Estimating Semivariograms to Build Correlation Matrices for *J*

Ralph F. Milliff and Jan Morzel
NWRA/CoRA Boulder, CO

Brian Powell and Dax Matthews
Univ. Hawaii, Honolulu, HI

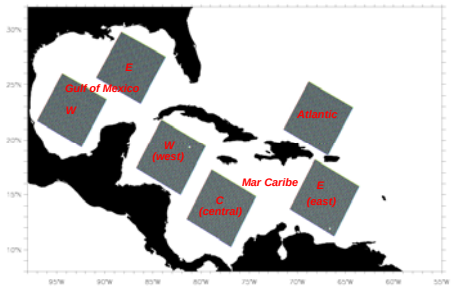
$$J = \mathbf{f} * C_f^{-1} * \mathbf{f} + \mathbf{l} \circ C_l^{-1} \circ \mathbf{l} + \mathbf{b} * C_b^{-1} * \mathbf{b}$$

where C_f is $N_{3D} \times N_{3D}$ for each forcing field, C_l is $N_{2D} \times N_{2D}$ for each state vector, and C_b is $N_{2D} \times N_{2D}$ for each state variable specified on the domain boundary $\partial\Omega$.

- Semivariograms (empirical, theoretical)
- Generalized Diffusion Equation methods

Banerjee et al., 2004: Hierarchical modeling and analysis for spatial data. *Monographs on Statistics and Applied Probability 101*, Chapman and Hall/CRC Press, Boca Raton, FL
Matthews, D. and B. Powell; 2010 in preparation.
Powell et al., 2008: 4DVAR data assimilation in the Intra-Americas Sea with the Regional Ocean Modeling System. *Ocean Modelling*, **23**, 130-145.
Powell et al., 2009: Near real-time assimilation and prediction in the Intra-Americas Seas with the Regional Ocean Modeling System. *Dyn. Atmos. Oceans*, **48**, 46-68.
Banerjee et al., 2004: Hierarchical modeling and analysis for spatial data. *Monographs on Statistics and Applied Probability 101*, Chapman and Hall/CRC Press, Boca Raton, FL
Weaver, A. and P. Courtier, 2001: Correlation modeling on the sphere using generalized diffusion equation. *Q.J.R. Meteorol. Soc.*, **127**, 1815-1846.

Regional Domains for Semivariogram Calculations



32x32 ROMS grid locations for each box

Sep 2003 (hurricane season), Apr 2004; QuikSCAT, NCEP, blended QSCAT-NCEP

Semivariogram Estimate for spatial data

$$\gamma(\mathbf{h}) = \frac{1}{N} \sum_{i=1}^{N} [\eta(x_i + h) - \eta(x_i)]^2$$

Relation to Spatial Covariance

$$\begin{aligned} 2\gamma(\mathbf{h}) &= \mathrm{Var}(\eta(\mathbf{x} + \mathbf{h}) - \eta(\mathbf{x})) \\ &= \mathrm{Var}(\eta(\mathbf{x} + \mathbf{h})) + \mathrm{Var}(\eta(\mathbf{x})) - 2\mathrm{Cov}(\eta(\mathbf{x} + \mathbf{h}), \eta(\mathbf{x})) \\ &= C(0) + C(0) - 2C(\mathbf{h}) \\ &= 2[C(0) - C(\mathbf{h})] \end{aligned}$$

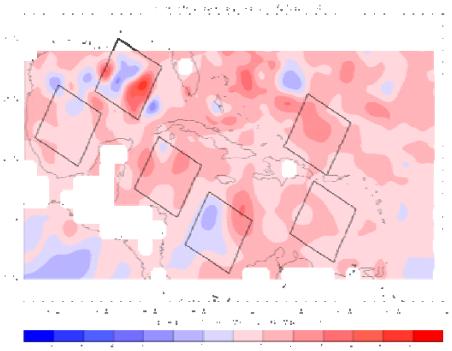
$$\gamma(\mathbf{h}) = C(0) - C(\mathbf{h})$$

Correlation Model Fit (exponential model)

$$C(\mathbf{h}) = \sigma^2 [1 - \exp(-\frac{\mathbf{h}}{\delta})]$$

where σ^2 is the sill and δ is the range

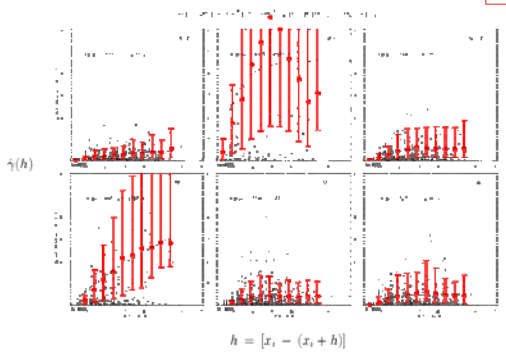
Monthly Average SSH Anomaly from AVISO *(September 2003)*



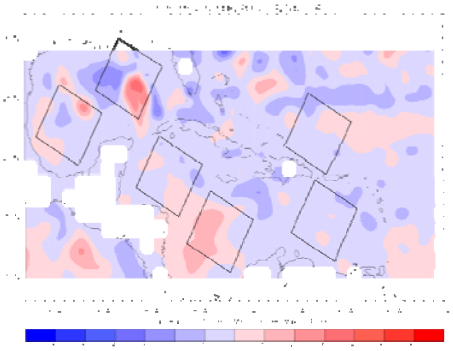
σ^2 *"sill"*

θ *"range"*

SSH; September 2003

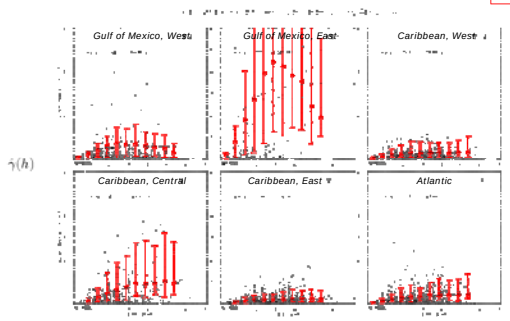


Monthly Average SSH Anomaly from AVISO *(April 2004)*



SSH; April 2004

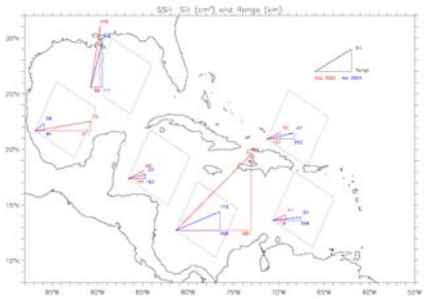
σ^2 "sill"
θ "range"



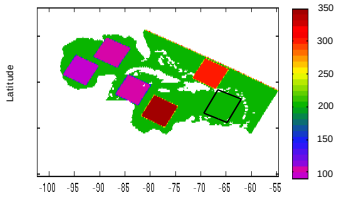
$$h = [x_i - (x_i + h)]$$

Range and Sill Parameters from Semivariogram Analysis

SSH Anomalies; Sep 2003 and Apr 2004



Distribute Covariance Parameters: *Laplacian Smoother (Range Example)*

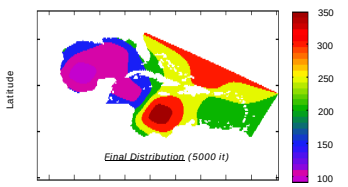


$$\theta^n = \nabla^2\theta^{n-1} + r\theta^n$$

$$r = \begin{cases} 1 & \text{if } \theta_j^n \le \theta_j^{n_0} - \varepsilon \\ 0 & \text{otherwise} \end{cases}$$

$$\varepsilon = 11.5\ \text{km}$$

θ^{n_0} is the initial distribution, θ_j, $j = 1, 2, \ldots, 6$ are the center values for each sub-domain.



Generalized Diffusion Equation Methods (Weaver and Courtier, 2001)

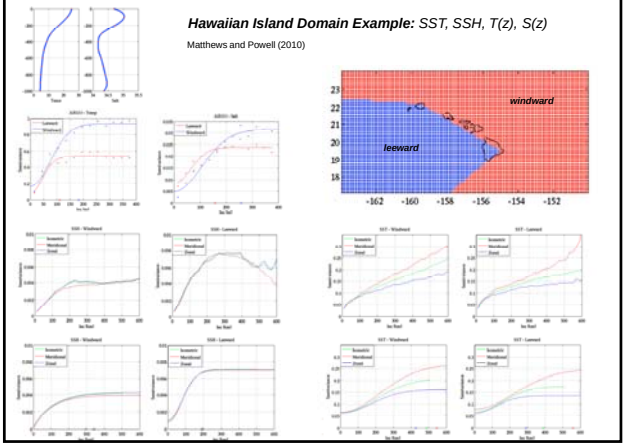
for a field variable η, model the correlation at each point in the field as:

$$\eta(t_{AF}) = \mathbf{L}_F\,\eta(t_0)$$

where

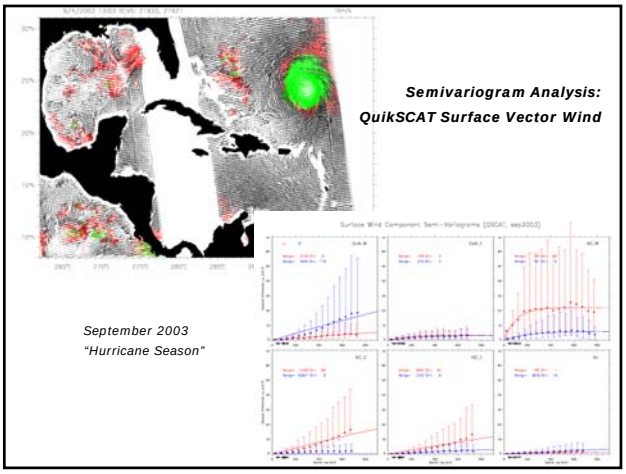
$$\mathbf{L}_F = \left[\mathbf{I} - \sum_{p=1}^{P} \kappa_p \Delta t \{-\mathbf{D}\}^p\right]^{AF}$$

$$\eta(t_m) = \eta(t_{m-1}) - \sum_{p=1}^{P} \kappa_p \Delta t \{-\nabla\}^p \eta(t_{m-1})$$



Hawaiian Island Domain Example: *SST, SSH, T(z), S(z)*

Matthews and Powell (2010)



Semivariogram Analysis:
QuikSCAT Surface Vector Wind

September 2003
"Hurricane Season"

EXTRA SLIDES