JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 99, NO. C12, PAGES 24,995–25,008, DECEMBER 15, 1994

The sea state bias in altimeter estimates of sea level from collinear analysis of TOPEX data

Dudley B. Chelton

College of Oceanic and Atmospheric Sciences, Oregon State University, Corvallis

Abstract. The wind speed and significant wave height ($H_{1/3}$) dependencies of the sea state bias in altimeter estimates of sea level, expressed in the form $\Delta h_{\mathrm{SSB}} = bH_{1/3}$, are examined from least squares analysis of 21 cycles of collinear TOPEX data. The bias coefficient b is found to increase in magnitude with increasing wind speed up to about 12 m s^{-1} and decrease monotonically in magnitude with increasing $H_{1/3}$. A parameterization of b as a quadratic function of wind speed only, as in the formulation used to produce the TOPEX geophysical data records (GDRs), is significantly better than a parameterization purely in terms of $H_{1/3}$. However, a four-parameter combined wind speed and wave height formulation for b (quadratic in wind speed plus linear in $H_{1/3}$) significantly improves the accuracy of the sea state bias correction. The GDR formulation in terms of wind speed only should therefore be expanded to account for a wave height dependence of b. An attempt to quantify the accuracy of the sea state bias correction Δh_{SSB} concludes that the uncertainty is a disconcertingly large 1% of $H_{1/3}$.

1. Introduction

The sea state bias (SSB) remains the altimeter range correction that is least understood theoretically. Qualitatively, it is known that altimetric estimates of sea level are biased low because of a greater backscattered power per unit surface area from wave troughs than from wave crests. Although there have been significant advancements in the theoretical understanding of this bias [e.g., *Jackson,* 1979; *Barrick and Lipa,* 1985; *Srokosz,* 1986, 1987; *Lagerloef,* 1987; *Rodríguez,* 1988; *Fu and Glazman,* 1991], empirical models derived from analyses of satellite, aircraft and tower-based altimeter measurements continue to offer the most accurate estimates of the sea state bias presently available. The uncertainty of the sea state bias is large (comparable to the orbit error for TOPEX/POSEIDON). Assessment of the accuracy of the sea state bias correction is therefore important to all scientific applications of the data. The objective of this study is to obtain an empirical estimate of the sea state bias from the NASA altimeter (hereinafter referred to as TOPEX) on board the TOPEX/POSEIDON satellite.

Because all empirical studies have found that the sea state bias increases with increasing wave height, the bias is traditionally formulated as a linear function of significant wave height $H_{1/3}$, which corresponds approximately to 4 times the standard deviation of the sea surface elevation within the altimeter footprint. *Barrick and Lipa* [1985] and *Srokosz* [1986] provide a theoretical basis for this formalism. Altimetric measurements of sea level can thus be written as

$$h_{\mathrm{meas}} = h_{\mathrm{true}} + h_{\mathrm{error}} + \Delta h_{\mathrm{SSB}}, \tag{1a}$$

where

$$\Delta h_{\mathrm{SSB}} = bH_{1/3} \tag{1b}$$

is the sea state bias, h_{true} is the true sea level, and h_{error} includes all sources of measurement error other than the sea state bias. As summarized by *Chelton* [1988], Δh_{SSB} expressed in the form (1b) includes the electromagnetic (EM) bias, the tracker bias, and any contribution of the skewness bias that is proportional to $H_{1/3}$. The sea state bias likely depends on characteristics of the wave field other than $H_{1/3}$, but this is the only wave characteristic easily and unambiguously extracted from altimeter data.

For the NASA production of TOPEX interim geophysical data records (IGDRs), *Hevizi et al.* [1993] developed an expression for the sea state bias coefficient b based on a synthesis of all available sea state bias studies from aircraft, tower, and Geosat altimeter data. The form adopted is a quadratic function of wind speed u,

$$b(u) = a_0 + a_1 u + a_2 u^2 \,. \tag{2}$$

On the basis of additional aircraft data and preliminary analyses of TOPEX IGDR data, the *Hevizi et al.* [1993] coefficients a_0, a_1, and a_2 were adjusted a small amount to obtain the values listed in Table 1 that were used in the production of the NASA geophysical data records (GDRs).

There is some lingering uncertainty about the relevance of the sea state bias in the ~10-m footprint radar

Paper number 94JC02113.
0148-0227/94/94JC-02113$05.00

Table 1. The Parameters of the Quadratic Wind Speed Dependent Model $b = a_0 + a_1 u + a_2 u^2$ for the Sea State Bias Coefficient

	GDR	Collinear Regression	Crossover Regression
a_0	-0.0029	-0.0047 ± 0.0086	0.0036 ± 0.007
a_1	-0.0038	-0.0038 ± 0.0007	-0.0045 ± 0.0008
a_2	1.55×10^{-4}	$(1.6 \pm 0.3) \times 10^{-4}$	$(1.9 \pm 0.3) \times 10^{-4}$
VAR	7.67	7.67	7.35

The three values for each parameter correspond to (1) the values used in the GDR processing, (2) the values derived here by regression analysis of collinear TOPEX Data, and (3) the values derived by *Gaspar et al.* [this issue] by regression analysis of crossover difference TOPEX data. The bottom row of the table gives the reduced variance (VAR) of sea level in cm^2 after removal of the sea state bias.

measurements from aircraft and towers to that in the much larger ~10-km footprint satellite altimeter measurements. Moreover, the parameter space spanned by the geographically and temporally limited aircraft and tower data does not include the full range of conditions sampled globally by a satellite altimeter. These concerns, plus the somewhat wizardly heritage of the GDR sea state bias algorithm, beg an empirical determination of the sea state bias from in-orbit TOPEX GDR data.

The objective of this study is thus to examine the validity of the GDR sea state bias algorithm by investigating the sea state bias as a function of both wind speed and wave height from collinear analysis of 21 repeat cycles of TOPEX data. This study compliments a similar empirical determination of sea state bias by *Gaspar et al.* [this issue] who used a completely independent approach based on crossover differences. Indeed, the present study was largely motivated by the surprising wind speed and wave height dependencies of the sea state bias coefficient obtained by Gaspar et al., as discussed in section 5. The collinear analysis presented here provides an independent corroboration of the Gaspar et al. crossover results.

2. Data Processing

The data used in this study were obtained from cycles 9–30 of the TOPEX GDRs. Cycles 1–8 were not considered because of concerns that the satellite attitude control problems experienced during this period might affect the sea state bias estimation. Data from the French POSEIDON altimeter, including all of repeat cycle 20, were also not considered because of insufficient data quantity to estimate the sea state bias separately for POSEIDON by the collinear method used here.

The standard environmental corrections (wet and dry tropospheric corrections, ionospheric correction, inverse barometer correction, solid Earth tide correction, and the extended Schwiderski ocean tide corrections, including the loading tide) were applied to each TOPEX sea level estimate; only the sea state bias correction was omitted. All flagged TOPEX data were eliminated.

The wind speeds used here to estimate the sea state bias coefficient were calculated from the normalized radar cross section (σ_0) using the modified Chelton and Wentz (MCW) wind speed model function developed for the Geosat altimeter by *Witter and Chelton* [1991a]. A 0.7-dB bias was subtracted from the TOPEX σ_0 values in the GDRs in order to obtain a probability distribution function of σ_0 consistent with that of Geosat.

The problem of uncertainties in the reference level (the marine geoid) in altimeter sea level estimates can be dealt with in collinear analysis by two methods. One approach is to eliminate the geoid by analyzing collinear differences between individual tracks. Another approach is to compute and remove the mean sea level at each observation location, thus eliminating the geoid. Although there are merits to both approaches, the latter method was used here. The sea state bias coefficient was then estimated from the residual data as described in detail in the following sections. The method thus requires interpolating the TOPEX data to a fixed grid along the repeating satellite ground track in order to facilitate calculation of the spatially dependent mean sea level. This was achieved by linear interpolation of the h_{meas}, $H_{1/3}$ and σ_0 data to approximately 7-km intervals, as previously described for Geosat by *Zlotnicki et al.* [1990].

With all previous altimeter data, errors in orbit height have been so large (of order 50 cm or more) that it has been necessary also to remove orbit errors from the sea level estimates. An unsettling result obtained independently by *Zlotnicki et al.* [1989] and B. Douglas (personal communication, 1990) is that estimates of the sea state bias coefficient b can be sensitive to the detailed form of the orbit error correction. Fortunately, this issue can be avoided with TOPEX/POSEIDON data because the orbit errors are only 3–4 cm [*Tapley et al.,* this issue], and can therefore be neglected in the estimation of sea state bias.

3. Wind Speed Dependence

3.1. Nonparametric Approach

Ideally, a wind speed dependence of the sea state bias coefficient b should be determined nonparametrically, rather than by presupposing a functional dependence (e.g., quadratic, as in the GDR algorithm) on wind speed and evaluating the coefficients of the function by regression analysis. This can be achieved by stratifying coincident estimates of h_{meas} and $H_{1/3}$ according to their corresponding wind speed u computed from σ_0 as described in section 2.

As noted in section 2, the mean value must be removed from h_{meas} separately at each grid location $k = 1, \ldots, K$ because of uncertainties in the marine geoid. In order to estimate the sea state bias empirically from altimeter data, the mean value must also be removed from the right side of (1a). Consider a wind speed bin of width Δu centered on u_i. The mean value of the measured sea level, for example, within this wind speed bin at location k can be expressed as

$$\langle h_{\mathrm{meas}}(u_i)\rangle_k = \left[N_k(u_i)\right]^{-1} \sum_{j=1}^{N_k(u_i)} h_{\mathrm{meas}}(t_{kj}, u_i). \quad (3)$$

The angle brackets with subscript k are shorthand notation for the mean value at location k and the wind speed dependence inside the angle brackets indicates that only the $N_k(u_i)$ observations at location k that fall within the wind speed bin are included in the mean. The mean value of (1) then becomes

$$\langle h_{\mathrm{meas}}(u_i)\rangle_k = \langle h_{\mathrm{true}}(u_i)\rangle_k + \langle h_{\mathrm{error}}(u_i)\rangle_k + \langle b(u_i)H_{1/3}(u_i)\rangle_k. \quad (4)$$

Define the residuals (denoted by primes) of each quantity in (1) at time t_{kj} to be the deviations from the local mean values in (4), for example,

$$h'_{\mathrm{meas}}(t_{kj}, u_i) = h_{\mathrm{meas}}(t_{kj}, u_i) - \langle h_{\mathrm{meas}}(u_i)\rangle_k. \quad (5)$$

The explicit dependence on wind speed indicates that the residuals are computed only from observations within the wind speed bin. Then (4) can be subtracted from (1) to obtain

$$h'_{\mathrm{meas}}(t_{kj}, u_i) = h'_{\mathrm{true}}(t_{kj}, u_i) + h'_{\mathrm{error}}(t_{kj}, u_i) + b(u_i)H_{1/3}(t_{kj}, u_i) - \langle b(u_i)H_{1/3}(u_i)\rangle_k. \quad (6)$$

In general, b cannot be passed through the angle brackets on the right side of (6). However, if b is a smoothly varying function of wind speed and the bin width Δu is sufficiently small, then the sea state bias can be considered constant over the wind speed bin. In this case $b(u_i)$ can be passed through the angle brackets and (6) becomes

$$h'_{\mathrm{meas}}(t_{kj}, u_i) = h'_{\mathrm{true}}(t_{kj}, u_i) + h'_{\mathrm{error}}(t_{kj}, u_i) + b(u_i)H'_{1/3}(t_{kj}, u_i). \quad (7)$$

For application here, wind speed bins of width $\Delta u = 2$ m s^{-1} were used.

For estimation of the sea state bias coefficient, the residual wave heights for each wind speed bin at the different locations k were then pooled over all K grid locations. The resulting total number of observation times t in wind speed bin u_i is thus $N(u_i) = \sum_{k=1}^{K} N_k(u_i)$. An estimate $\hat{b}$ of the sea state bias coefficient at wind speed u_i was then computed from the residual data for each wind speed bin by regressing the pooled residual sea level measurements on the pooled residual significant wave heights, as suggested by (7), that is,

$$h'_{\mathrm{meas}}(t, u_i) = \hat{b}(u_i)H'_{1/3}(t, u_i) + \epsilon(t, u_i). \quad (8)$$

The residual true sea level and residual mean errors in (7), as well as the misfit of the regression model, are included in the errors ϵ. Note that it is necessary to assume that $H'_{1/3}$ and ϵ are uncorrelated in order to evaluate $\hat{b}$ by minimizing the sum of squared errors in (8).

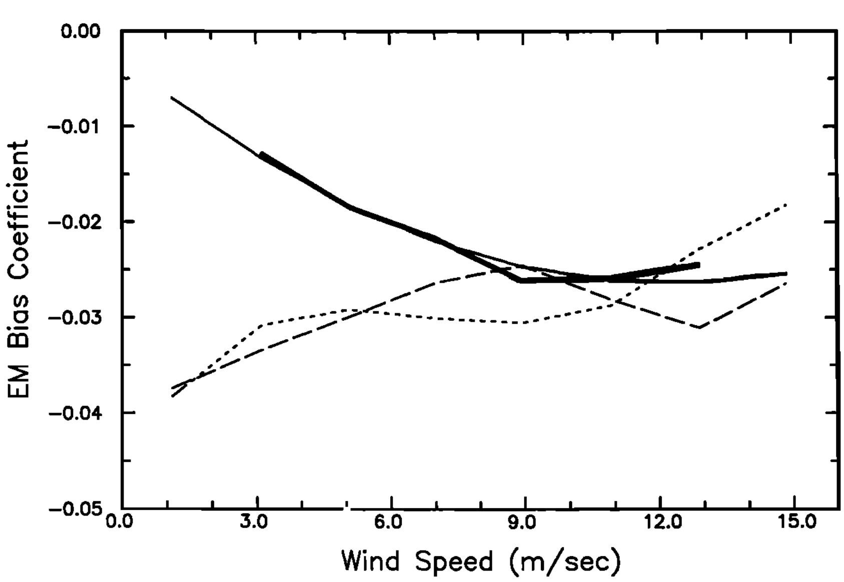

Figure 1. The wind speed dependence of the TOPEX GDR algorithm (thin solid line) and the nonparametric estimate obtained here (heavy solid line). The incorrect estimates based on the method of *Witter and Chelton* [1991b] are shown for TOPEX and Geosat by the dashed and dotted lines, respectively.

This nonparametric method of estimating the wind speed dependence of b has previously been applied to Geosat data by *Witter and Chelton* [1991b]. Unfortunately, that analysis is flawed in that the mean values in (4) were computed at each location k over all wind speeds, rather than just over the $N_k(u_i)$ wind speeds that fell within the bin of width Δu centered on u_i. The resulting *Witter and Chelton* [1991b] estimates of $b(u_i)$ for Geosat are shown by the dotted line in Figure 1. The TOPEX estimates of $b(u_i)$ computed by the same (incorrect) method are shown by the dashed line in Figure 1. Although neither of these estimates of the wind speed dependence of the sea state bias coefficient can be considered valid because of the error in the computation of the residuals, there is at least some comfort in noting a general consistency between the results obtained for Geosat and TOPEX.

For correct application of the nonparametric method, mean values $\langle h_{\text{meas}}(u_i)\rangle_k$ and $\langle H_{1/3}(u_i)\rangle_k$ at each location k must be computed for each wind speed bin i from only the $N_k(u_i)$ observations that fall within the bin, as described above. This clearly imposes serious sampling problems with a limited data set. For the 21 cycles of data analyzed here, there are a maximum of 21 total observations at each grid location k (fewer, in the event of data dropouts). This limited number of observations is apportioned among the different wind speed bins. Typically, several wind speed bins at a particular location are empty. The sampling problems were addressed here by discarding all observations that fell within wind speed bin i at location k for which there were fewer than $N_k(u_i) = 3$ observations over the 21 cycles of data. For bins with $N_k(u_i) \geq 3$ observations, residual sea level estimates and significant wave heights were computed and pooled for the regression model (8) for wind speed bin i.

The resulting numbers of pooled residual wave height values $N(u_i)$ in the lowest wind speed bin (0–2 m s^{-1}) and in wind speed bins higher than 14 m s^{-1} were too small to obtain reliable estimates of b. The regression estimates $\hat{b}(u_i)$ for the other wind speed bins are shown by the heavy solid line in Figure 1. For comparison, the quadratic GDR formulation is shown by the thin solid line. It is evident that the nonparametric estimates of the wind speed dependence of b for wind speeds in the range 2–14 m s^{-1} agree very well with the GDR algorithm. In contrast, the *Witter and Chelton* [1991b] estimates of $b(u)$ are clearly in error, even showing an incorrect tendency for decreasing sea state bias with increasing wind speed.

3.2. Parametric Approach

If an analytical form is assumed for the wind speed dependence of the sea state bias coefficient, the coefficients of the analytical expression can be estimated parametrically by regression analysis without the need to stratify the observations according to wind speed, as in section 3.1. Such a parametric estimate has the advantage that it is not limited to the restricted (2–14 m s^{-1}) range of wind speeds, as is the nonparametric estimate in Figure 1.

The sea state bias (1) expressed in terms of residuals is

$$h'_{\text{meas}} = h'_{\text{true}} + h'_{\text{error}} + bH_{1/3} - \langle bH_{1/3}\rangle_k . \quad (9)$$

The residuals in (9) differ from those in (6) in that, rather than being binned by wind speed, the mean values are computed over all N_k observations at location k, for example,

$$h'_{\text{meas}}(t_{kj}) = h_{\text{meas}}(t_{kj}) - \langle h_{\text{meas}}\rangle_k , \quad (10)$$

where the mean value is defined as

$$\langle h_{\text{meas}}\rangle_k = N_k^{-1} \sum_{j=1}^{N_k} h_{\text{meas}}(t_{kj}) . \quad (11)$$

The apparent success of the quadratic function of wind speed deduced from the nonparametric method in section 3.1 suggests adopting the GDR analytical form (2) for $b(u)$. Substitution into (9) gives

$$h'_{\text{meas}}(t_{kj}) = h'_{\text{true}}(t_{kj}) + h'_{\text{error}}(t_{kj}) + a_0 H'_{1/3}(t_{kj}) + a_1 \left[u(t_{kj})H_{1/3}(t_{kj})\right]' + a_2 \left[u^2(t_{kj})H_{1/3}(t_{kj})\right]' , \quad (12)$$

where the primed square brackets are the residuals of the products inside of the brackets, for example,

$$\left[u(t_{kj})H_{1/3}(t_{kj})\right]' = u(t_{kj})H_{1/3}(t_{kj}) - \langle uH_{1/3}\rangle_k . \quad (13)$$

The residuals h'_{meas}, $H'_{1/3}$, $\left[uH_{1/3}\right]'$ and $\left[u^2H_{1/3}\right]'$ were computed at each location k and then pooled over all K grid locations. The parameters a_0, a_1 and a_2 were then estimated from the pooled residuals by regressing h'_{meas} on the other three residuals, as suggested by (12), that is,

$$h'_{\text{meas}}(t) = \hat{a}_0 H'_{1/3}(t) + \hat{a}_1 \left[u(t)H_{1/3}(t)\right]' + \hat{a}_2 \left[u^2(t)H_{1/3}(t)\right]' + \epsilon(t) . \quad (14)$$

Here the estimates of the true parameters a_0, a_1, and a_2 have been denoted with circumflexes. As in (8), ϵ includes the residual true sea level and residual measurement errors in (12), as well as the misfit of the regression model.

The least squares estimates of the parameters obtained by regression analysis of the full 21 cycles of data considered here are listed in Table 1. The corresponding quadratic dependence of the sea state bias coefficient is shown graphically by the heavy solid line in Figure 2a. It is difficult to quantify uncertainties in these least squares estimates. The classical formalism for determining confidence intervals for the regression parameters assumes that the residuals $\epsilon(t)$ are statistically independent. This is clearly not the case in the 7-km gridded data analyzed here since $\epsilon(t)$ includes large-scale true sea level and residual measurement errors.

In principal, the effective number of degrees of freedom in the regression model can be estimated from

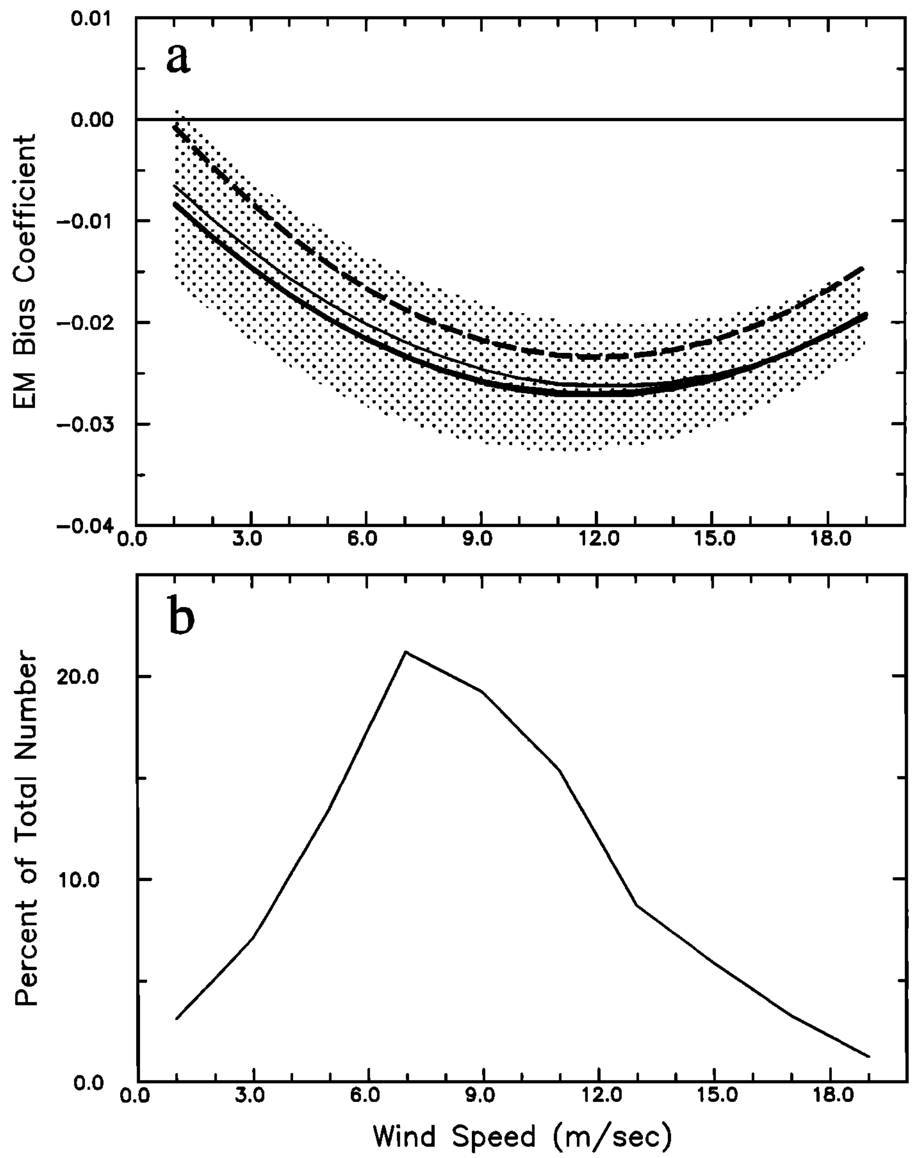

Figure 2. (a) The three-parameter model equation (2) for the sea state bias coefficient used in the TOPEX GDR algorithm (thin solid line), obtained here from global regression analysis of residual collinear TOPEX data (heavy solid line), and obtained by *Gaspar et al.* [this issue] from regression analysis of crossover difference TOPEX data (heavy dashed line). The gray shaded region represents the ± 1 standard deviation uncertainty of the collinear regression estimates obtained here. (b) The percentages of global TOPEX observations in wind speed bins of width 2 m s^{-1}.

the decorrelation length scale of the residuals $\epsilon(t)$ as described, for example, by *Chelton* [1983]. Such an approach could be applied to account for the "redundancy" of the data along the satellite ground track. Adjusting the effective degrees of freedom to account for correlations between $\epsilon(t)$ from neighboring ground tracks that arise from the large zonal scales of the true sea level and residual measurement errors in $\epsilon(t)$ is more problematic, as this requires knowledge of the spatial and temporal decorrelation scales (equivalent to specifying the full three-dimensional wavenumber-frequency spectrum of the residuals).

A more practical and conservative approach was used here to estimate the uncertainties of the regression parameters. Each of the 21 cycles of data analyzed in this study was considered to be an independent realization and the regression (14) was performed separately for each cycle. The error bars listed in Table 1 represent the standard deviation of the 21 individual estimates of the parameters. These error bars thus represent the stability of the regression model from one repeat cycle to the next.

The overall uncertainty of the sea state bias coefficient as a function of wind speed was obtained by computing the values of b from each of the 21 regression models at wind speed intervals of 1 m s^{-1}. The standard deviation of the 21 estimates in each wind speed bin, shown as a function of wind speed by the shaded area in Figure 2a, is approximately 0.008 near the peak of the wind speed distribution shown in Figure 2b. This corresponds to a sea state bias uncertainty of 0.8% of $H_{1/3}$, which represents an uncertainty of 1.6 cm in the sea state bias correction Δh_{SSB} for a typical wave height of 2 m (see Figure 3b below).

The agreement between the estimates of the parameters obtained here from regression analysis of global residual collinear TOPEX data and the values used in the GDR algorithm (listed in Table 1 and shown by the thin solid line in Figure 2a) is remarkable, if not somewhat coincidental given the large uncertainty of the least squares estimate of a_0; the GDR values of all three parameters are very nearly the same as the collinear least squares estimates, easily falling within the ranges of uncertainty of the least squares estimates obtained here. The reduction of sea level variance (bottom row of Table 1) is the same for both models.

For comparison, the quadratic wind speed dependence of the sea state bias coefficient estimated by *Gaspar et al.* [this issue] from regression analysis of crossover difference TOPEX data is also listed in Table 1 and shown by the heavy dashed line in Figure 2a. The crossover estimate is everywhere smaller in magnitude than the GDR and collinear estimates but falls within the ±1 standard deviation range of uncertainty

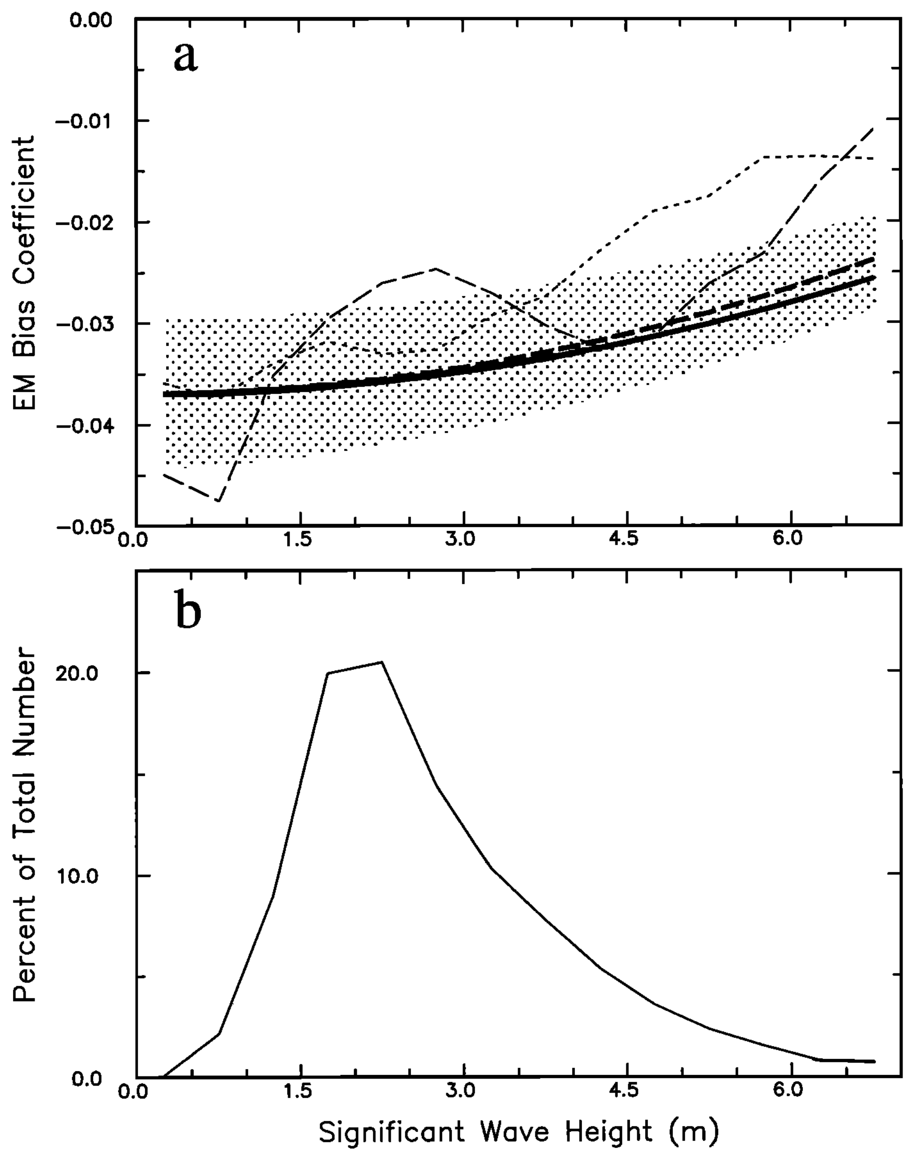

Figure 3. (a) The two-parameter model equation (15) for the sea state bias coefficient obtained here from regression analysis of residual collinear TOPEX data (heavy solid line) and obtained by *Gaspar et al.* [this issue] from regression analysis of crossover difference TOPEX data (heavy dashed line). The incorrect estimates based on the method of *Witter and Chelton* [1991b] are shown for TOPEX and Geosat by the thin dashed and dotted lines, respectively. The gray shaded region represents the ±1 standard deviation uncertainty of the collinear regression estimates obtained here. (b) The percentages of global TOPEX observations in wave height bins of width 0.5 m.

of the collinear least squares estimates obtained here. The difference is about 0.005 near the peak of the wind speed distribution. It is noteworthy that the uncertainties of the regression coefficients obtained by Gaspar et al. are nearly identical to those obtained here. The two totally independent approaches for estimating the sea state bias thus produce very similar results.

The large uncertainty of the collinear regression estimate $\hat{a}_0$ merits additional discussion. This constant offset term in the regression is apparently much less robust than the wind speed coefficients $\hat{a}_1$ and $\hat{a}_2$. This was investigated in greater detail by separating the TOPEX data set into two subsets and performing the three-parameter regression on each half of the data. Bias coefficients estimated separately from observations in the northern and southern hemispheres had virtually the same values of $\hat{a}_1$ and $\hat{a}_2$, but the two estimates of $\hat{a}_0$ differed by about 0.01. Similar results were obtained when bias coefficients were estimated separately from the even and odd cycles of data. For a typical significant wave height of 2 m, a change of 0.01 in $\hat{a}_0$ corresponds to a 2-cm change in the sea state bias correction Δh_{SSB}.

4. Wave Height Dependence

Because the quadratic wind speed formulation for the sea state bias coefficient b is purely empirical, it is prudent also to investigate the possibility of a wave height dependence. *Witter and Chelton* [1991b] have previously derived nonparametric estimates of $b(H_{1/3})$ for Geosat. As shown by the dotted line in Figure 3a, they obtained an approximate constant −0.035 for wave heights smaller than about 3 m; for larger wave heights, b decreased in magnitude with increasing $H_{1/3}$. The TOPEX estimates of $b(H_{1/3})$ computed by the same method are shown by the thin dashed line in Figure 3a. The TOPEX and Geosat results differ in detail, but the two estimates are qualitatively consistent: both are similar in magnitude and both generally decrease in magnitude with increasing wave height.

As discussed in section 3.1, the residuals were calculated incorrectly by *Witter and Chelton* [1991b]. The dotted and thin dashed lines in Figure 3a therefore cannot be considered valid. Unfortunately, the nonparametric method outlined in section 3.1 cannot be used to determine the $H_{1/3}$ dependence of the sea state bias; stratifying coincident estimates of h_{meas} and $H_{1/3}$ into $H_{1/3}$ bins yields residual $H_{1/3}$ values with dynamic range limited to the width of the $H_{1/3}$ bins. This results in very unstable least squares estimates of b within each $H_{1/3}$ bin.

A wave height dependence of the sea state bias coefficient b in (9) must therefore be investigated parametrically for a specified analytical form for $b(H_{1/3})$. *Gaspar et al.* [this issue] have proposed a two-parameter model of the form

$$b(H_{1/3}) = a_0 + a_1 H_{1/3}^2 . \qquad (15)$$

The estimates of the parameters of this model obtained here from regression analysis of the residual collinear TOPEX data using the method described in section 3.2 are listed in Table 2. The resulting $H_{1/3}$ dependence of the sea state bias, shown by the heavy solid line in Figure 3a, decreases in magnitude with increasing $H_{1/3}$. The ±1 standard deviation uncertainty of the two-parameter bias coefficient (shown by the shaded area in Figure 3a) is about 0.007 near the peak of the wave height distribution shown in Figure 3b.

Table 2. The Parameters of the Wave Height Dependent Model $b = a_0 + a_1 H_{1/3}^2$ for the Sea State Bias Coefficient

	Collinear Regression	Crossover Regression
a_0	-0.037 ± 0.007	-0.037 ± 0.004
a_1	$(2.5 \pm 0.8) \times 10^{-4}$	$(2.9 \pm 0.7) \times 10^{-4}$
VAR	7.30	7.16

The two values for each parameter correspond to (1) the values derived here by regression analysis of collinear TOPEX data and (2) the values derived by *Gaspar et al.* [this issue] by regression analysis of crossover difference TOPEX data. The bottom row of the table gives the reduced variance (VAR) of aea level in cm^2 after removal of the sea state bias.

For comparison the estimates of the parameters obtained by *Gaspar et al.* [this issue] from regression analysis of crossover difference TOPEX data are listed in Table 2 and the resulting sea state bias coefficient is shown by the heavy dashed line in Figure 3a. The collinear and crossover regression estimates of the parameters of the model (15) are virtually identical. The collinear estimates account for a slightly greater (but probably not statistically significant) reduction of collinear sea level variance (bottom row of Table 2).

It is evident from Figure 3 that a linear dependence of b on $H_{1/3}$ would have worked nearly as well as the second-order dependence considered here. A similar conclusion was reached from the crossover analysis by *Gaspar et al.* [this issue]. The second-order dependence was used here to enable a comparison of the collinear results with the crossover results presented by Gaspar et al. This two-parameter model will be replaced with an improved regression model in the next section.

An interesting point that can be noted from Figure 3a is that, although somewhat smaller in magnitude, the incorrect nonparametric estimates of $b(H_{1/3})$ obtained by *Witter and Chelton* [1991b] are fortuitously similar to the two parametric estimates. They speculated that the apparent decrease in sea state bias with increasing wave height was attributable to the large attitude errors in the Geosat data. As attitude errors for TOPEX are very small during the cycles 9–30 analyzed here, the wave height dependence deduced from both Geosat and TOPEX data evidently cannot be attributed to attitude errors. It must represent a real wave height dependence of the sea state bias.

The usual method of evaluating the relative accuracy of two regression models is to compare the reduction of

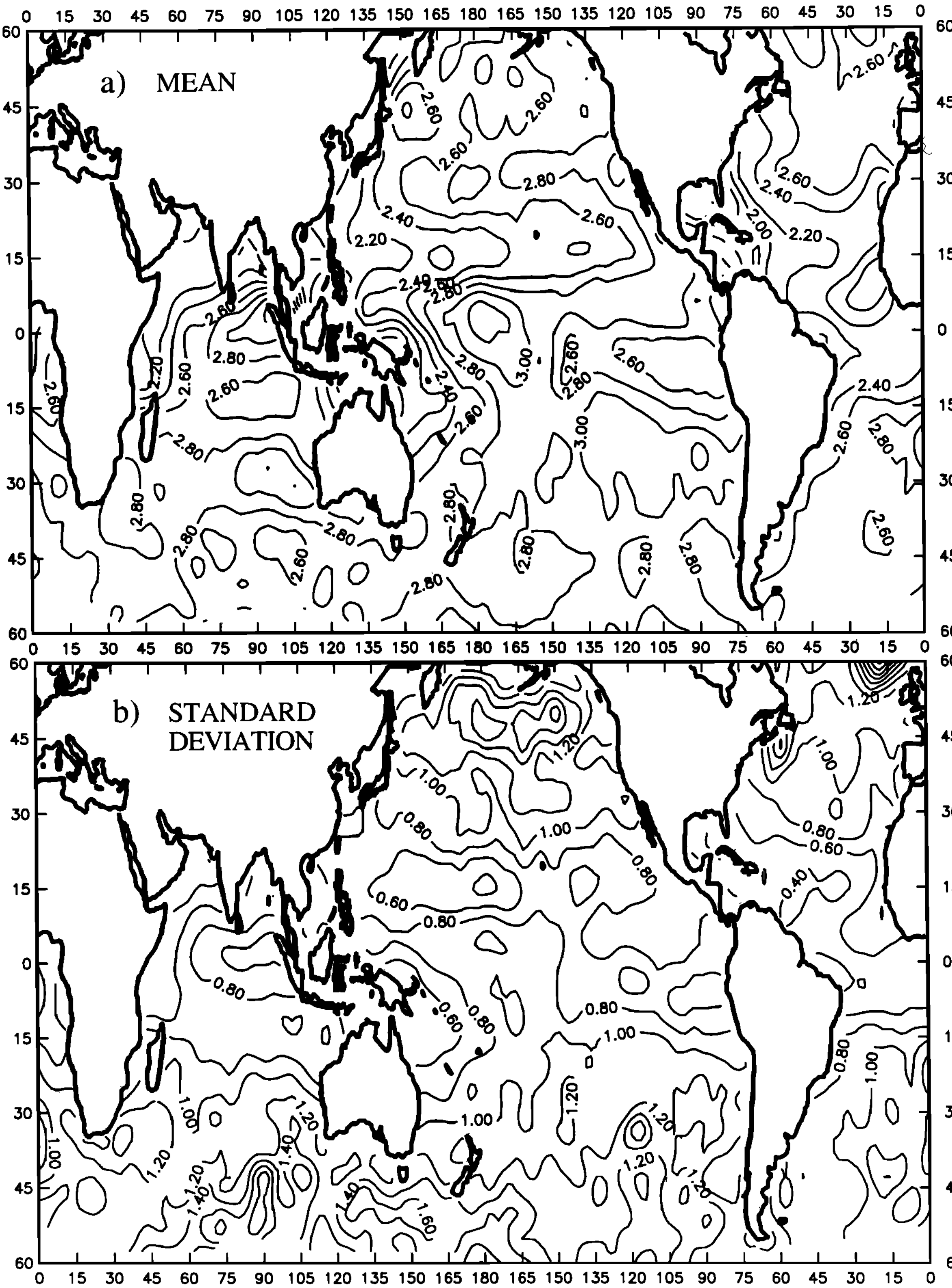

Figure 4. Maps of (a) the mean and (b) the standard deviation of the differences between the sea state bias corrections $\Delta h_{SSB} = bH_{1/3}$ in centimeters obtained using the collinear estimates of the three-parameter model equation (2) and the two-parameter model equation (15) for the bias coefficient b.

variance obtained using each model (see bottom rows of Tables 1 and 2). The reduction of variance for the three-parameter wind speed dependent model (2) considered in section 3.2 is 5% greater than that of the two-parameter wave height dependent model (15). Without knowing the effective number of degrees of freedom of the statistical estimates (for the reasons discussed in section 3.2), it is difficult to ascertain whether this difference is statistically significant. A pragmatic approach is to examine the differences between the sea state bias corrections Δh_{SSB} obtained using the two models. As a somewhat arbitrarily adopted threshold, differences exceeding 1 cm will be considered statistically significant. This value was chosen to be commensurate with the mission objectives of achieving better than 1-cm accuracy for each of the TOPEX range corrections.

The mean and standard deviation of the differences between the sea state bias corrections obtained using the two models were computed at each grid location. Contour maps of the results are shown in Figure 4. It is evident from Figure 4a that the three-parameter corrections are consistently smaller in magnitude than the two-parameter corrections (resulting in positive differences everywhere, since both sea state bias coefficients are negative); the mean differences range from about 2.4 cm to 3.2 cm, with the largest differences occurring in regions of low wind speed. A geographically uniform mean difference would have no effect on geostrophic velocities estimated from altimeter data. The magnitude of the mean difference is therefore not nearly as important as spatial variability of the mean difference. In regions of strong horizontal gradients of the mean differences, which generally coincide with regions of strong gradients in the wind field (Figure 4a), the two sea state bias corrections result in different mean geostrophic velocity fields.

For altimetric studies of sea level variability, the mean difference between the two sea state bias corrections is not a concern because it is removed along with all other time-invariant contributions to the altimetric estimates of sea level. For variability studies, only the standard deviation shown in Figure 4b is important. The standard deviation ranges from a minimum of about 0.4 cm in the tropics to a maximum of 1.8 cm in the westerly wind belts.

In consideration of the geographical variability of the mean differences in Figure 4a and the fact that the standard deviations in Figure 4b are in excess of 1 cm everywhere outside of the tropics, it can be concluded that the models (2) and (15) are significantly different for altimetric studies of both the mean and time-varying ocean circulation. Because it accounts for a greater reduction of the variance, the three-parameter model (2) is the better model for the sea state bias coefficient b. The differences in Figure 4 can therefore be interpreted as indicative of errors in the two-parameter model (15).

Table 3. The Symmetric 4 × 4 Matrix of Cross Correlations Between Wind Speed (u), Wind Speed Squared, Significant Wave Height ($H_{1/3}$), and Significant Wave Height Squared

	u	u^2	$H_{1/3}$	$H_{1/3}^2$
u	1.00	0.96	0.74	0.68
u^2		1.00	0.78	0.75
$H_{1/3}$			1.00	0.95
$H_{1/3}^2$				1.00

5. Combined Wind Speed and Wave Height Dependence

The attributes of the wind speed and wave height dependent models for the sea state bias coefficient b presented in sections 3 and 4 are very perplexing. The wind speed dependent model exhibits an increase in b with increasing wind speed up to about 12 m s^{-1}, consistent with previous aircraft and tower radar measurements. Wind speed and significant wave height are positively correlated with a correlation of 0.74 (Table 3). One would therefore expect b also to increase with increasing $H_{1/3}$, but this is opposite what is observed. The same results were obtained by *Gaspar et al.* [this issue], although they did not discuss the paradox of the seemingly inconsistent dependence of b on wind speed and wave height. Evidently, important independent information about the sea state bias is contained in the wind speed and wave height data.

This suggests considering higher-order models for b that include both wind speed and significant wave height. Least squares estimation of the parameters of such higher-order models is tricky, however, because of the high correlations between the various linear and nonlinear combinations of wind speed and significant wave height (see Table 3). The uncertainties in the regression estimates of the parameters increase with the number of parameters included in the model or when the correlations between regression variables are high (e.g., section 7 of *Chelton* [1983]). The number of parameters in these higher-order models must therefore be restricted.

Gaspar et al. [this issue] conducted a thorough search of parameter space and concluded that the best four-parameter model for the sea state bias coefficient b in (9) based on their crossover analysis was

$$b(u, H_{1/3}) = a_0 + a_1 u + a_2 u^2 + a_3 H_{1/3}. \qquad (16)$$

This model is a defensible choice as it combines the quadratic dependence on wind speed evident from Figures 1 and 2 and the approximate linear dependence on wave height evident from Figure 3. The intent here is not to duplicate the *Gaspar et al.* [this issue] screening of regression variables based on collinear analysis. Indeed, strong arguments can be made against the virtue of systematically examining all possible four-parameter regression models to find the one model that yields the smallest mean square residuals [e.g., *Davis,* 1977; *Barnett and Hasselmann,* 1979; *Hasselmann,* 1979].

Table 4. The Parameters of the Wind Speed and Wave Height Dependent Model $b = a_0 + a_1 u + a_2 u^2 + a_3 H_{1/3}$ for the Sea State Bias Coefficient

	Collinear Regression	Crossover Regression
a_0	-0.027 ± 0.016	-0.019 ± 0.009
a_1	-0.0028 ± 0.0009	-0.0037 ± 0.0008
a_2	$(1.0 \pm 0.3) \times 10^{-4}$	$(1.4 \pm 0.3) \times 10^{-4}$
a_3	0.0028 ± 0.0012	0.0027 ± 0.0011
VAR	8.04	7.87

The two values for each parameter correspond to (1) the values derived here by regression analysis of collinear TOPEX data and (2) the values derived by *Gaspar et al.* [1994] by regression analysis of crossover difference TOPEX data. The bottom row of the table gives the reduced variance (VAR) of sea level in cm^2 after removal of the sea state bias.

The problem is that the likelihood of finding a regression variable that is correlated with the sea state bias purely by chance over the sample data set increases with the number of potential regression variables considered. This complicates determination of the effects of sampling errors on the uncertainties of the regression parameters. The objective here is therefore to adopt the Gaspar et al. four-parameter model (16) developed from crossover analysis of TOPEX data and use collinear analysis to provide totally independent estimates of the regression parameters a_0, a_1, a_2, and a_3.

The estimates of the parameters of the model (16) obtained here from regression analysis of the residual collinear TOPEX data are listed in Table 4 and the corresponding bias coefficient is contoured as a function of wind speed and wave height in Figure 5a. The four-parameter model combines the attributes of the separate wind speed and wave height models considered in sections 3.2 and 4: b increases with increasing wind speed up to about 12 m s^{-1} and decreases monotonically with increasing wave height. The ± 1 standard deviation uncertainty of the four-parameter bias coefficient (contoured in Figure 5b) is about 0.011 near the peak of the joint distribution of wind speed and wave height shown in Figure 5c. The resulting uncertainty of 1.1% of $H_{1/3}$ for the sea state bias correction Δh_{SSB} is larger than the uncertainty of 0.8% of $H_{1/3}$ obtained for the three-parameter model. This reflects the above noted effects of the increased number of parameters and the high correlation between the regression variables (see Table 3).

For comparison, the estimates of the four parameters obtained by *Gaspar et al.* [this issue] from regression analysis of crossover difference TOPEX data are also listed in Table 4 and the corresponding bias coefficient is contoured in Figure 6a. As with the models considered in sections 3.2 and 4, the collinear and crossover regression estimates of the four-parameter model for b are very similar. The difference (Figure 6b) is only about 0.003 near the peak of the joint distribution of wind speed and wave height, which easily falls within the ± 1 standard deviation uncertainty (Figure 5b) obtained here for the collinear estimates. As with the other two regression models considered here, the uncertainties of the regression coefficients obtained by *Gaspar et al.* [this issue] from crossover analysis are very similar to those obtained here by collinear analysis.

The collinear three- and four-parameter models (16) and (2) can be evaluated as in section 4 by examining the differences between the sea state bias corrections Δh_{SSB} obtained using the two models. Contour maps of the mean and standard deviation of the differences are shown in Figure 7. Similar to the comparison in Figure 4 between the two- and three-parameter models, the differences are dominated by a mean difference that ranges from about 2.2 cm in the western tropics within each ocean basin to about 3.0 cm in the westerly wind belts (Figure 7a). The spatial gradients of this mean difference are generally smaller than in Figure 4. Consequently, the mean geostrophic velocity fields are only moderately sensitive to the differences between the three- and four-parameter sea state bias models.

Equatorward of about 30°, the standard deviation of the differences between the three- and four-parameter models (Figure 7b) is spatially uniform with a value of about 0.4 cm. At higher latitudes, the standard deviation increases to about 1.2 cm in the westerly wind belts of both hemispheres. It can be concluded that the greater reduction of variance for the four-parameter model (16) compared with the three-parameter model (2) (see bottom rows of Tables 4 and 1) is statistically significant according to the 1-cm difference criterion adopted in section 4. The four-parameter model for the sea state bias coefficient b is therefore significantly better than the three-parameter GDR model.

6. Summary and Conclusions

The sea state bias in altimeter estimates of sea level has been modeled in this study in the traditional way as $\Delta h_{SSB} = bH_{1/3}$, where $H_{1/3}$ is the significant wave height. Three different parameterizations of the bias coefficient b were considered and the parameters of the various models were estimated by least squares regression of residual collinear TOPEX data after removing appropriate mean values at each gridded data location.

The three-parameter model in terms of wind speed only (section 3.2) is significantly better than the two-parameter model in terms of $H_{1/3}$ only (section 4). The two separate models indicate that b increases with increasing wind speed up to about 12 m s^{-1} and decreases monotonically with increasing $H_{1/3}$. This conclusion is evidently robust, as the same results have been obtained by *Gaspar et al.* [this issue] from a completely independent analysis using TOPEX crossover differences. In consideration of the high positive correlation between wind speed and $H_{1/3}$, the opposite dependencies of b on increasing wind speed and increasing $H_{1/3}$ are perplexing. The sea state bias coefficient must depend on both wind speed and wave height. Purely empirical

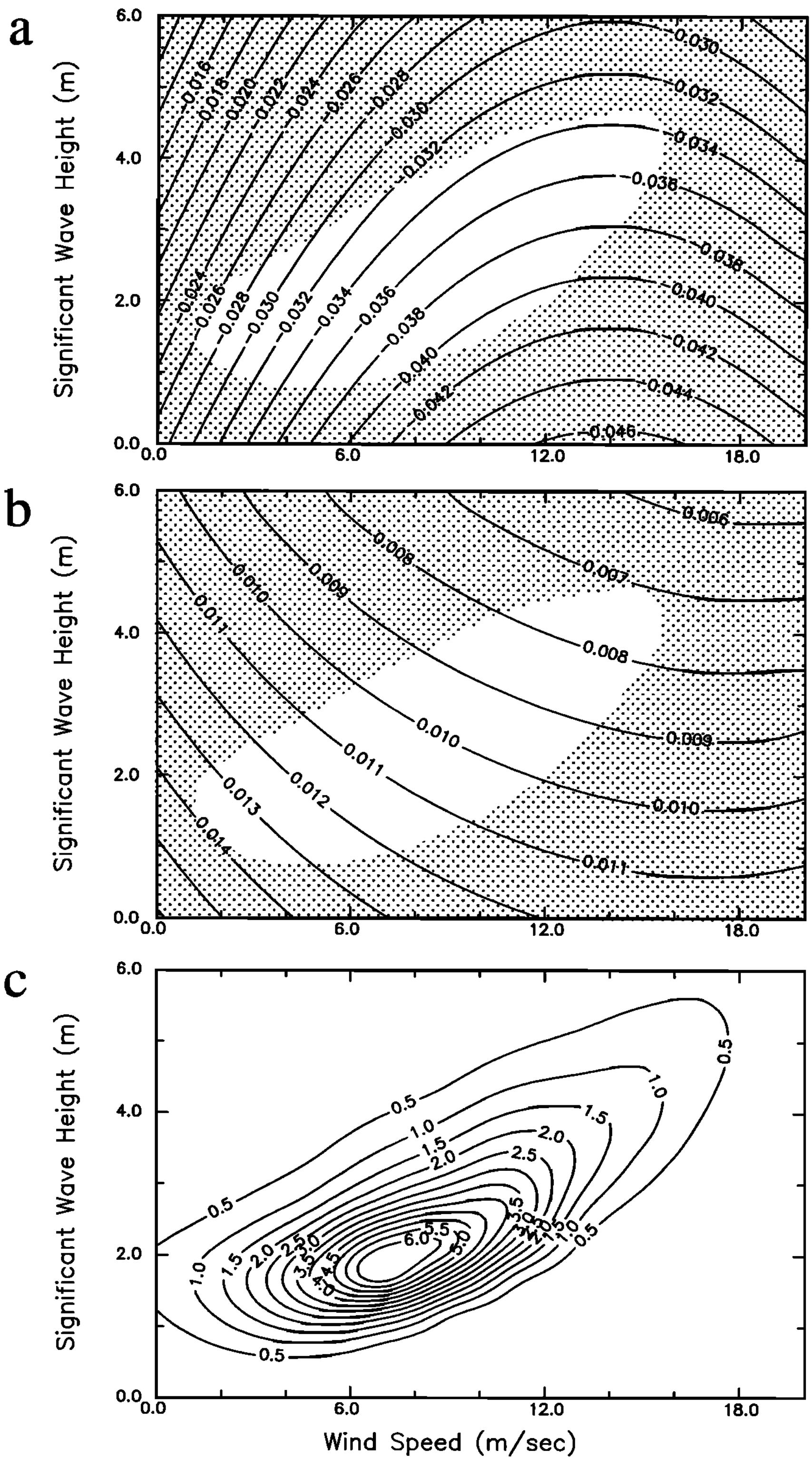

Figure 5. Contour plots of the wind speed and significant wave height dependence of (a) the four-parameter model equation (16) for the sea state bias coefficient obtained here from regression analysis of residual collinear TOPEX data, (b) the ±1 standard deviation uncertainty of the collinear regression estimates obtained here, and (c) the percentages of global TOPEX observations for wind speed bins of width 2 m s^{-1} and wave height bins of width 0.5 m. Shading in Figures 5a and 5b represents the region in Figure 5c with wind speed and wave height bins containing less than 1% of the global TOPEX data.

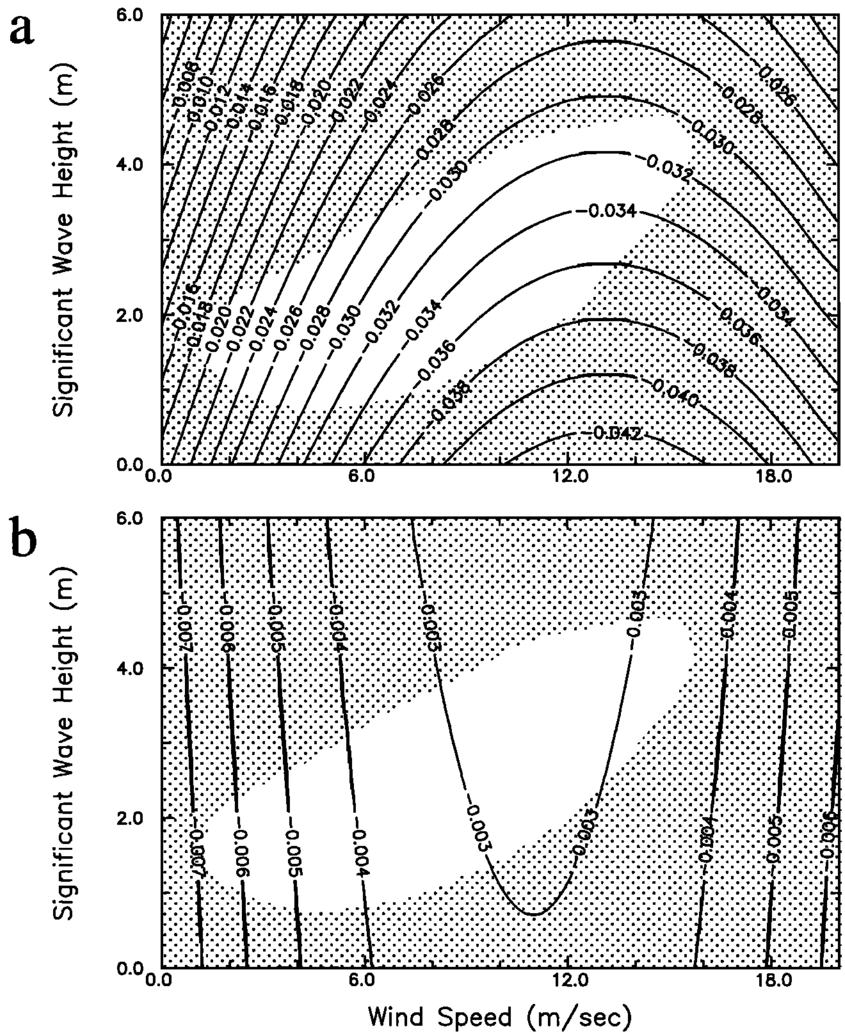

Figure 6. Contour plots of the wind speed and significant wave height dependence of (a) the four-parameter model equation (16) for the sea state bias coefficient obtained by *Gaspar et al.* [this issue] from regression analysis of crossover difference TOPEX data. (b) The difference between the collinear and crossover estimates of the four-parameter model shown in Figures 5a and 6a, respectively. The shading in both plots represents the region in Figure 5c with wind speed and wave height bins containing less than 1% of the global TOPEX data.

models such as those considered here cannot elucidate the physical basis for these effects.

A four-parameter model for b that incorporates the combined effects of wind speed and wave height was considered in section 5. The estimates of the four parameters obtained here by collinear regression analysis are very nearly the same as those estimated independently by *Gaspar et al.* [this issue] by regression analysis of crossover differences. The four-parameter model was shown to be significantly better than the three-parameter model (which is the formulation used to produce the geophysical data records). Over much of the world ocean, the standard deviation of the differences between the bias corrections Δh_{SSB} obtained using the two models is small. In regions of high wind speed and wave height, however, the differences exceed 1 cm. The GDR formulation for the sea state bias should therefore be expanded to account for the wave height dependence of b.

Considerable attention has been devoted in this study to the uncertainties of the empirical models for b. As discussed in section 3.2, classical techniques are not appropriate since the residuals of the regression models are not all statistically independent. The approach used here was to partition the full data set into 21 subsets, defined by the boundaries of the 10-day TOPEX repeat cycles. The uncertainties of the parameter estimates were then determined from the variability of the parameters obtained from the 21 subsets of data. This approach effectively treats each cycle as an independent realization, thus providing a conservative measure of the stability of the parameter estimates.

The conclusions of the error analyses are that the uncertainty of the sea state bias correction Δh_{SSB} is about 1% of $H_{1/3}$. Most of this uncertainty appears to be in the constant offset term a_0 in the regression. Uncertainties of this magnitude are rather disturbing. For a typical significant wave height of 2 m, the uncertainty

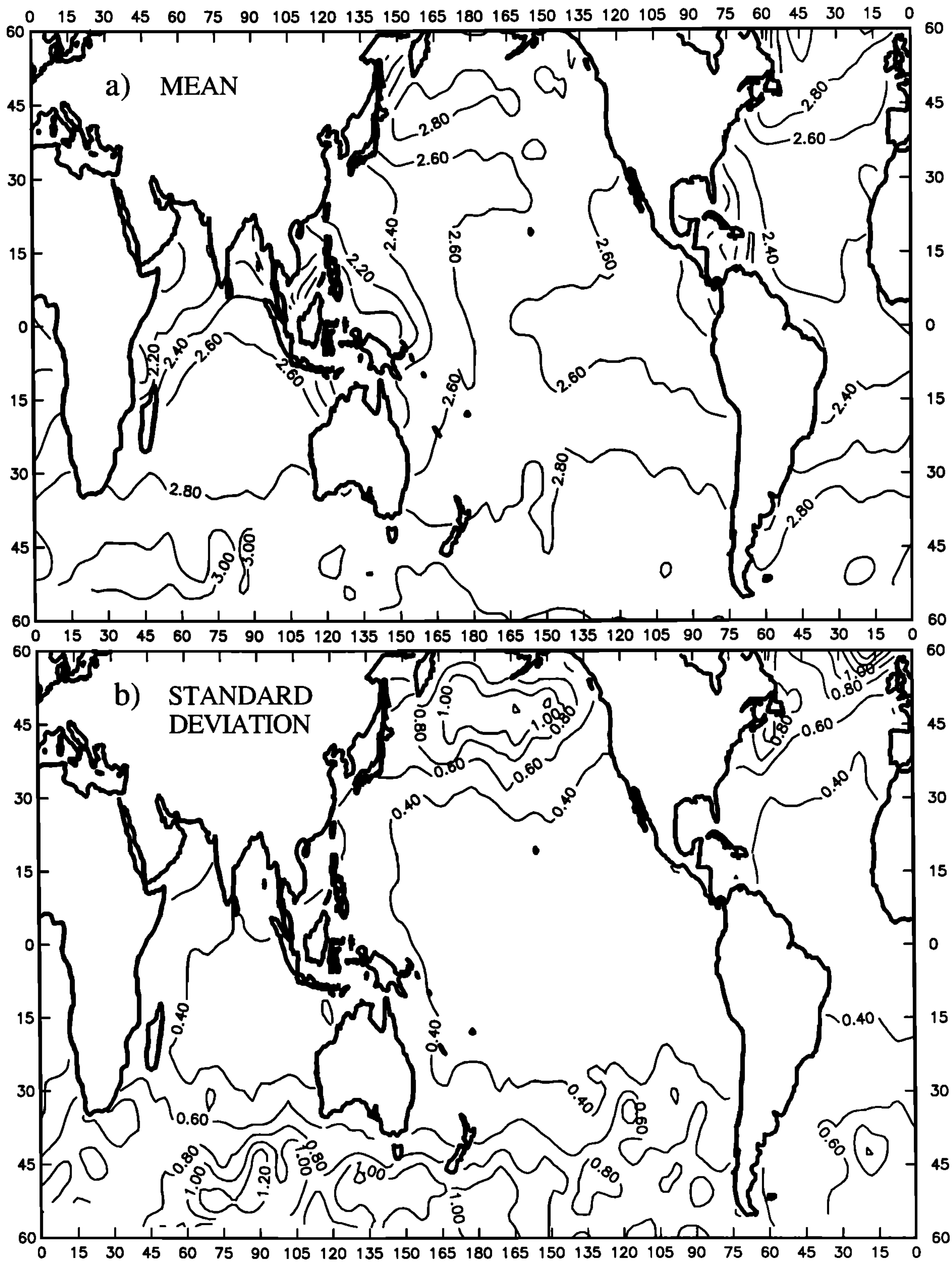

Figure 7. Maps of (a) the mean and (b) the standard deviation of the differences between the sea state bias corrections $\Delta h_{SSB} = bH_{1/3}$ in centimeters obtained using the collinear estimates of the three-parameter model equation (15) and the four-parameter model equation (16) for the bias coefficient b.

of Δh_{SSB} is 2 cm, which represents the largest error in the TOPEX error budget for the media corrections of the range measurements. Indeed, this error is nearly as large as the 3 to 4-cm time dependent orbit errors in TOPEX data.

The ability to unravel the wind speed and wave height dependencies of the sea state bias from empirical regression analyses is severely limited by the high cross correlations between the various wind speed and wave height variables (see Table 3). An improved understanding of the physical basis for the sea state bias may provide some insight into the reasons for the large uncertainty of the sea state bias, as well as the paradox of the seemingly inconsistent dependence on wind speed and wave height.

Acknowledgments. I thank Ernesto Rodríguez for drawing my attention to a flaw in the *Witter and Chelton* [1991b] implementation of the nonparametric method of estimating the sea state bias coefficient, and Michael Schlax for help resolving the detailed nature of the flaw, as explained in section 3.1. I also thank Michael Schlax for performing the computer calculations for this study. This research was supported by contract 958127 from the Jet Propulsion Laboratory funded under the TOPEX Announcement of Opportunity.

References

Barnett, T. P., and K. Hasselmann, Techniques of linear prediction, with application to oceanic and atmospheric fields in the tropical Pacific, *Rev. Geophys., 17,* 949–968, 1979.

Barrick, D. E., and B. J. Lipa, Analysis and interpretation of altimeter sea echo, *Adv. Geophys., 27,* 60–99, 1985.

Chelton, D. B., Effects of sampling errors in statistical estimation, *Deep Sea Res., 30,* 1083–1103, 1983.

Chelton, D. B., WOCE/NASA altimeter algorithm workshop, *U.S. WOCE Tech. Rep. 2,* 70 pp., U.S. Planning Office for WOCE, College Station, Tex., 1988.

Davis, R. E., Techniques for statistical analysis and prediction of geophysical fluid systems, *Geophys. Astrophys. Fluid Dyn., 8,* 245–277, 1977.

Fu, L.-L., and R. Glazman, The effect of the degree of wave development on the sea state bias in radar altimetry measurement, *J. Geophys. Res., 96,* 829–834, 1991.

Gaspar, P., F. Ogor, P.-Y. Le Traon, and O.-Z. Zanife, Estimating the sea state bias of the TOPEX and POSEIDON altimeters from crossover differences, *J. Geophys. Res.,* this issue.

Hasselmann, K., Linear statistical models, *Dyn. Atmos. Oceans, 3,* 501–521, 1979.

Hevizi, L. G., E. J. Walsh, R. E. McIntosh, D. Vandemark, D. E. Hines, R. N. Swift, and J. F. Scott, Electromagnetic bias in sea surface range measurements at frequencies of the TOPEX/POSEIDON satellite, *IEEE Trans. Geosci. Remote Sens., 31,* 376–388, 1993.

Jackson, F. C., The reflection of impulses from a nonlinear random sea, *J. Geophys. Res., 84,* 4939–4943, 1979.

Lagerloef, G. S. E., Comment on "On the joint distribution of surface elevations and slopes for a nonlinear random sea, with an application for radar altimetry" by M. A. Srokosz, *J. Geophys. Res., 92,* 2985–2987, 1987.

Rodríguez, E., Altimetry for non-Gaussian oceans: Height biases and estimation of parameters, *J. Geophys. Res., 93,* 14,107–14,120, 1988.

Srokosz, M. A., On the joint distribution of surface elevations and slopes for a nonlinear random sea, with an application for radar altimetry, *J. Geophys. Res., 91,* 995–1006, 1986.

Srokosz, M. A., Reply, *J. Geophys. Res., 92,* 2989–2990, 1987.

Tapley, B., et al., Precision orbit determination for TOPEX/POSEIDON, *J. Geophys. Res.,* this issue.

Witter, D. L., and D. B. Chelton, A Geosat wind speed algorithm and a method for altimeter wind speed algorithm development, *J. Geophys. Res., 96,* 8853–8860, 1991a.

Witter, D. L., and D. B. Chelton, An apparent wave height dependence in the sea state bias in Geosat altimeter range measurements, *J. Geophys. Res., 96,* 8861–8867, 1991b.

Zlotnicki, V., L.-L. Fu, and W. Patzert, Seasonal variability in global sea level observed with Geosat altimetry, *J. Geophys. Res., 94,* 17,959–17,969, 1989.

Zlotnicki, V., A. Hayashi, and L.-L. Fu, The JPL-Oceans 8902 version of Geosat altimetry data, *Rep. JPL D-6939,* 17 pp. plus figures, Jet Propul. Lab., Pasadena, Calif., March 1990.

D. B. Chelton, College of Oceanic and Atmospheric Sciences, Oregon State University, Oceanography Administration Building 104, Corvallis, OR 97331-5503.

(Received February 7, 1994; revised July 26, 1994; accepted August 16, 1994.)