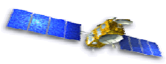
Developing a spline-based nonparametric estimator for the altimeter sea state bias (SSB) problem



H. Feng[1], S. Yao[2], L. Li[2], N. Tran[3], S. Labroue[3], D. Vandemark[1]

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

I. Abstract

In this study, we develop and evaluate an alternative NP SSB model based on the spline regression (SP) method. A direct advantage of the SP NP estimation approach is its quickness of computation speed, particularly in a multi-variable estimation problem.

The sea state bias (SSB), a cm-scale range error induced by ocean surface gravity waves, remains one of the largest sources of uncertainty in altimeter range measurements. Classically, the altimeter SSB has been estimated by a parametric form: SSB=f(x, $\boldsymbol{\theta}$) where $\boldsymbol{\theta}$ is a vector of model parameters and f is a pre-specified function of x, a vector with SSB–related variables, including significant wave height (Hs), surface wind speed (U10) or some combination of the two (Gaspar et al. 1994; Chelton, 1994). More recently, more accurate SSB estimation has been proposed using nonparametric (NP) estimation approaches. A direct advantage of NP estimators is that f is determined implicitly by the data without imposition of a pre-specified function. The NP SSB estimator used to date in the 2D space of {Hs, U10} which are significant wave height and wind speed, respectively, is based on a so-called kernel smoothing (Gaspar et al., 1998). Its improved version, called a local linear kernel (LK) estimator (Gaspar et al., 2002), is now used to provide operational SSB models for Jason-1 and Jason-2. Several issues with this approach pose some limitations. First, the LK approach is computationally expensive. Second, creation of multi-dimensional SSB solutions using more inputs than the standard 2D{Hs, U10} is difficult though recently a 3D SSB has been proposed (Tran et al., this meeting). Finally, the accuracy of LK estimators has not been assessed against other well-known NP estimation approaches. The overall project goal is to develop a computationally-efficient and accuracy-preserved alternative for developing altimetric NP SSB estimators. 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 SSB solution in 2D {U10, Hs} and 3D {U10, Hs, X} space.
- **S**ummarization 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 SSB information contained variable.

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 implies different context, and so the corresponding SSB estimates have some systematic difference due to distinct data 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., this meeting)
- Spline (SP) regression /smoothing model (we are developing in this work)

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

The 2D {Hs ,U10} SSB problem:

$$\begin{aligned} Y_i &= SSB_{NP}(\mathbf{X_i}) + \varepsilon_\mathbf{i} \\ &= f_1(x_{1,i}) + f_2(x_{2,i}) + \varepsilon_\mathbf{i} && \text{(no interactions)} \\ &= f_1(x_{1,i}) + f_2(x_{2,i}) + \text{interactions} + \varepsilon_\mathbf{i} && \text{(with interactions)} \end{aligned}$$

where x1 and x2 indicates Hs and U10, respectively, and Y is SSHA. f_s can be some SP model (linear, quadratic, cubic).

Cubic SP Basis: $\boldsymbol{R_X}(i) = 1, x_{1i}, x_{1i}^2, R(x_{1i}, \kappa_{11}),\ldots, R(x_{1i}, \kappa_{1q}), x_{2i}, x_{2i}^2, R(x_{2i}, \kappa_{21}),\ldots R(x_{2i}, \kappa_{2q}),$

$$\left\{ \begin{array}{l} x_{1i}x_{2i}, \\ x_{1i}R(x_{2i}, \kappa_{21}),\ldots, x_{1i}R(x_{2i}, \kappa_{2q}), \\ x_{2i}R(x_{1i}, \kappa_{21}),\ldots, x_{2i}R(x_{1i}, \kappa_{2q}), \end{array} \right\} \text{(interactive terms)}$$

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

SP Fitting : $\hat{Y} = \boldsymbol{R_X}\,\boldsymbol{\beta}$

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 $\| \mathbf{Y} - \boldsymbol{R_X}\,\boldsymbol{\beta} \|^2$

Development consideration: 1) SP function selection, 2) Number of knots , 3) interactive terms

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)
- sHs: Swell Height (WW3)
- Tm: Wave period (WW3)

sHs and Tm are from wave model (WW3, wavewatch 3) outputs, temporally/spatially interpolated to J1 ground measurement points (Feng et al., 2006). Note there are approximately 16M data points in total for a year (2002), A subset of it randomly sampled to 1.0Million, used for this work

III.3 Results

Quadratic Spline	vs.	Cubic Spline

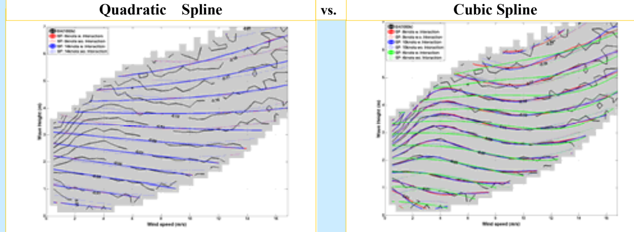


Figure 1. Impacts of number of knots and interactions for Quadratic (left) and Cubic (right) SP-based SSB estimates (in m) in the 2D domain of {U10,Hs}. A color-line contours indicate one set of the SP estimates from a specified number of knots, and the corresponding solid and dash ones stand for SP fitting with and without interaction terms, respectively. As a reference, the bin-averaged (BA) SBB are also given by the black contours

III.4. Brief Summary of SP-based SSB development

- Require a cubic SP regression at least to fit the SSB(**X**)
- Require number of knots > 8 at least
- Adding interaction terms in the SP models is critical
- SP fitting computation cost is very low!
- SP apparently fits to the data (BA) very well

IV. Comparison of Spline (SP) and Local linear Kernel (KL) SSBs

2D SSB (Hs,U10) case

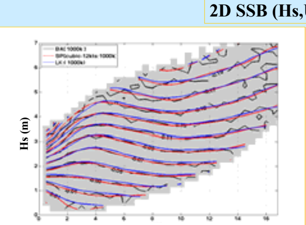


Figure 2a: Comparison of 2D SSB (Hs,U10} estimates: Spline (SP: red), Local linear Kernel (LK, blue) , and Bin-averaged (BA, black). SP-based SSB is developed on 12 knots with full interactive terms considered. The data-shown region at least 100samples)

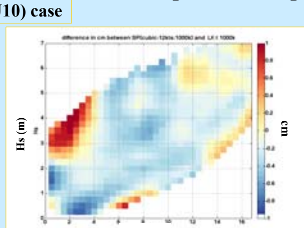


Figure 2b: Difference of 2D SSB (Hs,U10} estimates between Spline and Local linear Kernel (LK) in cm. (SP based SSB developed on 12 knots with full interactive terms considered, and the data-shown region with at least 100samples)

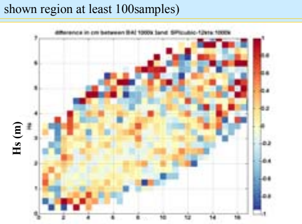


Figure 2c: Difference between bin-average and SP SSB estimates in cm. (SP based SSB developed on 12 knots with full interactive terms considered and the data region with at least 100 samples)

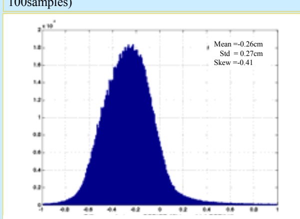


Figure 2d: Distribution of the difference between SP-and LK- based SSB (Hs,U10} estimates in cm. (SP based SSB developed on 12 knots with full interactive terms considered)

3D SSB (Hs, U10, Tm) case

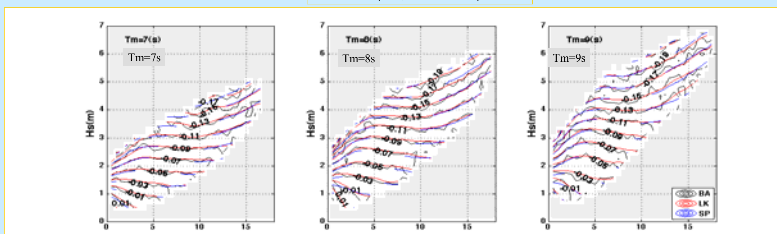


Figure 3. 3D -SSB(U10,Hs,Tm) estimates by SP and KL vs. BA. SSB (Hs, U10) at Tm (wave period) =7second (left), Tm=8second (middle) and Tm=9second (right) (SP based on 10 knots with full interactive terms considered and the data-shown region with at least 80 samples)

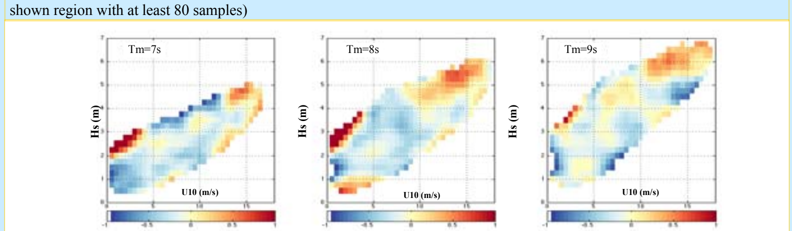


Figure 4. difference of 3D-SSB(U10,Hs,Tm) estimates in cm between SP and KL SSB (Hs, U10) at Tm (wave period) =7second (left), Tm=8second (middle) and Tm=9second (right). (SP based SSB developed on 10 knots with full interactive terms considered and the data-shown region with at least 80 samples)

V. Concluding remarks

Spline (SP) regression based SSB models have been developed, and thoroughly evaluated against widely-used local linear kernel (LK) SSBs in both 2D {Hs, U10} and 3D {Hs, U10, Tm} domains under an identical direct SSHA dataset. Preliminary conclusions are

- SP-based approach is a capable and computationally-efficient alternative for the NP SSB estimation problem, and particularly easier to implement and be adapted to higher dimension SSB estimation
- SP regression apparently fits to the data being used very well (Figures 2a and 2c, and Figure 4)
- The SP and KL-based SSB estimates are nearly equivalent although there is a systematic but small difference within a range of 2-3mm , LK SSB higher than SP SSB in magnitude (Figures 2b,2d and 4) most likely due to distinct NP approaches

References

- 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.
- Feng, H., D. Vandemark, Y. Quilfen, B. Chapron, and B. Beckley. Assessment of wind forcing impact on a global wind-wave model using the TOPEX altimeter, *Ocean Engineering*, 2006
- Gaspar et al., Estimation of the sea state bias in radar altimeter measurements of sea level: Results from a new nonparametric method, J. Geophys. Res., 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.
- Vandemark D., N. Tran, B. Beckley, B. Chapron, and P. Gaspar, 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

Nice, France, 10-12 November, 2008