A Three-Dimensional Variational Data Assimilation Scheme for the Regional Ocean Modeling System: Implementation and Basic Experiments

Zhijin Li[1] and Yi Chao

Jet Propulsion Laboratory, California Institute of Technology

Pasadena, California, USA

James C. McWilliams and Kayo Ide

Department of Atmospheric and Oceanic Sciences

& Institute of Geophysics and Planetary Physics

University of California, Los Angeles, California, USA

Submitted to Journal of Geophysical Research: Ocean

Revised: September 18, 2007

[1]Corresponding author: Jet Propulsion Laboratory, M/S 300-323, 4800 Oak Grove Drive, Pasadena, CA 91109; zhijin.li@jpl.nasa.gov

Abstract

A three-dimensional variational data assimilation scheme for the Regional Ocean Modeling System (ROMS), named ROMS3DVAR, has been described in Li et al. [2007]. In this paper, ROMS3DVAR is applied to the central California coastal region, an area characterized by inhomogeneity and anisotropy, as well as by dynamically unbalanced flows. A method for estimating the model error variances from limited observations is presented, and the construction of the inhomogeneous and anisotropic error correlations based on the Kronecker product is demonstrated. A set of single observation experiments illustrates the inhomogeneous and anisotropic error correlations and weak dynamic constraints used. Results are presented from the assimilation of data gathered during the Autonomous Ocean Sampling Network (AOSN) experiment during August 2003. The results show that ROMS3DVAR is capable of reproducing complex flows associated with upwelling and relaxation, as well as the rapid transitions between them. Some difficulties encountered during the experiment are also discussed.

Keywords: coastal oceans, three-dimensional variational data assimilation, anisotropic error covariance, weak constraints, upwelling.

Index terms: 4217 Coastal processes; 4260 Ocean data assimilation and reanalysis; 4262 Ocean observing systems; 4263 Ocean predictability and prediction; 4255 Numerical modeling.

1 Introduction

We recently developed a three-dimensional variational data assimilation scheme (3DVAR) for the Regional Ocean Modeling System and named the scheme ROMS3DVAR. The theoretical formulation of the scheme was presented in Li et al. [2007]. In this paper, we describe the implementation of the proposed scheme and illustrate the importance of the major formulations proposed in Li et al. [2007].

The performance of 3DVAR hinges on the specification of the background error covariance matrices and the incorporation of dynamic constraints. In Li et al. [2007], we proposed particular formulations for constructing the background error correlations and incorporating dynamic constraints. A Kronecker product method was proposed to construct background error correlations. This formulation allows us to construct three-dimensional (3D) correlations, which enables us to incorporate some major aspects of inhomogeneity and anisotropy in coastal oceans, while rendering ROMS3DVAR computationally efficient and reliable. Concerning dynamic constraints, the weak geostrophic and hydrostatic formulations are used. Using these formulations, ROMS3DVAR demonstrated the capability to assimilate a variety of observations simultaneously, including satellite altimetry sea surface heights (SSHs), satellite sea surface temperatures (SSTs), temperature/salinity (T/S) profiles from various observational platforms, Acoustic Doppler Current Profiler (ADCP) current data, and high-frequency (HF) radar sea surface current data.

For the purposes of illustration and evaluation, we apply the scheme to the central California coastal region. Although the ROMS3DVAR scheme is formulated as generally as possible, and hence can be applied to any coastal region, there are several reasons for us

to focus on central California in this study. It is one of the most sampled and investigated coastal regions [e.g., Hickey, 1998]; the unbalanced properties of the flow related to strong spring and summer upwelling [e.g., Hickey, 1998] make it dynamically very interesting; we have extensive experience modeling the central California coastal ocean [e.g., Marchesiello et al., 2003]; and the Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS) [Hodur, 1997] provides us with continuous high resolution atmospheric forcing [Chao et al., 2003; Chao et al., 2007]. Further, ROMS3DVAR was used in the Autonomous Ocean Sampling Network (AOSN) Monterey Field Experiment undertaken during August 2003 in the area surrounding Monterey Bay, California (for details, see http://aosn.mbari.org). This experiment demonstrated an unprecedented *in situ* observational capability for coastal oceans and collected large amounts of data from a variety of observational platforms, including satellites, airplanes, ships, drifters, buoys, autonomous underwater vehicles, and, in particular, a fleet of undersea gliders.

For evaluation and validation, we focus on the following three aspects. First, the error covariances are examined in detail. A method for estimating error covariances with model generated data is described, and the features of the constructed error covariances are discussed to highlight their anisotropy, inhomogeneity and inseparability. Also, the importance of the weak dynamic constraints is examined. Second, we will present a set of single-observation influence experiments that exhibit the structures of the specified correlations, as suggested by Parrish and Derber [1992] and Thepaut et al. [1996]. The analysis increments from single-observation influence experiments are proportional to the covariances of the background errors, allowing us to examine how ROMS3DVAR responds to observations. Finally, the performance of ROMS3DVAR during August 2003 is evaluated against

the AOSN data.

The outline of this paper is as follows: In section 2, we briefly describe the basic configuration of ROMS and ROMS3DVAR as well as the atmospheric forcing. Section 3 describes the estimation of the background error covariance. In section 4 we present the single-observation influence experiments to examine the behavior of ROMS3DVAR. Section 5 presents experiments that assimilate observations from the AOSN experiment. Our focus here is on examining the performance of ROMS3DVAR and investigating some major issues of coastal data assimilation, such as the weak geostrophic constraint. Finally, section 6 discusses and summarizes the main results obtained in this study.

2 ROMS3DVAR Configuration for the U.S. West Coastal Ocean

2.1 ROMS Nested Configuration and Atmospheric Forcing

ROMS is a free-surface, hydrostatic, three-dimensional primitive equation regional ocean model [Shchepetkin and McWilliams, 2005, 2006; Marchesiello et al., 2001]. The vertical discretization uses a stretched terrain-following coordinate (S-coordinate) on a staggered grid over variable topography [Song and Haidvogel, 1994]. The stretched coordinate allows increased resolution in areas of interest, such as the thermocline and bottom boundary layers. The horizontal discretization uses a boundary-fitted, orthogonal curvilinear coordinate on a staggered Arakawa C-grid [Arakawa and Lamb, 1977]. Coastal boundaries are specified as a finite-discretized grid via land/sea masking. The lateral boundary condition is no-normal-

flow at the coast.

ROMS has a one-way nesting capability [Blayo and Debreu, 1999]. As described in Chao et al. [2007], a three-level, nested configuration is used in this study: the largest domain (L0) covers the U.S. West Coastal Ocean at a horizontal resolution of 15km; the next (L1) is nested in L0 covering the central California coastal ocean at a horizontal resolution of 5km; the last (L2) is nested in L1 and focuses on the region around Monterey Bay at a horizontal resolution of 1.5km. The open ocean boundary condition for L2 is determined by L1, that of L1 by L0, and that of L0 by a climatological simulation of the Pacific ocean.

All three nested levels have 32 layers in the vertical. Figure 1 shows the bathymetry and horizontal domain of L2, which is bounded by the California coast at the eastern edge and open ocean at the other three edges. A prominent feature of the bathymetry is the narrow shelf whose width varies along the coast. The slope is particularly steep along the Monterey Submarine Canyon. The numerical algorithms of ROMS are specifically designed to reduce pressure gradient errors associated with steep topography [Shchepetkin and McWilliams, 2003].

In this study the atmospheric forcing is provided by hourly output from the Coupled Ocean and Atmosphere Mesoscale Prediction System (COAMPS), operated by the Naval Research Laboratory [Hodur, 1997]. COAMPS has a quadruple nested-grid configuration with horizontal resolutions of 81, 27, 9 and 3km for the largest to the smallest nested domain for the atmospheric model. The outputs from the 27km-, 9km-, and 3km-resolution COAMPS are used to force the L0, L1 and L2 domains of ROMS, respectively. The surface latent and sensible heat fluxes, as well as surface evaporation rates, are derived from sea surface air temperatures, sea surface relative humidities, 10m winds and sea surface tem-

peratures (SSTs) from the ocean model, using the bulk formulae proposed by Kondo [1975]. The fresh water flux is computed as the calculated evaporation rate minus the COAMPS precipitation rate (E-P). The wind stress is derived from the 10m winds using the formula of Large and Pond [1982]. The COAMPS winds show good agreement with the observed winds from the Monterey Bay Aquarium Research Institute (MBARI) M2 mooring (Fig. 2).

2.2 ROMS3DVAR Configuration

2.2.1 Incremental 3DVAR Scheme

ROMS3DVAR is an incremental 3DVAR scheme, and it incorporates weak constraints based on hydrostatic balance and geostrophic balance. As such, ROMS3DVAR defines the cost function as follows:

$$\begin{aligned} J_c(\delta\mathbf{x}_c) &= J^{\mathrm{b}}_{\zeta'}(\delta\mathbf{x}_{\zeta'}) + J^{\mathrm{b}}_{\psi''\chi''}(\delta\mathbf{x}_{\psi''\chi''}) + J^{\mathrm{b}}_{TS}(\delta\mathbf{x}_{TS}) \\ &+ J^{\mathrm{o}}_{\zeta}(\delta\mathbf{x}_{\zeta'}, \delta\mathbf{x}_{TS}) + J^{\mathrm{o}}_{uv}(\delta\mathbf{x}_{\zeta'}, \delta\mathbf{x}_{\psi''\chi''}, \delta\mathbf{x}_{TS}) + J^{\mathrm{o}}_{TS}(\delta\mathbf{x}_{TS}) \ . \end{aligned} \tag{1}$$

In the cost function, there are five different control variables, $\delta\mathbf{x}_{\zeta'}$, $\delta\mathbf{x}_{\psi''}$, $\delta\mathbf{x}_{\chi''}$, $\delta\mathbf{x}_T$, and $\delta\mathbf{x}_S$. Following the notation introduced in Li et al. [2007], $\mathbf{x}$ and $\delta\mathbf{x}$ are the model state and its increment on the ROMS3DVAR grid (defined in section 2.2.2), respectively; the subscripts ζ', ψ'', χ'', T, and S correspond to the non-steric SSH, ageostrophic streamfunction, ageostrophic velocity potential, temperature, and salinity. All control variables are 3D, except the non-steric 2D SSH. The background cost function consists of

$$\begin{aligned} J^{\mathrm{b}}_{a\zeta'} &= \frac{1}{2}\delta\mathbf{x}^T_{a\zeta'}\mathbf{B}^{-1}_{a\zeta}\delta\mathbf{x}_{a\zeta'} \\ J^{\mathrm{b}}_{\psi''\chi''} &= \frac{1}{2}\delta\mathbf{x}^T_{\psi''\chi''}\mathbf{B}^{-1}_{\psi''\chi''}\delta\mathbf{x}_{\psi''\chi''} \end{aligned}$$

$$J^{\mathrm{b}}_{TS} = \frac{1}{2}\delta\mathbf{x}^T_{TS}\mathbf{B}^{-1}_{TS}\delta\mathbf{x}_{TS}, \tag{2}$$

where $\mathbf{B}_{\zeta'}$, $\mathbf{B}_{\psi''\chi''}$, and $\mathbf{B}_{TS}$ are the background error covariances corresponding to non-steric SSH, ageostrophic streamfunction and velocity potential, and T and S. The observational cost function consists of

$$\begin{aligned}
J^{\mathrm{o}}_{\zeta'} &= \frac{1}{2}(\mathbf{\Pi}\delta\mathbf{x}_{TS} + \delta\mathbf{x}_{\zeta'} - \mathbf{d}_{\zeta})^T\mathbf{R}^{-1}_{\zeta}(\mathbf{\Pi}\delta\mathbf{x}_{TS} + \delta\mathbf{x}_{\zeta'} - \mathbf{d}_{\zeta}) \\
J^{\mathrm{o}}_{uv} &= \frac{1}{2}(\mathbf{\Gamma}_{\zeta'}\delta\mathbf{x}_{\zeta'} + \mathbf{\Gamma}_{TS}\delta\mathbf{x}_{TS} + \mathbf{\Phi}\delta\mathbf{x}_{\psi''\chi''} - \mathbf{d}_{uv})^T\mathbf{R}^{-1}_{uv}(\mathbf{\Gamma}_{\zeta'}\delta\mathbf{x}_{\zeta'} + \mathbf{\Gamma}_{TS}\delta\mathbf{x}_{TS} + \mathbf{\Phi}\delta\mathbf{x}_{\psi''\chi''} - \mathbf{d}_{uv}) \\
J^{\mathrm{o}}_{TS} &= \frac{1}{2}(\delta\mathbf{x}_{TS} - \mathbf{d}_{TS})^T\mathbf{R}^{-1}_{TS}(\delta\mathbf{x}_{TS} - \mathbf{d}_{TS}) \ ,
\end{aligned} \tag{3}$$

where $\mathbf{d}_{\zeta}$, $\mathbf{d}_{uv}$ and $\mathbf{d}_{TS}$ are the innovation vectors of SSH, the velocity components and T/S, and $\mathbf{R}_{\zeta}$, $\mathbf{R}_{uv}$, and $\mathbf{R}_{TS}$ are the corresponding observational error covariances.

2.2.2 ROMS3DVAR Implementation

ROMS3DVAR executes data assimilation for each nested level. Because all levels use the same formulation. we focus in the subsequent discussions on L2, which has the smallest domain but the highest horizontal resolution. For consistency with the ROMS open boundary conditions, ROMS3DVAR takes the forecast and returns the analysis at all ROMS L2 model grid points but not at those determined by L1.

To handle the large-dimension background error covariance matrices while taking anisotropy and inhomogeneity into account, ROMS3DVAR computes a three-dimensional (3D) correlation matrix as a Kronecker product $\mathbf{C}^{\xi\kappa} \otimes \mathbf{C}^{\eta}$ of a two-dimensional (2D) matrix $\mathbf{C}^{\xi\kappa}$ in the vertical and cross-shore directions and a one-dimensional (1D) matrix $\mathbf{C}^{\eta}$ in the alongshore direction [Li et al., 2007]. Here ξ, η and κ stand for the alongshore, cross-shore and vertical

directions, respectively. ROMS3DVAR also computes the 2D surface correlation matrix $\mathbf{C}^{\xi\eta}$ for non-steric SSH as a Kronecker product, that is, $\mathbf{C}^{\xi} \otimes \mathbf{C}^{\eta}$.

Any $\mathbf{C}^{\xi\kappa}$, $\mathbf{C}^{\xi}$ or $\mathbf{C}^{\eta}$ is grid-based, i.e., correlations are defined between two grid points. A difficulty arises, however, in constructing $\mathbf{C}^{\xi\kappa}$ using the S-coordinate in the presence of variable topography. To illustrate this, let us consider an error correlation $C^{\xi\kappa}(i_1{:}k_1, i_2{:}k_1)$ between a nearshore grid point $(i_1 : k_1)$ and a grid point $(i_2 : k_1)$ 15km offshore and on the same S-level k_1 near the ocean bottom in two different vertical cross-shore sections. In a section south of Monterey Bay, the depth at $(i_1{:}k_1)$ is about 10m and at $(i_2{:}k_1)$ about 500m (Fig. 1). Because the depth of thermocline and mixed layer in this region is about 30m in late summer and less than 100m in winter, $C^{\xi\kappa}(i_1 : k_1, i_2 : k_1)$ between these two 2D grid points should be near zero. In a section north of the bay, however, $C^{\xi\kappa}(i_1{:}k_1, i_2{:}k_1)$ may not necessarily be zero because the depth of the grid point $(i_2{:}k_1)$ is much shallower than 500m. This variation in the correlations presents a fundamental difficulty in constructing $\mathbf{C}^{\xi\kappa}$ using the S-coordinate.

To circumvent the difficulty, the ROMS3DVAR grid employs a Z-coordinate for the vertical discretization. The same staggered ROMS C-grid is used in the horizontal so that the resulting analysis will easily satisfy the lateral boundary condition. In this study the ROMS3DVAR grid uses $N^{\xi} \times N^{\eta} = 82 \times 178$ horizontal curvilinear grid points, and $N^{\kappa} = 24$ vertical Z-levels in L2.

Accordingly, the execution of the incremental 3DVAR given a ROMS forecast on the S-coordinate is accomplished through the following procedure. Using spline interpolation [Akima, 1970], the background state on the Z-coordinate is first formed by transforming the ROMS forecast. After performing the incremental 3DVAR, the lateral boundary condition

along the coastline is imposed on the analysis increment in the Z-coordinate. ROMS3DVAR produces the final analysis by transforming and adding the analysis increment to the original ROMS forecast on the S-coordinate. The transformation back to the S-coordinate is performed on the analysis increment (not the full analysis) to reduce the interpolation errors and thus help maintain the delicate dynamical balance that is attained by forward model integration in the ROMS forecast. A concern may arise, however, concerning whether the final analysis satisfies the lateral boundary condition on the S-coordinate. To address this concern, we examined the final analysis in various experiments and found that the lateral boundary condition is, in general, adequately satisfied.

For the estimation of coastal ocean dynamics with time scales ranging from hours to days, we adopted for ROMS3DVAR a six-hour (6h) assimilation cycle. The first cycle of the day begins by performing a 6h ROMS forecast using the analysis valid at 03UTC as an initial condition; 03UTC corresponds to 7pm local standard time (LST). Once the valid 09UTC 6h ROMS forecast is completed and all observations in the 6h time-bin between 06 and 12UTC are collected, ROMS3DVAR executes the incremental 3DVAR and computes the analysis valid at 09UTC by treating all observations as if they were taken at 09UTC. This completes the first 6h cycle. ROMS3DVAR repeats the cycle four times a day and produces analyses valid at 03, 09, 15 and 21UTC.

3 Estimation of Standard Deviation and Correlation Matrices

Having described the general ROMS-DAS procedure and its grid arrangement, we turn our attention to the issues concerning estimation of the standard-deviation matrices and the correlation matrices for the background errors. This section presents a general method for estimation of the standard deviation and correlation matrices associated with $\mathbf{B}_{\zeta'}$, $\mathbf{B}_{\psi''\chi''}$, and $\mathbf{B}_{TS}$, along with its application to ROMS3DVAR.

3.1 Generation of the Surrogate Data Set

Estimation of the 3DVAR standard deviation and correlation matrices requires a large data set that represents the forecast-error statistics of the data assimilation system. Two methods have been suggested for constructing such a data set. One method uses observations as a sampling of the true state [Hollingsworth and Lonnberg, 1986; Phillips, 1986]. The other method uses the difference between two forecasts valid at the same time but with distinct lead times as a representation of the forecast error [Parrish and Derber, 1992]. This is the so-called NMC method, where NMC stands for the National Meteorological Center (now the National Center for Environmental Prediction). An advantage of the NMC method is the proficient generation of a large data set for any control variable. While Parrish and Derber [1992] cautiously described it as a crude first step, the NMC method has been used operationally at major meteorological centers [e.g., Rabier et al., 1998; Barker et al., 2004]. The usefulness of the data sets generated by the NMC method depend implicitly on the quality and number of observations used in the data assimilation system.

Although the California coastal ocean is one of the most sampled coastal oceans, almost all observations are still sporadic, sampled monthly or even seasonally at irregular locations. An example is the California Cooperative Ocean Fisheries Investigation (CalCOFI) [http://calcofi.org], which provides only seasonal sampling. For ROMS3DVAR, which uses a high-resolution model to forecast coastal ocean dynamics with time scales ranging from hours to days, such observations are simply inadequate for use in either of the two above-mentioned methods of building large, reliable data sets.

We therefore employ an alternate method. With an ensemble of long-time ROMS simulations, our method assembles a surrogate data set where the standard deviation matrices and the correlation matrices are estimated using a limited number of observations. The ensemble is designed to account for forecast errors due to atmospheric forcing, model errors, and initial conditions. For errors in atmospheric forcing, we use two types of atmospheric forcing derived from the hourly COAMPS output for the period from July 20, 2003 to October 31, 2004: 3h means and daily means. For model errors, we use two vertical resolutions in ROMS: 20 levels and 32 levels. For errors in the initial condition, we use six different initial conditions in the four ensembles (see Table 1). These six initial conditions are randomly selected from six independent ROMS simulations. To assemble a surrogate data set that contains time scales ranging from hours to days, we sample at 03, 09, 15 and 21UTC daily, the same times as the ROMS analysis, from November 1, 2003 to October 31, 2004. The two time levels at 03 and 15UTC correspond to 7pm and 7am LST and represent the diurnal variation.

With the ROMS ensemble simulations in hand, a surrogate data set is built using the following procedure. First, we compute the twelve monthly means and form a perturbation

data set by removing the corresponding monthly means. Then the monthly means and the perturbation data set are interpolated to the ROMS3DVAR grid on the Z-coordinate (section 2). Next, the perturbation data set is transformed to the ROMS3DVAR control variables: non-steric SSH is obtained from temperature, salinity, and SSH based on the hydrostatic relation [Eq. (14), Li et al., 2007]. The geostrophic streamfunction and velocity potential are derived from the non-steric SSH, temperature, salinity, and horizontal velocity based on the geostrophic relation [Eq. (17), Li et al. 2007]. The hydrostatic and geostrophic operators used in these transformations are computed using the corresponding monthly means [Li et al., 2006; Appendix A, Li et al., 2007]. Finally, we remove the ensemble monthly mean of the perturbation to enforce the mean-zero condition on the perturbation of the ROMS3DVAR control variables.

Due to the seasonal transition of upwelling in the region, the thermocline shows a significant seasonal variability, which leads to a seasonal variability of correlations, particularly vertical correlations. As such, ROMS3DVAR uses monthly standard deviation matrices and correlation matrices. Each monthly matrix is derived using three months of data including the previous and subsequent months. In this way, transitions from one month to the next are relatively smooth. For the rest of this paper, we discuss August only because our experiments in the following sections use observations taken in August 2003 during the AOSN experiment (see section 5).

For simplicity in presentation, we formulate the standard deviation and correlation matrices for the case where the ocean domain is a 3D box with a flat bottom, but the actual ROMS3DVAR uses an extension of the formulae to an irregular domain. The five control variables that make up the surrogate data are $\triangle x_{\zeta'}^{\xi\eta}$, $\triangle x_{\psi''}^{\xi\eta\kappa}$, $\triangle x_{\chi''}^{\xi\eta\kappa}$, $\triangle x_{T}^{\xi\eta\kappa}$, and $\triangle x_{S}^{\xi\eta\kappa}$. We

index a control variable by $\{\triangle x^{\xi\eta\kappa}(i:j:k,t,m)\}$: $i \in [1, N^{\xi}]$, $j \in [1, N^{\eta}]$, and $k \in [1, N^{\kappa}]$ for the 3D space; $t \in [1, N_t]$ for time with $N_t = 4 \times 92$; and $m \in [1, M_{en}]$ for the ensemble member with $M_{en} = 6$. All formulae also apply to 2D variables $\triangle x^{\xi\eta}_{\zeta'}$, simply suppress the vertical direction, that is, $N^{\kappa} = 1$.

3.2 Standard Deviation Matrices

We first present a method for estimating a 3D standard deviation matrix $\mathbf{\Sigma}^{\xi\eta\kappa}$ from the surrogate data $\{\triangle x^{\xi\eta\kappa}(i:j:k,t,m)\}$: $\mathbf{\Sigma}^{\xi\eta\kappa}$ represents any of $\mathbf{\Sigma}^{\xi\eta}_{\zeta'}$, $\mathbf{\Sigma}^{\xi\eta\kappa}_{\psi''}$, $\mathbf{\Sigma}^{\xi\eta\kappa}_{\chi''}$, $\mathbf{\Sigma}^{\xi\eta\kappa}_{T}$, or $\mathbf{\Sigma}^{\xi\eta\kappa}_{S}$ the five control variables.

3.2.1 Formulation

There are two reasons why the standard deviations computed from the surrogate data cannot be used directly as the standard deviations of the forecast errors. One is that the statistics of the surrogate data do not necessarily represent those of the forecast errors. As we continue to assimilate, the actual forecast errors should become smaller and the deviation from the surrogate data statistics increases. The forecast errors also depend on the spatial distribution of observations over time. A third reason is the presence of small spatial scales in the standard deviation computed from the surrogate data at each grid point (Fig. 3). These small spatial scales may be caused by the topographic variations or noise due to a finite sampling of the surrogate data. Either way, the presence of small spatial scales in the standard deviation introduces undesirable small spatial scales in the analysis increment.

To construct $\mathbf{\Sigma}^{\xi\eta\kappa}$, we first compute the "structure function" $\hat{\sigma}^{\xi\kappa}(i:k)$, which is an

alongshore average of the standard deviation of the surrogate data:

$$(\hat{\sigma}^{\xi\kappa}(i{:}k))^2 \;=\; \frac{1}{N^{\eta}M_tM_{en}}\sum_{j=1}^{N^{\eta}}\sum_{t=1}^{M_t}\sum_{m=1}^{M_{en}}(\triangle x^{\xi\eta\kappa}(i{:}j{:}k,t,m)-\triangle\overline{x}^{\xi\kappa}(i{:}k))^2\ , \tag{4}$$

where

$$\triangle\overline{x}^{\xi\kappa}(i{:}k) \;=\; \frac{1}{N^{\eta}M_tM_{en}}\sum_{j=1}^{N^{\eta}}\sum_{t=1}^{M_t}\sum_{m=1}^{M_{en}}\triangle x^{\xi\eta\kappa}(i{:}j{:}k,t,m)\ . \tag{5}$$

We choose the alongshore average because there are relatively few alongshore variations in the standard deviation of the surrogate data (e.g., Fig. 3). This is also consistent with the separability assumption of alongshore correlations from vertical cross-shore correlations used in the construction of correlation matrices [Li et al., 2006; see also section 3.3 here]. Figure 4 shows the structure function $\hat{\sigma}^{\xi}_{\zeta'}(i)$ of the 2D non-steric SSH, obtained by suppressing the vertical direction in (4) and (5). The properties of the non-steric SSH structure function are discussed further in section 3.2.2.

The standard deviation matrix $\mathbf{\Sigma}^{\xi\eta\kappa}$ is obtained by applying a tuning (scaling) parameter α. A diagonal entry of $\mathbf{\Sigma}^{\xi\eta\kappa}$ corresponding to a grid point $(i{:}j{:}k)$ is given by $\alpha\hat{\sigma}^{\xi\kappa}(i{:}k)$. The tuning parameter α is sought to represent the ratio of actual forecast errors to those of the surrogate data. In practice, α can be estimated using empirical rules or comparisons between forecasts and observations. When the number of observations is limited, a single value of α should be used to define $\mathbf{\Sigma}^{\xi\eta\kappa}$ uniformly over all grid points. As more observations become available, α can become dependent on location to better represent the generally complex distribution of forecast errors.

3.2.2 Application

Due to the limited number of observations, ROMS3DVAR currently uses a single $\alpha_{\zeta'}$ for $\mathbf{\Sigma}^{\xi\eta}_{\zeta'}$, a single $\alpha_{\psi''\chi''}$ for both $\mathbf{\Sigma}^{\xi\eta\kappa}_{\psi''}$ and $\mathbf{\Sigma}^{\xi\eta\kappa}_{\chi''}$, and a single α_{TS} for both $\mathbf{\Sigma}^{\xi\eta\kappa}_{T}$ and $\mathbf{\Sigma}^{\xi\eta\kappa}_{S}$.

The structure function $\hat{\sigma}^{\xi}_{\zeta'}(i)$ for the non-steric SSH is shown in Fig. 4. Far offshore, the steric SSH is dominant, which is consistent with large scale model results [e.g., Fukumori, 1998]. The amplitude of the non-steric SSH decreases as the cross-shore distance from the coast increases. This implies that nearshore SSH changes are due to the water mass convergence and offshore changes due to the T/S changes. We choose $\alpha_{\zeta'} = 0.5$ for $\mathbf{\Sigma}^{\xi\eta}_{\zeta'}$ based on a comparison of the ROMS3DVAR forecasts with altimetry observations.

Figure 5 shows the structure functions $\hat{\sigma}^{\xi\kappa}_{\psi''}(i{:}k)$ and $\hat{\sigma}^{\xi\kappa}_{\chi''}(i{:}k)$ of ageostrophic streamfunction and ageostrophic velocity potential. Note that the streamfunction and velocity potential are not directly observed, but the velocity components are. To determine the scaling parameter $\alpha_{\psi''\chi''}$, we then use an empirical formula for the horizontal velocity suggested by Daley [Chap. 5.2, Daley, 1991]:

$$\frac{1}{2}(E_v^2 + E_u^2) = \alpha_{\psi''\chi''}(L_{g\psi}^{-2}\hat{\Sigma}^2_{g\psi} + L_{\psi''}^{-2}\hat{\Sigma}^2_{\psi''} + L_{\chi''}^{-2}\hat{\Sigma}^2_{\chi''}) \ . \tag{6}$$

The left-hand side represents the variance of differences between the model forecasts and observations in both velocity components. The right-hand side represents the variance of the velocity components computed from the streamfunction and velocity potential of the surrogate data set under the following assumptions: no correlation between ageostrophic streamfunction and ageostrophic velocity potential; no correlation between the errors of the geostrophic velocity and ageostrophic velocity; and isotropy of the ageostrophic streamfunction and the ageostrophic velocity potential. $L_{g\psi}$, $L_{\psi''}$ and $L_{\chi''}$ are decorrelation length scales

for the standard deviations $\hat{\Sigma}_{g\psi}$, $\hat{\Sigma}_{\psi''}$, and $\hat{\Sigma}_{\chi''}$, respectively. Based on the ROMS3DVAR forecasts and the Acoustic Doppler Current Profile (ADCP) measurements at the M2 mooring (Fig. 1), we choose $\alpha_{\psi''\chi''} = 0.38$ for $\mathbf{\Sigma}^{\xi\eta\kappa}_{\psi''}$ and $\mathbf{\Sigma}^{\xi\eta\kappa}_{\chi''}$.

Figure 6 shows the structure functions $\hat{\sigma}^{\xi\kappa}_{T}(i:k)$ and $\hat{\sigma}^{\xi\kappa}_{S}(i:k)$ for temperature and salinity in a vertical cross-shore section. The relatively large values of $\hat{\sigma}^{\xi\kappa}_{T}(i:k)$ near the surface are due to variability forced by the atmosphere. In contrast, the relatively large values of $\hat{\sigma}^{\xi\kappa}_{S}(i{:}k)$ are associated with thermocline variability. Based on a comparison of the ROMS3DVAR forecasts and available mooring observations, we choose $\alpha_{TS} = 0.35$ for $\mathbf{\Sigma}^{\xi\eta\kappa}_{T}$ and $\mathbf{\Sigma}^{\xi\eta\kappa}_{S}$.

3.3 Correlation Matrices

Having constructed $\mathbf{\Sigma}^{\xi\eta\kappa}$, we now present a method for estimating the self-correlation $\mathbf{C}^{\xi\eta\kappa}$ from the surrogate data: $\mathbf{C}^{\xi\eta\kappa}$ represents any of $\mathbf{C}^{\xi\eta}_{\zeta'\zeta'}$, $\mathbf{C}^{\xi\eta\kappa}_{\psi''\psi''}$, $\mathbf{C}^{\xi\eta\kappa}_{\chi''\chi''}$, $\mathbf{C}^{\xi\eta\kappa}_{TT}$, or $\mathbf{C}^{\xi\eta\kappa}_{SS}$, though the vertical correlation are suppressed for the 2D correlation matrix $\mathbf{C}^{\xi\eta}_{\zeta'\zeta'}$. Because ROMS3DVAR uses a Kronecker product-based algorithm $\mathbf{C}^{\xi\eta\kappa} = \mathbf{C}^{\xi\kappa} \otimes \mathbf{C}^{\eta}$ to construct the self-correlation matrix [Li et al., 2007], we describe a method for computing $\mathbf{C}^{\xi\kappa}$ and $\mathbf{C}^{\eta}$. We also present a method for estimating the error cross-correlation matrix $\mathbf{C}^{\xi\eta\kappa}_{TS}$ between temperature and salinity. Although the current ROMS3DVAR neglects the error cross-correlation between ageostrophic streamfunction and velocity potential, it can be incorporated by considering the cross-correlation $\mathbf{C}^{\xi\eta\kappa}_{\psi''\chi''}$.

3.3.1 Formulation

To compute the ROMS3DVAR self-correlation matrices, we assume that the correlations from the surrogate data have the same structure as the forecast errors. This allows us to compute $\mathbf{C}^{\xi\kappa}$ and $\mathbf{C}^{\eta}$ for the five self-correlations directly from the surrogate data.

An entry of $\mathbf{C}^{\xi\kappa}$, corresponding to the correlation between two 2D grid points $(i_1:k_1)$ and $(i_2{:}k_2)$ in the vertical cross-shore section, is

$$C^{\xi\kappa}(i_1{:}k_1, i_2{:}k_2) = \frac{1}{\hat{\sigma}^{\xi\kappa}(i_1{:}k_1)\hat{\sigma}^{\xi\kappa}(i_2{:}k_2)} \frac{1}{N^{\eta}M_tM_{en}} \sum_{j=1}^{N^{\eta}} \sum_{t=1}^{M_t} \sum_{m=1}^{M_{en}} \left(\triangle x^{\xi\eta\kappa}(i_1{:}j{:}k_1, t, m) - \triangle\overline{x}^{\xi\kappa}(i_1{:}k_1)\right) \left(\triangle x^{\xi\eta\kappa}(i_2{:}j{:}k_2, t, m) - \triangle\overline{x}^{\xi\kappa}(i_2{:}k_2)\right) , \quad (7)$$

where $\hat{\sigma}^{\xi\kappa}(i{:}k)$ and $\triangle\overline{x}^{\xi\kappa}(i{:}k)$ are given in (4) and (5). Similarly, an entry of $\mathbf{C}^{\eta}$ between grids j_1 and j_2 in the alongshore direction is

$$C^{\eta}(j_1, j_2) = \frac{1}{\hat{\sigma}^{\eta}(j_1)\hat{\sigma}^{\eta}(j_2)} \frac{1}{N^{\xi}N^{\kappa}M_tM_{en}} \sum_{i=1}^{N^{\xi}} \sum_{j=1}^{N^{\kappa}} \sum_{t=1}^{M_t} \sum_{m=1}^{M_{en}} \left(\triangle x^{\xi\eta\kappa}(i{:}j_1{:}k, t, m) - \triangle\overline{x}^{\eta}(j_1)\right) \left(\triangle x^{\xi\eta\kappa}(i{:}j_2{:}k, t, m) - \triangle\overline{x}^{\eta}(j_2)\right) \quad (8)$$

where

$$(\hat{\sigma}^{\eta}(j))^2 = \frac{1}{N^{\xi}N^{\kappa}M_tM_{en}} \sum_{i=1}^{N^{\xi}} \sum_{j=1}^{N^{\kappa}} \sum_{t=1}^{M_t} \sum_{m=1}^{M_{en}} (\triangle x^{\xi\eta\kappa}(i{:}j{:}k, t, m) - \triangle\overline{x}^{\eta}(j))^2 \quad (9)$$

$$\triangle\overline{x}^{\eta}(j) = \frac{1}{N^{\xi}N^{\kappa}M_tM_{en}} \sum_{i=1}^{N^{\xi}} \sum_{j=1}^{N^{\kappa}} \sum_{t=1}^{M_t} \sum_{m=1}^{M_{en}} \triangle x^{\xi\eta\kappa}(i{:}j{:}k, t, m) \ . \quad (10)$$

If the surrogate data set has a large enough sample, then the resulting correlations are locally smooth between neighboring grid points. However, spurious correlations can occur between remote grid points, leading to undesirably long tails in $\mathbf{C}^{\xi\kappa}$ and $\mathbf{C}^{\eta}$ [Gaspari and Cohn, 1999]. Such long tails can cause noisy analysis increments. A common technique to address this problem is to apply a localization function that retains correlations computed

from the surrogate data within a local neighborhood but suppresses all correlations at large distances. For localization of $\mathbf{C}^{\xi\kappa}$ and $\mathbf{C}^{\eta}$, we use a Gaussian matrix. The final correlation matrix is then given by the Hadamard (or Schur) product of a correlation matrix computed from the surrogate data and the Gaussian matrix.

When using the Hadamard product to construct a correlation matrix, caution should be exercised: The Hadamard product produces a positive definite matrix only if both matrices are positive definite [e.g. Horn and Johnson, 1994]. While the correlation matrix estimated from the surrogate data is generally positive definite, the Gaussian matrix is not necessarily so, especially for large length scales. In ROMS3DVAR a Cholesky factorization is applied to the correlation matrices for preconditioning [section 4.1, Li et al., 2007]. The use of the LAPACK code for the Cholesky factorization acts as a verification of the positive definiteness of $\mathbf{C}^{\xi\kappa}$ and $\mathbf{C}^{\eta}$ [Anderson et al., 1999].

3.3.2 Application

For the localization of $\mathbf{C}^{\xi\kappa}$ and $\mathbf{C}^{\eta}$ in ROMS3DVAR, we use a Gaussian matrix with a horizontal length scale of 50km and a vertical length scale of 400m. For the non-steric SSH, the 2D correlation $C^{\xi\eta}_{\zeta'\zeta'}(i_1 : j_1, i_2 : j_2)$ is described in Li et al. [2007]. The decorrelation length scale of $\mathbf{C}^{\xi}_{\zeta'\zeta'}$ is about 10km nearshore, increases to about 35km further offshore and remains the same beyond 80km offshore. Here the decorrelation length scale is estimated as the distance to the point where the correlation reduces to e^{-1}. In contrast, $\mathbf{C}^{\eta}_{\zeta'\zeta'}$ is basically homogeneous with the decorrelation length of about 30km, although the decorrelation length scale decreases slightly around the latitude of Monterey Bay.

Figure 7 shows the 2D correlations $C^{\xi\kappa}(i_1 : k_1, i_2 : k_2)$ of ageostrophic streamfunction

$\mathbf{C}^{\xi\kappa}_{\psi''\psi''}$ and ageostrophic velocity potential $\mathbf{C}^{\xi\kappa}_{\chi''\chi''}$, for fixed $(i_1{:}k_1)$ at either an offshore or a nearshore location and varying $(i_2{:}k_2)$ in the vertical cross-shore section. The property of the 1D alongshore correlations $\mathbf{C}^{\eta}_{\psi''\psi''}$ and $\mathbf{C}^{\eta}_{\chi''\chi''}$ are similar to those of $\mathbf{C}^{\eta}_{\zeta'\zeta'}$, that is, they are basically homogeneous with the decorrelation length of about 30km.

A slight difference between $\mathbf{C}^{\xi\kappa}_{\psi''\psi''}$ and $\mathbf{C}^{\xi\kappa}_{\chi''\chi''}$ appears in Fig. 7 in the decorrelation length scale: $\mathbf{C}^{\xi\kappa}_{\psi''\psi''}$ has slightly larger vertical but smaller cross-shore decorrelation length scale than $\mathbf{C}^{\xi\kappa}_{\chi''\chi''}$. Otherwise $\mathbf{C}^{\xi\kappa}_{\psi''\psi''}$ and $\mathbf{C}^{\xi\kappa}_{\chi''\chi''}$ have quite similar spatial patterns. In general, the cross-shore decorrelation length scale decreases as the distance from the coast increases, while the vertical decorrelation length scale does not change in the cross-shore direction.

Figure 8 shows the 2D correlation $C^{\xi\kappa}(i_1{:}k_1, i_2{:}k_2)$ of temperature $\mathbf{C}^{\xi\kappa}_{TT}$ for $(i_1{:}k_1)$ fixed at offshore and nearshore locations at depths of 50m and 200m. The behavior of $\mathbf{C}^{\xi\kappa}_{TT}$ is quite similar to that of $\mathbf{C}^{\xi\kappa}_{\psi''\psi''}$ and $\mathbf{C}^{\xi\kappa}_{\chi''\chi''}$, but the decorrelation length scales of $\mathbf{C}^{\xi\kappa}_{TT}$ are slightly larger in the cross-shore direction and smaller in the vertical direction than those of $\mathbf{C}^{\xi\kappa}_{\psi''\psi''}$ and $\mathbf{C}^{\xi\kappa}_{\chi''\chi''}$. Moreover, the cross-shore variation of the vertical decorrelation length scale is fairly small. Variations in the decorrelation length scale – smaller near the surface – can be explained by the dominance of baroclinic normal modes. The properties of the 1D alongshore correlation $\mathbf{C}^{\eta}_{TT}$ are similar to those of $\mathbf{C}^{\eta}_{\zeta'\zeta'}$, as well as those of $\mathbf{C}^{\eta}_{\psi''\psi''}$ and $\mathbf{C}^{\eta}_{\chi''\chi''}$.

The vertical cross-shore correlation of salinity $C^{\xi\kappa}_{SS}(i_1{:}k_1, i_2{:}k_2)$ is shown in Fig. 9. While decorrelation length scales are very similar to those of $\mathbf{C}^{\xi\kappa}_{TT}(i_1{:}k_1, i_2{:}k_2)$, the slant structure of $C^{\xi\kappa}_{SS}(i_1 : k_1, i_2 : k_2)$ nearshore suggests a strong influence of the thermocline variability as observed in $\mathbf{\Sigma}^{\xi\kappa}_{SS}$ (Fig. 6). The larger vertical decorrelation length scales nearshore for salinity are understandable. They reflect the fact that, during upwelling, salinity near the surface is controlled by the amount of upwelled deep water. This feature contributes to the

significance of the inseparability in the cross-shore and vertical directions. The properties of the 1D alongshore correlation $\mathbf{C}^{\eta}_{SS}$ are similar to those of $\mathbf{C}^{\eta}_{\zeta'\zeta'}$ and $\mathbf{C}^{\eta}_{TT}$.

From Figs. 8 and 9, we can see the variation of the vertical decorrelation length scale with depth. The vertical decorrelation length scale becomes larger with depth. This variation is also similar for the ageostrophic streamfunction and velocity potential, though their vertical decorrelation length scales are larger overall than those of temperature and salinity.

In ROMS3DVAR, the cross-correlation $\mathbf{C}^{\xi\eta\kappa}_{TS}$ between temperature and salinity is parameterized by the local cross-correlation vector $\mathbf{r}_{TS}$ and the average of the corresponding self-correlations $\mathbf{C}^{\xi\eta\kappa}_{TT}$ and $\mathbf{C}^{\xi\eta\kappa}_{SS}$ [Eq. (24), Li et al.,2007] . Figure 10 shows $\mathbf{r}_{TS}$ at the surface and at a depth of 75m.

4 Single-Observation Experiments

Experiments using a single observation are simple yet they can yield insight into how a data assimilation system works [Parrish and Derber, 1992; Thepaut et al., 1996]. They demonstrate the function of the background error covariance matrices, the role of the constraints, and the sensitivity of the system to the type and location of observations. They also illustrate how observed information propagates vertically, which is an intricate task for any ocean data assimilation system because of the extremely sparse observations in the deep ocean and the presence of the mixed layer and thermocline. In ROMS3DVAR, this task is undertaken by the hydrostatic balance and vertical error correlations. In addition, single observation experiments are quite effective in verifying the computer code.

4.1 SSH Observation

SSHs are observed with satellite altimetry and tide gauges. SSH observations have been a major data source for ocean data assimilation owing to, not only wide availability, but also their ability to constrain the temperature and salinity profiles through the hydrostatic balance [e.g., Cooper and Haines, 1996]. A weak hydrostatic balance constraint is essential for coastal data assimilation. As shown in Fig. 3 for the standard deviation, the non-steric SSH dominates steric SSH near the coast. This suggests that the use of a strong constraint is unreasonable there. In contrast, the steric SSH dominates in the offshore region. This implies that, without the weak or strong constraint, the background error covariance matrix must include the cross-correlations between total SSH, temperature, and salinity. These cross-correlations are inhomogeneous and have long vertical decorrelation length scales. It is difficult to reconcile such complex relations in the cross-correlations in a way that is computationally feasible for an incremental 3DVAR algorithm.

In the absence of other types of observations, a SSH observation induces analysis increments in the three ROMS3DVAR control variables, $\delta\mathbf{x}^{\mathrm{a}}_{\zeta'}$, $\delta\mathbf{x}^{\mathrm{a}}_{T}$, and $\delta\mathbf{x}^{\mathrm{a}}_{S}$, by invoking the weak constraint for hydrostatic balance. Geostrophic balance is, however, applied as a strong constraint. Thus, the velocity increment is strictly determined by $\delta\mathbf{x}^{\mathrm{a}}_{\zeta'}$, $\delta\mathbf{x}^{\mathrm{a}}_{T}$, and $\delta\mathbf{x}^{\mathrm{a}}_{S}$.

Figure 11 shows the analysis increment using a single SSH observation at an offshore location with a positive innovation, i.e., the value of the observation is higher than that of the ROMS3DVAR forecast. Both steric and non-steric SSH increments are positive, leading to geostrophic velocity increments that have an anti-cyclonic eddy-like structure. The positive steric SSH increment is related to a positive temperature increment and a negative salinity increment. An anisotropic effect of the decorrelation length scale in the

error correlations is visible in the length scales of the increments that are larger in the alongshore direction. If the observation were taken exactly at a ROMS3DVAR grid point, then the non-steric SSH increment would have a similar spatial pattern to the horizontal error correlation $\mathbf{C}^{\xi\eta} = \mathbf{C}^{\xi} \otimes \mathbf{C}^{\eta}$ that assumes horizontal separability [Fig.3, Li et al., 2007]. Therefore, an observation function can act as a smoother when the corresponding observation is not taken at a ROMS3DVAR grid point.

The vertical spread of the observed information is shown in Fig. 12 for the same experiment. The increments are mainly contained above a depth of 150m. The maxima of the temperature and salinity increments are centered at about 50m, which is around the bottom of the mixed layer. Both geostrophic velocity increments have their maximum at the surface but penetrate deeper than the temperature and salinity increments. This is due to the barotropic pressure gradient induced by the non-steric SSH increments.

To examine the sensitivity to observation location, we carry out another experiment with a single nearshore observation. The analysis increments are shown in Fig. 13. All increments have smaller length scales than in the offshore experiments due to the cross-shore inhomogeneity of the decorrelation length scales (section 3.3.2).

The steric SSH increment is much less dominant in comparison with the offshore observation. The steric SSH increment accounts for more than 80% of the total SSH increment in the offshore case, while it accounts for only about 30% in the nearshore case. The non-steric SSH increment is more important than the steric SSH increment nearshore, which is consistent with the ratio of the non-steric and steric error variances as can be inferred from Fig. 3.

4.2 Along-Shore Velocity Observation

In the absence of other types of observations, a velocity observation induces analysis increments in the four ROMS3DVAR control variables, $\delta\mathbf{x}^{\mathrm{a}}_{\psi''}$, $\delta\mathbf{x}^{\mathrm{a}}_{\chi''}$, $\delta\mathbf{x}^{\mathrm{a}}_{T}$, and $\delta\mathbf{x}^{\mathrm{a}}_{S}$, through the weak constraint for the geostrophic balance. Total velocity increments have both geostrophic and ageostrophic components and hence depend on all four control variables. The relative amplitude of the geostrophic and ageostrophic increments depends on the relative amplitude of the corresponding standard deviations: in general, ageostrophy is significant near the coast while geostrophy rules in the offshore region. The SSH increment is completely determined by $\delta\mathbf{x}^{\mathrm{a}}_{T}$ and $\delta\mathbf{x}^{\mathrm{a}}_{S}$ because hydrostatic balance is applied as a strong constraint.

Figure 14 shows the analysis increment using a single observation of alongshore velocity at a nearshore location with a positive innovation. The increment of the total cross-shore velocity shows a classic "butterfly-like" structure. One interesting result is the relative value of the geostrophic velocity and ageostrophic velocity. The geostrophic velocity accounts for about 60% of the total.

4.3 Temperature Observation

Temperature is the most observed variable in the ocean. In the absence of other types of observations, a temperature observation induces analysis increments in two control variables, $\delta\mathbf{x}^{\mathrm{a}}_{T}$ and $\delta\mathbf{x}^{\mathrm{a}}_{S}$; the latter is caused by the cross-correlation $\mathbf{C}^{\xi\eta\kappa}_{TS}$ between them. Both hydrostatic balance and geostrophic balance are applied as strong constraints, hence $\delta\mathbf{x}^{\mathrm{a}}_{T}$ and $\delta\mathbf{x}^{\mathrm{a}}_{S}$ completely determine the SSH and velocity increments.

Figure 15 shows the analysis increment for a single observation of temperature at a

nearshore location with a positive innovation. The increment structures of the velocity components are classic ones. The increment of salinity is obviously due to the cross-correlation. Consistent with Fig. 10, the cross-correlation is negative nearshore, and thus the increments of salinity and temperature are opposite in sign.

5 Application to the AOSN Experiment

This section demonstrates the performance of ROMS3DVAR using observations from the AOSN field experiment in August 2003. ROMS3DVAR was executed in real-time during the AOSN experiment, but the results presented here are from the post-experiment analysis, which we call reanalysis. After describing the AOSN experiment, we mainly focus on two aspects. For demonstration of the robustness, we present results from a rapid flow transition from upwelling to relaxation forced by a change in atmospheric wind stress. For an objective evaluation of the ROMS3DVAR, we compare the results with two sets of independent observations that have not been assimilated. A more detailed description of the reanalysis results can be found in Chao et al. [2007].

5.1 AOSN and ROMS3DVAR

The AOSN experiment in Monterey Bay, California, August 2003, provided an unprecedented number of *in situ* observations by a variety of means and instruments; details of the experiment can be found at the WWW site (http://www.mbari.org/aosn). During August 2003, ROMS3DVAR generated analyses and forecasts in real-time using the 6h forecast cycle [section 2.2]. After the experiment, additional observations became available while

some of the original observations were upgraded. The results presented in this section are obtained using all observations available. The observations include the NAVOCEANO MC-SST Level 2 High Resolution Picture Transmission and Local Area Coverage (HRPT/LAC) 2.2km sea surface temperature dataset (Level 2), underwater glider temperature/salinity profiles, aircraft SSTs, ship temperature/salinity/depth (CTD), and Automatic Underwater Vehicle (AUV) temperature/salinity profiles (http://www.mbari.org/aosn for details about observations).

In this reanalysis experiment we adjusted the weak geostrophic balance constraint. The geostrophic velocity is computed after smoothing is applied to $\delta\mathbf{x}_{\zeta'}$, $\delta\mathbf{x}_T$, and $\delta\mathbf{x}_S$. The smoothing is a spatial average with the weight determined by the Gaussian function e^{-d^2/r_0^2}, where d is the distance and $r_0 = 3$km (approximately two times the model grid size). This additional smoothing is used to reduce an overestimation of the geostrophic velocity at small scales due to the small decorrelation length scales nearshore. Also, it is a practical consideration. It has been shown that a minimization process generally acts first on the larger scales. Thus, the small scales are dealt with mainly during the last stage of the minimization process [Veerse and Thepaut, 1998]. In this experiment, we allow a maximum of 40 iterations for the minimization. In this case, it is possible that the minimization is terminated before it converges to the minimum of the cost function. In this case, some small scale noise may remain. The smoothing can help reduce such small scale noise.

5.2 Transition Between Upwelling and Relaxation

5.2.1 Synoptic Analysis

The ocean responds quickly to changes in the atmosphere. In this region, atmospheric conditions are subject to strong diurnal variations as well as rapid weather changes. The ocean is particularly responsive to wind direction changes. Equatorward wind stress favors the upwelling of the cold water along the coastline, while poleward wind stress causes relaxation, leading to warm surface and sub-surface water. In August 2003, there were two major upwelling events (Fig. 2). The first started on August 6, and was followed by relaxation from August 20 to 22. The second upwelling event started on August 23 and ended by August 30.

The transition from the first upwelling state to the relaxation is shown in Fig. 16 via the sea surface velocity and temperature for August 18 to 21, 2003. Nearshore there is a cold belt with temperatures as low as 10 to 12°C on August 18. The surface currents are systematically equatorward near the coast. During the relaxation event the SST increases, and the cold upwelling centers both north and south of Monterey Bay are replaced by warm water by August 21, 2003. The SST increases are larger than 5°C in some locations.

In association with the transition, the surface current also underwent a significant change. During the upwelling there is a systematic equatorward velocity jet, which follows closely the relatively large SST gradient zone nearshore. This jet is interrupted during the transition. The most significant change occurred within Monterey Bay, where the velocity reverses direction from equatorward to poleward. This direction reversal is well captured by ROMS3DVAR as seen in the comparison with the HF radar measurements presented next.

5.2.2 Comparison with HF Radar Observations

A comparison with HF-radar velocity observations is ambitious for the present ROMS3DVAR because no velocity observations were assimilated. Therefore, the velocity analysis increment is strictly the result of a geostrophic balance [section 4, Li et al., 2007]. In contrast, the HF radar covers mainly Monterey Bay where ageostrophic velocities are significant. Moreover, the tidal velocities in this region can be as large as $0.1-0.2$m/s [Petruncio et al., 1998], which is comparable to the velocities arising from dynamic processes. ROMS does not include tidal forcing yet.

For comparison purposes, we suppress the tidal signal in the HF radar observations by taking the average over August 12 and 19, 2003 for the first upwelling event and over August 20 to 22, 2003 for the relaxation event. The average also reduces the HF radar observational error. The observational error is generally considered to be 5-10cm/s, but an averaged field should have much smaller errors. The corresponding composite maps are made from the ROMS3DVAR reanalysis. The results are shown in Fig. 17.

During the upwelling period, the HF radar observations show a jet-like structure that swerves in and out of the bay with a high speed. It also shows a cyclonic circulation within the bay. During the relaxation period the flow pattern changes drastically. The HF radar shows a cyclonic circulation dominating the entire bay. Differences from the upwelling flow pattern are most striking near the head of the Monterey Peninsula where the flow direction is reversed. These features in the upwelling and relaxation events are captured fairly well in the ROMS3DVAR reanalysis.

To quantify the agreement of ROMS3DVAR with these observations, we use the correlation of the flow patterns observed by the HF radar and the ROMS3DVAR reanalysis. The

correlation is defined by

$$corr = \frac{\Sigma_{m=1}^{M^{HF}}[(u_m^{HF} - \overline{u}^{HF})(u_m^{\mathrm{a}} - \overline{u}^{\mathrm{a}}) + (v_m^{HF} - \overline{v}^{HF})(v_m^{\mathrm{a}} - \overline{v}^{\mathrm{a}})]}{\sqrt{\Sigma_{m=1}^{M^{HF}}[(u_m^{HF} - \overline{u}^{HF})^2 + (v_m^{HF} - \overline{v}^{HF})^2]}\sqrt{(\Sigma_{m=1}^{M^{HF}}[(u_m^{\mathrm{a}} - \overline{u}^{\mathrm{a}}) + (v_m^{\mathrm{a}} - \overline{v}^{\mathrm{a}})^2]}} \ . \quad (11)$$

M^{HF} is the number of the HF radar observation for the corresponding event, the superscript HF stands for the HF radar, and $\overline{\{\cdot\}}$ for the average over M^{HF} observation. The correlation is 0.82 for the upwelling period and 0.77 for the relaxation period. Overall, ROMS3DVAR has the capability to reproduce major features of the velocity field during this particular transition.

Here a natural question is how much the data assimilation improves the representation of the sea surface current described above. To answer this question, a simulation experiment is performed, starting from August 1, 2003. The model configuration is the same as the one used for the data assimilation experiment. The results are shown in the third row of Fig. 17. The simulation produces a realistic southward jet-like current during the upwelling event as shown by the spatial correlation of 0.83 with the HF radar surface currents, but its velocity is over-predicted and the strong flow intrudes into the bay slightly too far. As a consequence, the current direction did not reverse within the bay during the relaxation. In contrast, in the reanalysis the upwelling front is well constrained, and thus the velocity and location of the jet-like current and the reversal of the current direction is realistically reproduced.

5.3 Evaluation with T/S Profiles at Mooring M2

While a variety of evaluations were performed, we present here an evaluation of the ROMS3DVAR reanalysis using independent observations of T/S profiles at Mooring M2.

The M2 mooring records profiles of T/S, velocities, and winds every 10 minutes. M2 is

located near the boundary of warm offshore water and cold upwelled nearshore water (Figs. 1 and 16). Hence, the ocean state at M2 is sensitive to the location of this boundary, which depends on the strength of upwelling. Because of the continuous nature of the measurements and their sensitivity to upwelling, we do not assimilate them but use them instead as an independent data set for evaluation.

A comparison of the ROMS3DVAR reanalysis and the M2 observations is shown in Fig. 18 for August 2003. The M2 observations are averages over each 24-hour time window, while the ROMS3DVAR data is the average of the four analyses produced each day.

During the first upwelling event, the observed M2 temperatures became warmer near the surface, also more gradually below, until August 20, 2003 when the relaxation begins. This suggests that the M2 location was not affected by the upwelled cold water during this upwelling event. In fact, it was controlled by a warm tongue, which is the consequence of a complex "bifurcated-flow" structure of the cold water [see Aug. 18 and 20 in Fig. 16; Rosenfeld et al., 1994]. During the relaxation period, temperatures became cooler, rather than warmer as might have been expected. During the second upwelling event, temperatures rapidly became warmer. Overall, the observations suggest that warm temperatures are generally associated with lower salinity and cold temperatures with higher salinity. The ROMS3DVAR reanalysis shows a good agreement with these observed phenomena. The maximum root-mean squared error (RMSE) of the ROMS3DVAR reanalysis and the observation is centered around the bottom of the mixed layer at about 40m for both temperature and salinity. The maximum RMSE is less than 1.0°C in temperature and 0.15psu in salinity, in comparison with a maximum STD for the M2 observations of 2.3°C in temperature and 0.28psu in salinity.

6 Summary and Discussions

In Li et al. [2007], we proposed a 3DVAR scheme for the coastal ocean. This scheme has the capability of incorporating inhomogeneous, anisotropic and inseparable correlations, and also weak geostrophic and hydrostatic balances. Here we demonstrated ways to take advantage of these capabilities and discussed their importance when the scheme is applied to the California coastal ocean, where inhomogeneity and anisotropy are significant.

Concerning inhomogeneity, it was shown that the horizontal decorrelation length scale offshore can be as much as two times larger than that nearshore for the California coastal ocean. This is true for all the control variables. The decorrelation length scale turns out to be less than the Rossby radius of deformation nearshore, while it is larger than the Rossby radius offshore. One dynamic consequence of this inhomogeneity is that there are significant ageostrophic velocities nearshore. If the decorrelation length scales specified are too large, the ageostrophic velocity may be underestimated nearshore.

The anisotropy was not significant offshore, but it was nearshore. For our case, the alongshore decorrelation length scale can be as large as the alongshore dimension of the entire model domain nearshore (not shown). Dynamically this can be explained by the fact that the wind stress is systematically coherent in the region. However, such a large decorrelation length scale is not desirable in data assimilation [Gaspari and Cohn, 1999]. This complexity of the anisotropy requires a scheme to adjust to it. In ROMS3DVAR we conservatively used Schur production to limit the decorrelation length scale, and thus anisotropy was only partially taken into account.

The standard deviation and correlation of the background error were estimated based

on an ensemble of model simulations. A concern is that the method is an *ad hoc* one without solid theoretical and practical justification. One supportive argument is that the model forecast has larger errors in regions with larger variability. On the other hand, the inhomogeneity and anisotropy may be overestimated. We consider that the error variance structure and correlations obtained are only a preliminary estimate. They need to be improved upon when appropriate observations become available.

ROMS3DVAR assimilated a variety of observational data collected during the AOSN experiment. These data included satellite and aircraft SSTs, ship CDT, T/S profiles from a fleet of gliders and AUVs. These observations came with differing degrees of accuracy and different time and space resolutions. All these observations were assimilated simultaneously, which demonstrated that ROMS3DVAR has a strong capability of handling a large amount of observational data from different sources. On the other hand, ROMS3DVAR can assimilate any subset of the AOSN data. Actually, ROMS3DVAR was used for an Observational System Experiment (OSE) with the AOSN data. The OSE allowed users to assimilate any selected subset of the data. ROMS3DVAR worked reliably during AOSN, and also within the OSE framework. We attribute the reliability and robustness of the developed ROMS3DVAR to the particular 3D error correlations and weak dynamic constraints used.

The experiment with the AOSN observations further demonstrated the capability of the system to reproduce complex flow patterns during upwelling and relaxation, as well as the transition between the upwelling and relaxation regimes. It is especially encouraging that the results from ROMS3DVAR showed reasonable accuracy in comparison with independent observations from the M2 mooring and HF radar. Note that the mooring is located in the Monterey Submarine Canyon, while the HF radar measures the area over the Canyon head.

These locations always present great challenges to coastal ocean modelers. The success of the data assimilation in these locations indicates that ROMS3DVAR has the capability of coping with complex bathymetry and dynamics.

During the AOSN reanalysis, only temperature and salinity observations were assimilated, and the current was constrained through the weak geostrophic and hydrostatic balance. Since the velocity field during the transition between the upwelling and relaxation is reasonably reproduced, the weak geostrophic and hydrostatic balance formulation appears to work reasonably well.

In this paper we presented the overall performance of ROMS3DVAR. Prediction skill is not addressed here, but extensive evaluations are presented in Chao et al. [2007]. The sensitivity to and influence of individual data sets were not addressed. These are crucial and complicated issues, especially for sea surface height observations from satellite altimetry [Fu and Cazenave, 2001] and *in situ* tidal gauge and sea surface velocity observations from HF radar [Paduan and Grabe, 1997]. The particularly challenging difficulty in assimilation of SSH and velocity observations arises from the strong tides in coastal oceans. These issues will be addressed in subsequent papers.

Acknowledgments

The research described in this publication was carried out, in part, at the Jet Propulsion Laboratory (JPL), California Institute of Technology, under a contract with the National Aeronautics and Space Administration (NASA). Computations were performed on an SGI Origin computer through the JPL Supercomputing Project. Support from the Office of Naval Research (ONR) under grant numbers N00014-04-1-0166, N00014-04-1-0191, N00014-04-1-

0401 and N00014-05-1-0293 is gratefully acknowledged. The discussion with Drs. J. Doyle and X. D. Hong at the Navy Research Laboratory is acknowledged. We appreciate Drs. P. Marchesiello, X. Capet, and A. Shchepetkin for setting up the nested ROMS configuration and for many stimulating discussions. We also thank Dr. J. Paduan at the Navy Post graduate School and Dr. I. Shumann at the Navy Research Laboratory for discussions related to the HF radar.

REFERENCES

Akima, H. (1970), A new method of interpolation and smooth curve fitting based on local procedures, *J. the ACM, 4*, 589-602.

Anderson, E., Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen (1999), *LAPACK Users' Guide*, 3rd ed., 407 pp., SIAM, Philadelphia.

Arakawa, A., and V. R. Lamb (1977), Computational design of the basic dynamical processes of the UCLA general circulation model, in *Methods in Computational Physics*, vol 17, edited by J. Chang, pp. 173-265, Academic Press, New York.

Barker, D. M., W. Huang, Y. R. Guo, A. J. Bourgeois, Q. N. Xiao (2004), A three-dimensional variational data assimilation system for MM5: Implementation and initial results, *Mon. Weather Rev., 132*, 897-914.

Blayo, E. and L. Debreu (1999), Adaptive mesh refinement for finite-difference ocean models: first experiments, *J. Phys. Oceanogr., 29*, 1239-1250.

Chao, Y., Z. Li, J.C. McWilliams, and co-authors (2007), Development, implementation and evaluation of a data-assimilative ocean forecasting system off the central California coast, *Deep Sea Res.*, submitted.

Chao, Y., Z. Li, J. Kindle, F. Chavez, and J. Paduan (2003), A high-resolution surface vector wind product for coastal oceans: Blending satellite scatterometer measurements with regional mesoscale atmospheric model simulations, *Geophys. Res. Lett., 30*, 1013, doi10.1029/2002GL015729.

Cooper, M., and K. Haines (1996), Altimetric assimilation with water property conservation. *J. Geophys. Res., 101*, 1059-1077.

Daley, R. (1991), *Atmospheric Data Assimilation, Cambridge atmospheric and space science series*, 457 pp., Cambridge University Press, Cambridge, UK.

Fu, L.-L., and A. Cazenave (Eds.) (2000), *Satellite Altimetry and Earth Sciences: A Handbook of Techniques and Applications*, 463 pp., Academic Press, New York.

Fukumori, I., R. Raghunath, L.-L. Fu (1998), Nature of global large-scale sea level variability in relation of atmospheric forcing: A modeling study, *J. Geophys. Res., 103, C3*, 5493-5512.

Gaspari, G., and S. E. Cohn (1999), Construction of correlation functions in two and three dimensions, *Quart. J. Roy. Meteorol. Soc., 125*, 723-757.

Hickey, B. (1998), Coastal oceanography of Western North America from the tip of Baja California to Vancouver Island. *The Sea, 11*, 345-393.

Hodur, R. M. (1997), The Naval Research Laboratory's Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS^{tm}), *Mon. Weather Rev., 125*, 1414-1430.

Horn, R. A., and C. R. Johnson (1994), *Topics in Matrix Analysis*, 615 pp., Cambridge University Press, New York.

Hollingsworth, A., and P. Lonnberg (1986), The statistical structure of short-range forecast error as determined from radiosonde data. Part I: The wind fields, *Tellus, 38A*. 111-136.

Kondo, J. (1975), Air-sea bulk transfer coefficients in diabatic conditions, *Boundary-Layer Met., 9*, 91-112.

Large, W. G., and S. Pond (1982), Sensible and latent heat flux measurements over the ocean, *J. Phys. Oceanogr., 12*, 464-482.

Li, Z., Y. Chao, and J.C. McWilliams (2006), Computation of the streamfunction and velocity potential for limited and irregular domains, *Mon. Weather Rev., 134*, 3384-3394.

Li, Z., Y. Chao, J.C. McWilliams, and K. Ide (2007), A three-dimensional variational data assimilation scheme for the Regional Ocean Modeling System *J. Atmos. Oceanic Tech.*, submitted.

Marchesiello, P., J. C. McWilliams, and A. F. Shchepetkin (2001), Open boundary conditions for long-term integration of regional ocean models, *Ocean Modelling, 3*, 1-20.

Marchesiello, P., J. C. McWilliams, and A. F. Shchepetkin (2003), Equilibrium structure and dynamics of the California Current System, *J. Phys. Oceanogr., 33*, 753-783.

Parrish, D. F., and J. C. Derber (1992), The National Meteorological Center's spectral-interpolation system, *Mon. Weather Rev., 120*, 1747-1763.

Petruncio, E. T., L. K. Rosenfeld, and J. D. Paduan, (1998), Observations of the internal tide in Monterey Canyon, *J. Phys. Oceanogr., 28*, 1873-1903.

Phillips, N. A. (1986), The spatial statistics of random geostrophic modes and first-guess error, *Tellus, 38A*, 137-161.

Paduan, J. D., and H. C. Graber (1997), Introduction to high-frequency radar: Reality and myth, *Oceanography, 10*, 36-39.

Rabier, F, A. McNally, E. Andersson, P. Courtier, P. Unden, J. Eyre, A. Hollingsworth, and F. Bouttier (1998), The ECMWF implementation of three-dimensional variational assimilation (3D-Var). II: Structure functions, *Quart. J. Roy. Meteorol. Soc., 124*, 1809-1829.

Rosenfeld, L. K., F. B. Schwing, N. Garfield, and D. C. Tracy (1994), Bifurcated flow from an upwelling center: a cold water source for Monterey Bay, *Cont. Shelf Res., 14*, 931-964.

Shchepetkin, A. F., and J. C. McWilliams (2003), A method for computing horizontal pressure-gradient force in an ocean model with a non-aligned vertical coordinate, *J. Geophys. Res., 108(C3)*, 3090, doi:10.1029/2001JC001047.

Shchepetkin, A. F., and J. C. McWilliams (2005), The Regional Ocean Modeling System: A split-explicit, free-surface, topography-following-coordinate ocean model, *Ocean Modeling, 9*, 347-404.

Shchepetkin, A. F., and J. C. McWilliams (2006), Computational kernel algorithms for fine-scale, multi-process, long-time oceanic simulations, in *Handbook of Numerical Analysis: Special Volume: Computational Methods for the Atmosphere and the Oceans*, edited by R. Temam and J. and Tribbia, Elsevier, North-Holland.

Song Y. T., and D. B. Haidvogel (1994), A semi-implicit ocean circulation model using a generalized topography-following coordinate system, *J. Comput. Phys., 115*, 228-244.

Thepaut, J.-N., P. Courtier, P. G. Belaud, and G. Lemaitre (1996), Dynamical structure functions in a four-dimensional variational assimilation: A case study. *Quart. J. Roy. Meteorol. Soc., 122*, 535-561.

Veerse, F., and J.-N. Thepaut (1998), Multiple-truncated incremental approach for four-dimensional variational assimilation, *Quart. J. Roy. Meteor. Soc., 124*, 1889-1908.

Figure Captions

Figure 1: The L2 domain and bathymetry. The bathymetry contour interval is 250m. The black circle shows the location of the M2 mooring at 122.378°W and 36.697°N. The sharp trough near M2 is the Monterey Submarine Canyon.

Figure 2: Hourly alongshore wind stress derived from 10m winds at M2 for August, 2003: observation (top) and COAMPS (bottom). The unit is N/m^2. For the location of M2, see Fig. 1.

Figure 3: Standard deviation of the non-steric SSH (left) and steric SSH (right) computed from the surrogate data at each grid point. The unit is m.

Figure 4: The structure function $\hat{\sigma}^{\xi}_{\zeta'}(i)$ of the non-steric SSH as a function of cross-shore distance from the coast, corresponding to the left panel of Fig. 3.

Figure 5: The structure functions $\hat{\sigma}^{\xi\kappa}_{\psi''}(i:k)$ and $\hat{\sigma}^{\xi\kappa}_{\chi''}(i:k)$ of ageostrophic streamfunction (left) and ageostrophic velocity potential (right). The unit is s^{-1}. The white areas are the locations where there is no single grid point in the water along the entire shore.

Figure 6: The structure functions of temperature $\hat{\sigma}^{\xi\kappa}_{T}(i{:}k)$ (left) and salinity $\hat{\sigma}^{\xi\kappa}_{S}(i{:}k)$ (right). The units are °C for temperature and psu for salinity.

Figure 7: Vertical cross-shore correlation of ageostrophic streamfunction $C^{\xi\kappa}_{\psi''\psi''}(i_1{:}k_1, i_2{:}k_2)$ (top) and ageostrophic velocity potential $C^{\xi\kappa}_{\chi''\chi''}(i_1{:}k_1, i_2{:}k_2)$ (bottom) for fixed $(i_1{:}k_1)$ at an offshore location (left) and a nearshore location (right) and varying $(i_2:k_2)$ in the vertical cross-shore section. The depth of $(i_1{:}k_1)$ is 50m.

Figure 8: Vertical cross-shore correlation of temperature $C_{TT}^{\xi\kappa}(i_1:k_1, i_2:k_2)$ for fixed $(i_1:k_1)$ at an offshore (left) and nearshore (right) location at a depth of 50m (top) and at a depth of 200m (bottom) and varying $(i_2{:}k_2)$.

Figure 9: Same as Fig. 8 but for salinity $C_{SS}^{\xi\kappa}(i_1{:}k_1, i_2{:}k_2)$.

Figure 10: Local cross-correlation $\mathbf{r}_{TS}$ between the temperature and salinity at the surface (left) and at a depth of 75m (right).

Figure 11: Analysis increments for total SSH, steric SSH, temperature, salinity, and geostrophic velocity components at the sea surface using a SSH observation at an offshore location, 123.0°W and 36.8°N. The unit is m for SSH, °C for temperature, psu for salinity, and m/s for velocity. The innovation is −0.048m, i.e., the observed value is lower than the ROMS forecast. The standard deviation of the observation error is 0.025m.

Figure 12: Analysis increments in a vertical section at 36.8°N corresponding to Fig 11.

Figure 13: Same as Fig. 11, except the location of the observation is 122.2°W and 36.8°N nearshore.

Figure 14: Analysis increments for total velocity components, geostrophic velocity components, temperature and salinity at the sea surface using a single observation of the alongshore velocity at a nearshore location, 122.2°W and 36.8°N. The innovation is −0.03m/s. The standard deviation of the observation error variance is 0.1m/s. The units are °C for temperature, psu for salinity, and m/s for velocity.

Figure 15: Analysis increments for temperature, salinity and geostrophic velocity components at the sea surface using a single offshore temperature observation at a nearshore location at 122.2°W and 36.8°N. The innovation is 1.3°C. The standard deviation of the observation error is 0.35°C. The units are °C for temperature, psu for salinity, and m/s for velocity.

Figure 16: Temporal evolution of sea surface temperature (left) and velocity (right) in the ROMS3DVAR reanalysis during the transition period from upwelling (August 18 to 19, 2003) to relaxation (August 20 to 21, 2003).

Figure 17: Comparison of the averaged HF radar surface velocity (top) and the ROMS3DVAR reanalysis (middle) during the upwelling event between August 12 and 19, 2003 (right) and the relaxation event between August 20 and 22, 2003 (left). The corresponding currents from a simulation without data assimilation are also shown (bottom). The color represents the speed, and the unit is m/s.

Figure 18: Comparison of the ROMS3DVAR reanalysis and M2 observations for August 2003: the temporal evolution (left two columns) of temperature (top) and salinity (bottom) for the ROMS3DVAR reanalysis (left) and the M2 observations (center), and the corresponding root-mean-square error (RMSE) and standard deviation (STD) of the M2 observations (right). Both the mooring and ROMS data are averaged over a time window of 24h. Time is in days of August, 2003. The unit is °C for temperature and psu for salinity.

		Model	
		20 levels	32 levels
Atmospheric	3h	2	2
forcing	daily	1	1

Table 1: Allocation of six ensemble members to four settings based on two types of atmospheric forcing and two levels of vertical resolution.

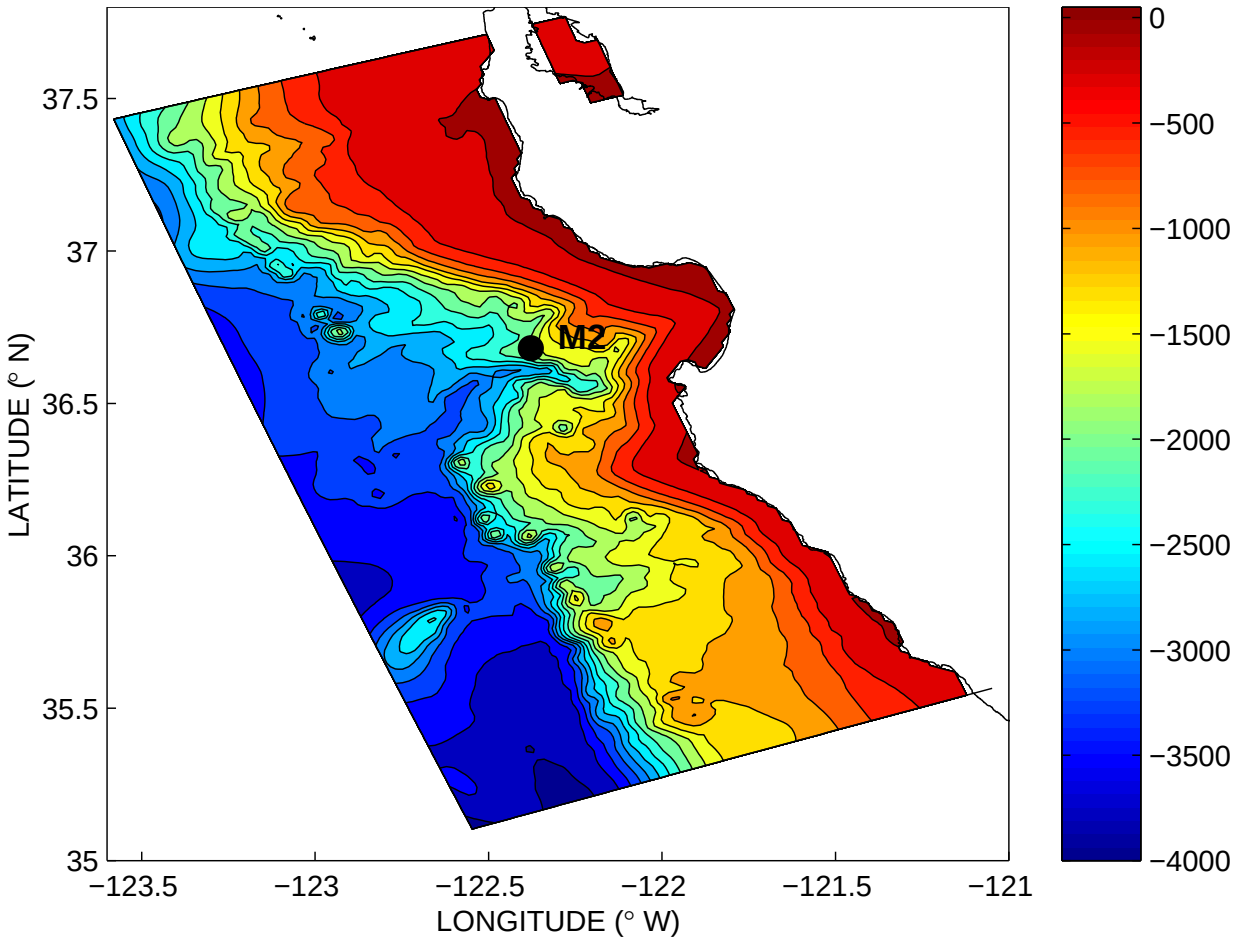

Figure 1: The bathymetry contour interval is 250m. The black circle shows the location of the M2 mooring at 122.378°W and 36.697°N. The sharp trough near M2 is the Monterey Submarine Canyon.

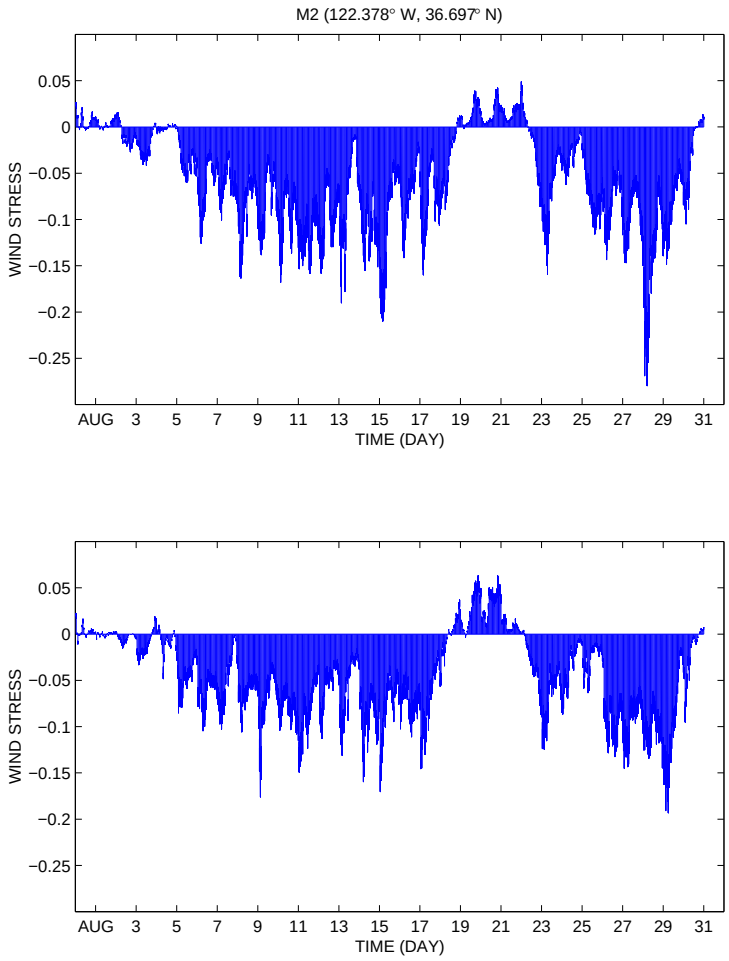

Figure 2: Hourly alongshore wind stress derived from 10m winds at M2 for August, 2003: observation (top) and COAMPS (bottom). The unit is N/m^2. For the location of M2, see Fig. 1.

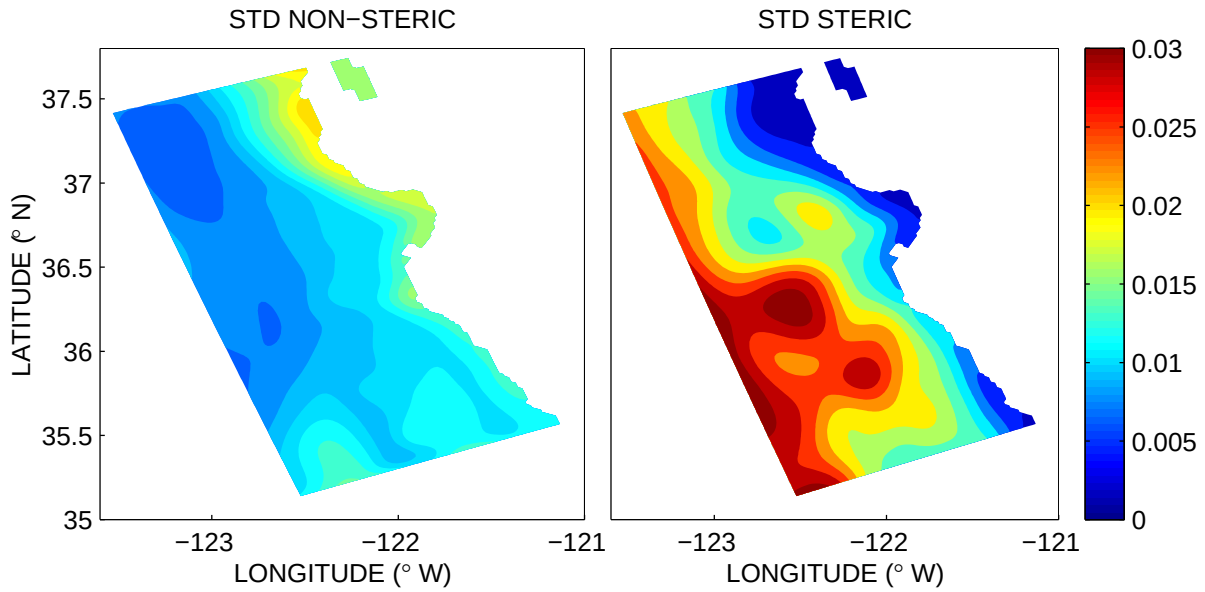

Figure 3: Standard deviation of the non-steric SSH (left) and steric SSH (right) computed from the surrogate data at each grid point. The unit is m.

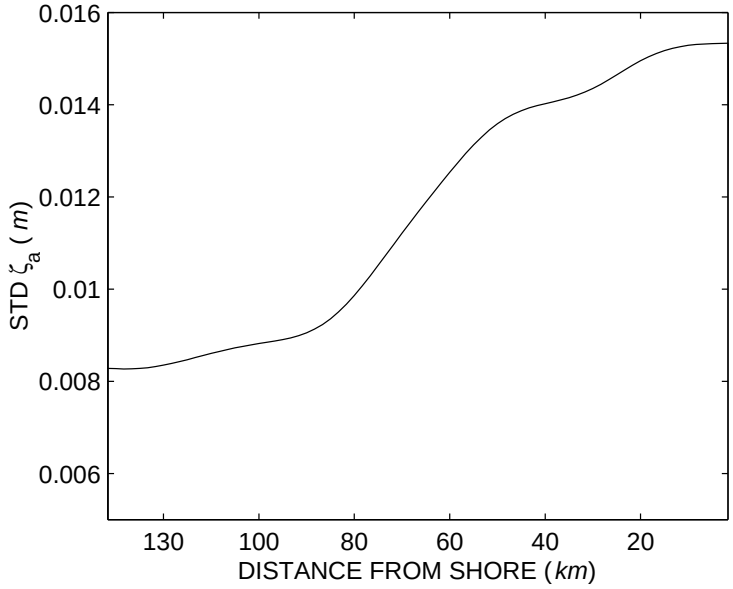

Figure 4: The structure function $\hat{\sigma}^{\xi}_{\zeta'}(i)$ of the non-steric SSH as a function of cross-shore distance from the coast, corresponding to the left panel of Fig. 3.

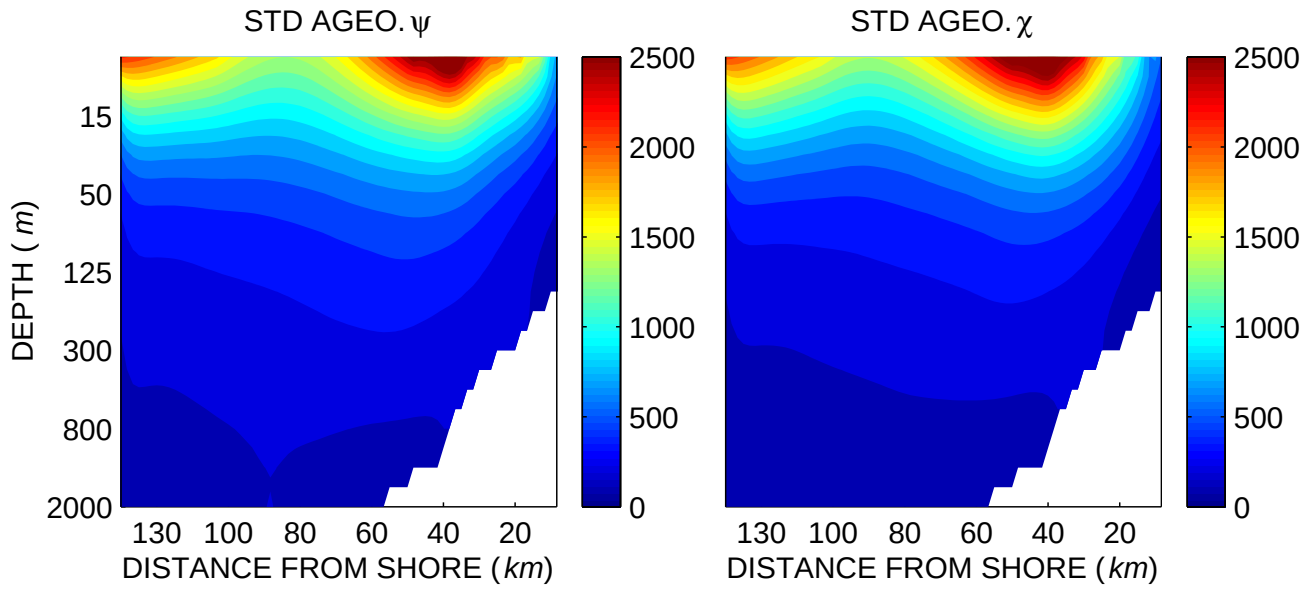

Figure 5: The structure functions $\hat{\sigma}^{\xi\kappa}_{\psi''}(i:k)$ and $\hat{\sigma}^{\xi\kappa}_{\chi''}(i:k)$ of ageostrophic streamfunction (left) and ageostrophic velocity potential (right). The unit is s^{-1}. The white areas are the locations where there is no single grid point in the water along the entire shore.

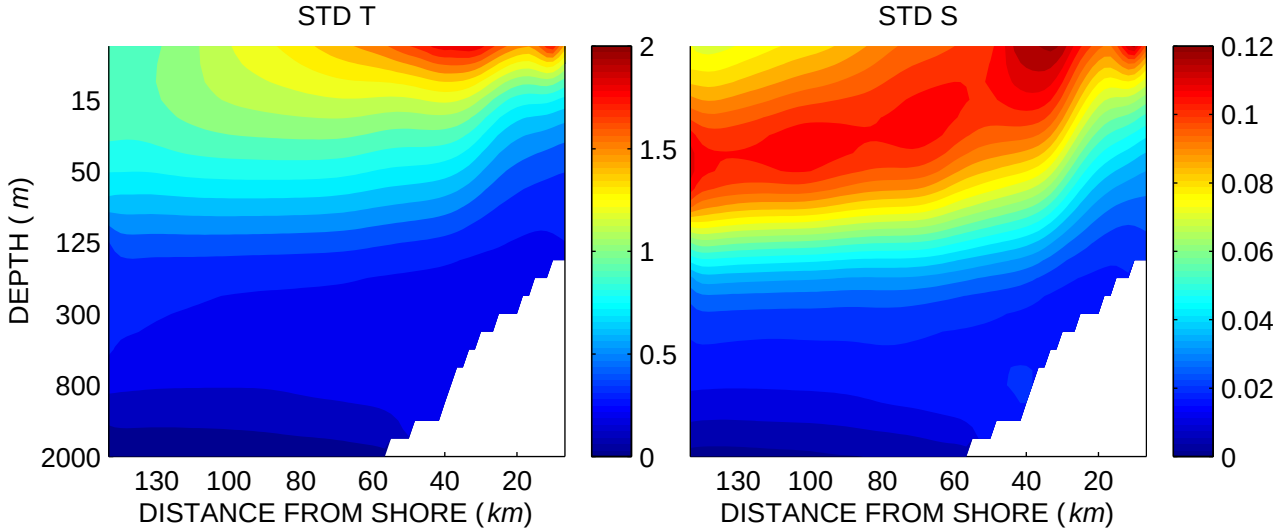

Figure 6: The structure functions of temperature $\hat{\sigma}^{\xi\kappa}_{T}(i{:}k)$ (left) and salinity $\hat{\sigma}^{\xi\kappa}_{S}(i{:}k)$ (right). The units are °C for temperature and psu for salinity.

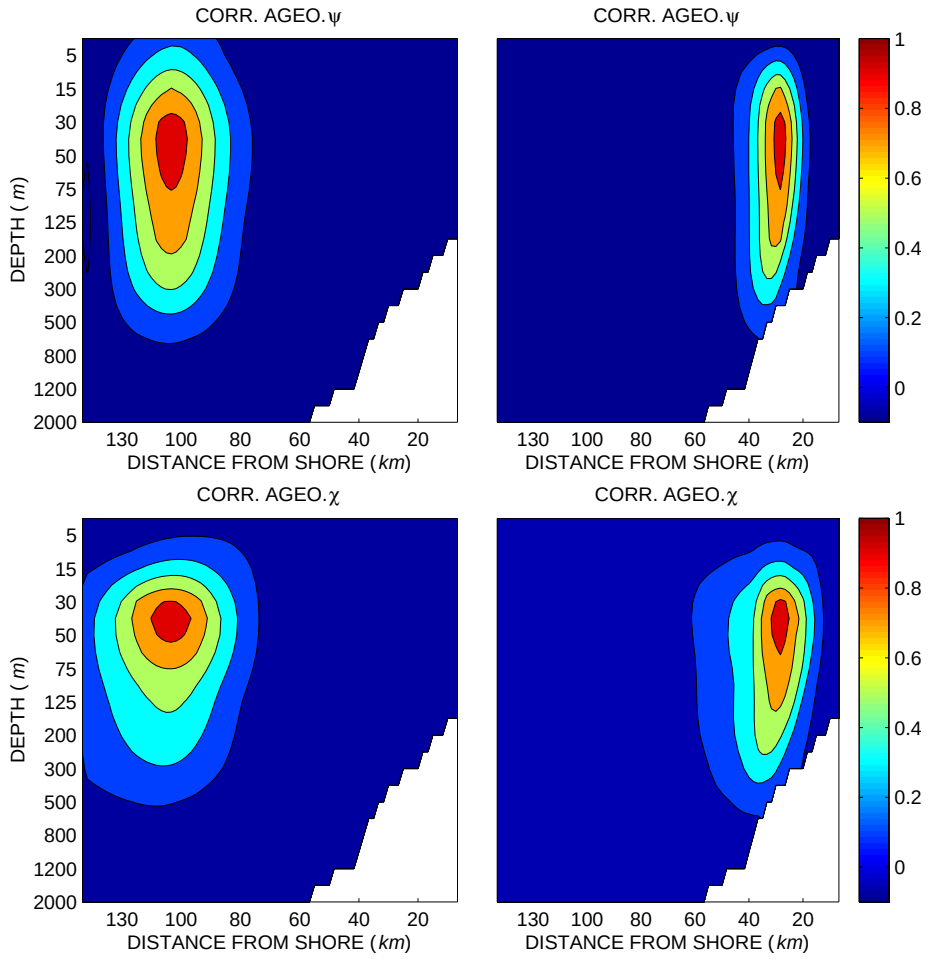

Figure 7: Vertical cross-shore correlation of ageostrophic streamfunction $C^{\xi\kappa}_{\psi''\psi''}(i_1{:}k_1, i_2{:}k_2)$ (top) and ageostrophic velocity potential $C^{\xi\kappa}_{\chi''\chi''}(i_1{:}k_1, i_2{:}k_2)$ (bottom) for fixed $(i_1{:}k_1)$ at an offshore location (left) and a nearshore location (right) and varying $(i_2{:}k_2)$ in the vertical cross-shore section. The depth of $(i_1{:}k_1)$ is 50m.

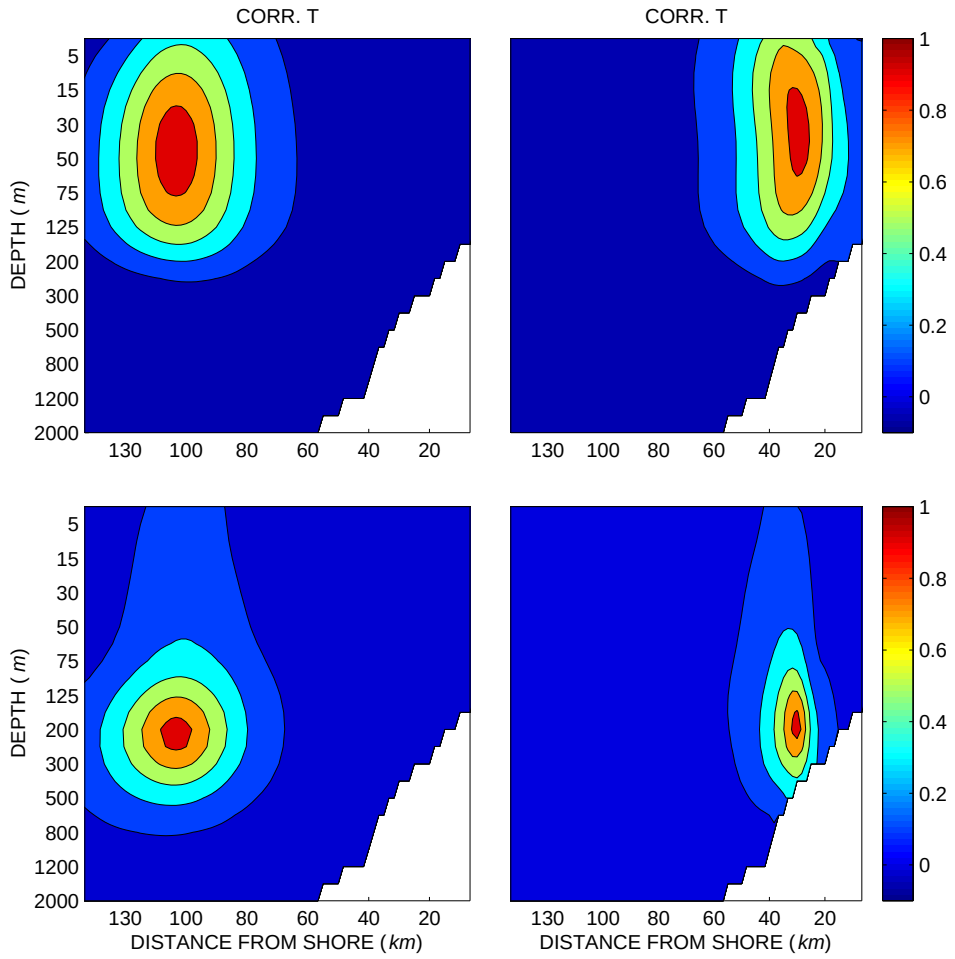

Figure 8: Vertical cross-shore correlation of temperature $C_{TT}^{\xi\kappa}(i_1:k_1, i_2:k_2)$ for fixed $(i_1:k_1)$ at an offshore (left) and nearshore (right) location at a depth of 50m (top) and at a depth of 200m (bottom) and varying $(i_2:k_2)$.

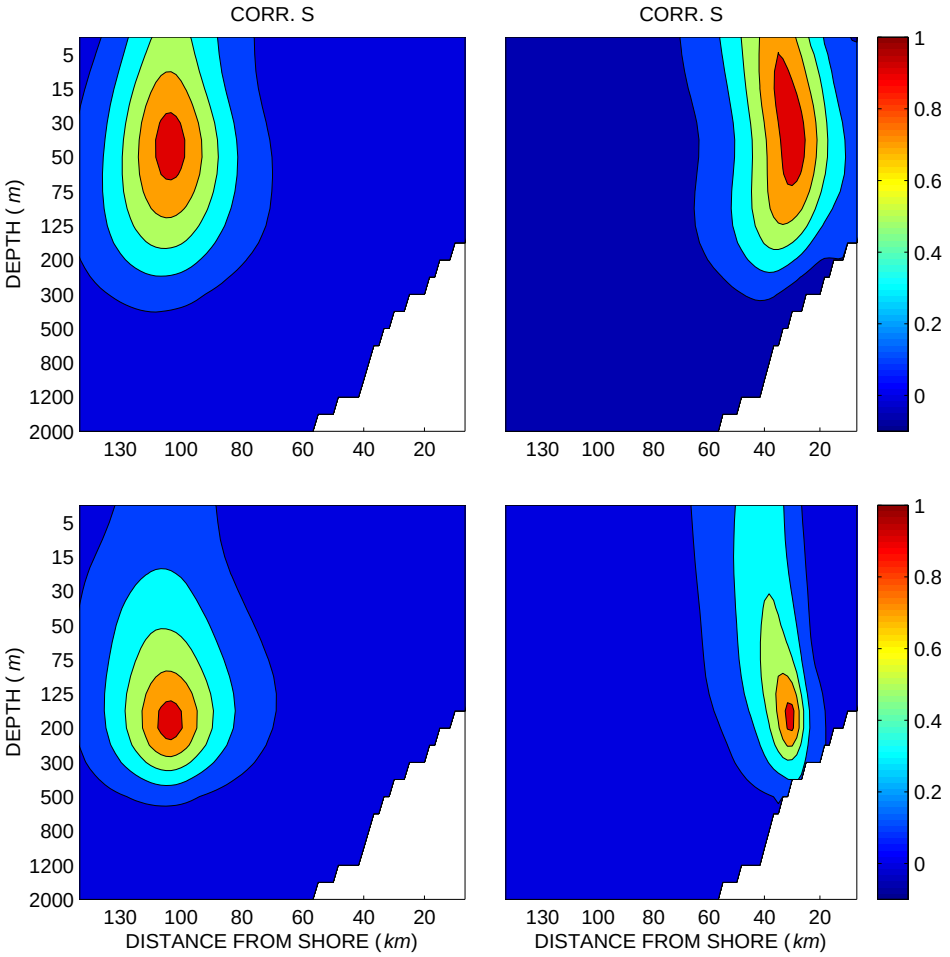

Figure 9: Same as Fig. 8 but for salinity $C_{SS}^{\xi\kappa}(i_1{:}k_1, i_2{:}k_2)$.

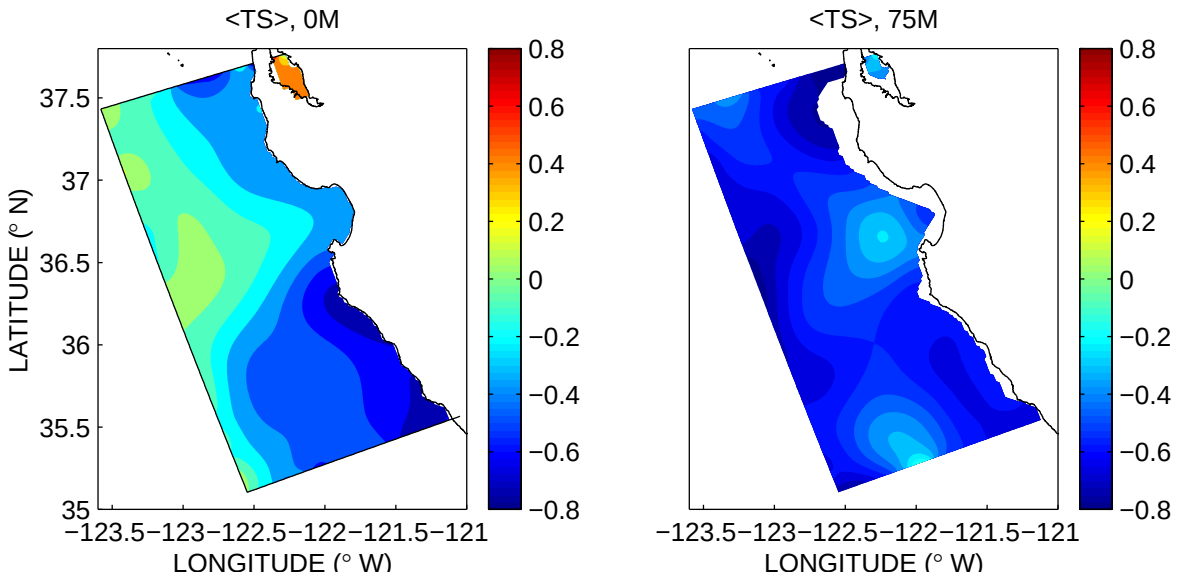

Figure 10: Local cross-correlation $\mathbf{r}_{TS}$ between the temperature and salinity at the surface (left) and at a depth of 75m (right).

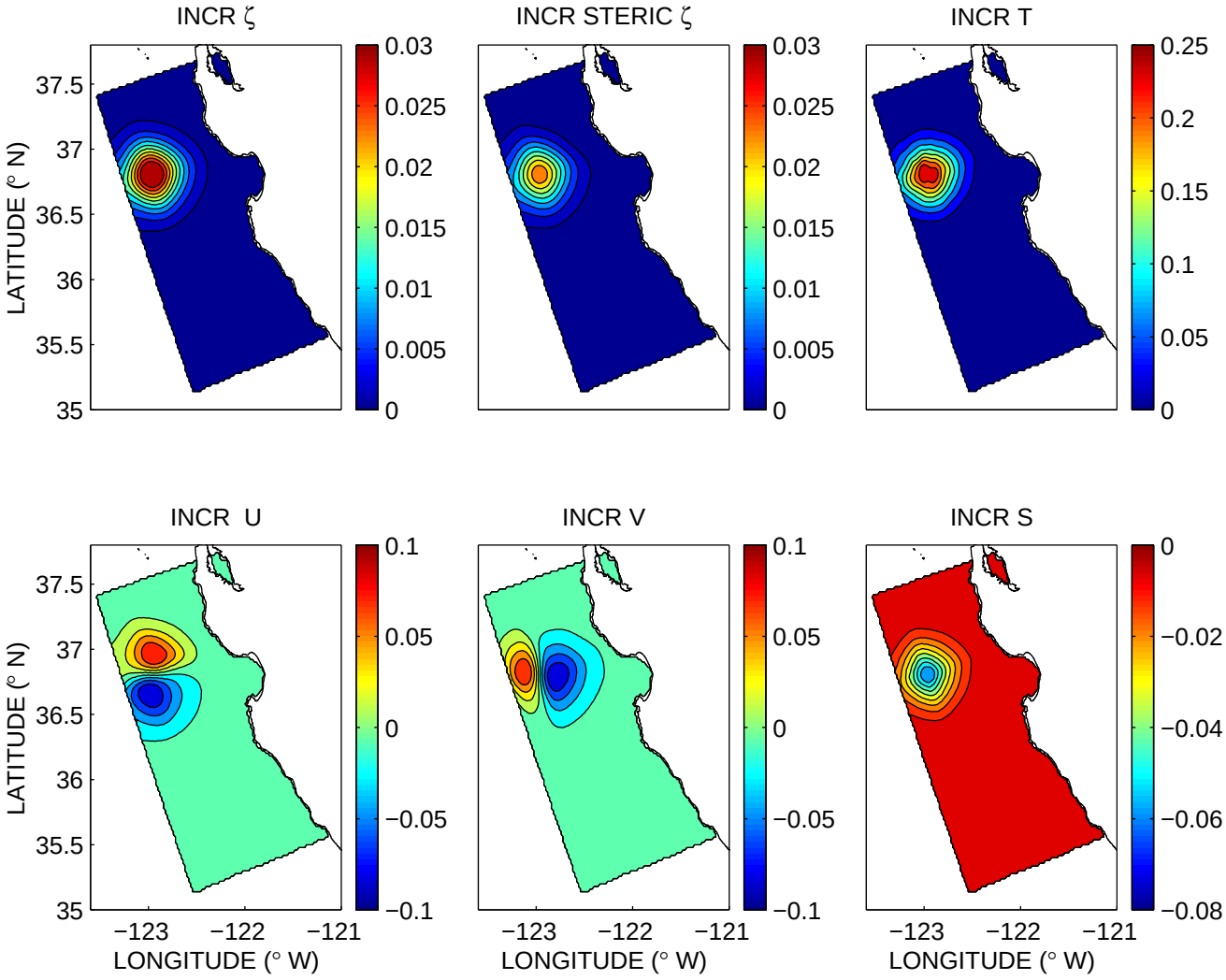

Figure 11: Analysis increments for total SSH, steric SSH, temperature, salinity, and geostrophic velocity components at the sea surface using a SSH observation at an offshore location, 123.0°W and 36.8°N. The unit is m for SSH, °C for temperature, psu for salinity, and m/s for velocity. The innovation is 0.048m, i.e., the observed value is higher than the ROMS forecast. The standard deviation of the observation error is 0.025m.

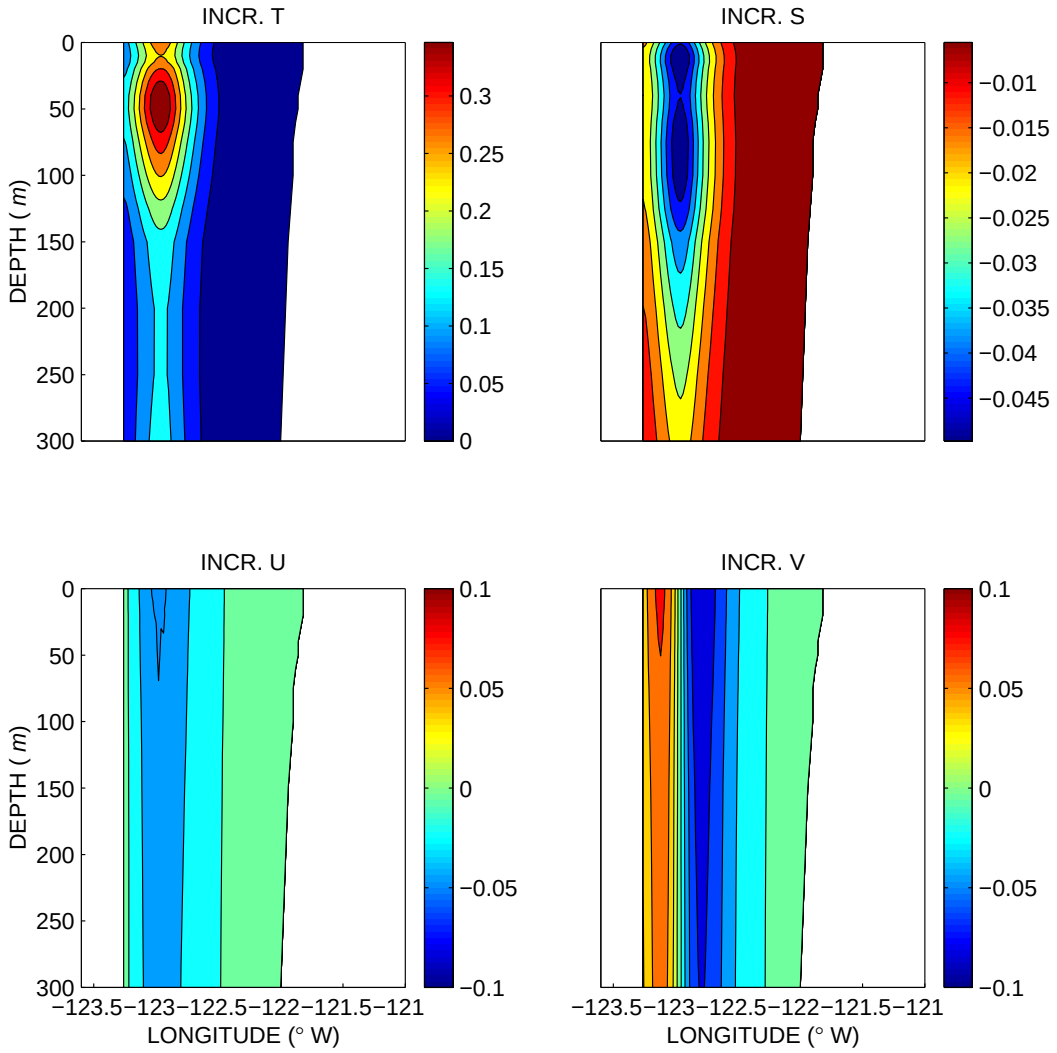

Figure 12: Analysis increments in the vertical section at 36.8°N corresponding to Fig 11.

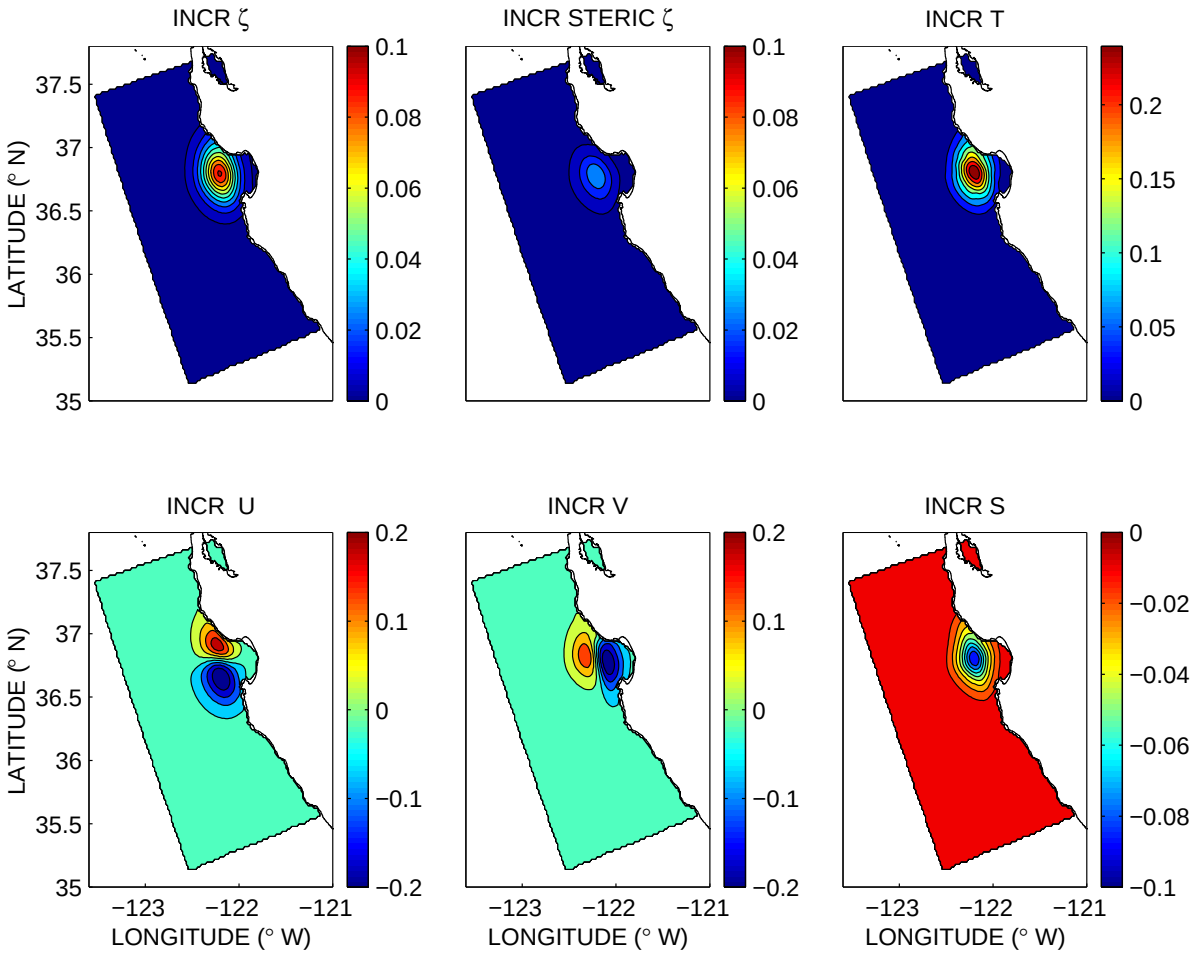

Figure 13: Analysis increments using a SSH observation at a location, 122.2°W and 36.8°N nearshore. The innovation is 0.115m. Others are the same as Fig. 11.

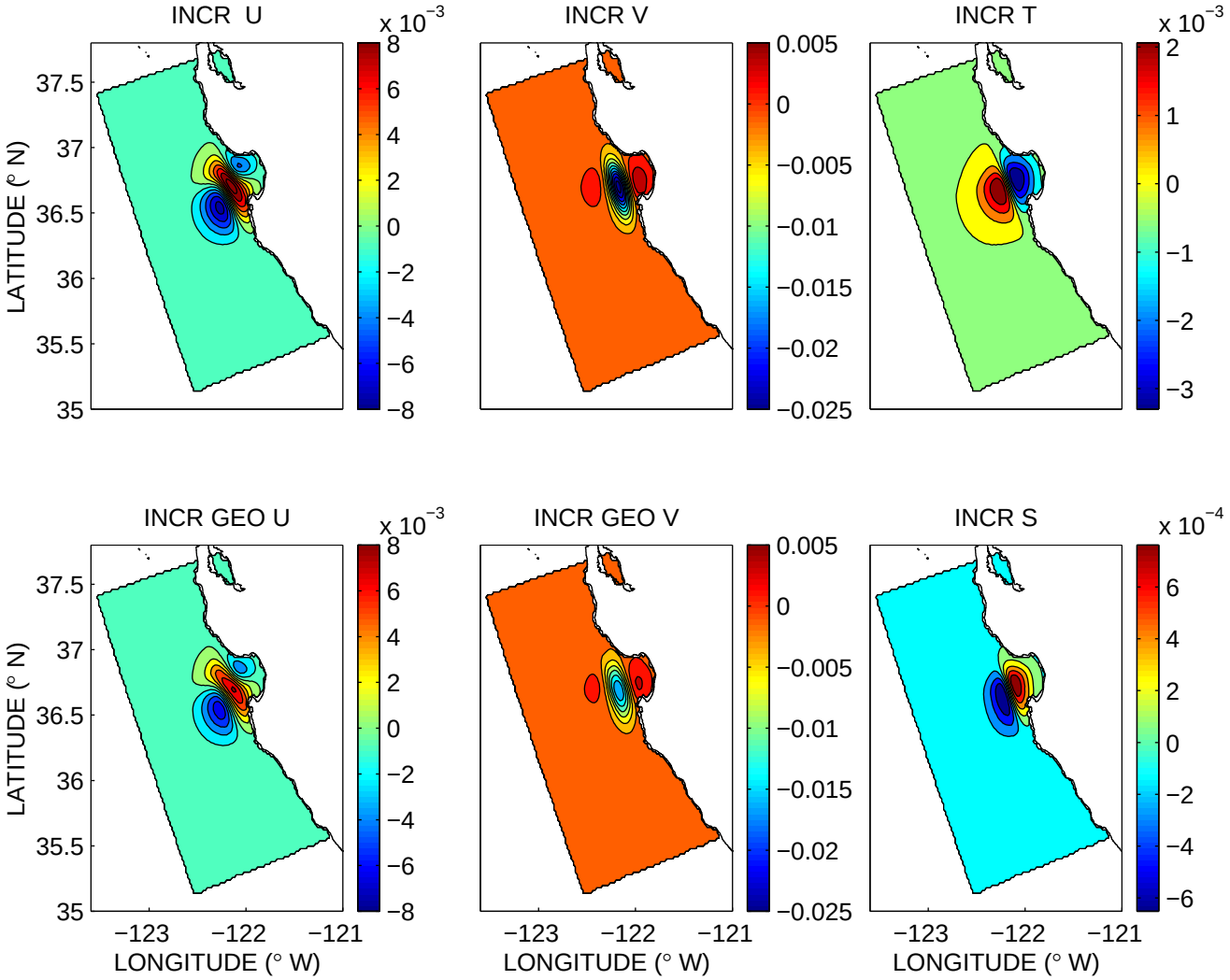

Figure 14: Analysis increments for total velocity components, geostrophic velocity components, temperature and salinity at the sea surface using a single observation of the alongshore velocity at a nearshore location, 122.2°W and 36.8°N. The innovation is −0.03m/s. The standard deviation of the observation error variance is 0.1m/s. The units are °C for temperature, psu for salinity, and m/s for velocity.

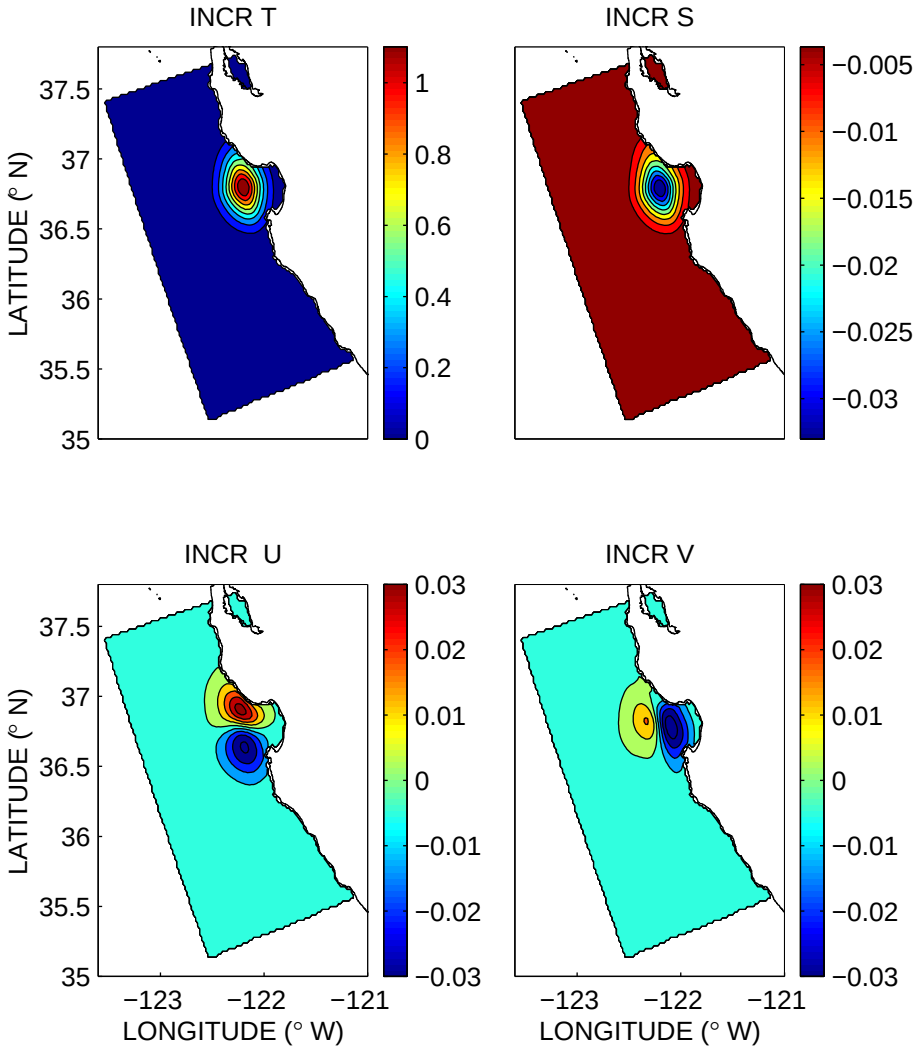

Figure 15: Analysis increments for temperature, salinity and geostrophic velocity components at the sea surface using a single offshore temperature observation at a nearshore location at 122.2°W and 36.8°N. The innovation is 1.3°C. The standard deviation of the observation error is 0.35°C. The units are °C for temperature, psu for salinity, and m/s for velocity.

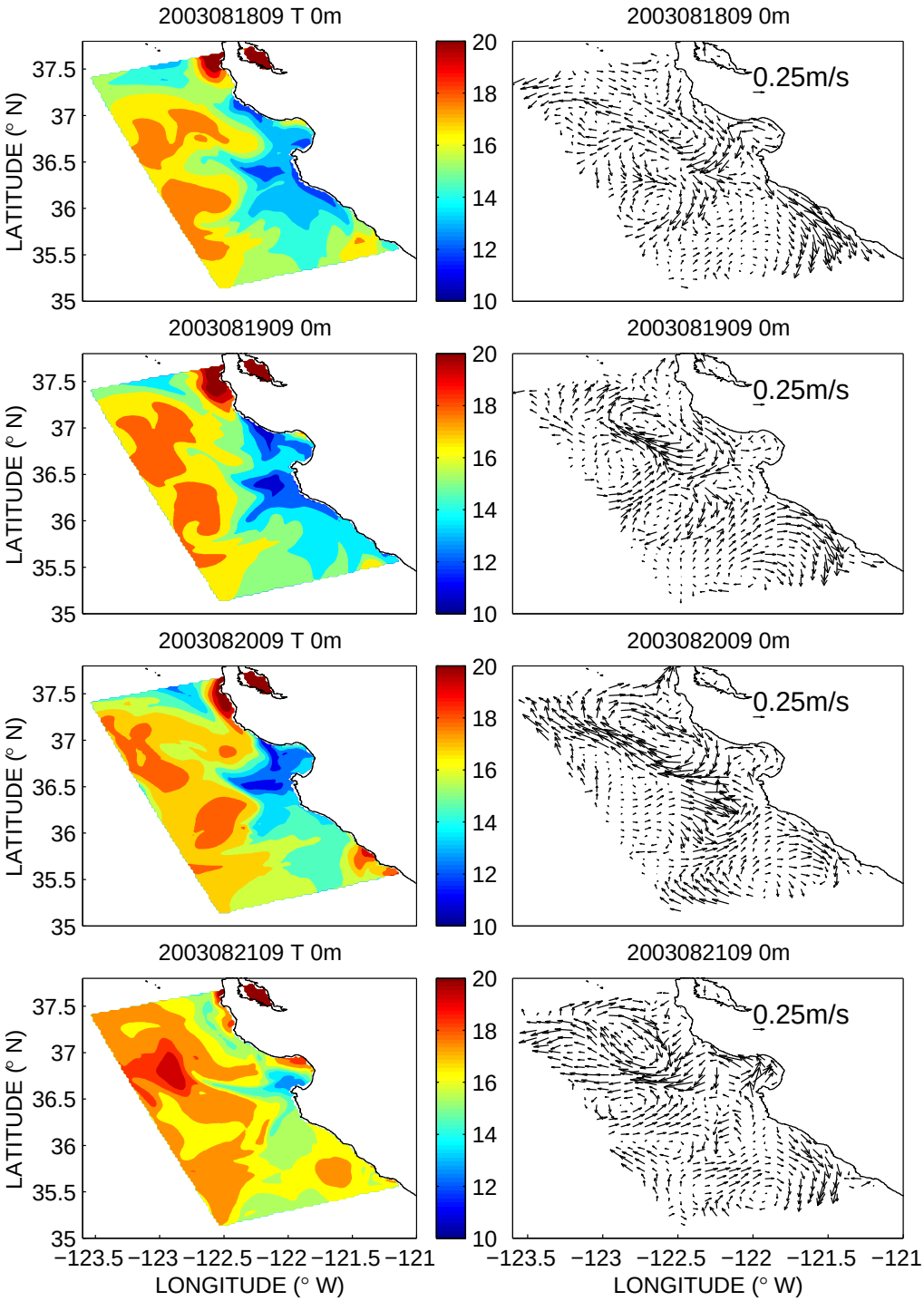

Figure 16: Temporal evolution of sea surface temperature (left) and velocity (right) in the ROMS3DVAR reanalysis during the transition period from upwelling (August 18 to 19, 2003) to relaxation (August 20 to 21, 2003).

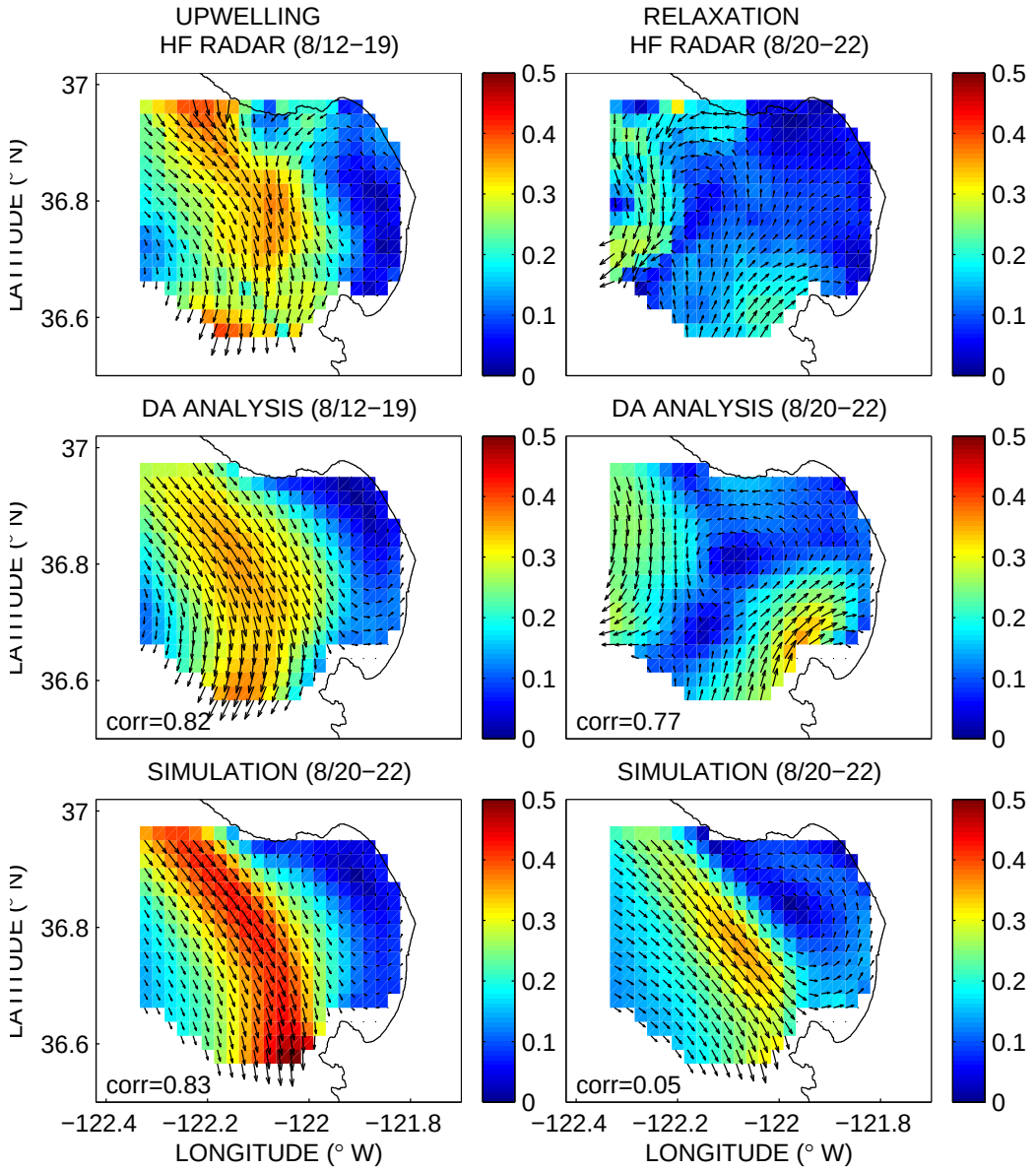

Figure 17: Comparison of the averaged HF radar surface velocity (top) and the ROMS3DVAR reanalysis (middle) during the upwelling event between August 12 and 19, 2003 (right) and the relaxation event between August 20 and 22, 2003 (left). The corresponding currents from a simulation without data assimilation are also shown (bottom). The color represents the speed, and the unit is m/s.

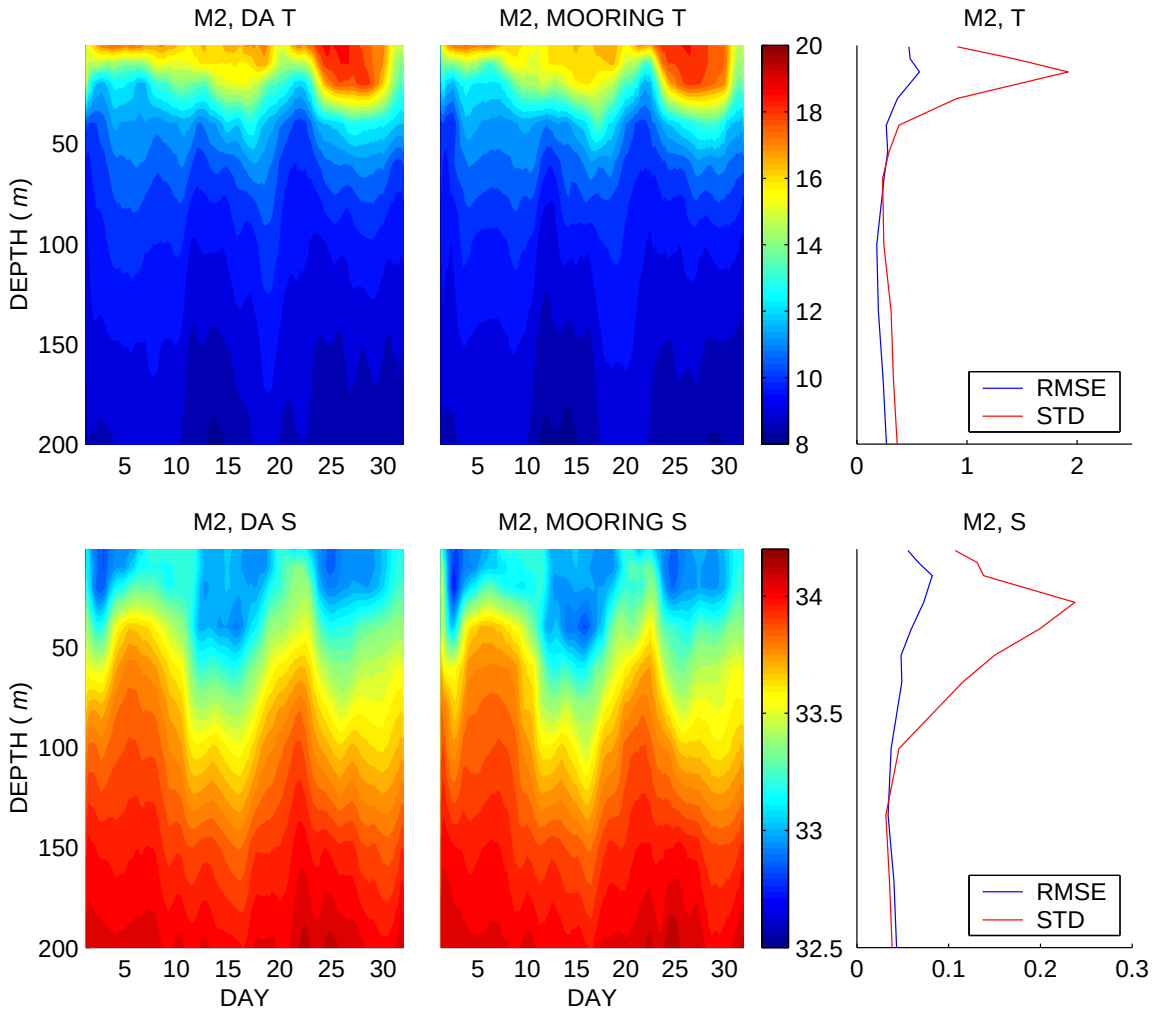

Figure 18: Comparison of the ROMS3DVAR reanalysis and M2 observations for August 2003: the temporal evolution (left two columns) of temperature (top) and salinity (bottom) for the ROMS3DVAR reanalysis (left) and the M2 observations (center), and the corresponding root-mean-square error (RMSE) and standard deviation (STD) of the M2 observations (right). Both the mooring and ROMS data are averaged over the time window of 24h. Time is in days of August, 2003. The unit is °C for temperature and psu for salinity.