

# **Variational assimilation of HF radar surface currents in a coastal ocean model off Oregon**

P. Yu, A. L. Kurapov, G. D. Egbert, J. S. Allen, P. M. Kosro

*College of Oceanic and Atmospheric Sciences,*

*Oregon State University, Corvallis, OR, 97331*

[pyu@coas.oregonstate.edu](mailto:pyu@coas.oregonstate.edu)

tel: (626) 696-5272

fax: (541) 737-2064

Jan. 17, 2012

## **Abstract**

The impact of assimilation of sea surface velocity fields observed by a set of high-frequency (HF) radars is studied using a three-dimensional ocean circulation model configured along the Oregon coast. The study period is June-July 2008 featuring upwelling and separation of the coastal currents into the adjacent interior ocean. The nonlinear model is based on the Regional Ocean Modeling System (ROMS) and the data assimilation (DA) component on the AVRORA system utilizing the representer-based variational algorithm. Assimilation proceeds in a series of 3-day windows, providing an analysis solution in each window and a 3-day forecast into the next window. Experiments with two different initial condition error covariances are compared (one is dynamically balanced, based on the linearized equation of state, temperature-salinity relation, and geostrophic and thermal wind balance relations and the other is multivariate unbalanced). While the assimilation impact is statistically better in the case of the balanced covariance, the case with the unbalanced covariance also provides sensible improvement in terms of surface velocity and sea surface temperature (SST) model-data forecast statistics. The analysis of representer functions shows that even if the initial condition error covariance is unbalanced, the correction fields at the model initial time are partially balanced after each dynamical field is smoothed independently, due to inherent dynamical properties of the adjoint model. Assimilation of the HF radar surface currents improves not only surface velocity forecasts, but also geometry of the upwelling SST front and the sea surface height (SSH) slope near the coast, as verified against unassimilated satellite SSH and SST data. The assimilation also alters the latitudinal distribution of the time-averaged offshore transport. Combined HF radar velocity and other observations, *e.g.*, altimetry, is needed to better constrain surface geostrophic currents in the entire model domain, including the area not covered by the HF radar data.

## 1. Introduction

Along the Oregon shelf, a strongly wind-driven upwelling system is present in summer (Huyer and Smith, 1978, Springer *et al.*, 2009, Koch *et al.*, 2010). A real-time data assimilation (DA) and forecast system has been implemented in that area as a part of a regional ocean observing system (<http://www-hce.coas.oregonstate.edu/~orcoss/ACTZ/SSCforecast.html>;<http://www.nanoos.org>). It has assimilated along-track sea surface height (SSH) from satellite altimetry, sea surface temperature (SST), and high-frequency (HF) radar surface current observations. In a series of hindcast studies we have assessed the impact of some of the assimilated data types, in particular by comparison with unassimilated data and by analyzing dynamical features revealed by the model-data synthesis. Kurapov *et al.* (2011) provides a comprehensive description of the assimilation system and demonstrates the effect of assimilation of along-track altimetry and SST data. In this study, we use the same DA system to investigate the impact of assimilation of HF radar surface currents.

Coastally-based HF radars (Kosro *et al.*, 1997) measure ocean surface currents at a spatial resolution of several kilometers, capturing essential processes in the coastal ocean, including wind-driven (Oke *et al.*, 2002a, Kosro, 2005) and tidal currents (Erofeeva *et al.*, 2003, Kurapov *et al.*, 2003, O’Keefe, 2005). Data from a combination of six long-range and five standard-range HF radars along the Oregon coast have been available in near-real time (<http://bragg.coas.oregonstate.edu/ORCoast/>) extending over a distance of 650 km in the alongshore direction and as far as 200 km offshore (Figure 1). The HF radar data cover both the shelf, where the dynamics are predominantly wind-driven, and the adjacent coastal transition zone (CTZ), where the dynamics are more dominated by nonlinear interactions of jets and eddies, fed by the coastal

current instabilities and separation (Strub *et al.*, 1991, Brink and Cowles, 1991, Koch *et al.*, 2010). Similar systems have been in operation elsewhere, including California and the US East Coast (Harlan *et al.*, 2010). An extensive surface current observation network such as the one off Oregon provides a unique opportunity to constrain model estimates of processes that determine shelf - interior ocean exchange.

Earlier studies have shown advantages of assimilating HF radar data (Lewis *et al.*, 1998, Breivik and Sætra, 2001, Oke *et al.* 2002a, Shulman and Paduan, 2009, Barth *et al.*, 2008, Wilkin *et al.*, 2005, Zhang *et al.*, 2010). To the best of our knowledge, there has not been a study showing a positive impact of surface velocity assimilation on the geometry of the SST front, which will be one of the points of this presentation. One of the features of our assimilation approach is the use of the four-dimensional variational (4DVAR) reparameter-based DA method (Bennett, 2002, Kurapov *et al.*, 2007, 2009, 2011). It is implemented in a series of time intervals, providing both time and space interpolation of the assimilated data sets in each window, using interpolation (covariance) structures that depend on the time-varying background solution.

## **2. The model**

The nonlinear forecast model component of the DA system is based on the Regional Ocean Modeling System (ROMS, version 3.2, [www.myroms.org](http://www.myroms.org)), a hydrostatic, Boussinesq, primitive equation ocean model with a free surface and a terrain-following vertical  $s$ -coordinate, featuring advanced numerics (Shchepetkin and McWilliams, 2005). The Mellor-Yamada Level-2.5 subgrid turbulence scheme is utilized in the vertical (Mellor and Yamada, 1982). Harmonic

horizontal tracer diffusion and momentum dissipation are implemented along  $s$ -surfaces with the diffusion and dissipation coefficients equal to  $10 \text{ m}^2/\text{s}$ .

The model domain extends from  $40.5^\circ\text{N}$  to  $47.5^\circ\text{N}$ , covering the entire Oregon coast and parts of the Washington and California coasts, and from  $123.7^\circ\text{W}$  to  $129^\circ\text{W}$ , extending more than 300 km offshore (see Figure 1). The regular horizontal grid is defined in spherical polar coordinates. The resolution is approximately 6-km in each horizontal direction with 15 levels in the vertical. The minimum water depth along the coast is set at 40 m. The resolution of this model is modest compared to those used in our recent process-oriented studies of circulation off Oregon [3-km (Springer *et al.*, 2009, Koch *et al.*, 2010) and 1-km (Osborne *et al.*, 2011)] to facilitate faster turnout of variational DA runs (which can be on the order of 10-100 times as expensive in terms of computational time as a single nonlinear model run).

No-normal-flow and free-slip boundary conditions apply to velocities along the coast. The other three boundaries (north, south, and west) are open with conditions provided by the daily-averaged outputs of the regional Navy Coastal Ocean Model of the California Current System (NCOM CCS; Shulman *et al.*, 2004). NCOM had a 9-10 km resolution in the horizontal and utilized hybrid coordinates in the vertical (with 19 terrain-following layers in the upper 138 m and  $z$ -coordinates below). Atmospheric forcing for NCOM was obtained from the Coupled Ocean-Atmosphere Mesoscale Prediction System (COAMPS, Doyle *et al.*, 2008). Differences in our free-run ROMS solutions and NCOM can in part be explained by the different atmospheric forcing and vertical discretization. Note that NCOM-CCS did assimilate data through a multi-stage process involving empirical three-dimensional (3D) projection of satellite SSH and SST maps using archives of temperature and salinity (Fox *et al.*, 2002) and model nudging.

The bulk flux formulation (Fairall *et al.*, 1996) is used to calculate the surface momentum and heat fluxes in ROMS. Atmospheric fields, including wind speed, air temperature, relative humidity, and atmospheric pressure, are obtained from the 12-km resolution North America Mesoscale Model (NAM) forecast archives (<http://nomads.ncdc.noaa.gov/data.php>). A 40-hour low-pass filter was applied to the atmospheric parameters used for surface forcing computation, in part for consistency with the available boundary data time series. Filtering the wind speed can reduce the peak stress values. While this is admittedly a shortcoming in our formulation, no model bias was found compared to the surface velocity data (Figure 2).

Although the tangent linear (TL) and adjoint (ADJ) components are now included in the standard ROMS distribution (Moore *et al.*, 2011), in this application we have used our own, stand-alone TL and ADJ codes AVRORA [Advanced Variational Regional Ocean Representer Analyzer, see (Kurapov *et al.*, 2009, 2011)], which are dynamically and algorithmically consistent with ROMS. Using AVRORA, instead of the DA component integrated in ROMS, has allowed us additional flexibility implementing and studying effects of different initial condition error covariances, data functionals (see Kurapov *et al.*, 2011), and other aspects of the DA system. ROMS, AVRORA, routines facilitating model error covariance implementation, and other elements of the variational method implementation are utilized as stand-alone executables coupled via C shell scripts. TL&ADJ AVRORA allows finding corrections to the initial conditions (including SSH  $\zeta$ , two components of horizontal velocity  $u$  and  $v$ , temperature  $T$ , and salinity  $S$ ) and wind stress (not corrected in this study). Similar to ROMS, the output state of AVRORA is defined as fields of SSH,  $u$ ,  $v$ ,  $T$ , and  $S$  provided at all the grid points at equal time intervals (every 4 hours in our case). Using AVRORA, any observation, defined as a linear combination of the elements of the multivariate time-and-space-discrete model state, can be

assimilated without making any changes to the TL model and its ADJ counterpart. For instance, daily-averaged surface velocity data can be easily matched to the daily-averaged model output. Also, the SSH slope from altimetry can be assimilated to correct model surface slopes and hence surface geostrophic currents without necessarily affecting the mean SSH level (Kurapov *et al.*, 2011).

### **3. The data**

Daily surface current maps are derived from the combination of six long-range and five standard-range HF radars along the Oregon coast from June 1 to July 30, 2008. Each radar measures a component of the surface current in a radial direction from the radar. Original radial component data have been processed using a least square procedure to obtain a map of orthogonal  $u$  and  $v$  components on a  $6 \times 6$  km grid. Although assimilation of the original radial component data can potentially be more beneficial (see discussion in section 9), we use the mapped daily-averaged data, which eased pre-DA quality control of the data.

SST and SSH observations are used to verify the assimilation results. Blended  $0.1^\circ$  resolution SST maps (provided by D. Foley, NOAA-CoastWatch) combine data observed by the Advanced Microwave Scanning Radiometer (AMSR-E) instrument, the Advanced Very High Resolution Radiometer (AVHRR), the geostationary GOES Imager, and the Moderate Resolution Imaging Spectrometer (MODIS). The images are available daily representing five-day weighted averages centered on each day. These will be compared to the daily-averaged model SST. The SSH product is the AVISO Jason-1 along-track absolute dynamical topography, smoothed by the data provider and available at the 18-km along-track resolution (<http://www.aviso.oceanobs.com/>).

When comparing these data to the model, the along-track means have been removed from both the observations and their model counterpart since we are mostly interested in whether the model reproduces correctly the SSH slope, which is dynamically related to the geostrophic component of the surface currents.

A free-run model solution is obtained implementing ROMS without DA, starting on May 1, 2008 from initial conditions obtained by interpolation from NCOM-CCS. The solution for June-July is analyzed. Important features associated with coastal upwelling are present in this solution, including intensified offshore transport of cold upwelled waters in the area near Cape Blanco (43°N), also observed in satellite SST (Figure 3a and b). However, details of the SST front differ between the observations and the free-run model. For instance, the modeled separated cold jet extends farther offshore than observed at 43°N. The two-month-averaged observed HF radar surface currents compare qualitatively well to similarly averaged free-run model currents (Figure 2a and b), showing southward and offshore motion of the surface waters. Figure 2c shows the time-mean difference between the model and observed currents. Relatively larger misfits are found on the shelf near 46°N and offshore near the Heceta Bank complex (44-45°N) and Cape Blanco (43°N). Near 46°N, the Columbia River plume may affect circulation patterns (Banas *et al.*, 2008, Liu *et al.*, 2009) and contribute to the error in our model which does not include the river discharge. Larger misfits south of Cape Blanco (43°N) and offshore in the CTZ can be associated with limited model predictability of jets and eddies (Kim *et al.*, 2011). The map of the time-averaged root mean square difference (or “error”, RMSE) between the model and observed current magnitudes (Figure 2d) shows larger values near Cape Blanco, where complex time-dependent flow processes associated with separation of the coastal current occur during summer.



#### 4. The data assimilation approach

In our tests, assimilation proceeds in a series of 3-day time windows (Figure 4). The daily-averaged surface currents are utilized by the AVRORA assimilation system to find corrections to the initial conditions at the beginning of each window. As a combined effect of the adjoint model dynamics and the assumed model error covariance (see below), the correction is obtained for every initial field (SSH,  $u$ ,  $v$ ,  $T$ , and  $S$ ). The nonlinear ROMS is run for 6 days starting from the corrected initial conditions, providing a 3-day analysis (a nonlinear solution in the given assimilation window) and a 3-day forecast, which is used as the background solution for model linearization in the next window. It is verified that the nonlinear ROMS analysis fits the assimilated data better than the prior model (forecast from previous window). The following cost function is minimized in each window:

$$\mathbf{J}(\mathbf{u}) = [\mathbf{u}(0) - \mathbf{u}_0^b]' \mathbf{C}_0^{-1} [\mathbf{u}(0) - \mathbf{u}_0^b] + [\mathbf{d} - \mathbf{L}\mathbf{u}(t)]' \mathbf{C}_d^{-1} [\mathbf{d} - \mathbf{L}\mathbf{u}(t)], \quad (1)$$

where  $\mathbf{C}_0$  and  $\mathbf{C}_d$  are the initial condition and data error covariance matrices, respectively,  $\mathbf{u}(0)$  is the estimated initial condition,  $\mathbf{d}$  is the data vector including all assimilated observations,  $\mathbf{L}$  is a measurement operator matching the model state and the observations, and  $\mathbf{u}_0^b$  is the prior (background) initial condition. The minimum to (1) is found using the indirect representer method (Chua and Bennett, 2001, Bennett, 2002, Kurapov *et al.*, 2007, 2011). It is only briefly outlined here. Our implementation is very similar to Kurapov *et al.* (2011), where additional details can be found. In general, a representer is a multivariate solution of a linearized model that “represents” an observational functional (a sampling rule) for a given observation in a linear vector space, in which the scalar product is defined by the model part of the cost function (Friedman, 1982, Bennett, 2002). A representer for the  $k$ -th observation is obtained as a result of

one ADJ model computation (run backward in time forced with an impulse at the observation location and time), followed in our case by convolution of the ADJ solution at the initial time  $\lambda_k(0)$  with  $\mathbf{C}_0$ , and then one TL model computation (run forward in time using  $\mathbf{C}_0\lambda_k(0)$  as initial conditions). The optimal initial condition can be written as:

$$\mathbf{u}(0) = \mathbf{u}_0^b + \mathbf{C}_0 \sum_k b_k \lambda_k(0) \quad , \quad (2)$$

where  $b_k$  can be determined using the representers and  $\mathbf{d} - \mathbf{L}\mathbf{u}^b(t)$ .

If the number of observations is large, computation of every representer is impractical. An iterative procedure is applied instead to determine the correction as an optimal linear combination (the second term in (2)). To speed up the minimization, a preconditioner (Egbert *et al.*, 1994, Bennett, 2002) is formed by computation of a small subset of all representers. The representers from this subset are sampled at the observation locations and times and used to build an effective preconditioner for the iterative solver. Then the conjugate-gradient algorithm (CGA) can be used to find  $b_k$  and the optimal linear combination (2). Each step of the CGA is again accomplished using an ADJ model run (forced with a linear combination of impulse functions), implementation of  $\mathbf{C}_0$  to the adjoint solution at the initial time, and one TL model run. In our tests, approximately 200 representers were computed in each window to form the preconditioner. Then, the preconditioned CGA converged in approximately 10–20 iterations. Similar convergence rates were obtained assimilating along-track SSH and SST (see Kurapov *et al.*, 2011). The computer processor unit (CPU) time required to obtain the optimal solution would be comparable to that of 400 forecast model runs. The computational time is reduced considerably by distributing the representer runs (independent from each other) to different processors of a cluster.

## 5. Error covariances

The data errors are assumed to be uncorrelated, and the corresponding data penalty term in the cost function is

$$[\mathbf{d} - \mathbf{L}\mathbf{u}(t)]' \mathbf{C}_d^{-1} [\mathbf{d} - \mathbf{L}\mathbf{u}(t)] = \sigma_d^{-2} \sum_{k=1}^K (x_k^{obs} - \overline{x_k^{mod}(t)})^2, \quad (3)$$

where the  $u$  and  $v$  velocity components at each location are considered to be two separate observations,  $x_k^{obs}$  is an observed velocity value, and the overbar denotes the time average of the model velocity over a corresponding 1-day period;  $\sigma_d=0.05$  m/s. Note that this simple choice of the data error covariance is done for convenience. In practice, the data error variance and correlations would depend on many factors, including the velocity magnitude, spatial scales of variability, data location, etc. (O'Keefe, 2005).

The initial condition error covariance  $\mathbf{C}_0$  can provide smoothing (filtering small spatial scales in the adjoint solution) and also introduce dynamical balances relevant to ocean flows. In our initial tests (section 6), the importance of using a dynamically balanced covariance is studied. Cases with two different covariances will be compared.

The first covariance assumes that errors in model variables are mutually uncorrelated:

$$\mathbf{C}_0 = \begin{pmatrix} \mathbf{C}_\zeta & 0 & 0 & 0 & 0 \\ 0 & \mathbf{C}_u & 0 & 0 & 0 \\ 0 & 0 & \mathbf{C}_v & 0 & 0 \\ 0 & 0 & 0 & \mathbf{C}_T & 0 \\ 0 & 0 & 0 & 0 & \mathbf{C}_S \end{pmatrix}, \quad (4)$$

where  $\mathbf{C}_x = \langle \delta x \delta x' \rangle$ ,  $\delta x$  denotes an error in each model variable ( $\zeta$ ,  $u$ ,  $v$ ,  $T$ , or  $S$ ), and  $\langle \dots \rangle$  is the statistical ensemble average (expected value). The covariance of initial errors in each model variable  $\mathbf{C}_x$  is defined as the product of two (for  $\zeta$ ) or three (for every 3D variable) bell-shaped covariances, separable in longitude  $\varphi$ , latitude  $\theta$ , and the vertical (terrain-following)  $s$ -coordinate used in ROMS. For an element of  $\mathbf{C}_x$  corresponding to model nodes  $(\varphi_1, \theta_1, s_1)$  and  $(\varphi_2, \theta_2, s_2)$ ,

$$\begin{aligned} \langle \delta x(\varphi_1, \theta_1, s_1) \delta x(\varphi_2, \theta_2, s_2) \rangle &= \sigma_x(\varphi_1, \theta_1, s_1) \sigma_x(\varphi_2, \theta_2, s_2) \\ &\times \exp\left(-\frac{(\varphi_1 - \varphi_2)^2}{2l_\varphi^2}\right) \exp\left(-\frac{(\theta_1 - \theta_2)^2}{2l_\theta^2}\right) \exp\left(-\frac{(s_1 - s_2)^2}{2l_s^2}\right), \end{aligned} \quad (5)$$

where  $\sigma_x$  is the error standard deviation, and  $l_\varphi$ ,  $l_\theta$ , and  $l_s$  are the decorrelation scales in the zonal, meridional, and vertical directions, respectively. Note that although matrix  $\mathbf{C}_0$  is very large, we do not have to store it. Only the rule by which  $\mathbf{C}_0$  multiplies a vector of initial condition adjoint sensitivity values  $\lambda(0)$  is needed. The vertical decorrelation scale is based on  $s$ -coordinate for simplicity of implementation. The actual vertical scale (in meters) will thus be proportional to the sea depth. This is not an unreasonable assumption, as it reduces the direct impact of assimilation of surface data near bottom in shallow waters, where flows are more directly wind-driven.

The second covariance uses the balance operator approach of Weaver *et al.* (2005). The linear balance operator  $B$  is defined using basic diagnostic relationships (including linearized forms of the equation of state and T-S relation, geostrophy, and thermal wind balance – see Kurapov *et al.* 2011, appendix B). It allows computation of perturbations in all the initial fields given a subset of fields that can be assumed to be mutually uncorrelated. In our case:

$$\delta \mathbf{u} = \begin{pmatrix} \delta \zeta \\ \delta u \\ \delta v \\ \delta T \\ \delta S \end{pmatrix} = B \begin{pmatrix} \delta T \\ \delta \Psi \end{pmatrix}, \quad (6)$$

where  $\delta \Psi$  is the depth-integrated transport stream function perturbation. Technically,  $B$  is provided as a *FORTRAN* code. Then,

$$\mathbf{C}_0 \lambda(0) = \langle \delta \mathbf{u} \delta \mathbf{u}' \rangle \lambda(0) = B \begin{pmatrix} \mathbf{C}_T & 0 \\ 0 & \mathbf{C}_\Psi \end{pmatrix} B' \lambda(0), \quad (7)$$

where  $B' \lambda$  is written as the computer code adjoint to  $B$ . A bell-shaped covariance of type (5) is assumed for  $\mathbf{C}_T$ . Since our understanding of errors in  $\Psi$  is limited, we assume  $\mathbf{C}_\Psi = 0$  in our computations. Still, all cross-correlations between model state variables are nonzero due to the effect of  $B$ .

For the experiments in this study, the following covariance parameters are chosen. The decorrelation scales in each horizontal direction are nominally equal to  $l_\phi = l_\theta = 25$  km. The vertical length scale, based on  $s$ -coordinate, is 100 m in a 3000-m water column. For the unbalanced covariance,  $\sigma_\zeta = 0.02$  m; for the 3D variables,  $\sigma_T = 0.5^\circ$ ,  $\sigma_S = 0.11$  psu, and  $\sigma_u = \sigma_v = 0.08$  m/s at the surface, all reduced with depth as  $\exp(z/z_0)$ , where  $z_0 = 100$  m ( $z$  is 0 at the surface, and negative at depth). Note that in the unbalanced case,  $\sigma_S$  is chosen to be consistent with  $\sigma_T$  based on the average  $T$ - $S$  relation in the upper water layer. Choices for  $\sigma_u$  and  $\sigma_v$  are consistent with  $\sigma_\zeta$  and the assumed horizontal decorrelation scale by geostrophy, *i.e.*,  $\sigma_u = (g\sigma_\zeta)/(lf)$ , where  $g$  is the gravity and  $f$  is the Coriolis parameter. In the balanced case,  $\sigma_T$  is similarly  $0.5^\circ\text{C}$  at the surface and is decreased with increasing depth. Error levels for other

components of the ocean state as well as their 3D correlation structure are fully determined by the balance operator and its adjoint.

## **6. Initial experiments: effect of $C_0$ .**

The sensitivity of the DA system to HF radar surface current assimilation was initially studied in a three-day, single assimilation window experiment, July 19-21, 2008. The free-run solution for this period is used as the prior (background) solution. After the corrections to the initial conditions are obtained, we run the model for a period of 15 days (3-day analysis plus 12-day forecast), to assess the long term response of the system to assimilation in a single window. This study included comparisons of cases with the balanced and unbalanced covariances. Figure 5 shows the impact of DA in terms of the area-averaged model-data RMSE and correlation, separately for HF radar surface currents (left) and unassimilated SST (right). In terms of the surface velocity statistics (Figure 5a and c), the results for the first three days show an improved fit to the assimilated data. The positive impact on the surface velocities in terms of the RMSE is still apparent at the end of the 12-day forecast period, while correlation of model and observed currents returns to the prior model level faster, on the 5<sup>th</sup> day of the forecast. Velocity statistics in the cases with the two different covariances are similar. Comparison of the model solution with SST (Figure 5b and d) shows improvement both in terms of RMSE and correlation over the entire 15-day period with somewhat better performance from the balanced covariance.

Although the case with the balanced covariance outperforms that with the unbalanced covariance, the latter also provides a useful correction. DA performance in the unbalanced covariance case exceeded our expectations and suggested that the correction to the initial

condition may exhibit some features associated with large-scale dynamic balances, even without help from the balanced  $\mathbf{C}_0$ . These features are potentially provided by the adjoint model. To investigate this issue, we will examine initial fields of adjoint sensitivity  $\lambda_k(0)$  and  $\mathbf{C}_0 \lambda_k(0)$  corresponding to the observation of the daily-averaged surface zonal velocity component  $v$  (Figures 6-8). Before we discuss results, let us recall that  $\mathbf{C}_0 \lambda_k(0)$  is the initial condition for the representer  $\mathbf{r}_k(t)$ . In the statistical interpretation of the variational method (Bennett, 2002),  $\mathbf{r}_k(t) = \langle L_k \delta \mathbf{u}(t), \delta \mathbf{u}(t) \rangle$  is the covariance, between the error in the prior model solution for every component (SSH,  $u$ ,  $v$ ,  $T$ , and  $S$ ) at every point in space and time ( $\delta \mathbf{u}(t)$ ) and the error in that same prior model solution sampled at the observation location and time  $L_k \delta \mathbf{u}(t)$ . If  $\mathbf{r}_k(0) = \mathbf{C}_0 \lambda_k(0)$  is plotted, corresponding to an observation of the surface velocity  $v$ , it can be expected that its SSH component would have a slope in the zonal (cross-shore) direction, meaning that the initial errors in the surface pressure gradient are in an approximate geostrophic balance with the errors in the model surface  $v$ , sampled at a later time. If the error in the surface  $v$  is associated with baroclinic dynamics on geostrophic scales, we could expect that the SST component of  $\mathbf{C}_0 \lambda_k(0)$  to the east (west) of the observation location is positively (negatively) correlated with the surface  $v$  error. These arguments suggest that the SSH and SST components of  $\mathbf{C}_0 \lambda_k(0)$  might be expected to look like dipoles, oriented in the direction of  $v$ . Analysis below confirms these expectations.

For this analysis, we choose a time window of July 19-21. Figure 6 shows the 3-day average background surface velocity field. The observation location is chosen on the path of the jet separated from the coast, to additionally study the impact of advection on  $\lambda_k(0)$ . In Figure 7 showing  $\lambda_k(0)$ , the top, middle, and bottom rows correspond to three different observations

obtained at the same location but different times (days 1, 2, and 3); columns (left to right) show the SSH component of  $\lambda_k(0)$ , and the temperature component at the surface, 50 m, and 500 m depth. In all fields, we find a variety of larger and smaller scale structures that might be associated with different dynamical mechanisms affecting co-variability of the sampled variable and initial fields.

The areas of maximum sensitivity in all three cases are displaced to the northeast of the observation location, seen in the SSH, SST, and  $T$  fields at the 50-m depth. The fields corresponding to the observation on day 3 show the farthest displacement. However, no displacement is found at 500 m depth (Figure 7d, h, and l). These differences are consistent with advection by a relatively shallow baroclinic coastal jet present in the background solution. Note that the effect of advection on adjoint computations (of sensitivity of the initial fields to the area integrated SST) has also been demonstrated by Zhang *et al.* (2009) in a study off the US East Coast.

The dipole structure suggested by our qualitative study is hard to identify in the SSH sensitivity fields (Fig. 7, column 1), partly due to the presence of small scale features, which are likely associated with inertia-gravity waves (although we have not analyzed their characteristics in detail). In the SST sensitivity fields (column 2), short-scale (20-km) variability along the jet path (particularly seen in Fig. 7j, day 3 observation) may potentially be associated with jet instabilities. These instability patterns are mostly shallow and not as evident in the initial temperature sensitivity fields at 50 m depth (column 3). At this depth, as well as at 500 m (column 4), the dipole structure is clear in the  $T$ -component of  $\lambda_k(0)$  (particularly, Fig. 7 c, g, h, and l). We also find that fields at the 500 m depth exhibit small scale wave-like patterns, which



are not as pronounced at 50 m. Their origin and role, and relation to the SSH wave-like patterns are potentially interesting topics for future research.

Upper (lower) plots in Figure 8 show  $C_0\lambda_k(0)$  for the cases of the balanced (unbalanced) covariance, as well as evolution of the representer solutions starting from those initial conditions. Only results for the day 1 observation (corresponding to sensitivity fields at the top of Figure 7) are shown. In the balanced covariance case, the SSH (Figure 8a) and SST (Figure 8d) components of the representer initial conditions show a dipole structure, in approximate geostrophic balance with velocities. In this case, positive correction to the meridional surface velocity component (at the time of measurement on day 1) would be associated with the plotted correction in the SSH slope (lowering to the west and rising to the east) at the initial time. A cyclonic (anticyclonic) eddy to the west (east) of the observation location is associated with cooling (warming) of the surface. Note that the optimal correction to the initial conditions is obtained as a linear combination of structures like this, corresponding to each observation. Figure 8 b, c, e, and f show the evolution of the representer, for the balanced  $C_0$  case. The SSH component is relatively unchanged after 24 h, and the SST component shows some variability, potentially associated with baroclinic instabilities and effects of the background currents.

The SSH field from the unbalanced case (Figure 8, two lower rows) does not show a strong dipole structure in SSH at the model initial time, as this is filtered (along with the small scale wave patterns) by the assumed SSH covariance. However initial fields of surface velocity and SST are qualitatively similar to the balanced case. Since the TL model describes the natural development of balanced dynamics, the SSH component of the representer adjusts quickly, in about 6 hours, to the velocity and SST fields and soon after that all the representer components look qualitatively similar to the balanced case. Based on this result we conclude that the

unfiltered sensitivity  $\lambda_k(0)$  already provides information about dynamical balances. Filtering (and scaling) each component independently, applying the unbalanced  $\mathbf{C}_0$ , may disrupt these balances at some level, but if length and amplitude scales are dynamically sensible, negative effects will not be too severe, and balances will be quickly restored. For example, the assumption that the initial errors in each field are auto-correlated on a scale comparable to the Rossby radius of deformation will allow a quasi-geostrophic balance between field components to persist after filtering. It is possible that for this reason, the DA improvement in the unbalanced covariance case is sensible and comparable to the balanced covariance case.

## 7. A 60-day DA experiment

A 60-day (20 three-day cycles) assimilation experiment was implemented for the period of June 1 - July 30. Only the case with the balanced covariance is described here. Statistical comparisons involving surface velocities and SST are shown in Figure 9, 10, and 11. These will be discussed together with monthly mean fields of observed and model SST and model volume transports integrated between the surface and 200 m depth from the free-run, analysis, and NCOM solutions for June and July (Figures 12 and 13, respectively).

As a result of DA, the surface currents are improved in both the analyses and 3-day forecasts within the entire 60-day interval, in terms of both area-averaged model-data RMSE and correlation (Figure 9a and c). We also compute the model-data RMSE by averaging both in space and in time separately for days 1, 2, and 3 of the analysis and forecast (combining in each case misfits from 20 days) (Figure 10a). This also shows that analyses and forecasts are better than that for the free-run solution. In this plot, we note that the change in RMSE from analysis

day 3 ( $0.08 \text{ m s}^{-1}$ ) to forecast day 1 ( $0.11 \text{ m s}^{-1}$ ) is larger than the change from analysis day 1 to 3, or forecast day 1 to 3. The nonlinear ROMS solution fits velocity observations uniformly well on analysis days 1-3, but experiences transition to a less satisfactory solution during forecast day 1 (which is still statistically better than the free-run model). Maps of velocity RMSE, averaged in time point-by-point separately for each day of the analysis and forecast (Figure 11a-f), show that this jump in RMSE occurs mainly in the jet separation area south of Cape Blanco (43N) and along the coast north of 44.5N.

The SST RMSE for the free-run solution (Figure 9b, grey solid line) grows on average with time, associated with limited predictability of jets and eddies in the CTZ as the SST front moves westward. The SST correlation of the free-run model and the satellite data is generally smaller in June than in July (see Figure 9d, grey solid line). This may be associated with a relatively weaker contrast of warmer offshore and colder shelf waters in early summer. Also, warmer SST is observed in the area of the Columbia River plume (not modeled in our case), more apparent in June (Figure 12a) than in July (Figure 13a). The analysis and forecast SST RMSE are worse than in the free-run model during the first 20 days of June, but show significant improvement over the free-run model after that. In our interpretation, after the influence of the Columbia River plume has diminished, the DA model describes the offshore translation of colder shelf waters in the CTZ better than the free-run model. The SST model-data correlation is moderately better for the analysis and forecast fields compared to the free-run model over most of the study period (Figure 9d). Analysis of the SST RMSE values obtained by averaging in space and in time separately for each day of the analysis and forecast (Figure 10b) reveals that both the analysis and forecast are better than the prior model SST for each day. However, curiously, the analysis SST RMSE gradually decreases from day 1 to 3 and this tendency continues for the forecast SST RMSE,

such that the SST forecasts are statistically better than SST analyses (which can be also seen in the time-series RMSE, see Figure 9b). Maps of SST RMSE, obtained by averaging in time separately for days 1, 2, and 3 of the analyses and forecasts (Figure 11, bottom), show larger SST analysis errors on day 1 in areas where the differences between forecast and observed velocities were larger, and where the DA impact was stronger. Correction of SST resulting from assimilation of velocities introduces inconsistencies that are improved dynamically with time.

Improvement in the SST model-data statistics resulting from surface velocity assimilation is associated with the corrected geometry of the SST front. This is illustrated in Figure 3, in which the daily mean SST from the free-run, analysis, and forecast solutions are plotted next to the satellite image and the free-run model SST on July 23, 2008 (note the forecast field is a result of the nonlinear run starting from improved initial conditions in the previous assimilation window). These plots also include the observed and model surface currents (free-run, analysis, and forecast) at the observation locations. Qualitatively, the analysis and forecast SST fields look to be in better agreement with the observed SST than the free-run solution. For instance, very intense separation of the colder shelf waters off Cape Blanco ( $43^{\circ}\text{N}$ ) is seen in the free-run model (Figure 3b), but not in the data (Figure 3a). The extent and shape of the cold water front at that latitude in the analysis and forecast solutions (Figure 3c and d) are much closer to the observations. The satellite data also suggest the presence of a large anti-cyclonic eddy between  $41^{\circ}\text{N}$  and  $43^{\circ}\text{N}$ , not reproduced in the free-run solution. After assimilation, this eddy is apparent in surface currents and SST model fields.

The geostrophic component of the surface currents is proportional to the SSH slope. The impact of velocity assimilation on SSH is verified against Jason-1 along-track altimeter data (Figure 14, track 206 (upper plots) and 247 (lower plots); see Figure 1 for track locations). On

average, within the area of HF radar data coverage (between 126°W and the coast), the SSH slope in the assimilation solution follows the observed SSH slope qualitatively better than the free run model. For instance, in Figure 14h there is an abrupt drop in SSH between 125.5-125°W, which is captured in the analysis, but not in the free-run solution. To provide a more quantitative assessment, we compute the model-data SSH RMSE errors averaged over along-track segments between 126°W and the coast and over each track. Before the RMSE is computed, the model and data were demeaned independently for each pass. For track 206, the RMSE for the free-run (analysis) solution is 6.6 (4.4) cm; for track 247, the RMSE for the free-run (analysis) solution is 3.8 (2.7) cm. Along both tracks, the RMSE is reduced in the analysis solutions.

We have not had enough subsurface data during this period to verify whether surface velocity assimilation has a positive impact at depth. Data from several conductivity-temperature-depth (CTD) surveys have been available with a 1-m vertical resolution (courtesy of W. Peterson and J. Peterson, NOAA Northwest Fisheries Science Center), but these extend only over the shelf. Representative examples of the observed and model cross-shore potential density sections are given in Figure 15 (see Figure 1 for the section locations). The section along the Newport (NH) line, 44.65°N (Figure 15a-d) is in the area of relatively along-shore uniform slope, where shelf dynamics dominated by wind-driven upwelling and downwelling are generally well reproduced by a free-run baroclinic primitive equation model (Oke *et al.*, 2002b, Kurapov *et al.*, 2005). The section off Crescent City (section CC, 42.6°N, Figure 12e-h) is in the jet separation area south of Cape Blanco where variability is less predictable, more dominated by jet and eddy activity in the CTZ (Koch *et al.*, 2010). For each cross-shore section we provide (see Figure 15, left to right) the observed, free-run model, analysis, and NCOM-CCS potential density fields (recall, the latter are the fields from a lower resolution data assimilative model that provided

boundary conditions for our study). The plotted model sections extend to  $126^{\circ}\text{W}$  (the zone of HF radar coverage), about twice as far as the observed sections go. The extent of the observed CTD sections is shown as the vertical white line in the model sections.

Our free-run model as well as NCOM-CCS indeed reproduces the hydrographic structure over the shelf along the NH line qualitatively correctly. Assimilation of surface velocities essentially does not change the hydrographic structure associated with the wind-driven upwelling on the shelf. Note that in this cross-shore section, the isopycnals in the analysis solution (Figure 15c) are lowered near  $126^{\circ}\text{W}$  compared to the free-run solution or NCOM-CCS. A similar effect was found on this date in the case using the unbalanced covariance as well. Additional analysis shows that this effect is local, associated with a pool of warmer-than-observed water and an anticyclonic eddy formed in the analysis solution in the area not covered by the HF radar data (Figure 16). Assimilation of additional data sources covering the whole model domain is expected to improve the model solution beyond the HF radar coverage area. We have tested whether combined assimilation of the HF and satellite altimetry SSH data (Jason-1 and Envisat) can help (assimilation of these additional data proceeded as described in Kurapov *et al.*, 2011). The assimilation of altimetry does have a positive impact on SSH along the tracks (Figure 17). In particular, the strong northward flow along  $127^{\circ}\text{W}$  between about  $43^{\circ}\text{N}$  and  $46^{\circ}\text{N}$  is reduced, primarily as a result of the impact of the data from the Envisat track 284 (see Figure 17b). However, it turns out that no track from either satellite passes through the core of the possibly spurious or misplaced eddy (Figure 16d and e) centered near  $126.5^{\circ}\text{W}$  and  $44.5^{\circ}\text{N}$ . This eddy is still found in the new analysis solution, with a smaller impact area. We believe that surface currents should be better constrained when wide swath altimetry data become available in the future (Hénaff *et al.*, 2008).

The sections for the CC line (Figure 15e-h) show that neither our free-run model nor NCOM-CCS reproduced the isopycnal structure correctly in this area. The observations show an upwelling front at 125°W and a strong subsurface horizontal density gradient on the shelf. The free-run model shows the front farther offshore and an area of relatively flat isopycnals inshore of this front. The free-run model isopycnals near the shelf bottom suggest a strong northward flow, which is not supported by the observed density sections. The NCOM-CCS, which assimilated SST, shows the front at the correct location. However, the density structure inside of the front is not correct, and is more similar to our free-run solution (showing an area of low cross-shore density gradient and deepening of isopycnals near the bottom associated with the northward flow). While our analysis solution, constrained by assimilation of only HF radar surface currents, is unable to fix all of those deficiencies, it does move the front inshore (although the front is weaker). Also, the area inshore of the front shows upwelled isopycnals, more consistent with the observations than the free-run model. Future research is needed to determine whether any combination of surface data can constrain the subsurface density field better.

## **8. Assimilation effect on volume transports**

A coastal model constrained by DA can potentially provide improved estimates of shelf-interior ocean transports of volume, heat, and materials. This information would provide guidance for climate models that do not resolve these fluxes. Here we analyze the impact of HF radar surface current assimilation on the volume transport across and along the shelf.

The two-month time-averaged vertical integral of the zonal velocity along 125°W is shown in Figure 18a. These estimates from the free-run model and analysis are qualitatively similar in magnitude, but have considerable differences in latitudinal distribution. In Figure 18b, the vertically integrated volume flux is separated into westward and eastward flow components. To obtain those components, the zonal velocity profiles  $u(z)$  were integrated separately for  $u > 0$  and  $u < 0$ . The DA has affected the westward flux more than eastward flux. The two-month meridional volume transport from 125°W to the coast (Figure 19) shows that DA changes the northward alongshore transport. South of 43°N the transport is primarily northward. The abrupt drop in magnitude of the northward transport changes latitudinal location from 42.5°N in the free-run solution to between 41.5°N and 42°N in the DA analysis. That change is consistent with the change of location of maximum westward offshore transport shown in Figure 18.

Figure 20 shows time- and area-averaged meridional currents ( $v$ ) over the shelf and slope, as functions of the latitude. The upper plots are for June, and the lower for July. Values on the left (Figure 20a and c) are computed by averaging over the top 200 m between the coast and the location of the 2000-m isobath and plots on the right (Figure 20b and d) by averaging in a sector bounded by the bottom, the 200-m horizontal depth level, and the vertical line at the location of the 2000-m isobath. Results for the free-run ROMS, analysis, and NCOM-CCS are shown. Even though NCOM-CCS provides boundary conditions for our ROMS model, considerable differences are found in the meridional fluxes obtained from the free-run ROMS and NCOM estimates. DA affects the upper 200-m meridional flux more than the subsurface along-slope flux and brings these estimates closer to NCOM-CCS. An increase in the southward currents above 200 m between 42°N and 43°N in the DA analysis compared to the free-run solution is apparent and is consistent with the changes in the monthly mean transport in that region shown in Figure



12 and 13. The ROMS estimate of the subsurface meridional flux is larger (more positive) than NCOM in June, and lower (negative or less positive) in July. While we do not have data to assess which solution is closer to the truth, this issue suggests interesting questions for future research, in particular with regard to the role of the poleward undercurrent (Pierce *et al.*, 2000), seasonal variability of the subsurface flows in this area, and ability of the DA system to constrain those.

We find it interesting that during July (Figure 13b, c, and d) the intensified offshore flux near 42°N, not found in our free-run model, is obtained in both our DA analysis and NCOM fields. The location of this intensive westward flow is constrained by assimilation of different types of data in the two models (HF radar surface currents in our case and SSH and SST in the case of NCOM).

## **9. Summary**

Assimilation of surface currents from a set of standard and long-range HF radars into the ROMS model coupled with the variational AVRORA DA system improves the quality of near-surface velocity forecasts and has a positive effect on the geometry of the SST front. In the area covered by the HF radar data, there is a moderate improvement in the SSH slope (by comparison with along-track altimetry). HF radar observations cover areas very close to the coast and could be potentially useful in bridging the data gaps in altimetry. Assimilation has a stronger impact on near-surface horizontal volume fluxes than subsurface flows. Reasons for differences in subsurface meridional fluxes in NCOM and ROMS (which was nested in NCOM) will require additional, more focused studies.

Results with two different initial condition error covariances were implemented and compared. The case with the unbalanced covariance showed a better performance than expected. Analysis of the representers has suggested that using the adjoint model, plus the assumption that the initial errors in each oceanic field are correlated over a horizontal scale comparable to the Rossby radius of deformation, result in a dynamical balance between the oceanic fields in the DA correction even without using the balanced initial condition error covariance. The case with the balanced covariance still showed a moderate improvement over the case with the unbalanced covariance in terms of SST (which was not assimilated).

Surface maps of two orthogonal components of surface velocity have been assimilated. These maps are routinely used in our real-time assimilation system since their visual inspection can be readily done as a matter of the pre-DA data quality control. However, assimilation of original, radial component data could be more advantageous (Kurapov *et al.*, 2003, Barth *et al.*, 2008, Shulman and Paduan, 2009, Powell, pers. comm.), in particular, because the accuracy of the radial component data is spatially more uniform. Also, in the areas sampled by only one HF radar, where mapping using traditional methods for data analysis would not be possible, the radial component data would still provide a useful constraint on the simulated flow. Several tests, not discussed above, have been performed assimilating the radial component data. In those cases, DA served as a dynamically-based mapping tool that yielded model fields of the two orthogonal components of surface currents. Those have been found to be close to maps obtained directly from the observations (more traditional data interpolation). The radial current component assimilation showed effects on SST similar to the case assimilating the mapped orthogonal currents (Figure 3e).

In this study, the wind forcing was assumed to have no error. Allowing for errors in wind forcing as well as the initial conditions may improve the forecast skill of the DA system (Kurapov *et al.*, 2009, Zhang *et al.*, 2010).

Although the Columbia River discharge is neglected in our model, our study suggests its potential impact on the SST. Inclusion of the fresh water plume in the DA system would be a logical next step, but it should be taken carefully. First, our implementation of the balanced covariance assumed a simple linear relation between  $T$  and  $S$ . This formulation will have to be revised in the presence of a thin surface layer of a fluid with anomalous  $T$ - $S$  properties, and impact of assimilation on plume structure would have to be evaluated, *e.g.*, using hydrographic section data. Second, inclusion of the Columbia River fresh water discharge could be most conveniently (and realistically) accomplished if adding tidal variability, which would help to mix the river and ocean waters in the estuary. Assimilation in the presence of tides, in particular, the impact of re-initialization on the internal tide field (Osborne *et al.*, 2011), raises additional research questions.

Based on our study, the HF radar measurements are a valuable data source for constraining circulation both over the shelf and in the adjacent interior ocean. Assimilation of this data type in the real-time model will be most beneficial in combination with other observations, including SST and alongtrack SSH.

**Acknowledgments.** This research was supported by the Office of Naval Research (ONR) Physical Oceanography Program (grants N000140810942, N000141010745), NOAA-CIOSS, NOAA-IOOS (NANOOS), and NSF (grants OCE-0648314, OCE-1030922). The authors thank Dr. Igor Shulman for providing the boundary conditions for the simulations, and for valuable

discussions. The authors also thank the anonymous reviewers for several thoughtful and constructive comments.

## References

- Banas, N.S., P. MacCready, and B.M. Hickey, 2008: The Columbia River plume as cross-shelf exporter and along-coast barrier. *Continental Shelf Research*, doi:10.1016/j.csr.2008.03.011
- Barth, A., Alvera-Azcárate, A., and Weisberg, R., 2008: Assimilation of high-frequency radar currents in a nested model of the West Florida Shelf. *J. Geophys. Res.*, **113**, C08033, doi:10.1029/2007JC004585, 2008. 839
- Bennett, A. F., 2002: Inverse Modeling of the Ocean and Atmosphere. *Cambridge University Press*, 234 pp.
- Breivik, Ø., and Ø. Sætra, 2001: Real time assimilation of HF radar currents into a coastal ocean model, *J. Mar. Syst.*, **3–4**, 161–182.
- Brink, K. H., and T. J. Cowles, 1991: The Coastal Transition Zone Program. *J. Geophys. Res.*, **96**, 14,637-14,647.
- Chua, B., and A. F. Bennett, 2001: An inverse ocean modeling system. *Ocean Modelling*, **3**, 137-165.
- Doyle, J. D., Jiang, Q., Chao, Y., Farrara, J., 2008. High-resolution real-time modeling of the marine atmospheric boundary layer in support of the AOSN-II field campaign. Deep-Sea Research II, this issue [doi:10.1016/j.dsr2.2008.08.009].
- Egbert, G. D., A. F. Bennett, and M. G. G. Foreman, 1994: TOPEX/POSEIDON tides estimated using a global inverse model. *J. Geophys. Res.*, **99(C12)**, 24,821-24,852.
- Egbert, G. D., and A.F Bennett, 1996: Data Assimilation Methods for Ocean Tide. Modern Approaches to Data Assimilation in Ocean Modelling. *Elsevier* 147-180.
- Erofeeva, S. Y., G. D. Egbert, and P. M. Kosro, 2003: Tidal currents on the central Oregon shelf: Models, data, and assimilation, *J. Geophys. Res.*, **108(C5)**, 3148, doi:10.1029/2002JC001615
- Fairall, C. W., E. F. Bradley, D. P. Rogers, J. B. Edson, and G. S Young, 1996: Bulk parameterization of air-sea fluxes in TOGA COARE. *J. Geophys. Res.*, **101**, 3747-3767.
- Fox, W. J. Teague, C. N. Barron, M. R. Carnes, and C. M. Lee, 2002: The Modular Ocean Data Assimilation System (MODAS). *J. Atmos. Oceanic Technol.*, **19**, 240–252

Friedman, A., 1982: Foundations of modern analysis, *Dover Publ.*, 256 pp.

Harlan, J., E. Terrill, L. Hazard, C. Keen, D. Barrick, C. Whelan, S. Howden, and J. Kohut, 2010: The Integrated Ocean Observing System High-Frequency Radar Network: Status and Local, Regional, and National Applications. *Mar. Technol. Soc. J.*, **44(6)**, 122-132.

Hénaff, M.L., P.D. Mey, B. Moure, and P.L. Traon, 2008: Contribution of a Wide-Swath Altimeter in a Shelf Seas Assimilation System: Impact of the Satellite Roll Errors. *J. Atmos. Oceanic Technol.*, **25**, 2133–2144. doi: 10.1175/2008JTECHO576.1

Huyer, A. and R.L. Smith. 1978: Physical characteristics of Pacific northwestern coastal waters. The Marine Plant Biomass of the Pacific Northwest Coast, R.W. Krauss, ed., Oregon-State University, Corvallis, Oregon, 37-55.

Kim, S., R.M. Samelson, and C. Snyder, 2011: Toward an uncertainty budget for a coastal ocean model. *Mon. Wea. Rev.*, **139**, 866-884.

Koch, A. O., A. L. Kurapov and J. S. Allen, 2010: Modeling analysis of a separated jet in the Coastal Transition Zone off Oregon. *J. Geophys. Res.*, **115**, C08020, doi:10.1029/2009JC005704.

Kosro, P.M., J.A. Barth, and P.T. Strub, 1997: The coastal jet: Observations of surface currents along the Oregon Continental Shelf from HF radar. *Oceanography*, **10(2)**, 53-56.

Kosro, P.M., 2005: On the spatial structure of coastal circulation off Newport, Oregon, during spring and summer 2001, in a region of varying shelf width. *J. Geophys. Res.*, **110**, C10S06, doi:10.1029/2004JC002769.

Kurapov, A. L., G. D. Egbert, J. S. Allen, R. N. Miller, S. Y. Erofeeva, and P. M. Kosro, 2003: M2 internal tide off Oregon: inferences from data assimilation. *Journal of Physical Oceanography*, **33**, 1733-1757.

Kurapov, A. L., J. S. Allen, G. D. Egbert, R. N. Miller, 2005: Modeling bottom mixed layer variability on the mid-Oregon shelf during summer upwelling, *J. Phys. Oceanogr.*, **35**, 1629-1649.

Kurapov, A. L., G. D. Egbert, J. S. Allen, and R. N. Miller, 2007: Representer-based variational data assimilation in a nonlinear model of nearshore circulation. *J. Geophys. Res.*, **112**, C11019, doi:10.1029/2007JC004117

Kurapov, A. L., G. D. Egbert, J. S. Allen, and R. N. Miller, 2009: Representer-based analyses in the coastal upwelling system. *Dynam. Atmos. Ocean*, **48(1-3)**, 198-218. doi:10.1016/j.dynatmoce.2008.09.002.

- Kurapov, A. L., D. Foley, P. T. Strub, G. D. Egbert, and J. S. Allen, 2011: Variational assimilation of satellite observations in a coastal ocean model off Oregon. *J. Geophys. Res.*, **116**, C05006. doi:10.1029/2010JC006909
- Le Hénaff, Matthieu, Pierre De Mey, Baptiste Mourre, Pierre-Yves Le Traon, 2008: Contribution of a Wide-Swath Altimeter in a Shelf Seas Assimilation System: Impact of the Satellite Roll Errors. *J. Atmos. Oceanic Technol.*, **25**, 2133–2144. doi: 10.1175/2008JTECHO576.1
- Lewis, J. K., I. Shulman, and A. F. Blumberg, 1998: Assimilation of CODAR observations into ocean models. *Cont. Shelf Res.*, **18**, 541– 559.
- Liu, Y., P. MacCready, B. M. Hickey, E. P. Dever, P. M. Kosro, and N. S. Banas (2009), Evaluation of a coastal ocean circulation model for the Columbia River plume in summer 2004, *J. Geophys. Res.*, **114**, C00B04. doi:10.1029/2008JC004929.
- Mellor, G., and T. Yamada (1982), Development of a turbulence closure model for geophysical fluid problems. *Rev. Geophys.*, **20(4)**, 851– 875.
- Moore, A. M., H. G. Arango, G. Broquet, B. S. Powell, J. Zavala-Garay, and A. T. Weaver, 2011: The Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilation systems, Part I: Formulation and Overview, *Prog. Oceanogr.*, **91(2011)**, 34-49.
- O'Keefe, S., 2005: Observing the coastal ocean with HF radar, MS Thesis, Oregon State University.
- Oke, P. R., J. S. Allen, R. N. Miller, G. D. Egbert, and P. M. Kosro, 2002a: Assimilation of surface velocity data into a primitive equation coastal ocean Model. *J. Geophys. Res.*, **107(C9)**, 3122, doi:10.1029/2000JC000511.
- Oke, P. R., J. S. Allen, R. N. Miller, and G. D. Egbert 2002b: A modeling study of the three-dimensional continental shelf circulation off Oregon: Part II. Dynamical Analysis. *J. Phys. Oceanogr.*, **32**, 1383-1403.
- Osborne, J. J., A. L. Kurapov, G. D. Egbert and P. M. Kosro, 2011: Spatial and temporal variability of the M2 internal tide generation and propagation on the Oregon shelf. *J. Phys. Oceanogr.*, **41**, 1733-1757. doi: 10.1175/2010JPO4298.1
- Pierce, S. D., R. L. Smith, P. M. Kosro, J. A. Barth, and C. D. Wilson, 2000: Continuity of the poleward undercurrent along the eastern boundary of the mid-latitude North Pacific, *Deep-Sea Res. II*, **47**, 811-829.
- Shchepetkin, A. F., and J. C. McWilliams, 2005: The Regional Ocean Modeling System: A split-explicit, free-surface, topography following coordinates ocean model. *Ocean Modelling*, **9**, 347-404.

Shulman, I., J.C. Kindle, S. deRada, S.C. Anderson, B. Penta, and P.J. Martin, 2004: Development of a Hierarchy of Nested Models to Study the California Current System. Estuarine and Coastal Modeling 2003, Estuarine and Coastal Modeling, 8th International Conference on Estuarine and Coastal Modeling, Malcolm L. Spaulding - Editor, November 3–5, 2003, Monterey, California, 682 USA.

Shulman, I., and J. D. Paduan, 2009: Assimilation of HF radar-derived radials and total currents in the Monterey Bay area. *Deep Sea Research II*, **56(3-5)**, 149-160. [doi:10.1016/j.dsr2.2008.08.004].

Springer, S. R., R.M. Samelson, J. S. Allen, G. D. Egbert, A. L. Kurapov, R. N. Miller, and J. C. Kindle, 2009: A Nested Grid Model of the Oregon Coastal Transition Zone: Simulations and Comparisons with Observations During the 2001 Upwelling Season. *J. Geophys. Res.*, **114**, C02010, doi:10.1029/2008JC004863.

Strub, P. T., P. M. Kosro, A. Huyer, *et al.*, 1991: The Nature of the Cold Filaments in the California Current System, *J. Geophys. Res.*, **96(8)**, 14743–14768.

Weaver, A. T., Deltel. C., Machu, E., Ricci, S. and N. Daget, 2005: A multivariate balance operator for variational ocean data assimilation. *Q. J. Roy. Meteorol. Soc.*, **131**, 3605-3625.

Wilkin, J. L., H. G. Arango, D. B. Haidvogel, C. S. Lichtenwalner, S. M. Glenn, and K. S. Hedström, 2005: A regional ocean modeling system for the long-term ecosystem observatory, *J. Geophys. Res.*, **110**, C06S91, doi: 10.1029/2003JC002218.

Zhang, W. G., J. L. Wilkin, J. C. Levin, and H. G. Arango, 2009: An Adjoint Sensitivity Study of Buoyancy- and Wind-driven Circulation on the New Jersey Inner Shelf, *Journal of Physical Oceanography*, **39**, 1652-1668.

Zhang, W. G., J. L. Wilkin, and H. G. Arango, 2010: Towards an integrated observation and modeling system in the New York Bight using variational methods. Part I: 4DVAR data assimilation, *Ocean Modelling*, **35**, 119-133.

## Figure Captions

**Figure 1.** Model bathymetry and observation locations (arrows: mapped surface currents from HF radars averaged over June and July 2008, dots: Jason-1 altimeter tracks; also shown are NH and CC hydrographic survey sections). Locations of radars are shown as circles (long-range) and asterisks (standard-range). Bathymetric contours are 200, 1000, and 2000 m.

**Figure 2.** Two-month time-averaged surface velocity statistics at HF data locations: (a) mean observed currents, (b) mean currents from the free-run model, (c) the difference between the

mean observed and free-run model currents, and (d) the RMS difference between the observed and model fields.

**Figure 3.** SST (color) and surface currents (arrows), daily averages on July 23, 2008: (a) observed, (b) free-run model solution, (c) assimilation analysis, (d) assimilation forecast, and (e) analysis from assimilation of HF radial component data.

**Figure 4.** The flowchart of the data assimilation system. Assimilation proceeds in a series of 3-day windows. In each window, the forecast solution from the previous window is used as the background field. The initial condition is updated as a result of the assimilation, running the TL&ADJ AVRORA repeatedly. The forecast nonlinear model (NL ROMS) is then run started from the improved conditions to provide a 3-day analysis and a 3-day forecast, which serves as the prior solution in the next window.

**Figure 5.** Model-data statistics in the initial 3-day assimilation test: (upper) RMSE and (lower) correlation; (left) for surface currents and (right) for SST.

**Figure 6.** The 3-day average background surface velocity vectors (arrows) and speed (color,  $\text{ms}^{-1}$ ) (7/19-21). Observation location used in representer analyses (Figure 7-8) is shown as the white circle.

**Figure 7.** The components of the adjoint sensitivity  $\lambda(0)$ , corresponding to the observation of surface  $v$  on days 1 (row 1), 2 (row 2), and 3 (row 3) (the observation location is marked as the white circles). Left to right: the SSH component ( $\text{s}^{-1}$ ) (column 1) and the temperature component ( $^{\circ}\text{C}^{-1}\text{ms}^{-1}$ ) at the surface (column 2), 50m depth (column 3), and 500m depth (column 4).

**Figure 8.** The representer solutions for SSH and SST at time steps of 0, 12h, and 24h from the balanced case (upper two rows) and unbalanced case (lower two rows) corresponding to the observation of surface  $v$  on day 1 (the observation location is marked as the white circle). The SSH (row 1 and the row 3, units: m) and SST (row 2 and row 4, units:  $^{\circ}\text{C}$ ) fields are shown. Solutions are scaled by the representer  $v$  value at the observed location and time to obtain conventional units.



**Figure 9.** The RMSE (upper) and correlation (lower) of the model solution against observations of surface currents (left) and SST (right) in the 60-day (20 assimilation cycles) case.

**Figure 10.** Model-data RMSE averaged separately in space and in time for forecast and analysis day 1, 2, and 3. (a) surface currents and (b) SST.

**Figure 11.** The RMS difference between the model solution and the observed surface currents at different analysis days ((a) day 1, (b) day 2, and (c) day 3)) and forecast days ((d) day 1, (e) day 2, and (f) day 3).

**Figure 12.** Monthly mean (June, 2008) SST (color) and transport of the upper 200m (vectors): (a) observed SST, (b) free-run model solution, (c) assimilation analysis, and (d) NCOM solution. Note: the ROMS free-run and analysis vectors are shown at 12 km resolution (every other grid point). The NCOM vectors are interpolated to that same grid.

**Figure 13.** Similar to Figure 9, for July, 2008.

**Figure 14.** Comparison of the model SSH against the Jason-1 along-track altimetry observations. Locations of Track 206 (a-f), and Track 247 (g-l) are shown in figure 1.

**Figure 15.** Comparison of the model potential density fields against the CTD observations (NH and CC lines). The locations of the CTD lines (black triangles) are shown in Figure 1. The white lines on the model fields correspond to the offshore extent of the CTD data.

**Figure 16.** SST (color), surface current (vectors), and SSH (black contours) in the area near the NH section on Jul 12, 2008: (a) observations, (b)-(c) free run, and (d)-(e) analysis.

**Figure 17.** SSH analysis solution on Jul 12, 2008 from assimilation of HF only (left) and combined SSH and HF (right). Dotted lines show tracks from Jason-1 and star lines from Envisat. Note the HF radar data extends offshore to about 126W.

**Figure 18.** Two-month temporal mean of the vertically-integrated zonal velocity along 125W (a: total, b: eastward (positive) and westward (negative)).

**Figure 19.** Two-month temporal mean of the meridional volume transport from 125W to the coast (a: total, b: northward (positive) and southward (negative)).

**Figure 20.** Monthly mean meridional current (June and July, 2008) within the 2000m isobath. The left panels show the currents averaged over the depth range from the surface to 200m (or to the bottom where the depth is less than 200m). The right panels show the currents averaged over the depth range from 200m to the bottom. Both averages extend offshore to the 2000 m isobath.