Toward the Development of a Coupled COAMPS-ROMS Ensemble Kalman Filter and Adjoint with a focus on the Indian Ocean and the Intraseasonal Oscillation

PI: Andrew M. Moore[1], Tel: 831-459-4632, e-mail: ammoore@ucsc.edu
Co-PIs: James Doyle[2], Tel: 831-656-4716, e-mail: james.doyle@nrlmry.navy.mil
Jim Hansen[2], Tel: 831-656-4741, e-mail: james.hansen@nrlmry.navy.mil
Saša Gaberšek[2], Tel: 831-656-4882, e-mail: sasa.gabersek@nrlmry.navy.mil
Sue Chen[2], Tel: 831-656-4737, e-mail: sue.chen@nrlmry.navy.mil
Sergey Frolov[2], Tel: 831-656-4050, e-mail: sergey.frolov@nrlmry.navy.mil
Alex Reinecke[2], Tel: 831-656-4645, e-mail: alex.reinecke@nrlmry.navy.mil
Christopher A. Edwards[1], Tel: 831-459-3734, e-mail: cedwards@ucsc.edu
Hernan G. Arango[3], Tel: 848-932-3368, e-mail: arango@marine.rutgers.edu
Javier Zavala-Garay[3], Tel: 848-932-3457, e-mail: jzavala@marine.rutgers.edu
Jeffrey Anderson[4], 303-497-8991, e-mail: jla@ucar.edu

1 Dept of Ocean Sciences, 1156 High Street, University of California Santa Cruz, Santa Cruz CA 95064
2 Naval Research Laboratory, 7 Grace Hopper Ave, Monterey CA 93943-5502
3 Department of Marine and Coastal Sciences, Rutgers, The State University of New Jersey
71 Dudley Road, New Brunswick, NJ 08901-8521
4 National Center for Atmospheric Research, P.O. Box 3000, Boulder, Colorado 80307-3000

Award Number: N00014-15-1-2300

Performance period: 16 June 2016 to 15 June 2017
ONR Program Officer: Daniel Eleuterio (daniel.eleuterio@navy.mil)

LONG-TERM GOALS

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

1. Develop a regional coupled ocean-atmosphere model comprising the Navy's Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS) and the Regional Ocean Modeling System (ROMS).
2. Explore data assimilation strategies in regional coupled ocean-atmosphere models with a specific focus on air-sea interaction in the tropical Indian Ocean.

OBJECTIVES

The project has the following objectives:

(1) Interface ROMS and COAMPS with the Data Assimilation Research Testbed (DART) system at NCAR.

(2) Compare the performance of the Ensemble Kalman Filter (EnKF) using the Data Assimilation Research Testbed (DART) and 4-dimensional variational (4D-Var) approaches to data assimilation in ocean only and atmosphere only experiments.

(3) Develop a coupled model capability based on the pre-existing COAMPS and ROMS systems using the Earth System Modeling Framework (ESMF) NUOPC interface.

(4) Investigate the merits of the EnKF approach for assimilating data into a coupled model, particularly in the highly dynamic tropical Indian Ocean where air-sea interactions play a fundamental role shaping the monsoon circulation.

APPROACH

As part of this project, eight major research tasks are envisaged:

Task 1: Development of a coupled COAMPS-ROMS code using the ESMF NUOPC interface.
Task 2: Configuration of ROMS for the Indian Ocean and U.S. west coast.
Task 3: Configuration of COAMPS for the Indian Ocean and U.S. west coast.
Task 4: Interfacing ROMS with DART.
Task 5: Interfacing COAMPS with DART.
Task 6: Testing of the coupled model via a series of idealized problems.
Task 7: Configuration of the coupled COAMPS-ROMS system with DART.
Task 8: Preparation of observations from the Indian Ocean for data assimilation.

WORK COMPLETED

Task 1: Development of a coupled COAMPS-ROMS code using the ESMF

During year 2 of the project, NRL personnel established a repository for the coupled codes that can be shared amongst the project collaborators. A mirror of the repository has also been set up on a computer system at UC Santa Cruz to allow easier access to non-NRL project personnel.

The first prototype of a coupled ROMS-COAMPS system, with supporting data, within the ESPC-regional framework was also developed. This prototype uses the same coupling framework as the global Navy coupled model developed under the ESPC project, which will allow NRL in the future to couple the regional models (COAMPS and ROMS) directly within NRL global modeling infrastructure. At the present moment, NRL scientists are working on debugging the prototype system.

NRL also hosted an on-site visit by the ROMS and UCSC development teams to continue collaborative work on developing the interfaces necessary for the implementation of the coupled model. In addition, NRL developed a dataset of concurrent input files that allows testing of the developing coupled model against stand-alone versions, with a specific focus initially on a low resolution model of the US west coast for which existing ROMS and COAMPS configurations are available.

During year 2, Rutgers personnel worked extensively on the ROMS ESMF/NUOPC coupling framework. The ROMS coupling algorithms using the Earth System Modeling Framework (ESMF, version 7 or higher) with the National Unified Operational Prediction Capability (NUOPC) layer were re-written from scratch to facilitate compliance between coupling component interactions. The NUOPC layer is a simplified infrastructure on top of the ESMF library that provides conventions and templates to facilitate the easy coupling between Earth System Models (ESMs). The ESMF/NUOPC coupling algorithms in ROMS allows for both *driver* and *component* modes of operation. In the *driver* mode, ROMS provides all the interfaces needed to couple to other ESM components including the main driver, NUOPC-based generic ESM component services, model gridded components or NUOPC Model cap files, connectors between components for re-gridding source and destination fields, and input scripts and coupling metadata management. A NUOPC Model cap is a Fortran module layer that sits on top of the ESM component, making calls into it (*initialize*, *run*, and *finalize* phases). Alternatively, in the *component* mode, a NUOPC ROMS cap module is provided which can be adapted and incorporated into other NUOPC-based coupling systems.

Figure 1 shows a diagram for the ROMS native *driver* mode of operation including four ESM components: (i) Ocean Model, (ii) Atmosphere Model, (iii) Sea Ice model, and (iv) Wave Model. Obviously, the ocean gridded component is ROMS. Currently, three NUOPC Atmosphere Model cap modules have been coded including the Coupled Ocean-Atmosphere Mesoscale Prediction System (COAMPS, developed at the U.S. Naval Research Laboratory), the Regional Climate Model (RegCM Version 4.5; released by the International Center for Theoretical Physics, Trieste, Italy), and the Weather Research and Forecasting model (WRF; developed by a consortium of US institutions). The Sea Ice component is the Los Alamos sea ice model (CICE). The wave model components include the Combined Refraction/Diffraction Model (REF/DIF; develop at University of Delaware), the Simulating Waves Nearshore (SWAN, developed at Delft University of Technology), and the Wave Model (WAM, developed at the European Centre for Medium-Range Weather Forecasts).

For this project, we are only interested in the COAMPS-ROMS coupling. The NUOPC COAMPS cap file developed for the ROMS framework can be used in either the restricted and public (5.2.2) versions of COAMPS. We are currently debugging and testing this coupling interface.

The ROMS coupling framework via the ESMF/NUOPC library can be executed sequentially or concurrently. In sequential mode, all the coupled model components are executed on all of the specified Persistent Execution Threads (PETs). Alternatively, in concurrent mode, each coupled model component is executed on a non-overlapping set of PETs. The coupling type can be explicit or implicit. In explicit coupling, the exchange fields at the next time-step (n+1) are defined using known values from the time-step (n) before it. Explicit methods require less computational effort and are accurate for small coupling time windows. In implicit coupling, the exchange fields at the next time-step (n+1) are defined by including at the next time-step (n+1). Implicit methods are stable and allow longer coupling time-step but are expensive. Usually, the atmosphere-ocean coupling is semi-implicit: the exchange in one direction is explicit (ocean-to-atmosphere) and in the reverse direction is implicit (atmosphere-to-ocean). The REGRID interpolation method between source and destination fields can be bilinear or conservative. A nearest neighbor interpolation is possible between land/sea boundaries. Also, there is an option for two-step extrapolation support during regrinding.

The ESMF library does not provide a direct and elegant way for coupling nested grids. There are contributed extensions to the library that facilitate one-to-one nesting but it is messy. Since ROMS nesting is quite unique and complex, each nested grid is considered as separated components. We still need to test the nesting coupling capabilities in the future.

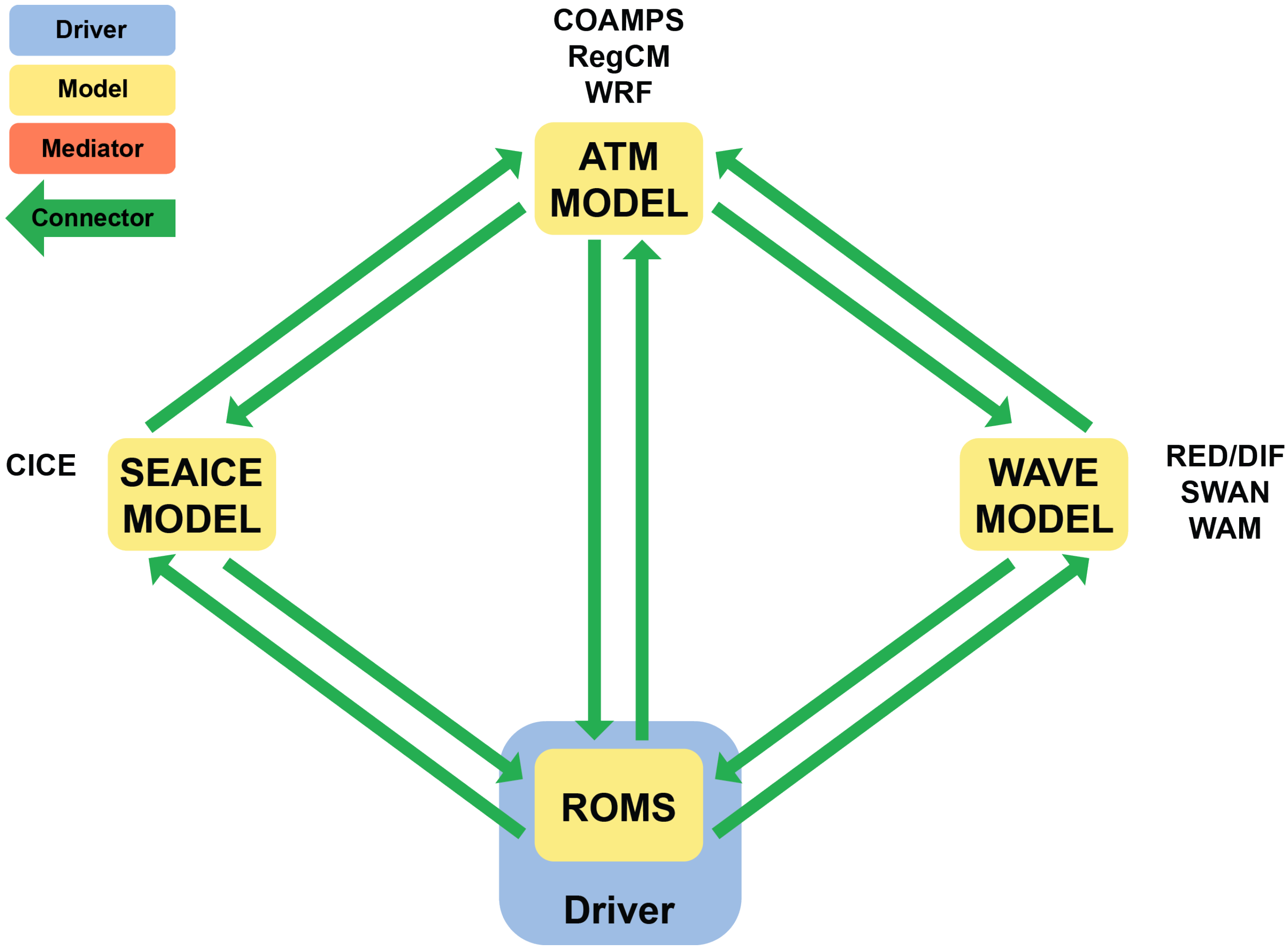

Figure 1: ROMS native ESMF/NUOPC coupling framework showing the Driver, the Model components, and the Connectors. The system includes four gridded components: Ocean Model (ROMS), Atmosphere Model (COAMPS, RegCM, or WRF), Sea Ice Model (CICE), and Wave Model (REF/DIF, SWAN, or WAM). This coupling system does not use the Mediators because custom coupling code (flux calculations, etc.) between multiple models is not needed.

Task 2: Configuration of ROMS for the Indian Ocean

During year 2, the ROMS 4-dimensional variational (4D-Var) data assimilation was implemented and calibrated using the most skillful Indian Ocean application identified in the previous report. The final framework adopted for 4D-Var was the dual-space algorithm which considerably reduces the computational cost when compared with the primal formulation of 4D-Var used previously. Special emphasis was placed on modeling of the background error covariance which demanded an extension of the original forward model integration from two to 10 years. The calibrated 4D-Var system was then used to produce an ocean analysis for the DYNAMO observational campaign period, and the skill of the analysis in hindcasting non-assimilated observations was evaluated.

In addition, Zavala-Garay visited PI Moore to work with the current implementation of DART in ROMS for the California Current System as the first step towards the implementation of DART for the Indian Ocean ROMS application.

Task 3: Configuration of COAMPS for the Indian Ocean and U.S. west coast.

The ROMS Indian Ocean grid is a subset of an existing configuration of COAMPS for the Indian Ocean. Similarly, the west coast ROMS and COAMPS grids overlap. These grids will form the foundation of the coupled model.

Task 4: Interfacing ROMS with DART

During year 2, development of the ROMS-DART interface has continued.

NCAR personnel continued to refine and extend the support and integration of ROMS and DART, particularly the support of the ROMS Verification Observation data format. In addition, the DART development branch was updated to track the enhancements provided by the Manhattan release of DART. (i.e., staying current). In addition, NCAR personnel assisted the UCSC group with parallel implementations of the DART ensemble generation using the SLURM job scheduler installed on one of the UCSC computer clusters.

At UCSC, a final suite of tests was performed to confirm the efficacy of the ROMS-DART interface. At the present time, a series of data assimilation experiments is underway comparing various configurations of the ROMS-DART Ensemble Adjustment Kalman Filter (EAKF) with the ROMS 4D-Var system in the California Current system. These experiments use a ROMS configuration with 10-km horizontal resolution (three times higher than that used in the original system tests), and ocean surface forcing fields derived from COAMPS. The data assimilation experiments will involve both identical twin experiments and real data experiments.

Task 5: Interfacing COAMPS with DART

NRL project personnel are reviewing the existing implementation for the DART-COAMPS interface. In collaboration with NCAR, the changes needed to bring DART-COAMPS into compliance with the latest release of DART ("Manahattan") are being addressed, and a site visit by NCAR team members to NRL will take place soon.

Task 6: Testing of the coupled model via a series of idealized problems.

At the present time, the coupled COAMPS-ROMS interface is being developed and tested using a low resolution version of a US west coast configuration for which existing ROMS and COAMPS grids are available.

Task 8: Preparation of observations from the Indian Ocean for data assimilation

This task has largely been completed.

RESULTS

4D-Var in the Indian Ocean

ROMS 4D-Var was implemented and calibrated using the most skillful Indian Ocean ROMS application, which is forced at the surface by the ERA-interim atmospheric analysis from the European Centre for Medium-range Weather Forecasts (ECMWF), and at the boundaries by the Hybrid Coordinate Ocean Model (HYCOM) analysis. Data streams being considered in 4D-Var are satellite sea surface temperature, satellite sea level anomalies, and the in-situ observations presented in Fig. 2.

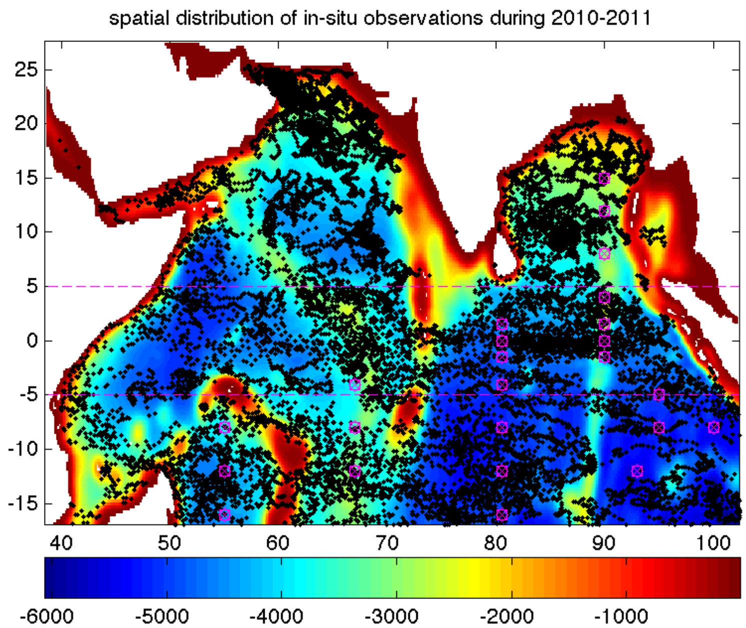

Figure 2: Model bathymetry (color) and spatial distribution of all the available in-situ observations during 2010 and 2011 (black dots). The magenta squares indicate the position of the RAMA moorings that recorded data during this period.

A crucial component of the 4D-Var setup is the adequate modeling of the background error covariance, and in particular the standard deviations used to penalize any increments to the initial conditions or boundary conditions and surface forcing. Early estimates of the standard deviations were based on just two years of model simulation (2011 and 2012) which resulted in a poor estimate of the corresponding background error covariance, and therefore a poor skill in the 4D-Var setup. Therefore, in order to improve this estimate, the length of the model simulation was extended to 10 years (2003-2012). This allowed us to obtain a better estimate of the seasonally-dependent Probability Density Function (PDF) of the model using the following approach. (i) The first two harmonics of the annual cycle were fitted to the 10 years of the modeled state vector, obtaining in this way a model for the seasonal cycle. (ii) The seasonal cycle was then removed from the modeled state vector and the resulting anomalies were used to construct a model for the seasonally-dependent standard deviations. This model was obtained by fitting the first two annual harmonics to the 3-month running standard deviation of the anomalies. As an example, Fig. 3 compares the model seasonal PDF (red lines) with the zonal current, temperature and salinity observed (blue lines) at a buoy from the Research Moored Array for the African-Asian-Australian Monsoon Analysis (RAMA) located at (0°,80.5° E). Similar comparisons at other depths and geographical locations suggested that the improved model PDF adequately captured most of the observed variability.

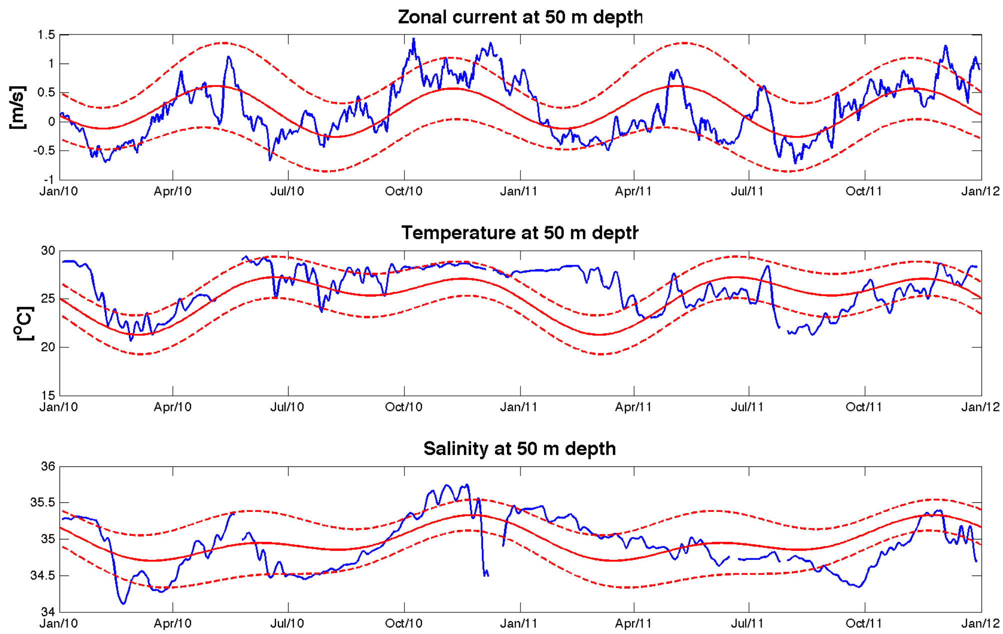

Figure 3: Comparison between the model seasonal PDF and the observed zonal current, temperature and salinity at (0°,80.5° E). The observed values are shown with blue lines while the model PDF is depicted as the seasonal mean (continuous red line) and the seasonal mean plus or minus 2 standard deviations (dashed red lines).

Once calibrated, the 4D-Var setup was used to produce an ocean analysis during the time period of the DYNAMO observational campaign (from Sept 2011 to January 2012). To evaluate the skill added by the 4D-Var we compared the model predictions with and without data assimilation with all the available observations shown in Fig. 2 in the latitudinal band 5°S-5°N. The skill of the model in hindcasting the observed variability was quantified by computing the correlation, root mean square (rms) and bias between the observed and modeled values at different depths, and the results are summarized in Fig. 4.

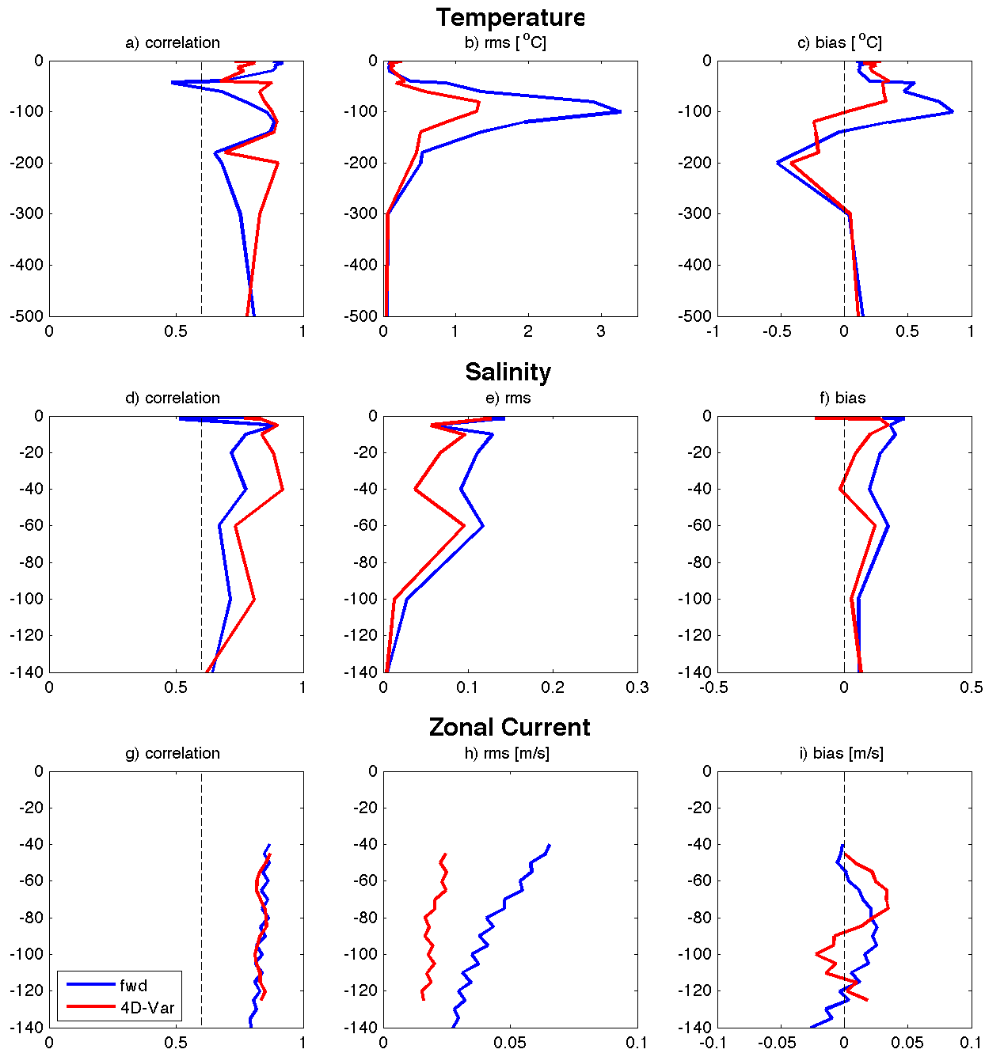

Figure 4: Comparison of the model skill with (red lines) and without (blue lines) data assimilation. The skill is quantified in terms of the correlation (a, d, g), rms difference (b, e, h) and bias (c, f, i) at different depths in the latitudinal band 5°S-5°N between non-assimilated observations and model predictions.

In general terms, for all the variables (temperature, salinity, and zonal current) the 4D-Var improved the ocean estimate by increasing the correlations at almost all depths (Figs. 4 a,d,g), reducing the rms of the difference between the model predictions and observations (Figs. 4 b,e,h), and modestly reducing the biases (Figs. 4 c,f,i). In particular, temperature anomalies around the thermocline (at approximately 100 m depth) are highly correlated with the SSH anomalies and therefore the assimilation of SSH resulted in a substantial reduction in the temperature rms errors around these depths as shown in Fig. 4b.

DART EAKF vs 4D-Var in the California Current

As noted above, a series of tests is currently underway in which circulation analyses of the California Current system computed from the DART-ROMS EAKF and the ROMS 4D-Var system are being compared. The cases run so far use actual ocean observations in the form of satellite SST, gridded Aviso SSH, and in situ observations of temperature and salinity from a variety of different platforms (e.g. XBTs, CTDs, Argo profiling floats). The EAKF was configured to assimilate observations once per day using the First-Guess at Appropriate Time (FGAT) approach, with localization and time-dependent covariance inflation. Here FGAT refers to the midnight on each day, although another time of day may be more appropriate. The period considered is 4 Jan – 15 Feb 2003. An ensemble size of 50 members was chosen, and the perturbations for the initial ensemble on 4 Jan 2003 were computed by drawing, at random, January analysis increments from a 31-year sequence of 4D-Var analyses for the same configuration of ROMS, excluding of course analysis increments for 2003. The same observations were assimilated into ROMS using 4D-Var and a 1-day assimilation window, but in this case the observations were assimilated at the actual observation times. In the 4D-Var experiments two outer-loops and seven inner-loops were employed.

A useful first-look metric for comparing the two sequences of analyses is the weighted squared difference between 1-day forecasts initialized from the analyses, and the observations collected during the forecast period before they are assimilated. Specifically, we consider $J_o=(\mathbf{y}^o-\mathbf{y}^m)^T\mathbf{R}^{-1}(\mathbf{y}^o-\mathbf{y}^m)$, where $\mathbf{y}^o$ denotes the vector of observations, $\mathbf{y}^m$ is the model forecast evaluated at the observation times and locations, and $\mathbf{R}$ is the observation error covariance matrix. In the case of the EAKF, $\mathbf{y}^m$ was computed from the mean of the corresponding $\mathbf{y}^m$ for each ensemble member, and can be viewed as the ensemble mean evaluated at the observation points. The metric J_o forms one component of the cost function that is minimized during 4D-Var. During both the EAKF and 4D-Var calculations, observations that yield innovations that depart by more than three standard deviations from the expected value were rejected from the analysis.

Figure 5a shows a time series of J_o for the 1-day forecasts from the ROMS-DART EAKF and 4D-Var experiments. In general, J_o is lower for the 4D-Var cases, although J_o computed for the EAKF also includes observations that were subsequently rejected from the analyses, so this will tend to increase J_o in this case. Figures 5b-5d show the rms difference between the 1-day forecast fields initialized from each analysis and the unassimilated observations. Errors in SST are generally a little smaller during the EAKF experiments (Fig. 5b), although subsurface temperature errors are smaller during the 4D-Var cases but with same caveat that the EAKF statistics include observations that were subsequently rejected. SSH and salinity errors are significantly smaller during the 4D-Var forecasts than those arising from the EAKF (Figs. 5c and 5d), although in the case of salinity some of the large peaks in EAKF forecast error (e.g. during Feb. 2003) are doubtless due to the inclusion of rejected salinity observations in the computation of the statistics. Figure 5 represents a first-look comparison, and these statistics will be recomputed taking into account observations rejected by the EAKF as the PIs continues to familiarize themselves with the EAKF diagnostic output.

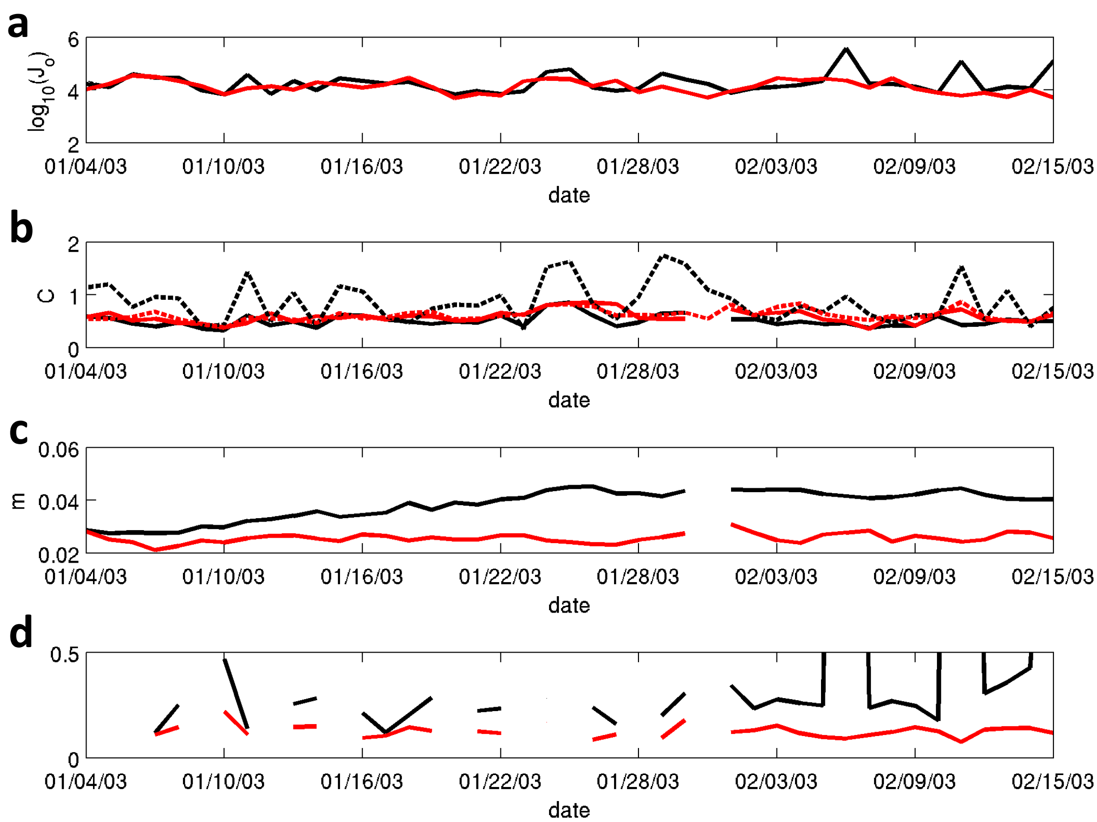

Figure 5: (a) Time series of J_o the EAKF ensemble mean (black) and 4D-Var (red) for 1-day forecasts initialized from each data assimilation estimate. The rms difference between the 1-day forecast fields and the observations for the EAKF ensemble mean (black) and 4D-Var (red) for (b) SST (solid lines) and in situ temperature (dashed lines), (c) SSH, and (d) in situ salinity.

PLANS FOR FY 2018

NRL-Monterey, Rutgers University, UCSC:
Development, testing and debugging of the ESMF-NUOPC interface for ROMS and COAMPS will continue, and the first tests of the coupled system will get underway.

NRL-Monterey:
Debugging and testing of the coupled model framework will be completed and the skill of a five-day coupled forecast compared to stand-alone models will be explored. In addition, the COAMPS-DART interface will be updated in accordance with the Manhattan release of DART. Scripting necessary to execute cycling of the COAMPS-DART system will also be developed.

Rutgers University:
We will continue with the implementation of the DART-EAKF in ROMS. Of particular interest is the comparison of the skill performance of the EAKF versus the 4D-Var, and the possibility of

combining them in a hybrid data assimilation scheme. We plan to document the Indian Ocean ROMS application and this comparison using all the available observations in a scientific manuscript to be submitted for publication.

NCAR:
NCAR project personnel will visit NRL to refine the COAMPS-DART system and develop run-time strategies for the fully coupled system for the target architectures.

UC Santa Cruz:
Analysis of the EAKF and 4D-Var comparison experiments for the ROMS California Current model will continue, and it is anticipated that preliminary results will be presented at the GODAE OceanView Data Assimilation Task Team (GOV DA-TT) workshop in Oct. 2017. UCSC project personnel will also assist in the testing of the coupled COAMPS-ROMS system as it becomes available, and provide assistance with ROMS-DART implementations in the Indian Ocean. The UCSC group will also transition to the Manhattan release of DART.

IMPACT/APPLICATIONS

This project will contribute significantly to the functionality and utility of COAMPS, ROMS, and DART which are all widely used and important community models, tools and resources.

TRANSITIONS

The new ROMS utilities developed as part of this project are already available from the ROMS web site http://myroms.org and can be actively used and further developed by other research groups in the U.S. and elsewhere as user competence increases. Similarly, the ROMS interface for DART is now freely available as part of the Manhattan release of the DART system. Improvements to the COAMPS systems will be transitioned through the Navy's 6.4 Small-Scale Modeling program.

RELATED PROJECTS

COAMPS and COAMPS-TC will be used in related 6.1 projects within PE 0601153N that include studies of tropical cyclone dynamics, air-ocean coupling, and boundary layer studies, and in related 6.2 projects within PE 0602435N that focus on the development of the atmospheric components (QC, analysis, initialization, and forecast model) of COAMPS and COAMPS-TC.

REFERENCES

None.