The ROMS IAS Data Assimilation and Prediction System: Quantifying Uncertainty

Andrew M. Moore
Department of Ocean Sciences, 1156 High St.
University of California
Santa Cruz CA 95064
Phone: (831) 459-4632 fax: (831) 459-4882 email: ammoore@ucsc.edu

Brian Powell
Department of Oceanography
University of Hawai`i at Manoa
1000 Pope Road
Honolulu HI 96822
Phone: (808) 956-6724 fax: (808) 956-2352 email: powellb@hawaii.edu

Award Number: N00014-08-1-0556

LONG-TERM GOALS

The long-term scientific goals of this research project are:

1. To develop a state-of-the-art ocean 4-dimensional variational (4D-Var) data assimilation and ocean forecasting system for the Regional Ocean Modeling System (ROMS);
2. To develop a state-of-the-art suite of post-processing and diagnostic tools in support of ROMS 4D-Var;
3. To gain the necessary experience using the ROMS 4D-Var systems in complex circulation environments;
4. To train the next generation of users of the ROMS 4D-Var system.

OBJECTIVES

The main objectives of this project are: (**i**) to assess the impact of observations on ocean state estimates and the ensuing forecasts; (**ii**) to quantify the expected errors in 4D-Var ocean circulation estimates; and (**iii**) to develop multimodel ensemble and superensemble methods for ocean models.

APPROACH

The primary tool used is the Regional Ocean Modeling System (ROMS. To address the aforementioned goals and objectives, we are using a recently developed suite of tools that utilize the tangent linear (TL), adjoint (AD), and finite-amplitude tangent linear (RP) versions of the ROMS code. ROMS, TLROMS and ADROMS were developed under the support of previous ONR funding, while the development of RPROMS was supported by NSF. Three 4D-Var data assimilation systems have been developed for ROMS (Moore et al., 2010a,b,c), one based on the primal formulation and two based on the dual formulation.

WORK COMPLETED

During the current reporting period we have completed the following tasks:

1. **Observation impact:** The transpose of the gain matrix derived from each 4D-Var assimilation cycle provides information about the impact of each observation on the circulation estimate. A separate driver is available as part of ROMS for computing observation impacts, and we have applied it to explore the impact of observations in two different ROMS configurations: the Intra-Americas Sea (IAS) and the California Current System (CCS).
2. **Observation sensitivity:** The adjoint of the entire 4D-Var system can be used to assess the impact of changes in the observations or the observation network for a single 4D-Var assimilation cycle without the need to recompute the 4D-Var analysis. This is extremely useful, and is a very efficient method for performing Observation System Experiments (OSEs). A separate driver is available as part of ROMS for computing observation sensitivities, and has been applied to ROMS CCS.
3. **Analysis and forecast error estimates based on adjoint 4D-Var:** Ensemble 4D-Var methods have been proposed in meteorology as a way of computing flow-dependent background error covariances (Belo Pereira and Berre, 2006). However, computing an ensemble of 4D-Var analyses is computationally very demanding. Using the adjoint of the 4D-Var system, we have developed a new method for efficiently computing background, analysis and forecast error variances without the need to explicitly generate the 4D-Var ensemble members. At the present time, we are restricted to computing covariance information for linear functions of the state-vector (e.g. space-time averages, transport, heat content, etc), but the method holds promise for extracting more detailed information.
4. **An improved ROMS configuration for the IAS:** Analysis of the water mass properties and inter-passage transports in ROMS IAS has revealed several deficiencies in the model simulation of the circulation which have necessitated the expansion of the model domain to include a large part of the North and South Atlantic, hereafter ROMS ATL. ROMS ATL is currently being used for data assimilation.
5. **4D-Var Workshop:** A one week, hands-on, ROMS 4D-Var workshop was held at UC Santa Cruz 12-16 July, 2010, attended by 37 scientists and students from seven different countries. The aim of the workshop was to provide help and training for expert ROMS users in running the ROMS 4D-Var drivers. The workshop was comprised of a combination of formal lectures, tutorials, and hands-on exercises based on ROMS CCS. Lecture notes are available at http://myroms.org/index.php?page=4DVAR_2010_agenda, while https://www.myroms.org/wiki/index.php/4DVar_Tutorial_Introduction provides detailed information about the hands-on exercises.
6. **Manuscript preparation:** A series of three manuscripts are also in the final stages of preparation and describe the entire ROMS 4D-Var system. Part I describes the mathematical and technical aspects of ROMS 4D-Var and supporting diagnostic algorithms, while Parts II and III demonstrate the performance of the systems, and present examples of calculations from all of the 4D-Var system components. These manuscripts will be submitted to *Progress in Oceanography* in the Fall of 2010. A fourth manuscript describing the use of the adjoint of 4D-

Var for computing error covariance information is also in preparation, with a planned submission to *Monthly Weather Review* in Fall/Winter 2010.

RESULTS

(a) <u>**Observation Impact**</u>

Here we show an example of observation impact results from the IAS domain only, since they have motivated further modifications to the model domain and assimilation system as outlined in 4 above. Our primary focus in the IAS has been the transport through the Yucatan Channel. Direct observations of Yucatan transport by Sheinbaum et al (2002) during the period 1999-2001 revealed a mean transport of 23.1 Sv with a variance of 3.6 Sv. In the absence of data assimilation, ROMS IAS yields a transport of 24.9 Sv and a variance of 3.2 Sv. When 4D-Var is applied sequentially using 7 day assimilation cycles and all available observations (satellite SST, SSH, and in situ data), the model mean transport is 23.5 Sv with a variance of 4.6 Sv. Therefore, 4D-Var improves the mean transport of each assimilation compared to the observations, although the variance is somewhat overestimated. Figure 1 shows a time series of the Yucatan transport increment associated with each observing platform providing observations within the model domain. The transport increment is defined as the difference in the *posterior* and *prior* transport estimates, and prior to June 1999 only satellite observations were assimilated. Figure 1 shows that all observation types contribute significantly to the transport increment. The in situ observations that have the largest impact are hydrographic measurements of temperature and measurements of velocity. The latter are associated with the ADCP measurements of Sheinbaum et al (2002) on which the observational estimates of transport are based.

Figure 2 shows the root mean square (rms) impact of all in situ observations on the Yucatan transport increments averaged over all 4D-Var cycles. Observations within the Yucatan Channel and downstream in the Gulf of Mexico exert a significant influence on the transport. However, there is a single XBT line that runs just east of the Bahamas and down through the Windward Passage which exerts a disproportionately large impact on the Yucatan transport. These large impacts are due to deficiencies in the model which are unable to accommodate the observations east of Bahamas because this is a shallow water region that is poorly represented in ROMS. When these data are assimilated into the model they generate large amplitude initialization shocks that affect the Yucatan Channel transport. This then highlights an additional utility of the observation impact calculations in that they can be used to identify problem areas in the model (as well as problems with the observations) that contribute to initialization shocks.

(b) <u>Observation Sensitivity</u>

The adjoint of 4D-Var can be use very efficiently for OSEs, and an example is presented in Figure 3 for ROMS CCS for the period July 2002-Dec 2004 using 7 day assimilation cycles. In this example we are considering the transport increments (*posterior* minus *prior*) across 37N from the coast to 127W, 0-500m, during each 4D-Var cycle. Each cycle is treated independently, and Fig. 3 shows the results from three calculations: (i) a case in which either all Argo observations or all satellite SSH observations are withheld, and each 4D-Var analysis cycle is repeated; (ii) a case where the change in the transport increment is predicted using the outcome of the observation sensitivity calculations using the adjoint of 4D-Var; and (iii) a case where the change in transport increment is predicted based on the observation impact using the transpose of the gain matrix. Figure 3 demonstrates that there is excellent agreement between the changes in transport directly computed by rerunning 4D-Var (blue curves) and those predicted by observation sensitivity calculations (red curves). On the other hand, the observation impacts are a poor predictor of the expected changes (black curves). The observation sensitivity requires only one additional run of the adjoint of the 4D-Var system, which can then be used to predict changes in circulation associated with any change in the observation array.

(c) The North Atlantic ROMS Configuration

Our previous data assimilation experiments using ROMS IAS have revealed a number of significant issues associated with modeling of this region. In particular, in order to obtain good *prior* estimates of the deep water masses and inter-passage transports, it is necessary to extend the model domain to include much of the North Atlantic and part of the South Atlantic. The new ROMS ATL domain has ~60 km horizontal resolution and 20 levels in the vertical, and the spatial extent of the model grid is shown in Figure 4. All of the available data set available for assimilation in the model are summarized in Table 1. Figure 4 also shows spatial coverage of the in situ temperature and salinity observations that are available at each model grid cell location during the period 1993-2005. The subsurface data coverage is excellent over much of the model domain, so we anticipate that this will help to constrain the *posterior* circulation estimates from 4D-Var, and alleviate many of the problems that we experienced using the ROMS IAS configuration. 4D-Var experiments using this model configuration are currently under way, and observation impact, observation sensitivity, and analysis error calculations that target the Yucatan and inter-passage transports of the Caribbean Sea will be performed using the results of 4D-Var.

IMPACT/APPLICATIONS

This project contributes significantly to the functionality and utility of the Regional Ocean Modeling System (ROMS), a popular and important community model and resource. ROMS is in fact unique in that of all the community ocean models that are available, ROMS is the only model that possesses such a wide range of 4D-Var algorithms, analysis tools, and diagnostic capabilities. The posterior analysis error, observation impact and observation sensitivity tools that are have been developed as a part of this project have advanced ROMS to a state where it is comparable to the most sophisticated operational systems currently available at several premier numerical weather prediction centers worldwide.

TRANSITIONS

The new ROMS utilities developed as part of this project are freely available from the ROMS web site http://myroms.org and will be actively used and further developed by other research groups in the U.S. and elsewhere as user competence increases.

RELATED PROJECTS

The work described here is closely related to the following ONR supported projects:

"A community Terrain-Following Ocean Model (ROMS)", PI Hernan Arango, grant number N00014-08-1-0542.

"Bayesian Hierarchical Models to Augment the Mediterranean Forecast System", PI Ralph Miliff, grant number N00014-05-C-0198.

"Understanding Predictability of the Ocean", PI Brian Powell, grant number N00014-09-1-0939.

REFERENCES

Belo Pereira, M. and L. Berre, 2006: The use of an ensemble approach to study the background error covariances in a global NWP model. *Mon. Wea. Rev.*, **134**, 2466-2498.

Sheinbaum, J., J. Candela, A. Badan and J. Ochoa, 2002: Flow structure and transport in the Yucatan Channel. *Geophys. Res. Letters*, **29**, doi:10.1029/2001GL013990.

PUBLICATIONS

Moore, A.M., H.G. Arango, G. Broquet, B.S. Powell, J. Zavala-Garay and A.T. Weaver, 2010a: The Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilation systems. Part I: System overview and formulation. *Progress in Oceanography*, In preparation for submission in Oct 2010.

Moore, A.M., H.G. Arango, G. Broquet, C. Edwards, M. Veneziani,
B. Powell. D. Foley, J. Doyle, D. Costa and P. Robinson, 2010b: 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*, In preparation for submission in Oct 2010.

Moore, A.M., H.G. Arango, G. Broquet, C. Edwards, M. Veneziani, B. Powell. D. Foley, J. Doyle, D. Costa and P. Robinson, 2010c: 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*, In preparation for submission in Oct 2010.

Moore, A.M., H.G. Arango and G. Broquet, 2010: Analysis and forecast error estimates derived from the adjoint of 4D-Var. *Monthly Weather Review*, In preparation for submission in Fall 2010.

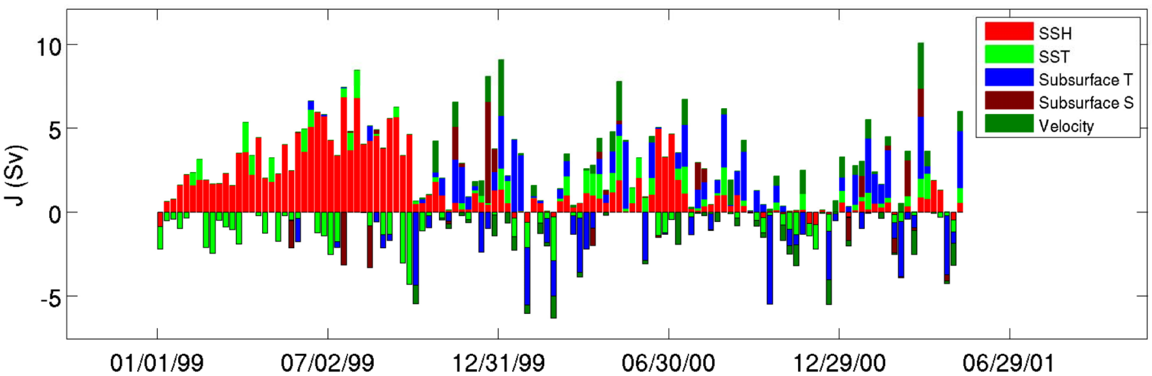

Figure 1: A time series of the Yucatan transport increment (i.e. posterior transport minus prior transport) for each 4D-Var assimilation cycle. The contribution of each observing platform to the transport increment is indicated by the different colored bars, while the sum of all the bars represents the total transport increment.

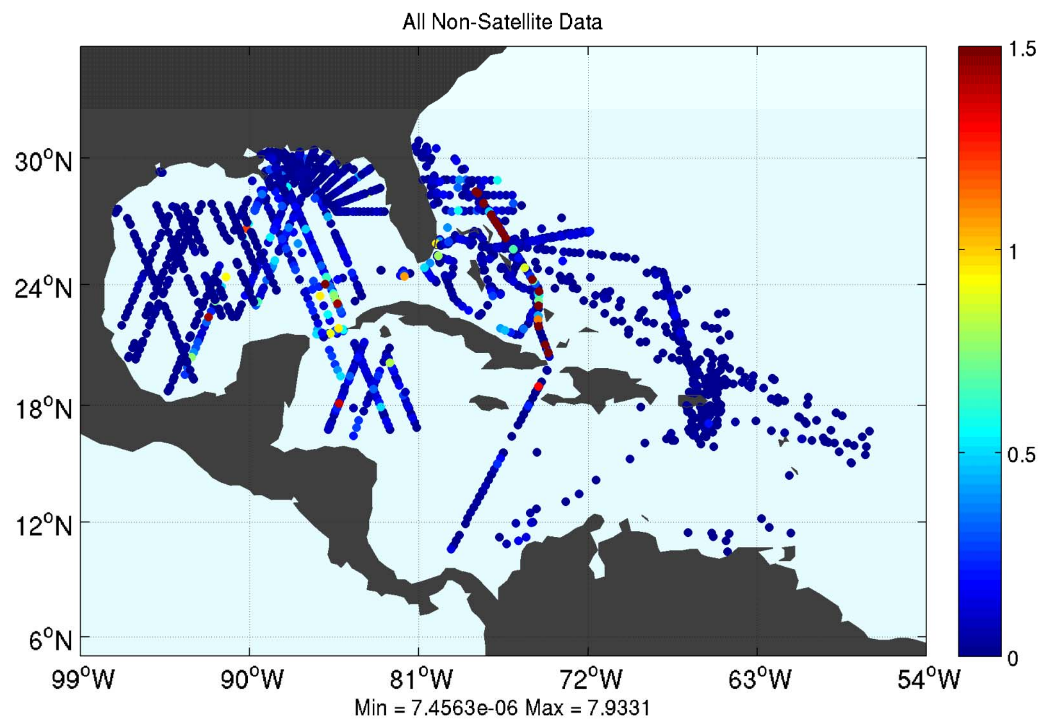

Figure 2: The rms impact of in situ observations on Yucatan transport increments averaged over all data assimilation cycles.

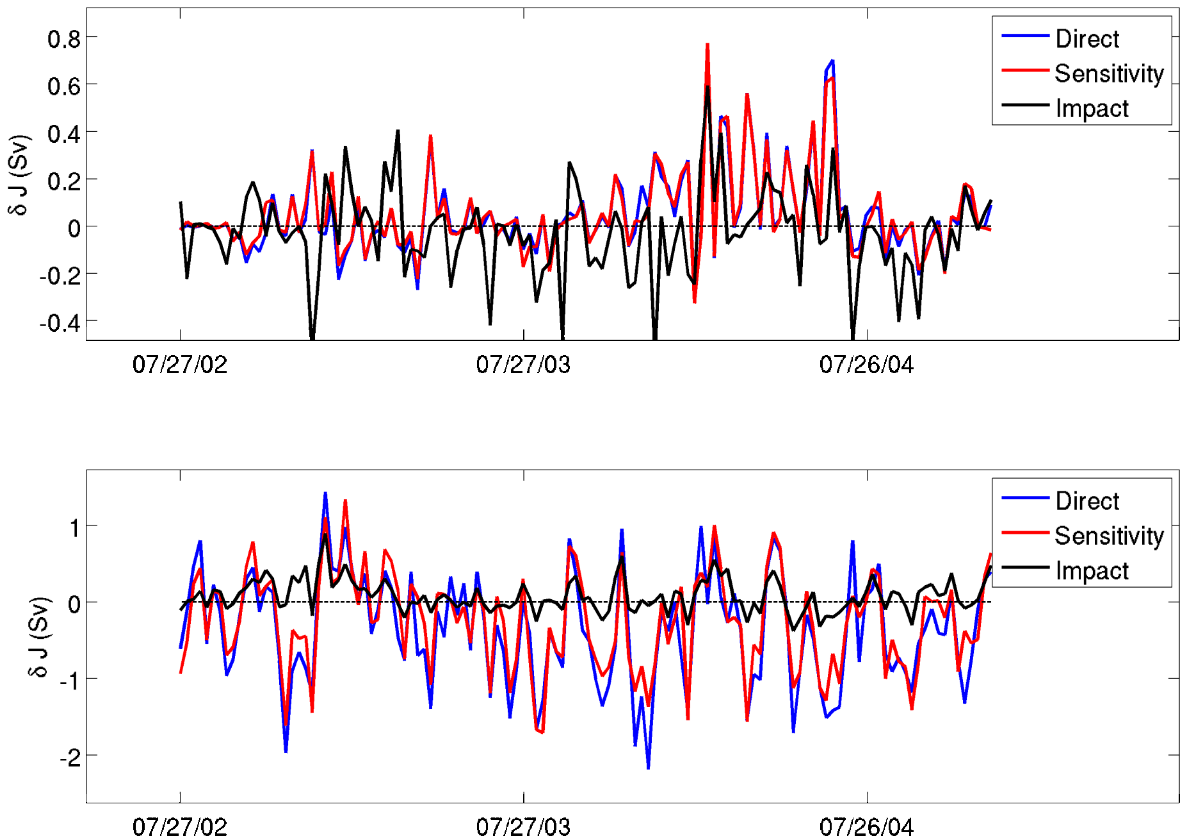

Figure 3: Time series (blue curves) of the change in 37N transport computed directly from the prior and posterior circulations estimates of each 7-day 4D-Var cycle when (a) all Argo observations are withheld, and (b) all satellite SSH observations are withheld. Also shown are time series (red curves) of the change in 37N transport based on the predictions of observation sensitivity analyses. Time series based on the predictions of observation impact analyses are shown by the black curve.

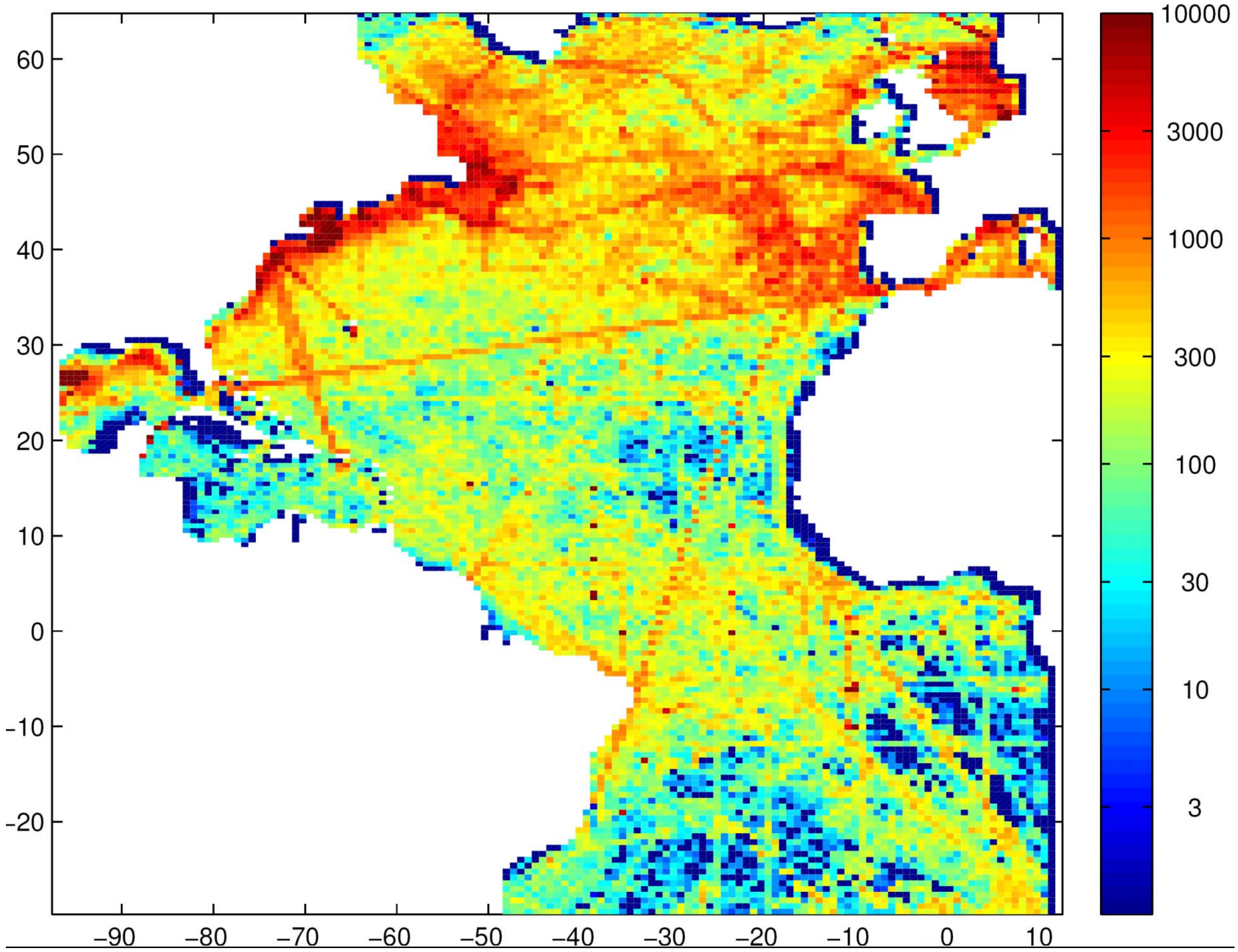

Figure 4: The number of in situ hydrographic observations available during the period 1993-2005 that fall within each grid cell of the new ROMS ATL model domain.

Data Source	***Type***	***Date Range***	***Number of Observations***
GHRSST	SST	1993-2006	66,797,303
AVISO	SSH	1993-2006	62,525,941
Levitus Climatology	Climatology	1993-2006	8,286,628
MetOffice Hadley Centre EN3 Dataset	CTD, XBT, ARGO	1993-2006	6,365,950
WOCE (World Ocean Circulation Experiment)	ADCP	1993-1999	1,349,386
MMS ADCP (Department of Interior's Minerals Management Service)	ADCP	2005	1,078,420
CICESE	CTD, Thermistor, ADCP	1999-2006	1,025,426
SAIC	CTD	2004-2005	683,309
JASADCP (Joint Archive for Shipboard ADCP)	ADCP	1993-2006	185,682
NDBC (National Data Buoy Center)	ADCP	2002-2006	124,124

Table 1: A summary of the ocean observations that will be assimilated into ROMS ATL.