

## A Two-Parameter Wind Speed Algorithm for Ku-Band Altimeters

J. GOURRION

*IFREMER, Département d'Océanographie Spatiale, Plouzane, France*

D. VANDEMARK AND S. BAILEY

*NASA Goddard Space Flight Center, Wallops Island, Virginia*

B. CHAPRON

*IFREMER, Département d'Océanographie Spatiale, Plouzane, France*

G. P. GOMMENGINGER, P. G. CHALLENGOR, AND M. A. SROKOSZ

*The James Rennell Division, Southampton Oceanography Centre, Southampton, United Kingdom*

(Manuscript received 5 July 2001, in final form 15 March 2002)

### ABSTRACT

Globally distributed crossovers of altimeter and scatterometer observations clearly demonstrate that ocean altimeter backscatter correlates with both the near-surface wind speed and the sea state. Satellite data from TOPEX/Poseidon and NSCAT are used to develop an empirical altimeter wind speed model that attenuates the sea-state signature and improves upon the present operational altimeter wind model. The inversion is defined using a multilayer perceptron neural network with altimeter-derived backscatter and significant wave height as inputs. Comparisons between this new model and past single input routines indicates that the rms wind error is reduced by 10%–15% in tandem with the lowering of wind error residuals dependent on the sea state. Both model intercomparison and validation of the new routine are detailed, including the use of large independent data compilations that include the SeaWinds and ERS scatterometers, ECMWF wind fields, and buoy measurements. The model provides consistent improvement against these varied sources with a wind-independent bias below  $0.3 \text{ m s}^{-1}$ . The continuous form of the defined function, along with the global data used in its derivation, suggest an algorithm suitable for operational application to Ku-band altimeters. Further model improvement through wave height inclusion is limited due to an inherent multivaluedness between any single realization of the altimeter measurement pair  $[\sigma^o, H_s]$  and observed near-surface winds. This ambiguity indicates that  $H_s$  is a limited proxy for variable gravity wave properties that impact upon altimeter backscatter.

### 1. Introduction

Empirical models have been devised to improve satellite altimeter ocean wind speed retrieval using many differing numerical approaches and datasets (Brown et al. 1981; Dobson et al. 1987; Chelton and McCabe 1985; Witter and Chelton 1991; Glazman and Greysukh 1993; Young 1993; Lefèvre et al. 1994; Freilich and Challenor 1994). The global altimeter ocean wind product is mostly limited to validation and climatological usage (e.g., Young 1999) because the altimeter's nadir-pointing geometry only permits estimates of surface wind speed along a narrow ( $\approx 2 \text{ km}$ ) swath and precludes wind direction detection. However, accurate wind speed estimates are also important because they are used in the

point-by-point correction of an altimeter's estimate of mean sea surface height via the electromagnetic bias algorithm. Freilich and Challenor (1994) and Glazman and Greysukh (1993) expand on these points and suggest that objective model improvement metrics should include the minimization of wind speed biases and root-mean-square error, removal of nonwind geophysical impacts such as sea state, and functional continuity (finite first derivative) such that the wind speed histogram is not distorted.

The operational altimeter wind speed product for TOPEX is derived from interpolation over the look up table known as the modified Chelton–Wentz algorithm (MCW) of Witter and Chelton (1991). This model directly maps measured Ku-band altimeter backscatter ( $\sigma^o$ ) to the wind speed 10 m above the ocean ( $U_{10}$ ). The overall bias,  $\langle U_{\text{err}} \rangle$ , for this algorithm is suggested to be  $0.48 \text{ m s}^{-1}$  ( $U_{\text{err}} = U_{\text{altim}} - U_{\text{insitu}}$ ) and the root-mean-

---

*Corresponding author address:* Mr. Douglas C. Vandemark, NASA GSFC, Bldg. N-159, Room E222, Wallops Island, VA 23337.  
E-mail: vandemark@gsfc.nasa.gov

square (rms) error lies between 1.5 and 2.0  $\text{m s}^{-1}$  (e.g., Witter and Chelton 1991; Gower 1996; Freilich and Challenor 1994; Wu 1999). Numerous studies have suggested that the form for this single parameter algorithm could be improved upon (e.g., Freilich and Challenor 1994). However, the limited amount of validation data combined with the generally small level of improvement in algorithm performance leaves MCW as the current choice for new altimeters such as the GEOSAT Follow-On (GFO) and *Jason-1*.

A goal that remains of interest for altimeter wind retrieval is the detection and correction of wind speed errors associated with longer ocean waves that are not necessarily closely coupled to the local wind field. Evidence for a sea-state effect on altimeter-derived wind has been addressed in several studies (Monaldo and Dobson 1989; Glazman and Pilorz 1990; Glazman and Greysukh 1993; Lefèvre et al. 1994; Freilich and Challenor 1994; Hwang et al. 1998). Reported results range from substantial impacts to no impact (cf. Wu 1999). The central and unique factor here is an altimeter's coincident and accurate measure of significant wave height,  $H_s$ .

Motivation for the present study follows from compilation of a large number of TOPEX altimeter observations made coincident with National Aeronautics and Space Administration (NASA) scatterometer (NSCAT) surface wind estimates. The global coverage, fidelity, and volume of this dataset leads to a much clearer picture of  $H_s$  variation impacts upon altimeter backscatter and wind inversion over a range of wind speeds from 1 to 20  $\text{m s}^{-1}$ . We develop two models, a forward and an inverse solution, using neural network methods to map between altimeter and scatterometer observations and incorporating a globally derived correction for sea-state impacts using the altimeter-derived  $H_s$  estimate. Numerous independent datasets containing collocation between altimeter and ancillary wind estimates are then used to evaluate a best-choice routine and its applicability for operational usage. This paper is derived from an earlier report by Gourrion et al. (2000). That effort includes wind speed models for the C-band altimeter aboard TOPEX, but the present effort concentrates solely on Ku-band model definition.

## 2. Collocated datasets for training and validation

A choice is usually made in wind speed algorithm developments between evaluating a limited amount of measurements collected over fixed-location ocean buoys and methods that utilize global wind statistics as the basis for tailoring the model (cf. Witter and Chelton 1991; Freilich and Challenor 1994). A new option has arrived because ocean observing satellites are now often in orbit at the same time. Many of these satellites share wind observing attributes that can be exploited when their respective ocean footprints cross. This study follows the collocation approach and we make use of the

unprecedented ability to combine TOPEX/Poseidon (termed TOPEX hereafter) satellite altimeter measurements with those made at the same time and place by satellite scatterometers. The resulting crossover datasets are very large and form the basis for our algorithm development and validation. The TOPEX/NSCAT dataset is chosen for algorithm training and will be detailed below. Six additional datasets are assembled for subsequent altimeter wind speed model validation. These data sets contain similar TOPEX crossover compilations using winds from other satellites [European Remote-Sensing satellite (ERS) and SeaWinds], surface models [European Centre for Medium-Range Weather Forecasts (ECMWF)], and buoys. We also include a dataset from the *ERS-2* altimeter. The varied data sources offer the opportunity to assess altimeter algorithm stability with respect to overall wind bias and rms errors, as well as removal of wave height impacts. Moreover, inherent dataset differences provide insight into issues such as intercomparison noise sources and time-dependent sensor calibration adjustments.

### a. TOPEX and NSCAT

One central assumption in this study is that surface wind estimates obtained using NSCAT can serve as the reference for those points where near-simultaneous TOPEX observations occurred during the NSCAT mission, September 1996 to June 1997. The collocated dataset has the attributes of high spatial and temporal correlation, large data population, and global coverage.

Past studies and the known characteristics of wind wave growth suggest that long wave impacts may vary over fairly small spatial scales of order 10–100 km. For instance, the most common fetch for a wind sea on the ocean is about 70 km (Tournadre and Blanquet 1994). Not surprisingly, the horizontal length scale for atmospheric fronts falls into this range as well. Thus to probe the long wave impacts on altimeter winds, we desire a finescale reference wind product. An attractive feature for the case of NSCAT is the availability of a special high resolution scatterometer wind product produced by S. Dunbar at the Jet Propulsion Laboratory [High Resolution-Merged Geophysical Data Records (HR-MGDR); Dunbar 1997]. The surface resolution is reduced to  $25 \times 25$  km for each wind estimate, a so-called wind vector cell.

Wind retrieval test results can differ depending on intercomparison product resolutions and spatial and temporal separation. This fact is already well documented from an altimeter wind inversion standpoint (e.g., Monaldo and Dobson 1989; Gower 1996; Hwang et al. 1998; Freilich and Dunbar 1993). To keep intercomparison errors to a minimum we desire equal sensor surface coverage between the altimeter and scatterometer. The high-resolution NSCAT product is attractive for this purpose as the TOPEX wind resolution cell is relatively small ( $O(2 \times 6)$  km). Larger footprint ( $50 \times$

50 km) products such as from Special Sensor Microwave Imager (SSM/I), NSCAT, or *ERS-1* or *-2* encompass four times greater area.

The model function used to derive NSCAT wind speed is NSCAT-1 (Wentz and Smith 1999). A new algorithm, NSCAT-2, is now in existence. Its inclusion into our dataset is not straightforward but changes between NSCAT-1 and 2 have been evaluated and found to be small in their wind speed impacts for the data range ( $1\text{--}20\text{ m s}^{-1}$ ) of interest here. Multiple independent validations presented in section 6 suggest this is the case. There we include evaluation of results against a collocation of TOPEX with NASA's Ku-band SeaWinds scatterometer that uses the QSCAT-1 model function to derive its wind. The potential for a nearly constant  $0.3\text{ m s}^{-1}$  NSCAT bias is discussed in that section, but data within this study do not confirm or deny this discrepancy. The largest distinction between NSCAT-1 and QSCAT-1 models comes at very high wind speeds, above  $15\text{--}20\text{ m s}^{-1}$ . An appendix addresses altimeter inversion for speeds above  $20\text{ m s}^{-1}$ .

A given collocation point within this dataset can be characterized as follows. TOPEX altimeter data come from the TOPEX/Poseidon Merged Geophysical Data Records, generation B (MGDRB; Benada 1997). A valid TOPEX/NSCAT crossover occurs when the time difference is less than  $\pm$  one hour and wind cell spatial intersection is within  $\pm 12\text{ km}$  of their respective centers. Flags within each sensor's data record at that point must indicate deep ocean, no sea ice, and high data quality per handbook recommendations. Our resulting global dataset consists of over 245 000 points spanning the nine month NSCAT mission.

This study assumes that the scatterometer reference winds are free of error due to sea state impacts. NSCAT winds are derived from antenna incidence angles spanning  $16^\circ\text{--}54^\circ$  from vertical. Sea-state impacts are not considered within the NSCAT wind inversion scheme and are assumed negligible based on numerous studies (see Nghiem et al. 1997). Recent global study (Queffeuou et al. 1999) does imply that a measurable sea state effect can be seen in lower-angle NSCAT data, but also that the larger-angle scatterometer data do not exhibit significant correlation. Our datasets were used to affirm the Queffeuou et al. (1999) findings (not shown here) and lead us to conservatively accept only those NSCAT winds where the observation angle exceeded  $40^\circ$ . The expectation is that this will reduce possible sea-state impacts on scatterometer winds to levels acceptable for altimeter model study. Buoy and wind forecast validations in sections 3 (e.g., Fig. 5) and 6 indicate that this is the case. The same angular cutoff criterion is applied to all scatterometer data in this study.

To bring the TOPEX spatial resolution as near as possible to NSCAT we include an average over those TOPEX data points (one point every 6 km along track) that fall within the given NSCAT wind cell. Thus the TOPEX "wind cell" characteristics vary from  $2 \times 6$

km to  $2 \times 25\text{ km}$ . An additional minimal filtering of outliers includes the removal of those points where  $\sigma^o$  is below 5.0 or above 30.0 dB and those cases where the TOPEX liquid water estimate exceeds 500 microns—a likely rain event. The composite dataset contains 96 436 samples. The main parameters of interest at each point are the NSCAT 10-m wind estimate, TOPEX normalized radar cross section ( $\sigma^o$ ) at Ku-band, and TOPEX  $H_s$ . TOPEX in this case denotes only the NASA altimeter aboard the satellite; we do not use data from the Poseidon altimeter for this study.

#### b. TOPEX and ERS C-band scatterometers

The *ERS-1* and *-2* satellites' Active Microwave Instrument included a C-band scatterometer that provided another opportunity for TOPEX collocation. TOPEX/ERS crossovers were compiled by CERSAT (Centre ERS d'Archivage et de Traitement) with details provided in a data report (CERSAT 1996). To encompass the 1996–97 NSCAT period, and to acquire a large sampling, data include both *ERS-1* and *-2* platforms over the 2-yr period extending from November 1995 to November 1997. The period November 1995 to May 1996 corresponds to *ERS-1*, with *ERS-2* data after this. The scatterometers are virtually identical and their accurate cross calibration was accounted for within the CERSAT database.

ERS backscatter measurement processing results in a wind inversion with surface resolution cell of  $50 \times 50\text{ km}$ . Version 3.2 of the CMOD wind inversion algorithm was used. Flagging of both TOPEX and ERS data was conservatively performed with the same criterion as for the TOPEX/NSCAT dataset.

Collocation was first done for a  $\pm$  one hour time lag and a  $\pm 50\text{ km}$  spatial separation. This resulted in more than 300 000 data points. A subset was extracted to keep only those crossovers with 30 min and 15-km time and space separations. The final dataset also contains only those cases where all ERS data are for an incidence angle higher than  $40^\circ$ . This leads to a globally distributed total of 70 500 samples.

#### c. TOPEX and SeaWinds

Global Ku-band scatterometer-derived winds from the SeaWinds sensor aboard the Quik-SCAT satellite (QSCAT) have been available since July of 1999. SeaWinds has two rotating antennas operating at fixed incidence angles ( $46^\circ$  and  $52^\circ$ ). Collocations of TOPEX and QSCAT data have been performed by CERSAT for the twelve month period from July (1999) to August 2000. The Ku-band wind speed algorithm for SeaWinds is the QSCAT-1 model function. Note also that there was a substantial change in TOPEX between the NSCAT and QuikSCAT periods in that the original NASA altimeter (side A) aboard TOPEX was switched off and replaced with a redundant system (side B) in February

1999. The changes associated with this TOPEX switch are not expected to significantly impact this study.

For the TOPEX/SeaWinds dataset the crossovers constraints for time and space separation are less than 30 min and 15 km, respectively. As with NSCAT, a high-resolution processing was applied, resulting in  $25 \times 25$  km wind cells. This product provides two different TOPEX measurement resolutions to remain consistent with the collocated datasets detailed above. The first is the single point MGDR-B data with  $2 \times 6$  km resolution. The second is a 13-s average along the TOPEX track, corresponding to a  $2 \times 70$  km ground cell. The former is used in the validations of section 6. Filtering of the two data sources was performed as described for TOPEX/NSCAT with the exception being that all SeaWinds wind vector estimates are implicitly for incidence angles above  $40^\circ$ . The dataset used in our study contains more than 330 000 data points.

#### d. TOPEX and ECMWF winds

Atmospheric model estimates of surface wind vector from the ECWMF surface analysis are provided for each point in the three collocation datasets described above. These model-derived winds are used for altimeter algorithm validation purposes. Note that these model winds do not come from the ECMWF reanalysis. The ECMWF analysis provides near-surface wind vector estimates for an altitude of 10 m above the ocean and for neutral stability. Their output is on a  $1.125^\circ \times 1.125^\circ$  grid every 6 h. Model output is interpolated in space and time to derive a wind estimate collocated with the altimeter observation. This results in a maximum time lag of  $\pm 3$  h. Ground resolution for ECMWF is then  $125 \times 60$ – $125$  km, depending on latitude.

This type of wind product has proven to be quite useful for past global development of the average wind characteristics of a sensor such as the satellite scatterometer or altimeter. However, the inherent smoothness associated with the relatively larger spatial and temporal scales of such a model does introduce variance into any comparison with a finer-resolution observations. For example, Freilich and Dunbar (1993) estimate this inter-comparison noise factor to be roughly 30% in their altimeter wind model study.

#### e. ERS-2 altimeter and NSCAT

Observations from the Ku-band altimeter aboard the ERS-2 satellite are provided as one independent assessment of the results derived using the TOPEX-based crossover datasets. Crossovers between the ERS-2 altimeter and NSCAT have been obtained from CERSAT and the compilation covers the whole NSCAT period (CERSAT 1997; Queffeuou 1999). Our dataset was assembled and filtered for the cases above. One difference here is that NSCAT wind estimates are for  $50 \times 50$  km (not 25 km) wind cells. The ERS-2 altimeter measure-

ments are averaged over 11 s, representing an along track distance of about 70 km. One-second averaged altimeter measurements are also provided. The globally distributed dataset contains 129 000 crossovers.

#### f. TOPEX and NDBC buoys

A dataset of collocated buoy/altimeter measurements has been developed at the Southampton Oceanography Centre to investigate sea-state effects on retrieved altimeter wind speed (e.g., Gommenginger et al. 2002). Featuring nearly 4500 data points, it is the largest dataset to-date of collocated buoy/altimeter measurements in the open ocean, and represents a significant increase with respect to previous collocated datasets featuring only a few hundred points (Monaldo and Dobson 1989; Glazman and Greysukh 1993). The buoy data originate from the U.S. National Data Buoy Center (NDBC), the Canadian Meteorological Data Services (CMEDS), the UK Met Office (UKMO), and the Japanese Maritime Agency (JMA). These moored buoys were selected for their location in open waters and their proximity to the TOPEX altimeter tracks. Here, the space-time collocation criteria were set to select data within 50 km and 30 min of a TOPEX overpass, in accordance with the criteria commonly used in this type of study (e.g., Hwang et al. 1998). With these criteria and the application of the standard ice and rain flags in the Aviso MGDR products, the collocation yielded a dataset of 4444 measurements obtained between September 1992 and December 1998.

The buoy wind speed represents a 8-min average. It is provided as a neutral stability 10-m estimate, adjusted using the air-sea temperature dependent correction factors of Dobson (1991). No attempt was made to compensate for the small systematic drift in TOPEX  $H_s$  estimates toward the end of 1998 (end of the side A mission) as this represented a small proportion of our dataset. Basic outlier removal eliminated any data for which the altimeter  $U_{10}$  (retrieved with MCW) and the buoy  $U_{10}$  differed by more than  $5 \text{ m s}^{-1}$ .

### 3. The training set—NSCAT and TOPEX

Figure 1 presents TOPEX altimeter parameters versus NSCAT-derived wind speed for the TOPEX/NSCAT collocation dataset. Only one tenth of the total sampling are represented in Figs. 1a–c. Figure 1a shows TOPEX Ku-band  $\sigma^\circ$  versus NSCAT wind speed. Their relation appears to be close to linear above  $5$ – $6 \text{ m s}^{-1}$ , while, with decreasing wind speeds,  $\sigma^\circ$  increases rapidly. One can note that  $\sigma^\circ$  scatter is on the order of 1–2 dB for a given wind speed. This scatter also increases with decreasing wind speed. An obvious feature of the large dataset is measurement across a range of ocean wind speeds but note that the population lessens above  $15 \text{ m s}^{-1}$ .

TOPEX  $H_s$  measurements are shown in Fig. 1b. The

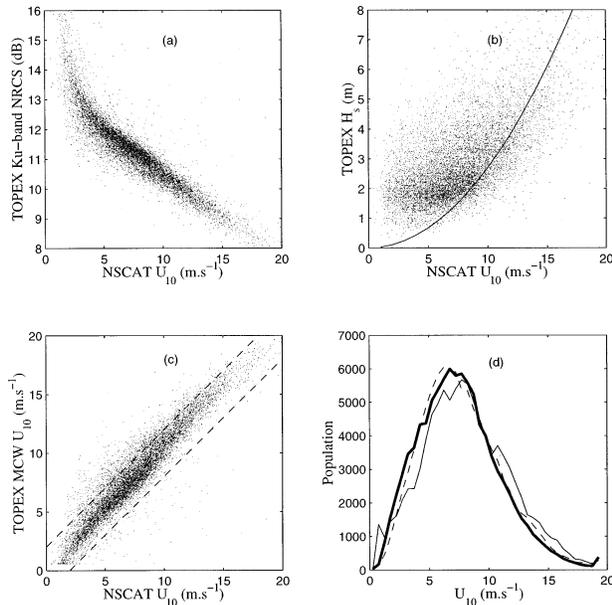


FIG. 1. (a) TOPEX Ku-band  $\sigma^0$ , (b) TOPEX  $H_s$ , and (c) TOPEX MCW-derived wind speed—all versus the reference NSCAT 10-m wind speed. (d) Histograms for  $U_{\text{NSCAT}}$  (thick),  $U_{\text{ECMWF}}$  (dashed), and  $U_{\text{MCW}}$  (thin). The model curve in (b) represents fully developed seas for a given NSCAT wind speed using the Sverdrup–Munk model.

observed range of sea-state conditions and  $H_s$  wind dependence are consistent with past global TOPEX observations (e.g., Callahan et al. 1994). A model curve is superposed on the graph representing  $H_s$  for a wind sea that has reached full development under a given wind (Sverdrup and Munk 1947). Observed  $H_s$  values

falling below this curve likely indicate cases where a wind sea is building. For light to moderate winds  $H_s$  data usually exceed this curve indicating that a swell component is common. It follows that idealized fetch-limited wind sea situations are infrequent on the open ocean. This holds except for cases of high wind speeds where swell becomes a much smaller contributor within the total wave height.

Figure 1c presents altimeter wind speed estimated from the operational TOPEX algorithm (Witter and Chelton 1991), showing that wind differences frequently exceed  $\pm 2 \text{ m s}^{-1}$ . About 15% of the TOPEX-derived wind estimates exceed these bounds within the complete TOPEX/NSCAT dataset. Furthermore, a systematic overestimation of wind speed by about  $0.5 \text{ m s}^{-1}$  is observed over the whole wind speed range. This bias is confirmed by the wind speed histograms presented in Fig. 1d. This same level of bias has also been pointed out in past studies (e.g., Freilich and Challenor 1994). Figure 1d also provides the wind speed histogram for collocated ECMWF wind model data. There is excellent agreement between the NSCAT and ECMWF estimation. Note that the global wind distributions show prevailing ocean wind speeds fall between 3 and  $12 \text{ m s}^{-1}$ . This strong weighting of the sampled population towards a median value near  $7 \text{ m s}^{-1}$  needs consideration in wind model developments and validation.

Figure 2a gives some sense of the global distribution of this compilation. Data for both satellites were conservatively filtered for sea-ice occurrence and so this should not be a noise factor. A salient characteristic of the crossings for these polar-orbiting platforms was the

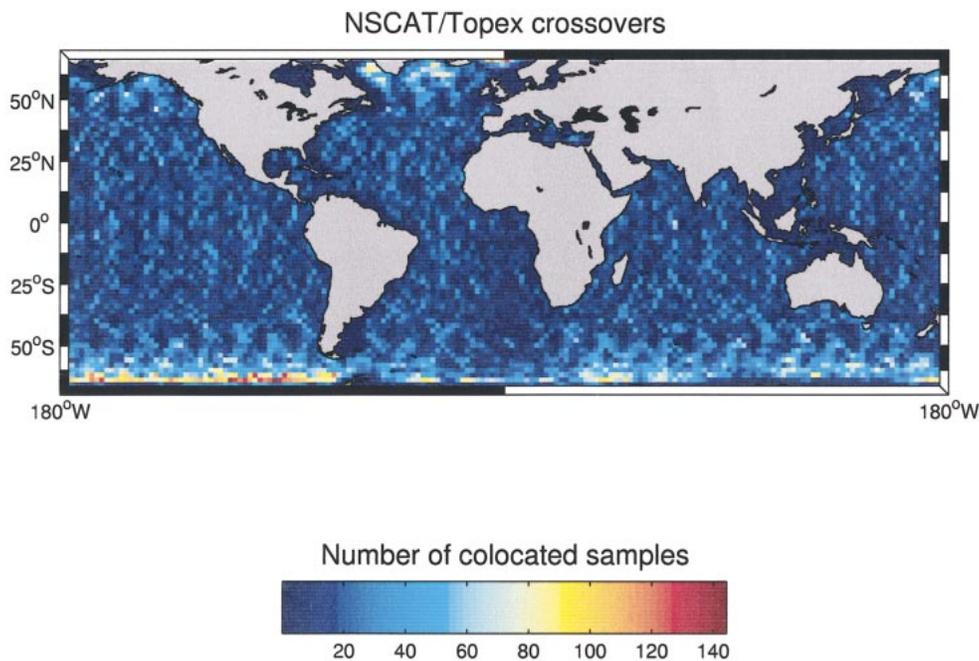


FIG. 2. Spatial density of TOPEX/NSCAT satellite crossover points used in this study.

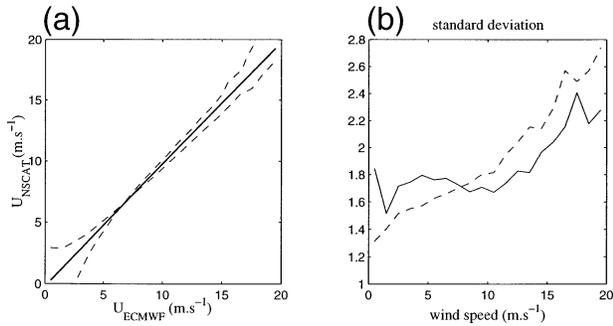


FIG. 3. NSCAT and ECMWF wind speed comparisons: (a) bin-averaged relations (dashed) and symmetrical linear regression (solid); (b) standard deviation of the  $U_{NSCAT} - U_{ECMWF}$  wind difference, as a function of NSCAT wind (dashed) and ECMWF wind (solid).

increased likelihood of high-latitude intersections. This same polar weighting is evident in the TOPEX/ERS and TOPEX/QSCAT data.

Next is an assessment of the chosen reference wind product for model developments—NSCAT’s 10-m wind derived from a 25-km wind vector cell. As noted earlier, scatterometer data are restricted to winds derived using radar incidence angles above  $40^\circ$  to ameliorate possible wind speed errors associated with long waves or sea state. For indepth validation of the NSCAT winds and the NSCAT-1 geophysical model function see Freilich and Dunbar (1999) and Wentz and Smith (1999).

Figure 3a presents bin-averaged relations (1  $m s^{-1}$  bin width) and an orthogonal linear regression between the ECMWF model wind and the NSCAT product. As mentioned earlier the model winds have been interpolated and are a lower resolution product to begin with. Both products will have substantial rms about a “true” wind measurement as well as possible biases. An orthogonal linear regression is provided to show results are invariant with choice of regressor. The comparison between the two parameters shows excellent agreement with a slight bias that never exceeds  $0.3 m s^{-1}$  in the average of one estimate when averaged over binned subsets of the other.

Referencing Fig. 3b one sees that standard deviation (std) of the  $U_{NSCAT} - U_{ECMWF}$  wind difference as a function of ECMWF or NSCAT wind speed is about constant at  $1.7 m s^{-1}$  for winds less than  $10 m s^{-1}$ . For higher wind speeds the std increases strongly—up to  $2.4 m s^{-1}$  at  $16 m s^{-1}$ . This increase is likely associated with the different resolution cells for NSCAT (25 km) and ECMWF (120 km). This becomes more important for stronger winds because they are likely to be associated with smaller-scale meteorological features.

The cross correlation between significant wave height, altimeter backscatter and surface wind is now addressed. Figure 4 displays a gridding of  $\sigma^\circ$  versus  $U_{NSCAT}$  with color representing the average value for altimeter-derived  $H_s$  at a given wind speed and  $\sigma^\circ$ . One can observe a clear dependence of altimeter  $\sigma^\circ$  upon

$H_s$  variation for a given  $U_{10}$ . The relative magnitude of the change decreases with increasing wind speed. However, the variance is quite strong for common ocean wind speeds, such as the 2-dB range at  $6 m s^{-1}$ . Such a 2-dB range translates to large wind error for the MCW  $\sigma^\circ$  wind speed algorithm where wind sensitivity is about  $3 m s^{-1} dB^{-1}$  at  $6 m s^{-1}$ .

It is important to recognize that the pattern that emerges in Fig. 4 is due to the very large amount and global nature of the assembled dataset. Much of the space mapped here represents infrequent events and high latitude observations that might not be seen in smaller and more localized composite datasets (e.g., altimeter collocations with the NDBC buoy network observations). One can expect such a mapping to gain even better definition with a long-term multiyear data compilation where the sampling of sparse regions and infrequent events increases (e.g., continued collocation between TOPEX/*Jason-1* and the SeaWinds sensors).

The observed correlation between altimeter  $\sigma^\circ$  and  $H_s$  for any fixed wind speed confirms, at least in some respects, numerous past studies suggesting sea-state impacts on altimeter wind estimations as discussed in our introduction and first noted by Monaldo and Dobson (1989). cursory inspection of the GEOSAT/NDBC dataset used for many past sea state impact studies indicated little of the variations seen in Fig. 4 and as such, inconclusive findings (cf. Wu 1999) are not surprising.

Another means of illustrating the sea-state impact is in terms of the operational altimeter wind product. Figure 5 (left panel) provides that TOPEX estimate versus the coincident ECMWF model wind speed for varied  $H_s$  levels. The TOPEX wind bias is systematically above or below the reference wind for a given  $H_s$ . NSCAT wind data (right panel) for the same conditions indicate that the scatterometer wind product shows a negligible relation to  $H_s$  variation. These data affirm that the altimeter exhibits this mean  $H_s$  dependence, whereas the high angle scatterometer winds do not.

We reemphasize that the results shown here depict the average, not instantaneous, globally derived signatures. Moreover, significant wave height is the lowest-order moment of the directional wave height spectrum, and  $H_s$  does not necessarily covary with the long wave slope statistics that correlate more closely with  $\sigma^\circ$  measurements. Therefore, just as for the backscatter alone, one anticipates no unique pairing of  $\sigma^\circ$  and  $H_s$  that maps to a unique  $U_{10}$ . Ambiguity is likely to remain due to the indeterminate ratio of sea versus swell within  $H_s$ , the indeterminate direction of the sea versus the swell, unknown fetch and duration, etc. The notion that  $H_s$  is a limited surrogate for actual long wave impacts in altimeter  $U_{10}$  inversion should be kept in mind when developing a two-parameter ( $\sigma^\circ$  and  $H_s$ ) wind speed algorithm for the altimeter.

#### 4. Two-parameter model functions

As discussed, a sea-state impact upon  $\sigma^\circ$  measurements has been suspected for some time, but the pre-

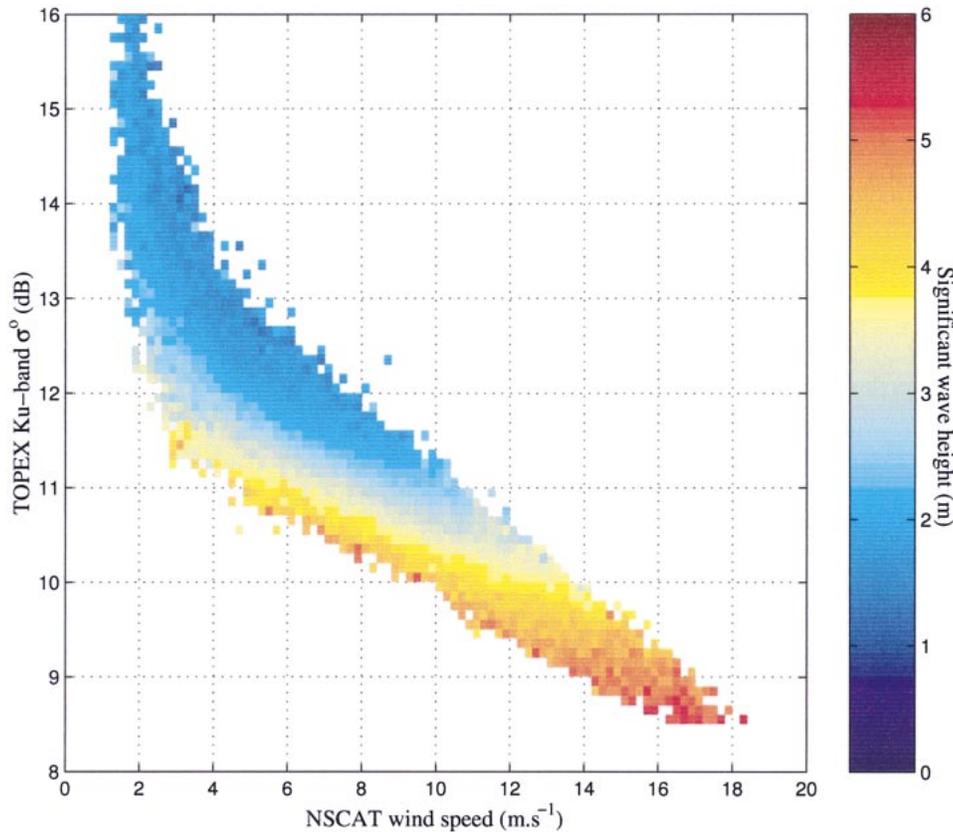


FIG. 4. Grid for TOPEX  $\sigma^\circ$  vs NSCAT wind speed and  $H_s$ , where color represents the average value for  $H_s$ . Bin width : 0.2 m s<sup>-1</sup>, 0.1 dB.

vailing wind speed inversion for TOPEX still follows the single parameter MCW model:

$$\sigma^\circ = f_{\text{MCW}}(U_{10}). \quad (1)$$

The NSCAT/TOPEX data offers a new opportunity to develop a globally based altimeter wind speed solution relevant for operational usage. That is, invert  $U_{10}$  using solely those two products measured by any Ku-band altimeter:

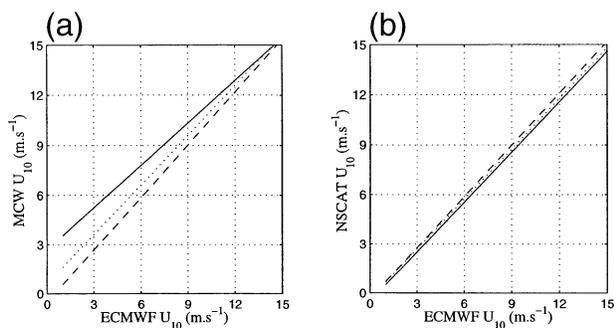


FIG. 5. Symmetrical linear regression between (a)  $U_{\text{mcw}}$  and  $U_{\text{ecmwf}}$ , (b)  $U_{\text{nscat}}$  and  $U_{\text{ecmwf}}$  vs ECMWF  $U_{10}$  for different SWH classes: 1 m (dashed line), 3 m (dotted line), 5 m (solid line). The  $H_s$  range is  $\pm 0.5$  m about the indicated value.

$$U_{10} = f_1(\sigma^\circ, H_s). \quad (2)$$

Referring back to Fig. 4 one can see that the three-dimensional representation indicates some nonlinearity, particularly when considering both high and low wind speed regimes. For this reason we choose artificial neural networks to develop the empirical model relating these three parameters,  $\sigma^\circ$ ,  $H_s$ , and  $U_{10}$ . This choice assures a robust nonparametric mapping, with one ramification being that a derivative sea-state parameter such as pseudowave age (e.g., Glazman and Greysukh 1993) that combines wind speed (or  $\sigma^\circ$ ) with  $H_s$  should be encompassed by this solution.

The overall objective in this nonparametric regression is to generate a globally faithful and continuous representation of the observed relation between a reference  $U_{10}$  and TOPEX measurements. A most direct inversion is given by  $f_1$  [Eq. (2)]. However, from a physical or sensor perspective one anticipates that Ku-band  $\sigma^\circ$  observations are correlated with the local wind waves ( $U_{10}$ ) but also with the long wave roughness and its surrogate,  $H_s$ . As such, a more physically relevant (forward) model treats those latter variables as network inputs and  $\sigma^\circ$  as the output:

$$\sigma^\circ = f_2(U_{10}, H_s). \quad (3)$$

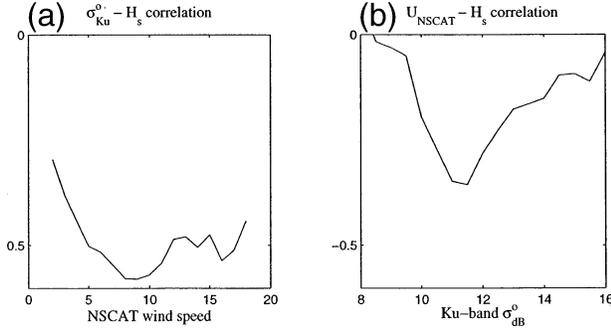


FIG. 6. Correlation of the dependent variable in (a)  $f_2$  and (b)  $f_1$  with variation in  $H_s$ : y axis depicts the value of the linear correlation coefficient computed at each respective  $x$ -axis location.

Wind speed is readily derived from  $f_2$  by inversion using a look-up table. These two functions [Eqs. (2) and (3)] will not necessarily provide identical  $U_{10}$  values for identical  $\sigma^o$  and  $H_s$  inputs.

This point is deduced from the relatively high correlation observed between the radar backscatter and  $H_s$  in comparison to the correlation between NSCAT's wind speed and  $H_s$  (see Fig. 6). At a given value of its first input parameter  $\sigma^o$  ( $U_{10}$ ), the  $f_1$  ( $f_2$ ) mapping has a relatively lower (higher) correlation between its output variable  $U_{10}$  ( $\sigma^o$ ) and  $H_s$ . The input variables [ $U_{10}$ ,  $H_s$ ] for  $f_2$  are less self-correlated than those for  $f_1$ . Thus, the  $f_2$  mapping is expected to weight the sea-state parameter more strongly than  $f_1$ .

Detail for  $f_1$  and  $f_2$  model developments are provided in appendix A. In initial studies (Gourrion et al. 2000) two differing neural techniques were applied to the TOPEX/NSCAT data. The first approach used a general regression neural network (GRNN) and the second approach used the multilayer perceptron (MLP). That report concludes that the two approaches give results that are in close agreement. This paper reports only the MLP solutions as they provide closed-form solutions that are most readily disseminated.

#### a. Model definition

Given the input pair  $[\sigma^o, H_s]$  or  $[U_{10}, H_s]$  the following equations and coefficients (Tables 1, 2, and 3) define the analytical neural network solutions  $f_1$  and  $f_2$ , respectively.

The  $f_1$  solution provides an altimeter-derived 10-m neutral-stability wind speed trained to the NSCAT-1 model function output:

TABLE 1. Input and output data scaling coefficients needed for both neural network models.

Parameter	$a$	$b$
$\sigma^o$	-0.34336	0.06909
$H_s$	0.08725	0.06374
$U_{10}$	0.10000	0.02844

TABLE 2. Coefficients for the  $f_1$  model.

Parameter	Matrix elements	
$\mathbf{W}_x$	-33.95062	-11.03394
	-3.93428	-0.05834
$\mathbf{W}_y$	0.54012	10.40481
$\mathbf{B}_x$	18.06378	-0.37228
$B_y$	-2.28387	...
$\mathbf{P}$	$a_{\sigma^o} + b_{\sigma^o}\sigma^o$	$a_{H_s} + b_{H_s}H_s$

$$U_{10} = f_1 = \frac{Y_{f_1} - a_{U_{10}}}{b_{U_{10}}}, \quad (4)$$

while  $f_2$  predicts the normalized radar cross section at Ku-band and for vertical incidence:

$$\sigma^o = f_2 = \frac{Y_{f_2} - a_{\sigma^o}}{b_{\sigma^o}}. \quad (5)$$

The variable  $Y$  (for either  $f_1$  or  $f_2$ ) is derived as

$$Y = [1 + \exp^{-(\mathbf{W}_x \mathbf{X}^T + \mathbf{B}_x^T)}]^{-1}, \quad (6)$$

with  $\mathbf{X}$  defined as

$$\mathbf{X} = [1 + \exp^{-(\mathbf{W}_y \mathbf{P}^T + \mathbf{B}_y^T)}]^{-1}. \quad (7)$$

The subscripts on scaling coefficients ( $a$  and  $b$ ) correspond to the appropriate altimeter parameter, boldface variables are vectors, and  $^T$  denotes the vector transpose. Here,  $\mathbf{P}_{f_1}$  (respectively  $\mathbf{P}_{f_2}$ ) is the two-row input matrix to the  $f_1$  ( $f_2$ ) transfer function given by

$$\begin{pmatrix} \widetilde{\sigma^o} \\ \widetilde{H_s} \end{pmatrix}, \quad \text{respectively} \quad \begin{pmatrix} \widetilde{U_{10}} \\ \widetilde{H_s} \end{pmatrix},$$

the  $\sim$  denoting data normalized with the appropriate  $a$  and  $b$  coefficients. Equations (4) and (5) provide the network outputs once renormalized using these coefficients.

Both Eqs (6) and (7) are log-sigmoid transfer functions, commonly used in MLP neural network developments. Output from each spans from 0 to 1. Among MLP nonlinear transfer functions, these are easiest to compute both in the forward and backpropagation (partial derivative) modes of operation. The log-sigmoid transfer function is chosen because outputs become asymptotic at the input extremes. Inputs closer to zero cause the output to act more linearly. These properties lead to an ideal amplifier when coupled together in stages. Each stage can only saturate to 0 or 1, thereby not overwhelming inputs of the following MLP stage. The

TABLE 3. Coefficients for the  $f_2$  model.

Parameter	Matrix elements	
$\mathbf{W}_x$	-43.39541	-6.92550
	2.78612	1.22293
$\mathbf{W}_y$	1.18281	-3.30096
$\mathbf{B}_x$	7.83459	-1.46489
$B_y$	1.13906	...
$\mathbf{P}$	$a_{U_{10}} + b_{U_{10}}U_{10}$	$a_{H_s} + b_{H_s}H_s$

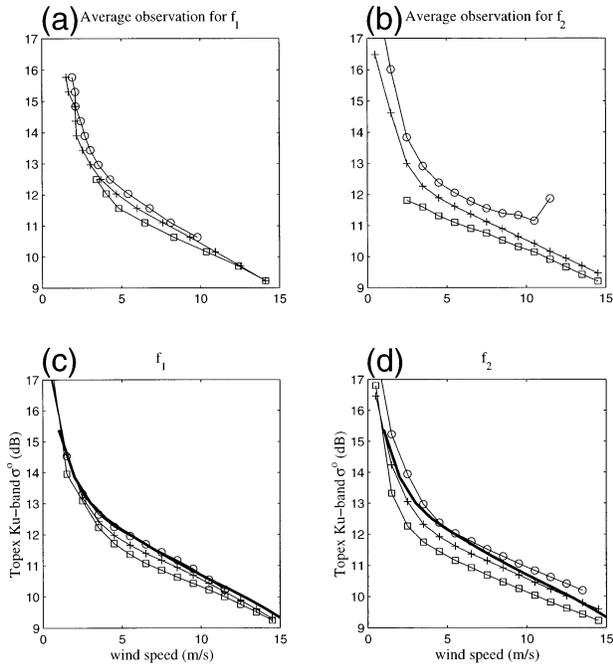


FIG. 7. (a),(b) represent the average over the input data domain for the  $f_1$  (left) and  $f_2$  (right) models as derived from TOPEX/NSCAT dataset. (c),(d) represent the overall behavior of the neural network solutions. Curves depict data or model results about three different SWH values: 1 m (circles), 3 m (pluses), and 5 m (squares). The MCW model is overlaid for comparison (thick line).

attenuating property of the log-sigmoid is what allows us to represent nonlinear input-output relationships within our training and test data.

As described in appendix A, we empirically determined that adequate characterization is obtained with a three-layer MLP network. Model outputs are readily calculated with the simple matrix operations defined in Eqs. (4)–(7).

*b. Model results*

Figure 7 illustrates model response versus  $\sigma^o$  and versus wind speed for three separate  $H_s$  values. Upper panels depict the average TOPEX/NSCAT observations and are representative of the  $f_1$  and  $f_2$  training sets. The standard TOPEX routine (MCW) is displayed for reference. Both neural models clearly differ from this MCW result. The observations and models also make it clear that the  $f_1$  mapping differs substantially from  $f_2$ . As already discussed, this may lead to differences with respect to  $H_s$  impacts on wind inversion. Differences are particularly evident in Fig. 7 at low and high  $U_{10}$ . Overall, it is apparent that the  $f_1$  model represents a weaker departure from MCW.

Histograms of NSCAT and model-derived winds are shown in Fig. 8. Wind speed is retrieved from  $f_2$  using a look-up table. The histogram bin size here is  $1.0 \text{ m s}^{-1}$ . Agreement between the three products is close, with

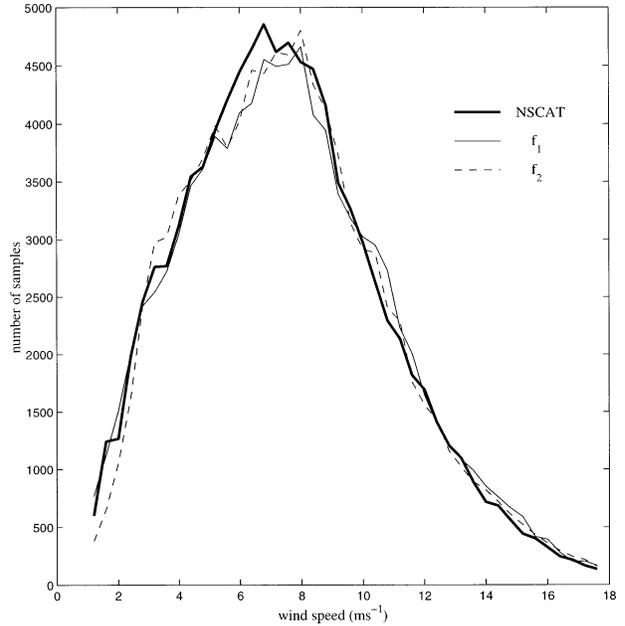


FIG. 8. Frequency distribution of wind speeds over the entire NSCAT/TOPEX dataset.

the models missing only slightly in the range of  $5\text{--}7 \text{ m s}^{-1}$ . Referring back to Fig. 1, one observes that the new functions appear to improve over MCW.

Finally, recall that multivaluedness within this particular three parameter dataset thwarts efforts to further refine model performance by various methods such as changing the training set makeup, network size, or data limits. That is, given altimeter measurement pairs [ $\sigma^o$ ,  $H_s$ ] do not always map to a unique wind speed. This property is partially attributed to the fact that  $H_s$  is a limited proxy for the actual long wave conditions dictating the relation between  $\sigma^o$  and  $U_{10}$ . Illustration is provided in Fig. 9, where the left panel shows observed  $U_{10}$  ranging from 4 to 8  $\text{m s}^{-1}$  for fixed values of  $H_s$  (1.7 m) and  $\sigma^o$  (12.2 dB). The right-hand panel provides evidence that the multivaluedness is geophysical in na-

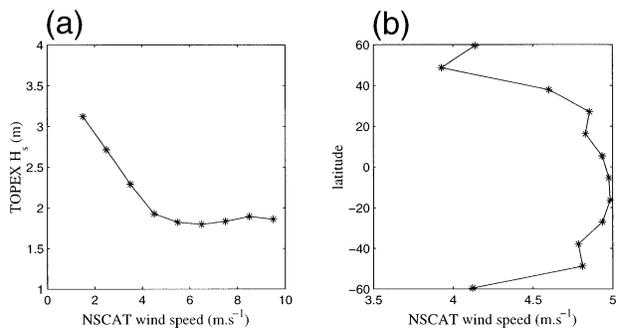


FIG. 9. (a) The average of TOPEX  $H_s$  values for a given NSCAT  $U_{10}$  ( $\pm 0.5 \text{ m s}^{-1}$ ) and TOPEX  $\sigma^o = 12.2$  ( $\pm 0.2 \text{ dB}$ ). (right panel) (b) The average NSCAT  $U_{10}$  at differing zonal bands for the fixed TOPEX data pair [ $\sigma^o = 12.2$  ( $\pm 0.2$ ),  $H_s = 1.7$  ( $\pm 0.2$ )].

ture. This plot shows systematic  $U_{10}$  variance versus latitude for the specified  $[\sigma^o, H_s]$  pair. To first order, this latitudinal dependence indicates some regional change in the long wave climate. Toward the poles there is typically a larger ratio between swell and wind waves within the total sea state. Wind speed is systematically  $0.6 \text{ m s}^{-1}$  lower than at midlatitudes. Regardless of the source, a multivalued behavior manifests itself in empirical model training as a fundamental error source that the network cannot resolve without additional information. This is true for both  $f_1$  and  $f_2$  mappings. The latitudinal dependence evident in Fig. 9 also suggests that the improved two-parameter altimeter wind models will still exhibit residuals in regional or seasonal evaluations. Further work on these points is warranted but deemed beyond the scope of the present effort.

### 5. Wind speed model intercomparison

This section uses the total TOPEX/NSCAT data to compare model-derived results with previously published altimeter wind speed routines. A following section documents additional independent validation. A key study objective is to provide wind estimates with improved overall statistical performance when applied to global open-ocean observations. Past efforts have emphasized that the measure for algorithm accuracy should not solely be the global wind speed error (cf. Glazman and Greysukh 1993). The large amount of data in our training and validation datasets permits detailed wind error evaluation. This includes the ability to assess error at light, moderate and high wind speeds, as well as the identification of residual correlation with the reference wind speed and/or  $H_s$ . Evaluation criteria values both bias and rms minimization. Due to operational and climate study considerations, a continuous function is also desired. Moreover, estimates should produce a faithful global wind speed histogram.

There are many published altimeter wind speed models that use only the Ku-band  $\sigma^o$ . A review is found in Chelton et al. (2001). Here, the choice of single-parameter algorithms for intercomparison is limited to Brown et al. (1981), Witter and Chelton (1991), and Freilich and Challenor (1994). These are noted as B81, MCW, and FC, respectively, in the text to follow. The contention is that these models encompass most other variations—B81 represents a buoy-tuned discontinuous (three branch) function that has often produced the lowest overall rms, whereas MCW and FC represent smoother, statistically derived functions that are well-validated and robust.

Two wind speed models that utilize both the Ku-band  $\sigma^o$  and  $H_s$  are also assessed. Lefèvre et al. (1994) provides a closed-form parametric solution as a function of the two altimeter inputs based on comparison between TOPEX and wind model (ARPEGE, the Meteo-France atmospheric model) data. Glazman and Greysukh (1993, GG2 hereafter) developed a classification approach

TABLE 4. Wind speed error statistics for TOPEX altimeter  $U_{10}$  derived using the specified algorithms. Computations are made over the 96 436 samples within the collocated dataset TOPEX/NSCAT where the reference NSCAT wind speed fell between 1 and  $17 \text{ m s}^{-1}$ . Error trends are the slopes  $\alpha$  in the linear regression model  $U_{\text{err}} = \beta + \alpha X$  where  $X$  is  $U_{\text{NSCAT}}$  or  $H_s$ . The separation time between altimeter and scatterometer estimates is  $\pm 1 \text{ h}$ .

Altimeter model	Bias (m s <sup>-1</sup> )	Std	Rms	% error >2 m s <sup>-1</sup>	U error trend	H <sub>s</sub> error trend
B81	0.36	1.16	1.22	9	-0.22	-0.04
MCW	0.61	1.19	1.34	12	0.03	0.48
FC	0.02	1.22	1.22	7	0.03	0.48
GG2	0.04	1.07	1.07	5	-0.06	0.17
Lefèvre	0.28	1.44	1.47	15	-0.28	0.17
<i>f</i> <sub>1</sub>	0.04	1.05	1.05	5	0.00	0.20
<i>f</i> <sub>2</sub>	0.01	1.10	1.10	6	-0.02	-0.06

where wind speed is estimated using one of two distinctly separate single parameter ( $\sigma^o \rightarrow U_{10}$ ) models. This work was based on comparisons between the GEOSAT altimeter and buoy-derived wind and wave measurements. For operational inversion, the classification relies on altimeter  $H_s$  and  $\sigma^o$  estimates. The two GG2 wind models are rooted in the concept of pseudowave age,  $\zeta[=f(H_s/U^2)]$ , where the first model is for “young” or fetch-limited wind wave conditions and the second covers all other situations (i.e., covering fully developed seas and the ubiquitous mixed seas). As noted in GG2, the classification is inherently discontinuous and leads to bimodality in the output  $U_{10}$  distribution.

Table 4 presents wind speed differences between the altimeter and scatterometer for the commonly cited case where statistics are computed over all wind values. Computations for each altimeter algorithm are included. The reference (NSCAT)  $U_{10}$  range used here is 1–17  $\text{m s}^{-1}$ . The limits encompass the range of conditions over which most of the models were developed. Note that, where applicable, TOPEX  $\sigma^o$  values are adjusted down by 0.63 dB before use in GEOSAT-era routines (see Callahan et al. 1994).

The wind speed difference  $U_{\text{err}}$  is defined as

$$U_{\text{err}} = U_{10}^{\text{altimeter}} - U_{10}^{\text{reference}}, \quad (8)$$

where the reference is NSCAT in this case and the neutral stability wind speed is given in meters per second. Wind error standard deviation is simply

$$\text{std} = (\langle U_{\text{err}}^2 \rangle - \langle U_{\text{err}} \rangle^2)^{1/2} = (\text{rms}^2 - \text{bias}^2)^{1/2}. \quad (9)$$

This factor provides bias-independent rms assessment and is of some value in the following discussion where small differences become significant.

A first observation regarding Table 4 is the rms error level of 1.0–1.4. This is low compared to past published values of 1.6 to 2.0 associated with buoy or model wind speed comparisons. This decrease is most likely attributed to the tight constraint on spatial collocation distance (<15 km) and similar spatial footprints. Such rms levels agree with altimeter and buoy comparisons where the spatial separation is limited to 15 or 20 km (e.g.,

TABLE 5. Results follow the same format as table 4, but estimates are now localized to three separate wind speeds (4, 8, and 12 m s<sup>-1</sup>) as indicated.

Altimeter model	Bias	Std	Rms	% error >2 m s <sup>-1</sup>	H <sub>s</sub> error trend
3–5 m s <sup>-1</sup>					
B81	1.20	1.02	1.57	16	0.72
MCW	0.62	1.28	1.43	3	0.91
FC	-0.04	1.24	1.24	9	0.89
GG2	0.43	0.92	1.01	6	0.61
Lefèvre	1.52	1.08	1.87	26	1.05
f <sub>1</sub>	0.09	1.07	1.07	7	0.50
f <sub>2</sub>	0.15	0.99	1.00	6	0.17
7–9 m s <sup>-1</sup>					
B81	0.30	0.81	0.86	3	0.50
MCW	0.59	1.14	1.28	12	0.78
FC	-0.03	1.14	1.14	7	0.78
GG2	-0.15	1.14	1.15	7	0.26
Lefèvre	-0.12	1.07	1.08	6	0.99
f <sub>1</sub>	0.04	1.00	1.00	5	0.26
f <sub>2</sub>	-0.02	1.01	1.01	5	-0.05
11–13 m s <sup>-1</sup>					
8311 samples					
B81	-0.84	0.91	1.24	6	0.38
MCW	0.85	1.19	1.46	14	0.53
FC	0.28	1.23	1.26	7	0.56
GG2	0.01	1.12	1.12	5	0.44
Lefèvre	-0.88	1.20	1.49	15	0.93
f <sub>1</sub>	0.06	1.15	1.16	5	0.28
f <sub>2</sub>	-0.07	1.36	1.36	10	-0.26

Hwang et al. 1998). The largest temporal separation here is one hour. Total sample population is more than 96 000, thus parameter noise levels are negligible.

The overall bias for any altimeter routine is well below 1 m s<sup>-1</sup> with only MCW output showing a value above 0.5. Slopes computed for the linear trend of  $U_{err}$  versus wind speed are provided as some indication of bias variability with wind speed. It is clear that B81 and Lefevre model bias trends are much higher than for the other routines. The most commonly reported assessment parameter, global rms, shows values spanning from 1.05 to 1.47. The Lefevre algorithm exhibits the highest error. The single parameter models (B81, MCW, and FC) exhibit similar rms and std estimates (1.2) while the 3 two-parameter algorithms (GG2,  $f_1$ ,  $f_2$ ) are below 1.10.

Table 4 also provides a column representing the wind error slope versus  $H_s$  for consistency with Freilich and Challenor (1994). But this parameter is of questionable meaning when taken over all wind speeds and a much clearer picture emerges when viewing that error trend for specific  $U_{10}$  levels as shown in Table 5. This table summarizes local error estimates near 4, 8, and 12 m s<sup>-1</sup> that provide measures at the most populated moderate wind region, and for low and high wind levels where there is still a substantial data population. Continuous detail versus the reference  $U_{10}$  is provided in Fig. 10. The panels show wind error bias, standard deviation, and rms along with the linear regression correlation coefficient ( $R$ ) for wind error versus  $H_s$ .

Residual statistics versus wind speed are not typically cited in altimeter wind studies, often due to the limited

data samples and hence high uncertainty. But this information helps to ascertain small but measurable algorithm differences. Referring to Table 5 and Fig. 10, it becomes evident that overall B81 statistics of Table 4 can be misleading. The figure shows that B81 has the lowest rms error in the range of 7–9 m s<sup>-1</sup>. But highest rms levels then occur for B81 at low and high winds in step with the bias magnitude. B81 also presents lower wind error standard deviation than the other single parameter algorithms, again related to wind-dependent bias. B81 bias variation relates directly to the -0.22 m s<sup>-1</sup> error trend of Table 4 and seems consistent with Freilich and Challenor (1994) who noted that the chosen three-branch form of the Brown model leads directly to these characteristics. All indications are that B81 is a weaker choice for wind inversion than, for example, MCW or FC.

The MCW statistics exhibit constancy (bias, std,  $H_s$ -related residual) versus  $U_{10}$ , but a constant positive bias of about 0.5 to 0.6 m s<sup>-1</sup> produces a systematically elevated rms. This bias agrees with other studies (e.g., Glazman and Greyzikh 1993; Freilich and Challenor 1994). The FC model presents a lower bias as well as stable rms and bias variation over all winds. In fact, we conclude that FC is nearly equivalent to MCW aside from this 0.5 m s<sup>-1</sup> constant. For example,  $H_s$  correlation for MCW and FC in Fig. 10 is almost identical.

Perhaps the most notable observation regarding single parameter models is the measurable correlation between wind error and the altimeter-derived  $H_s$  for any wind speed. The error slope (and linear regression correlation coefficient) is strongest in the range 4–9 m s<sup>-1</sup> with values of 0.8–0.9. Correlation decreases slightly for the lowest and highest  $U_{10}$  and reaches a minimum at highest winds.

Tables 4 and 5 and Fig. 10 show that the two input algorithms provide low bias levels and consistently lower standard deviations aside from Lefèvre et al. (1994). Considering all factors, it is evident that the Lefevre model performs poorly and this algorithm will not be discussed further.

Data indicate that the GG2 model leads to wind error statistics that often match the low levels obtained with the neural network solutions. But the bias is somewhat erratic versus  $U_{10}$ . One feature of this model is the abrupt lowering of the  $H_s$  error trend for moderate wind speeds of 7–9 m s<sup>-1</sup> seen in Fig. 10. Aside from this region, the  $H_s$  impact (error trend or correlation) is nearly identical to the single parameter algorithms. This localized use of the  $H_s$  input is a feature of the two-class scheme used in that algorithm combined with the large range of possible [ $\sigma^o$ ,  $H_s$ ] combinations within this moderate wind range. Basically, the GG2 leads to a bimodal wind histogram (not shown) due to its discontinuous nature, much as for B81.

As alluded to earlier, the  $H_s$  error trend for a given  $U_{10}$  will, by design, be significantly reduced in the neural models relative to routines such as MCW or FC. Table

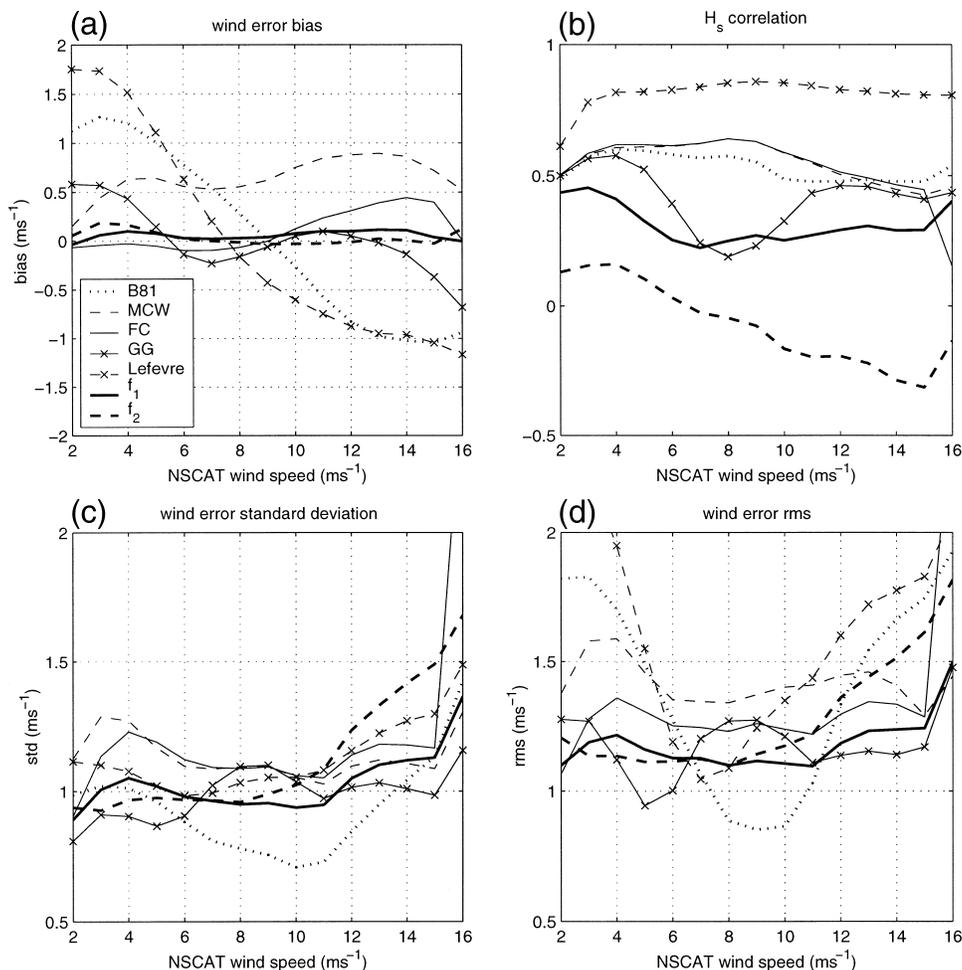


FIG. 10. Altimeter wind speed error statistics for the cited algorithms as a function of NSCAT  $U_{10}$ : (a) bias, (c) standard deviation, (b) SWH correlation with wind error, and (d) rms. Calculations are made over  $2 \text{ m s}^{-1}$  bins.

5 and Fig. 10 clearly indicate that the  $f_2$  routine strongly attenuates the  $H_s$  residual (to less than  $0.1 \text{ m s}^{-1} \text{ m}^{-1}$  while MCW and FC were generally higher than  $0.6 \text{ m s}^{-1} \text{ m}^{-1}$ ). The  $f_1$  model provides a reduction but not removal of this correlation. Recall that these two algorithms carry different weighting of  $H_s$  due to their alternate training. Both neural solutions provide a very small and stable bias across the complete wind speed range. The rms for the two solutions is nearly identical up to  $10 \text{ m s}^{-1}$  and at that point the  $f_2$  begins to increase substantially with increasing wind speed. This increase is partially related to the aforementioned multivaluedness that the neural network encounters when attempting to optimize the  $[U_{10}, H_s] \rightarrow \sigma^\circ$  mapping over both moderate and high wind speeds. The forward network model,  $f_2$ , upon inversion to wind speed, can actually lead to a negative correlation with  $H_s$  at high winds as seen in Table 5. Thus while  $f_2$  best removes the  $H_s$  dependence and may represent a better physically based model for Ku-band  $\sigma^\circ$  as a function of  $H_s$  (surrogate for wave

climate), the model is not optimal for point-to-point  $U_{10}$  inversion. This observation is strong evidence in support of the idea that  $H_s$  does not directly describe true long wave conditions that cloud the mapping between  $U_{10}$  and  $\sigma^\circ$ .

Considering all algorithms and observations including their bias, functional continuity and rms error minimization, it appears that  $f_1$  provides the best overall altimeter wind speed model. The improvement (relative to the operational MCW model) is on the order of 10%–15% in terms of the global or local reduction in the wind error standard deviation or rms. In an absolute sense, the new routine appears to lower the rms error by  $0.1\text{--}0.2 \text{ m s}^{-1}$ . Improvements are negligible for the high winds above  $12\text{--}14 \text{ m s}^{-1}$ . The bias is minimized to less than  $0.1 \text{ m s}^{-1}$  over all wind speeds. These observations are derived for the NSCAT scatterometer reference winds. Recall that though these data were also used to develop the model, the training method used only 5% of the total dataset. This insures a high level

TABLE 6. Altimeter  $U_{\text{err}}$  statistics as for Table 4, but where the reference  $U_{10}$  now changes with each row as indicated. The three TOPEX/ECMWF subsets come respectively from the TOPEX/NSCAT, TOPEX/QSCAT, and TOPEX/ERS compilations as discussed in section 2. Separation times for altimeter comparisons with scatterometer and buoy observations is  $\pm 30$  min in all cases. Computations are made over the reference wind range of 1 to 17  $\text{m s}^{-1}$ . The slope of the error trend with  $H_s$  (see Table 4) is given as  $\alpha_{H_s}$ .

Dataset	Time period	Extent (month)	No. of samples	FC			$f_1$		
				Bias	Std	$\alpha_{H_s}$	Bias	Std	$\alpha_{H_s}$
TOPEX/NSCAT	1996–97	9	48 331	0.02	1.16	0.49	0.05	0.99	0.20
TOPEX/QSCAT	1999–00	12	88 324	−0.48	1.18	0.44	−0.53	0.97	0.15
ERS/NSCAT	1996–97	9	129 701	−0.70	1.02	0.47	−0.56	0.87	0.16
TOPEX/ERS	1995–97	24	55 765	−0.07	1.08	0.27	−0.08	1.01	−0.03
TOPEX/ECMWF1	1996–97	9	231 102	−0.24	1.84	0.38	−0.20	1.77	0.14
TOPEX/ECMWF2	1999–00	12	312 102	0.03	1.46	0.37	0.03	1.43	0.08
TOPEX/ECMWF3	1995–97	24	208 518	−0.06	1.64	0.25	−0.05	1.56	0.01
TOPEX/BUOY	1992–98	75	4380	−0.32	1.45	0.56	−0.132	1.33	0.26

of independence and confidence for these conclusions within the assumption that NSCAT winds represent a valid  $U_{10N}$  reference. Attention is now given to additional algorithm validation using the numerous independent data products discussed earlier.

## 6. Further model validation

Model intercomparisons presented above suggest that the two-input altimeter algorithm,  $f_1$ , provides the best performance. The model is now evaluated against the

seven additional collocation datasets detailed earlier. These sets include independent altimeter (*ERS-2*) measurements as well as reference wind speed estimates from C- and Ku-band scatterometers, the ECMWF model, and a large open-ocean buoy compilation. The varied reference winds encompass most sources used in past altimeter studies and they each present differing spatial sampling, TOPEX time period, and scattering or model physics. The main goal here is to determine if the algorithm provides consistently improved results across these varied sources. Expanded detail in support of the following findings can be found in Gourrion et al. (2000).

Initial altimeter algorithm assessment against these data indicates that relative differences amongst the algorithms discussed in the previous section (e.g.,  $f_1$  performance vs  $f_2$ ) remain consistent with the TOPEX/NSCAT findings. Therefore presentation is condensed to compare and contrast the preferred single and two-parameter routines: FC and  $f_1$ . As mentioned earlier, FC is essentially the operational MCW routine without a 0.5  $\text{m s}^{-1}$  overall bias.

Altimeter wind error statistics are presented in Tables 6 and 7. Sample population and time period covered are listed in Table 6. Note that the buoy dataset covers the longest time period, yet contains the fewest points.

Look first to  $U_{10}$  bias values. It is evident that both the FC and  $f_1$  models provide biases that fall below 0.5  $\text{m s}^{-1}$  for all TOPEX cases. Given that the combined periods span from 1992 to 2000, this is, first of all, an indication of long-term stability for the satellite's  $\sigma^o$  calibration. For reference, a  $\sigma^o$  change of 0.1 dB corresponds to roughly 0.3  $\text{m s}^{-1}$  in wind speed for the altimeter. The *ERS-2* altimeter (*ERS/NSCAT* in the tables) exhibits a consistent 0.5 to 0.9  $\text{m s}^{-1}$  bias below TOPEX for both the FC and  $f_1$  models. This translates to a small +0.2 to 0.3 dB  $\sigma^o$  bias range between the *ERS-2* altimeter and TOPEX. Returning to the TOPEX table entries, the bias differences between FC and  $f_1$ , for any given dataset in Tables 6 and 7, are very small; in almost all cases less than 0.1  $\text{m s}^{-1}$ . Agreement between the present operational algorithm and the pro-

TABLE 7. Results follow the same format as table 6, but estimates are now localized to three separate wind speeds (4, 8, and 12  $\text{m s}^{-1}$ ) as indicated.

Dataset	FC			$f_1$		
	Bias	Std	$\alpha_{H_s}$	Bias	Std	$\alpha_{H_s}$
3–5 $\text{m s}^{-1}$						
TOPEX/NSCAT	−0.03	1.23	0.92	0.10	1.05	0.52
TOPEX/QSCAT	−0.65	1.07	0.70	−0.50	0.89	0.39
ERS/NSCAT	−0.92	0.86	0.53	−0.72	0.76	0.29
TOPEX/ERS	0.08	1.12	0.79	0.21	0.94	0.38
TOPEX/ECMWF1	−0.01	1.85	0.93	0.13	1.83	0.59
TOPEX/ECMWF2	−0.01	1.46	0.66	0.18	1.40	0.28
TOPEX/ECMWF3	0.05	1.63	0.81	0.19	1.57	0.44
TOPEX/BUOY	−0.34	1.49	0.80	−0.03	1.43	0.51
7–9 $\text{m s}^{-1}$						
TOPEX/NSCAT	−0.07	1.09	0.78	0.03	0.95	0.27
TOPEX/QSCAT	−0.47	1.10	0.72	−0.38	0.96	0.17
ERS/NSCAT	−0.77	0.89	0.67	−0.53	0.79	0.11
TOPEX/ERS	−0.11	0.91	0.54	−0.09	0.84	−0.03
TOPEX/ECMWF1	−0.34	1.72	0.94	−0.27	1.68	0.47
TOPEX/ECMWF2	−0.04	1.38	0.66	0.02	1.34	0.14
TOPEX/ECMWF3	−0.04	1.52	0.76	−0.01	1.47	0.25
TOPEX/BUOY	−0.46	1.35	0.86	−0.18	1.19	0.34
11–13 $\text{m s}^{-1}$						
TOPEX/NSCAT	0.31	1.14	0.57	0.10	1.05	0.30
TOPEX/QSCAT	−0.49	1.14	0.54	−0.70	1.04	0.20
ERS/NSCAT	−0.29	1.10	0.48	−0.40	1.01	0.18
TOPEX/ERS	−0.20	0.97	0.44	−0.43	0.90	0.12
TOPEX/ERS2	−0.14	0.76	0.39	−0.38	0.68	0.05
TOPEX/ECMWF1	−0.39	1.78	0.68	−0.58	1.69	0.45
TOPEX/ECMWF2	0.13	1.53	0.59	−0.14	1.50	0.29
TOPEX/ECMWF3	−0.21	1.52	0.51	−0.47	1.50	0.20
TOPEX/BUOY	−0.08	1.59	0.88	−0.05	1.41	0.63

posed  $f_1$  implies that a switch to use of the latter model should be seamless.

One possible outlier with respect to the altimeter wind speed bias (for both FC and  $f_1$ ) comes in the QSCAT comparison. Here the TOPEX bias is consistently below other sets by about  $0.5 \text{ m s}^{-1}$ . Recall that QSCAT is a Ku-band scatterometer but differs from NSCAT in several aspects, the most important being the scatterometer model function. We noted in section 2 that the NSCAT-1 model function was purported to be  $0.3 \text{ m s}^{-1}$  too low and that the QSCAT-1 model is without this bias. Raising the NSCAT-1 model by  $0.3 \text{ m s}^{-1}$  would bring TOPEX closer to QSCAT but also, consequently, bias TOPEX above most other table entries. The weight of the other data presented here, including possible time-dependent sensor and model bias variations, is not adequate to accept or dismiss bias adjustments at this small level. Averaging the bias values over all TOPEX entries of Table 7, one concludes that the absolute wind speed bias level is below  $0.3 \text{ m s}^{-1}$  for both  $f_1$  and FC algorithms.

Reduction in the error variation versus  $H_s$  is the prime indicator of a difference between the FC and  $f_1$  algorithms. The reduction is evident for all  $\alpha_{H_s}$  entries in Table 7. This is consistent with the results of the previous section and with the goal of attenuating long-wave impact on the altimeter wind speed. Some disparity between particular reference sets is evident, but the reduction is not. The values of FC  $\alpha_{H_s}$  range from 0.6 to 0.9 and the  $f_1$  model lowers these levels on average by factors of 1.8, 3.5, and 2.4 at 4, 8, and  $12 \text{ m s}^{-1}$ , respectively. Thus, regardless of possible sampling or product disparities, the sea state-related impact is always measurably decreased in a global average sense. Similar results hold for those wind levels not shown, consistent with Fig. 10.

Finally, the rms wind error is assessed using the standard deviation to directly compare (without bias) FC and  $f_1$  results. Reference winds lead to some clear differences here for the TOPEX data sets. Scatterometer-based comparisons using NSCAT, QSCAT, and ERS, all provide the lowest std levels, generally between  $0.9$  and  $1.2 \text{ m s}^{-1}$ . As before,  $f_1$  output reduces the std by 10% to 15% overall and at the noted wind speeds. The TOPEX-ECMWF comparisons show std levels of 1.5–1.8, an increase of 30%–45% over results obtained using the scatterometers. This rise is consistent with the expected increase in intercomparison noise due to ECMWF's inherently larger time and space averaging as predicted by Freilich and Dunbar (1993). Moreover, the std reduction obtained using  $f_1$  is not so dramatic, only 3%–5%. So while the  $H_s$  dependence ( $\alpha$ ) is clearly attenuated in the ECMWF comparisons using the  $f_1$  algorithm, the elevated intercomparison noise tends to mask the precision gained. Note that the ERS/NSCAT comparison usually indicates the lowest std levels, and these values are also comparable to those of the TOPEX/ERS computations. The common factor between these two

datasets, not found in the others, is the use of a 50-km scatterometer wind vector cell. The lower std levels for these table entries may be associated with a possible reduction in the scatterometer's inherent wind estimate noise. Regardless, the nominal 10%–15% reduction in std is still found when using  $f_1$  in lieu of the FC algorithm.

#### a. Buoy comparison

Comparison between TOPEX-derived wind speed and buoy measurements represents our only in situ validation. For this reason, greater detail is presented here. Further information on this particular data set can be found in Gommenginger et al. (2002). Findings to follow were also affirmed by comparing GEOSAT altimeter data to NDBC buoy observations in Gourrion et al. (2000). Comparison statistics are listed in Tables 6 and 7. The results are generally consistent with the findings above. Note however that std values are now of the order of  $1.4 \text{ m s}^{-1}$  instead of the  $1.0 \text{ m s}^{-1}$  levels found for the altimeter/scatterometer validation datasets. This may be attributed to differences in measurement techniques, sampling methods and/or in the time-space collocation criteria. The std and residual  $H_s$  trend ( $\alpha_{H_s}$ ) display a reduction between the one- and two-parameter algorithms, which is consistent with previous validation results. In this case, the improvement in the overall std with the additional parameterization on  $H_s$  is of the order of 10%.

Figure 11 presents statistics versus buoy wind speed over the range of  $1$ – $17 \text{ m s}^{-1}$  calculated in  $2 \text{ m s}^{-1}$  wide wind bins incremented by  $0.5 \text{ m s}^{-1}$ . Display format follows Fig. 10 and results are shown for four altimeter model functions as indicated. Comparison of Figs. 11 and 10 suggests the altimeter wind error is nearly invariant between these two reference wind products for wind speeds above  $4$ – $5 \text{ m s}^{-1}$ . As before, one sees the FC and MCW models track each other aside from a  $\sim 0.5 \text{ m s}^{-1}$  bias. GG2 compares well with the  $f_1$  results in all aspects. These two input models once again outperform the single parameter mapping. In Fig. 11 and Tables 6 and 7 the  $f_1$  model returns the smallest and most stable wind error bias over the full wind speed range, with retrieved winds only slightly underestimated (by  $0.2 \text{ m s}^{-1}$ ) at intermediate wind speeds. The std for GG2 is lowest of all models at low wind speed, although the neural network-derived model is alone in achieving a near-zero bias. Here again we observe that the two-parameter models reduce, but do not eliminate, the residual dependence on  $H_s$ .

Beside the statistical tests presented above, the ability of altimeter wind speed algorithms to reproduce the histogram of the original validation dataset is a good assessment of the models' overall performance. Accurate recovery of the wind histogram is particularly important for climatology studies. Hence, the performance of the two-parameter  $f_1$  model is compared to the MCW mod-

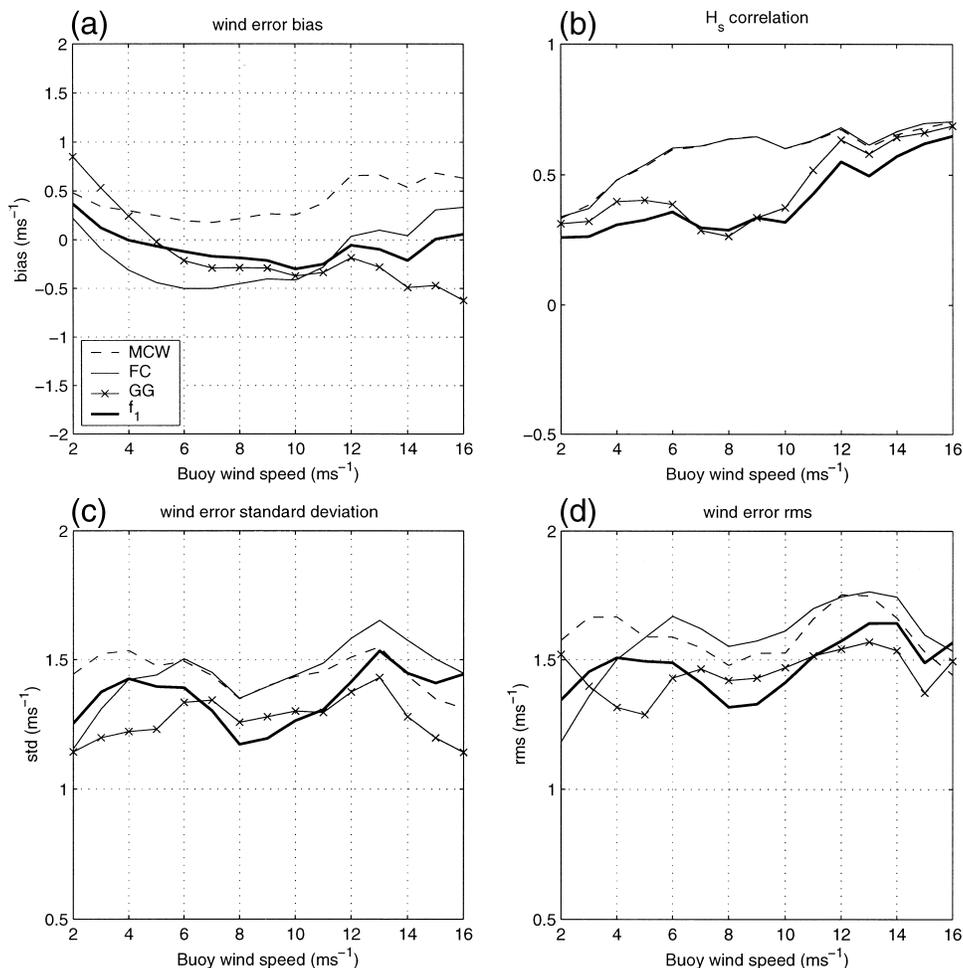


FIG. 11. Altimeter wind error statistics using buoy measurements as the reference wind. Computations and data display follow those for Fig. 10.

el, as this is the model currently used for operational wind speed retrieval on TOPEX and ERS altimeters.

The retrieved wind speed histograms are compared to the buoy result in Fig. 12. Visual improvement is evident for  $f_1$  with respect to MCW. The goodness of fit is estimated via the computed correlation between

the buoy and model wind speed histograms. One finds a correlation of 0.989 for MCW and 0.994 for  $f_1$ . The improved fit is most noticeable at intermediate to high wind speeds ( $U_{10} > 8 \text{ m s}^{-1}$ ). Consequently, the two-parameter  $f_1$  model is expected to return more accurate global wind fields than the current operational MCW algorithm.

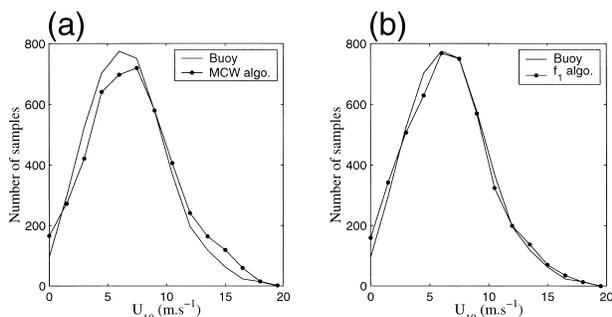


FIG. 12. Frequency distributions of wind speed retrieved with (a) MCW and (b)  $f_1$  along with the buoy-derived result. Bin size is  $1.5 \text{ m s}^{-1}$ .

*b. Residual sea-state effects*

The availability of in situ wave period measurements within the collocated TOPEX/buoy dataset permits further investigation of the effect of sea state on altimeter retrieved wind speed. From this dataset a dependence of altimeter wind speed on wave age has been reported (Gommenginger et al. 2002) and MCW winds shown to systematically underestimate winds in underdeveloped sea conditions (wave age  $< 1.5$ ). Here we assess if wave age dependence is ameliorated by the parameterization with  $H_s$  in the  $f_1$  model. Figure 13 represents the retrieved wind error  $U_{err}$  calculated for MCW and

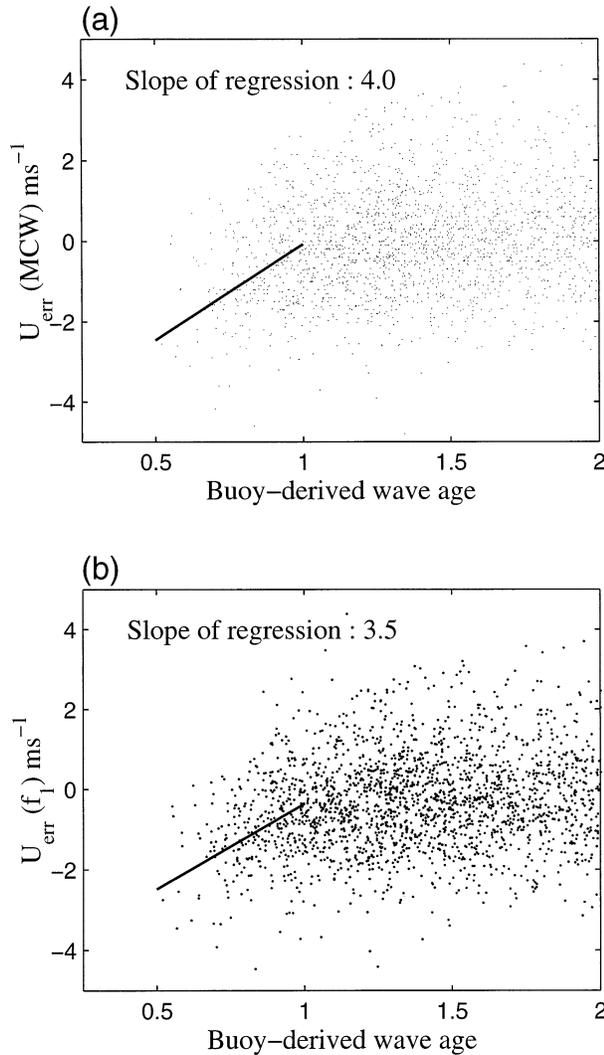


FIG. 13. Residual dependence on buoy-derived wave age ( $\zeta$ ) observed in the retrieved wind error calculated with (a) MCW and with (b)  $f_1$ .

$f_1$  against the buoy-derived wave age calculated from the peak wave period,  $T_p$ , as  $\zeta = gT_p(2\pi U_{10})^{-1}$ . A dependence of the wind error on wave age is clearly observed for young seas ( $\zeta < 1$ ) for both altimeter wind speed algorithms. No residual dependence on wave age is observed for either model for wave age greater than 1.5.

The residual wave age dependence for young seas is quantified with the linear regression coefficient of  $U_{\text{err}}$  against  $\zeta$  calculated for  $\zeta < 1.0$ . We find that the parameterization with  $H_s$  used in the  $f_1$  model yields a marginal reduction in the wave age trend, from 4.0 for MCW to 3.5  $\text{m s}^{-1}$  per unit of wave age for  $f_1$ . However, the dependence of altimeter wind speeds on wave age in underdeveloped sea conditions certainly remains. Thus the addition of  $H_s$  into the algorithm falls far short of correcting the known underestimation associated with

young seas. It is again notable that only 7% of the total TOPEX/buoy samples over this period of 1992–98 have  $\zeta < 1.0$ . This fraction is consistent with the global observations (see Fig. 1) discussed earlier.

## 7. Summary

This study defines and validates a two-input altimeter wind speed algorithm applicable for operational use, where a Ku-band altimeter's coincident  $\sigma^o$  and  $H_s$  estimates are utilized in the point-to-point inversion. An analytical formulation (termed  $f_1$ ) is prescribed with nine coefficients as detailed in section 4. Motivation comes from the new capability to assemble large, globally distributed and high fidelity model training sets composed of coincident satellite altimeter and scatterometer crossovers. The dataset chosen for model training and subsequent validation is a 96 000 sample compilation of TOPEX and NSCAT crossings. Limiting NSCAT usage to only higher incidence angle retrievals strengthens our assumption that the scatterometer wind product is itself free of sea-state impacts. Subsequent validations using buoy and ECMWF winds provide further support.

The empirical development is focused to define an improved and robust wind inversion that incorporates  $H_s$  into the solution. This routine should be applicable for all Ku-band altimeters such as those aboard the ERS, ENVISAT, GFO, and *Jason-1* platforms.  $f_1$  intercomparison to past altimeter models and numerous independent validations demonstrates modest, but measurable, success in improving upon the current operational MCW model. These independent data sources include an extensive buoy compilation, the ERS scatterometer, the SeaWinds scatterometer, and the ECMWF model. The  $f_1$  inversion ( $[\sigma^o, H_s] \rightarrow [U_{10N}]$ ) delivers an overall rms improvement of 10–15%, 0.1 to 0.2  $\text{m s}^{-1}$  in absolute terms. The domain for model application covers all values of  $H_s$  and wind speeds ranging from 1 to 20  $\text{m s}^{-1}$ . Error statistics were evaluated over the range of 1–17  $\text{m s}^{-1}$ . Wind speed bias is below 0.3  $\text{m s}^{-1}$  throughout this range. Improvement in rms error is significant up to winds of about 12  $\text{m s}^{-1}$  and equivalent to MCW above this point. The weighting of  $H_s$  within the model becomes negligible at these high wind levels. While wind speeds above 20  $\text{m s}^{-1}$  are infrequent, a slight modification of  $f_1$  that aligns the altimeter inversion with that predicted by the *QSCAT-1* model function is proposed in appendix B. Statistically, the GG2 algorithm provides similar improvement, but we recall that this classification scheme leads to point-to-point estimate discontinuities and a bi-modal wind speed distribution. TOPEX wind speed histograms, derived using  $f_1$ , provide marked improvement over the MCW result in comparison to either buoy, model or scatterometer results. This affirms the continuous nature of the network solution and its applicability for operational use. The TOPEX-generated model is also shown to work

well when applied to *ERS-2* altimeter data. The model is adjusted for differing altimeters using a constant  $\sigma^\circ$  bias and is also relatively insensitive to  $H_S$  estimate errors.

Physically, one expects that  $H_S$  is a limited proxy for actual gravity wave slope variations that affect the nominal relation between the observed radar cross section and wind speed. Observations and model functions depict measurable correlations amongst the three variables [ $\sigma^\circ$ ,  $H_S$ ,  $U_{\text{NSCAT}}$ ] but they also exhibit a multivaluedness that inhibits further wind estimate improvement in the absence of additional surface roughness information. This ambiguity leads to an increased altimeter wind speed noise if one attempts to use the forward ( $f_2$ ) model for  $U_{10}$  inversion. This effect occurs even though  $f_2$  is the most effective at removing residual error associated with  $H_S$ . Another repercussion is the inability of the  $f_1$  model to correct altimeter wind underestimation during fetch-limited events. These events are rare within the global data set and the neural network minimization solution gives little weight (by design) to this mapping within the domain of possible outcomes.

These points emphasize that present empirical and global-mean model functions do not fully capture the scattering physics. Their application to specific case study will not dramatically improve upon results derived using the MCW model. It is clear, however, that the  $f_1$  algorithm is a measurable improvement that can directly replace the single parameter routine. As importantly, the documented correlation between  $\sigma^\circ$  and  $H_S$  at any chosen wind speed needs to be considered within the empirical modeling of the altimeter sea state bias correction (cf. Chapron et al. 2001). Future work combining the altimeter's unique coincident measure of  $H_S$  with multifrequency  $\sigma^\circ$  (e.g., at S-, C-, Ka- and/or Ku-band) signatures is certain to bring further refinement to these geophysical inversions on both global and local scales.

*Acknowledgments.* Our thanks to the reviewers and to N. Tran for help in the neural network development. This work was supported by ESA, CNES, and NASA's Office of Earth Science as part of TOPEX/Poseidon and *Jason-1* science working team efforts.

## APPENDIX A

### Neural Network Methodology

A multilayer perceptron (MLP) is a well documented neural architecture that, when combined with backpropagation of errors, provides a powerful and compact solution to many intractable problems (e.g., Hagan et al. 1996; Sterling and Pollack 1996). Advantages of the backpropagation MLP include a small solution network and quick (forward) computational speed that permits training over a large input vector set. The main disadvantage is the many variables which must be considered when constructing a MLP. This includes the number of

hidden layers, the type of transfer function(s), the initial conditions, and the types of backpropagation MLPs available. One must also consider the training time which is a direct function of training set size and MLP chosen for the task.

For this study, it was empirically determined that adequate characterization is found using a three layer MLP consisting of an input layer, a single hidden layer, and an output layer. Two nodes represent the two inputs [ $U_{10}$ ,  $H_S$ ] for  $f_1$  or [ $\sigma^\circ$ , and  $H_S$ ] for  $f_2$ . The hidden layer has two nodes and the output layer has one,  $U_{10}$  or  $\sigma^\circ$ . A log-sigmoid transfer function,  $f_{\text{net}}$ , of the form

$$f_{\text{net}} = [1 + \exp^{-(WA+B)}]^{-1} \quad (\text{A1})$$

is used between both the hidden and output layers where  $\mathbf{A}$  is the input vector and  $W$  and  $B$  are the network weights and biases. A requisite rescaling of all input and output variables normalizes the data to run between 0.1 and 0.9 due to the asymptotic limits of this function.

The MLP network developed here is not trained in the traditional manner of dividing the set in two and using one set for training and the other set for independent testing. This is due to the sparse distribution of data at the extremes in the data grid and how these points are underrepresented during neural training. Since a function's sensitivities (partial derivatives with respect to all network weights and biases) are added during training (Hagen et al. 1996), large differences in data distribution across the training domain can potentially lead to solutions that are biased towards higher density regions. It is clear from the wind speed distributions shown in section 2 that there are many more training samples near  $7 \text{ m s}^{-1}$  than, for example, near  $2$  or  $18 \text{ m s}^{-1}$ . In turn, it was empirically determined that the nonunique nature (see below) of the proposed mapping between these three chosen parameters leads to a solution that favors minimization near the densest data population at the expense of low and high wind regimes. To compensate for these effects, we equalize the distribution of points placed into any given training set, effectively compressing the training set size at the same time. The size reduction enables a large portion of data to be held aside for validation and permits use of the following backpropagation method to speed solution.

Training set equalization involves partitioning the data into 21 equally spaced wind speed bins spanning from  $0$  to  $20 \text{ m s}^{-1}$ . Only 200 randomly-selected sample data pairs from each bin are allowed into any given training vector set. Any bin with fewer than 200 samples has all its points included. All data samples for wind speeds above  $20 \text{ m s}^{-1}$  are also included in any given training set. This equalization procedure focuses solely on the wind speed rather than, for example, a three-dimensional equalization across the triplet [ $\sigma^\circ$ ,  $H_S$ ,  $U_{10}$ ] dataset. This more rigorous partitioning was performed and no measureable difference was found. In fact, slight degradation in model performance generally resulted.

An inherent multivaluedness, to be discussed later, is responsible for much of this behavior.

Final network coefficients were determined using the following procedure. A loop is initialized that converges to an optimized solution by creating network coefficients for each of many random draws of equalized training subsets (5% of the total set), each subset extracted as described above. An enhanced error back-propagation technique, called the Levenberg–Marquardt algorithm, is used to increase training speed. For each training set this Levenberg–Marquardt algorithm is run towards a convergence limit easily obtained after 10 epochs. At this point training is halted and a “goodness-of-fit” (Kohler 1988) test computed using the  $\chi^2$  test statistic,  $K$ , of the form

$$K = \sum_{i=1}^n (y_i - x_i)^2/x_i, \quad (\text{A2})$$

where  $y$  is the model output data distribution,  $x$  is the known output distribution, and  $i$  represents a distribution bin.

The Levenberg–Marquardt algorithm alone does a good job minimizing the msse (mean sum squared error). For almost any randomly drawn data set the msse will converge to nearly the same minimum value. However, what we desire here is a mapping that both minimizes the global error variance as well the error bias and skewness across the entire output variable domain. This additional step of computing a  $\chi^2$  test statistic between the known and neural-calculated distributions serves as an additional optimization criterion to the Levenberg–Marquardt training. The combined  $\chi^2$  goodness-of-fit and msse optimizations lead to computed and known data distributions that are equalized and a solution where bias and variance variation are minimized. Note that this training method provides a solution that is ultimately derived from a small subset of the total data; that is, any given model solution is only derived from one equalized training subset containing roughly 4500 data points.

## APPENDIX B

### Correction to $f_1$ for High Wind Speeds

Young (1993) addressed the special case of altimeter wind inversion at high winds by examining those rare events when the GEOSAT altimeter swath crossed over strong cyclones. Ground truth was inferred using modeled cyclone winds for six such storms within the period of 1987–89. The resulting formulation, applicable above  $20 \text{ m s}^{-1}$ , relates GEOSAT Ku-band backscatter to wind speed under a simple linear model:

$$U_{10} = -6.4\sigma^o + 72. \quad (\text{B1})$$

This model was developed to extend an altimeter’s wind inversion domain of validity beyond limits obtainable under past studies such as Witter and Chelton

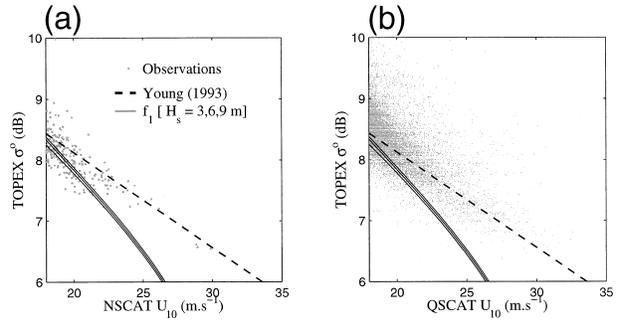


FIG. B1. TOPEX altimeter  $\sigma^o$  measurements versus scatterometer wind speed for winds above  $18 \text{ m s}^{-1}$ . (a) NSCAT and (b) QSCAT observations are presented. Satellite measurements were coincident to within 15 km and 30 min.

(1991). Typical upper limits for acquiring reliable reference winds in these efforts has been  $15\text{--}20 \text{ m s}^{-1}$ . Scatterometer data sets within the present study offer the opportunity to corroborate the Young findings. In particular, the winds derived from SeaWinds using the QSCAT-1 model function are purported to be a validated improvement (e.g., Donnelly et al. 1999) over past scatterometer inversions at these high wind speed levels. Assuming this is the case, high wind data from both the TOPEX/NSCAT and TOPEX/QSCAT compilations are presented in Fig. B1. In total, the 30-min filtered observations yield 247 points for the former and 4800 samples for the latter for events where winds exceed  $18 \text{ m s}^{-1}$ . These amounts represent less than 0.3% and 1.5% of the respective global data sets. The discrepancy between the number of high wind samples for NSCAT and QSCAT sets is unresolved at this point pending further validation of the scatterometer models at high winds. Note that these data are for clear sky conditions as indicated by the TOPEX radiometer and scatterometer data flags.

Substantial scatter is evident in the TOPEX  $\sigma^o$  observations versus  $U_{10}$ , but a linear decay with increasing wind is apparent. Young (1993) and  $f_1$  models are also shown. Note that three  $f_1$  curves are given, spanning  $H_s$  of 3–9 m, and that this changes little in the model output at these winds.

The neural network training used to develop  $f_1$  was presented with the NSCAT data of Fig. B1 but model constraint was weighted strongly towards wind speeds less than  $20 \text{ m s}^{-1}$ . Thus, it is not unexpected to see divergence between  $f_1$  and these observations. What is remarkable is the agreement between data (for both NSCAT and QSCAT references) and the model of Young (1993). Computation of a linear least squares fit through the TOPEX/NSCAT and TOPEX/QSCAT data sets for these high wind values agrees well with Eq. (B1). Standard GEOSAT-to-TOPEX comparison procedure dictates adjusting  $\sigma^o$  levels by 0.63 dB when applying a GEOSAT-era model like Young’s to TOPEX data. But these results do not include that bias. This  $\sigma^o$  discrepancy falls within the cyclone wind model’s es-

timated margin of error (Young 1993) and agreement here suggests that no adjustment be made for TOPEX application.

Conclusions are then that for high wind speeds above  $18 \text{ m s}^{-1}$ , the  $f_1$  model will give increasingly negative wind speed biases; spanning from below 0.5 to as much as  $5 \text{ m s}^{-1}$  for mean values from 18 to  $30 \text{ m s}^{-1}$ , respectively. The model of Young (1993) can be directly applied to TOPEX data to remedy this trend when looking at cyclone-level values up to at least  $30 \text{ m s}^{-1}$ .

## REFERENCES

- Benada, R., 1997: Merged GDR (TOPEX/Poseidon) generation B: User's handbook. D-11007, Jet Propulsion Laboratory, PO-DAAC, 172 pp.
- Brown, G., H. Stanley, and N. Roy, 1981: The wind speed measurement capability of spaceborne radar altimeters. *IEEE J. Oceanic Eng.*, **6**, 59–63.
- Callahan, P. S., C. S. Morris, and S. V. Hsiao, 1994: Comparison of Topex/Poseidon  $\sigma^0$  and scientific wave height distributions to Geosat. *J. Geophys. Res.*, **99** (C12), 25 015–25 025.
- CERSAT, 1996: Off-line wind scatterometer ERS products. User's handbook, IFREMER, C2-MUT-W-01-IF, 86 pp.
- , 1997: Collocated NSCAT/ERS products. User's handbook, IFREMER, NSCAT-MUT-S-01-IF, 20 pp.
- Chapron, B., D. Vandemark, T. Elfouhaily, D. Thompson, P. Gaspar, and S. Labroue, 2001: Altimeter sea state bias: A new look at global range error estimates. *Geophys. Res. Lett.*, **28**, 3947–3950.
- Chelton, D., and P. McCabe, 1985: A review of satellite altimeter measurement of surface wind speed: With a proposed new algorithm. *J. Geophys. Res.*, **90**, 4707–4720.
- , J. C., Ries, B. J. Haines, L.-L. Fu, and P. S. Callahan, 2001: Satellite altimetry. *Satellite Altimetry and Earth Sciences*, L.-L. Fu and A. Cazenave, Eds., Academic Press, 1–131.
- Dobson, E., F. Monaldo, J. Goldhirsh, and J. Wilkerson, 1987: Validation of Geosat altimeter derived wind speeds and significant wave heights using buoy data. *J. Geophys. Res.*, **92**, 10 719–10 731.
- Dobson, F., 1991: Review of reference height and averaging time of surface wind measurements at sea. Rep. 3, Marine Meteorology and Related oceanographic Activities, WMO, 64 pp.
- Donnelly, W. J., J. R. Carswell, R. E. McIntosh, P. S. Chang, J. Wilkerson, F. Marks, and P. G. Black, 1999: Revised ocean backscatter models at C and Ku band under high-wind conditions. *J. Geophys. Res.*, **104** (C5), 11 485–11 497.
- Dunbar, R. S., 1997: NASA scatterometer, high-resolution merged geophysical data product: User's guide. Ver 1.1, Jet Propulsion Laboratory, PO-DAAC, 17 pp.
- Freilich, M., and R. Dunbar, 1993: Derivation of satellite wind model functions using operational surface wind analyses: An altimeter example. *J. Geophys. Res.*, **98**, 14 633–14 649.
- , and P. Challenor, 1994: A new approach for determining fully empirical altimeter wind speed model functions. *J. Geophys. Res.*, **99**, 25 051–25 062.
- , and R. Dunbar, 1999: The accuracy of the NSCAT-1 vector winds: Comparisons with NDBC buoys. *J. Geophys. Res.*, **104**, 11 231–11 246.
- Glazman, R., and S. Pilorz, 1990: Effects of sea maturity on satellite altimeter measurements of surface wind. *J. Geophys. Res.*, **95**, 2857–2870.
- , and A. Greysukh, 1993: Satellite altimeter measurements of surface wind. *J. Geophys. Res.*, **98**, 2475–2483.
- Gommenginger, C., M. Srokosz, P. Challenor, and P. Cotton, 2002: Development and validation of altimeter wind speed algorithms using an extended collocated buoy/TOPEX dataset. *IEEE Trans. Geosci. Remote Sens.*, **40**, 251–260.
- Gourrion, J., D. Vandemark, S. Bailey, and B. Chapron, 2000: Satellite altimeter models for surface wind speed developed using ocean satellite crossovers. IFREMER Tech. Rep. DROOS-2000-02, 61 pp.
- Gower, J., 1996: Intercalibration of wave and wind data from TOPEX/Poseidon and moored buoys off the west coast of Canada. *J. Geophys. Res.*, **101**, 3817–3830.
- Hagan, M., H. Demuth, and M. Beale, 1996: *Neural Network Design*. PWS Publishing, 637 pp.
- Hwang, P., W. Teague, G. Jacobs, and D. Wang, 1998: A statistical comparison of wind speed, wave height and wave period derived from satellite altimeters and ocean buoys in the gulf of Mexico region. *J. Geophys. Res.*, **103**, 10 451–10 468.
- Kohler, H., 1988: *Essentials of Statistics*. Scott, Foresman and Company, 526 pp.
- Lefèvre, J., J. Barckicke, and Y. Ménard, 1994: A significant wave height dependent function for TOPEX/Poseidon wind speed retrieval. *J. Geophys. Res.*, **99**, 25 035–25 046.
- Monaldo, F., and E. Dobson, 1989: On using significant wave height and radar cross section to improve radar altimeter measurements. *J. Geophys. Res.*, **94**, 12 699–12 701.
- Nghiem, S. V., F. K. Li, and G. Neumann, 1997: The dependence of ocean backscatter at Ku-band on oceanic and atmospheric parameters. *IEEE Trans. Geosci. Remote Sens.*, **35**, 581–600.
- Queffeuilou, P., B. Chapron, and A. Bentamy, 1999: Comparing Ku band NSCAT scatterometer and ERS-2 altimeter winds. *IEEE Trans. Geosci. Remote Sens.*, **37**, 1662–1670.
- Sterling, T., and S. Pollack, 1996: *Introduction to Statistical Processing*. Prentice-Hall, 663 pp.
- Sverdrup, H., and W. Munk, 1947: Wind, sea, and swell: Theory of relations for forecasting. Tech. Rep. 1, H.O. Publ. 601, U.S. Hydrographic Office, 44 pp.
- Tournadre, J., and S. Blanquet, 1994: Wind and wave mesoscale variability from in situ and altimeter data. *Global Atmos. Ocean Syst.*, **2**, 221–247.
- Wentz, F., and D. Smith, 1999: A model function for the ocean normalized radar cross section at 14 ghz derived from NSCAT. *J. Geophys. Res.*, **104**, 11 499–11 514.
- Witter, D., and D. Chelton, 1991: A Geosat altimeter wind speed algorithm and a method for altimeter wind speed algorithm development. *J. Geophys. Res.*, **96**, 8853–8860.
- Wu, J., 1999: On wave dependency of altimeter sea returns—Weak fetch influence on short ocean waves. *J. Atmos. Oceanic Technol.*, **16**, 373–378.
- Young, I., 1993: An estimate of the Geosat altimeter wind speed algorithm at high wind speeds. *J. Geophys. Res.*, **98**, 20 275–20 286.
- , 1999: Seasonal variability of the global ocean wind and wave climate. *Int. J. Climatol.*, **19**, 931–950.