. The latter problem is not only due to the lack of measurements in those regions, but also to some deficiencies in the physical description of the processes, because the results also differ in the 1990s, when both reanalysis models benefit from the assimilation of satellite measurements in those regions.

In our study the validation of 6-hourly fields was

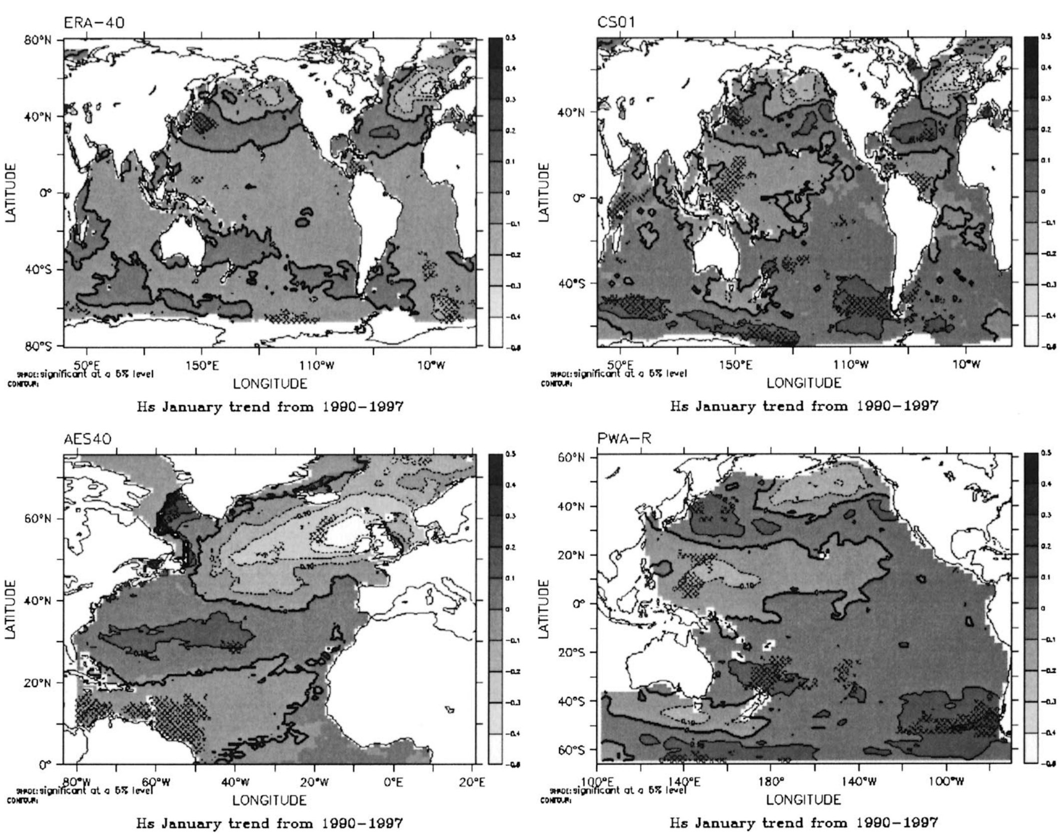
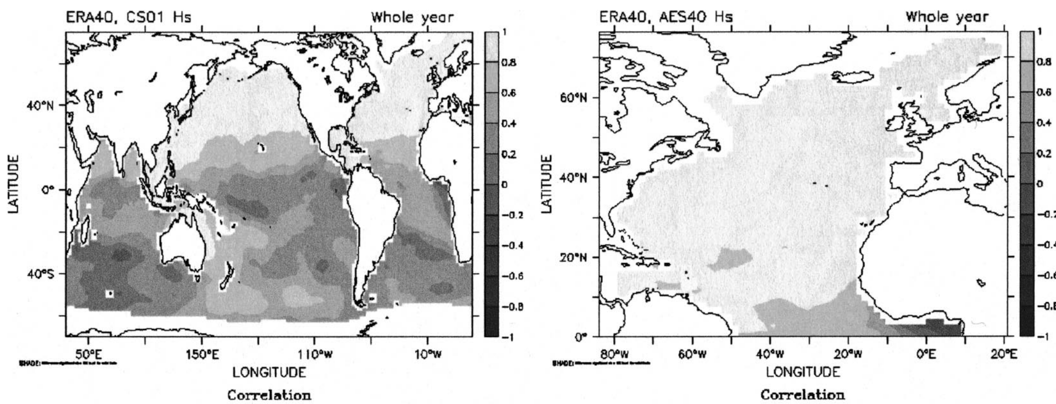


FIG. 8. Surface plots of the trend (m yr^{-1}) in the Jan monthly mean significant wave height data of the different reanalyses from 1990 to 1997, with the areas where the trend is significant at a 5% level (shaded).



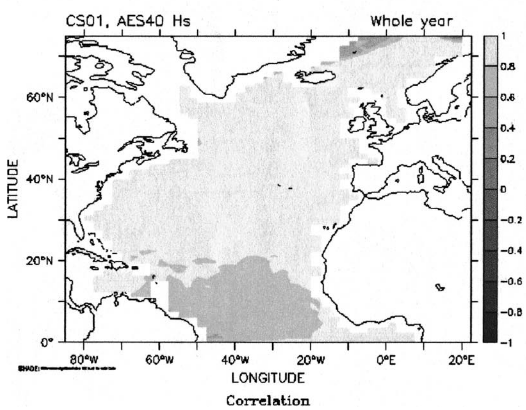


FIG. 9. Correlation between the different H_s reanalysis monthly means anomalies from 1958 to 1967.

restricted to 4 yr: 1978, 1988, 1994, and 1997. These years were, however, strategically chosen to give a description of all of the error characteristics of the datasets. This is because error characteristics are only expected to change due to changes in the reanalysis observing systems, because the models and assimilation techniques do not change. The results we have obtained for the 1978 data characterize the errors of the datasets until the mid-1980s, the results obtained for 1988 characterize the errors of the datasets until the use of *ERS-1* observations, and the results obtained for 1994 and 1997 characterize the errors of the datasets during the assimilation of the *ERS-1* and *ERS-2* observations, respectively.

At short time scales, the differences between the various reanalysis datasets of winds and waves are large. In terms of monthly means, the differences between the wind fields of AES40 and CS01 are almost nowhere significant, and the ERA-40 monthly means differ from those datasets mainly south of 30°N. The various H_s datasets differ both in short and medium time scales. On the other hand, the long-time-scale behavior of both winds and waves in the various datasets, as measured by trends, is quite similar—an indication that the long-time-scale features are equally present in all datasets. Regarding long-term variability of the data, discrepancies occur in the Tropics during the two periods considered, and in the Southern Hemisphere during the initial period.

The lessons/recommendations to be taken from this study are the following:

- For detailed descriptions of U_{10} and H_s data and their variability, we recommend using the ERA-40 dataset for global studies and the AES40 dataset for studies restricted to the North Atlantic.
- For studies in which a good description of high H_s values is important, the PWA-R or the CS01 datasets should be chosen instead of the ERA-40 data, which do quite a poor job in depicting high quantiles.
- The quality of the H_s fields from the reanalyses can still be improved by using better-tuned wave models.
- For studies of long-term trends in the datasets it is not important which dataset is chosen because long-term features seem to be equally present in all datasets.
- Although in terms of the usual statistical measures the ERA-40 data compare better than the NCEP–NCAR reanalysis data with the observations in the Tropics, extra attention should be given to the validation and, if possible, correction of any of these datasets in studies involving data from that region because their U_{10} fields differ significantly.

Acknowledgments. We are thankful to Helen Snaith of the Southampton Oceanography Centre for the altim-

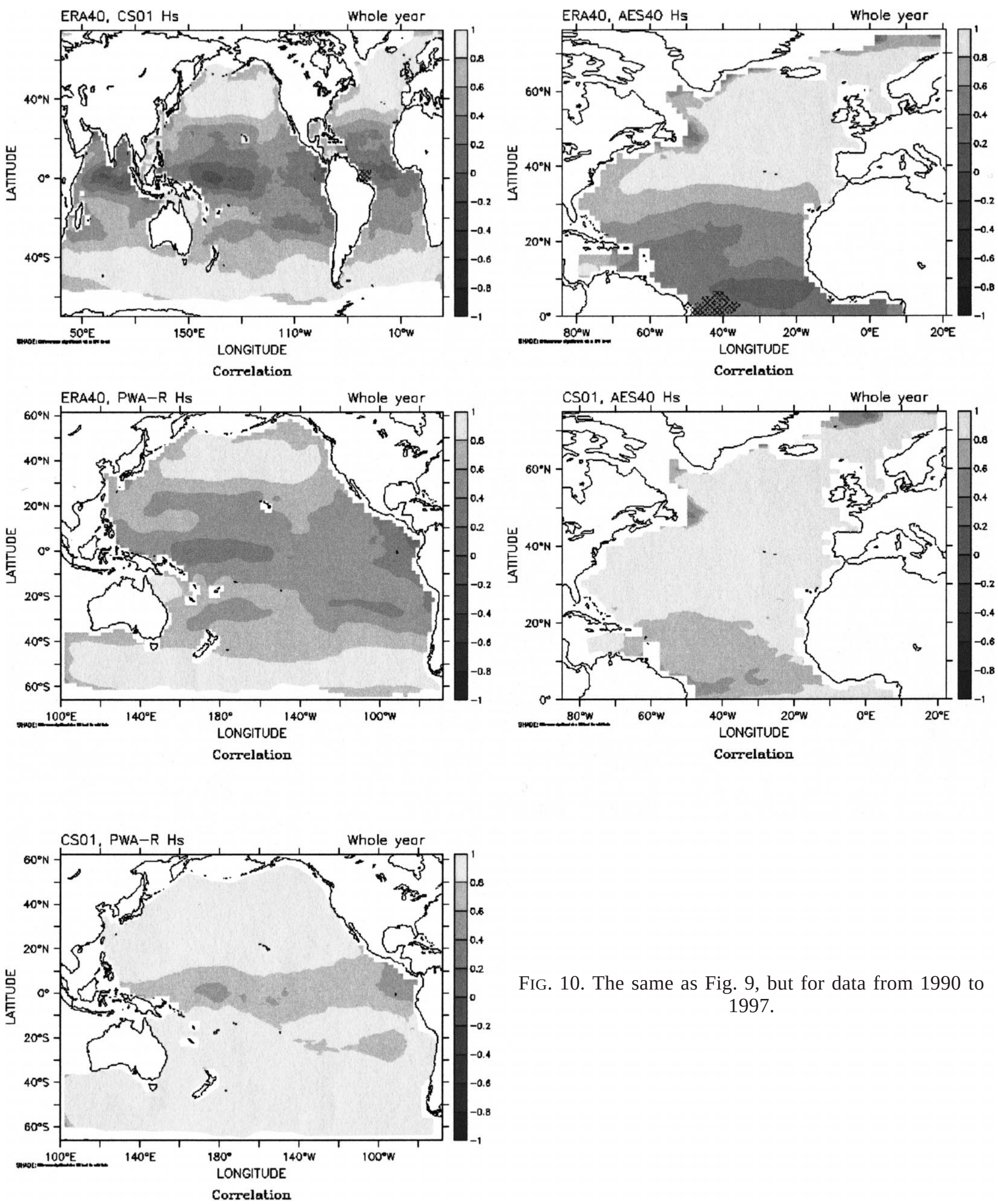


FIG. 10. The same as Fig. 9, but for data from 1990 to 1997.

eter data, to the American National Oceanographic Data Center for the available buoy data, and to Camiel Severijns for the field library and technical support. The two referees' comments/suggestions, which greatly improved our initial manuscript, are also acknowledged. We also thank INTAS (Grant 96-2089) for facilitating a discussion with Sergey Gulev, David Woolf, and Roman Bortkovsky. This work was funded by the EU-funded ERA-40 Project (No. EVK2-CT-1999-00027).

REFERENCES

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

Caires, S., and A. Sterl, 2001: Comparative assessment of ERA-40 ocean wave product. *Proc. ECMWF Workshop on Re-analysis,* Reading, United Kingdom, ECMWF, ERA-40 Project Report Series 3, 353–368.

——, and ——, 2003: Validation of ocean wind and wave data using triple collocation. *J. Geophys. Res.,* **108**, 3098, doi:10.1029/2002JC001491.

Challenor, P., and D. Cotton, cited 1999: Trends in TOPEX significant wave height measurements. [Available online at http://www.soc.soton.ac.uk/JRD/SAT/TOPtren/TOPtren.pdf.]

Cotton, P. D., and D. J. T. Carter, 1994: Cross-calibration of TOPEX, ERS-1, and GEOSAT wave heights. *J. Geophys. Res.,* **99** (C12), 25 025–25 033.

Cox, A. T., and V. R. Swail, 2001: A global wave hindcast over the period 1958–1997: Validation and climate assessment. *J. Geophys. Res.,* **106** (C2), 2313–2329.

Gibson, J. K., P. Kållberg, S. Uppala, A. Nomura, A. Hernandez, and E. Serrano, 1997: ERA description. ECMWF Re-analysis Final Rep. Series 1, 71 pp.

Gilhousen, D., 1999: Improvement in National Buoy Center measurements. *Achievements in Marine Climatology,* V. Swail, Ed., Environment-Canada, 79–89.

Gourrion, J., D. Vandemark, S. Bailey, B. Chapron, C. P. Gommenginger, P. G. Challenor, and M. Srokosz, 2002: A two parameter wind speed algorithm for Ku-band altimeters. *J. Atmos. Oceanic Technol.,* **19,** 2030–2048.

Graham, N. E., and H. F. Diaz, 2001: Evidence of intensification of North Pacific winter cyclones since 1948. *Bull. Amer. Meteor. Soc.,* **82,** 1869–1893.

——, R. R. Strange, and H. F. Diaz, 2002: Intensification of North Pacific winter cyclones 1948–98: Impacts on California wave climate. *Proc. 7th Int. Workshop on Wave Hindcasting and Forecasting,* Banff, AB, Canada, U.S. Army Engineer Research and Development Center, 60–69.

Gulev, S. K., V. Grigorieva, A. Sterl, and D. Woolf, 2003: Assessment of the reliability of wave observations from voluntary observing ships: Insights from the validation of a global wind wave climatology based on voluntary observing ship data. *J. Geophys. Res.,* **108**, 3236, doi:10.1029/2002JC001437.

Janssen, P. A. E. M., D. Doyle, J. Bidlot, B. Hansen, L. Isaksen, and P. Viterbo, 2002: Impact and feedback of ocean waves on the atmosphere. *Atmosphere–Ocean Interactions*, W. Perrie, Ed., Vol. 1, *Advances in Fluid Mechanics*, WIT Press, 155–197.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.,* **77,** 437–471.

Kistler, R., and Coauthors, 2001: The NCEP–NCAR 50-year reanalysis: Monthly means CD-ROM and documentation. *Bull. Amer. Meteor. Soc.,* **82,** 247–267.

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.

Mood, A. M., F. A. Graybill, and D. C. Boes, 1974: *Introduction to the Theory of Statistics.* McGraw-Hill, 564 pp.

Snaith, H. M., 2000: Global altimeter processing scheme user manual: v1. Southampton Oceanography Centre Internal Document 53, Southampton, United Kingdom, 44 pp.

Sterl, A., G. J. Komen, and P. D. Cotton, 1998: Fifteen years of global wave hindcasts using winds from the European Centre for Medium-Range Weather Forecast reanalysis: Validating the reanalyzed winds and assessing the wave climate. *J. Geophys. Res.,* **103** (C3), 5477–5494.

Swail, V. R., and A. T. Cox, 2000: On the use of NCEP–NCAR reanalysis surface marine wind fields for a long-term North Atlantic wave hindcast. *J. Atmos. Oceanic Technol.,* **17,** 532–545.

Tolman, H. L., 1999: User manual and system documentation of WAVEWATCH III version 1.18. NOAA/NWS/NCEP Ocean Modeling Branch Contribution 166, 112 pp.

——, 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 models at NCEP. *Wea. Forecasting,* **17,** 311–333.

Uppala, S., 2001: ECMWF reanalysis, 1957–2001. *Proc. ECMWF Workshop on Reanalysis,* Reading, United Kingdom, ECMWF, ERA-40 Project Report Series 1–10.

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

Wang, X. L., and V. R. Swail, 2001: Changes of extreme wave heights in Northern Hemisphere oceans and related atmospheric circulation regimes. *J. Climate,* **14,** 2204–2221.

——, and ——, 2002: Trends of Atlantic wave extremes as simulated in a 40-yr wave hindcast using kinematically reanalyzed wind fields. *J. Climate,* **15,** 1020–1035.

Woodruff, S. D., H. F. Diaz, J. D. Elms, and S. J. Worley, 1998: COADS Release 2 data and metadata enhancements for improvements of marine surface flux fields. *Phys. Chem. Earth,* **23,** 517–527.

Young, I. R., 1993: An estimate of the Geosat altimeter wind speed algorithm at high wind speeds. *J. Geophys. Res.,* **98,** 20 275–20 285.