Weak and Strong Constraint 4DVAR in the Regional Ocean Modeling System (ROMS): Development and Applications

Di Lorenzo, E.
Georgia Institute of Technology

Arango, H.
Rutgers University

Moore, A. and B. Powell
UC Santa Cruz

Cornuelle, B and A.J. Miller
Scripps Institution of Oceanography
❖ Short review of development and theory
   (an alternative derivation of the representer method)

❖ Current applications
   (some pending issues)
Inverse Ocean Modeling System (IOMs)
Chua and Bennett (2001)
To implement a representor-based generalized inverse method to solve weak constraint data assimilation problems

NL-ROMS, TL-ROMS, REP-ROMS, AD-ROMS
Moore et al. (2004)

Inverse Regional Ocean Modeling System (ROMS)
Di Lorenzo et al. (2007)
a representor-based 4D-variational data assimilation system for high-resolution basin-wide and coastal oceanic flows
ROMS Block Diagram

NEW Developments

Stability Analysis Modules
Non Linear Model
Tangent Linear Model
Representer Model
Adjoint Model
Sensitivity Analysis

Data Assimilation
1) Incremental 4DVAR
   Strong Constrain
2) Indirect Representer
   Weak and Strong
   Constrain
3) PSAS

Ensemble Ocean Prediction

KERNEL
NLM, TLM, RPM, ADM

physics
biogeochemical
sediment
sea ice

TLCHECK_OCEAN
PICARD_OCEAN
GRAD_OCEAN
PERT_OCEAN

Optimal perturbations
ADM eigenmodes
TLM eigenmodes
Forcing singular vectors
Stochastic optimals
Pseudospectra

Regional Ocean Modeling System
Research Community

Arango et al. 2003
Moore et al. 2004
Di Lorenzo et al. 2007
Download:

ROMS components
http://myroms.org
Arango H.

IOM components
http://iom.asu.edu
Muccino, J. et al.
ASSIMILATION Goal

\[ u(t) + \text{Use Observations} \rightarrow \hat{u}(t) \]

- Initial Guess
- Best Model Estimate (consistent with observations)

\[(A)\]
**STRONG Constraint**

\[
\begin{align*}
\frac{\partial u}{\partial t} &= N(u) + F(t) \\
u(0) &= u_0 + e_0
\end{align*}
\]

\[(B)\]
**WEAK Constraint**

\[
\begin{align*}
\frac{\partial u}{\partial t} &= N(u) + F(t) + e(t) \\
u(0) &= u_0 + e_0
\end{align*}
\]

...we want to find the corrections \( e \)
Quadratic Linear Cost Function for residuals $J[e_0]$.

$$
J[e_0] = \left[ \hat{d} - \int_{t_0}^{T} H(t')R(t_0, t')dt' e_0 \right]^T C^{-1}_e \left[ \hat{d} - \int_{t_0}^{T} H(t')R(t_0, t')dt' e_0 \right] + e_0^T P^{-1} e_0
$$
Quadratic Linear Cost Function for residuals \(J[e_0]\)

\[
J[e_0] = \begin{bmatrix}
\hat{d} - \int_{t_0}^{T} H(t')R(t_0, t')dt' e_0 \\
+ e_0^T P^{-1} e_0
\end{bmatrix}^T \begin{bmatrix}
C_e^{-1} \\
\hat{d} - \int_{t_0}^{T} H(t')R(t_0, t')dt' e_0
\end{bmatrix}
\]

1) corrections should reduce misfit within observational error

2) corrections should not exceed our assumptions about the errors in model initial condition.
ASSIMILATION

Cost Function

\[ J[e_0] = \begin{bmatrix} \hat{d} - \int_{t_0}^{T} H'(t')R(t_0,t')dt' e_0 \end{bmatrix}^T \begin{bmatrix} \hat{d} - \int_{t_0}^{T} H'(t')R(t_0,t')dt' e_0 \end{bmatrix} C^{-1}_e \]

\[ + e_0^T P^{-1} e_0 \]
\( G \) is a mapping matrix of dimensions observations X model space

def: 
\[
G = \int_{t_0}^{T} H(t')R(t_0, t')dt'
\]

\[
J[e_0] = \left[\hat{d} - \int_{t_0}^{T} H(t')R(t_0, t')dt'e_0\right]^T C_e^{-1} \left[\hat{d} - \int_{t_0}^{T} H(t')R(t_0, t')dt'e_0\right] + e_0^T P^{-1} e_0
\]
\( \hat{J}[e_0] = \left[ \hat{d} - \hat{G}e_0 \right]^T \hat{C}_\varepsilon^{-1} \left[ \hat{d} - \hat{G}e_0 \right] + e_0^T P^{-1} e_0 \)

\[ G = \int_{t_0}^{T} H(t') R(t_0, t') dt' \]

\( G \) is a mapping matrix of dimensions observations X model space.

**Assimilation**

**Cost Function**
\[
\left( G^T C_{\epsilon}^{-1} G + P^{-1} \right) e_0 - G^T C_{\epsilon}^{-1} \hat{d} = 0
\]

\[
\frac{\partial J[e_0]}{\partial e_0} = 0
\]

\[
J[e_0] = \left[ \hat{d} - \int_{t_0}^{t} H(t')R(t_0, t')dt' e_0 \right]^T C_{\epsilon}^{-1} \left[ \hat{d} - \int_{t_0}^{t} H(t')R(t_0, t')dt' e_0 \right] + e_0^T P^{-1} e_0
\]

\[
J[e_0] = \left[ \hat{d} - Ge_0 \right]^T C_{\epsilon}^{-1} \left[ \hat{d} - Ge_0 \right] + e_0^T P^{-1} e_0
\]
4DVAR inversion

\[
\left( G^T C_\varepsilon^{-1} G + P^{-1} \right) \hat{e}_0 - G^T C_\varepsilon^{-1} \hat{d} = 0
\]

**Hessian Matrix**

\[
\hat{H}
\]

**def:**

\[
G = \int_{t_0}^{T} H(t') R(t_0, t') dt'
\]
4DVAR inversion

\[
\left( G^T C_\varepsilon^{-1} G + P^{-1} \right) e_0 - G^T C_\varepsilon^{-1} \hat{d} = 0
\]

Hessian Matrix

\[
\begin{align*}
\hat{H} = & \frac{dGPCG eGP}{G = \int_{t_0}^{T} H(t')R(t_0, t')dt'}
\end{align*}
\]

IOM representer-based inversion

\[
\left( \hat{\beta}^T + C_\varepsilon \right) \left( \beta^T \right)^{-1} e_0 = \hat{d}
\]
4DVAR inversion

\[
\left( \hat{H} \right) \left( G^T C_\varepsilon^{-1} G + P^{-1} \right) e_0 - G^T C_\varepsilon^{-1} \hat{d} = 0
\]

Hessian Matrix

IOM \textit{representer-based} inversion

\[
\left( \hat{P} \right) \left( GPG^T + C_\varepsilon \right) \left( PG^T \right)^{-1} e_0 = \hat{d}
\]

Stabilized \textit{Representer Matrix}

\[
\hat{R} \equiv GPG^T
\]

\text{def:}

\[
G = \int_{t_0}^{T} H(t') R(t_0, t') dt'
\]
4DVAR inversion
\[
\int_{t_0}^{T} \left[ G^T(t) \mathbf{C}_e^{-1} G(t') + \mathbf{C}^{-1}(t, t') \right] e(t') dt' - G^T(t) \mathbf{C}_e^{-1} \hat{d} = 0
\]
\[
\hat{H}(t, t')
\]

Hessian Matrix

IOM representer-based inversion
\[
\int_{t_0}^{T} \int_{t_0}^{T} \left[ G(t') \mathbf{C}(t', t'') G^T(t'') + \mathbf{C}_e \right] dt' dt'' \int_{t_0}^{T} \int_{t_0}^{T} \left[ \mathbf{C}(t', t'') G^T(t'') \right]^{-1} e(t') dt' dt'' = \hat{d}
\]
\[
\hat{P}
\]

Stabilized Representer Matrix

Representer Coefficients

\[
\hat{R} \equiv \int_{t_0}^{T} \int_{t_0}^{T} \left[ G(t') \mathbf{C}(t', t'') G^T(t'') + \mathbf{C}_e \right] dt' dt''
\]
def:
\[
G(t) = \int_{t}^{T} H(t') R(t, t') dt'
\]
An example of *Representer Functions* for the Upwelling System

Computed using the TL-ROMS and AD-ROMS

\[
\langle s(\hat{t})s^T(\hat{t}) \rangle = \\
= \langle R(t_0, \hat{t})s(t_0)(R(t_0, \hat{t})s(t_0))^T \rangle \\
= R(t_0, \hat{t}) \langle s(t_0)s^T(t_0) \rangle R^T(\hat{t}, t_0) \\
= R(t_0, \hat{t})C(t_0, t_0)R^T(\hat{t}, t_0)
\]
An example of **Representer Functions** for the Upwelling System

Computed using the TL-ROMS and AD-ROMS

$$\langle s(\hat{t})s^T(\hat{t}) \rangle =$$

$$= \langle R(t_0,\hat{t})s(t_0)(R(t_0,\hat{t})s(t_0))^T \rangle$$

$$= R(t_0,\hat{t}) \langle s(t_0)s^T(t_0) \rangle R^T(\hat{t},t_0)$$

$$= R(t_0,\hat{t})C(t_0,t_0)R^T(\hat{t},t_0)$$
Applications of inverse ROMS:

- **Baroclinic coastal upwelling**: synthetic model experiment to test the development

- **CalCOFI Reanalysis**: produce ocean estimates for the CalCOFI cruises from 1984-2006. *Di Lorenzo, Miller, Cornuelle and Moisan*

- **Intra-Americas Seas Real-Time DA**
  *Powell, Moore, Arango, Di Lorenzo, Milliff et al.*
Coastal Baroclinic Upwelling System Model Setup and Sampling Array
Applications of inverse ROMS:

- **Baroclinic coastal upwelling**: synthetic model experiment to test the development

1) The representer system is able to initialize the forecast extracting dynamical information from the observations.

2) Forecast skill beats persistence
SKILL of assimilation solution in Coastal Upwelling
Comparison with independent observations

Di Lorenzo et al. 2007; Ocean Modeling
Assimilation solutions

Day=0

Day=2

Day=6

Day=10
Day=14

Day=18

Day=22

Day=26
Intra-Americas Seas Real-Time DA

Powell, Moore, Arango, Di Lorenzo, Milliff et al.

www.myroms.org/ias
CalCOFI Reanalysis: produce ocean estimates for the CalCOFI cruises from 1984-2006.

Di Lorenzo, Miller, Cornuelle and Moisan
Things we struggle with …

- Tangent Linear Dynamics can be very unstable in realistic settings.

- Background and Model Error COVARIANCE functions are Gaussian and implemented through the use of the diffusion operator.

- Fitting data vs. improving the dynamical trajectory of the model.
Assimilation of surface Salinity $t_N$
Assimilation of surface Salinity $t_N$

True Initial Condition

Which model has correct dynamics?

$t_N$

True

Model 1

Model 2
Time Evolution of solutions after assimilation

Wrong Model

Good Model
Time Evolution of solutions after assimilation

Wrong Model

Good Model

DAY 1
Time Evolution of solutions after assimilation

Wrong Model

DAY 2

Good Model
Time Evolution of solutions after assimilation

Wrong Model

Good Model

DAY 3
Time Evolution of solutions after assimilation

Wrong Model

Good Model

DAY 4
RMS difference from TRUE

Observations

Days

RMS

Less constraint

More constraint
Applications of inverse ROMS (cont.)

- **Improve model seasonal statistics** using surface and open boundary conditions as the only controls.

- **Predictability of mesoscale flows in the CCS:** explore dynamics that control the timescales of predictability.

*Mosca et al. – (Georgia Tech)*
inverse machinery of ROMS can be applied to regional ocean climate studies ...
inverse machinery of ROMS can be applied to regional ocean climate studies ...

**EXAMPLE:**
Decadal changes in the CCS upwelling cells

Chhak and Di Lorenzo, 2007; *GRL*
SSTa Composites

Observed PDO index
Model PDO index

Cold Phase
Warm Phase

Chhak and Di Lorenzo, 2007; GRL
Tracking Changes of CCS Upwelling Source Waters during the PDO using adjoint passive tracers ensembles

COLD PHASE
ensemble average

APRIL UPEWLLING SITE

WARM PHASE
ensemble average

Chhak and Di Lorenzo, 2007; GRL
Changes in depth of Upwelling Cell (Central California) and PDO Index Timeseries

Model PDO
PDO lowpassed
Surface
0-50 meters
(-) 50-100 meters
(-) 150-250 meters

Chhak and Di Lorenzo, 2007; GRL
References


Adjoint passive tracers $P(t)$ ensembles

\[
\frac{\partial P}{\partial t} = -\mathbf{u} \cdot \nabla P + K \frac{\partial^2 T}{\partial z^2}
\]

$P(t_0) = P_0$

\[\mathbf{u}\text{ physical circulation independent of } P(t)\]
What if we apply more smoothing?

True Initial Condition

Wrong Model

Good Model

True

Model 1

Model 2
Assimilation of data at time $t_N$

**True Initial Condition**

**True**

**Model 1**

**Model 2**
April Upwelling Site

COLD PHASE
ensemble average

WARM PHASE
ensemble average

Chhak and Di Lorenzo, 2007; GRL
What if we really have substantial model errors?

\[
\frac{\partial P}{\partial t} + u \cdot \nabla P = K \frac{\partial^2 T}{\partial z^2}
\]

\[
P(t_0) = P_0
\]
Current application of inverse ROMS in the California Current System (CCS):

1) **CalCOFI Reanalysis**: produce ocean estimates for the CalCOFI cruises from 1984-2006.  
   *NASA - Di Lorenzo, Miller, Cornuelle and Moisan*

2) **Predictability of mesoscale flow in the CCS**: explore dynamics that control the timescales of predictability.  
   *Mosca and Di Lorenzo*

3) **Improve model seasonal statistics** using surface and open boundary conditions as the only controls.
Comparison of SKILL score of IOM assimilation solutions with independent observations

\[
SKILL(s) = 1 - \frac{(s_{\text{true}} - s)^T (s_{\text{true}} - s)}{(s_{\text{true}} - s_{\text{clima}})^T (s_{\text{true}} - s_{\text{clima}})}
\]

**HIRES**: High resolution sampling array

**COARSE**: Spatially and temporally aliased sampling array
Instability of the Representer Tangent Linear Model (RP-ROMS)

RP-ROMS WEAK constraint solution

RP-ROMS with TRUE as BASIC STATE

RP-ROMS with CLIMATOLOGY as BASIC STATE
ASSIMILATION Setup
California Current

**Sampling:**
(from CalCOFI program)
5 day cruise
80 km stations spacing

**Observations:**
T,S CTD cast 0-500m
Currents 0-150m
SSH

**Model Configuration:**
Open boundary cond.
nested in CCS grid
20 km horiz. Resolution
20 vertical layers
Forcing NCEP fluxes
Climatology initial cond.
ASSIMILATION Results

SSH [m]

STARK
day=5

TRUE
day=5

WEAK
day=5

1st GUESS
day=5
ASSIMILATION Results

SSH [m] ERROR or RESIDUALS

1st GUESS day=5

WEAK day=5

STRONG day=5
Reconstructed Initial Conditions

WEAK
day=0

TRUE
day=0

STRONG
day=0

1st GUESS
day=0
Normalized Observation-Model Misfit

Error Variance Reduction
STRONG Case = 92%
WEAK Case = 98%

Assimilated data:
TS 0-500m
Free surface
Currents 0-150m