Representer-Based Variational Data Assimilation in a Nonlinear Model of Nearshore Circulation

Alexander L. Kurapov, G. D. Egbert, J. S. Allen, R. N. Miller (*College of Oceanic and Atmospheric Sciences, Oregon State University*)

With support from ONR Physical Oceanography Program

INTRODUCTION

Variational data assimilation (DA):

-finds the ocean state estimate by fitting dynamical models to observations in a least squares sense, in a finite time interval

-adjusts model inputs (initial and boundary conditions, forcing, parameters)

-interpolates between observations (sparse in space and time) utilizing a dynamically consistent state-dependent model solution error covariance

Challenges with nonlinear models:

-Tangent linear (TL) and adjoint (ADJ) models are required; optimization algorithms are complicated and computationally intense

-Growing instabilities in the TL model may pose a threat to convergence

-Variability associated with nonlinear eddy interactions is not necessarily present in the model inputs (forcing). However, variational DA finds the optimal solution by adjusting model inputs.

Can we correctly estimate the ocean state dominated by these interactions?

SHALLOW-WATER NEARSHORE CIRCULATION MODEL

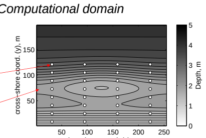
(after Slinn, Allen & Holman, JGR 2000)

- Forcing: steady, represents the effect of breaking waves
- Linear bottom drag, bi-harmonic horizontal dissipation
- Alongshore periodic boundary conditions

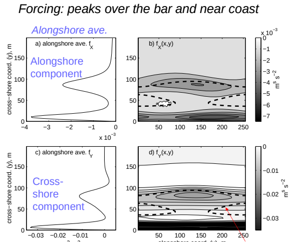
Computational domain

circles: observation locations

sandbar



Forcing: peaks over the bar and near coast



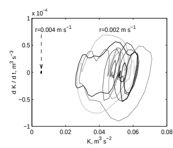
H=2 m

Model response to steady forcing depends on the magnitude of bottom friction coefficient (r):

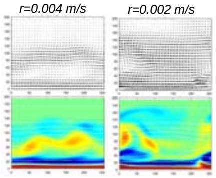
r=0.004 m/s: equilibarated waves

r=0.002 m/s: more strongly nonlinear, irregular eddy flow

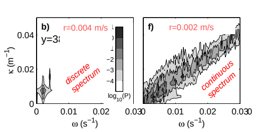
K vs. dK/dt, where K is the area-averaged kinetic energy of deviations from the time-mean



Snapshots of velocity and vorticity fields



Wave-number/frequency (κ/ω) power spectrum of vorticity at y=38 m (inshore of the sandbar)



DATA ASSIMILATION METHOD (ref: Chua and Bennett, Oc. Mod. 2001)

Nonlinear model: $\frac{\partial \mathbf{u}}{\partial t} = N(\mathbf{u}) + \mathbf{f}(t) + \boldsymbol{\varepsilon}_f(t)$

Initial condition: $\mathbf{u}(\mathbf{x},0) = \mathbf{i}(\mathbf{x}) + \boldsymbol{\varepsilon}_0$

Data: $\mathbf{g}_k \bullet \mathbf{u} = d_k + e_k \quad (k=1,\ldots,K)$, or $L(\mathbf{u}) = \mathbf{d} + \mathbf{e}$

Penalty functional: $J(\mathbf{u}) = \underbrace{\boldsymbol{\varepsilon}_0 * \mathbf{C}_0^{-1} * \boldsymbol{\varepsilon}_0}_{\text{integration over the area}} + \underbrace{\boldsymbol{\varepsilon}_f \bullet \mathbf{C}_f^{-1} \bullet \boldsymbol{\varepsilon}_f}_{\text{integration over the area and time [0..T]}} + \mathbf{e}'\mathbf{C}_d^{-1}\mathbf{e}$

The n-th inverse solution: $\mathbf{u}^n_{INV} = \mathbf{u}^n_{FWD} + \boldsymbol{\psi}^n$

Prior (fwd) model (full state TL model):

$$\frac{\partial \mathbf{u}^n_{FWD}}{\partial t} = N(\mathbf{u}^{n-1}_{INV}) + \mathbf{A}[\mathbf{u}^{n-1}_{INV}](\mathbf{u}^n_{FWD} - \mathbf{u}^{n-1}_{INV}) + \mathbf{f}$$

$$\mathbf{u}^n_{FWD}(\mathbf{x},0) = \mathbf{i}$$

TL operator, dependent on previous iteration solution

The representer method: nonlinear Euler-Lagrange equations (necessary conditions for min J(u)) are solved iteratively, as a series of linearized optimization problems ("outer loop iterations", $\{\mathbf{u}^n_{INV}\}$)

To compute the n-th correction field (ψ^n), utilize the ADJ and "perturbation" TL models repeatedly, searching for optimal b_k in the data subspace ("inner loop iterations"):

ADJ model:

$$-\frac{\partial \boldsymbol{\lambda}^n}{\partial t} = \mathbf{A}'[\mathbf{u}^{n-1}_{INV}]\boldsymbol{\lambda}^n + \sum_{k=1}^{K} b_k \mathbf{g}_k$$

$$\boldsymbol{\lambda}^n(\mathbf{x},T) = 0$$

Perturbation TL model:

$$\frac{\partial \boldsymbol{\psi}^n}{\partial t} = \mathbf{A}[\mathbf{u}^{n-1}_{INV}]\boldsymbol{\psi}^n + \mathbf{C}_f \bullet \boldsymbol{\lambda}^n$$

$$\boldsymbol{\psi}^n(\mathbf{x},0) = \mathbf{C}_0 * \boldsymbol{\lambda}^n(\mathbf{x},0)$$

The convergence of the nonlinear optimization algorithm over long integration times and the accuracy of the forcing estimate may depend on the choice of the assumed forcing error covariance $\mathbf{C}_f(\mathbf{x}_1,t_1,\mathbf{x}_2,t_2)=\mathbf{C}_x(x_1,x_2)\mathbf{C}_T(t_1,t_2)$ (particularly, the choice of $\mathbf{C}_T$)

THE VALIDITY LIMIT OF THE TL MODEL

Dataless Iterations:

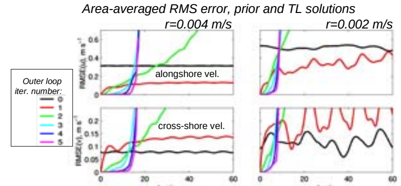
- Use the full state TL model (linearized w/ respect to previous iter TL solution or 0 on the 1st iteration)

- Run with true forcing and IC

TL solutions converge to the truth only in a limited time interval t_{lim}

Can DA be approached in a time interval T >> t_{lim}?

Area-averaged RMS error, prior and TL solutions



equilibrated waves regime: t_{lim}=15 min

irregular eddy regime: t_{lim}=7 min

DA EXPERIMENT SET-UP (WITH SYNTHETIC DATA):

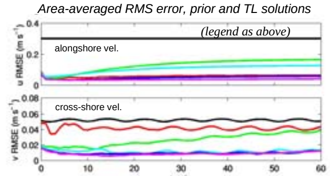
- True solution: fully developed eddy flow
- Observations: true surf. elevation and velocities sampled at the 32 locations (see map on the left), every min; random noise added (0.02 m, 0.02 m/s)
- Prior=0 (deliberately, the prior is chosen to be substantially different from the truth, to explore the nonlinear part of the algorithm)

ASSIMILATION IN THE WEAKLY NONLINEAR (EQULIBRATED WAVES) REGIME (r=0.004 m/s):

Outer loop iterations converge to the true solution in the 1 h interval with the steady $\mathbf{C}_f$ (C_T = const)

In this regime, initial conditions and steady forcing provide effective means of control of the time-variable flow. Forcing provides a correct steady energy input and initial conditions define the initial phase of the shear wave

Area-averaged RMS error, prior and TL solutions

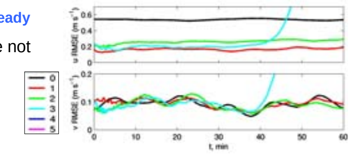


ASSIMILATION IN THE STRONGLY NONLINEAR (IRREGULAR FLOW) REGIME (r=0.002 m/s):

Case 1. Try C_T = const, since true forcing is steady

Then, scales of observed flow variability are not represented in the forcing correction.

Outer loop iterations do not converge

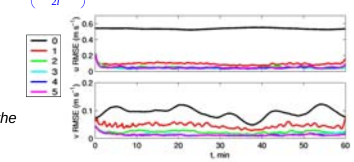


Case 2. Standard bell-shaped C_T: $C_T(t_1-t_2) = C_T^{BELL} = \exp\left(\frac{-(t_1-t_2)^2}{2l^2}\right)$

l=5 min: no convergence

l=1 min: convergence (TL solution is close to the true NL solution)

In this regime, it is important to allow variability in forcing associated with rapid changes inshore of the sand bar



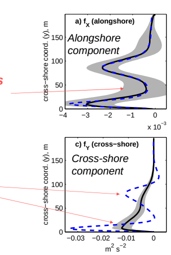
Forcing estimate: not quite accurate with $C_T=C^{BELL}$

Alongshore and time ave. forcing ± temporal StD

(blue line is true frc)

Temporal variations comparable to the time mean

Underestimated peak forcing

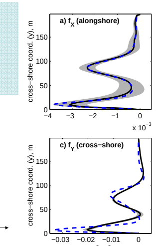


Case 3: composite C_T

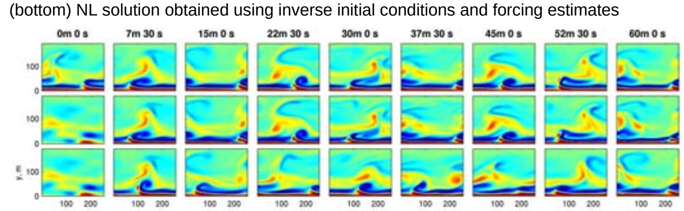
(C_T=1+ *a* C^{BELL}, where *a*<<1)

Accurate and convergent TL solution

Better, than in case 2, forcing estimate



Vorticity snapshots (in case 3): (top) truth, (middle) iteration 7 TL inverse solution, and (bottom) NL solution obtained using inverse initial conditions and forcing estimates



CONCLUSIONS:

In a strongly nonlinear flow regime, the representer method converges to an accurate solution over a time interval that is substantially longer than the validity limit of the TL model or the eddy time scale

To stabilize inversion, time variability in the forcing (on eddy time scale) is allowed by a choice of the forcing error covariance