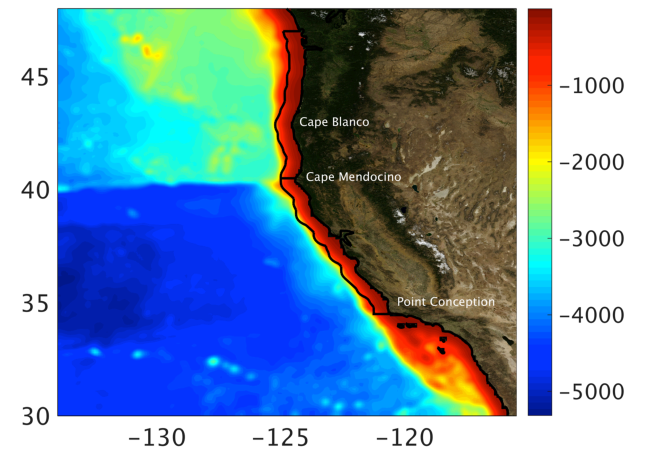

Weak Constraint 4-Dimensional Variational Data Assimilation in a model of the California Current System

William J. Crawford[1], Andrew M. Moore[1], Polly Smith[2], Ralph Milliff[3], Jérôme Fiechter[1], Christopher Wikle[4] & Christopher A. Edwards[1]

[1]Department of Ocean Sciences, University of California, Santa Cruz, CA [2]Department of Mathematics and Statistics, University of Reading, UK
[3]Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO [4]Department of Statistics, University of Missouri, Columbia, MO

Introduction

- We present a methodology for estimating the model error covariance matrix for use in the Weak Constraint formulation of the Regional Ocean Modeling System (ROMS) 4-dimensional variational (4D-Var) data assimilation system.
- Numerical models have intrinsic error due to discretization, parameterization, etc.
- Intrinsic model error is often overlooked due to a lack of understanding regarding the size and nature of the errors.
- We estimate the model error covariance, $\mathbf{Q}$, using a combination of twin-model experiments, linear inverse modeling and Bayesian Hierarchical Modeling.
- Model configuration:
  - Computed using the ROMS 4D-Var[1] data assimilation system
  - Domain: 30-48ºN, 116-135ºW & 42 vertical levels
  - 1/10º horizontal resolution (~10km)
  - Surface forcing from COAMPS[2]
  - Boundary conditions from SODA[3]
  - 8-day consecutive overlapping assimilation windows

Hypothesis and Experimental Setup

◆ **Imagine two ocean forecasts:**

1) A perfect forecast subject to error free forcing - $\frac{d\mathbf{x}_t}{dt} = \mathcal{N}(\mathbf{x}_t) + \mathbf{w}_t(t)$
2) An imperfect forecast subject to error free forcing - $\frac{d\mathbf{x}}{dt} = \mathcal{N}(\mathbf{x}) + \mathbf{w}_t(t) + \epsilon(t)$

The evolution of the error in the forecast can be approximated as:

$$\frac{d\Delta\mathbf{x}}{dt} = \mathcal{N}(\mathbf{x}) - \mathcal{N}(\mathbf{x}_t) + \epsilon(t) \simeq \mathbf{N}\Delta\mathbf{x} + \epsilon(t)$$

◆ **Application**

During strong-constraint 4D-Var, increments are added to the surface forcing. If the forcing is perfect, the increments are assumed to be due to model error

We can then run two separate non-linear model integrations:

1) Using the prior surface forcing yielding $\mathbf{x}_1$
2) Using the 4D-Var corrected forcing yielding $\mathbf{x}_2$

We hypothesize that the differences, $\delta\mathbf{x} = \mathbf{x}_2 - \mathbf{x}_1$, that arise due the differences in forcing will mimic the influence the model error, ϵ, will have on $\Delta\mathbf{x}$

◆ **Estimating $\mathbf{Q}$ in a Linear Inverse Model.**

If we can model $\delta\mathbf{x}$ as a 1st order Markov process, then the evolution of $\delta\mathbf{x}$ can be estimated in a linear inverse model as:

$$\delta\mathbf{x}(t_{i+1}) = \mathbf{A}\delta\mathbf{x}(t_i) + \boldsymbol{\xi}(t_i)$$

From this model we can solve for $\boldsymbol{\xi}$ and compute the model error covariance as:

$$\mathbf{Q} = E\{\boldsymbol{\xi}\boldsymbol{\xi}^T\}/dt$$

◆ **Bayesian Hierarchical Model[4] (BHM) to refine twin model experiment**

Our surface forcing product is NOT perfect as assumed in theory

More accurate to use BHM to decide where we can trust 4D-Var adjusted winds

Use a mix of model prior winds and 4D-Var adjusted winds for a new non-linear run to compare with COAMPS prior case

Compare weak constraint model run using the Bayesian Hierarchical Model to compute $\mathbf{Q}$ with a solution computed without (see **Results**)

Twin model setup

- COAMPS[2] used as prior wind stress in strong constraint 4D-Var calculations where increments are added to produce the best analysis
- We compute two sequence of non-linear ROMS integrations spanning 2003-2004:

  1) Model initialized with the 4D-Var background circulation for each 8 day assimilation window, and integrated forward in time subject to the COAMPS background surface forcing

  2) Model initialize with the same background circulation as in (1), but in this case the strong constraint 4D-Var analysis estimates of the wind forcing were used to integrate the model forward in time
- The background fluxes of heat and freshwater were constrained the same in both sequences

Linear Inverse Model & Estimation of Model Error Covariance

- Following von Storch (1995)[5], the linear propagator matrix, $\mathbf{A}$, is computed as:

$$\mathbf{A} = \mathbf{C}_1\mathbf{C}_0^{-1}$$

- Not practical, however, to compute $\mathbf{A}$ directly since $\mathbf{C}_1$ and $\mathbf{C}_0^{-1}$ contain ~10^6 elements
- We estimate $\mathbf{A}$ instead using the 50 leading 3D multivariate EOFs of $\delta\mathbf{x}$ where $\mathbf{C}_1$ and $\mathbf{C}_0^{-1}$ are lag-0 and lag-1 covariance matrices of the PC time series[5,6]
- We can then calculate the stochastic forcing as:

$$\boldsymbol{\xi}(t_i) = (\delta\mathbf{x}(t_{i+1}) - \mathbf{A}\delta\mathbf{x}(t_i))$$

- Estimate of $\mathbf{Q}$ highlights areas where the model is known to perform poorly[7] (Figure 1)

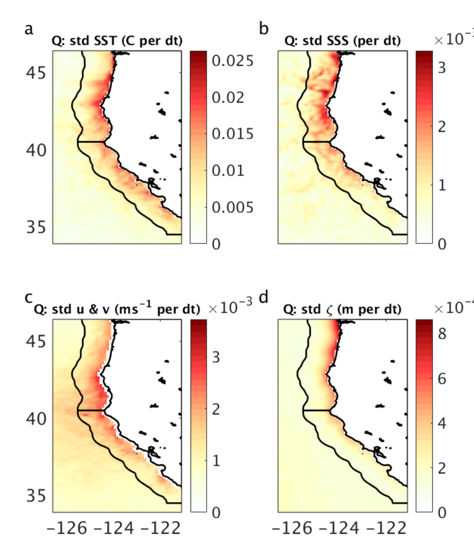

*Figure 1: Standard deviations of **Q***

Bayesian Hierarchical Model

- We can never know the surface forcing exactly
- Utilize a Bayesian Hierarchical Model to determine where 4D-Var has made adjustments that can be trusted
- If the 4D-Var winds lie within 2σ of the mean of the BHM wind realizations, the 4D-Var wind is considered trustworthy, and not due to intrinsic model error
- On a point-by-point basis, if 4D-Var wind is trusted, replace it with the prior COAMPS wind
- Compute new non-linear model integration using a set of wind forcing containing a mix of unadjusted and 4D-Var adjusted COAMPS winds
- Compute new model differences, EOFs and estimate of $\mathbf{Q}$

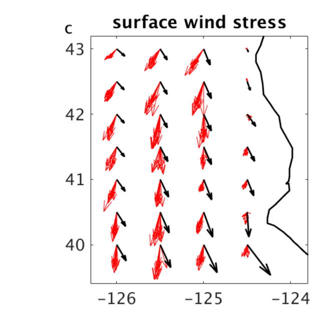

Figure 2: 10 realizations of BHM winds (red) and 4D-Var winds (black)

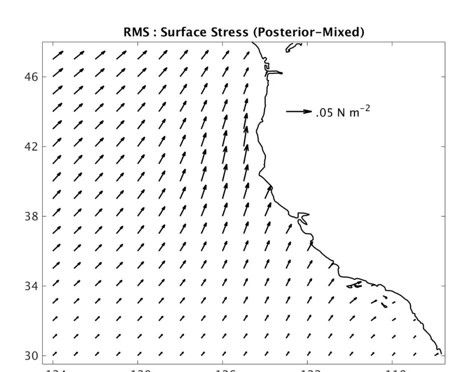

Figure 3: RMS difference between COAMPS background wind stress and the "mixed wind" stress

Results

- Two sets of data assimilation experiments computed for each of the years 2003 and 2005:
  1) 4D-Var adjusted winds used in the twin-model experiment
  2) BHM used to test the 4D-Var wind corrections (i.e. "mixed winds" case)
- The 2003 cases will not be independent of the observations since they are used in the twin model experiments, however, 2005 does represents an independent period apart from any propagation of information in time from the observations in the 2003-2004 time period
- Time series of the observational component of the cost function, $J_o = (\mathbf{y} - H(\mathbf{z}_a))^T\mathbf{R}^{-1}(\mathbf{y} - H(\mathbf{z}_a))$ indicates the case utilizeing the BHM to estimate $\mathbf{Q}$ outperforms the case of strong constraint 4D-Var and the weak constraint case without use of the BHM for both 2003 and 2005 (Figure 4).
- When allowance is made for model error during weak constraint 4D-Var, we expect the corrections to the surface forcing to be smaller than those subject to the strong constraint (Figure 5)
- We attribute much of this correction to the influence of model error since COAMPS verifies well against independent observations[2], and ROMS CCS is known to possess errors in this region[8]

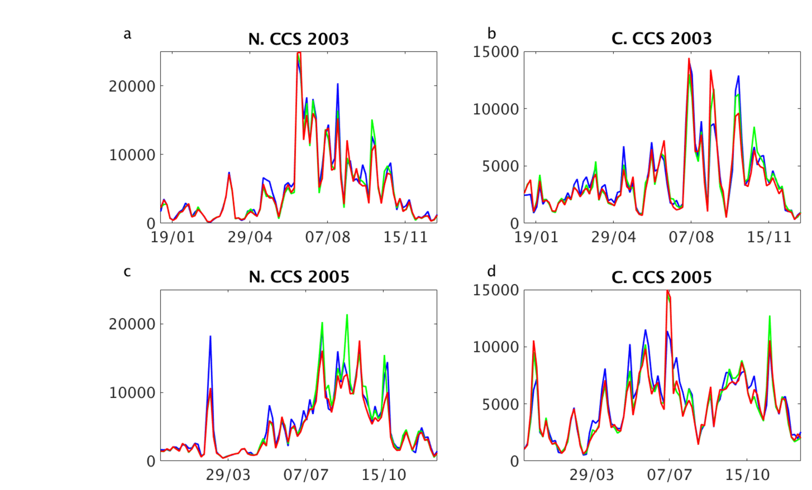

Figure 4: Observational component of the cost function within coastal regions in the northern CCS and the central CCS to 100 km offshore (shown in domain figure in the upper left hand corner) during 2003 (a and b) and 2005 (c and d). Strong constraint case (blue curve), and the weak constraint w/ BHM (red curve) and Weak Constraint w/o BHM (green curve).

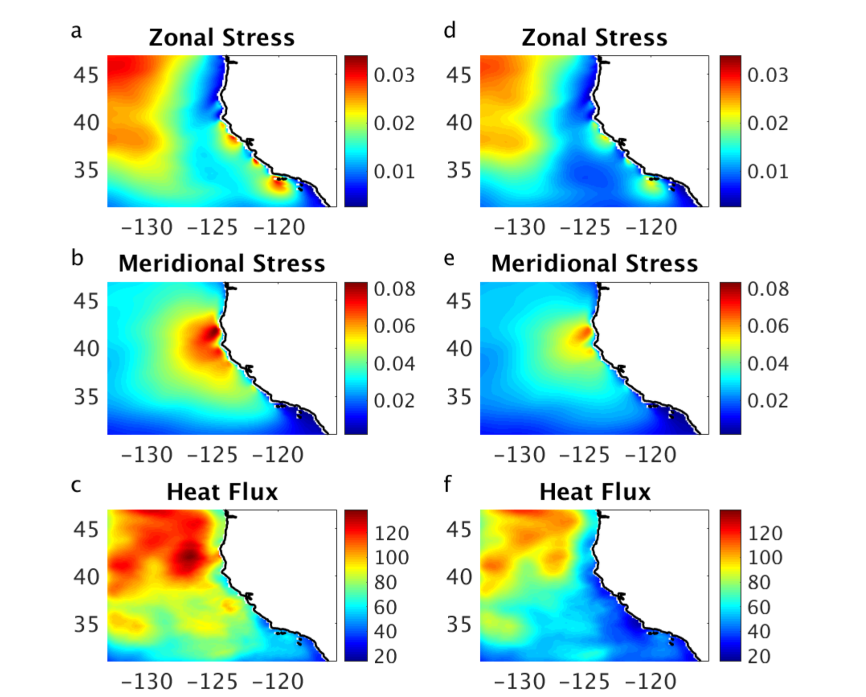

Figure 5: RMS difference between analysis and background 4D-Var estimates in 2005 of (a) zonal wind stress (N m[2]), (b) meridional wind stress, (N m[2]) and (c) surface heat flux (W m[2]) from the strong constraint calculations. Panels (d)-(f) show the corresponding differences for the weak constraint 4D-Var experiment BHM.

- During weak constraint 4D-Var, corrections are made to account for model error in every prognostic state variable
- RMS of the corrections per time step during the BHM experiment in 2005 correspond to the same geographic locations as the largest standard deviations in Fig. 1 (Figure 6)
- The corrections are also similar for 2003, an indication of the validity of the $\mathbf{Q}$ derived from the period 2003-2004 to other independent years

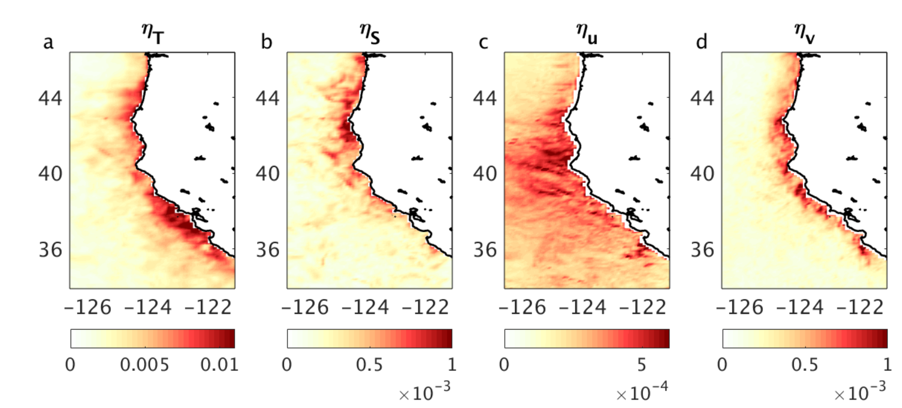

Figure 6: The RMS correction made by weak constraint 4D-Var per time step during the BHM experiment in 2005 for (a) SST, (b) SSS, (c) surface zonal velocity and (d) surface meridional velocity.

References:

[1] Moore, et al., *Prog. Oceanogr.* **2011a**, *91*, 34-49 [2] Doyle, et al., *Deep Sea Res. II*, **2009**, *56*, 87–99 [3] Carton & Giese, *Mon. Wea. Rev.* **2008**, *136*, 2999-3017 [4] Milliff, et al., *Q. J. R. Meteorol. Soc.*, **2011** *137*, 858–878 [5] von Storch, et al. *J. Climate.* **1995**, *8*, 377-400 [6] Penland & Sardeshmukh, *J. Climate,* **1995**, *8*, 1999–2024 [7] Veneziani, *J. Geophys. Res.* **2009**, C04015 [8] Broquet, et al., *Mercator Ocean Quarterly Newsletter,* **2011**, *91*, 34-49