CHAPTER 21

A Coastal Ocean Forecast System for the U.S. Mid-Atlantic Bight and Gulf of Maine

John Wilkin, Julia Levin, Alexander Lopez, Elias Hunter,
Javier Zavala-Garay, and Hernan Arango

*Department of Marine and Coastal Sciences, Rutgers, The State University of New Jersey,
New Brunswick, NJ, USA*

Coastal ocean models that downscale global operational models are widely used to study regional circulation at enhanced resolutions. When operated as nowcast/forecast systems, these models offer predictions that can provide actionable guidance for maritime applications. A nowcast/forecast system for the northeast U.S. coastal ocean is described in this chapter to illustrate, by example, the many practical issues to be considered when configuring such a model for operational oceanography applications. The system uses the Regional Ocean Modeling System (ROMS) and four-dimensional variational data assimilation of observations from a comprehensive network of in situ platforms, coastal radars, and satellites. The emergence of open access web data services that adhere to community conventions for metadata descriptions for coordinate systems and geo-scientific data types, and support geospatial search and sub-setting, are shown to foster inter-operability of data and model usage, accelerate the test, validate and acceptance cycle for modeling system enhancements, streamline the addition of new data streams, facilitate operational monitoring of the system, and enable novice users to view and download model outputs to underpin the generation of higher level ocean information products.

Introduction

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 nowcast or forecast systems, these models offer predictions that assist in 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 an ocean modeling nowcast and forecast system, with data assimilation, for the northeast North American coast extending from Cape Hatteras, North Carolina, northward to near Halifax on the Scotian Shelf of Canada. This domain encompasses two very different dynamical regimes in the Mid-Atlantic Bight (MAB) and the Gulf of Maine (GOM).

In the MAB (Cape Hatteras to Cape Cod, Massachusetts; Fig. 21.1), a permanent front at the shelf-break separates relatively fresh and cool waters on the broad (~100 km-wide) shelf from

Wilkin, J., et al., 2018: A coastal ocean forecast system for the U.S. Mid-Atlantic Bight and Gulf of Maine. In "*New Frontiers in Operational Oceanography*", E. Chassignet, A. Pascual, J. Tintoré, and J. Verron, Eds., GODAE OceanView, 593-624, doi:10.17125/gov2018.ch21.

saltier, warmer Slope Sea water (Mountain, 2003). This shelf-break front is prone to instabilities with wavelengths on the of order 40 km that evolve on timescales of a few days (Fratantoni and Pickart, 2003; Gawarkiewicz et al., 2004; Linder and Gawarkiewicz, 1998), and sustain along-shelf currents that reach the seafloor driving significant flow-bathymetry interactions. Eddy shelf interactions tied to Gulf Stream-induced warm core rings (Zhang and Gawarkiewicz, 2015) also lead to cross-shelf exchange with surface and sub-surface structure at scales of 10-30 km and days to weeks. Across-shelf fluxes of heat, freshwater, nutrients, and carbon control water mass characteristics and impact ecosystem processes in the MAB.

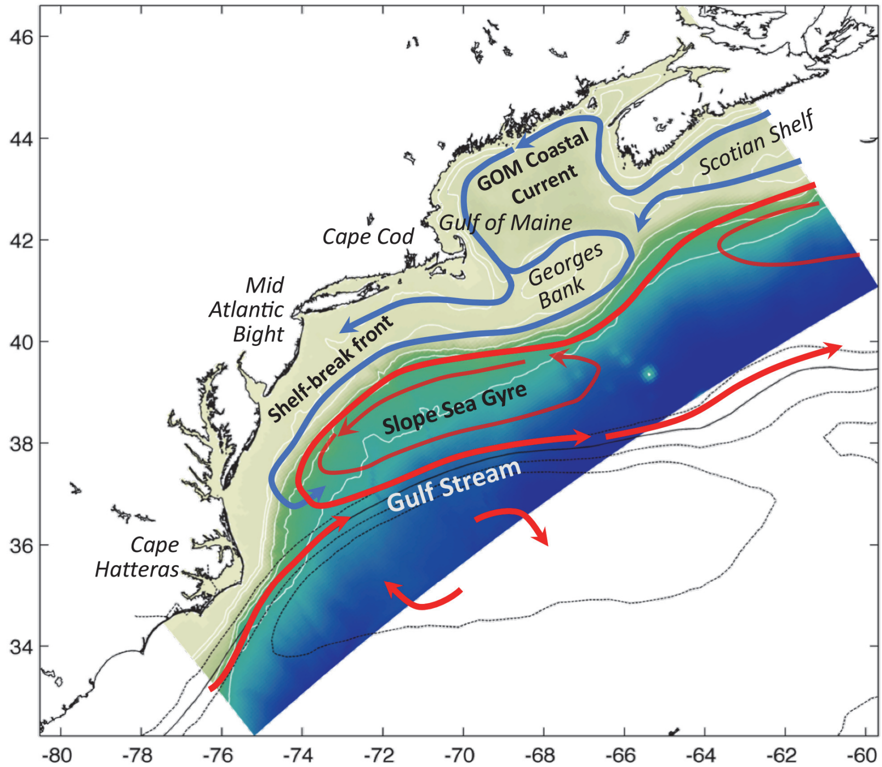

Figure 21.1. Bathymetry (colors) in the MARACOOS ROMS Doppio ocean model domain. Heavy lines and arrows show, schematically, the mean circulation of cool (blue) and warmer (red) currents. White lines show major bathymetric contours. Dotted black lines show major contours of the mean dynamic topography; solid black line is the north wall of the Gulf Stream.

The GOM is a relatively shallow, semi-enclosed marginal sea with local depth maxima of ~400 m in three distinct sub-basins. This basin topography exerts an influence on the overlying pattern of geostrophic flow. Water mass characteristics in the Gulf are determined by significant river inflows and multi-layer exchange flows through two narrow passages that flank Georges Bank. The region has famously strong tides that cause vigorous vertical mixing.

Therefore, all of the following processes influence circulation in this region: winds, tides, buoyancy input from rivers, air-sea heat fluxes, strong inflows from boundary currents at both the southern and northern extremities of the domain, and mesoscale Gulf Stream eddies that impinge upon the shelf edge. A large-scale, along-shelf pressure gradient is also significant in the along-shelf momentum budget (Csanady, 1976; Lentz, 2008).

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. However, a great advantage to studying and modeling this region is that the MAB and GOM are quite 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) and the Northeastern Regional Association for Coastal Ocean Observing Systems (NERACOOS; neracoos.org), both members of the U.S. Integrated Ocean Observing System (IOOS) network of coastal regional associations, which are ad hoc consortia of federal, state, academic and commercial partners. MARACOOS operations include an extensive CODAR (Coastal Ocean Dynamics Applications Radar) HF-radar network that observes surface currents from the coast to the shelf edge, and deployments of autonomous underwater glider vehicles to acquire subsurface temperature, salinity and biogeochemical data along transects throughout the MAB. NERACOOS operations emphasize a network of several telemetering buoys that observe atmospheric conditions and surface and subsurface ocean conditions, and prototype and sustained programs observing biogeochemical properties of the Gulf. The MAB is also home to the National Science Foundation Ocean Observatories Initiative’s (http://oceanobservatories.org) coastal Pioneer Array of seven profiling moorings and multiple deployments of autonomous vehicles, which provide very high-resolution observations in the vicinity of the MAB shelf-break front.

Traditionally, federal agencies have been the primary organizations implementing operational models in the U.S., while academic institutions have concentrated on process studies and model development and experimentation. A global example, operated by NOAA National Centers for Environmental Prediction (NCEP), is the Real-Time Ocean Forecasting System (RTOFS; http://polar.ncep.noaa.gov/global), a hybrid coordinate, 1/12° global ocean model that runs once a day and serves 3-hour interval forecasts of surface values and daily interval full-volume forecasts from the initial time out to 144 hours (six days). The NOAA Center for Operational Oceanographic Products and Services (CO-OPS) also operates forecast systems for several critical ports, harbors, bays, and estuaries on the U.S. coastline (https://tidesandcurrents.noaa.gov/models.html) with robust and automated model runs, quality checking, and forecast product generation.

But increasingly, non-federal regional association (RA) partners in IOOS are running ocean forecast systems. These systems might not meet federal requirements for operational robustness and reliability; nevertheless, many user communities find the immediate environmental information served by regional association models to be valuable (Wilkin et al., 2017). This may be because: the systems are superior in intrinsic skill having been carefully configured with local knowledge; they assimilate local datasets not adopted for operations by global centers; and they offer regional products, higher resolution, or local interpretive expertise that are not matched by larger domain systems.

Regional forecasting systems require open boundary condition information at the domain perimeter, which is typically drawn from global- or basin-scale model-based analysis systems, such as the products delivered by the Mercator-Océan system of the Copernicus Marine Environmental Monitoring Service (CMEMS) or the HYCOM system operated by the U.S. Naval Research Laboratory and NOAA. Both Mercator-Océan and HYCOM are elements in the international GODAE OceanView (Bell et al., 2015) program, developing and evaluating global ocean forecast systems. Regional modeling is thus a downscaling exercise, and a measure of success is whether the resulting ocean state analyses and forecasts achieve skill that exceeds that of the driving model in the same region. Other performance metrics will be based on the utility of valued-added products designed to meet specific stakeholder needs such as actionable guidance for marine operations, hazards, water quality, and regional ecosystems and fisheries.

The Rutgers University Ocean Modeling Group has operated a real-time forecasting system for the MARACOOS region since September 2009, assimilating in situ temperature and salinity data, CODAR velocities, satellite sea surface height (SSH) and satellite surface temperature (SST) for all available platforms. The system uses the Regional Ocean Modeling System (ROMS) model and its 4D-Var (four-dimensional variational) data assimilation system.

This article documents the set-up and operation of the MARACOOS ROMS forecast system, and is intended as a tutorial, by example, for others who might wish to emulate our effort to establish a similar downscaling model system in other coastal regions. The principal features of ROMS are described in the second section of this chapter. The third section outlines the particular configuration of ROMS for the MAB and GOM domains, including the information used for meteorological, river, and boundary forcing. The fourth section describes our choices in configuring 4D-Var for this application, the datasets assimilated, steps taken in the preparation and pre-processing of the observations, a discussion of practical issues regarding data access and monitoring the real-time system, and preliminary results of the performance of the prototype system still under development. We conclude with summary remarks on general lessons and challenges for the development and operation of models that enhance the value of observations through model-based synthesis and data assimilation to provide robust and reliable past, present, and forecasted ocean conditions.

Regional Ocean Modeling System – ROMS

Dynamical and numerical core

The ROMS computational kernel is described in detail elsewhere (Shchepetkin and McWilliams, 2005; Shchepetkin and McWilliams, 2009) and need not be reiterated here, but several aspects of the kernel are notable for the advantages they bring to coastal ocean and shelf sea simulation.

ROMS solves the hydrostatic, Boussinesq, Reynolds-averaged Navier-Stokes equations in terrain-following vertical coordinates. It employs a split-explicit formulation wherein the two-dimensional depth-integrated continuity and momentum equations are advanced using a much smaller time step than the three-dimensional baroclinic momentum and tracer equations. Time-weighted averaging of the barotropic mode prevents aliasing of unresolved signals into the slow baroclinic mode while accurately representing barotropic motions resolved by the baroclinic time step (e.g., tides and coastal-trapped waves). A formulation of the Equation of State and the density Jacobian is implemented that minimizes the pressure gradient truncation error that can be problematic in other terrain-following coordinate models. The terrain-following coordinate can be stretched vertically to better resolve surface and bottom boundary layers. Collectively, these features enhance the representation of friction, baroclinicity, and the vortex stretching of flow adjacent to steep bathymetry that are fundamental to steering sub-inertial frequency circulation in the coastal ocean and continental shelf-break region.

A finite-volume, finite-time-step discretization for the tracer equations improves integral conservation and constancy preservation properties associated with the variable free surface, which is important in coastal applications where the free surface displacement represents a significant fraction of the water depth.

By virtue of these many features, ROMS is a particularly attractive choice as a hydrodynamic modeling platform for achieving accurate, efficient, high-resolution ocean simulations of mesoscale processes in the coastal ocean and adjacent deep sea.

Vertical turbulent mixing closure

ROMS presents users with several options for how vertical eddy viscosity for momentum and eddy diffusivity for tracers are parameterized. The many coastal applications of the model have used them all: (i) a k-profile parameterization (KPP) in surface- and bottom-boundary layers (Large et al., 1994; Durski et al., 2004), (ii) Mellor-Yamada level 2.5 (MY25) (Mellor and Yamada, 1982), or (iii) the generic length-scale (GLS) method (Umlauf and Burchard, 2003), which includes several sub-options for closure and stability function.

The KPP scheme specifies turbulent mixing coefficients in the boundary layers based on Monin-Obukhov similarity theory, and in the interior principally as a function of the local gradient Richardson number (Large et al., 1994; Wijesekera et al., 2003). The KPP method is diagnostic in the sense it does not solve a time-evolving (prognostic) equation for any of the elements of the turbulent closure, whereas the MY25 and GLS schemes are of the general class of closures where two prognostic equations are solved – one for turbulent kinetic energy and the other related to turbulence length scale.

The implementation of GLS in ROMS is described by Warner et al. (2005), who also contrasted the performance of the various GLS sub-options and the historically widely used MY25 scheme. While the differing schemes lead to differences in the vertical eddy mixing profiles, the net impact on profiles of model state variables (velocities and tracers) is relatively minor. Similar conclusions were reached by Wijesekera et al. (2003). In the model set-up described here we use the GLS k-kl closure option, which is essentially an implementation of MY25 within the GLS conceptual framework. Our experience is that the model solutions are not particularly sensitive to the choice of sub-options within GLS, but the k-kl closure appears to be somewhat more stable in routine operations.

Two-way nesting

The ROMS code allows for one-way and two-way nesting for refinement grids (https://www.myroms.org/wiki/Nested_Grids) to provide increased resolution in specific regions, with future code developments planned to extend these capabilities to composite grids that only partially overlap, or are not aligned in their local grid coordinates.

The methodology for two-way nesting follows the same paradigm as the information exchange methodology used by ROMS to evaluate horizontal advection and diffusion operators across periodic boundaries or parallel subdomain partitions (in the MPI coarse-grained parallel execution option). In refinement nesting, a so-called "child receiver" grid obtains the information it needs to complete the high-order spatial stencils for the advection and diffusion operators surrounding the grid perimeter by having this information interpolated from the "parent donor" grid into a "contact region. This exchange is made on every model time-step, in both the predictor and corrector partial time steps. Lateral open boundary conditions are not applied since the extension of the numerical stencil into the contact region evaluates the full primitive equations horizontal operators at the perimeter of the receiver grid. This approach is preferable to providing donor grid information to the receiver strictly at the perimeter via open boundary conditions because open boundary conditions formulations are inevitably lacking in aspects of the ocean dynamics. Indeed, if the donor and receiver grids are identical (i.e., there is no refinement and all grid points coincide) then this approach is exact because it emulates the regular ROMS MPI parallel tiled domain decomposition (Warner et al., 2010).

This configuration can be strictly one-way (downscaling) with information flowing only from parent to child, or two-way. To achieve two-way nesting, i.e., including upscaling of information from the fine to coarse grid, the roles of donor and receiver are reversed in the complimentary step. The child grid becomes the donor; its solution is averaged to the parent grid resolution and replaces the parent solution where they overlap. Thus, the added physical realism achieved by increased resolution in the refined child grid is communicated back to the larger domain.

Variational data assimilation

ROMS supports a suite of differing implementations of 4D-Var data assimilation that are complemented by post-processing and analysis tools for a variety of applications. The system is described in detail by Moore et al. (2011a; 2011b; 2011c), and only a brief review of the important features will be presented here.

If we denote by **x** the state vector of the ocean (i.e., *T*, *S*, *u*, *v,* and *SSH*), the goal of 4D-Var is to identify the model initial conditions, surface forcing, and open boundary conditions (collectively referred to as the control vector **z**) that yield the "best" ocean circulation estimate, $\mathbf{x_a}$. The "best" circulation estimate is that associated with the **z** that minimizes a cost function given by $J_{NL} = (\mathbf{z}-\mathbf{z_b})^T \mathbf{B}^{-1}(\mathbf{z}-\mathbf{z_b}) + (\mathbf{y}-H(\mathbf{z}))^T \mathbf{R}^{-1}(\mathbf{y}-H(\mathbf{z}))$ where $\mathbf{z_b}$ is a *prior* or background estimate of the control vector, **y** is the vector of observations, **B** and **R** are respectively the background error and observation error covariance matrices, and *H* is the observation operator that maps **z** to the space-time observation locations. In the case of 4D-Var, the operator H includes the ROMS model and information is dynamically interpolated in time via the model dynamics. According to Bayes' theorem, the cost function J_{NL} corresponds to the logarithm of the *posterior* probability distribution of **z**, so finding the minimum of J_{NL} is equivalent to identifying the most likely ocean circulation state, described by $\mathbf{z_a}$, given the prior circulation estimate $\mathbf{z_b}$ and the observations **y**. The topology of J_{NL} may be very complicated so in general the minimum of J_{NL} is found using an iterative truncated Gauss-Newton method, which takes the form of a sequence of linear minimization problems. During each sequence of linear minimizations, the cost function that is actually minimized is given by:

$$J_k = \delta\mathbf{z}_k^T \mathbf{B}^{-1} \delta\mathbf{z}_k + (\mathbf{H}_{k-1}\delta\mathbf{z}_k - \mathbf{d})^T \mathbf{R}^{-1} (\mathbf{H}_{k-1}\delta\mathbf{z}_k - \mathbf{d}) \qquad (1)$$

where $\delta\mathbf{z}_k = \sum_{i=1}^{k-1} \delta\mathbf{z}_i = \mathbf{z}_k - \mathbf{z_b}$ represents the departure of the kth iterate from the background, $\mathbf{d} = (\mathbf{y} - H(\mathbf{z_b}))$ is referred to as the innovation vector, and $\mathbf{H}_{k-1}$ is the tangent linearization of the observation operator H linearized about $\mathbf{z}_{k-1}$. Linearization of the minimization problem in this way is referred to as the "incremental" approach (Courtier et al., 1994) and is used in ROMS 4D-Var. The primary workhorse algorithm that will be used here is the incremental, strong constraint, dual formulation of ROMS 4D-Var in which J is minimized directly in the space spanned by the observations.

The minimization of (1) proceeds via a Lanczos formulation of the restricted B-preconditioned conjugate gradient (CG) method (Gürol et al., 2014), with each iteration in the sequence of so-called inner loops identifying a new search direction in the control vector space that is orthogonal to prior search directions. When the $\delta\mathbf{z}_k$ that minimizes J_k has been identified, the estimate $\mathbf{z}_k$ about which **H** is linearized is updated (a so-called outer loop), and minimization of (1) proceeds again. The inner and outer loops are continued until J_{NL} is reduced to an acceptable level or until the iterative procedure has converged to the point where further reductions in are negligible.

Details of specific algorithmic choices regarding background error and observation error for our application are detailed in the fourth section.

ROMS configuration for MAB and GOM – Doppio

Doppio – A double espresso

The model configuration used here builds upon experience with an established model of the MAB region termed ESPreSSO (Experimental System for Predicting Shelf and Slope Optics; Zavala-Garay et al., 2014) that has underpinned numerous regional studies related to ecosystems (Hu et al., 2012; Xu et al., 2013), biogeochemical cycles (Mannino et al., 2016), sediment transport (Dalyander et al., 2013; Miles et al., 2015), storm-driven circulation (Miles et al., 2017; Seroka et al., 2017), and underwater acoustics (Lin et al., 2017), as examples. In a comparison of seven real-time models encompassing the MAB region (Wilkin and Hunter, 2013), no model was more skillful than data assimilative ESPreSSO in capturing MAB circulation.

Model resolution

The present domain, depicted in Fig. 21.1, is twice the size of the ESPreSSO grid – hence the moniker "Doppio" that we use to refer to the model. The grid has uniform horizontal resolution of 7 km. Finer resolution (at ~2 km, or less) would be desirable from the perspective of allowing the emergence of submesoscale variability, but the modest resolution greatly facilitates experimentation with 4D-Var assimilation. Given the effective resolution of CODAR data, along-track altimeter data, other satellite data, and the relatively sparse in situ observations, this resolution is adequate for the length scales that data assimilation might reasonably be expected to constrain.

Two-way nesting grid refinement offers the capability to refine model forecast resolution in selected regions using initial conditions drawn from the data assimilative analysis, through this is not currently an approach we implement. We are experimenting with two-way nested 4D-Var (which must apply nesting also in the adjoint and tangent linear models as they are iterated by the inner and outer loops of 4D-Var) as a rigorous approach to propagating the information content of local high-resolution observations through the grid hierarchy, such as the closely spaced moorings and multiple concurrent underwater vehicles operating as part of the Ocean Observatories Initiative Pioneer Coastal Array.

The model vertical resolution is 40 terrain-following levels, with stretching chosen such that throughout the continental shelf ocean the vertical resolution is finer than 1.5 m at the sea surface and better than 3 m in the bottom boundary layer.

Forcing

Air-sea fluxes of momentum and heat are computed from atmospheric conditions in the marine boundary layer (air pressure, temperature, relative humidity, and 10 m winds) using standard bulk formulae (Fairall et al., 2003). The SST predicted by ROMS is used in the calculation of out-going long-wave radiation, sensible heat exchange, and the transfer coefficients for boundary layer turbulence. In the forecast system, meteorological conditions are specified using 3-hour interval fields from the NCEP North American Mesoscale (NAM) weather forecast model. In longer retrospective simulations we use fields from the North American Regional Reanalysis (Mesinger et al. 2006). The net shortwave radiation is known to have a significant positive bias in these models (Kennedy et al., 2011) so is decreased by a factor of 25%. The diurnal cycle of net shortwave is poorly resolved by 3-hour interval data, so these are converted to daily average values that ROMS modulates internally with an idealized local diurnal cycle that is a function of longitude, latitude and year-day.

The direct influence of sea level atmospheric pressure gradient enters the momentum equations via the sea surface boundary condition to the hydrostatic balance, i.e. the model simulates its own dynamical response to the inverse barometer effect.

The NAM forecast we use is the CONUS (Continental U.S.) nest computed at 12 km resolution in a Lambert-conic projection, though the data we access via NOMADS (NOAA Operational Model Archive and Distribution System https://nomads.ncep.noaa.gov) is re-mapped to uniform 0.11°

longitude/latitude rectangular grid. The NAM forecast runs every six hours, but our ocean forecast only runs daily so we choose to access only the 00:00 UTC forecast cycle since this is reliably loaded on NOMADS by late evening local time.

The calculation of air-sea momentum flux (i.e., stress) in ROMS can be configured to account for the relative speed of the air and water by subtracting the ocean surface vector current from the 10 m boundary layer wind prior to applying the bulk formulae following Bye et al. (1979), but with no explicit account taken of the influence of Stokes drift associated with a wave field. This option was not active in the Doppio configuration for which we show results below, but following a systematic analysis showing it adds skill, the option is now standard in the latest prototype system.

Air-sea freshwater flux is set by the precipitation from NAM or North American Regional Reanalysis subtracted from the evaporation associated with the latent heat loss computed by the bulk formulae.

Open boundary conditions are always a vexing issue for regional ocean models. We draw information on the 3-D ocean state at the model open boundary from the Mercator-Océan system (Drévillon et al., 2008) using daily mean fields from the CMEMS PSY4QV3R1 Global Ocean Physics Analysis and Forecast. In the forecast window these fields are refreshed daily with each new Mercator-Océan cycle. An appealing feature of CMEMS is the provision of a consistent product suite spanning several years back into the past. This allowed us to compute a long-term mean of the PSY4QV3R1 results, which we adjust to remove moderate biases (most noticeable in the salinity of the inflow from the Labrador Sea that enters our model northern boundary) by replacing the Mercator-Océan mean with our own Mid-Atlantic region Ocean Climatological Hydrographic Analysis (MOCHA) annual mean (Fleming and Wilkin, 2010; Fleming, 2016). The MOCHA climatology is based on hydrographic observations from the World Ocean Database (WOD) (Boyer et al., 2009) augmented by CTD data from the NOAA Northeast Fisheries Science Center (NEFSC) and inner shelf CTD observations acquired by MARACOOS institutions, mapped to an equal angle 0.05° grid (~5 km) on 57 standard depths by Fleming (2016) using an adaptation of the weighted least squares method of Ridgway et al. (2002). Bias adjustments to mean velocity and mean dynamic topography (MDT) are made using dynamically balanced velocity and sea level consistent with MOCHA computed by the same climatological mean 4D-Var approach used by Levin et al. (2018) for ESPreSSO. These adjustments to Mercator-Océan conditions affect the long-term mean only; mesoscale variability is retained. The boundary data are used in ROMS via a combination of active/passive perimeter radiation and nudging (Marchesiello et al., 2001) and flow relaxation (Blayo and Debreu, 2006) in a "nudging zone" some 70 km wide around the perimeter.

The Mercator-Océan model does not include tides, so harmonic tidal sea level and current variability derived from a regional model (Mukai et al., 2002) is added to the depth-averaged velocity and sea level boundary data, and imposed on the dynamics using methods adapted from Flather (1976) and Chapman (1985) following Mason et al. (2010).

River inflows are important sources of buoyancy that contribute to coastal current dynamics and the salt budget throughout the GOM and MAB. While the continental U.S. and Canada are relatively well-instrumented with real-time monitoring stream gauges, still an appreciable portion of the watershed is un-gauged and it is necessary to account for this to capture the full measure of freshwater inflow. We have taken a statistical approach to rescaling real-time gauge data using results from an analysis of 11 years (2000-2010) of river flow and precipitation data, coupled with water balance and water transport models to infer the daily discharge from land to sea at 3 arc minute latitude and longitude resolution (Stewart et al., 2013). At this resolution there are 403 sources along the coast from Nova Scotia to Cape Hatteras that we aggregated into 22 major discharge sites coinciding with significant named rivers. Maximum covariance analysis was used to correlate the net river discharge with observations at stream gauges operated by the USGS and Water Service of Canada that reliably report in near real-time. This gave us a statistical model to infer the full discharge from the watershed based on the partial direct observations.

Free-running model performance

The set-up described above is our prior configuration for Doppio running freely as a "forward" model without data assimilation. Using the data sets to be described in the next section, and that we ultimately will use for data assimilation, we next present some results on the skill of this forward model configuration.

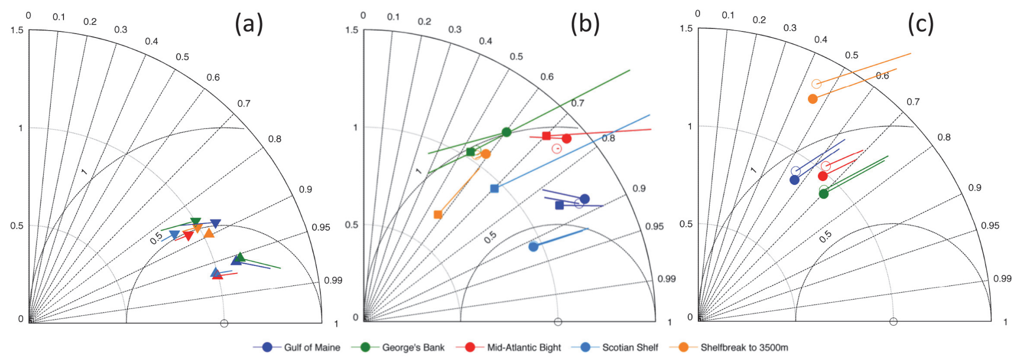

Figure 21.2. Taylor diagrams for model skill. Colors indicate sub-region of domain as shown in key. (a) Temperature at surface ▲ and sub-surface ▼ for control case. (b) Salinity for satellite precipitation forcing ●, HYCOM open boundary conditions □□, and control case (open circles). (c) Surface current skill for wind minus current surface stress calculation ●, and control case (open circles).

The skill of the free-running forward model, which we term the Doppio *control* case, is shown in a set of Taylor diagrams (Taylor, 2001) in Fig. 21.2. The radial distance to the symbols is the model standard deviation normalized by observation standard deviation, the azimuth is the arc cosine of the correlation, and the distance to the point (1,0) on the x-axis is therefore the normalized centered RMS error. Less conventionally for Taylor diagrams, we add sticks whose length is the normalized mean bias of the model; then the distance from the end of the stick to (1,0) is the overall normalized RMS error including bias. The closer to (1,0), the better the performance.

The number of satellite versus in situ observations of temperature differs by orders of magnitude, so surface and sub-surface skill is shown separately in Fig. 21.2a. Symbols are color-coded by sub-regions within the model domain. Skill is consistently higher for surface than sub-surface temperature, but we make little of this. Generally speaking, SST is a poor metric of model skill because it is so strongly constrained by the sea surface boundary condition imposed by air temperature – it's relatively easy to get SST right. Sub-surface temperature, on the other hand, is the consequence of vertical mixing and lateral transport processes acting over quite some time and distance and is a more demanding skill metric. We see that the sub-surface temperature skill of the control simulation is consistently high in all geographic sub-regions, and almost without bias.

Our model development and assessment framework makes use of the ROMS observation file format for 4D-Var assimilation (described more fully later on) through a convenient model option (VERIFICATION; https://www.myroms.org/blog/archives/128) that samples the model at the 4-D (space and time) locations of each datum in the unified observation file using the same observation operator, *H*, used by 4D-Var. When enabled, this option causes the model to compute, as it runs, a full suite of model minus observation statistics ready for analysis. This accelerates the experimentation cycle of model change, quantitative evaluation, and acceptance/rejection of configuration modifications.

To illustrate, Fig. 21.2b shows skill for in situ salinity for runs with different precipitation forcing data sets: the control case with rainfall from the NAM analysis (unfilled circles) and a run using rainfall derived from satellite (Huffman et al., 2007; filled circles). The two cases are almost indistinguishable, so we elected to retain the model-based precipitation data rather than further pursue use of satellite derived products.

Also shown in Fig. 21.2b (filled squares) is the salinity skill of a run using output from the GODAE HYCOM model for open boundary conditions. Centered RMS error is comparable in the two cases with the exception of a decrease in correlation on the Scotian Shelf. But when the bias is considered (the end of the sticks) there is a more noticeable decrease in skill for the case using HYCOM, hence our choice to retain Mercator-Océan for the control configuration.

Fig. 21.2c illustrates another configuration test, which is to take account of the relative speed of water and air in the bulk formula for surface stress. When this modification is applied (filled circles) there is a modest but consistent shift toward higher skill, and therefore this change is being adopted in future model simulations.

As a final comment on model performance we compare the along-isobath component of velocity between the model and two NERACOOS moored current meters in the GOM (Figure 21.3). The blue line is the squared coherency in the time series, with faint red lines indicating 95% confidence limits. Where the lower limit falls to zero, the coherence is not statistically significant and the blue line is erased. At the western site, within the GOM coastal current, the time series are coherent across all timescales. At the eastern site in the Northeast Channel, variability in the mesoscale band (20–50 days period) is not coherent, which encourages us that the assimilation of observations (which are ample to capture this timescale) from current meters and coastal altimetry has the potential to bring model variability into mesoscale event-wise agreement with the data. The coherence in variability at high frequencies at both mooring sites is likely testament that the local meteorological forcing, which dominates ocean surface current response at these time scales, is sufficiently accurate.

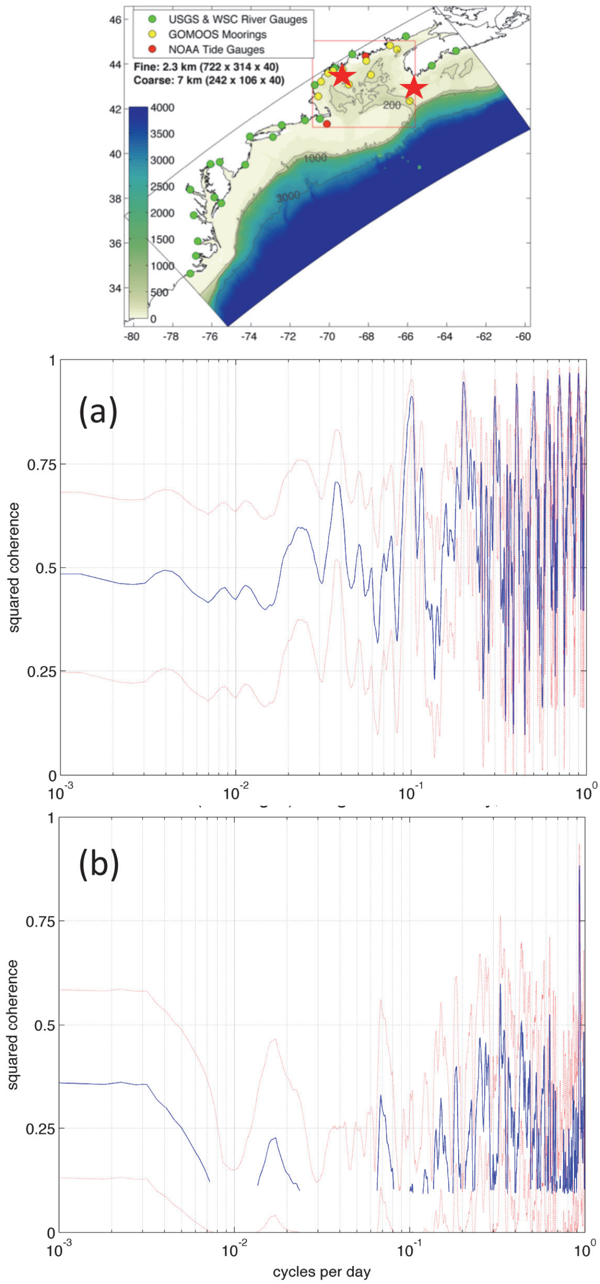

Figure 21.3. Squared coherency between along-isobath component of near-surface velocity and NERACOOS mooring data for (a) western site in GOM Coastal Current, and (b) eastern site in Northeast Channel entrance to GOM. Red stars in map show mooring locations.

Doppio Data Assimilation

4D-Var error hypothesis

4D-Var requires a prior estimate of the model background error covariance for initial, boundary and air-sea forcing conditions. ROMS specifies these as a univariate correlation matrix scaled by the square of standard deviations provided by the user. We computed background standard deviations from the mesoscale variability in a multi-year forward run (i.e., without data assimilation). The user also provides initial condition and boundary condition error de-correlation scales that determine the univariate correlation. In Doppio, we use 50 km in the horizontal, and 50 m in the vertical; surface forcing de-correlation scale is 100 km.

Observation errors are estimated from accepted practice with respect to the various observation platforms. For example, we would anticipate infrared satellite SST imagery to have the widely accepted expected error of 0.4 °C. Likewise, CODAR currents errors are nominally 0.1 m s^{-1}, but we inflate this by the mapping error returned by the optimal interpolation step that combines radial components into velocity vectors in CODAR data processing. CTD and current-meter errors for temperature and velocity, respectively, are expected to be substantially less than these. In practice, we find convergence of 4D-Var is aided by scaling these observation errors by the corresponding background error variance. This is an acknowledgment that our error model is imperfect, and most likely due to an incomplete account of representation error, which is the inability of the model to represent processes that are not captured by the model resolution or possibly aliased by the data sampling.

Observations for assimilation

Temperature and salinity

We have gone to extensive lengths to assemble the most comprehensive set possible of all observations of ROMS state variables – temperature, salinity, velocity and sea level – in the MAB and GOM. For in situ temperature and salinity, MARACOOS and NERACOOS are key providers on the MAB continental shelf shoreward of the shelf-break, and within the GOM. We augment these data streams with further in situ observations from the National Data Buoy Center, National Marine Fisheries Ecosystem Monitoring voyages, and data reported by in situ platforms of opportunity (Argo profiling floats, Expendable Bathythermographs [XBTs], drifting buoys, vessel underway thermo-salinograph). Contrasting the distribution of data accessible to us in near real time with the comparable data set assimilated in Mercator-Océan (Fig. 21.4) is readily apparent that our data assembly activity has very effectively harvested a wealth of information with the potential to significantly improve state estimation by data assimilation.

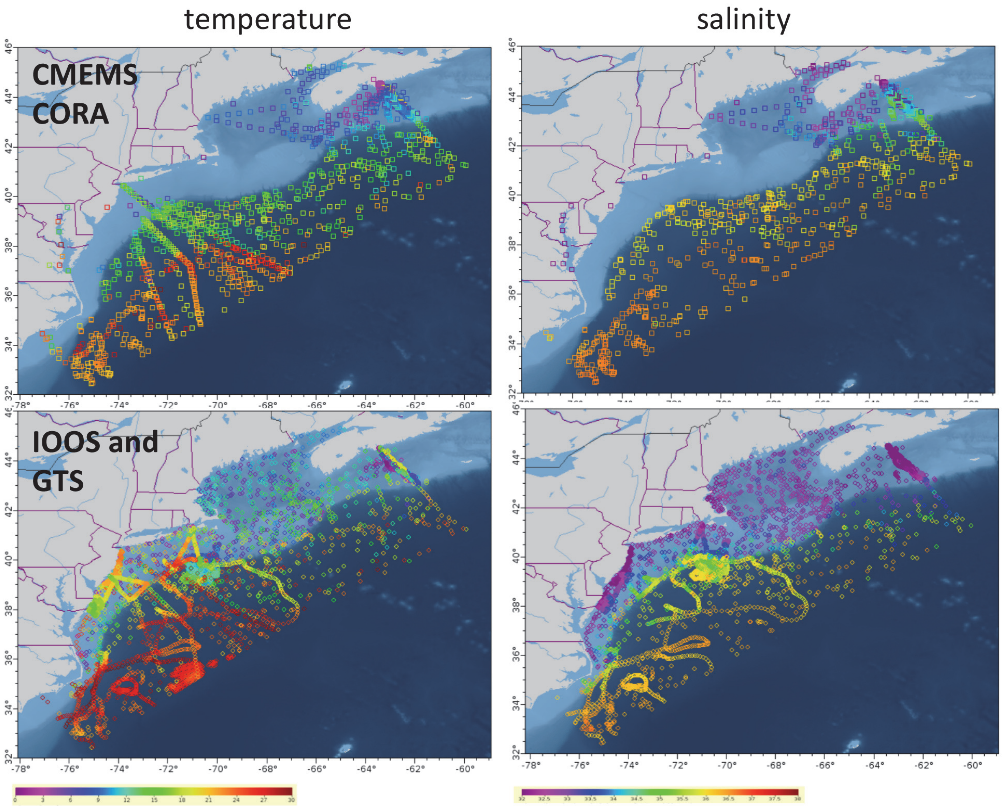

Figure 21.4. All sub-surface observations of (left) temperature and (right) salinity during 2015. Top row: Distribution of data in the CMEMS CORA database assimilated in Mercator-Océan. Bottom row: Distribution of data assembled for Doppio from U.S. IOOS data sources and the WMO Global Telecommunication System (GTS).

Though not presently available as near real time data streams, we have also expanded the data available in delayed mode (for retrospective reanalyses or skill assessment) by incorporating data from the eMOLT (Environmental Monitoring on Lobster Traps) project that returns bottom temperature data from sensors on lobster traps, water column and bottom temperature from the Northeast Cooperative Research Study Fleet Program derived from sensors mounted on fishing trawl doors, and water column temperatures (surface to 200 m depth) from sensor tags on loggerhead turtles that migrate throughout the estuaries of the MAB and across the continental shelf ocean. The wide geographic spread of these data is shown in Fig. 21.5.

The various observational assets we access, their approximate resolution, their latency, and the near real time sources from which we acquire the data operationally, are summarized in Table 21.1. From experience, we cannot overemphasize the value of more and varied sources of data in a coastal ocean analysis system.

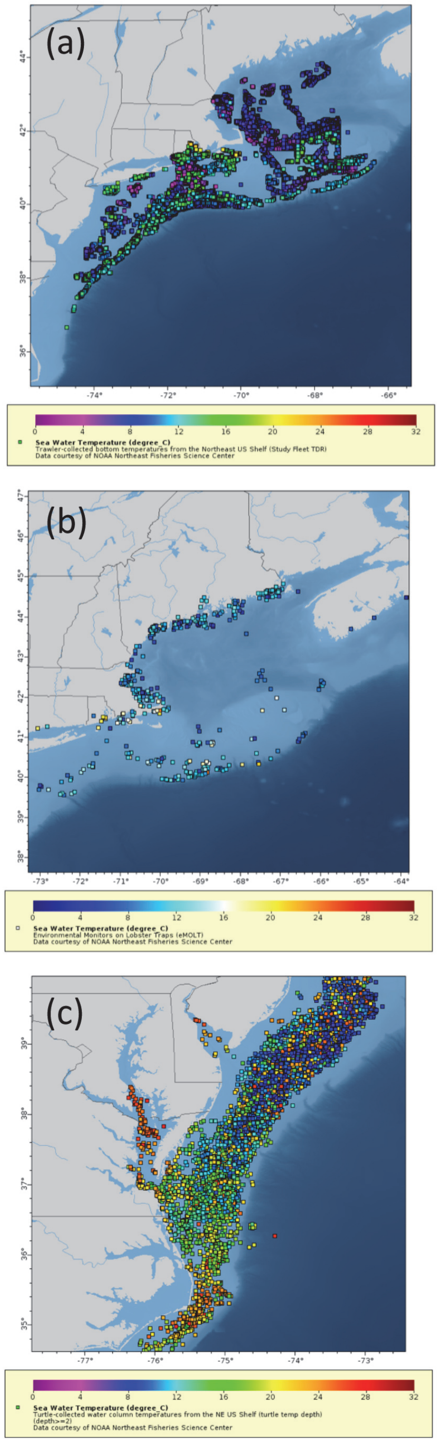

Figure 21.5. Observations of sub-surface temperature from sensors on novel observing platforms. (a) Northeast trawl fishing fleet. (b) Lobster traps. (c) Sea turtles. At publication these data are not accessible to the Doppio forecast system, but will soon become accessible in near real time through the NOAA Ocean Technology Transfer program for operational oceanography applications.

Observation type and platform	Source	Sampling frequency and resolution	Latency
AVHRR infrared SST	MARACOOS.org and NOAA CoastWatch	4 passes per day, 1 km	2 hr
GOES infrared SST	NOAA Coast-Watch	hourly, 6 km	12 hr
AMSR2 and WindSat microwave SST	NASA JPL PO-DAAC	daily, 15 km	24 hr
SSH: 4 satellite altimeters: Jason, AltiKa, CryoSat and Sentinel-3 with coastal corrections	Radar Altimeter Database System at TU Delft	~1 pass daily in do-main, ~4 km	4 hr
in situ T, S on GTS from National Data Buoy Center buoys, Argo floats, shipboard XBT, surface drift-ers	NOAA Observing System Monitor-ing Center (OSMC)	varies with platform	~12 hr
Surface currents from CODAR HF-radar	MARACOOS.org	hourly, 1km	4 hr
Glider T, S ~1-2 deployments per month in domain by MARACOOS	IOOS Glider Data Assembly Center	dense along trajectory	2 hr

Table 21.1. A summary of the observational data streams accessed for the ROMS Doppio near real time data assimilation system.

Where the observation sampling in space and/or time is higher than the model spatial resolution and model time step, observations should be combined to form "super-observations," a standard practice in data assimilation (Daley, 1991). Super-observations are data averages, weighted by inverse observation error, within chosen space and time bins. The formation of super-observations reduces data redundancy and poor conditioning of the cost function with respect to minimization.

The preparation of temperature and salinity data for ROMS 4D-Var from satellite and in situ platforms is relatively straightforward, as is their merger into super-observations. There are some challenges and subtleties, however, to the use of satellite altimeter sea level and HF-radar currents due to high frequency motions – principally tides. We detail these next.

Sea level and velocity

The Jason series of radar altimeter satellites measure sea surface height (SSH) along ~12 ground-tracks that traverse the Doppio domain with a 10-day repeat cycle. Adding in the other satellites in the altimeter constellation (presently CryoSat, AltiKa and Sentinel-3A), this coverage is complemented by a mix of different ground-track patterns and repeat cycles that combine to form a very comprehensive data set.

Historically, a great deal has been inferred about coastal ocean dynamics from the analysis of coastal sea level data from tide gauges, yet relatively little similar analysis has been conducted using altimetry. This is due in large measure to assertions that errors in altimeter data near the coast render the data unusable, yet significant progress has been made over the past decade in extending the validity of altimeter data to within a few kilometers of the coast by the appropriate application of altimeter radar range corrections and re-tracking of radar waveforms proximate to land (Cipollini et al., 2017; Vignudelli et al., 2011). This opens up to coastal oceanographers the opportunity to exploit the dynamical information content of so-called “coastal corrected altimeter” data.

In the MAB and GOM, altimeter data that would ordinarily be rejected by conventional quality control in coastal regimes can be reclaimed by judicious application of the data error flags and a revised wet tropospheric radar range correction (Feng and Vandemark, 2011). We extract 1 Hz along-track (approximately 6 km interval) Jason data from the Radar Altimeter Database System (rads.tudelft.nl) (Scharroo et al., 2013), making coastal corrections that retain data close to land (up to the 25-m isobath). These are to (i) use the European Centre for Medium-Range Weather Forecasts Wet Troposphere radar range correction in place of the onboard microwave radiometer correction that is contaminated by land within 50 km of the coast, and (ii) to ignore entirely the rain error flag that rejects numerous valid observations in the GOM. Data from CryoSat, AltiKa and Sentinel-3A are similarly downloaded and coastal-corrected for assimilation.

Our ROMS configuration simulates its own response to atmospheric pressure at the sea surface, so we do not make the dynamic atmosphere correction to altimeter range because that would be dynamically inconsistent and would inflate the model-data error.

In the coastal ocean where steep and variable bathymetry exacerbates uncertainty in the geoid and mean sea surface at short length scales (several tens of kilometers) there is an acute need to improve the precision of the mean dynamic topography (MDT) that is summed with altimeter sea level anomaly data to give an absolute dynamic topography (sea level above geoid) for assimilation that corresponds to the ROMS sea surface height prognostic variable.

Unfortunately, global MDT products such as the CNES-CLS13 MDT (Rio et al., 2014) (also generically referred to as “AVISO MDT” – Archiving, Validation and Interpretation of Satellite Oceanographic data – www.altimetry.fr) exhibit features in the MAB and GOM that oceanographers familiar with the locale recognize as unrealistic. These include contours of MDT strongly orthogonal to the coast that indicate landward geostrophic flow, some closed contours that imply isolated recirculation in shelf waters, and an intense boundary current adjacent to the coast of northern Virginia. Instead of AVISO MDT, we use a mean sea surface height computed by the climatological mean 4D-Var analysis (Levin et al., 2018) mentioned earlier in the context of bias correction of the open boundary condition data from Mercator-Océan. This approach has the added advantage that the mean sea level in the assimilated altimeter data is consistent with the mean dynamical balance of a free-running Doppio model.

MARACOOS CODAR systems began observing surface currents in the MAB in 2001, with the network reaching near complete coverage of the region by 2009. Radial component data from more than a dozen sites are gridded by optimal interpolation into a 6-km resolution vector velocity product with mapping error depending on the number, extent of overlap, and relative direction of the individual radial current observations (Roarty et al., 2010). In preparation for data assimilation, these data were further binned to 15 km resolution, but with velocities with large normalized optimal interpolation mapping errors ignored. The size of the bins was chosen to provide independent super-observations within the background error de-correlation scale.

Concern that small phase errors in the barotropic tide might dominate model-data misfit for sea level and velocity prompted us to apply pre-processing steps to reduce this possibility. Using harmonic analysis of long time series of CODAR data we de-tide those observations and replace the tidal variability with a signal computed from tidal analysis of a long free run of Doppio. Through this action it is our conjecture that the model-data misfit will not be dominated by tidal energy but instead will be due principally to dynamical responses that are within the scope to be modified by adjustment of the 4D-Var control variables. We adopt a similar approach in the pre-processing of altimeter sea level for assimilation, but where de-tiding of the observations is by application of the GOT4.10 harmonic tide altimeter range correction (Ray, 2013) in the extraction of data from the Radar Altimeter Database System.

This step is admittedly ad hoc, and has not been rigorously evaluated for its impact. There is the possibility that it carries little advantage because the tidal response in the free-running model is quite skillful.

A final point of detail in the handling of altimeter data is that early on in developing the prototype ESPreSSO system, when experimenting with assimilating satellite data only (SST and altimetry), we encountered a tendency for 4D-Var increments to excite unnaturally energetic surface gravity waves. Though unphysical to an oceanographer, these waves are nevertheless valid solutions to the ROMS governing equations. They can arise if not explicitly penalized in the cost function minimization. In the absence of subsurface temperature and salinity observations that would be inconsistent with simple gravity wave dynamics and discourage their excitation, and without the implementation of a thermal wind balance constraint (Weaver et al., 2005) on the multi-variate component of the error covariances, we adopted a simple approach aimed at suppressing the generation of these waves. This was to repeat the satellite SSH observations one hour prior to, and one hour following, the actual observation time. For these "pseudo-observations" the ROMS harmonic tidal sea level at the appropriate phase is added to the de-tided satellite value. Augmenting the observational data set in this way has the effect of encouraging 4D-Var to choose a solution more in accord with slowly varying sub-tidal sea level dynamics than with gravity waves excited by an impulse in the model-data misfit.

Practical operational cycle

Daily forcing and observation data ingest

Our near real time 4D-Var analysis system runs daily. Each evening (U.S. local time) a set of automated scripts runs to acquire forecast meteorology and open boundary data, and conduct the various pre-processing steps noted above. The stream gauge data are also acquired nightly, inflated by the maximum covariance analysis, and extrapolated into the forecast interval by persisting the last observation. The scripts that execute these tasks employ a mix of software tools including Matlab, Python, NCO toolbox, and Perl. The daily timeline of data gathering, processing, analysis, and output, is illustrated in Fig. 21.6.

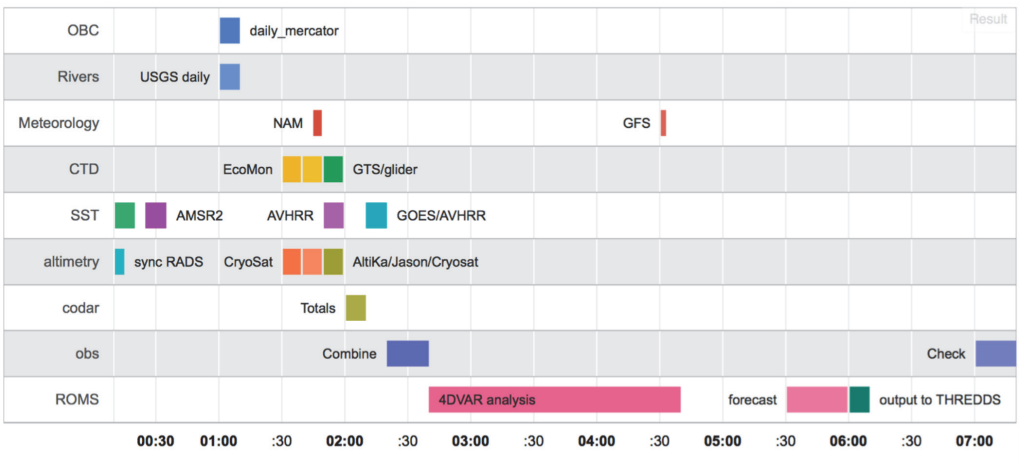

Figure 21.6. Schedule of data assembly steps that proceed daily via automated scripts to acquire open boundary conditions from Mercator-Océan, river discharge data from USGS and Water Service of Canada, metrological forcing data from NOAA NOMADS, in situ CTD data from numerous sources, satellite SST from infrared and microwave platforms, altimeter SSH (following local synchronization with the Radar Altimeter Database System database), and CODAR surface currents. The combine step forms super-observations. 4DVAR analysis takes approximately 155 minutes. All times are local U.S. Eastern Standard Time.

For some data, the latency from observation time to availability is predictable (e.g. polar orbiting satellites and HF-radar), while for others the schedule is more erratic. For these, remote data servers are polled a relatively short time in advance of the merger of common observation types into super-observations – effectively operating a "last chance" for inclusion strategy.

The data assimilation analysis step runs daily, but it uses three days of data. Data that are delayed additions to the database (more than 24 hours latency, but less than 60 hours) will therefore miss out on inclusion in the analysis at first, but could still subsequently enter the analysis on a following day. Therefore, all data acquisition queries are configured to request data for a full three days prior to analysis time.

Open data access and web services

The trend in the ocean science community toward providing access to data in a manner that follows the so-called FAIR data principles (findable, accessible, interoperable, reusable) (Wilkinson et al., 2016) has proven a great help to implementing the Doppio real-time forecast system. The vast majority of data we use are available via openly accessible web services such as THREDDS (Thematic Real-time Environmental Distributed Data Services) (Unidata 2018b) and ERDDAP (Simons 2018) and are formatted according to agreed conventions for metadata descriptions, e.g. the Climate-Forecast (CF) Conventions (http://cf-conventions.org) (Gregory, 2003). This convergence of standards makes it possible in many instances to re-use code with only minor modification to bring a new data set into the real-time data stream. Spatial and temporal sub-setting facilities in THREDDS and ERDDAP make software tools for data acquisition easily re-useable for new downscaling applications, requiring little more than the redefinition of the bounding box that encompasses the model domain.

Our experience is that searchable catalogs at data assembly centers such as NOAA CoastWatch (https://coastwatch.pfeg.noaa.gov/erddap/griddap) and the NASA Physical Oceanography Distributed Active Archive (PO.DAAC; https://podaac.jpl.nasa.gov) have made satellite data sets and services readily findable and accessible.

Near real time access to in situ data sets is less straightforward, especially coastal observations; applied coastal ocean modelers would enjoy greater access to existing data if there were more widespread embrace of the FAIR principles by the coastal ocean observing community. There is, however, a comprehensive near real time global aggregation of open-ocean in situ data that also encompasses many shelf regions, boundary currents and marginal seas. This is via the WMO Global Telecommunication System (GTS) that coordinates timely delivery of atmosphere and ocean observations to numerical weather prediction centers. By international convention (Resolution 40 of the Twelfth World Meteorological Congress, 1995), ocean physical observations (sea level, temperature, salinity and velocity) are considered to be meteorological data that may be acquired and exchanged without restriction due to maritime borders. Ocean observations in the GTS data stream are principally acquired beyond the continental shelf break using surface drifters, Argo profiling floats, or instrumentation mounted on or deployed by vessels participating in volunteer observing networks (XBT, XCTD, and underway thermo-salinograph). The data from shallow coastal waters that reach the GTS are predominantly from fixed moorings such as U.S. territorial sea observations from the NOAA data buoy network.

The CMEMS CORA (Cabanes et al., 2013) global data set of in situ ocean observations (CMEMS Product Identifier INSITU_GLO_TS_REP_OBSERVATIONS_013_001_b at https://marine.copernicus.eu) depicted in Fig. 21.4 that are assimilated in the Mercator-Océan model that Doppio uses for open boundary conditions comprises essentially the same data are as available in near real time via GTS, but with post-processing quality control and checking to reject bad profiles and apply delayed mode corrections.

Prior to launching the ROMS data assimilation step, the forcing and observation data assembly process (Fig. 21.6) concludes with the merger of observations of the same state variable into super-observations, for reasons noted previously. For satellite SST and SSH, super-observations are formed at the model grid resolution (here ~7 km) for each satellite pass. For in situ data, super-observations are grouped into 14 km spatial (2 model grid cells) and 12-minute time interval (2 model time steps) bins. These are then written to a single ROMS "obsfile" for 4D-Var analysis and assessment of the subsequent forecast, or to be used as verification data in freely running retrospective simulations for model experimentation such as presented earlier in the chapter. There are typically of order 200,000 independent super-observations in a three-day analysis interval.

Monitoring Doppio operational inputs

The daily process of assembling boundary condition and forcing inputs for Doppio, and aggregating observations for assimilation for skill assessment, requires continual monitoring to ensure continuity and integrity of the data streams. This is somewhat automated, since batch scripts and the ROMS model itself will register errors when web services are unavailable or file creation fails, but other failure modes are more subtle and require the vigilance of an operator to detect.

The ROMS forcing information (meteorology, river sources, etc.) netCDF file formats (Unidata, 2018c) are easily aggregated over time using THREDDS and made accessible in a graphical browse format using an ERDDAP Slide Sorter interface. Fig. 21.7 shows an example of a Slide Sorter web page configured to display the last day of NAM meteorology data, and the last three days of river discharge data. The date stamp on the NAM plot quickly reveals whether forecast data are present – which may not be so if NCEP services were unavailable when this job executed in the schedule (Fig. 21.6); the river discharge plot would display a warning that the query produced no matching results if the data acquisition failed because, for example, a gauge malfunctioned. The plots will reveal numerically valid but geophysical unreasonable data to an operator who takes time to browse the display. Any such adverse outcome would alert an operator that the Doppio system lacked valid inputs required to deliver a complete forecast.

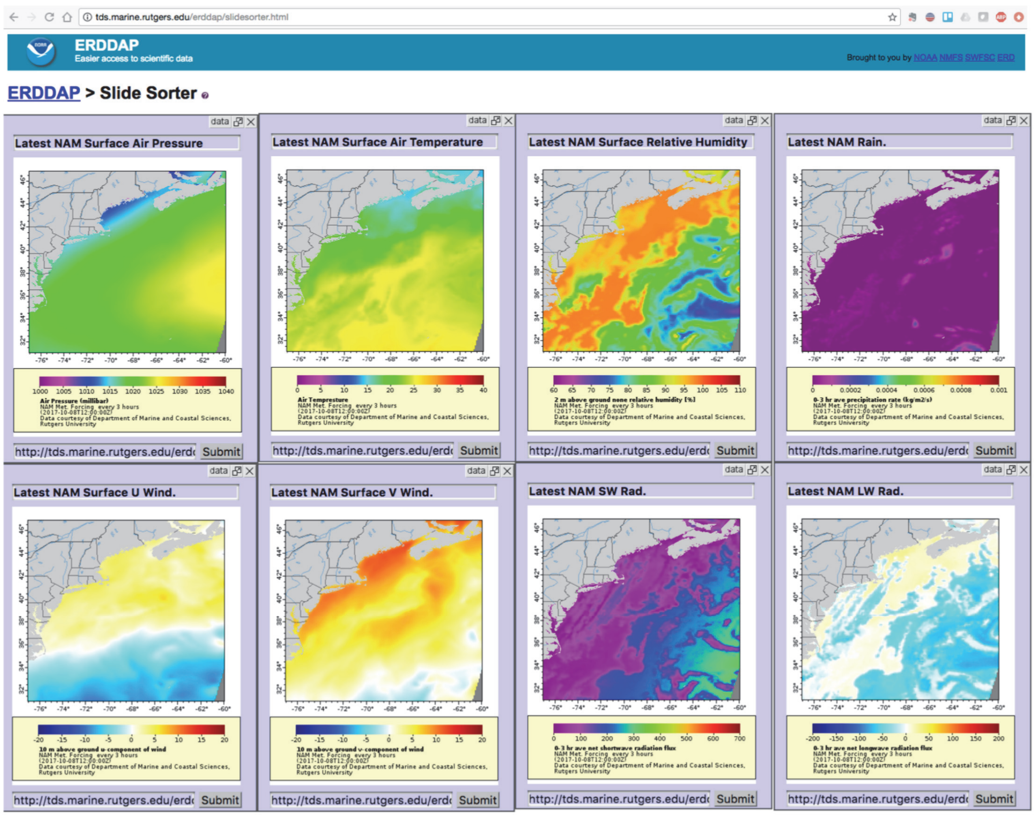

Figure 21.7. Partial view of the ERDDAP web interface that serves as an operator control panel to visualize inputs (meteorology and rivers) for the Doppio forecast system.

The ROMS netCDF *obsfile* format holds the observation geo-location and acquisition time, and also tracks the provenance of the observing platform and data provider. Making an aggregation of these files accessible with ERDDAP provides a convenient service to monitor the data entering the DA analysis. Fig. 21.8 shows an ERDDAP Slide Sorter web page configured to display subsets of the data for an example three-day analysis interval. In separate panels, user-defined optional constraints restrict views into the *obsfile* to highlight only subsurface observations, or certain provenance codes, to quickly reveal the presence or absence of anticipated data. Typical checks of the data stream we might use this system for include: If an autonomous underwater glider vehicle is known to have been recently deployed in the region it can be checked whether those data are entering the assimilation data stream via GTS; distinct provenance codes identify each altimeter satellite in the constellation, so it can be checked whether all anticipated data are flowing through the ground segment of each mission to the Radar Altimeter Database System; SST from different sensors and satellites can be visually browsed for consistency or noise that may be indicative of incomplete cloud-clearing.

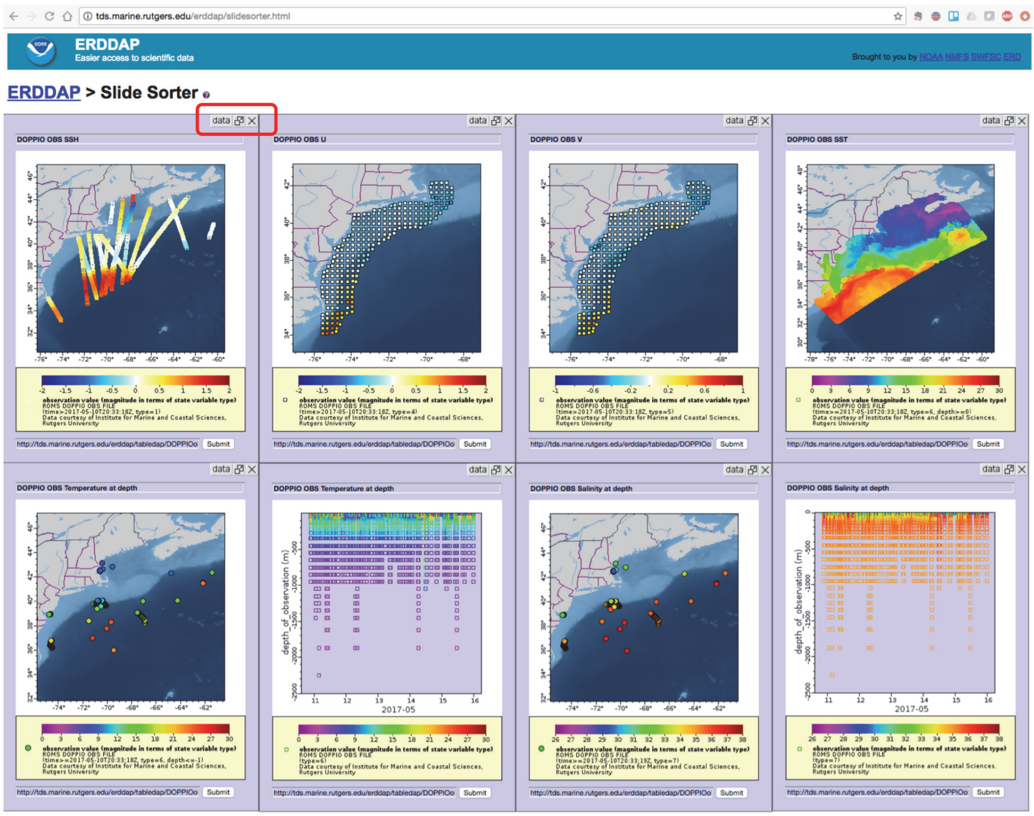

Figure 21.8. View of the ERDDAP web interface for browsing the data base of super-observations entering the Doppio data assimilation analysis. The example shows 6 days of data during May 2015. Top row, from left: altimeter SSH, u- and v-component of CODAR currents, and all satellite SST. Bottom row, from left: Sub-surface temperature as geo-locations and as a function of time and depth, all salinity as geo-locations and as a function of time and depth. The *data* button highlighted in the top left panel launches a data set browse page similar to Figure 21.9.

The Slide Sorter views are not previously created static images, but rather are generated anew from data in the *obsfile* when the page is loaded. An operator can modify the content displayed by selecting an individual browse plot to launch the underlying ERDDAP page in a new window. This allows customization of the geospatial search (i.e., zooming, or depth range constraints), constraining the display by data provenance, or modification of the image appearance (symbols, colors, etc.). This Slide Sorter view was easily configured to monitor near real time operational data streams for Doppio because ERDDAP accepts “now” as time query search, as in requesting “time>now-3days”.

A further useful feature of ERDDAP is the ability to immediately download the displayed data in formats readable by a wide range of scientific, GIS and spreadsheet software via a RESTful (Representational State Transfer) interface. This facilitates further analysis, such as the calculation of data statistics, and comparisons to independent observations or other models.

In an ERDDAP data view or download request the dataset identifier, variable name, search constraints and plot commands are fully described in the browser URL, so it can be bookmarked, shared or scripted via the UNIX “curl” command. These and many other features of ERDDAP are described by Simons (2018) and documented in web links included on every ERDDAP web page.

Analysis/forecast cycle – Doppio 4D-Var

Once the data ingest and preparation steps are complete, the ROMS 4D-Var sequence of inner and outer loops iterates toward an optimal solution. Upon convergence, the time varying 3-D ROMS solution through the three-day analysis window represents the maximum likelihood estimate of the ocean state during that interval. The conditions at the end of interval, notionally a “nowcast”, become the initial conditions for a single 72-hour forecast. The forecast horizon of 72 hours is set by the scope of the NAM meteorological forecast.

The 4D-Var analysis uses two outer and eight inner loops. Our experience with Doppio is that further iterations typically accomplish little in reducing the cost function. For this model grid of 240 by 104 horizontal points and 40 vertical levels, with 720 time steps in the three-day analysis interval, execution of the iterative 4D-Var analysis typically takes 155 minutes of walk clock time on 12 cores of a modest UNIX computer (3.5 GHz Intel Xeon processor). The subsequent forecast takes 15 minutes to complete on the same machine.

Operational system outputs

ROMS output files conform to CF-Conventions and the Common Data Model (CDM) API (Unidata, 2018a) that describes semantic layers for coordinate systems and scientific data types common in geophysics. Libraries for Python, Matlab, and many other scientific analysis software tools support the CF and CDM data models, and this standardization in output format greatly facilitates the inter-operability of the Doppio modeling system with partners in MARACOOS, IOOS and the broader user community when model outputs are made available via open access web services such as THREDDS.

We serve Doppio model output on the full ROMS 3-D grid at 1-hourly intervals of simulated time using the Forecast Model Run Collection (FMRC) facility in THREDDS. Every daily run generates three days of analysis and three days of forecast output, so there are multiple realizations of any given date. We choose to serve a FMRC “best time series” aggregation formulated as the concatenation of the central 24 hours of each three-day analysis, appended with the final day of the latest analysis and the forecast. With a data URL end-point that does not change, this FMRC “best time series” therefore provides users with a continuous, hourly, monotonic 3-D ocean state retrospective estimate, plus the forecast of the day. Users need only specify date and time in a data request, and FMRC will return an unambiguous result.

The “best time series” is the mostly widely used FMRC product, but FMRC preserves each individual cycle of analysis and forecast in its entirety and these are equally easily accessed as part of the THREDDS collection. Forecast skill can therefore be evaluated in posterior analyses targeted at appraising performance with respect to user-specific metrics, or against newly acquired delayed mode data that were not available to the data assimilation.

For the model grid dimensions noted above, 1-hourly interval 3-D model state snapshots and 24-hour average files add 7.2 Gb of output to the THREDDS collection each day. In addition, we add 6.1 Gb each day to an offline archive comprising the observation files and saved boundary conditions and forcing files for future re-runs and re-evaluations of the system.

The Doppio forecast is harvested by MARACOOS for ingest to the U.S. Coast Guard Environmental Data Server (EDS) that provides guidance to the USCG Search and Rescue Optimal Planning System (SAROPS), and by the National Marine Fisheries Service of NOAA for guidance on bottom temperatures in the MAB that strongly influence regional fisheries. MARACOOS also make views of the surface current and bottom temperature available through their *OceansMap* graphical browse service (http://oceansmap.maracoos.org), which allows easy qualitative comparison to many other observations and model-based products.

Supplementary special output products are three-day predictions of the drifting trajectories of mobile observing assets starting from their last reported position. These are computed for skill assessment versus the observed path of passive drifters, and to have advice ready in real time should an autonomous underwater glider vehicle become disabled and float passively at the surface; the prediction aids the mobilization of a vessel to rendezvous with a disabled vehicle for recovery. In addition, each day we predict drift trajectories originating at the locations of fixed moorings in the Ocean Observatories Initiative Pioneer Coastal Array in anticipation of the break out of a mooring; again, to provide guidance in mobilizing a recovery response.

Skill assessment – Prototype near real time Doppio output

The Doppio system described here was presented to students and lecturers at the GODAE International School on “New Frontiers in Operational Oceanography” in October 2017. At that time, the system operated in near real time on a “best effort” daily basis, generating ocean state analyses and 72-hour forecasts on most days, but accepting that occasional failures of either the data assimilation or forecast run are inevitable while in prototype.

Forecast failures typically stem from incomplete download of the meteorological forcing or river flow data – problems that can be identified but not remedied by the ERDDAP monitoring described earlier. Failures of the data assimilation analysis through instability of the 4D-Var iterations are rare, with most difficulties arising due to insufficiently quality-controlled data. Incomplete forecasts can also occur due to loss of power or network given that the system runs on a university computing infrastructure designed for research and education. Whatever the cause, once remedied the analysis and/or “forecast” (now actually a hindcast) are restarted so that the data can be added to the THREDDS FMRC catalog to complete best time series aggregation for later evaluation or applications.

The Doppio system could be hardened to forcing data interruptions by instituting redundant systems that enable a fail-over to alternative data streams or a fall back to a statistical model that combines climatology and persistence of prior valid data. The Mid-Atlantic Regional Association Coastal Observing Systems (MARACOOS) ROMS group may pursue such enhancements in the process of transitioning the Doppio system to near real-time sustained operation.

However, whatever steps are taken, vulnerabilities in the system will remain as long as there are dependencies on computing and cyberinfrastructure environments that have the potential to go down, and data streams that are operating on open research networks. Given these constraints, a system such as Doppio is unlikely to ever achieve "operational" status as would be defined by a national meteorological agency such as the National Centers for Environmental Protection (NCEP). Nevertheless, user communities exist for near real-time coastal ocean information products from research operators, whether they are model-based analyses or the underlying data stream themselves.

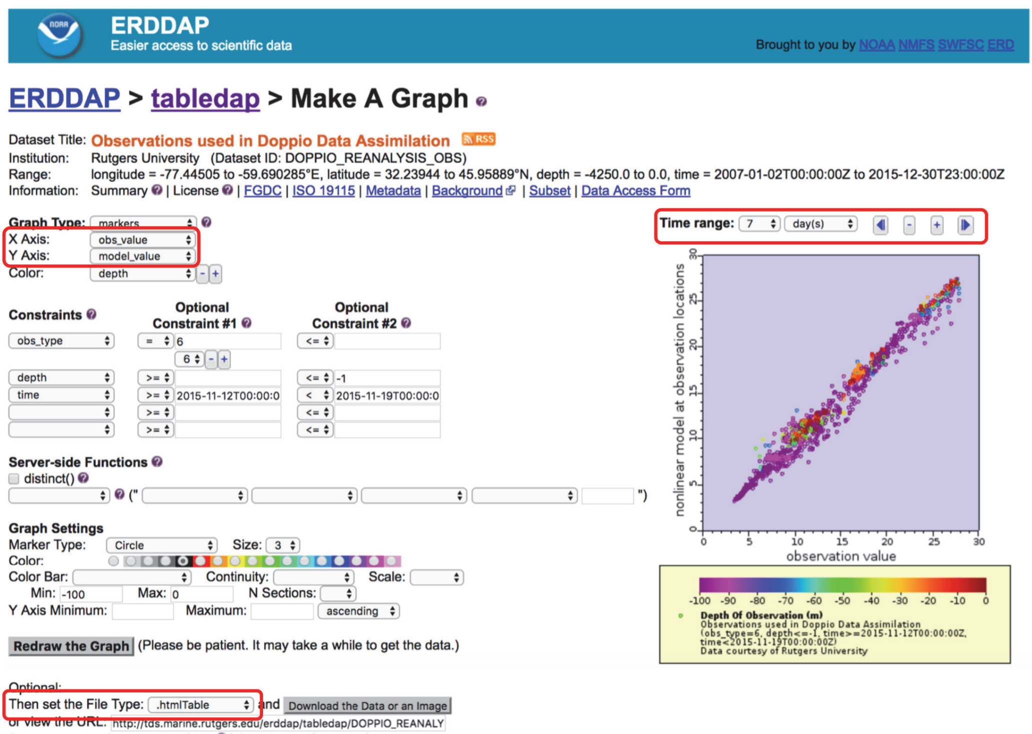

Figure 21.9. View of the ERDDAP web interface for browsing the data base of Doppio super-observations, but here set so as to display the ROMS 4DVAR analysis versus observation value as a scatter plot via the x-axis/y-axis controls (highlighted in the upper left red box). Also highlighted are the controls to select file type (lower left red box), where the RESTful interface offers netCDF, CSV, Matlab etc. download options, and the time range slider (upper right red box) which enables quick browse forward and backward through weekly time intervals.

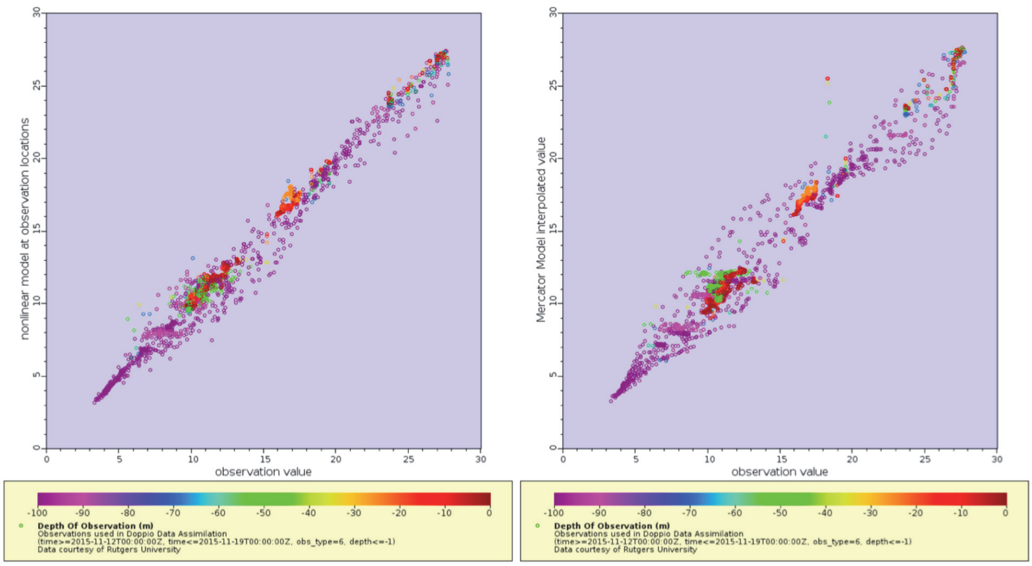

Figure 21.10. Model versus observed salinity comparison visualized with ERDDAP (from the FMRC best time series aggregation) for two 3-day analysis cycles in December 2015. Left: for Doppio ROMS. Right: for Mercator-Océan.

In a previous section, we showed that a free running Doppio model with bias-corrected open boundary condition data from Mercator-Océan retained useful skill throughout the model domain. Next we evaluate whether assimilation of the satellite and in situ data sets assembled for Doppio improves ocean state analyses compared to the free model. While this is to be expected – it is the principle of data assimilation to bring the analysis into agreement with data – what we wish to illustrate most is the ease with which the ROMS output formats and ERDDAP facilitate rapid browse and quantitative assessment of models and data for appraisal of the system in the context of operational oceanography.

It is fundamental to 4D-Var to minimize the innovation, this being the difference between the observations and the model state interpolated to the positions and times where the observations were made. In ROMS, this one-to-one match-up of model state to each observation is retained as output in a separate file. With ERDDAP, we virtually aggregate these two data sets (observations and corresponding ROMS estimate) and augment them with the independent Mercator-Océan analyses (without bias correction) also interpolated to the same observation locations. The merger of these products is accomplished in the back-end to ERDDAP; it does not require significant reformatting or re-writing of any of the ROMS output files.

It is then straightforward within ERDDAP to create scatter plots of observation value versus model, and to further focus the comparison using geospatial or time constraints or according to other metadata such as observing platform provenance. The RESTful interface enables download of the match-up data sets for further analysis. Some of these features are highlighted in Figure 21.9, which shows the ERDDAP interface one of our users would see.

Figure 21.10 shows examples of using this interface to quickly compare ROMS to observations and encode the plotted values according to observation depth or provenance. Comparison is also made to Mercator-Océan, from which it is evident that Doppio is much closer to the observations, though it must be recalled Mercator-Océan does not assimilate the vast majority of these data (Figure 21.4).

In closing, we note that the Doppio system as described here is still in prototype. Aspects of the system configuration may change before it enters sustained near real time operation, such as the background and observation error hypothesizes, data quality control and pre-processing practices, and the meteorological and river discharge inputs. The model and data browse interfaces described above are valuable for the way they accelerate the model prototype and update cycle by facilitating rapid qualitative and quantitative assessment of configuration changes and system performance within an operational oceanography environment.

Summary

We have described the configuration and operation of a modeling system that downscales output from a global GODAE forecast model in order to provide skillful estimates of ocean circulation in coastal, shelf and adjacent deep ocean waters of the northeast U.S. The Doppio modeling system is designed to assist the MARACOOS and NERACOOS Regional Associations of U.S. IOOS in the near real time delivery of coastal ocean information products in support of maritime safety, the marine economy, and the health and sustainable use of the coastal environment and coastal marine living resources.

Key elements of the system are (i) a regional ROMS model encompassing major estuaries and shallow coastal waters and extending far enough offshore to enable representation of oceanic mesoscale variability that drives coastal circulation, (ii) accurate surface meteorological forcing from the best available operational forecast of the national meteorological agency, (iii) attention to fully representing the buoyancy input from coastal river inflows, (iv) pre-processing to decrease biases in open boundary conditions by reference to a local, high resolution ocean hydrographic climatology, (v) assembly of a comprehensive suite of remote and in situ regional observations from all available platforms, and (vi) assimilation of these observations by 4D-Var.

Specifically, we have noted the essentials of our choices in configuring ROMS for the dynamical regime of the MAB and GOM region, and provided a brief overview of ROMS 4D-Var and how we have configured it for Doppio. At some length we have described the many observing platforms we access in near real time for assimilation prior to forecasting, and in delayed mode for further model evaluation. We have taken pains to detail many of the pre-processing steps required to adapt data streams typically utilized in mesoscale operational oceanography to the coastal environment. These details are documented for users who might wish to emulate our efforts and develop similar GODAE model downscaling system for other coastal regions globally.

To efficiently implement the many data ingest and model output steps that are part of Doppio, we make extensive use of web services that embrace community conventions for metadata descriptions and enable open access with geospatial searching and sub-setting capabilities, principally THREDDS and ERDDAP. When the providers of observations follow FAIR principles for serving their data they can be quickly incorporated into our near real time operations with minimal effort.

We have illustrated a number of instances of ERDDAP interfaces to observations and model outputs that enable rapid browse and data access and are readily customized for particular user purposes. We consider ERDDAP a transformational technology that through the versatility of the RESTful interface for data queries and download makes themed collections of observations and full model output data sets readily accessible to novice users with a minimum of training and little on-going intervention on the part of the data providers and model operators.

The data distribution plots in Fig. 21.4 reveal that, on this coastline at least, there exists a substantial data stream of observations that are available in near real time yet do not reach the global GODAE systems. It remains an open question the extent to which the skill of analyses from Mercator-Océan and similar systems might rival Doppio were they able to access this more extensive suite of observations. But even if this were so, there remains a strong case to undertake region specific downscaling with data assimilation endeavors such as Doppio in order to better link local coastal ocean knowledge to end users through communities such as the IOOS Regional Associations that can be more responsive to local stakeholder needs. A challenge for the existing Task Teams of GODAE, and a potential new frontier for a future GODAE School, is how outputs from the many near real time research systems like Doppio that are emerging globally might be communicated back to global GODAE and used to augment their skill locally through merged multi-model ensembles.

Acknowledgements

We thank Olmo Zavala, Peng Zhan, Xueming Zue and Misha Zujev for helpful comments on the manuscript. Grants NA11NOS0120038 and NA16NOS0120020 from the National Oceanic and Atmospheric Administration fund the implementation of ROMS for MARACOOS. Other grants from the Office of Naval Research, NASA, NSF and NOAA have contributed to elements of the Doppio system development, and to the ROMS model enterprise in general. The ROMS project benefits from the collaboration of the entire international ROMS user community at the www.myroms.org User Portal.

References

Bell, M., A. Schiller, P. Le Traon, N. SMith, E. Dombrowsky, and K. Wilmer-Becker, 2015: An introduction to GODAE OceanView. Journal of Operational Oceanography. Journal of Operational Oceanography, 8 (sup1), s2-11, doi:10.1080/1755876X.2015.1022041.

Blayo, E., and L. Debreu, 2006: Nesting Ocean Models. Ocean Weather Forecasting: An integrated view of oceanography, E. P. Chassignet, and J. Verron, Eds., Springer, 127-146.

Boyer, T. P., J. I. Antonov, O. K. Baranova, H. E. Garcia, D. R. Johnson, R. A. Locarnini, A. V. Mishonov, T. D. O'Brien, D. Seidov, I. V. Smolyar, and M. M. Zweng, 2009: World Ocean Database 2009, S. Levitus, Ed., NOAA Atlas NESDIS 66, U.S. Government Printing Office, Washington, D.C., 216 pp., DVDs.

Bye, J. A. T., R. A. Heath, and T. W. Sag, 1979: A Numerical Model of the Oceanic Circulation around New Zealand. Journal of Physical Oceanography, 9, 892-899.

Cabanes, C., A. Grouazel, K. von Schuckmann, M. Hamon, V. Turpin, C. Coatanoan, F. Paris, S. Guinehut, C. Boone, N. Ferry, C. de Boyer Montégut, T. Carval, G. Reverdin, S. Pouliquen, and P. Le Traon, 2013: The CORA dataset: validation and diagnostics of in-situ ocean temperature and salinity measurements. Ocean Science, 9, 1-18, doi:10.5194/os-9-1-2013.

Chapman, D., 1985: Numerical treatment of cross-shelf open boundaries in a barotropic coastal ocean model. Journal of Physical Oceanography, 15, 1060-1075.

Cipollini, P., J. Benveniste, F. Birol, J. Fernandes, E. Obligis, M. Passaro, P. T. Strub, G. Valladeau, S. Vignudelli, and J. Wilkin, 2017: Satellite Altimetry in Coastal Regions. Satellite altimetry over oceans and land surfaces, D. Stammer, and A. Cazenave, Eds., CRC Press, 343-380.

Courtier, P., J.-N. Thépaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4DVAR using an incremental approach. Quarterly Journal of the Royal Meteorological Society, 120, 1367-1388.

Csanady, G., 1976: Mean circulation in shallow seas. Journal of Geophysical Research, 81, 5389-5399.

Daley, R., 1991: Atmospheric Data Analysis. Cambridge University Press, 472 pp.

Dalyander, P. S., B. Butman, C. R. Sherwood, R. P. Signell, and J. Wilkin, 2013: Characterizing wave- and current--induced bottom shear stress: U.S. Middle Atlantic continental shelf. Continental Shelf Research, 52, 73-86, doi:10.1016/j.bbr.2011.03.031.

Drévillon, M., R. Bourdallé-Badie, C. Derval, J. M. Lellouche, E. Rémy, B. Tranchant, M. Benkiran, E. Greiner, S. Guinehut, N. Verbrugge, G. Garric, C. E. Testut, M. Laborie, L. Nouel, P. Bahurel, C. Bricaud, L. Crosnier, E. Dombrowsky, E. Durand, N. Ferry, F. Hernandez, O. Le Galloudec, F. Messal, and L. Parent, 2008: The GODAE/Mercator-Ocean global ocean forecasting system: results, applications and prospects. Journal of Operational Oceanography, 1, 51-57, doi:10.1080/1755876X.2008.11020095.

Fairall, C., E. Bradley, J. Hare, A. Grachev, and J. Edson, 2003: Bulk parameterization of air–sea fluxes: Updates and verification for the COARE algorithm. Journal of Climate, 16, 571-591.

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, doi:10.1080/01490419.2011.584828.

Flather, R. A., 1976: A tidal model of the northwest European continental shelf. Mem. Soc. Roy. Sci. Liege, Ser. 6, 10, 141-164.

Fleming, N., and J. Wilkin, 2010: MOCHA: A 3-D climatology of the temperature and salinity of the Middle Atlantic Bight. Eos, Trans. Amer. Geophys. Union, 91, Abstract PO35G-08.

Fleming, N. E., 2016: Seasonal and spatial variability in temperature, salinity and circulation of the Middle Atlantic Bight. PhD, 336 pp, Rutgers University, New Brunswick, NJ, doi:10.7282/T3XW4N4M.

Fratantoni, P., and R. Pickart, 2003: Variability of the shelf break jet in the Middle Atlantic Bight: Internally or externally forced. J. Geophys. Res, 108, 3166.

Gawarkiewicz, G., K. Brink, F. Bahr, R. Beardsley, M. Caruso, J. Lynch, and C. Chiu, 2004: A large-amplitude meander of the shelfbreak front during summer south of New England: observations from the Shelfbreak PRIMER experiment. Journal of Geophysical Research Oceans, 109, C03006, doi:10.1029/2002JC001468.

Gregory, J., 2003: The CF Metadata standard. CLIVAR Exchanges, 8, 4.

Gürol, S., A. T. Weaver, A. M. Moore, A. Piacentini, H. Arango, and S. Gratton, 2014: Preconditioned minimization algorithms for variational data assimilation with the dual formulation. Quarterly Journal of the Royal Meteorological Society, 140, 539-556, doi:10.1002/qj.2150.

Hu, J., K. Fennel, J. Mattern, and J. Wilkin, 2012: Data assimilation with a local Ensemble Kalman Filter applied to a three-dimensional biological model of the Middle Atlantic Bight. Journal of Marine Systems, 94, 145-156, doi:10.1016/j.jmarsys.2011.11.016.

Huffman, G. J., D. T. Bolvin, E. J. Nelkin, D. B. Wolff, R. F. Adler, G. Gu, Y. Hong, K. P. Bowman, and E. F. Stocker, 2007: The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. Journal of Hydrometeorology, 8, 38-55.

Kennedy, A. D., X. Dong, B. Xi, S. Xie, Y. Zhang, and J. Chen, 2011: A Comparison of MERRA and NARR Reanalyses with the DOE ARM SGP Data. Journal of Climate, 24, 4541-4557.

Large, W. G., J. C. McWilliams, and S. C. Doney, 1994: Oceanic vertical mixing: A review and a model with a nonlocal boundary layer parameterization. Reviews of Geophysics, 32, 363-403.

Lentz, S., 2008: Observations and a Model of the Mean Circulation over the Middle Atlantic Bight Continental Shelf. Journal of Physical Oceanography, 38, 1203-1221, doi:10.1175/2007JPO3768.1.

Levin, J., J. Wilkin, N. Fleming, and J. Zavala-Garay, 2018: Mean circulation of the Mid-Atlantic Bight from a climatological data assimilative model. Ocean Modelling, 128, 1-14, doi:10.1016/j.ocemod.2018.05.003.

Lin, Y.-T., A. Newhall, G. R. Potty, and J. H. Miller, 2017: A preliminary numerical model of three-dimensional underwater sound propagation in the Block Island Wind Farm area. Journal of the Acoustical Society of America, 141, 3933, doi:10.1121/1.4989146.

Linder, C., and G. Gawarkiewicz, 1998: A climatology of the shelfbreak front in the Middle Atlantic Bight. Journal of Geophysical Research, 103, 18,405-418,423.

Mannino, A., S. Signorini, M. Novak, J. Wilkin, M. Friedrichs, and R. Najjar, 2016: Dissolved organic carbon fluxes in the Middle Atlantic Bight: An integrated approach based on satellite data and ocean model products. Journal of Geophysical Research Biogeosciences, 121, doi:10.1002/2015JG003031.

Marchesiello, P., J. McWilliams, and A. Shchepetkin, 2001: Open boundary conditions for long-term integration of regional oceanic models. Ocean Modelling, 3, 1-20.

Mason, E., J. Molemaker, A. Shchepetkin, F. Colas, J. McWilliams, and P. Sangrà, 2010: Procedures for offline grid nesting in regional ocean models. Ocean Modelling, 35, 1-15.

Mellor, G. L., and T. Yamada, 1982: Development of a Turbulence Closure Model for Geophysical Fluid Problems. Reviews of Geophysics and Space Physics, 20, 851-875.

Mesinger, F., G. DiMego, E. Kalnay, K. Mitchell, P. E. Shafran, W, D. Jovic, J. Woollen, E. Rogers, E. E. Berbery, M, Y. Fan, R. Grumbine, W. Higgins, H. Li, Y. Lin, G. Manikin, D. Parrish, and W. Shi, 2006: North American Regional Reanalysis. Bulletin of the American Meteorological Society, 87, 343-360.

Miles, T., G. Seroka, and S. Glenn, 2017: Coastal ocean circulation during Hurricane Sandy. Journal of Geophysical Research, 122, 7095-7114, doi:10.1002/2017JC013031.

Miles, T., G. Seroka, J. Kohut, O. Schofield, and S. Glenn, 2015: Glider observations and modeling of sediment transport in Hurricane Sandy. Journal of Geophysical Research, 120, 1771-1791, doi:10.1002/2014JC010474.

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, 34-49.

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 II – Performance and application to the California Current System. Progress in Oceanography, 91, 50-73.

——, 2011c: 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.

Mountain, D., 2003: Variability in the properties of Shelf Water in the Middle Atlantic Bight, 1977–1999. Journal of Geophysical Research, 108, 3014.

Mukai, A. Y., J. J. Westerink, R. A. Luettich, and D. Mark, 2002: Eastcoast 2001, A tidal constituent database for the western North Atlantic, Gulf of Mexico and Caribbean Sea. Tech. Rep. ERDC/CHL TR-02-24, 196 pp.

Ray, R., 2013: Precise comparisons of bottom-pressure and altimetric ocean tides. Journal of Geophysical Research, 118, 4570-4584, doi:10.1002/jgrc.20336.

Ridgway, K. R., J. R. Dunn, and J. L. Wilkin, 2002: Ocean interpolation by 4-dimensional weighted least squares: Application to the waters around Australasia. Journal of Atmospheric and Oceanic Technology, 19, 1357-1375.

Rio, M.-H., S. Mulet, and N. Picot, 2014: Beyond GOCE for the ocean circulation estimate: Synergetic use of altimetry, gravimetry, and in situ data provides new insight into geostrophic and Ekman currents. Geophysical Research Letters, 41, 8918-8925, doi:10.1002/2014GL061773.

Roarty, H., S. Glenn, J. Kohut, D. Gong, E. Handel, E. Rivera Lemus, T. Garner, L. Atkinson, C. Jakubiak, W. Brown, M. Muglia, S. Haines, and H. Seim, 2010: Operation and application of a regional high frequency radar network in the Mid Atlantic Bight. Marine Technology Society Journal, 44, 133-145.

Scharroo, R., E. Leuliette, J. Lillibridge, D. Byrne, M. Naeije, and G. Mitchum, 2013: RADS: Consistent multi-mission products. in Proc. of the Symposium on 20 Years of Progress in Radar Altimetry, Venice, 20-28 September 2012, Eur. Space Agency Spec. Publ., ESA SP-710, 4 pp.

Seroka, G., T. Miles, Y. Xu, J. Kohut, O. Schofield, and S. Glenn, 2017: Rapid shelf-wide cooling response of a stratified coastal ocean to hurricanes. Journal of Geophysical Research, 122, 4845-4867, doi:10.1002/2017JC012756.

Shchepetkin, A., and J. McWilliams, 2005: The regional oceanic modeling system (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model. Ocean Modelling, 9, 347-404.

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, doi:10.1016/S1570-8659(08)01202-0.

Simons, R. A.: ERDDAP, NOAA/NMFS/SWFSC/ERD, Monterey, CA. [Available online at https://coastwatch.pfeg.noaa.gov/erddap/information.html.]

Stewart, R., W. Wollheim, A. Miara, C. Vorosmarty, B. Fekete, R. Lammers, and B. Rosenzweig, 2013: Horizontal cooling towers: riverine ecosystem services and the fate of thermoelectric heat in the contemporary Northeast US. Environmental Research Letters, 8, 025010, doi:10.1088/1748-9326/8/2/025010.

Taylor, K., 2001: Summarizing multiple aspects of model performance in a single diagram. Journal of Geophysical Research, 106, 7183-7192.

Umlauf, L., and H. Burchard, 2003: A generic length-scale equation for geophysical turbulence models. Journal of Marine Research, 61, 235-265.

Unidata, 2018a: Common Data Model (CDM) [software], UCAR/Unidata, Boulder, CO, https://www.unidata.ucar.edu/software/thredds/current/netcdf-java/CDM.

——, 2018b: Thematic Real-time Environmental Distributed Data Services (THREDDS) Data Server [software], UCAR/Unidata, Boulder, CO, doi:10.5065/D6N014KG, https://www.unidata.ucar.edu/software/thredds/current/tds/.

——, 2018c: Network Common Data Form (NetCDF) [software], UCAR/Unidata, Boulder, CO, doi:10.5065/D6H70CW6, https://www.unidata.ucar.edu/software/netcdf.

Vignudelli, S., A. Kostianoy, P. Cipollini, and J. Benveniste, 2011: Coastal Altimetry. Springer, 565 pp.

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, 81-113.

Warner, J. C., B. Armstrong, R. He, and J. B. Zambon, 2010: Development of a Coupled Ocean–Atmosphere–Wave–Sediment Transport (COAWST) Modeling System. Ocean Modelling, 35, 230-244, doi:210.1016/j.ocemod.2010.1007.1010.

Weaver, A., C. Deltel, E. Machu, S. Ricci, and N. Daget, 2005: A multivariate balance operator for variational ocean data assimilation. Quarterly Journal of the Royal Meteorological Society, 131, 3605-3625.

Wijesekera, H., J. Allen, and P. Newberger, 2003: Modeling study of turbulent mixing over the continental shelf: Comparison of turbulent closure schemes. Journal of Geophysical Research, 108, 3103.

Wilkin, J., and E. Hunter, 2013: An assessment of the skill of real-time models of Mid-Atlantic Bight continental shelf circulation. Journal of Geophysical Research, 118, 2919-2933, doi: 2910.1002/jgrc.20223.

Wilkin, J., L. Rosenfeld, A. Allen, R. Baltes, A. Baptista, R. He, P. Hogan, A. Kurapov, A. Mehra, J. Quintrell, D. Schwab, R. Signell, and J. Smith, 2017: Advancing coastal ocean modeling, analysis, and prediction for the U.S. Integrated Ocean Observing System. Journal of Operational Oceanography, doi:10.1080/1755876X.2017.1322026

Wilkinson, M., M. Dumontier, I. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J. Boiten, L. da Silva Santos, P. Bourne, J. Bouwman, A. Brookes, T. Clarke, and a. others, 2016: The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3, 160018, doi:10.1038/sdata.2016.18.

Xu, Y., B. Cahill, J. Wilkin, and O. Schofield, 2013: Role of wind in regulating phytoplankton blooms on the Mid-Atlantic Bight. Continental Shelf Research, 63, S26-S35, doi:10.1016/j.csr.2012.09.011.

Zavala-Garay, J., J. Wilkin, and J. Levin, 2014: Data assimilation in coastal oceanography: IS4DVAR in the Regional Ocean Modeling System (ROMS). Advanced Data Assimilation for Geosciences: Lecture Notes of the Les Houches School of Physics: Special issue June 2012, E. Blayo, M. Bocquet, E. Cosme, and L. Cugliandolo, Eds., Oxford University Press, 555-576, doi:10.1093/acprof:oso/9780198723844.003.0024.

Zhang, W., and G. Gawarkiewicz, 2015: Dynamics of the direct intrusion of Gulf Stream ring water onto the Mid-Atlantic Bight shelf. Geophysical Research Letters, 42, 7687-7695, doi:10.1002/2015GL065530.