Identifying Loci for Ocean Forecast Model Error from Ensembles of Surface Winds based on QuikSCAT and COAMPS

Polly Smith[1], Ralph F. Milliff[2,3], Andrew M. Moore[1] and Christopher A. Edwards[1]

[1] Ocean Sciences Department, University of California, Santa Cruz
[2] CoRA Division, NorthWest Research Associates, Boulder, CO
[3] Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO

Milliff, R.F., A. Bonazzi, C.K. Wikle, N. Pinardi and L.M. Berliner, 2011: Ocean Ensemble Forecasting, Part I: Ensemble Mediterranean Winds from a Bayesian Hierarchical Model., *Quarterly Journal of the Royal Meteorological Society*, **137**, Part B, 858-878, *doi: 10.1002/qj.767*.

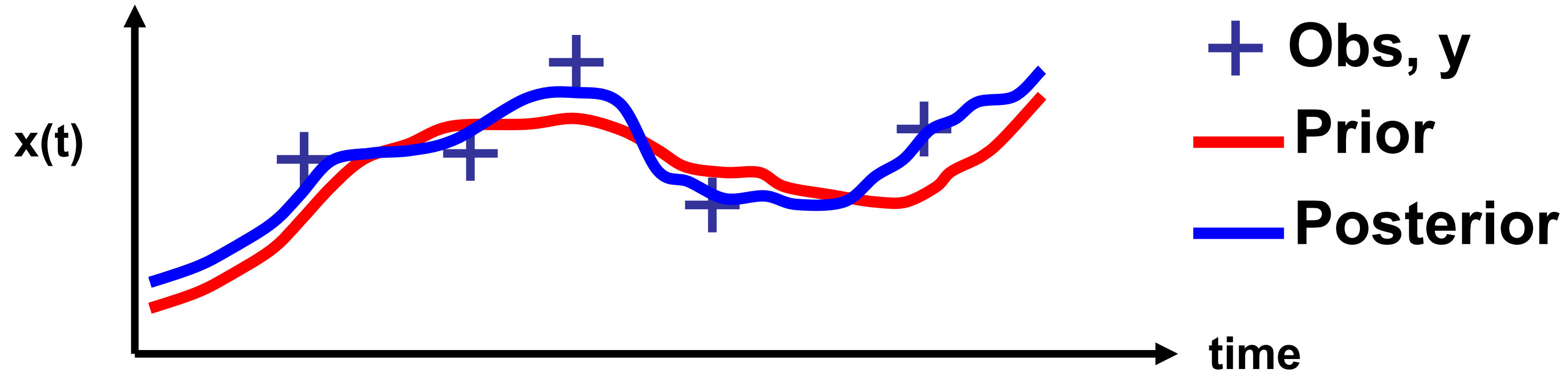
Moore, A.M., H.G. Arango, G. Broquet, B.S. Powell, J. Zavala-Garay and A.T. Weaver, 2011: The Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilation systems. Part I: System overview and formulation. *Progress in Oceanography*, **91**, 34-49.

Moore, A.M., H.G. Arango, G. Broquet, C. Edwards, M. Veneziani, B. Powell. D. Foley, J. Doyle, D. Costa and P. Robinson, 2011: The Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilation systems. Part II: Performance and application to the California Current System. *Progress in Oceanography*, **91**, 50-73.

Moore, A.M., H.G. Arango, G. Broquet, C. Edwards, M. Veneziani, B. Powell. D. Foley, J. Doyle, D. Costa and P. Robinson, 2011: The Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilation systems. Part III: Observation impact and observation sensitivity in the California Current System. *Progress Oceanography*, **91**, 74-94.

Ocean Data Assimilation (DA); Four-Dimensional Variational (4DVar) Method (in words):

Goal of DA, given state vector $\mathbf{x}(t-1)$, *is to adjust model trajectory by incorporating observations* $\mathbf{y}(t)$, *such that new estimate of* $\mathbf{x}(t)$ *is "optimal".*
So, $\mathbf{x}_a = \mathbf{x}_b + \mathbf{Kd}$ *where* $\mathbf{x}_b$ *is the background state (or prior),* $\mathbf{x}_a$ *is the new state (or posterior) and* $\mathbf{d} = (\mathbf{y} - \mathbf{H}\mathbf{x}_b)$ *are the innovations due to observations.*

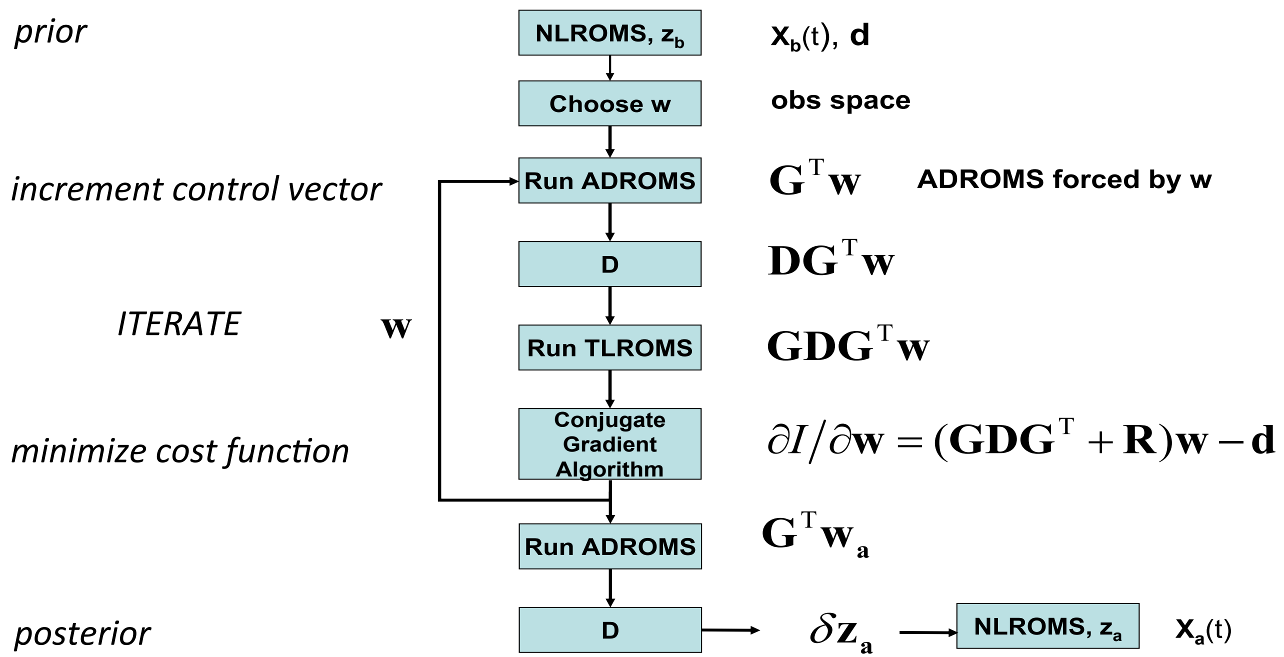


4DVar is a DA method wherein $\mathbf{Kd}$ *is expressed as a control vector of increments:*
$\delta\mathbf{z} = [\delta\mathbf{x}(0),\ \delta\mathbf{f}(t_i),\ \delta\mathbf{b}(t_i),\ \delta\eta(t_i)]^T$ *for increments in initial conditions* $\delta\mathbf{x}(0)$, *forcing* $\delta\mathbf{f}(t_i)$, *boundary conditions* $\delta\mathbf{b}(t_i)$ *and model error* $\delta\eta(t_i)$. *The model error increments are very hard to estimate and are neglected in "strong constraint" 4DVar.*

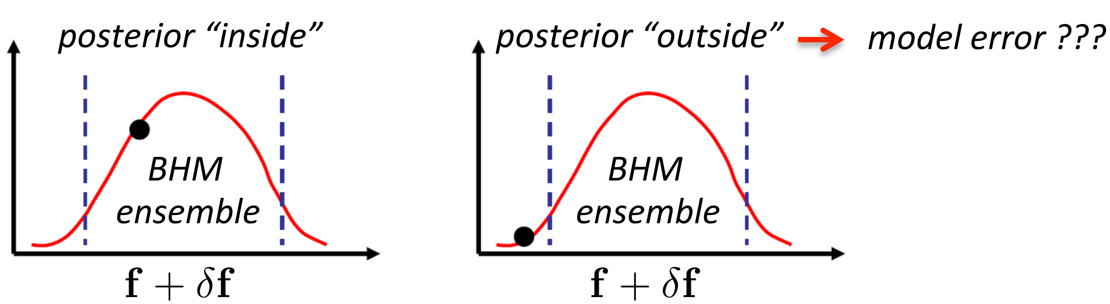
The posterior is "maximized" by minimizing a cost function. Minimization is achieved iteratively by incrementing the control vector. We will compare increments in the surface wind stress part of the control vector $\delta\mathbf{f}$ *with ensemble winds from a*

Bayesian Hierarchical Model (BHM)

Hypothesis: In a strong constraint 4DVar, places where the wind stress control vector iterations are driven outside a BHM wind ensemble are candidate loci for model error.



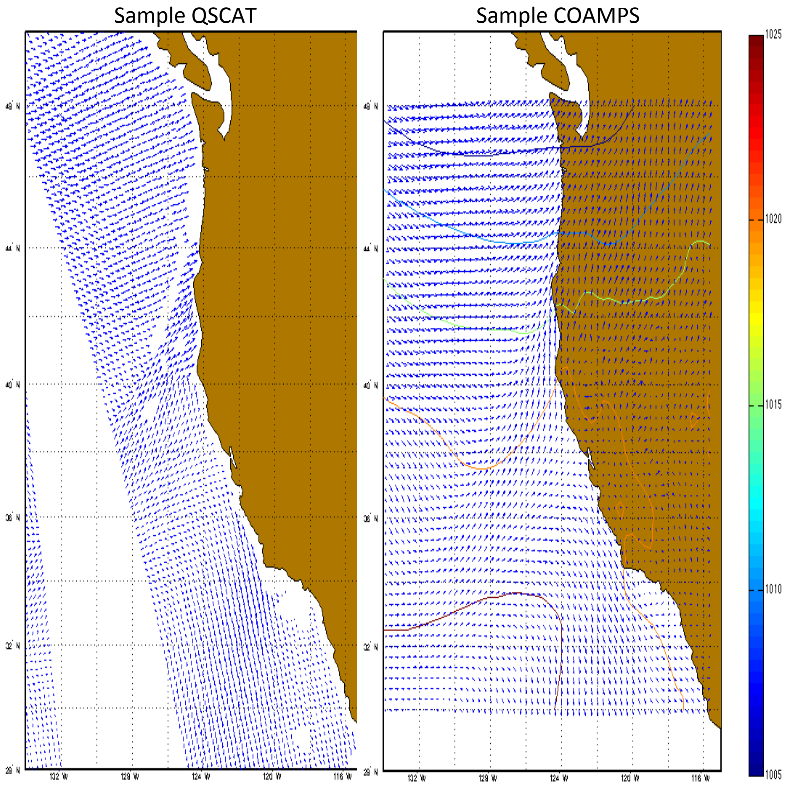
Schematic:



Ensemble Winds from a Bayesian Hierarchical Model (BHM)

$[X,\vartheta_p,\vartheta_d|D] \propto [\,D|X,\vartheta_d\,]\,[X|\vartheta_p]\,[\vartheta_p]\,[\vartheta_d]$

posterior *likelihood* *prior* *parameters*



Rayleigh Friction Equations

$$\frac{\partial u}{\partial t} - fv = -\frac{1}{\rho_o}\frac{\partial p}{\partial x} - \gamma u$$

$$\frac{\partial v}{\partial t} + fu = -\frac{1}{\rho_o}\frac{\partial p}{\partial y} - \gamma v$$

EOF expansion SLP anomaly

$$p(x,y,t) = \mu + \sum_{k=1}^{m} \mathrm{a}_k(t)\,\phi_k(x,y).$$

Geostrophic-ageostrophic approximation

$$u = -\frac{f}{\rho_o(f^2+\gamma^2)}\frac{\partial p}{\partial y} - \frac{\gamma}{\rho_o(f^2+\gamma^2)}\frac{\partial p}{\partial x}$$

$$v = \frac{f}{\rho_o(f^2+\gamma^2)}\frac{\partial p}{\partial x} - \frac{\gamma}{\rho_o(f^2+\gamma^2)}\frac{\partial p}{\partial y}$$

Stochastic Model

$$\mathbf{U}_t = a_{1,1}\mathbf{D}_y\mathbf{P}_t + a_{1,2}\mathbf{D}_x\mathbf{P}_t + \boldsymbol{\epsilon}_{u,t};$$

$$\mathbf{V}_t = b_{1,1}\mathbf{D}_x\mathbf{P}_t + b_{1,2}\mathbf{D}_y\mathbf{P}_t + \boldsymbol{\epsilon}_{v,t},$$

$$P_t(x,y) = \mu + \sum_{k=1}^{N} \alpha_{k,t}\,\phi_k(x,y).$$

Random variables, uncertainty model

$$a_{1,1} \sim N\left(-\frac{f}{\rho_o(f^2+\gamma^2)}, \sigma^2_{a_{11}}\right)$$

$$a_{1,2} \sim N\left(-\frac{\gamma}{\rho_o(f^2+\gamma^2)}, \sigma^2_{a_{12}}\right)$$

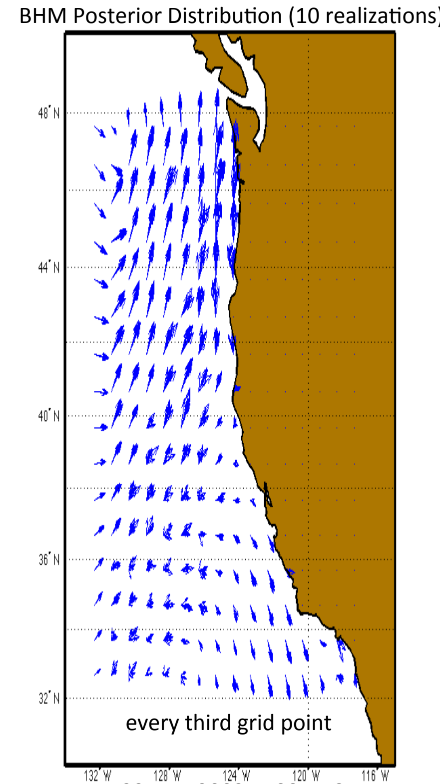
$$b_{1,1} \sim N\left(\frac{f}{\rho_o(f^2+\gamma^2)}, \sigma^2_{b_{11}}\right)$$

$$b_{1,2} \sim N\left(-\frac{\gamma}{\rho_o(f^2+\gamma^2)}, \sigma^2_{b_{12}}\right)$$

$$\alpha_{k,t} \sim N(0, \sigma^2_P).$$

$$\boldsymbol{\epsilon}_{u,t} = \sum_{k=1}^{n_\beta} \boldsymbol{W}_k \boldsymbol{\beta}^u_{k,t} + \tilde{\boldsymbol{\epsilon}}_{u,t}$$

$$\tilde{\boldsymbol{\epsilon}}_{u,t} \sim N(\mathbf{0}, \sigma^2_u\mathbf{I}).$$



Data Stage (likelihood)

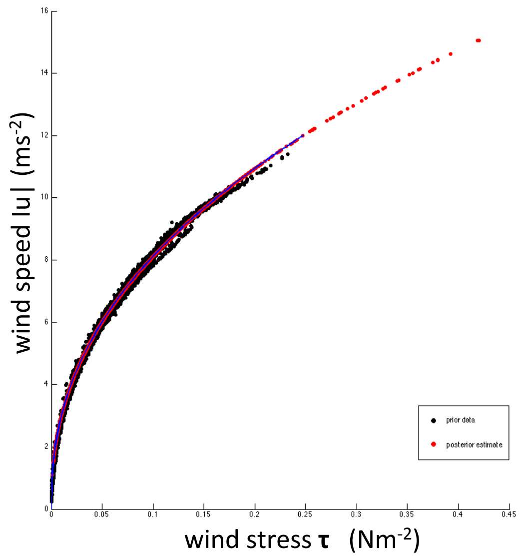
Process Model Stage (prior)

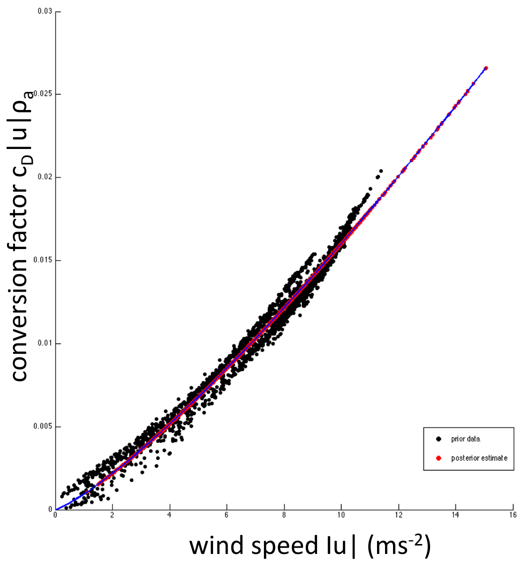
Posterior

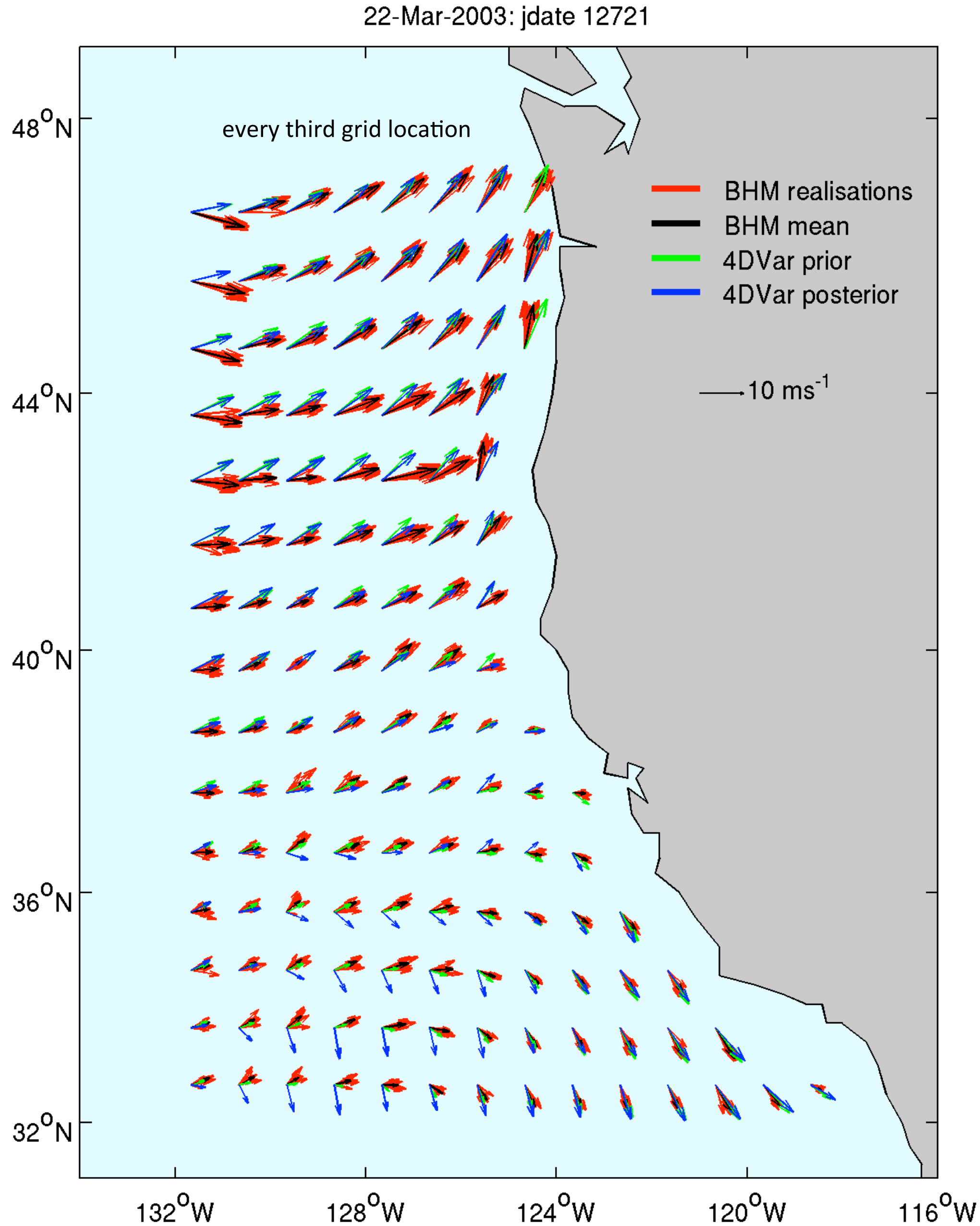
Comparing 4dVar Increments with BHM

Estimate u,v from τ every day

1. find |u| as function of τ
2. find c_D|u|ρ_a as function of |u|
3. estimate u,v from posterior $\tau_{x,y}$ / c_D|u|ρ_a



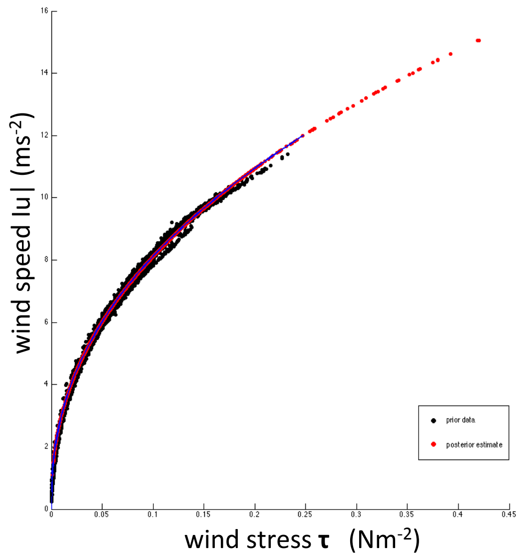


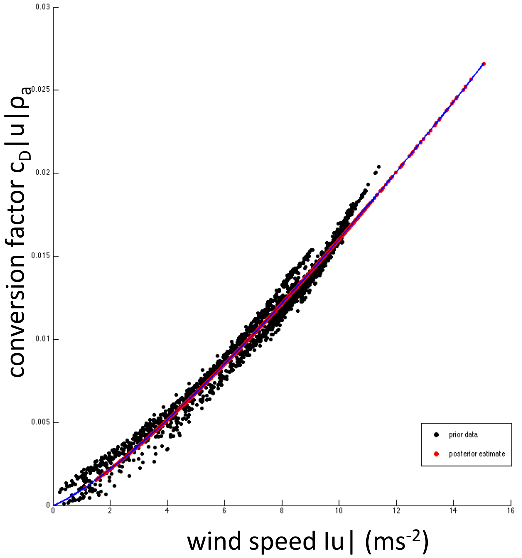


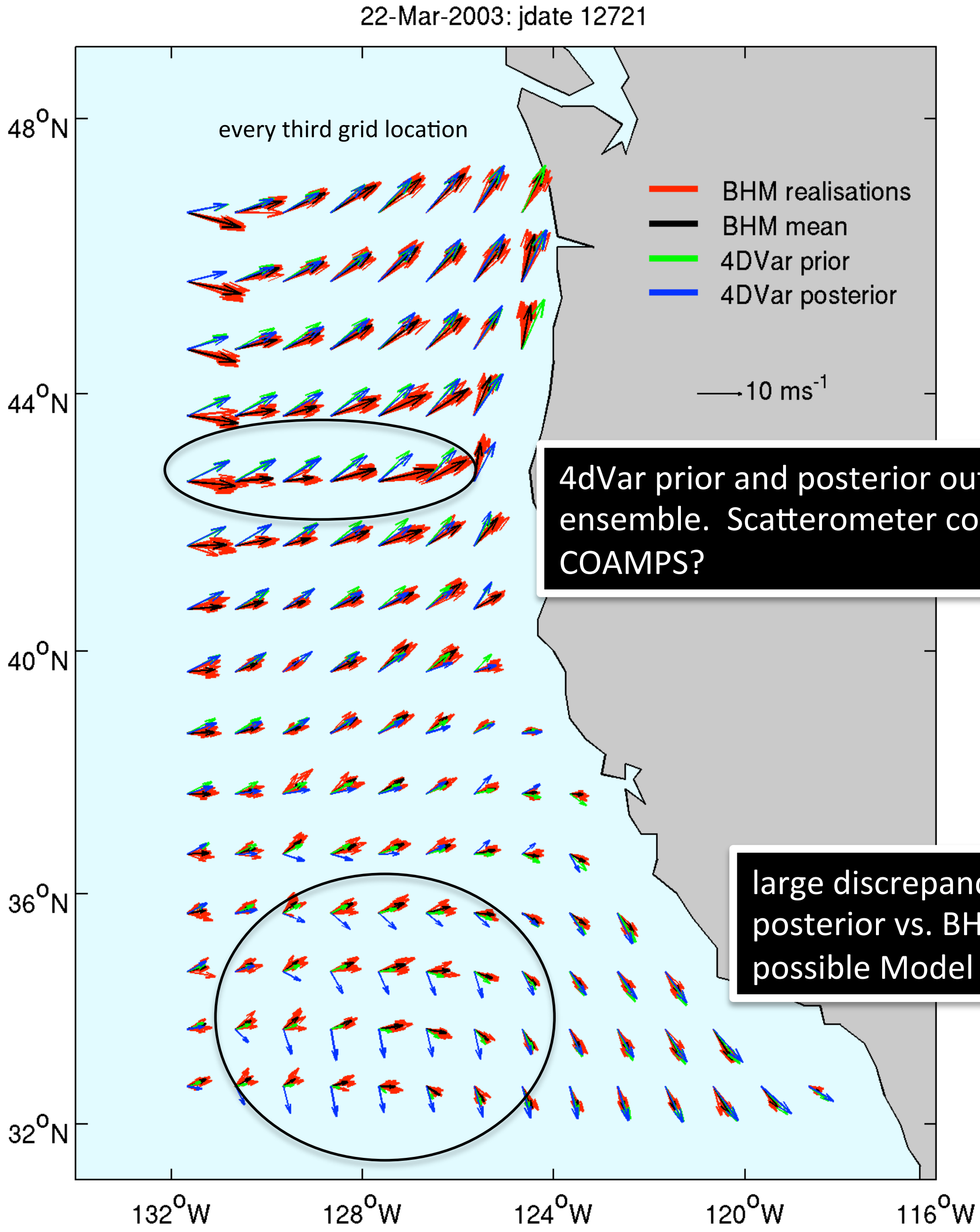
Comparing 4dVar Increments with BHM

Estimate u,v from τ every day

1. find |u| as function of τ
2. find $c_D|u|\rho_a$ as function of |u|
3. estimate u,v from posterior $\tau_{x,y} / c_D|u|\rho_a$



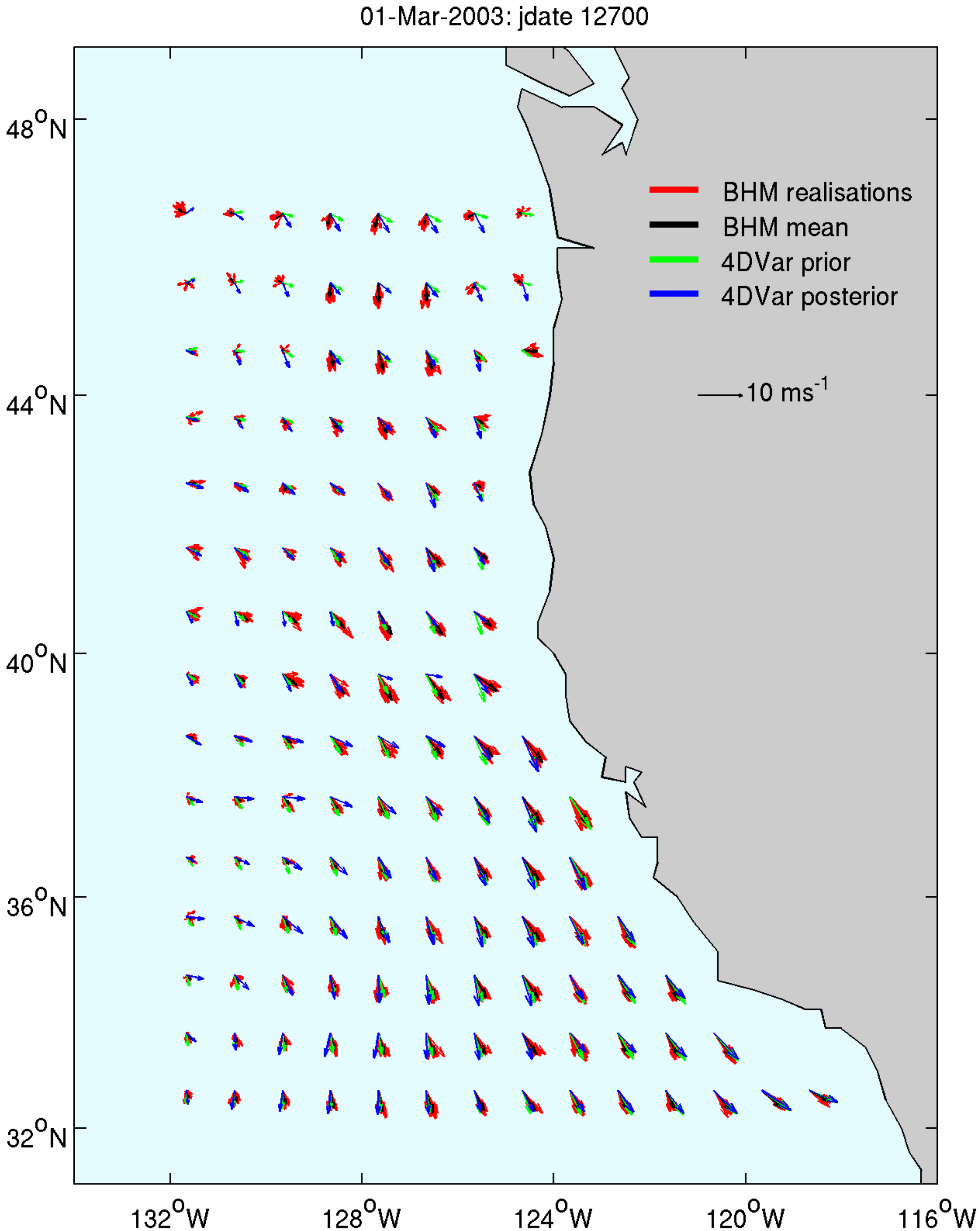




4dVar prior and posterior outside BHM ensemble. Scatterometer correction to COAMPS?

large discrepancies, 4dVar posterior vs. BHM ensemble; possible Model Error

Model Error as a dynamical function of space, time



Summary

- ❑ Ocean forecast model error estimates are important for assessing forecast accuracies and improving model and observation array designs (strong vs. weak constraint in 4dVar).

- ❑ Model error is hard to identify.

- ❑ Precisely specified error properties (uncertainty) in QuikSCAT winds have been used to help identify regions of probable model error in a state-of-the-art ocean forecast system.

- ❑ The probability distributions from a surface wind BHM (based on QuikSCAT and COAMPS) have quantitative value (see also Pinardi et al., 2011).

Work in Progress

- ✧ Wind stress distributions can be summarized directly from surface wind BHM posterior

- ✧ Other forcing control variable fluxes require BHMs as well (e.g. heat, fresh water)

- ✧ Model error dynamics can be modeled given data stages of the kinds suggested here.

EXTRAS

Temporal Autocorrelation Comparison: BHM vs. COAMPS (all of 2003)

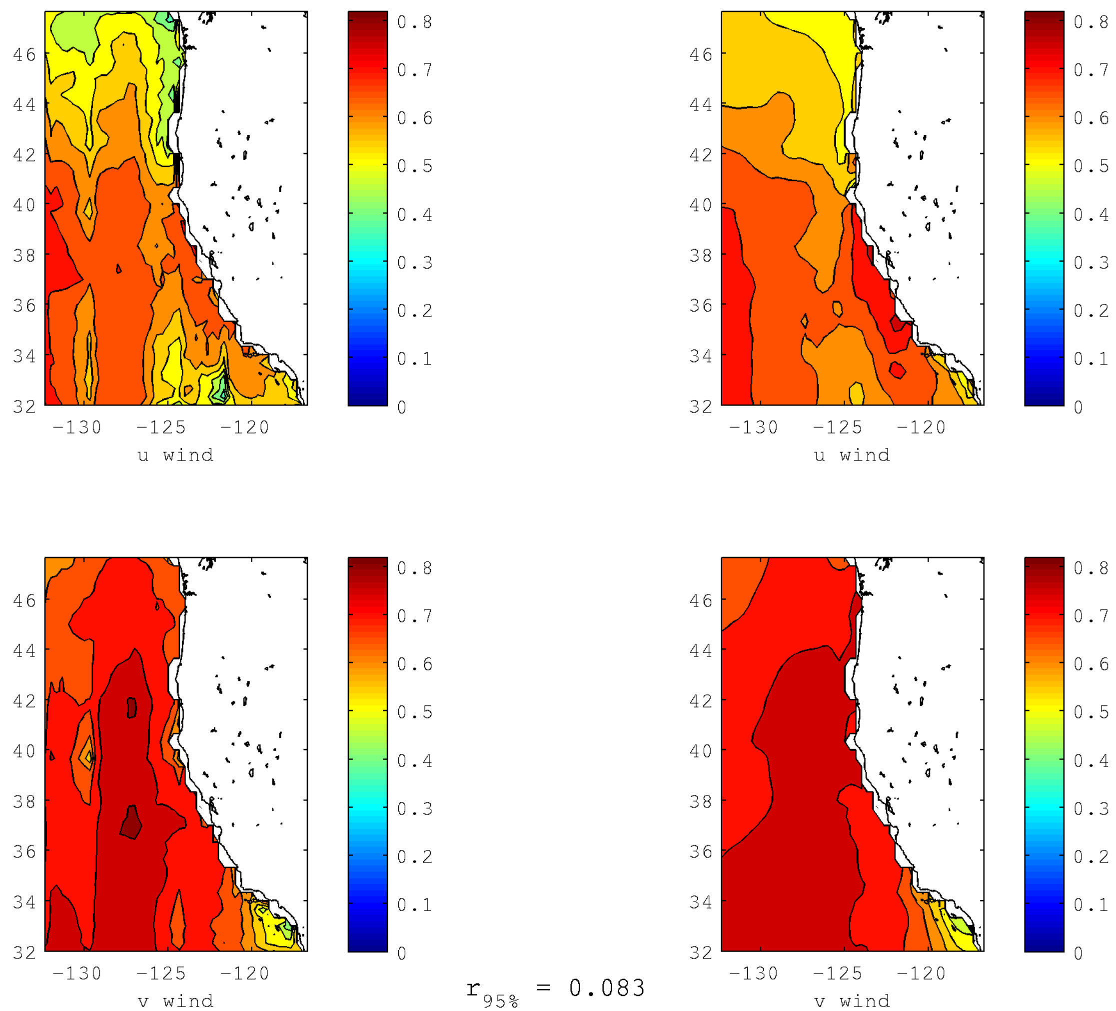
AR-1 model for temporal autocorrelation: $U(t;x,y) = \mathrm{r}_1(x,y)U(t-1;x,y) + \epsilon(t-1)$

$\mathrm{r}_{\mathbf{1}}(x,y)$

BHM ***COAMPS***

u wind

v wind



$r_{95\%} = 0.083$