

# Development and Validation of Altimeter Wind Speed Algorithms Using an Extended Collocated Buoy/Topex Dataset

Christine P. Gommenginger, Meric A. Srokosz, Peter G. Challenor, and P. David Cotton

**Abstract**—The development and validation of altimeter wind speed algorithms is investigated following the collation of the largest dataset to-date of coincident altimeter/buoy open ocean measurements. Nonlinear relationships between buoy wind and Topex backscatter are fitted to the 4500 points dataset using least-squares (LSQ). The addition of altimeter significant wave height (SWH) information causes a small but significant reduction of about 10% in root-mean-square (rms) error. The new LSQ algorithms yield significant improvement of the global wind speed bias and rms error compared to earlier models, but describe the wind to backscatter relationship poorly at extreme wind speeds. Best results are obtained with the Gourrion *et al.* (2000) model, improving on the Witter and Chelton (WC91) (1991) model used operationally. A residual dependence on sea state persists in all wind algorithms, which underestimate winds in young sea conditions on average by 1–1.5 m/s.

A case study confirms that ordinary LSQ attribute excessive weight to the peak of the wind speed histogram and yield algorithms with poor performance at extreme winds. Measurement errors are shown to greatly influence the fitted models performance, as accounting for normally distributed errors in both altimeter and buoy measurements with orthogonal distance regressions (ODRs) yields significant improvements. Hence, algorithms developed from relatively small collocated datasets (few thousand points) may perform as well as models developed from much larger datasets (tens of thousands of points) given adequate treatment of errors. However, it is anticipated that the ultimate accuracy of wind speed algorithms is still dependent on the quality of the fitted datasets.

**Index Terms**—Algorithm development, altimeter wind speed, collocated dataset, wave age.

## I. INTRODUCTION

FOR almost two decades now, satellite altimeter radars have provided quantitative information on wind speed, significant wave height (SWH), and sea surface height (SSH) on a global scale. Despite small along-track coverage (2–7 km diameter footprint), altimeter wind and wave measurements offer a valuable contribution to global ocean circulation and climatological studies. Accurate estimates of wind speed and SWH are particularly important as they currently form the basis for correcting the SSH measurements for sea state bias errors resulting from the presence of ocean waves on the surface [1]. However,

while it is generally accepted that altimeter SWH is now of comparable accuracy to that of moored buoys [2], [3], the issue of altimeter wind speed retrieval has remained an active area of research.

The retrieval of altimeter wind speed from microwave backscatter at nadir has been the object of numerous studies and many algorithms have been proposed in the past (e.g., [4]–[6]). The algorithm currently used to operationally provide wind estimates is the Witter and Chelton model (WC91)[7]. Derived empirically from a large dataset of Geosat altimeter data, it uses a lookup table to relate the wind speed at 10 m,  $U_{10}$ , singly to the altimeter backscatter coefficient at Ku-band,  $\sigma^0$ . Wind speed variability related to nonlocal wave effects has, however, long been reported [8] and there is recent evidence of seasonal biases [9]. With the availability of altimeter SWH, several attempts have been made to account for sea state effects using SWH [10], [11], but the evidence for any improvement has never been sufficiently compelling to justify the application of a SWH parameterization to operational purposes.

In this paper, the use of SWH as a sea state parameter is re-examined with a new Southampton Oceanography Centre (SOC) dataset of collocated buoy/Topex measurements in the open ocean. This dataset and its collation are described in Section II. The development of relationships between wind speed and the altimeter measurements using nonlinear least-square (LSQ) fitting is presented in Section III together with the evaluation of their global residual statistics with respect to published models. In Section IV, further tests establish the shortcomings of the LSQ approach and Section V looks at the impact of dataset composition and measurement errors on the performance of the derived algorithms.

## II. THE SOC COLLOCATED TOPEX/BUOY DATASET

Datasets of collocated buoy/altimeter measurements have previously been used to develop and study altimeter wind retrieval algorithms ([10], [12]) but have always been limited to only a few hundred data points. The difficulties in collating large collocated datasets arise primarily from the relatively short lifetime of satellite missions and the general scarcity of continuous buoys records in the open ocean. Also, wind speed studies require stringent time/space collocation criteria to capture the short-term/small-scale variability of the wind field, which further reduces the amount of data acceptable for this type of research.

Manuscript received February 23, 2001; revised August 14, 2001. This work was supported in part by ESA (ESTEC) under Contract 12934/98/NL/GD.

The authors are with the James Rennell Division for Ocean Circulation and Climate, Southampton Oceanography Centre, Southampton, SO14 3ZH, U.K. (e-mail: cg1@soc.soton.ac.uk).

Publisher Item Identifier S 0196-2892(02)02124-1.

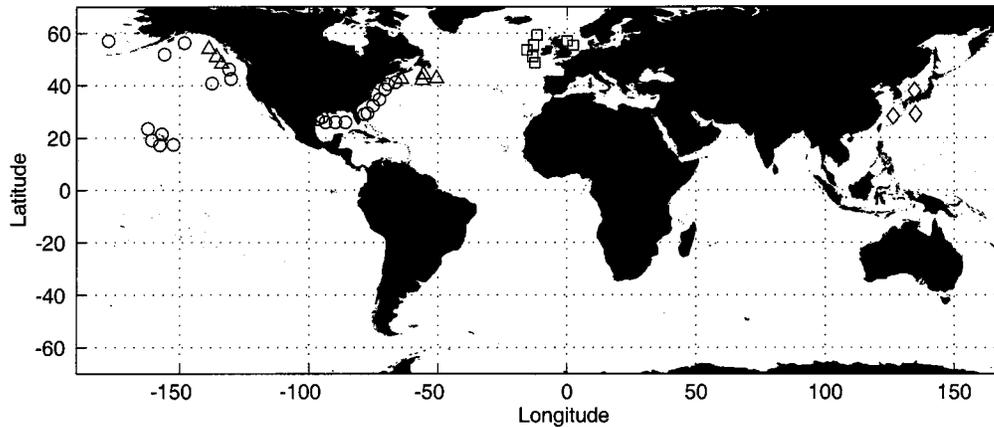


Fig. 1. Location of moored buoys used for collocation with Topex data (circles: U.S. National Data Buoy Center; triangles: Canadian Marine Environmental Data Service; squares: U.K. Meteorological Office; diamonds: Japanese Meteorological Agency).

The large collocated buoy/altimeter dataset used in this paper has been made possible following the relative longevity of the Topex altimeter and the steady increase of the number of long-term buoys in the open ocean providing continuous monitoring from which data is readily available for research. The buoy data used in this paper originates from the U.S. National Data Buoy Center (NDBC), the Canadian Marine Environmental Data Service (CMEDS), the U.K. Meteorological Office (UKMO) and the Japanese Meteorological Agency (JMA). The moored buoys used in this study were selected for their location in open water and their proximity to Topex tracks. The network of 41 moored buoys used aims to provide a representation of the global ocean (see Fig. 1), although information is still noticeably lacking in the southern hemisphere, and in particular in the Southern Ocean. We note that data are absent also in the Indian Ocean where seasonal effects are known to be significant.

A judicious choice of time/space collocation criteria is critical to ensure the quality of the final dataset. A compromise must be reached between 1) providing *in-situ* measurements representative of the conditions in the satellite footprint at the time of the overpass and 2) allowing a sufficiently large number of valid hits to allow meaningful statistical treatments. Here, the time and space separation criteria were set to select data within 50 km and 1 h of the Topex overpasses, in accordance with the criteria commonly used in this type of studies (e.g., [6], [13]). With these sampling criteria and the application of the standard ice and rain flags in the Aviso version of Topex geophysical data records (GDR), the collocation exercise yielded 4512 hits for the period between September 1992 and December 1998. Basic outlier removal was performed by eliminating any data for which the altimeter wind speed (retrieved with WC91) and the buoy wind speed differed by more than 5 m/s, resulting in a final dataset of 4444 collocated measurements.

The measurements in the collocated dataset consist of the Topex altimeter backscatter coefficient and the SWH at both Ku and C band, the buoy wind speed and direction, buoy SWH, mean and peak wave period, and air and sea temperatures. Buoy wind speed consists of the mean value computed for 8 min. Where the buoy anemometer height differed from the standard 10 m height, wind speeds were adjusted using the air-sea tem-

perature dependent correction factors reported by Dobson [14], [15]. The altimeter measurement for any given collocation point corresponds to the single 1 Hz altimeter data located closest to the buoy position within a 50-km radius. No attempt was made to compensate for the gradual drift in Topex's SWH estimates toward the end of 1998, as the magnitude of the error at the end of 1998 remained small ( $\sim 0.2$  m [16]) and affected only 10% of the dataset. Results were found to be consistent with similar computations performed previously on a smaller 1992–1997 dataset, thus confirming that the drift in Topex SWH has no perceptible impact on our analysis.

### III. EMPIRICAL WIND ALGORITHM DEVELOPMENT

Many methods have been used to derive the relationship between wind speed and altimeter backscatter. These range from multibranch polynomial fits [4] to theory-driven formulations [6], [17] or statistical histogram-matching approaches [5]. These algorithms form a disparate collection of relationships, which can be difficult to inter-compare. Here, we propose to fit our dataset with a number of simple functional forms inspired by previously suggested models. To avoid problems of spurious multipeaked histograms [18], only continuous and fully differentiable functions able to describe the full wind speed range from 0 to 20 m/s are considered. Results are presented for a small selection of the more successful models, chosen for their ease of implementation and the use of a minimum number of fitted coefficients.

#### A. Nonlinear Least-Square Fitting Procedure

The LSQ fitting to the collocated dataset of nonlinear functional forms was performed with a Levenberg–Marquardt optimization program available in the MATLAB Optimization toolbox. A bootstrap with replacement approach was adopted to provide confidence intervals for the fitted coefficients and the statistical parameters. A validation dataset of 882 points was generated by randomly extracting 20% of the dataset to enable independent assessment of the algorithms. The development dataset used for the LSQ fit consisted of the remaining 80% of the dataset augmented to the original dataset size by duplicating a random 882 of the remaining data points.

TABLE I  
WIND ERROR STATISTICS OVER THE FULL WIND SPEED RANGE FOR SINGLE PARAMETER ALGORITHMS

	Bias $\langle e \rangle$	Root-Mean-Square error $\langle e^2 \rangle^{1/2}$	Standard deviation $(\langle e^2 \rangle - \langle e \rangle^2)^{1/2}$
Witter & Chelton (1991; WC91)	0.28 [0.07]	1.50 [0.08]	1.47 [0.07]
Freilich & Challenor (1994; FC94)	-0.33 [0.07]	1.49 [0.06]	1.45 [0.07]
[M1] $a + b \sigma_{\text{Klin}}^0$	0.00 [0.11]	1.38 [0.07]	1.38 [0.07]
[M2] $a + b \exp(c \sigma_{\text{Klin}}^0)$	0.00 [0.10]	1.32 [0.07]	1.32 [0.07]
[M3] $a + b \sigma_{\text{Klin}}^0 + d \exp(e \sigma_{\text{Klin}}^0)$	0.00 [0.10]	1.32 [0.07]	1.32 [0.07]
[M4] $a + b \sigma_{\text{Clim}}^0$	0.00 [0.12]	1.46 [0.08]	1.46 [0.08]
[M5] $a + b \exp(c \sigma_{\text{Clim}}^0)$	0.00 [0.11]	1.43 [0.08]	1.43 [0.08]

The data splitting and the fitting procedure were repeated 400 times for each algorithm. In this case, the statistical parameters started to display a normal distribution after approximately 200 iterations.

For each realization, the parameters were evaluated also for four reference models: the  $\sigma^0$ -only algorithms of 1) WC91 [7] and 2) Freilich & Challenor (FC94) [5]; and the ( $\sigma^0$ , SWH) algorithms of 3) Glazman & Greysukh (GG93); and 4) Gourrion (Gr00) [19]. The GG93 algorithm refers to the “*j*-resolved” algorithm in the original paper where the altimeter backscatter coefficient and SWH serve to calculate the pseudowave age used to discriminate between two distinct  $\sigma^0$ -only models for young and mature seas. The Gr00 algorithm refers to the direct  $F1_{\text{MLP}}$  model, obtained by applying neural network techniques to a large dataset of collocated Topex altimeter/NSCAT scatterometer observations. Where algorithms were developed for Geosat data (WC91, GG93, FC94), the Topex backscatter coefficient was adjusted accordingly [20]. Similarly, where algorithms were developed for neutral stability wind speed measurements at 19.5 m height (FC94), the buoy wind speed was adjusted to 10 m using  $U_{10} = 0.93U_{19.5}$  [21].

### B. Global Wind Error Statistics

Following common practice, the performance of the LSQ algorithms was evaluated using the mean, root-mean-square error (rms) and standard deviation (std) of the wind error  $e$  defined as

$$e = U_{10\text{Alt}} - U_{10\text{Buoy}}. \quad (1)$$

The functional forms and global wind error statistics of the selected one-, two-, and three-parameter LSQ models are given in Tables I and II. Note that all functional forms presented make use of the backscatter coefficient in its linear form,  $\sigma_{\text{lin}}^0$ , which yields faster convergence (yet similar LSQ results) than the same formulations based on  $\sigma^0$  in decibels. Among the formulations in Tables I and II are model [M1] based on the power of  $\sigma^0$  law suggested by [18] and model [M3] inspired by the combination of a power law and an exponential proposed by [5] in their inverse model. The figures in brackets in Tables I and II represent the 95% confidence interval obtained from the bootstrap method. Similarly, the values of the fitted coefficients

and their 95% confidence intervals are given in Table III. The goodness of fit to the data of the various one-parameter models can be assessed from Fig. 2.

From the mean wind error (bias) results in Tables I and II, we find that all LSQ relationships remove the overall wind error bias of the order of 0.2 m/s otherwise observed in the reference models. Similarly, we find that even simple functional forms return significantly reduced wind residuals at the 95% confidence level. Similar functional forms applied to the backscatter coefficient at Ku-band ([M1] and [M2]) and at C-band ([M4] and [M5]) demonstrate the clearly degraded performance of the C-band based models, thereby confirming that, at nadir, Ku-band backscatter is better correlated with wind speed than C-band.

In Table II, the impact of the SWH parameterization is examined with algorithms [M6]–[M8] using the same backscatter dependence as  $\sigma^0$ -only models [M1]–[M3]. Here, the SWH information comes from the altimeter measurement at C-band,  $\text{SWH}_{\text{C}}$ , which correlates slightly better than  $\text{SWH}_{\text{Ku}}$  with the buoy SWH. The difference between  $\text{SWH}_{\text{Ku}}$  and  $\text{SWH}_{\text{C}}$  is within the magnitude of the error measurements so that little difference is observed when using  $\text{SWH}_{\text{Ku}}$  instead of  $\text{SWH}_{\text{C}}$ . The dependence on SWH in the LSQ models was introduced as a power law which describes well the variation of  $U_{10}$  with SWH [see Fig. 3(a)]. The SWH parameterization is seen to address some of the variability observed at intermediate wind speeds in the  $U_{10}$  to  $\sigma^0$  relationship [see Fig. 3(b)] when compared to the equivalent  $\sigma^0$ -only functional forms in Fig. 2. Best residual error results amongst two-parameter models were obtained with algorithm [M7] and [M8], which displayed an rms error reduction of the order of 10% with respect to the  $\sigma^0$ -only reference algorithms.

Finally, model [M9] represents one example of the combined use of Ku- and C-band backscatter in a three-parameter extension of the formulation used in model [M7]. From this and other tested forms (not shown), there is no evidence that any significant improvements can be achieved with an additional C-band backscatter parameterization.

On the basis of these results, best performance is achieved equally with two-parameter models [M7] and [M8]. Here,

TABLE II  
WIND ERROR STATISTICS OVER THE FULL WIND SPEED RANGE FOR TWO- AND THREE-PARAMETER ALGORITHMS

	Bias <e>	Root-mean-square error <e <sup>2</sup> > <sup>1/2</sup>	Standard deviation (<e <sup>2</sup> >-<e> <sup>2</sup> ) <sup>1/2</sup>
Glazman & Greysukh (1993; GG93)	-0.23 [0.06]	1.36 [0.06]	1.34 [0.06]
Gourrion <i>et al.</i> (2000; Gr00)	-0.12 [0.06]	1.35 [0.06]	1.34 [0.06]
[M6] $a + b \sigma_{\text{Klin}}^0 + d \text{ SWH}^e$	0.00 [0.10]	1.33 [0.07]	1.34 [0.07]
[M7] $a + b \exp(c \sigma_{\text{Klin}}^0) + d \text{ SWH}^e$	0.00 [0.10]	1.28 [0.07]	1.28 [0.07]
[M8] $a + b \sigma_{\text{Klin}}^0 + d \exp(e \sigma_{\text{Klin}}^0) + f \text{ SWH}^g$	0.00 [0.10]	1.28 [0.07]	1.28 [0.07]
[M9] $a + b \exp(c \sigma_{\text{Klin}}^0) + d \exp(e \sigma_{\text{Ctin}}^0) + f \text{ SWH}^g$	0.00 [0.10]	1.27 [0.07]	1.27 [0.07]

TABLE III  
COEFFICIENT OF THE LEAST-SQUARE FITTED FUNCTIONAL FORMS FORMULATED IN TABLES I AND II WITH 95% CONFIDENCE INTERVALS IN BRACKETS

	a	b	c	d	e	f	g
Single parameter: $\sigma_{\text{Ku}}^0$ only							
[M1]	-1.553 [0.20]	183.3 [8.6]	-1.161 [0.03]	-	-	-	-
[M2]	1.406 [0.07]	42.40 [0.7]	-0.145 [0.00]	-	-	-	-
[M3]	1.403 [0.08]	2.686 [3.3]	-11.890 [1.3]	42.38 [0.7]	-0.145 [0.00]	-	-
Single parameter: $\sigma_{\text{C}}^0$ only							
[M4]	0.341 [0.12]	3249 [368]	-1.806 [0.04]	-	-	-	-
[M5]	1.956 [0.07]	72.88 [1.1]	-0.087 [0.00]	-	-	-	-
Two parameters: $\sigma_{\text{Ku}}^0$ and SWH							
[M6]	-0.890 [0.2]	285.6 [19.6]	-1.347 [0.03]	-0.049 [0.01]	2.071 [0.12]	-	-
[M7]	1.928 [0.1]	52.835 [1.1]	-0.158 [0.00]	-0.308 [0.06]	1.145 [0.04]	-	-
[M8]	1.918 [0.1]	3.043 [5.2]	-12.17 [1.5]	52.83 [1.1]	-0.158 [0.00]	-0.299 [0.06]	1.162 [0.1]
Three parameters: $\sigma_{\text{Ku}}^0$ , $\sigma_{\text{C}}^0$ and SWH							
[M9]	2.031 [0.1]	47.16 [2.0]	-0.160 [0.00]	12.654 [3.6]	-0.095 [0.00]	-0.330 [0.06]	1.142 [0.07]

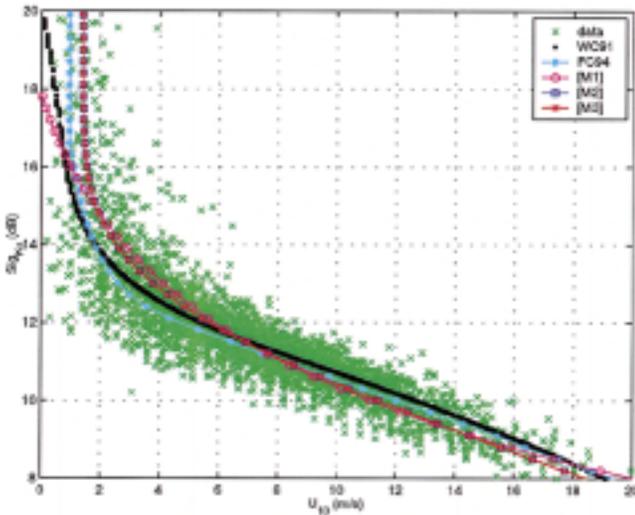


Fig. 2. Wind speed against altimeter Ku-band backscatter coefficient for collocated buoy/Topex dataset, two reference models (WC91, FC94) and three LSQ  $\sigma^0$ -only formulations.

model [M7] is shown a marginal preference given its smaller number of fitted parameters. Model [M7] is used in [22] as the empirical model selected for comparison with theoretical models predictions at nadir.

#### IV. WIND ALGORITHM PERFORMANCE: A CLOSER LOOK

Although global statistics are used traditionally to compare the performance of models, further tests are advisable to assess

the validity of altimeter wind speed algorithms. In particular, it is important to ensure that the performance of LSQ models are maintained over the whole range of wind speeds.

##### A. Local Wind Error Statistics

We start by looking at the wind error statistics calculated locally for individual wind speed bins. Fig. 4 represents the wind bias and std calculated over 1.5 m/s wide wind speed bins for three  $\sigma^0$ -only models (WC91, FC94, and model [M3]) and three ( $\sigma^0$ , SWH) models (GG93, Gr00, and model [M7]).

In Fig. 4(a), the Gr00 model produces the most consistent near-zero bias over the full wind speed range. The WC91 model displays a positive bias ranging from 0.2 m/s at low winds (<10 m/s) to 0.7 m/s at high winds (>10 m/s), thus confirming previous findings that the WC91 algorithm systematically overestimates all wind speeds. In contrast, the FC94 model underestimates wind speeds by as much as 0.5 m/s for low winds between 4 and 10 m/s, while giving satisfactory results for higher winds. Model GG93 display a monotonically decreasing trend which underestimates wind speeds by 0.5 m/s or more for most of the wind range ( $U_{10} > 6$  m/s). LSQ models [M3] and [M7] display biases at low winds markedly larger than any other algorithms. This can be traced to the poorer fit of the curvature of the  $U_{10}$  to  $\sigma^0$  relationship around  $U_{10} = 4$  m/s [see Figs. 2 and 3(b)].

The wind error std in Fig. 4(b) broadly falls between 1.0 and 1.7 m/s for all models, and for all models the larger variability is observed at wind speeds around 13 m/s (for reasons that are not clear). The LSQ models [M3] and [M7] display the smallest std

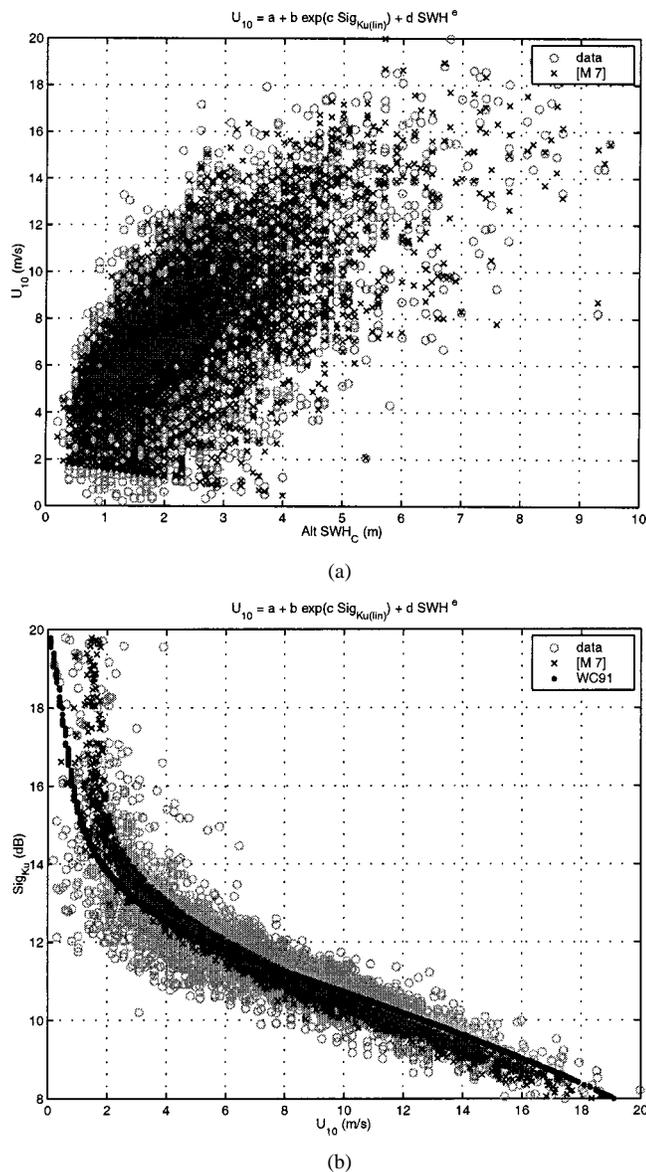


Fig. 3. (a) Wind speed against altimeter SWH<sub>C</sub> for collocated buoy/Topex dataset and altimeter wind speed retrieved with LSQ model [M7] featuring a sea state parameterization as a power of SWH<sub>C</sub>. (b) Wind speed against altimeter  $\sigma_{Ku}^0$  showing the variability in the  $U_{10}$  to  $\sigma^0$  relationship due to SWH at intermediate wind speeds.

values of about 1 m/s for wind speeds around  $U_{10} = 7$  m/s, thus highlighting the strong weight given to minimizing errors in this high data density region with the LSQ technique. The WC91 and FC94 models show consistent std over the wind speed range but are generally higher than all other models. The best overall performance in terms of consistency and low value of the std may arguably be attributed to the GG93 model, which also returns the smallest std at extreme wind speeds.

### B. Residual Dependence on Wind Speed

Another way to look at algorithm performance is through the residual dependence of the wind error on other variables. The dependence of the wind error on the buoy wind speed is shown in Fig. 5 as calculated for the same six algorithms as before. The contour lines represent the density of points and help highlight any residual dependence in the wind error.

Here, model WC91 and Gr00 produce the best results as witnessed by the isotropy of the contour lines. In contrast, model GG93 and LSQ models [M3] and [M7] produce clear warped results, indicating a residual trend with buoy wind speed. In this case, model GG93, [M3] and [M7] all appear to overestimate low and underestimate high wind speeds. Here, we shall just note that these latter models were all developed from small collocated buoy/altimeter dataset, while models WC91, Gr00, and FC94 which do not display the same trend, were derived from more extensive global datasets.

### C. Retrieved Wind Histograms

The agreement of the retrieved wind histogram with the original buoy wind speed histogram provides one more test of the validity of the algorithms. This is particularly important for wind climatology studies of regional/global and inter-annual/decadal changes in wind forcing. Fig. 6 compares the histograms of the wind speed retrieved for the six selected models with the histogram of the buoy wind measurements. The coefficient  $r^2$  represent the correlation between the altimeter retrieved and the buoy wind speed histograms shown in each subplot.

Here again, models WC91 and Gr00 produce the best fit between the retrieved and buoy wind histograms, with a correlation in excess of 0.99. LSQ models [M3] and [M7] display clear signs of the shortcomings of the LSQ approach in the present case, with retrieved wind histograms appearing narrower and peakier and enhanced density maxima around the dominant wind speed of  $U_{10} \sim 7$  m/s. The GG93 algorithm results in a similar albeit less marked problem and a clear shift of the density maximum toward lower wind speed [as suggested in Fig. 4(b)]. The FC94 wind speed displays a Rayleigh-type distribution which differs markedly from the original buoy wind speed histogram, thus highlighting that 1) our 4500 points dataset is NOT representative of the global wind field, and 2) by way of its derivation, the FC94 model is designed to reproduce the Rayleigh-type wind distribution of the global ocean and thus performs less satisfactorily in a typical conditions.

### D. Residual Dependence on Wave Age

The availability of *in-situ* wave period measurements from the buoys collocated with the altimeter enables us to examine any dependence of the altimeter wind speed on sea state maturity as suggested by [12]. Here, we will simply look at the residual wind calculated for the six chosen models against real wave age,  $\xi$ , defined as  $\xi = gT_p/(2\pi U_{10})$ , with  $T_p$  the peak wave period measured by the buoys. A more detailed investigation of the effect of wave development (e.g., through fetch) on altimeter winds can be found in [23].

The residual wind results are shown in Fig. 7 and have been fitted with a 4th degree polynomial (depicted by the starred line) to capture the changes in trend over the wave age range. From this, a clear residual dependence on real wave age is observed for all wind speed models, including the SWH-dependent two-parameter models. The relative trends for the different models are more easily inter-compared in Fig. 8, where the sea maturity effect causes a 1–1.5 m/s mean underestimation of wind speeds in young seas conditions ( $\xi < 1.5$ ) for all models.

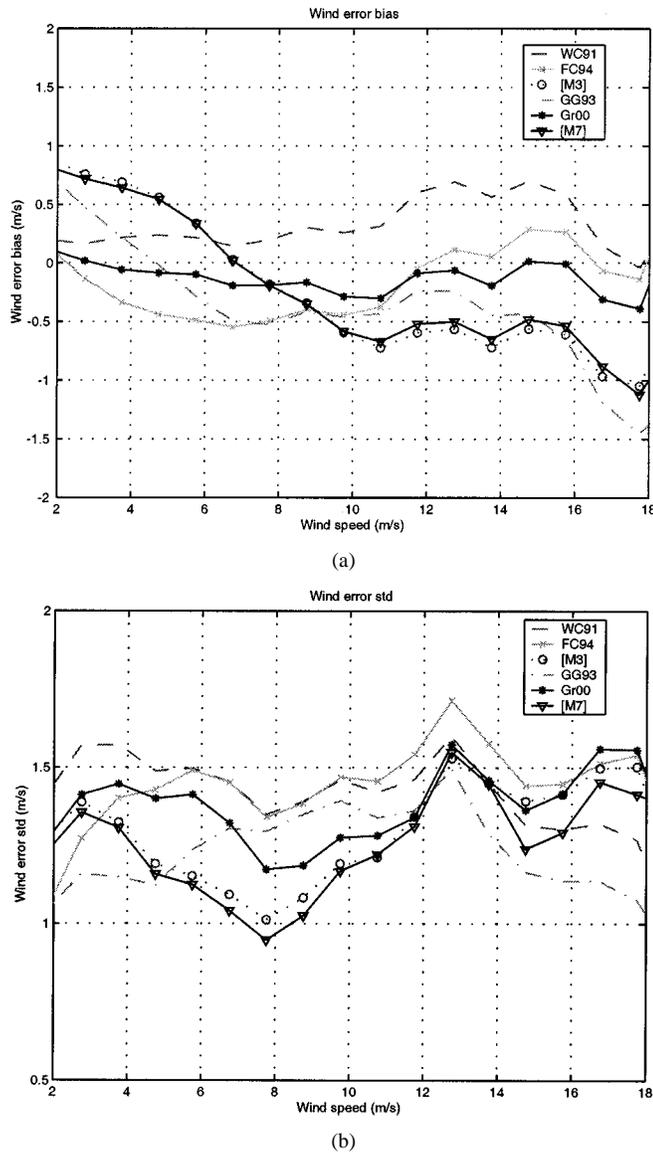


Fig. 4. (a) Wind error bias and (b) std calculated for 1.5 m/s wide wind speed bins for three  $\sigma^0$ -only models (WC91, FC94, and [M3]) and three ( $\sigma^0$ , SWH) models (GG93, Gr00, and model [M7]).

We note from Fig. 8 that the Gr00 model produces the smallest residual dependence on wave age, closely followed by the GG93 model. The LSQ models [M3] and [M7] display the largest residual dependence on wave age, although the reduction in magnitude of the wind error for model [M7] for both young and mature seas is consistent with the inclusion of a SWH parameterization. For young seas, the  $\sigma^0$ -only WC91 model unexpectedly returns the smallest wind error. This is attributed however to WC91 generally overestimating all wind speeds and thus providing a closer fit than other models to the higher wind speeds seen for young seas. This is confirmed by the results for the  $\sigma^0$ -only FC94 model which shows a trend of similar amplitude to WC91 but shifted toward more negative wind errors.

#### E. Discussion

A closer look at the performance of various algorithms revealed that minimizing the global wind error statistics is not

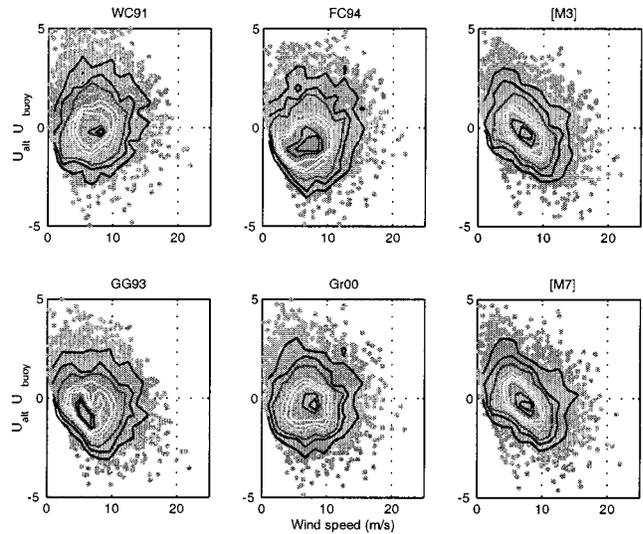


Fig. 5. Wind error,  $e$ , against buoy wind speed for three  $\sigma^0$ -only models (WC91, FC94, and [M3]) and three ( $\sigma^0$ , SWH) models (GG93, Gr00, and model [M7]).

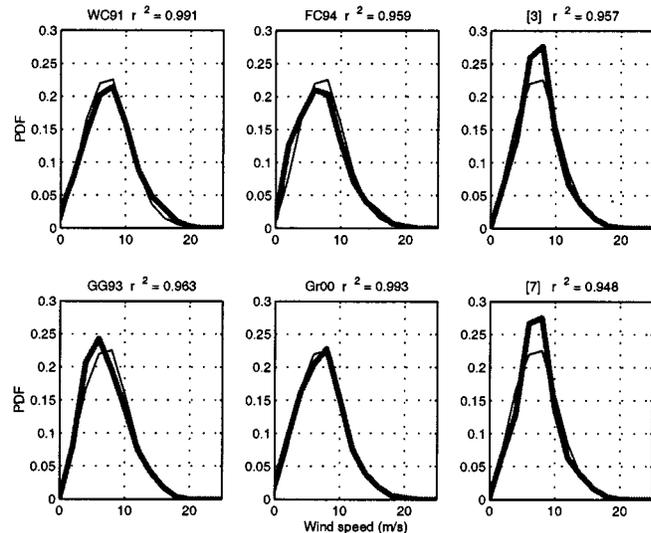


Fig. 6. Altimeter retrieved (thick line) and buoy wind speed (thin line) histograms over 1.5 m/s wide wind bins for three  $\sigma^0$ -only models (WC91, FC94, and [M3]) and three ( $\sigma^0$ , SWH) models (GG93, Gr00, and model [M7]).

sufficient to assess the overall validity of altimeter wind speed algorithms. Instead, the various characteristics of each model described in Sections II and III need to be accounted for. The pluses and minuses of each of the six models considered so far can be quantified and summarized as in Table IV to help make a balanced decision. Here, the performance of the algorithm is scored in every category by star ratings as poor (\*) to moderate (\*\*) and good (\*\*\*) .

The scoring in Table IV allows us to clearly identify the Gr00 model as the best model overall with average to good performance in all categories considered. Next best algorithm according to this classification is the WC91 model, whose major drawback lies in the systematic overestimation of the wind error (positive bias) and the higher global and local std of  $\sigma^0$ -only models. At the other end of the performance scale, model LSQ [M3], GG93, and LSQ [M7] are performing much

TABLE IV  
SUMMARY OF PERFORMANCE FOR THREE  $\sigma^0$ -ONLY MODELS (WC91, FC94, AND [M3]) AND THREE ( $\sigma^0$ , SWH) MODELS (GG93, GR00, AND MODEL [M7]). THE STAR RATING IDENTIFIES \* FOR POOR, \*\* FOR AVERAGE AND \*\*\* FOR GOOD

	WC91	FC94	LSQ [3]	GG93	Gr00	LSQ [7]
Global bias magnitude	*	*	***	*	**	***
Global std magnitude	*	*	***	*	**	***
Local bias consistency	**	**	*	*	***	*
Local std consistency	**	**	*	**	**	*
Residual dependence on buoy $U_{10}$	***	***	*	*	***	*
Quality of retrieved $U_{10}$ histogram	***	**	*	**	***	*
Residual dependence on wave age	*	*	*	**	**	*

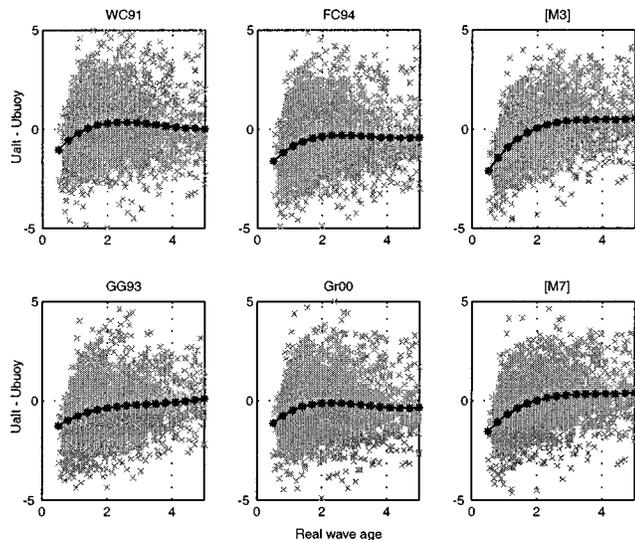


Fig. 7. Wind error  $e$  against real wave age,  $\xi$ , for three  $\sigma^0$ -only models (WC91, FC94, and [M3]) and three ( $\sigma^0$ , SWH) models (GG93, Gr00, and model [M7]). The starred line represents a 4th degree polynomial fit representing the change in trend for young ( $\xi < 1.5$ ) and mature seas ( $\xi > 1.5$ ).

less satisfactorily, particularly in terms of the poor description of the wind speed over the full wind range. As noted earlier, these problems seem to be specific to those models derived from fitting relatively small collocated buoy/altimeter datasets. This may point at possible effects related to the composition of collocated buoy/altimeter datasets or the existence of measurement errors. These issues are explored further in Section V. Overall, the addition of the SWH dependence in Gr00 allows this model to be preferred to the currently operational WC91 model, predominantly in view of its reduced wind error bias and its sensitivity to sea state development effects. We note however that the SWH parameterization is not sufficient to address fully the dependence of altimeter wind speed on sea state development.

## V. EFFECT OF DATASET COMPOSITION AND MEASUREMENT ERRORS ON ALGORITHM DEVELOPMENT

The LSQ fitted models developed from our collocated buoy/altimeter dataset have so far proven unable to produce satisfactory results over the full wind speed range, while producing exceptionally good results near the peak of the wind speed distribution (around  $U_{10} = 7$  m/s). The impact of the strongly peaked ( $U_{10}, \sigma^0$ ) distribution on the fitting

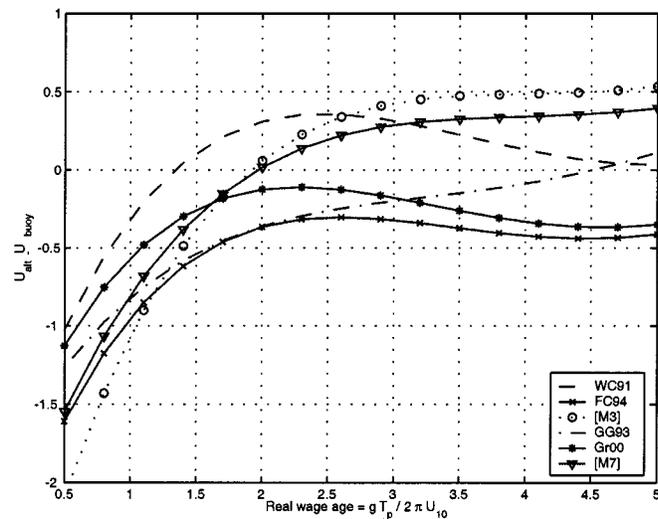


Fig. 8. Residual wind against real wave age trend from Fig. 7 for three  $\sigma^0$ -only models (WC91, FC94, and [M3]) and three ( $\sigma^0$ , SWH) models (GG93, Gr00, and model [M7]).

procedure is investigated here by comparing LSQ results obtained by fitting a measurement subset produced by equalizing the two-dimensional (2-D) ( $U_{10}, \sigma^0$ ) histogram of the full collocated dataset. For this, the data in each bin of the 2-D histogram are randomly sampled to retain a maximum of  $n$  samples (here,  $n = 3$ ), thereby flattening the histogram and artificially thinning the dataset.

The possible impact of measurement errors on the performance of the LSQ algorithms is explored with a different fitting procedure known as orthogonal distance regression (ODR) [24]. This technique allows nonlinear functional forms to be fitted to data while accounting for normally distributed errors in all variables. This technique is effectively a generalization of the LSQ fitting method, which also assumes normally distributed errors but only in the output variable (so far in this paper, buoy  $U_{10}$ ). The ODR technique thus allows uncertainties in the altimeter measurements to be accounted for as well. Note that this demonstration exercise is restricted to  $\sigma^0$ -only formulations (i.e., no SWH dependence included) as the primary aim is to understand what factors are responsible for the poor description of the  $U_{10}$  to  $\sigma^0$  relationship obtained by fitting our dataset.

The histogram-equalization and ODR exercises are carried out with a different functional form, proposed by Freilich and Challenor [5] to approximate the results of their statistical fit of a global dataset. Although the FC94 model did not stand out in Section IV as the optimal choice for a comparison, it is the only

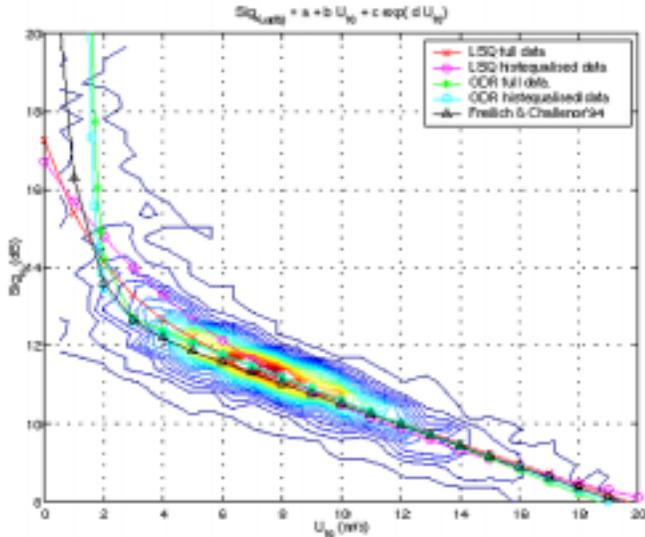


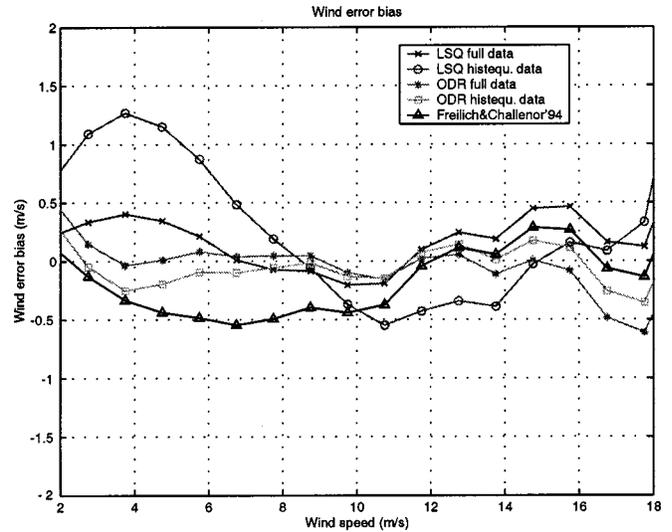
Fig. 9. Altimeter Ku-band backscatter to buoy  $U_{10}$  relationship for LSQ and ODR fits of functional form (2) to the full collocated buoy/altimeter dataset (4444 points; ODR3) and the histogram-equalized subset (479 points).

globally-derived model which results are approximated with a simple analytical formulation and which allow a direct evaluation of our fitting procedure. The analytical approximation of the FC94 model expressed as a function of wind speed at 19.5 m,  $U_{19.5}$ , reads

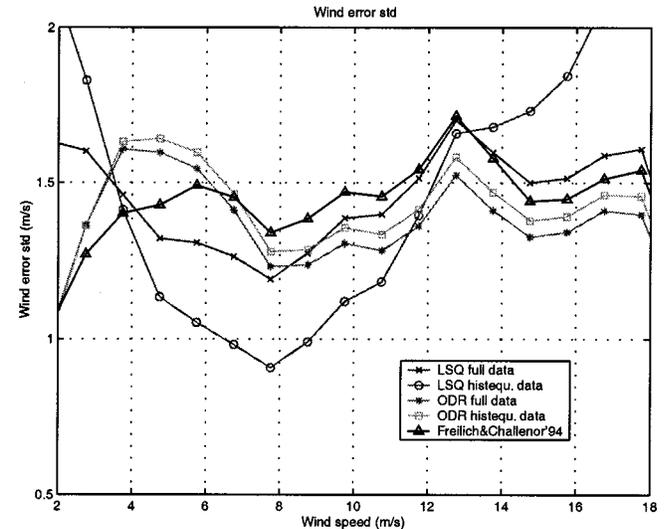
$$\sigma_{\text{dB}}^0 = a + bU_{19.5} + c \exp(dU_{19.5}). \quad (2)$$

The functional form in (2) is fitted to the full dataset and the histogram-equalized subset using the ordinary LSQ and the ODR methods initialized with the same parameters. The ODR procedure was run without a priori assumptions on the magnitude of the NRCS and wind speed errors. Estimates of the error were produced by the ODR in the course of the fit for each variable at each point. From this, it is possible to evaluate the std of the error distribution, which in the present case was of the order of 0.5 m/s and 0.5 dB for wind speed and NRCS, respectively. The four resulting models are shown in Fig. 9 for comparison with the FC94 model, and superimposed on the point density contour lines obtained for the full collocated buoy/altimeter dataset. Similarly, Fig. 10 shows the residual wind bias and std calculated over 1.5 m/s wind bins obtained for those same models.

Let us first consider the effect of dataset distribution on the quality of the models. We find that the LSQ results differ widely depending on which dataset is fitted. When fitting the full collocated dataset, the LSQ results closely trace the ridge of highest point density seen in the contour lines, but at the expense of the fit at extreme wind speeds. When applying LSQ to the histogram-equalized dataset, the resulting model diverges from the histogram peaks but remains unsatisfactory, as it becomes excessively sensitive to a typical data on the edges of the main body of the data. In contrast, the histogram-equalization has little effect on the results of the ODR fit.



(a)



(b)

Fig. 10. (a) Wind error bias and (b) std calculated for 1.5 m/s wide wind speed bins for  $\sigma^0$ -only models WC91, FC94, [3] and ODR3, and ( $\sigma^0$ , SWH) models GG93 and Gr00.

The ODR results, obtained with the same functional form as for LSQ above, present a much improved fit, which compares favorably with the global FC94 model for winds above 3 m/s. We note that the ODR results closely follow the high data density ridge also while offering a good approximation of the data at extreme wind speeds. The ODR models now adequately describe the sharp change of slope near  $\sigma^0 = 13$  dB [responsible for the increased bias in LSQ models at low winds; see Fig. 4(a)] and offer an even better approximation of the data at low winds ( $U_{10} < 3$  m/s) than the FC94 model. Fig. 10 supports these findings with the wind error bias [see Fig. 10(a)] and std [see Fig. 10(b)] for the ODR models now showing more consistent results over the range of wind speeds considered. Hence, accounting for normally distributed errors in both output and input variables with ODR permits the removal of undesirable properties of the LSQ fitted models, and returns results similar to those of models developed from large global dataset. Thus, altimeter quantization and buoy measurement errors play a major

role in defining the performance of the derived models, and this error-related problem can be resolved either by developing algorithms from as large as possible datasets (of the order of 100 000 data points as for WC91 and Gr00) representative of the global ocean, or by selecting statistical techniques which can account for the presence of errors in the fitted dataset.

## VI. CONCLUSION

A new dataset of collocated buoy/Topex measurements has been collated for the open ocean to investigate the relationship between wind speed and altimeter backscatter and SWH. The 4500-strong dataset was fitted using a LSQ method with a series of simple nonlinear functional forms inspired by previously published models. Comparison with four published models indicated that the addition of SWH information helps reduce the wind error std by about 10%. The LSQ models were found to produce the best results in terms in global wind bias and std when compared to the published models, but further analysis revealed shortcomings related to the poor description by LSQ models of the wind speed to backscatter relationship at extreme wind speeds. Best overall results were obtained for two-parameter algorithm by Gourrion *et al.* [19], although a residual dependence on sea state maturity remains for all models, resulting in wind speeds being underestimated by 1–1.5 m/s in young sea conditions.

An investigation into the impact on the LSQ fits of data distribution in small collocated datasets confirmed that the ordinary LSQ technique is not suitable for wind speed algorithm development as the principle of minimizing the distance between true and retrieved winds attributes excessive weight toward optimizing the fit near the peak of the wind distribution to the detriment of data at extreme wind speeds. A generalization of the LSQ approach known as ODR was implemented and established that the performance of models fitted to buoy/altimeter datasets can be vastly improved when accounting for normally distributed errors in both altimeter and buoy measurements. Hence, algorithms developed from relatively small collocated datasets (few thousand points) could match the performance of models developed from much larger datasets (few ten thousand points) if adequate statistical methods which account for measurement errors are used.

The application of ODR to two-parameter wind speed algorithm development is deferred to another paper as several issues need to be considered related to using ODR in three dimensions and to dealing with nonnormal wind error distributions. It is anticipated however that, in spite of more sophisticated statistical methods, the accuracy of wind speed algorithms is ultimately dictated by the quality of the fitted data. In the case of collocated buoy/altimeter datasets, it is conceivable that today's uncertainty in buoy measurements of wind speed at sea [25] already limit the possibility of further improvements with these methods.

## ACKNOWLEDGMENT

The authors are grateful to the U.S. National Data Buoy Center, the Canadian Marine Environmental Data Service, the

U.K. Meteorological Office, and the Japanese Marine Agency for providing the buoy data used in this work.

## REFERENCES

- [1] P. Gaspar and J. P. Florens, "Estimation of the sea state bias in radar altimeter measurements of sea level: Results from a new nonparametric method," *J. Geophys Res.*, vol. 103, pp. 15 803–15 814, 1998.
- [2] D. J. T. Carter, P. G. Challenor, and M. A. Srokosz, "An assessment of Geosat wave height and wind speed measurements," *J. Geophys Res.*, vol. 97, pp. 11 383–11 392, 1992.
- [3] J. F. R. Gower, "Intercalibration of wave and wind data from Topex/Poseidon and moored buoys off the West Coast of Canada," *J. Geophys Res.*, vol. 101, pp. 3817–3829, 1996.
- [4] G. S. Brown, H. R. Stanley, and N. A. Roy, "The wind speed measurement capability of spaceborne radar altimeters," *IEEE J. Oceanic Eng.*, vol. OE-6, pp. 59–63, 1981.
- [5] M. H. Freilich and P. G. Challenor, "A new approach for determining fully empirical altimeter wind speed model functions," *J. Geophys Res.*, vol. 99, pp. 25 051–25 062, 1994.
- [6] P. A. Hwang, W. J. Teague, G. A. Jacobs, and D. W. Wang, "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.*, vol. 103, pp. 10 451–10 468, 1998.
- [7] D. L. Witter and D. B. Chelton, "A Geosat altimeter wind speed algorithm and a method for altimeter wind speed algorithm development," *J. Geophys Res.*, vol. 96, pp. 8853–8860, 1991.
- [8] R. E. Glazman and S. H. Pilorz, "Effects of sea maturity on satellite altimeter measurements," *J. Geophys Res.*, vol. 95, pp. 2857–2870, 1990.
- [9] G. Chen, H. Lin, and J. Ma, "On the seasonal inconsistency of altimeter wind speed algorithms," *Int. J. Remote Sens.*, vol. 21, pp. 2119–2125, 2000.
- [10] F. Monaldo and E. Dobson, "On using significant wave height and radar cross section to improve radar altimeter measurements of wind speed," *J. Geophys Res.*, vol. 94, pp. 12 699–12 701, 1989.
- [11] J. M. Lefevre, J. Barckicke, and Y. Menard, "A significant wave height dependent function for TOPEX/POSEIDON wind speed retrieval," *J. Geophys Res.*, vol. 99, pp. 25 035–25 049, 1994.
- [12] R. E. Glazman and A. Greysukh, "Satellite altimeter measurements of surface wind," *J. Geophys Res.*, vol. 98, pp. 2475–2483, 1993.
- [13] F. J. Wentz, L. A. Mattox, and S. Peteherych, "New algorithms for microwave measurements of ocean winds with application to Seasat and SSM/I," *J. Geophys Res.*, vol. 91, pp. 2289–2307, 1986.
- [14] F. W. Dobson. (1991) Review of reference height and averaging time of surface wind measurements at sea. World Meteorological Organization. [Online]. Available: <http://www.wmo.ch>
- [15] S. D. Smith, "Coefficients for sea surface wind stress, heat flux, and wind profiles as a function of wind speed and temperature," *J. Geophys Res.*, vol. 93, pp. 15 467–15 472, 1988.
- [16] P. G. Challenor and P. D. Cotton. (1999) Trends in TOPEX significant wave height measurement. [Online]. Available: <http://www.soc.soton.ac.uk/JRD/SAT/TOPtren/main.html>
- [17] J. Wu, "On wave dependency of altimeter sea returns—Weak fetch influence on short ocean waves," *J. Atmos. Ocean. Technol.*, vol. 16, pp. 373–378, 1999.
- [18] D. B. Chelton and P. J. McCabe, "A review of satellite altimeter measurements of surface wind speed: With a proposed new algorithm," *J. Geophys Res.*, vol. 90, pp. 4707–4720, 1985.
- [19] J. Gourrion, D. Vandemark, S. Bailey, and B. Chapron. (2000) Satellite altimeter models for surface wind speed developed using ocean satellite crossovers. [Online]. Available: <http://topex.wff.nasa.gov/docs/docs.html>
- [20] P. S. Callahan, C. S. Morris, and S. V. Hsiao, "Comparison of Topex-Poseidon  $\sigma^0$  and significant wave height distributions to Geosat," *J. Geophys Res.*, vol. 99, pp. 25 015–25 024, 1994.
- [21] D. J. T. Carter, "Predictions of wave height and period for a constant wind velocity using the Jonswap results," *Ocean Eng.*, vol. 9, pp. 17–33, 1992.
- [22] C. Anderson, J. T. Macklin, C. P. Gommenginger, and M. A. Srokosz, "A comparison of the sea state sensitivity in empirical and theoretical models of altimeter ocean backscatter," *Can. J. Remote Sens.*, 2001, to be published.
- [23] C. P. Gommenginger, M. A. Srokosz, P. G. Challenor, V. Karaev, and P. D. Cotton, "Ocean wave effects on global altimeter wind climate retrieval," in *Proc. Waves 2001 Conf.*, San Francisco, CA, Sept. 3–5, 2001.

- [24] P. T. Boggs and J. E. Rogers, "Orthogonal distance regression," *Contemporary Mathematics*, vol. 112, pp. 183–194.
- [25] D. B. Gilhousen, "A field evaluation of NDBC moored buoy winds," *J. Atmos. Ocean Technol.*, vol. 4, pp. 94–104, 1986.



**Christine P. Gommenginger** received the M. S. and Technology of the Sea degree from Toulon University, France, in 1991, the D.E.A degree in electromagnetics, telecoms and remote-sensing from the University of Toulon/Nice-Sophia Antipolis, France, in 1992, and the Ph.D. degree in oceanography from Southampton University, U.K., in 1997.

Her main interests include interactions between microwave radiation and the sea surface, and the application of microwave remote sensing to oceanography and climate studies. Her work includes the use

of low-grazing angle navigational radar for dynamic ocean feature monitoring, the study of ocean roughness and SAR backscatter modulation by bathymetry in the coastal zone, and the effect of sea state on altimeter wind and sea level measurements. She is currently with the James Rennell Division for Ocean Circulation and Climate (JRD) at the Southampton Oceanography Centre (SOC), and is a member of the SOC Laboratory for Satellite Oceanography.



**Meric A. Srokosz** received the B.S. and Ph.D. degrees in mathematics from Bristol University, U.K., in 1976 and 1980, respectively.

He has over 20 years of experience in the fields of applied mathematics, wave studies, and remote sensing of the oceans. His research interests include rain from altimetry, plankton patchiness using ship and satellite sensors, data assimilation into biological models, altimeter sea state bias, ocean surface waves, salinity from space, and SST from passive microwave. He is PI or CoI on a number of

Earth Observation missions, including the ESA's ERS-1, ERS-2 and Envisat missions, the NASA/CNES Topex/Poseidon and Jason missions, the NASDA ADEOS and TRMM missions, and the ISRO OceanSat-1 mission. He has recently been appointed Scientific Co-ordinator of the Natural Environment Research Council Rapid Climate Change Thematic Programme. He is with the James Rennell Division for Ocean Circulation and Climate at the Southampton Oceanography Centre (SOC), U.K., and is a member of the SOC Laboratory for Satellite Oceanography.



**Peter G. Challenor** received the B.S. degree in mathematics from Exeter University, U.K., in 1976, and the M.S. degree in biometry from Reading University, U.K., in 1977. His research interests include satellite oceanography, in particular the use of altimetry for ocean circulation and the large scale processes, synergistic use of satellite data, surface waves statistics, statistics of extremes, data assimilation into models, and error estimation in models. He is currently team leader within the James Rennell Division for Ocean Circulation and Climate at the Southampton Oceanography Centre (SOC), U.K., and is a member of the SOC Laboratory for Satellite Oceanography.

Mr. Challenor is a Fellow of the Royal Statistical Society, a member of the Editorial Board of Applied Ocean Research, a Member of the American Geophysical Union, and a Member of ESA Radar Altimeter Science Advisory Group.

**P. David Cotton** received the B.S. degree in physics from the University of Sheffield, U.K., in 1982, and a the Ph.D. degree in ionospheric physics from the University of Sheffield in 1989.

His research interests include satellite measurements of sea state, in particular the verification and calibration of these measurements, the development of improved algorithms, and the investigation of interannual variability in wave climate. He has been accepted as Principal Investigator for the JASON and ENVISAT altimeter science programs. He is currently project manager for Satellite Observing Systems Ltd., U.K.

Dr. Cotton is a member of the U.K. Royal Meteorological Society, a member of the American Geophysical Union.