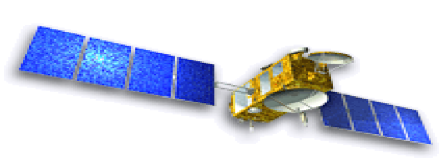
Developing spline-based nonparametric estimation for the altimeter sea state bias (SSB) problem



H. Feng[1], D. Vandemark[1] ,N. Tran[2], and L. Li[3]

[1]Ocean Process Analysis Lab ,University of New Hampshire, USA
[2]CLS, Ramonville StAgne, France
[3] Mathematics Department, University of New Hampshire, USA

I. Abstract

Feng et al. (2010) presents a new nonparametric approach, based on spline (SP) regression, for estimating the satellite altimeter sea state bias (SSB) correction. Model assessment is performed alongside SSB models computed using a local linear kernel (LK) smoothing, which is currently used to build operational altimeter SSB models. Key reasons for introducing this alternative for SSB application include simplicity in accurate model generation, ease in model replication among altimeter research teams, reduced computational requirements, and its suitability for higher dimensional SSB estimation. It is shown that the SP- and LK-based SSB solutions are effectively equivalent within the data dense portion, with an offset below 0.1 mm and rms difference of 1.9 mm for the 2D (i.e. wave height and wind speed) SSB model. Small differences at the 1-5 mm level do exist in the case of low data density, particularly at low wind speed and high sea state. Overall, the SP model appears to more closely follow the bin-averaged SSB estimates.

This poster presentation includes

- **D**evelopment and implementation details for the new SP-based NP SSB models.
- **T**horough assessment of the new SP-based SSB solution against the LK-based SSB solution in 2D [U10, Hs] and 3D [U10, Hs, x] space.
- **S**ummation of SP and LK approach benefits and limits for NP SSB modeling.

II. Overview of non-parametric (NP) SSB estimators

NP SSB estimation model :

$$Y = SSB_{NP}(\mathbf{X}) + \varepsilon$$

ε is a collective error term from various sources, assumed a zero mean noise.
X is a vector of SSB-correlated predictors, such as Hs and U10
Y is a response variable, the sea surface range error information.

Data availability of Y

- SSH (collinear or crossover) difference dataset: SSHD (Gaspar et al., 2002)
- Direct SSH anomaly dataset : SSHA (Vandemark et al.,2002)

Note that the error term ε of SSHD and SSHA are not equivalent, and so the corresponding SSB estimates can exhibit systematic difference due to distinct data sets being used (Labroue et al., 2004).

NP estimators for SSB

- Bin Averaged (BA): (Vandemark et al., 2002)
- Nadaraya-Watson Kernel Smoothing : (Gaspar et al., 1998)
- Local linear Kernel (LK):
  - SSB(U10,Hs) : (Gaspar et al., 2002)
  - 3P SSB(U10,Hs,Tm): (Tran et al., 2010)
- Spline (SP) regression in 2D/3D (Feng et al., 2010)

III.1 Spline-based SSB Regression Model (Ruppert et al., 2003)

The 2D {Hs ,U10} SSB problem:

$$Y = SSB_{\mathrm{NP}}(\boldsymbol{X}) + \varepsilon \qquad SSB_{\mathrm{NP}}(\boldsymbol{X}) = f_1(x_1) + f_2(x_2) + f_3(x_1, x_2)$$

Where f_1 and f_2 are smooth functions of x1 and x2, respectively, and function f_3 accounts for the coupled effect of x1 and x2 on SSB. for a cubic SP basis [1,x, B(x,k1)…], $f_1(x_1)$ and $f_2(x_2)$ can be represented as follows

$$f_1(x_1) = a_0 + a_1 x_1 + \sum_{j=1}^{q_1} a_{j+1} B(x_1, \kappa_{1j})$$
$$f_2(x_2) = b_0 + b_1 x_2 + \sum_{k=1}^{q_2} b_{k+1} B(x_2, \kappa_{2k})$$

where

$$B(x,z) = \left[\left(z-\frac{1}{2}\right)^2 - \frac{1}{12}\right]\left[\left(x-\frac{1}{2}\right)^2 - \frac{1}{12}\right]\Big/4 - \left[\left(|x-z|-\frac{1}{2}\right)^4 - \frac{1}{2}\left(|x-z|-\frac{1}{2}\right)^2 + \frac{7}{240}\right]\Big/24$$

q1 and q2 are the numbers of knots for f_1 and f_2, respectively, and κ refers to these knots. The interaction f_3 is modeled as a linear combination of [x1x2, x1B(x2, κ21), . . . , x1B(x2, κ2q2), x2B(x1, κ11), . . . ,x2B(x1, κ1q1)].
We assume that this cubic SP basis for f_3(x1, x2) leads to a linear SSB model **Y = Xβ + ε,** where the ith row of matrix X is

$$\begin{aligned} X_i = [&1, x_{1i}, B(x_{1i}, \kappa_{11}), \ldots, B(x_{1i}, \kappa_{1q_1}), \\ &x_{2i}, B(x_{2i}, \kappa_{21}), \ldots, B(x_{2i}, \kappa_{2q_2}), \\ &x_{1i}x_{2i}, x_{1i}B(x_{2i}, \kappa_{21}), \ldots, x_{1i}B(x_{2i}, \kappa_{2q_2}), \\ &x_{2i}B(x_{1i}, \kappa_{11}), \ldots, x_{2i}B(x_{1i}, \kappa_{1q_1})]. \end{aligned}$$

where the coefficient vector $\boldsymbol{\beta} = \{a_0, a_1, a_2 \ldots a_{q+3}, b_1, b_2 \ldots b_{q+2}, c_1, c_2, \ldots\}$, q is # of knots, and $\boldsymbol{\beta}$ minimizes $\|\mathbf{Y} - \boldsymbol{R}_X \boldsymbol{\beta}\|^2$
Solution considerations: 1) SP function selection, 2) Number of knots , 3) w or w/o interactive terms

III.3 Optimization of this spline regression

Cubic Spline

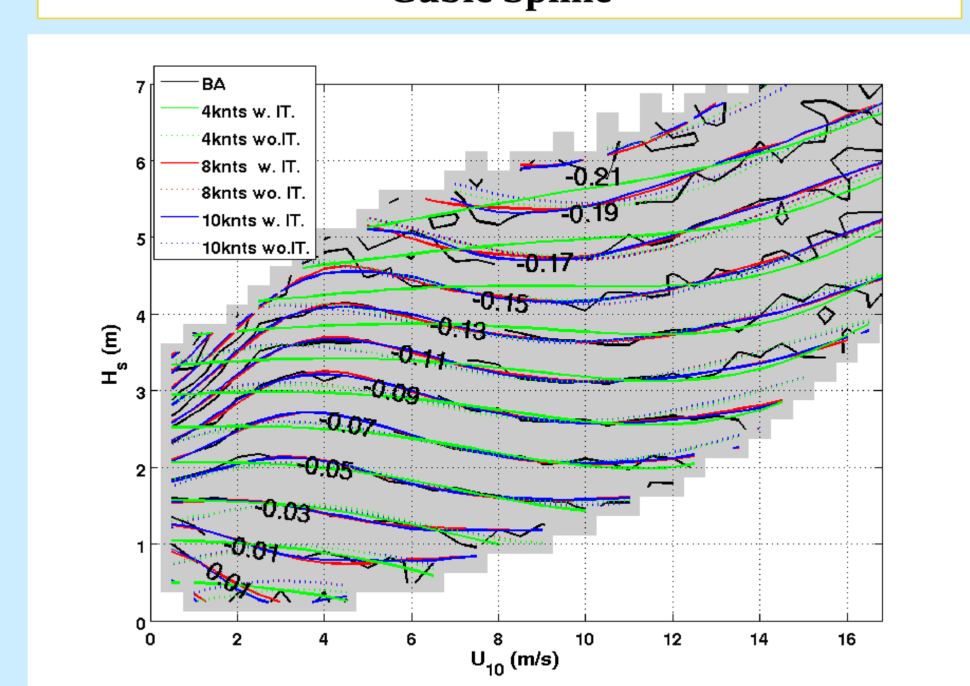


Figure 1 : Impacts of the number of knots and interaction terms for the cubic SP-based SSB estimates (in meters) over the 2-D SSB predictor *{U10,Hs} domain.* Three sets of the color-line contours indicate three distinct sets of the SP-based SSB estimates with respect to the specific number of knots (= 4, *8, 10), and* the corresponding solid and dash of the color-line contours stand for SP-based SSB with and without interaction terms, respectively. As a reference, the BA SSBs are also given by the black contour. Note that the shaded region indicates that at least 100 samples are found within each bin.

III.2. Data used

Y: Jason-1 (J1) Sea Surface Height Anomaly SSHA with all geophysical corrections applied but SSB

X: Hs: Significant wave height (J1)
U10: Wind speed (J1)
T_m: Mean Wave period (WW3)

T_m comes from wave model (WW3, WaveWatch 3) outputs, temporally/spatially interpolated to J1 ground measurement points (Feng et al., 2006). A 1 M pt subset randomly sampled from an annual total 16M data points is used for this work

III.4 Brief Summary of SP-based SSB development

- Require a cubic SP regression at adequately resolve the SSB(**X**)
- Required number of knots needs to exceed 8 if same # is used in all dimensions
- Adding interaction terms in the SP models is critical
- SP fitting computation cost is very low (x10)
- SP fits data (BA) very well

IV.1 2D SSB (Hs,U10) - spline and local linear kernel solutions

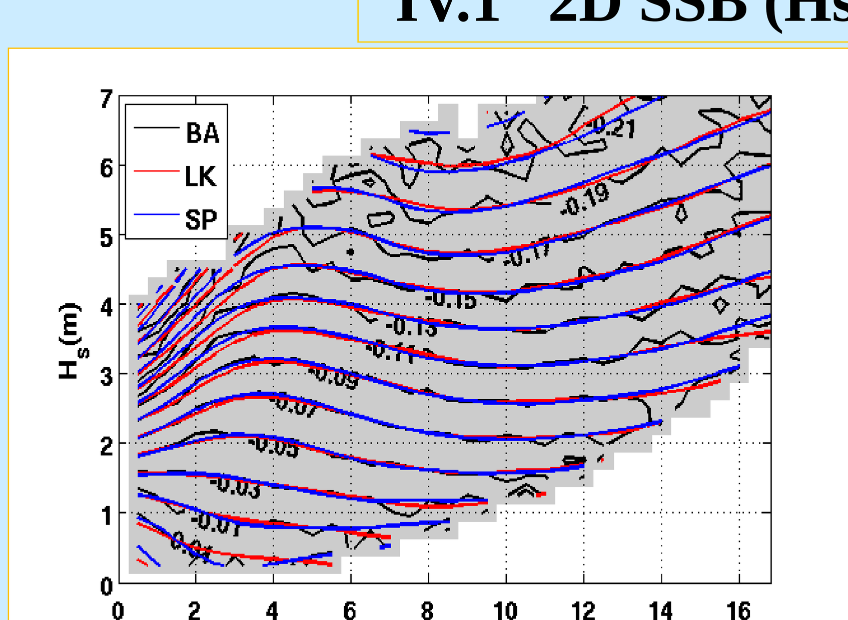


Figure 2: the 2-D SSB {U10,Hs} estimates (in meters) obtained by the SP-based (blue), the LK-based (red), and the simple BA (black)SSB. The SP model is developed with ten knots and interactive terms activated.

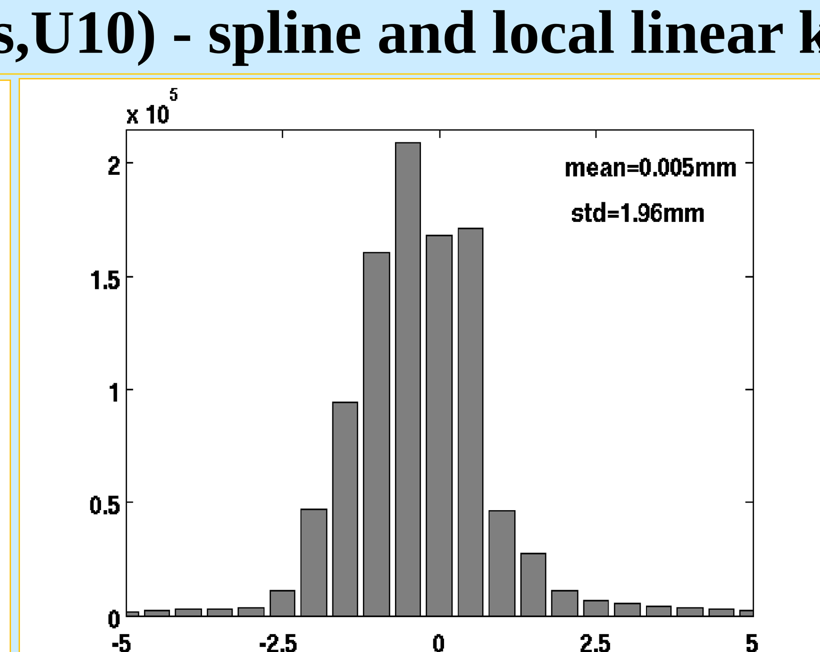


Figure 3: Distribution histogram of the difference (in millimeters) between the SP- and LK-based 2-D SSB {U10,Hs} estimates. The mean and standard deviation are noted as well.

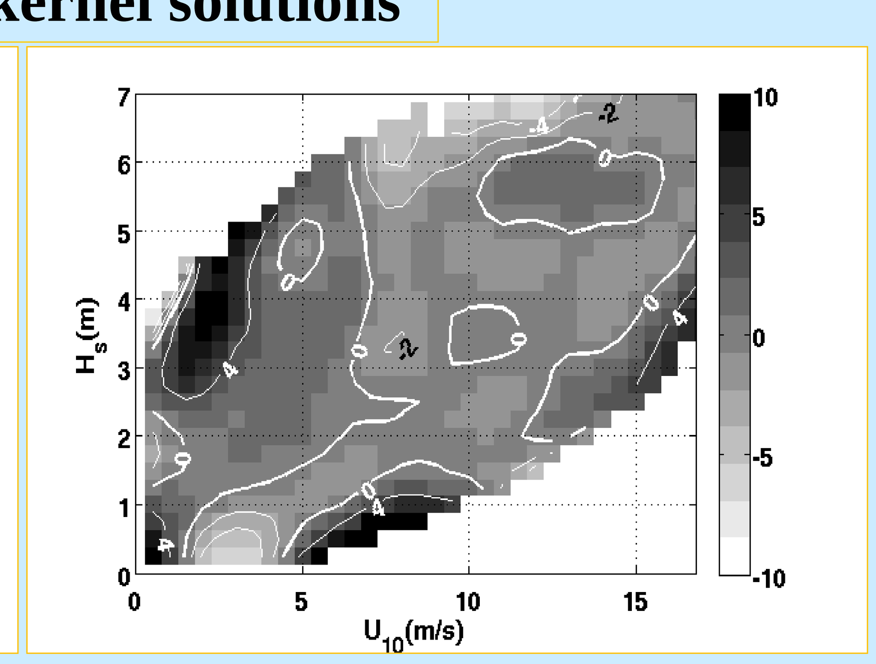


Figure4 : Difference (in mm) between SP- and LK-based SSB estimators over {U10, Hs} domain. Isopleths of (white) 0.00 cm, ±0.20 mm, and ±0.40 mm are indicated. The shaded region indicates >100 samples per bin.

IV.2 3D SSB (Hs,U10, T_m) comparisons

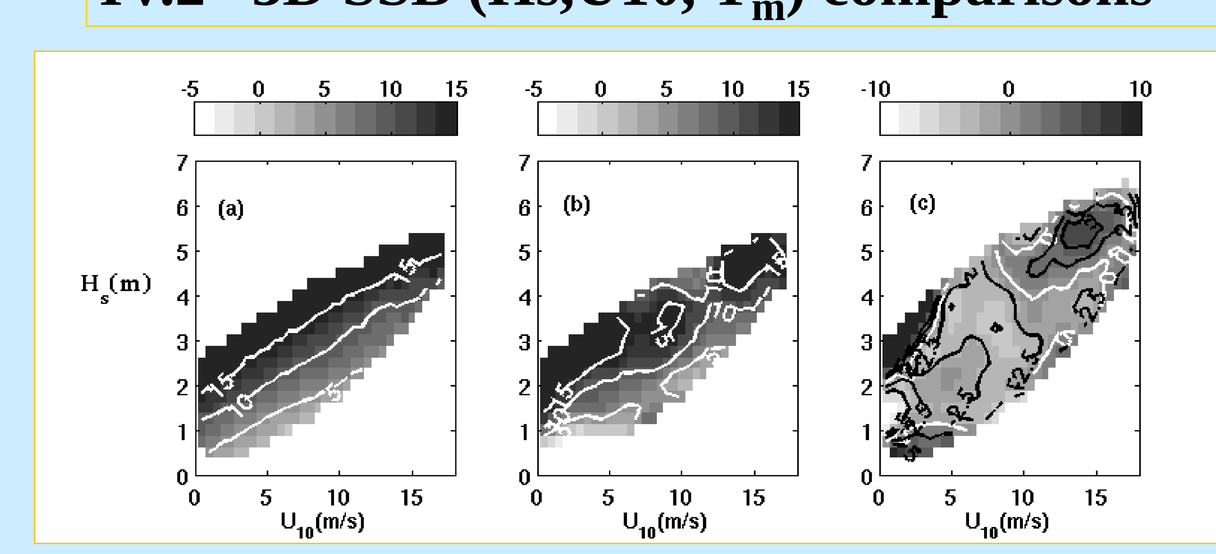


Figure 5: Difference (in millimeters) over the {U10,Hs} domain between (a) the 3-D SP SSB (U10,Hs, Tm = 8 s) and the 3-D SP SSB (U10,Hs, Tm = 7s), (b) the 3-D LK SSB (U10,Hs, Tm = 8s) and the 3-D LK SSB (U10,Hs, Tm = 7 s), and (c) the 3-D SP SSB and the 3-D LK SSB with both having Tm = 8 s. The shaded region indicates that at least 50 samples are found within each bin.

V. Concluding remarks

Spline (SP) regression based SSB models have been developed, and thoroughly evaluated against the altimeter science team standard local linear kernel (LK) SSB model approach in both 2D {Hs, U10} and 3D {Hs, U10, Tm} domains using an identical direct-SSHA Jasonn-1 study. Conclusions are :

- SP-based approach is a capable and computationally-efficient alternative for the NP SSB estimation problem, and is easier to implement and adapt to higher dimension SSB estimation
- The SP regression agrees well with bin-average data over the entire data data space in 2 and 3 dimensional solutions
- The SP and KL-based SSB estimates are effectively equivalent although small systematic differences occur below 2-3mm in the infrequently encountered data conditions.

References

- Feng, H., D. Vandemark, et al. Assessment of wind forcing impact on a global wind-wave model using the TOPEX altimeter, Ocean Engineering, 2006.
- Feng, H S.Yao, L. Li, N. Tran, D. Vandemark etal., Spline-Based Nonparametric Estimation of the Altimeter Sea-State Bias Correction , IGRSL, 7,3, 577-581,2010
- Gaspar et al., Estimation of the sea state bias in radar altimeter measurements of sea level: Results from a new nonparametric method, JGR., 103, 15,803,1998.
- Gaspar et al., Estimating the sea state bias of the TOPEX and POSEIDON altimeters from crossover differences, J. Geophys. Res., 99, 24,981– 24,994,1994
- Gaspar et at., Improving nonparametric estimates of the sea state bias in radar altimetry measurements of sea level, J. Atmos. Oceanic Tech., 19,1690–1707, 2002.
- Labroue, S, et al., Nonparametric estimates of the sea state bias for the Jason-1 radar altimeter, Marine Geodesy, 27: 531-481, 2004.
- Ruppert D. M.P. Wand, and R.J. Carroll, Semiparametric Regression, Cambridge University Press, 2003.
- Tran et al, The sea state bias in altimeter sea level estimates determined by combining wave model and satellite data, 115, C0320, *JGR-Oceans*, 115, 2010.
- Vandemark D. et al., Direct estimation of sea state impacts on radar altimeter sea level measurements, Geophys. Res. Lett., 29(24), 2148, 2002.

Acknowledgments: The research is sponsored by NASA's Science Directorate

Ocean Surface Topography Science Team Meeting

Lisbon, Portugal, 18-22 October, 2010