Estimation of the sea state bias in radar altimeter measurements of sea level: Results from a new nonparametric method

Article *in* Journal of Geophysical Research Atmospheres · July 1998

DOI: 10.1029/98JC01194

CITATIONS
76

READS
214

2 authors, including:

Philippe Gaspar
Mercator Ocean
122 PUBLICATIONS **4,466** CITATIONS

SEE PROFILE

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 103, NO. C8, PAGES 15,803–15,814, JULY 15, 1998

Estimation of the sea state bias in radar altimeter measurements of sea level: Results from a new nonparametric method

Philippe Gaspar
Collecte Localisation Satellites, Space Oceanography Division, Ramonville, France

Jean-Pierre Florens
Groupe de Recherche en Economie Mathematique et Quantitative–Institut d'Economie Industrielle
Université des Sciences Sociales, Toulouse, France

Abstract. The sea state bias (SSB) affecting altimetric measurements of the sea surface height (SSH) is classically estimated using empirical parametric models. The model parameters are determined to minimize the variance of the SSH differences at crossover points or along collinear tracks. It is shown here that so-calibrated models are not true least squares approximations of the SSB. Simulations indicate that the difference from a true least squares solution is typically a few millimeters to a few centimeters, varying with sea state. This difference proves to be the result of the inevitably imperfect specification of the model's parametric form, which corrupts the calibration process when performed on SSH differences rather than directly on SSH measurements. To avoid specification of a parametric form, we present a new nonparametric version of the SSB estimation problem and propose a general solution based on the statistical technique of kernel smoothing. This solution is then used to obtain the first fully nonparametric estimate of the TOPEX altimeter SSB as a function of both the wind speed and the wave height. The obtained estimate, though based on a limited data set, proves to be better than those obtained with the standard parametric SSB models featured in the TOPEX-POSEIDON Geophysical Data Records. The improvement is most significant in middle- and high-latitude oceans.

1. Introduction

Yaplee et al. [1971] first observed that the sea surface radar cross section per unit area is larger toward the wave troughs and smaller toward the crests. As a result, the sea level measured by radar altimeters appears to be lower than the true mean sea level. This effect is known as the electromagnetic bias. Technical estimation problems also give rise to two other sea-state-related biases: the skewness and tracker biases (for details, see *Chelton et al.* [1989]). These are generally lumped together with the electromagnetic bias to form the so-called sea state bias (SSB). The SSB is typically a few percent of the significant wave height (SWH) and thus ranges between a few centimeters and a few decimeters.

With a standard deviation as large as 2 cm, the SSB estimation error is, today, the largest error affecting the range measurements [*Fu et al.*, 1994] of the TOPEX and POSEIDON altimeters. The relatively poor quality of the SSB estimation is largely due to our limited understanding of the mechanisms that generate the SSB, and the electromagnetic bias, in particular. In the absence of well-established theoretical guidelines, a wide variety of empirical models have been developed, naturally leading to some controversy on the choice of the most appropriate one [e.g., *Rodriguez and Martin*, 1994; *Glazman et al.*, 1994]. Still, all empirical models rely on the hypothesis that the SSB can be modeled under the general form

$$\mathrm{SSB} = \mathrm{SWH}\, b(\mathbf{x}, \boldsymbol{\theta}) \tag{1}$$

Paper number 98JC01194.
0148-0227/98/98JC-01194$09.00

where b is a specified function of $\mathbf{x}$, a vector of pertinent sea-state-related variables, and $\boldsymbol{\theta}$, a vector of parameters. The function b is sometimes referred to as the relative bias. In practice, the components of $\mathbf{x}$ are systematically chosen among the few sea-state-related variables that can be directly measured by an altimeter: SWH, the wind speed (U), the backscatter coefficient (σ_0), or any combination thereof. Most models actually assume that b is a simple polynomial function of U and/or SWH [e.g., *Born et al.*, 1982; *Douglas and Agreen*, 1983; *Zlotnicki et al.*, 1989; *Melville et al.*, 1991; *Ray and Koblinsky*, 1991; *Gaspar et al.*, 1994; *Chelton*, 1994], but other parametric forms have been used [*Fu and Glazman*, 1991; *Glazman et al.*, 1994]. In all cases, the model parameters are determined to minimize the variance of the measured sea surface height (SSH) differences at crossover points or along collinear tracks. We show in this paper (section 2) that the set of model parameters that minimizes the variance of the crossover or collinear differences is not the set of model parameters that minimizes the distance between the true SSB and the model approximation. The obtained model is thus not a true least squares approximation of the SSB. It differs from it by, typically, a few millimeters to a few centimeters. This problem proves to be a consequence of the inevitably imperfect specification of the model's parametric form. An incomplete or not fully adequate model formulation generates autocorrelated errors which affect the model calibration obtained with SSH differences rather than directly with SSH measurements.

The search for a solution to this problem motivated the present work. Having recognized that the origin of the problem lies in the parametric approach, one is naturally led to consider use of nonparametric estimation methods. These methods

eliminate the rigidity of parametric methods, and the resulting errors, by removing the restriction that the fitted model belongs to a parametric family. Another advantage of these methods is that they leave no place for controversy about the choice of the functional form of the model. This choice is simply not needed.

Even if the above mentioned problem of the parametric approach had not been previously identified, a few authors had recognized that the use of parametric models might overly constrain the simulated behavior of the SSB and had developed nonparametric solutions. *Witter and Chelton* [1991a] first obtained a simple nonparametric estimation of the Geosat relative bias assuming that b was an unspecified function of either U or SWH. This function was estimated piecewise over discrete U or SWH intervals. The original implementation of the method proved to be incorrect but was subsequently corrected by *Chelton* [1994] and applied to TOPEX data. *Rodriguez and Martin* [1994] also estimated the TOPEX SSB in a partly nonparametric way. They assumed that the relative bias could be modeled as the sum of a nonparametric function of U and a linear function of SWH. The function of U was approximated by a sum of local basis functions. These models thus allow nonparametric variations of the SSB in only one dimension (U or SWH) and still conform to the usual formulation (equation (1)). Limitation to the one-dimensional case is more guided by numerical constraints than by the physics of the problem, and use of (1) appears to be a convenient, but not indispensable, choice. Indeed, it is now clear that the SSB dependence on SWH is not strictly linear. The results of *Witter and Chelton* [1991a], *Gaspar et al.* [1994], and *Chelton* [1994] show that the magnitude of the relative bias decreases when SWH increases. Theoretical models also suggest that the SSB dependence on SWH is weaker than linear [*Barrick and Lipa*, 1985; *Fu and Glazman*, 1991]. There is thus no compelling reason to formally single out SWH in the specification of a SSB model.

The SSB estimation method developed in this paper is multivariate and fully nonparametric. It avoids use of (1) and simply assumes that SSB is a (nonspecified) function of a set of pertinent variables. The SSB estimation problem is reformulated in section 3 using the now-classical concept of nonparametric regression. A general solution based on the statistical technique of kernel smoothing is presented. This solution is then used to estimate the SSB of the TOPEX Ku-band altimeter, based on a data set described in section 4. Section 5 deals with the numerical implementation of the solution. The results are presented in section 6 and compared with those obtained with commonly used parametric models.

This paper focuses on the practical implementation of the new estimation technique and on the discussion of the results obtained for the TOPEX SSB. A more theoretical analysis of the nonparametric SSB estimation problem and of the statistical properties of the obtained solution is given by J.-P. Florens and P. Gaspar (manuscript in preparation, 1998).

2. On the Use of Parametric Models for SSB Estimation

2.1. Basis for Parametric SSB Estimation

Let us denote h'_a the altimetric range measurement not corrected for the sea state bias. The difference between the satellite altitude above the reference ellipsoid and h'_a yields a biased sea surface height measurement (SSH′) that contains the geoid signal (h_g), the ocean dynamic topography (η), the SSB, and some measurement noise (w):

$$\mathrm{SSH}' = h_g + \eta + \mathrm{SSB} + w \tag{2}$$

The geoid signal, largely dominant but still imperfectly known, is easily eliminated from this equation by forming differences between two SSH′ measurements taken at different times but at the same geographic location (either along collinear tracks or at crossover points). Noting these two measurements SSH'_1 and SSH'_2, one obtains

$$\mathrm{SSH}'_2 - \mathrm{SSH}'_1 = (\mathrm{SSB}_2 - \mathrm{SSB}_1) + (\eta_2 - \eta_1) + (w_2 - w_1) \tag{3}$$

In the context of SSB estimation, the change in dynamic topography ($\eta_2 - \eta_1$) is considered as noise and lumped together with ($w_2 - w_1$) to form a single zero-mean noise term ε:

$$\mathrm{SSH}'_2 - \mathrm{SSH}'_1 = (\mathrm{SSB}_2 - \mathrm{SSB}_1) + \varepsilon \tag{4}$$

This observation equation is central to all SSB estimation techniques that rely on altimeter data only. On the basis of this equation, estimation of SSB is simple if a parametric form of the bias is assumed. The commonly used formulation is (1), but a more general expression is

$$\mathrm{SSB} = \varphi(\mathbf{x}, \boldsymbol{\theta}) \tag{5}$$

where φ is any specified function of the pertinent variables $\mathbf{x}$ and the parameters $\boldsymbol{\theta}$. Once the functional form of φ is chosen, sets of altimetric measurements can be used to determine the value of $\boldsymbol{\theta}$ that minimizes the variance of the sea surface height (SSH) differences after correction for the (modeled) SSB. This variance reads

$$E[\mathrm{SSH}_2 - \mathrm{SSH}_1]^2 = E[(\mathrm{SSH}'_2 - \varphi_2) - (\mathrm{SSH}'_1 - \varphi_1)]^2 \tag{6}$$

where E denotes an expectation and the abbreviated notation φ_i is used for $\varphi(\mathbf{x}_i, \boldsymbol{\theta})$. In most published models, φ is a linear function of $\boldsymbol{\theta}$ so that the minimization problem reduces to a simple, generally multivariate, regression problem [e.g., *Ray and Koblinsky*, 1991]. Standard minimization techniques also exist to deal with nonlinear models [e.g., *Fu and Glazman*, 1991].

A word of caution must be added about the use of the parametric form (5) instead of the more usual form (1). Indeed, the variance to minimize (6) only contains SSB differences. Therefore adding a constant to any SSB estimate that minimizes (6) will generate another minimum variance solution. To eliminate this undetermination, one particular value of the SSB must be specified. This must be explicitly done when using (5). It is implicitly done when using (1), as this particular model formulation imposes SSB = 0 when SWH = 0.

2.2. Properties of the Parametric SSB Estimates

Using (4), the variance of the SSH differences can be rewritten

$$E[\mathrm{SSH}_2 - \mathrm{SSH}_1]^2 = E[(\mathrm{SSB}_2 - \varphi_2) - (\mathrm{SSB}_1 - \varphi_1) + \varepsilon]^2 \tag{7}$$

Assuming, as usual, that ε and SSB are uncorrelated, (7) becomes

$$E[\mathrm{SSH}_2 - \mathrm{SSH}_1]^2 = 2E[\mathrm{SSB} - \varphi]^2 - 2E[(\mathrm{SSB}_1 - \varphi_1)(\mathrm{SSB}_2 - \varphi_2)] + E[\varepsilon]^2 \tag{8}$$

where the variance of the model error $E[\mathrm{SSB} - \varphi]^2 = E[\mathrm{SSB}_1 - \varphi_1]^2 = E[\mathrm{SSB}_2 - \varphi_2]^2$. Equation (8) clearly shows that the value of $\boldsymbol{\theta}$ that minimizes $E[\mathrm{SSH}_2 - \mathrm{SSH}_1]^2$ actually minimizes

$$E[\mathrm{SSB} - \varphi]^2 - E[(\mathrm{SSB}_1 - \varphi_1)(\mathrm{SSB}_2 - \varphi_2)] \quad (9)$$

In general, this value of $\boldsymbol{\theta}$ does not realize the least squares approximation of SSB by φ since it does not minimize $E[\mathrm{SSB} - \varphi]^2$. A special case where the same $\boldsymbol{\theta}$ minimizes both (9) and $E[\mathrm{SSB} - \varphi]^2$ is the case of a perfectly specified model, i.e., a model for which there exists a value of $\boldsymbol{\theta}$, say, $\boldsymbol{\theta}_0$, such that $\varphi(\mathbf{x}, \boldsymbol{\theta}_0)$ is the true SSB. In that case, both $E[\mathrm{SSB} - \varphi]^2$ and (9) are minimized (equal to zero) for $\boldsymbol{\theta} = \boldsymbol{\theta}_0$. This shows that the use of measurement differences allows correct calibration of a perfect SSB model. One can hardly claim, however, that parametric SSB models are perfect. Therefore the autocorrelation of the model error, $E[(\mathrm{SSB}_1 - \varphi_1)(\mathrm{SSB}_2 - \varphi_2)]$, affects the SSB estimation process and causes the SSB estimate to deviate from the true least squares solution. Is the deviation significant? This cannot be told from the altimeter data themselves, as the very large geoid signal in SSH′ measurements prevents any precise determination of the true least squares approximation. Still, a simple simulation can provide useful insights.

2.3. Evaluation of the Deviation From the True Least Squares Solution

The goal of the simulation is to obtain a reasonable estimate of the difference between a true least squares estimate of the SSB and the more usual solution obtained by fitting a parametric model on measurement differences. To avoid estimation accuracy problems, we generated a set of "idealized SSH′ measurements" containing only a known SSB plus some noise, but no geoid or dynamic topography signal. To obtain this, we took a set of about 130,000 TOPEX altimeter measurements at crossover points, representing the 20 first repeat cycles of the TOPEX data set completely described in section 4. We then replaced the SSH′ data field by a new field (z) made of a SSB simulated by the BM4 model [*Gaspar et al.*, 1994] plus a Gaussian random noise:

$$z = \mathrm{BM4}(U, \mathrm{SWH}) + w \quad (10)$$

where

$$\mathrm{BM4}(U, \mathrm{SWH}) = \mathrm{SWH}[a_1 + a_2 U + a_3 U^2 + a_4\,\mathrm{SWH}] \quad (11)$$

The values of the a_i parameters are those recommended by *AVISO* [1996] for TOPEX data. The standard deviation of the noise (w) is 6.3 cm, so that the standard deviation of the crossover differences after SSB correction ($w_2 - w_1$) is $6.3\sqrt{2} \approx 9$ cm, a realistic value for data of the TOPEX class.

On the basis of this data set, we proceeded with the calibration of an "imperfect" SSB model, in this case the widely used three-parameter model (BM3) of *Hevizi et al.* [1993]:

$$\mathrm{BM3}(U, \mathrm{SWH}) = \mathrm{SWH}[a_1 + a_2 U + a_3 U^2] \quad (12)$$

BM3 can obviously not provide a perfect fit to data generated with the more complete BM4 model. A first regression was performed using the z data themselves to obtain the true least squares approximation of BM4 by a BM3 model. A second regression was then performed using the crossover differences ($z_2 - z_1$). As expected, the two regressions yield different parameter estimates for BM3. The resulting differences between the two SSB estimates range from a few millimeters at low wave heights regularly increasing up to 3 cm for wave heights close to 10 m. Differences also vary slightly with wind speed. A similar pair of regressions was performed, this time with the most simple one-parameter model (BM1):

$$\mathrm{BM1}(\mathrm{SWH}) = a\,\mathrm{SWH} \quad (13)$$

In this case, the difference between the two model calibrations is equal to 0.4% of SWH, a 20% error for a typical SSB of −2% of SWH. It results in a 4-cm difference between the two SSB estimates for SWH = 10 m. This simple exercise thus shows that usual SSB models calibrated on crossover differences can be centimeters off the true least squares solution. In addition, this difference from the least squares solution varies with sea state. This can induce spurious SSH gradients between ocean regions with different sea surface conditions. Of course, this error is only a fraction of the total SSB estimation error, but it is a nagging one, as it is built in the parametric approach: The unavoidably imperfect specification of the model's parametric form corrupts its calibration when performed on SSH′ differences rather than directly on SSH′.

One way out of this problem is to avoid the specification of a parametric form, which naturally leads us to consider nonparametric estimation techniques. In the next section we reformulate the SSB estimation problem as a special nonparametric regression problem. It is special because the regression only uses measurement differences, not the measurements themselves. An original solution is presented, based on the now well-developed kernel smoothing technique.

3. Nonparametric SSB Estimation

3.1. Basis for Nonparametric Regression

Let us briefly recall the theoretical basis of nonparametric regression. Assume we have a random scalar variable ζ jointly distributed with a random vector $\mathbf{x}$. Given a set of n observations $(\zeta_i, \mathbf{x}_i)$, we want to estimate the regression function $r(\mathbf{x}) = E[\zeta|\mathbf{x}]$ without making assumptions concerning the functional form of r. Different technical solutions exist [e.g., *Härdle*, 1990]. Here we use the so-called kernel smoothing method that has the advantage of being quite intuitive and relatively simple to implement. Since the seminal papers of *Rosenblatt* [1956] and *Parzen* [1962], this technique has been extensively studied. It has now a strong theoretical basis [see, e.g., *Wand and Jones*, 1995].

The kernel estimator $\hat{r}(\mathbf{x})$ of the conditional expectation $E[\zeta|\mathbf{x}]$ is simply a weighted mean of the observed ζ_i, the weights being a decreasing function of a distance between $\mathbf{x}$ and $\mathbf{x}_i$:

$$\hat{r}(\mathbf{x}) = \sum_{i=1}^{n} \zeta_i \alpha_n(\mathbf{x} - \mathbf{x}_i) \quad (14)$$

where α_n is a weighting function whose general form reads

$$\alpha_n(\mathbf{x} - \mathbf{x}_i) = \frac{K\left(\dfrac{\mathbf{x} - \mathbf{x}_i}{\mathbf{h}_n}\right)}{\displaystyle\sum_{i=1}^{n} K\left(\frac{\mathbf{x} - \mathbf{x}_i}{\mathbf{h}_n}\right)} \quad (15)$$

K is a kernel function, $\mathbf{h}_n$ is the bandwidth vector, and the index n indicates that the bandwidth depends on the number of

measurements. A kernel is a symmetric scalar function satisfying $\int K(\mathbf{x})\, d\mathbf{x} = 1$. It is usually chosen to be a probability density function. The bandwidth vector has the same size as $\mathbf{x}$ and is made of positive numbers. The ratio $[\mathbf{x} - \mathbf{x}_i/\mathbf{h}_n]$ is simply defined component by component.

3.2. The SSB Case

Let us now concentrate on the specific SSB estimation problem. Simply start with the assumption that SSB is a (nonspecified) function φ of $\mathbf{x}$, a vector of p pertinent variables. As previously indicated, we will actually use $\mathbf{x} = (U, \mathrm{SWH})$, but this does not need to be specified at this stage. Noting $y = \mathrm{SSH}_2' - \mathrm{SSH}_1'$, equation (4) can be rewritten

$$y = \varphi(\mathbf{x}_2) - \varphi(\mathbf{x}_1) + \varepsilon \tag{16}$$

which implies

$$E[y|\mathbf{x}_2 = \mathbf{x}] = \varphi(\mathbf{x}) - E[\varphi(\mathbf{x}_1)|\mathbf{x}_2 = \mathbf{x}] \tag{17}$$

Having at our disposal a set of n measurements $(y_i, \mathbf{x}_{1i}, \mathbf{x}_{2i})$, the kernel estimator (14) can be used to evaluate the conditional expectations in (17). One obtains

$$\varphi(\mathbf{x}) = \sum_{i=1}^{n} y_i \alpha_n(\mathbf{x} - \mathbf{x}_{2i}) + \sum_{i=1}^{n} \varphi(\mathbf{x}_{1i})\alpha_n(\mathbf{x} - \mathbf{x}_{2i}) \tag{18}$$

This relation provides a practical means of estimating $\varphi(\mathbf{x})$, for any $\mathbf{x}$, provided that the values of $\varphi(\mathbf{x}_{1i})$ are known. In other words, the problem of estimating $\varphi(\mathbf{x})$ is now reduced to the problem of estimating φ for a finite set of values of $\mathbf{x}$. Equation (18) being valid for any of these values of $\mathbf{x}$, one can write

$$\varphi(\mathbf{x}_{1j}) = \sum_{i=1}^{n} y_i \alpha_n(\mathbf{x}_{1j} - \mathbf{x}_{2i}) + \sum_{i=1}^{n} \varphi(\mathbf{x}_{1i})\alpha_n(\mathbf{x}_{1j} - \mathbf{x}_{2i}) \tag{19}$$

$$\forall j = 1, \cdots, n$$

or, in matrix form,

$$(\mathbf{I} - \mathbf{A})\boldsymbol{\varphi}_1 = \mathbf{A}\mathbf{y} \tag{20}$$

where $\mathbf{I}$ is the $(n \times n)$ identity matrix, $\mathbf{A}$ is the $(n \times n)$ matrix with elements $a_{ji} = \alpha_n(\mathbf{x}_{1j} - \mathbf{x}_{2i})$, $\boldsymbol{\varphi}_1^T = [\varphi(\mathbf{x}_{11}), \cdots, \varphi(\mathbf{x}_{1n})]$, and $\mathbf{y}^T = [y_1, \cdots, y_n]$.

Unfortunately, the linear system (20) cannot be solved for $\boldsymbol{\varphi}_1$ since $\mathbf{I} - \mathbf{A}$ is singular. Indeed, the definition of the weights (equation (15)) implies that the sum of elements, in each line of this matrix, is zero. The rank of $\mathbf{I} - \mathbf{A}$ is thus equal to $n - 1$. We find here the numerical expression of the undetermination problem mentioned at the end of section 2.1. To fix it, we have to impose (rather than estimate) the value of one of the components of $\boldsymbol{\varphi}_1$, say, $\varphi(\mathbf{x}_{11})$, as the data set can always be rearranged to have the selected component in first place. The constraint is thus

$$\varphi(\mathbf{x}_{11}) = \varphi_0 \tag{21}$$

The choice of an appropriate $\mathbf{x}_{11}$ and φ_0 will be discussed in section 5.3. Given (21), equation (20) can be rewritten

$$\mathbf{B}_1\boldsymbol{\varphi} = \mathbf{A}\mathbf{y} - \mathbf{B}_0\varphi_0 \tag{22}$$

where $\boldsymbol{\varphi}$ is the vector made of the $n - 1$ components of $\boldsymbol{\varphi}_1$ that remain to be determined: $\boldsymbol{\varphi}^T = [\varphi(\mathbf{x}_{12}), \cdots, \varphi(\mathbf{x}_{1n})]$ and $(\mathbf{B}_0, \mathbf{B}_1)$ is a partition of $\mathbf{I} - \mathbf{A}$, $\mathbf{B}_0$ being simply the first column of $\mathbf{I} - \mathbf{A}$. We now have a system of n equations with $n - 1$ unknowns. A linear least squares solution is readily obtained:

$$\hat{\boldsymbol{\varphi}} = (\mathbf{B}_1^T\mathbf{B}_1)^{-1}\mathbf{B}_1^T(\mathbf{A}\mathbf{y} - \mathbf{B}_0\varphi_0) \tag{23}$$

This provides the missing estimates of $\varphi(\mathbf{x}_{1i})$, $i = 2, n$. These can then be plugged into (18) to estimate $\varphi(\mathbf{x})$ for any value of $\mathbf{x}$. The problem is thus formally solved. Provided that the bandwidth is properly chosen, it can be shown (J. P. Florens and P. Gaspar, manuscript in preparation, 1998) that the SSB estimator (equation (23)) is consistent and asymptotically normal. It is thus exempt from the problems affecting imperfectly specified parametric models.

The price to pay for the use of this nonparametric approach is the inversion of a $[(n - 1) \times (n - 1)]$ matrix, as required by (23). This becomes rapidly intractable as the number of measurements increases. We demonstrate in section 6 that useful solutions can be obtained with a reasonably small amount of data. Another potential problem is that there is no guarantee that $\mathbf{B}_1^T\mathbf{B}_1$ is not numerically singular or nearly so. We have not encountered this situation in the numerical experiments presented in section 6, but should this happen, the way to solve the problem is to impose the value of more than one of the components of $\boldsymbol{\varphi}_1$. The formal solution (equation (23)) is still valid but with a different $(\mathbf{B}_0, \mathbf{B}_1)$ partition of the $\mathbf{I} - \mathbf{A}$ matrix. If m components of $\boldsymbol{\varphi}_1$ have to take specified values, $\mathbf{B}_0$ is made of the m first columns of $\mathbf{I} - \mathbf{A}$.

4. Data Processing

To evaluate the SSB estimation technique presented in section 3, we used it to estimate the SSB of the TOPEX Ku-band altimeter. This estimation is based on 100 cycles of TOPEX measurements obtained between April 4, 1993, and April 15, 1996. This period actually covers 111 ten-day repeat cycles (cycles 21–131) from which we have discarded 10 POSEIDON cycles and a missing TOPEX cycle (cycle 118). Use is made of the crossover data sets provided by *AVISO* [1996] based on the recently reprocessed TOPEX-POSEIDON Geophysical Data Records (GDRs). The crossovers are computed cycle per cycle. The crossover differences ($y = \mathrm{SSH}_2' - \mathrm{SSH}_1'$) correspond to the descending arc measurement minus the ascending arc measurement. Therefore index 1 (index 2) denotes measurements taken on the ascending arc (descending arc) in the rest of this paper.

SSH′ is computed using the CNES orbit and the TOPEX Ku-band range measurements to which all standard corrections except the SSB are applied. More precisely, we apply the dry tropospheric and inverse barometer corrections deduced from the European Centre for Medium-Range Weather Forecasts (ECMWF) pressure analyses, the wet tropospheric correction deduced from TOPEX microwave radiometer measurements, the ionospheric correction deduced from the dual-frequency TOPEX altimeter measurements, the ocean tide computed with the Center for Space Research (CSR) 3.0 model, the solid Earth, and pole tides. The wind speed (U) is the 10-m wind speed computed by *AVISO* [1996] using the Ku-band backscatter coefficient (σ_0) and the modified Chelton-Wentz algorithm [*Witter and Chelton*, 1991b].

Stringent editing is applied to the data. All measurements with missing or unreliable corrections are eliminated. Also discarded are the measurements with $\sigma_0 < 7$ dB, $\sigma_0 > 30$ dB, or SWH > 12 m. After quality control a total of nearly 633,000 valid crossover differences are obtained.

5. Numerical Implementation

5.1. Choice of the Kernel Function and Bandwidth Vector

Estimation of the SSB based on (18) and (23) requires the specification of a kernel and the selection of an appropriate bandwidth vector. There is an infinite number of kernel functions, but, in practice, only a few are commonly used [see, e.g., *Wand and Jones*, 1995]. These have very similar statistical properties, and a specific choice is generally motivated by a specific requirement (such as, for example, the need to obtain estimators with continuous first derivatives). Such requirements have not (yet) been identified in the SSB estimation problem. We will therefore make the very common and conservative choice of a Gaussian kernel. For $\mathbf{x} = (U, \text{SWH})$ this kernel takes the form

$$K\left(\frac{\mathbf{x} - \mathbf{x}_i}{\mathbf{h}_n}\right) = \frac{1}{2\pi h_U h_{\text{SWH}}} \exp\left[\frac{-(U - U_i)^2}{2h_U^2}\right] \cdot \exp\left[\frac{-(\text{SWH} - \text{SWH}_i)^2}{2h_{\text{SWH}}^2}\right] \quad (24)$$

Choice of the bandwidth is more delicate. The bandwidth is clearly a smoothing parameter: A larger bandwidth yields a smoother estimate. Of course, the goal is not to overly smooth the solution but to select a bandwidth that minimizes a chosen measure of the distance between SSB and its estimate φ. Optimal bandwidth selection is generally an iterative process. An initial value of the bandwidth is used to obtain a first estimate of the function to be approximated. Information provided by this estimate is then used in the bandwidth tuning strategy to progressively improve the estimation. Bandwidth selection for nonparametric multivariate regressions is a complex and still very active field of research [e.g., *Vieu*, 1991; *Wand and Jones*, 1994; *Herrmann et al.*, 1995]. We will barely touch it here, since results obtained with a simple first guess of the bandwidth already prove to be very satisfactory (see section 6). This first guess is

$$h_x = C\sigma_x n^{-1/5} \quad (25)$$

where x is either U or SWH, σ_x is the standard deviation of x, and C is a theoretically determined constant that depends on the used kernel function (for a Gaussian kernel, $C = 1.06$). The theoretical basis for (25) is given, for example, by *Simonoff* [1996]. Leaving theory aside, it makes sense to take a bandwidth whose size decreases with the number of measurements (n) and is proportional to their spreading (σ_x).

In this work we will compute series of SSB estimates using data subsets limited to $n = 500$ (see next section). The standard deviations of U and SWH in our global data set being 3.7 m/s and 1.4 m, relation (25) yields $h_U \approx 1$ m/s and $h_{\text{SWH}} \approx 0.4$ m.

5.2. Determination of the Estimation Uncertainty

Nonparametric estimation theory offers asymptotic results concerning the uncertainty of the obtained estimates [e.g., *Härdle*, 1990]. In our case, however, this uncertainty can be evaluated without making use of the rather complex mathematical arsenal of asymptotic theories. The solution is simply to estimate the standard deviation of the different SSB estimates obtained with several data subsets of comparable size. The natural choice would be to perform a SSB estimation for each of the 100 repeat cycles of the global data set. This would require the inversion of a hundred $[(n - 1) \times (n - 1)]$ matrices, where n is typically between 6000 and 7000. This is feasible but represents a prohibitive amount of computer time (several days) on the workstation at our disposal. We thus adopted a purely pragmatic approach and tested the nonparametric estimation method with data sets limited to a reasonable number of crossover points, 500 in this case. The computer time needed to invert a $[500 \times 500]$ matrix is of the order of a minute. The whole SSB estimation procedure then takes a couple of hours. This procedure consists of the following steps:

1. Extract 500 randomly chosen crossover points from each repeat cycle.
2. Compute $\hat{\varphi}$ for each cycle. Then, using (18), compute $\varphi(\mathbf{x})$ over a regular grid of (U, SWH) values. The used grid size is 0.25 m/s in U and 0.25 m in SWH.
3. For each grid point, compute the mean and standard deviation of the individual SSB estimates $(\bar{\varphi}, \sigma_\varphi)$. The mean value $\bar{\varphi}$ is our best SSB estimate. Assuming independence of the individual estimates, the standard deviation of the estimation error on $\bar{\varphi}$ is

$$\sigma_{\bar{\varphi}} = \sigma_\varphi / \sqrt{100} \quad (26)$$

As will be demonstrated in section 6, the nonparametric SSB estimates obtained with such a very limited data subset are already of better quality than those obtained with parametric models fitted on the global data set.

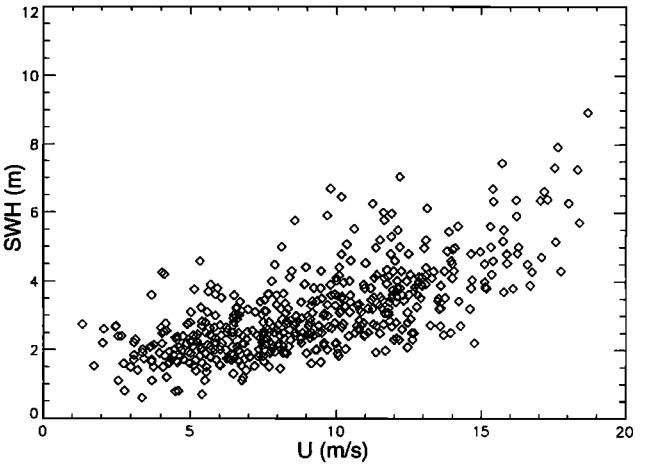

Figure 1. Scatterplot of the (U, significant wave height (SWH)) measurements in a set of 500 randomly chosen crossover points. The plotted values (squares) correspond to the ascending arc measurements ($\mathbf{x}_{1_i}$). Data come from TOPEX repeat cycle 46.

5.3. Imposing the Constraint

Finally, one has to decide at which ascending arc measurement one imposes $\varphi(\mathbf{x}_{11}) = \varphi_0$ and what the value of φ_0 is. The natural choice would be $\varphi(0, 0) = 0$, i.e., the SSB must be zero over a flat sea with no wind. Unfortunately, one never finds a measurement with strictly zero wave height and wind speed. Of course, one could impose $\varphi(\mathbf{x}_{11}) = 0$ for the $\mathbf{x}_{11}$ measurement closest to (0, 0), but this is not appropriate for the following two reasons:

1. The exact location of this $\mathbf{x}_{11}$ is quite variable from one data set to another, since the data density close to (0, 0) is low. Figure 1 shows the typical example of one of our 500-point

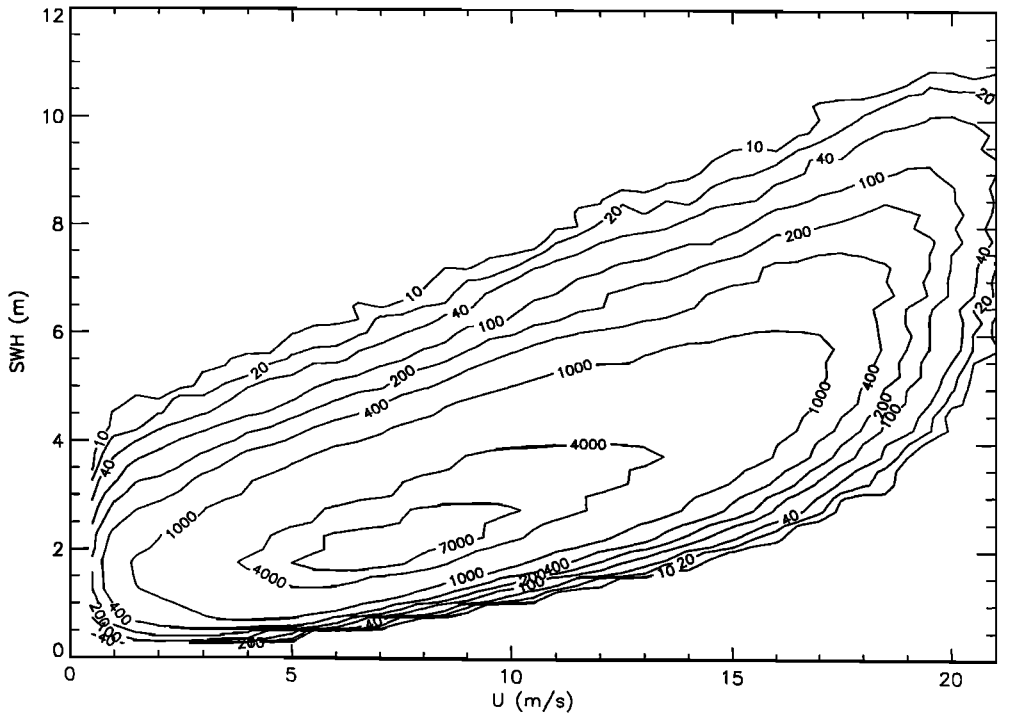

Figure 2. Distribution of altimetric measurements in the (U, SWH) plane. Measurements are binned into (U, SWH) boxes of width (0.5 m/s, 0.25 m). Isolines of the number of data per box are plotted.

data sets in which the $\mathbf{x}_{11}$ measurement closest to (0, 0) corresponds to U = 1.8 m/s and SWH = 1.6 m. With such a choice, estimates of φ obtained for individual cycles would be subject to a constraint applied at quite variable locations in the (U, SWH) plane. The mean $\bar{\varphi}$ would thus be of little significance.

2. Any $\mathbf{x}_{11}$ close to (0, 0) is far away from the mean of the (U, SWH) distribution and hence far from most other $\mathbf{x}_{2i}$. The imposed value of $\varphi(\mathbf{x}_{11})$ has thus little weight in the estimation of $\hat{\varphi}$. In the worst case the weight can be weak enough to make the problem numerically undetermined again.

The above discussion actually shows that selecting the value of $\mathbf{x}_{11}$ closest to the mean of the (U, SWH) distribution is a much better choice: The data density is generally large in the neighborhood of the mean, and a constraint imposed there has significant weight in the estimation of most components of $\hat{\varphi}$. In our global data set, the mean of the (U, SWH) distribution is at (8 m/s, 2.7 m). In all 500-point data sets, there is an $\mathbf{x}_{11}$ for which the distance to this value is less than 0.2 m/s in wind speed and less than 0.2 m in wave height. Numerical experiments actually show that the sensitivity of the final SSB estimate ($\bar{\varphi}$) to the choice of $\mathbf{x}_{11}$ is very small as long as $\mathbf{x}_{11}$ remains in a region of high data density (typically inside the 4000-isoline of Figure 2; see next section).

Having chosen $\mathbf{x}_{11}$, we must still pick a value for φ_0. Ideally, this value should be such that the final solution meets the natural requirement $\varphi(0, 0) = 0$. This is easily obtained in the following way: (1) Compute a solution $\bar{\varphi}$ using any reasonable value of φ_0 (for example, $\varphi_0 = -5$ cm), and (2) determine $\bar{\varphi}(0, 0)$ and subtract this value from the whole solution. The so-shifted solution is the final solution. It obviously satisfies $\varphi(0, 0) = 0$.

6. Results

6.1. Validity Domain of the SSB Estimates

The TOPEX SSB was estimated using the nonparametric technique implemented exactly as described in section 5. Before analyzing the results, it is important to realize that any statistical estimate of the SSB as a function of (U, SWH) is meaningful only in the region of the (U, SWH) plane where altimetric measurements are actually available. The distribution of these measurements is shown in Figure 2. This distribution was computed using both the ascending and descending arc measurements made at the crossover points in our global data set. Measurements were binned into (U, SWH) boxes of width (0.5 m/s, 0.25 m). Isolines of the number of data per box are plotted. Slightly over 95% of the data fall inside the 400-isoline. To help the reader visualize this data-rich region, the boxes containing less than 400 measurements will be systematically shaded in all subsequent figures presenting results in the (U, SWH) plane. Beware that this number of points concerns the global data set. The data subset with which SSB estimation is actually performed contains only a small fraction (about 8%) of the available measurements. The 400-isoline thus roughly corresponds to the 30-isoline for this subset.

6.2. Nonparametric Estimate of the TOPEX SSB and its Uncertainty

Let us first examine the standard deviation ($\sigma_{\bar{\varphi}}$) of the obtained SSB estimate (Figure 3a). The estimation uncertainty is minimum in the vicinity of the mean of the data distribution (8 m/s, 2.7 m), that is, in the region of maximum data density. This minimum is slightly below 1 mm. More generally, $\sigma_{\bar{\varphi}}$ remains below 4 mm in almost all of the nonshaded region, that is, for

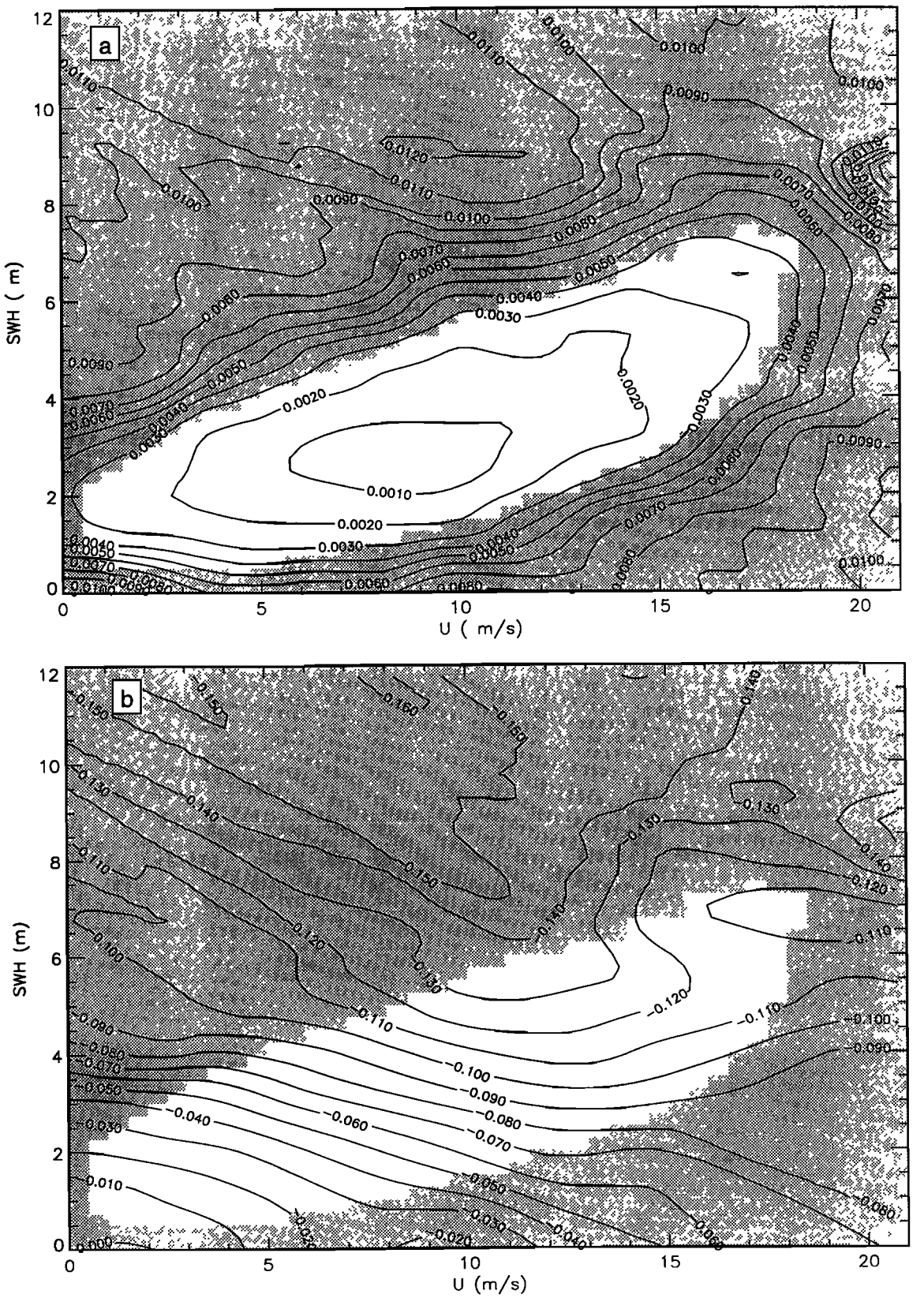

Figure 3. (a) Uncertainty of the nonparametric estimate of the TOPEX sea state bias (SSB) ($\sigma_{\bar{\varphi}}$, in meters) and (b) the corresponding mean estimate ($\bar{\varphi}$, in meters).

over 95% of the observed (U, SWH) values. This is a surprisingly good result, given that we only used 8% of the available data.

The mean estimate ($\bar{\varphi}$) of the SSB is shown in Figure 3b. For comparison, we also plotted (Figures 4a and 4b) the SSB estimates obtained by fitting the BM3 and BM4 models on our global data set:

$$\mathrm{BM3}(U, \mathrm{SWH}) = \mathrm{SWH}[0.0019 - 0.0044U + 0.00019U^2] \quad (27)$$

$$\mathrm{BM4}(U, \mathrm{SWH}) = \mathrm{SWH}[-0.021 - 0.0035U + 0.00014U^2 + 0.0027\,\mathrm{SWH}] \quad (28)$$

The values of the parameter estimates are very close to those previously obtained by *Gaspar et al.* [1994] and *Chelton* [1994].

Examination of the nonparametric solution (hereinafter denoted NP) shows that the nonparametric method has directly captured the main known characteristics of the SSB variability as a function of wind speed and wave height.

1. The estimated SSB is always negative, and its magnitude is essentially an increasing function of SWH. Variations with wind speed are less important. This is why the simple BM1 model (equation (13)) is so successful.

2. For a given SWH the magnitude of the SSB increases with U for wind speeds up to typically 12 m/s and then decreases at higher wind speeds. This is materialized by the downward and then upward slopes of the SSB isolines when going from left to right in the nonshaded region. This behavior was first clearly evidenced by *Hevizi et al.* [1993] and confirmed by *Arnold et al.* [1995]. It motivated use of the BM3 model (equation (12)) in which the relative bias is a quadratic function of U. Still, the comparison of Figures 3b and 4a indicates that the wind-related variability of the SSB, as captured by NP, is more complex than simply quadratic.

3. The rate of change of the SSB with SWH is not simply a constant as hypothesized in the BM1 or BM3 models. It appears to be relatively large for medium wave heights (SWH between 2 and 4 m), slightly weaker for small values of SWH, and markedly weaker for high waves. It even appears to be close to zero for the highest observed waves (the estimated SSB is nearly constant in the upper right corner of the nonshaded zone). This decrease in the magnitude of the relative bias with wave height was first detected by *Witter and Chelton* [1991a] in Geosat data. It was later observed in TOPEX and POSEIDON measurements [*Gaspar et al.*, 1994; *Chelton*, 1994]. The BM4 model (equation (11)) goes some way toward parameterizing this effect, since it allows the relative bias to decrease linearly with SWH. Figure 4b shows that the spacing between the SSB-isolines indeed becomes larger toward high SWH values but not enough to match the shape of the nonparametric solution.

6.3. Skill of the Nonparametric SSB Estimate

The skill of a SSB correction is generally measured in terms of explained variance. That is the reduction in the variance of SSH differences obtained when applying the given correction. As discussed in section 2.2, minimization of $E[\mathrm{SSH}_2 - \mathrm{SSH}_1]^2$ is a criterion that must be taken with care in the context of the calibration of imperfect parametric correction models. Still, good models will reduce this variance, the absolute minimum being reached for the exact model. The explained variance thus remains one legitimate criterion to assess the quality of a SSB model.

The explained variances for the BM3, BM4, and NP models are 9.43, 10.04, and 10.53 cm^2, respectively, again, a good result for the nonparametric solution that has been fitted on only 8% of the available data. Looking at the meridional distribution of the explained variance (Figure 5), one observes, like *Gaspar et al.* [1994], that BM4 does not explain more variance than BM3 at latitudes smaller than 30° but is significantly better at higher latitudes, where usually above-average wave heights and wind speeds are observed. The variance explained by NP is systematically larger than that explained by BM3 or BM4. The gain is modest in low latitudes and more significant at latitudes above 40°. This suggests that the accuracy gain provided by NP is particularly important for relatively high values of U and SWH.

The behavior of the mean crossover differences after SSB correction (often called residuals) is also interesting to analyze. Figure 6 shows the residuals for the NP model plotted as a function of the measured wind speed and wave height differences (ΔU, ΔSWH) at the crossover points. The residuals binned in ΔU boxes remain very small, a few millimeters at most. The residuals sorted according to ΔSWH are somewhat larger but still generally below 0.5 cm. For comparison, we also computed the residuals for the BM3 (equation (27)) and BM4 (equation (28)) models. The results are virtually identical to those shown by *Gaspar et al.* [1994]. For both BM3 and BM4 the residuals are largest when binned according to ΔSWH. For BM3 they reach 2.5 cm and show marked SWH-related variations. They are smaller for BM4 but still exceed 1 cm for $|\Delta\mathrm{SWH}| \geq 4$ m. The NP residuals are thus smaller but still show coherent variations with ΔSWH. This indicates that there is scope for further improving the present NP solution. Some improvements could probably be obtained through a fine tuning of the bandwidth vector. This (difficult) exercise is avoided here since there is little point in tuning a solution that makes use of only a small fraction of the available information. The first improvement shall certainly come from an improved numerical implementation of the NP estimation code. It will allow us to compute NP SSB estimates based on a considerably larger number of measurements. Bandwidth optimization will then be warranted.

6.4. Differences Between NP and the Parametric Models

NP, BM3, and BM4 all capture important variations of the SSB but also exhibit significant differences, plotted in Figure 7. First of all, these differences clearly have a nonzero mean. NP is typically 1 cm below BM3 and 2 cm above BM4. These differences mostly result from the significantly different shapes of the solutions in the vicinity of (0, 0), a data-poor region. Discarding these mean differences, it appears that the best match between NP and the two parametric models is obtained in the region centered roughly on the mean of the (U, SWH) distribution, that is, at (8 m/s, 2.7 m). The differences between NP and BM3 or BM4 exhibit the weakest gradients in this region. On the contrary, the largest variations in the model differences are observed in rough surface conditions ($U > 10$ m/s, SWH > 5 m) or under lighter winds but with SWH values close to the boundary of the nonshaded zone. The reason is the following. The NP estimate is designed to fit the observations over the whole data domain, without any model constraint. On the contrary, the parametric fits are heavily constrained by the numerous measurements made close to the mean of the (U, SWH) distribution. The parametric models are thus tuned to

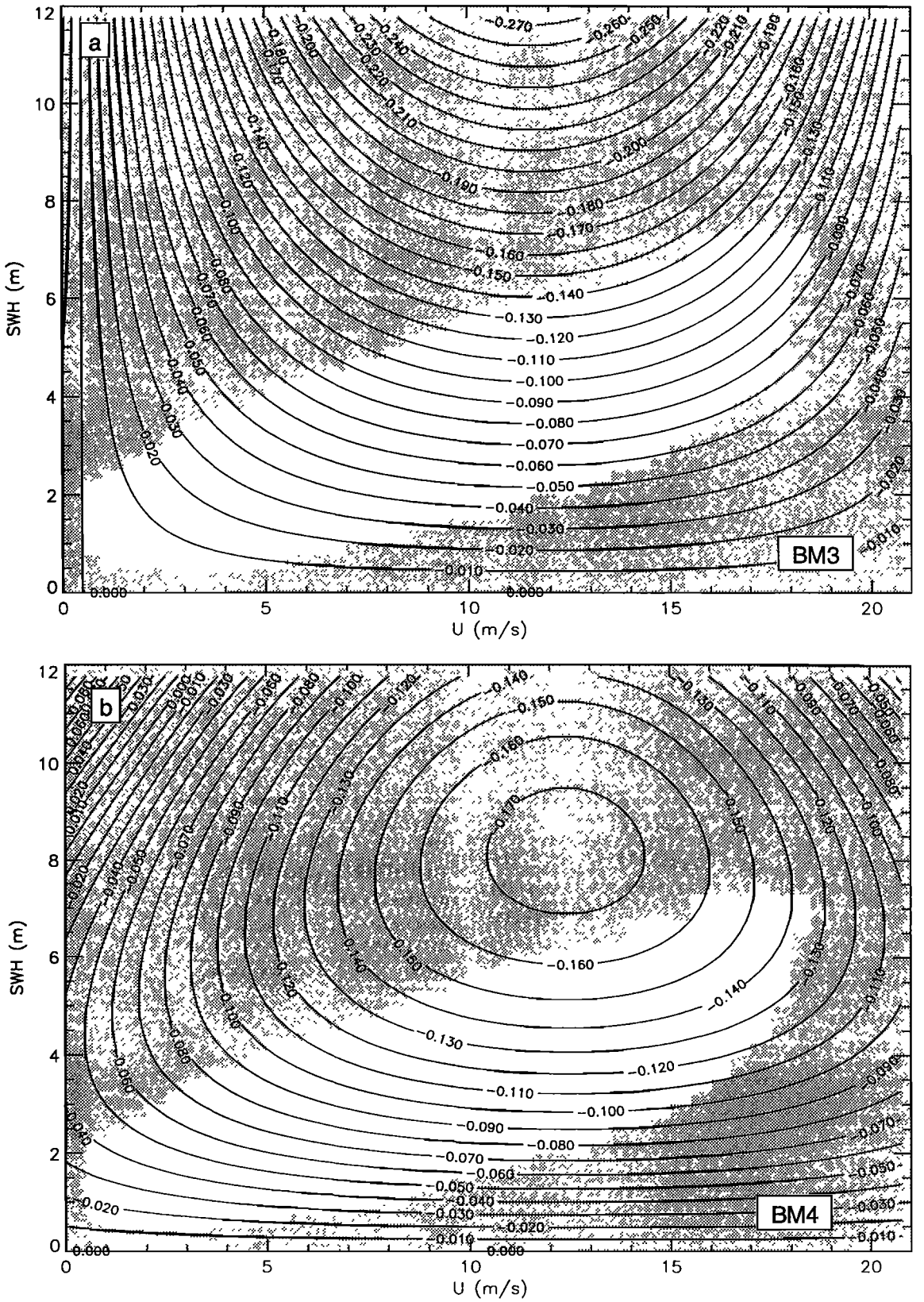

Figure 4. Estimates of the TOPEX Ku-band SSB (in meters) obtained by fitting (a) the BM3 and (b) the BM4 model, on our global crossover data set.

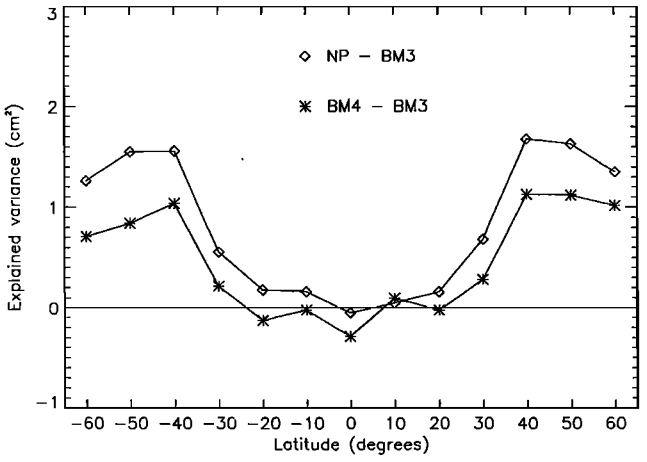

Figure 5. Meridional distribution of the variance explained by the BM4 and nonparametric (NP) models minus the variance explained by the BM3 model. These variances are computed using data binned into 10°-latitude bands.

best match the observations made in this data-rich region. They are close to the NP estimate in this zone, but away from it they become more constrained by their specified functional form than by the few available measurements. The parametric estimates can thus diverge from the NP solution if the specified parametric formulation is not totally adequate. This is typically what we observe in the high-wave/high-wind region. The difference between NP and BM3 varies most quickly, ranging from −1 cm for SWH = 5 m up to 3 cm for SWH = 7 m. In the same range of SWH values the difference between NP and BM4 exhibits somewhat smaller, but still centimetric, variations. These marked differences in the shape of BM3, BM4, and NP explain why the variances explained by these three SSB estimates differ most markedly in high-latitude regions, where high waves and high winds are more often observed. The differences between NP and the two parametric models are better behaved for lower values of U and SWH but can still vary by 1 or 2 cm, especially close to the boundaries of the nonshaded zone. These differences appear to be significant enough to allow NP to explain somewhat more variance than BM3 or BM4 in the calmer low-latitude oceans.

7. Conclusion

A reexamination of the widely used parametric SSB estimation method showed that parametric models calibrated to minimize the variance of SSH differences do not generally minimize the variance of the model error: $E[\mathrm{SSB} - \varphi]^2$. In other words, the so-calibrated models are not true least squares approximations of the SSB. Simulations show that for commonly used parametric models the difference from the true least squares solution is typically a few millimeters to a few centimeters, varying with wave height and wind speed. Such SSB estimation errors inevitably induce spurious gradients in the altimeter-derived SSH between ocean regions with different surface conditions. This problem proves to be the consequence of an imperfect specification of the model's parametric form that affects the model calibration when performed on measurement differences rather than on the measurements themselves.

One way out of this problem is to avoid the specification of the model's parametric form. This led us to develop a new nonparametric formulation of the SSB estimation problem and to propose a general solution for it, based on a kernel smoothing technique. This technique was then used obtain the first fully nonparametric estimate of the TOPEX SSB as a function of both U and SWH. The obtained solution is based on simple, conservative choices for the kernel function and its bandwidth vector.

The basis for the estimation is a global crossover data set including 100 TOPEX repeat cycles. To reduce the computational burden, however, only 8% of the available measurements were actually utilized. Still, the obtained SSB estimate (NP) proves to reduce the variance of the crossover differences more than the estimates obtained by fitting BM3 or BM4 models on the global data set. This is true for the globally averaged variance and also for the zonally averaged variance. The largest variance reduction, compared with BM3 or BM4, is

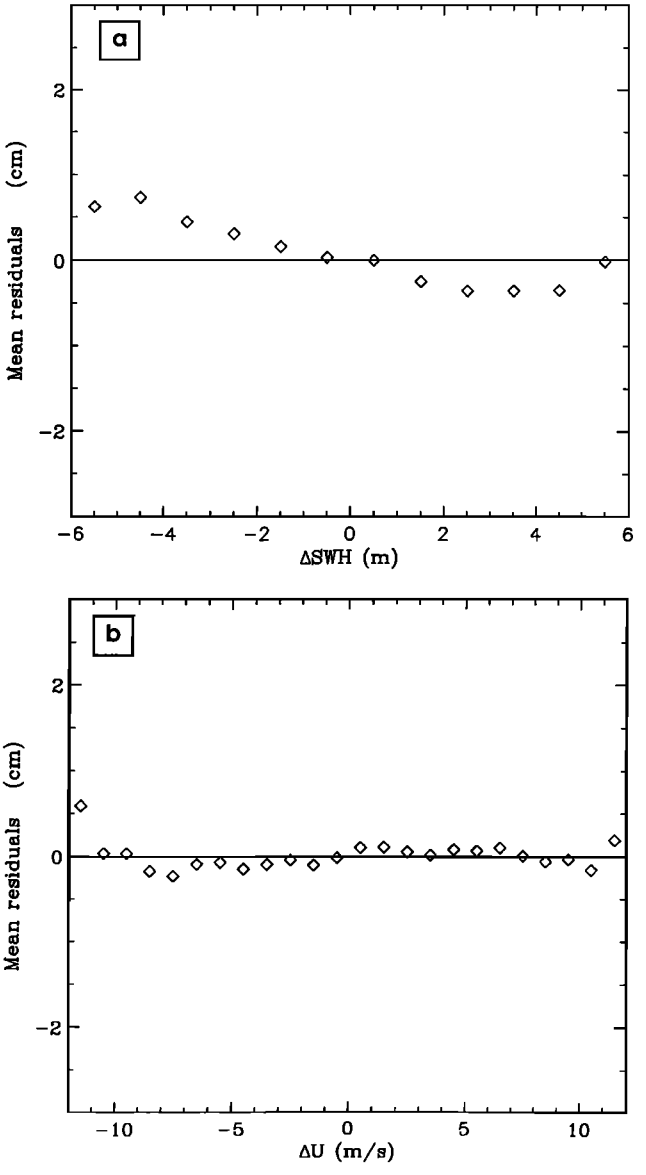

Figure 6. Mean residuals for NP as a function of (a) ΔSWH and (b) ΔU. The squares show averages computed on ΔSWH bins of width 1 m (Figure 6a) and on ΔU bins of width 1 m/s (Figure 6b).

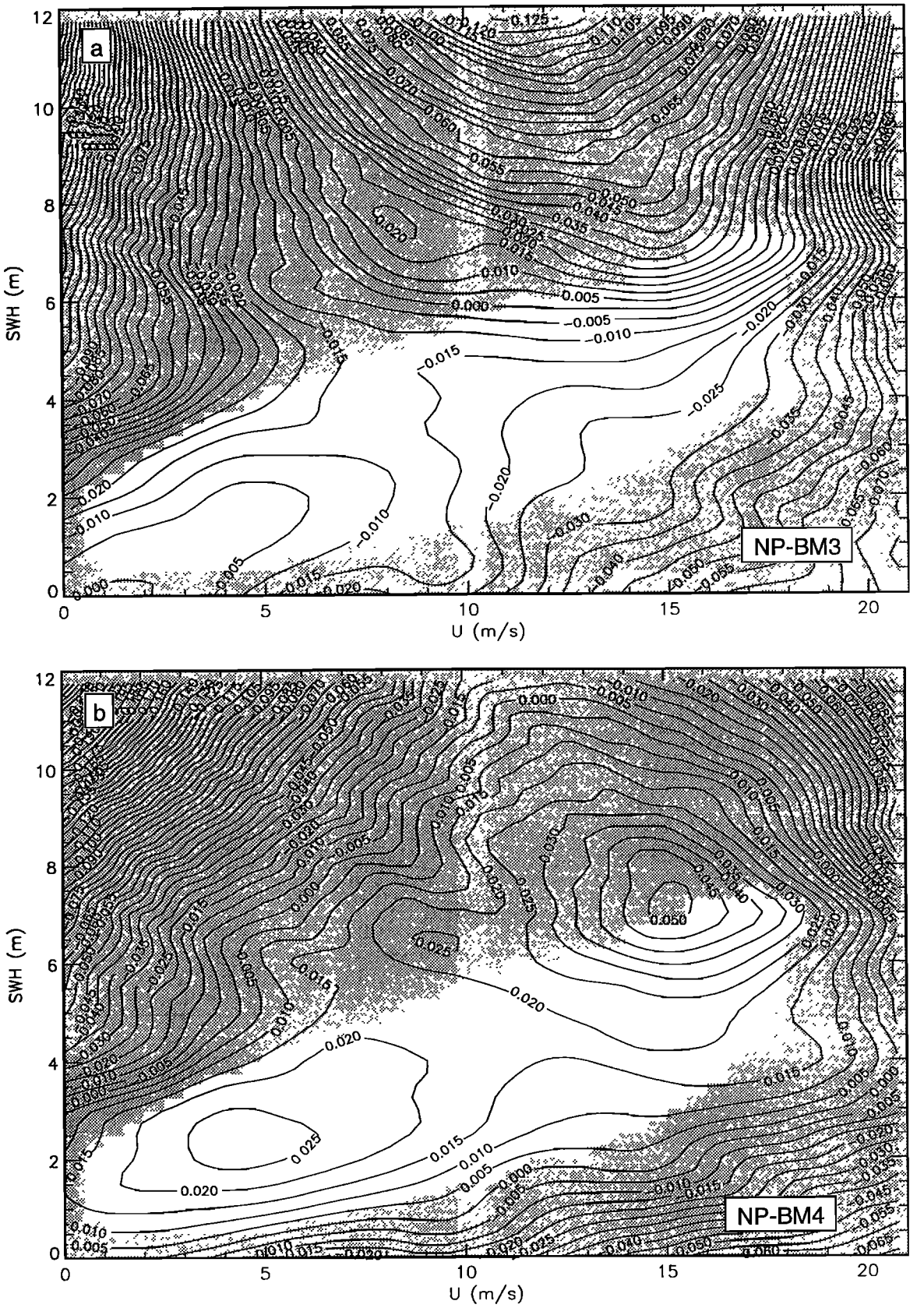

Figure 7. Differences (in meters) between NP and the SSB estimates obtained with (a) BM3 and (b) BM4.

obtained at latitudes above 40°, where high-wave/high-wind conditions are more frequent. It is also in such conditions that the differences between NP and the parametric models are the largest. Furthermore, the mean crossover residuals obtained with NP prove to be significantly smaller than those obtained with BM3 or BM4. They do not generally exceed 0.5 cm but still show SWH-related variations. This indicates that the present NP estimate can still be improved. To this aim, we first intend to improve the numerical implementation of the method. This will allow us to obtain nonparametric SSB estimates based on a larger number of measurements, while keeping the computation time within reasonable limits. In this way, we will be able to take full advantage of the information contained in the enormous, high-quality data set provided by the TOPEX/POSEIDON mission. Then we shall seek further improvement by fine tuning the bandwidth.

Acknowledgments. The authors thank Séverine Rouzaud, Françoise Ogor, and Dominique Bouniol, who developed the nonparametric estimation software and performed the numerical simulations. The authors also thank Bernadette Govaerts from the Catholic University of Louvain (Belgium), who suggested that the two of them collaborate on the subject of this paper. This work was supported by the Centre National d'Etudes Spatiales (CNES) under contracts 95/CNES/0436 and 856/CNES/96/0618. It is part of the activities devoted to the preparation of the forthcoming JASON altimetric mission.

References

Arnold, D. V., W. K. Melville, R. H. Stewart, J. A. Kong, W. C. Keller, and E. Lamarre, Measurements of electromagnetic bias at Ku and C bands, *J. Geophys. Res.*, *100*, 969–980, 1995.

AVISO, AVISO user handbook: Merged TOPEX-POSEIDON products, 3rd ed., *Rep. AVI-NT-02-101-CN*, 196 pp., Cent. Natl. d'Etudes Spatiales, Toulouse, France, 1996.

Barrick, D. E., and B. J. Lipa, Analysis and interpretation of altimeter sea echo, *Adv. Geophys.*, *27*, 61–99, 1985.

Born, G. H., M. A. Richards, and G. W. Rosborough, An empirical determination of the effects of sea state bias on Seasat altimetry, *J. Geophys. Res.*, *87*, 3221–3226, 1982.

Chelton, D. B., The sea state bias in altimeter estimates of sea level from collinear analysis of TOPEX data, *J. Geophys. Res.*, *99*, 24,995–25,008, 1994.

Chelton, D. B., E. J. Walsh, and J. L. MacArthur, Pulse compression and sea level tracking in satellite altimetry, *J. Atmos. Oceanic Technol.*, *6*, 407–438, 1989.

Douglas, B. C., and R. W. Agreen, The sea state correction for GEOS 3 and Seasat satellite altimeter data, *J. Geophys. Res.*, *88*, 1655–1661, 1983.

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

Fu, L.-L., E. J. Christensen, C. A. Yamarone, M. Lefebvre, Y. Ménard, M. Dorrer, and P. Escudier, TOPEX/POSEIDON mission overview, *J. Geophys. Res.*, *99*, 24,369–24,381, 1994.

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.*, *99*, 24,981–24,994, 1994.

Glazman, R. E., A. Greysuckh, and V. Zlotnicki, Evaluating models of sea state bias in satellite altimetry, *J. Geophys. Res.*, *99*, 12,581–12,591, 1994.

Härdle, W., *Applied Non-Parametric Regression*, Cambridge Univ. Press, New York, 1990.

Herrmann, E., M. P. Wand, J. Engel, and T. Gasser, A bandwidth selector for bivariate kernel regression, *J. R. Stat. Soc., Ser. B*, *57*, 171–180, 1995.

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.

Melville, W. K., R. H. Stewart, W. C. Koeller, J. A. Kong, D. V. Arnold, A. T. Jessup, M. R. Loewen, and A. M. Slinn, Measurements of electromagnetic bias in radar altimetry, *J. Geophys. Res.*, *96*, 4915–4924, 1991.

Parzen, E., On the estimation of a probability density function and the mode, *Ann. Math. Stat.*, *33*, 1065–1076, 1962.

Ray, R. D., and C. J. Koblinsky, On the sea state bias of the Geosat altimeter, *J. Atmos. Oceanic Technol.*, *8*, 397–408, 1991.

Rodriguez, E., and J. M. Martin, Estimation of the electromagnetic bias from retracked TOPEX data, *J. Geophys. Res.*, *99*, 24,971–24,979, 1994.

Rosenblatt, M., Remarks on some non-parametric estimates of a density function, *Ann. Math. Stat.*, *27*, 832–837, 1956.

Simonoff, J. S., *Smoothing Methods in Statistics*, 338 pp., Springer-Verlag, New York, 1996.

Vieu, P., Non-parametric regression: Optimal local bandwidth choice, *J. R. Stat. Soc., Ser. B*, *53*, 453–464, 1991.

Wand, M. P., and M. C. Jones, Multivariate plug-in bandwidth selection, *Comput. Stat.*, *9*, 97–116, 1994.

Wand, M. P., and M. C. Jones, *Kernel Smoothing*, 212 pp., Chapman and Hall, New York, 1995.

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, 1991a.

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

Yaplee, B. S., A. Shapiro, D. L. Hammond, B. D. Au, and E. A. Uliana, Nanosecond radar observation of the ocean surface from a stable platform, *IEEE Trans. Geosci. Electron.*, *GE-9*, 170–174, 1971.

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.

J.-P. Florens, GREMAQ-IDEI, Université des Sciences Sociales, Place Anatole France, 31042 Toulouse Cedex, France. (e-mail: florens@cict.fr)

P. Gaspar, Collecté Localisation Satellites (CLS), Space Oceanography Division, 31526 Ramonville, France. (e-mail: gaspar@cls.cnes.fr)

(Received September 11, 1997; revised March 17, 1998; accepted April 6, 1998.)