Integrating modeling and data assimilation using ROMS with a Coastal Ocean Observing System for the U.S. Middle Atlantic Bight

John Wilkin, Javier Zavala-Garay and Julia Levin

Institute of Marine and Coastal Sciences
Rutgers, the State University of New Jersey

Hydrodynamic models are used in coastal oceanography to simulate the circulation of limited-area domains for studies of regional ocean dynamics, biogeochemistry, geomorphology and ecosystem processes; for example, to deduce transport pathways for nutrients, sediments, pollutants or larvae. When operated as real-time now-cast or forecast systems, these models offer predictions that assist decision-making related to water quality and public health, coastal flooding, shipping, maritime safety, and other applications. Here we describe the configuration and operation of such a modeling system for the shelf waters of the Middle Atlantic Bight (MAB) – a region with a diversity of real-time models in operation and a dense *in situ* observational data set for assimilation and skill assessment.

The MAB spans the U.S. East Coast from Cape Hatteras, North Carolina, to Cape Cod, Massachusetts. A permanent front at the shelf-break separates relatively fresh and cool waters on the broad (~100 km wide) shelf from saltier, warmer Slope Sea water influenced by Gulf Stream eddies. This shelf-break front is prone to non-linear instabilities with wavelengths of order 40 km that evolve on time scales of a few days, and sustains along-shelf currents that reach the seafloor driving significant flow-bathymetry interactions. Appreciable across-shelf fluxes of heat, freshwater, nutrients, and carbon control water mass characteristics and impact ecosystem processes. The circulation is influenced by winds, tides, buoyancy input from rivers, a steady along-shelf sea level gradient, and mesoscale eddies that impinge upon the shelf edge. This spectrum of forcing mechanisms, and the dynamic shelf edge frontal zone, make the region a challenging laboratory for testing the skill of coastal ocean models and data assimilation methodologies.

The MAB is densely observed compared to coastal oceans globally, with much of the local data acquisition coordinated by the Mid-Atlantic Regional Association Coastal Ocean Observing System (MARACOOS; maracoos.org) – one element of a growing national network of regional observatories supported by ad hoc consortia of federal, state, academic and commercial partners. MARACOOS operates an extensive CODAR (Coastal Ocean Dynamics Applications Radar) network observing surface currents from the coast to the shelf edge, and deploys autonomous underwater glider vehicles (AUGV) to acquire subsurface temperature, salinity and biogeochemical data along transects throughout the MAB.

The Rutgers University Ocean Modeling Group (RU OMG!) has operated a real-time forecasting system for the MAB since September 2009, assimilating CODAR velocities, satellite sea surface height (SSH) and satellite surface temperature (SST) from multiple platforms. The assimilation of *in situ* temperature and salinity data from AUGV and ships of opportunity has been conducted experimentally. Further observations from National Marine Fisheries Ecosystem Monitoring (ECOMON) voyages and *in situ* data reported via the WMO GTS network are used in skill assessment. Climatological mean data analyses are utilized in a data-assimilative estimate of Mean Dynamic Topography that is combined with altimeter sea level anomalies.

The ocean circulation model is ROMS (Regional Ocean Modeling System; www.myroms.org), which solves the hydrostatic, Boussinesq, primitive equations in terrain-following vertical coordinates. Shchepetkin and McWilliams [2009] give a thorough review of the algorithmic elements that comprise the ROMS computational kernel. Features that make ROMS particularly attractive for continental shelf applications include: a formulation of the depth-integrated mode equations that suppresses aliasing into the slow 3-d baroclinic mode while retaining accuracy in representing barotropic motions due to tides and coastal-trapped waves; minimal pressure-gradient truncation error associated with the terrain following coordinates; a finite-volume, finite-time-step discretization for the tracer equations to improve conservation in shallow water applications where the free surface displacement is a significant fraction of the water depth; wetting and drying; and, optional positive-definite advection for biological tracers and sediment concentration. Companion applications include a sediment transport model and several ecosystem and bio-optical models, sea ice, and coupling to the SWAN wave model and WRF atmospheric model.

ROMS is unique in that three variants of 4-dimensional Variational (4DVar) data assimilation are supported in the code [*Moore et al.*, 2011a]: a primal formulation of incremental strong constraint 4DVar (I4DVAR), a dual formulation based on a physical-space statistical analysis system (4D-PSAS), and a dual formulation representer-based variant of 4DVar (R4DVar). In its general form, I4DVar seeks adjustments to the initial conditions, boundary conditions, and surface forcing in order to find the best estimate of the circulation in the least-square sense (model error can be also included in 4D-PSAS and R4DVar). In addition to 4DVar, other drivers in the ROMS suite utilize the adjoint and tangent linear models for formal analysis of observation impact and observational sensitivity [*Moore et al.*, 2011b].

The ROMS real-time data assimilative model for the MAB is termed ESPreSSO, for *Experimental System for Predicting Shelf and Slope Optics*. The horizontal resolution is 5 km, with 36 terrain following vertical levels. Open boundary data are adapted from the HYCOM-NCODA global ocean forecast system, and surface fluxes are derived from 12-km resolution 3-hourly forecast meteorology from the North American Mesoscale (NAM) model operated by NCEP. Observed daily-mean river discharges are applied at 7 sources. Harmonic tidal boundary conditions are taken from a regional model. The vertical turbulence closure is the *k-kl* option of the Generalized Length Scale (GLS) formulation [*Warner et al.*, 2005].

In ESPreSSO we use the I4DVAR formalism under the assumption that the surface and lateral boundary forcing are error-free and therefore the only control variables adjusted are the initial conditions of each analysis cycle, a procedure often referred to as “strong constraint” 4DVar. Daily, an I4DVAR analysis of 3 days of observations (CODAR, altimetry, SST) is performed to derive a now-cast of the ocean state – the *analysis*. From this analysis, a 72-hour forecast is launched using forecast meteorology and open boundary conditions, and persisted rivers. The first 24 hours of each forecast is retained as the “best estimate” of the ocean state for that day, with model output at 2 hour intervals added to an open access THREDDS (Thematic Real-time Environmental Distributed Data Services) server conforming to CDM (Common Data Model) standards. This facilitates inter-operability with MARACOOS partners, and ready access for remote users via OPeNDAP (Open-source project for a Network Data Access Protocol).

We perform substantial pre-processing and quality control of the data used for assimilation. For example, judicious application of the Jason-2 altimeter data error flags and revised wet tropospheric radar range correction [*Feng and Vandemark*, 2011] allows us to use these data in shallow water to within 10 km of the MAB coast. To altimeter SSH data we add Mean Dynamic Topography computed by a single 4DVar assimilation of an MAB region Ocean Climatological Hydrographic Analysis (MOCHA), mean wind stress, mean CODAR surface velocity, and mean currents from a decade of M/V Oleander ADCP data (on a New York to Bermuda transect).

To date, AUGV and ship of opportunity CTD data have largely been withheld for skill assessment, and do not yet enter the real-time data stream. In an inter-comparison of 7 ocean models that run in real-time and cover the MAB (3 global operational models, 3 MARACOOS models including ESPreSSO, and a coupled ROMS-SWAN-WRF model), only 3 of these models typically outperform a “prediction” based solely on climatology (MOCHA). With respect to model bias and centred root mean square error computed for 16 AUVG missions and 4 ECOMON cruises in 2010-2011, ESPreSSO is consistently among the three models with the lowest errors.

References

Feng, H., and D. Vandemark (2011), Altimeter Data Evaluation in the Coastal Gulf of Maine and Mid-Atlantic Bight Regions, *Marine Geodesy*, *34*, 340-363.

Moore, A. M., H. Arango, G. Broquet, B. Powell, A. T. Weaver, and J. Zavala-Garay (2011a), The Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilations systems, Part I - System overview and formulation, *Progress in Oceanography*, *91*(34-39).

Moore, A. M., H. Arango, G. Broquet, C. Edwards, M. Veneziani, B. Powell, D. Foley, J. Doyle, D. Costa, and P. Robinson (2011b), The Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilations systems, Part III - Observation impact and observation sensitivity in the California Current, *Progress in Oceanography*, *91*, 74-94.

Shchepetkin, A. F., and J. C. McWilliams (2009), Computational Kernel Algorithms for Fine-Scale, Multiprocess, Longtime Oceanic Simulations, *Handbook of Numerical Analysis, Computational Methods for the Atmosphere and the Oceans*, *14*, 121-183.

Warner, J., C. Sherwood, H. Arango, and R. Signell (2005), Performance of four turbulence closure models implemented using a generic length scale method, *Ocean Modelling*, *8*(1-2), 81-113.