Development, implementation and evaluation of a data-assimilative ocean forecasting system off the central California coast

Yi Chao [a,*], Zhijin Li [a], John Farrara [b], James C. McWilliams [c], James Bellingham [d], Xavier Capet [c], Francisco Chavez [d], Jei-Kook Choi [e], Russ Davis [f], Jim Doyle [g], David M. Fratantoni [h], Peggy Li [a], Patrick Marchesiello [i], Mark A. Moline [j], Jeff Paduan [k], Steve Ramp [k]

[a] *Jet Propulsion Laboratory, California Institute of Technology, M/S 183-601, 4800 Oak Grove Drive, Pasadena, CA 91109, USA*
[b] *Raytheon ITSS, Pasadena, CA, USA*
[c] *University of California, Los Angeles, CA, USA*
[d] *Monterey Bay Aquarium Research Institute, Moss Landing, CA, USA*
[e] *Naval Oceanographic Office, Stennis Space Center, MS, USA*
[f] *Scripps Institute of Oceanography, University of California, San Diego, CA, USA*
[g] *Naval Research Laboratory, Monterey, CA, USA*
[h] *Woods Hole Institution of Oceanography, Woods Hole, MA, USA*
[i] *Institut de Recherche pour le Développement (IRD), Noumea, New Caledonia*
[j] *California Polytechnic State University, San Luis Obispo, CA, USA*
[k] *Naval Postgraduate School, Monterey, CA, USA*

ARTICLE INFO

ABSTRACT

The development and implementation of a real-time ocean forecast system based on the Regional Ocean Modeling System (ROMS) off the coast of central California are described. The ROMS configuration consists of three nested modeling domains with increasing spatial resolutions: the US West coastal ocean at 15-km resolution, the central California coastal ocean at 5 km, and the Monterey Bay region at 1.5 km. All three nested models have 32 vertical sigma (or terrain-following) layers and were integrated in conjunction with a three-dimensional variational data assimilation algorithm (3DVAR) to produce snapshots of the ocean state every 6 h (the reanalysis) and 48-h forecasts once a day. This ROMS forecast system was operated in real time during the field experiment known as the Autonomous Ocean Sampling Network (AOSN-II) in August 2003. After the field experiment, a number of improvements were made to the ROMS forecast system: more data were added in the reanalysis with more careful quality control procedures, improvements were made in the data assimilation scheme, as well as model surface and side boundary conditions.

The results from the ROMS reanalysis are presented here. The ROMS reanalysis is first compared with the assimilated data as a consistency check. An evaluation of the ROMS reanalysis against the independent measurements that are not assimilated into the model is then presented. This evaluation shows the mean differences in temperature and salinity between reanalysis and observations to be less than 1 °C and 0.2 psu (practical salinity unit), respectively, with root-mean-square (RMS) differences of less than 1.5 °C and 0.25 psu. Qualitative agreement is found between independent current measurements and the ROMS reanalysis. The agreement is particularly good for the vertically integrated current along the offshore glider tracks: the ROMS reanalysis can realistically reproduce the poleward California Undercurrent. Reasonably good agreement is found in the spatial patterns of the surface current as measured by high-frequency (HF) radars. Preliminary results concerning the ROMS forecast skill and predictability are also presented. Future plans to improve the ROMS forecast system with a particular focus on assimilation of HF radar current measurements are discussed.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

Oceans, because of their vast volume will always be under-sampled. Traditional oceanographic instruments are deployed at either a single point (e.g., drifters, floats and moorings) or through a cross-section (e.g., following a ship track). It is difficult, if not

* Corresponding author. Tel.: +1 818 354 8168; fax: +1 818 354 0966.
E-mail address: Yi.Chao@jpl.nasa.gov (Y. Chao).

impossible, to sample the three-dimensional (3D) volume of the ocean, particularly over a large area. Satellite measurements can provide complete global coverage within a relatively short period of time (usually on the order of days); however, they only provide information at the ocean surface. In order to describe the 3D evolution of the ocean, 3D numerical ocean models have to be used in conjunction with an optimized sampling strategy.

A key challenge is to develop an end-to-end system integrating the observational assets with 3D predictive models. The Office of Naval Research (ONR) has organized the Autonomous Ocean Sampling Network (AOSN). The goal of AOSN is to develop an optimal sampling strategy, using a combined approach of 3D ocean assimilation models, and adapt it in real time. The core AOSN system consists of instruments (both *in situ* and satellite), numerical models with data assimilation capabilities, and adaptive sampling tools that will guide the deployment of observational assets (either fixed or mobile) for optimal performance. The strategy for AOSN is to develop the system incrementally, coupling proof of concept and tests of the system to specific scientific objectives.

The Monterey Bay (MB) has been selected for testing this modeling and sampling strategy because of its unique location and rich oceanographic features, as well as the existing infrastructure for both *in situ* and remote sensing (e.g., high-frequency (HF) radar) observations. MB is a broad and partly deep ($>1000\,\text{m}$), non-estuarine embayment in central California broadly open to the coast ocean. Its oceanography has received considerable study beginning in the early 1930s when MB was the center of a large sardine fishery, and continuing intermittently since the collapse of the fishery in the 1950s (Pennington and Chavez, 2000). The Monterey Bay Aquarium Research Institute (MBARI) initiated in 1989 a program of semi-monthly cruises to stations within and offshore of MB. In addition to the shipboard measurements, MBARI has maintained two moorings since 1989 (M1 and M2). HF radar measurements of MB are also available since 1995. MB is at the center of the central California upwelling region, which has received considerable oceanographic study because of its fisheries and proximity to large human populations. The region is strongly influenced by the process of coastal upwelling, which occurs along eastern ocean margins when southward winds force surface waters offshore, drawing deeper water towards the surface. The characteristic signature of upwelling is a cool band along the coast, typically tens of kilometers wide. This cool band is separated from warmer offshore waters by a series of fronts, plumes and eddies. These features can extend several hundred kilometers offshore. The nutrient-rich upwelling water supports high levels of phytoplankton and higher biological production.

A series of field experiments were conducted in MB in the past decade. The Monterey Upper-water Column Science Experiment (MUSE) was conducted in 2000 (Eriksen et al., 2001; Sherman et al., 2001). During MUSE 2000, the primary observational asset was an array of shallow-water gliders. The spatial scale of this glider network coverage was about 40-km alongshore and 10-km cross-shore. These gliders surveyed the upwelling source region, characterized the transport pathways, and quantified the transport fluxes. The primary observational asset of gliders had been successfully combined with several complimentary observational platforms (e.g., aircraft survey, HF radar, drifters, ship-board ADCP, moorings, and autonomous underwater vehicles or AUVs) using the adaptive sampling strategy. MUSE 2000 demonstrated the utility of real-time observations in providing information on wind forcing and the associated response. However, sea-surface temperatures (SSTs) measured from satellites were apparently insufficient for characterizing the oceanic response, whereas 3D ocean circulation models were not available at that time to predict

where the strongest lateral and vertical transports would develop and hence where circulation and properties of the ocean interior will be resolved by adaptive, multi-platform deployments. These field programs had helped create and demonstrate the capability of individual observational platforms and modeling systems.

The AOSN 2003 field experiment focused on the integration of these individual elements into an operational end-to-end system. The scientific focus of the AOSN experiment was on the day-to-day physical variability of an upwelling center off central California in relation to local atmospheric forcing and the general circulation of the California Current System. This paper describes the process of developing and implementing one of the two independent ocean modeling systems that were used during the 2003 AOSN field experiment.

Our modeling system is based on the Regional Ocean Modeling System (ROMS). The second modeling system is based on the Harvard Ocean Prediction System (HOPS). In addition to documenting the technical development and implementation, this paper will attempt to address three fundamental questions. The first question raised during the planning phase of the experiment was "Can one deliver a ROMS operational forecast in real time continuously during the experiment?" It was agreed that decisions for adaptive sampling would be made at a daily briefing in the "control room". Thus, the time requirement was to deliver the operational forecast within 24 h. Given the variety of observational data sets and demand of computing resources in real time, it was not obvious that this time requirement could be met on a continuous basis. The second question deals with the quality of the model analysis fields (or nowcast): "To what extent can our proposed system simulate the upwelling centers, including their locations, amplitudes and subsequent propagations?" Although the ROMS forecast system was operated in real time during the field experiment in August 2003, this paper only uses the reanalysis results because significant improvements have been made to the models and data assimilation schemes and an updated set of data have been assimilated in the hindcast mode. The updates to the data included the addition of the M1 and M2 mooring T and S data and use of a reprocessed version of the real-time data obtained using improved quality control methods. The updates to the model including using a bulk formulation to compute heat fluxes in place of the climatological fluxes used during the experiment, an increase in the number of vertical layers from 20 to 32 and a decrease in the minimum allowed model depth from 50 to 10 m. Improvements in the data assimilation scheme were achieved by implementing (1) less restrictive quality control criteria, (2) re-estimation of the model error correlations, (3) a reduction in model error variances, (4) a re-estimation of the correlation length scales based on the re-processed glider data and (5) a more sophisticated dynamic balance constraint. Initialized with the six-hourly nowcast fields, a short-term forecast would be made and potentially used to guide the adaptive sampling. Thus, the key question to be addressed was "What is the forecast skill for a given set of targeted variables?" This paper attempts to address these three questions.

2. Observational data sets

The 2003 AOSN field experiment brought together both fixed and mobile observational platforms with advanced data-assimilating 3D ocean models. The operational system included data collection by a variety of platforms and sensors that relayed information to shore within hours in real time where they were being assimilated into numerical models to create both the nowcast and forecast fields of the physical conditions in MB.

We separate the oceanographic data into two categories: one that is assimilated into ocean models, and the other as independent observations that will be used for model evaluations. This section describes the various observational data sets.

2.1. Oceanographic data sets for data assimilation

The data that are assimilated into ocean models consist of the following data types:

- SST from satellite (Advanced Very High Resolution Radiometer or AVHRR, TRMM Microwave Imager or TMI, Advanced Microwave Scanning Radiometer or AMSR) and aircraft operated by the Naval Postgraduate School (NPS),
- sea-surface height (SSH) from the TOEPX/Poseidon satellite altimeter, and
- vertical profiles of temperature and salinity measured by two moorings (M1: 36.75°N, −122.03°W; M2: 36.7°N, −122.39°W), ship CTDs (on R/V *Point Sur*, R/V *John Martin*, R/V *Point Lobos*, and R/V *Shana Rae*), gliders operated by the Scripps Institution of Oceanography (SIO) and Wools Hole Oceanographic Institution (WHOI), Dorado and Remus AUVs operated by MBARI and California Polytechnic State University (CalPoly).

The AVHRR SST data are collected from the Polar-orbiting Operational Environmental Satellites operated by NOAA (also known as NOAA-16 and NOAA-17). The local area coverage (LAC) SST data product is obtained from the Navy GODAE server. LAC SST has a spatial resolution of 1.47 km. Up to four AVHRR SST images could be available every day providing a detailed description of the fine structures in SST in the coastal oceans under cloud-free conditions. However, a significant problem with the AVHRR data is that there are frequently data gaps because the infrared sensor is unable to detect ocean temperature through cloud cover. Because of this limitation of the AVHRR data, a series of airborne mapping flights were carried out to provide the additional information of

SST. During August 2003, a total of 12 flights were conducted by NPS measuring SST and a number of other air–sea parameters. It will be shown in the next section that the aircraft SST measurements play an important role during a week-long period when there was no single AVHRR image available. The microwave SST data from TMI and AMSR have a spatial resolution of 25 km and are assimilated only in the larger domain ROMS.

In addition to remotely sensed SST data, the along-track (with a resolution of 7 km) SSH data are also used for ROMS data assimilation. However, due to the small domain size for the finest resolution ROMS and the land contamination (within 20–40 km of the coastline) of the near-shore SSH data, these data are only used in constraining the larger domain models for the broader-scale California Current System at relatively coarser spatial resolutions.

The major data sources for the ROMS data assimilation are the vertical profiles of temperature and salinity collected from a variety of platforms and sensors. Of particular interests are those vertical profiles collected from the mobile platforms. The Remus and Dorado AUVs, with their fast speed (on the order of 5 knots), can survey a given area within a relatively short period of time (on the order of a day). The gliders, with a typical speed of half a knot, have a much longer mission life on the order of weeks. The 10 gliders operated by WHOI operated for a few weeks and were deployed in the nearshore shallow waters, while the five gliders operated by SIO operated through the month and ranged as far as 100-km offshore and down to 600 m depth. The average number of vertical profiles per day was about 400 during the month-long experiment with the maximum daily profiles exceeding 800 during the mid-August (Fig. 1). All these vertical profiles of temperature and salinity were assimilated into ROMS within 24 h during the AOSN experiment.

2.2. Oceanographic data sets for model evaluations

In addition to the above oceanographic data sets that were assimilated into the ROMS ocean models, the following
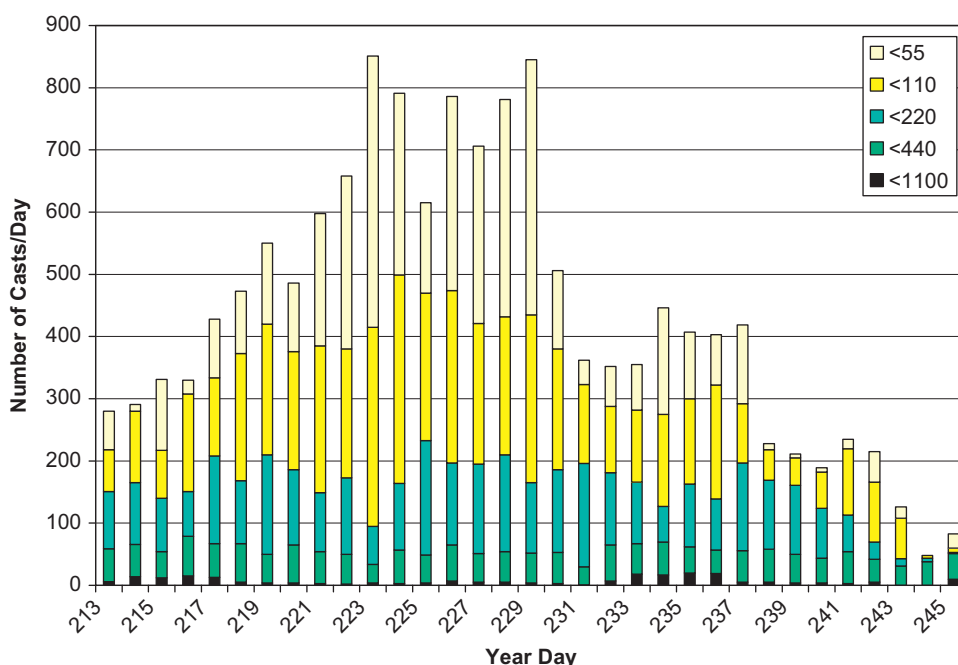

Fig. 1. Time series of the number of vertical temperature and salinity profiles collected during the August 2003 AOSN field experiment. The depth range (in unit of meters) of the measurements is listed in the box in the upper right corner.

measurements were not used for data assimilation and serve here as independent observations for model evaluations:

- current measurements at mooring locations M1 (36.75°N, −122.03°W) and M2 (36.7°N, −122.39°W),
- surface current measurements from HF radars, and
- the vertically integrated current derived from gliders.

The mooring data consist of vertical profiles of temperature, salinity and current in the upper few hundred meters. The surface currents are measured by four HF radar sites around the MB at Santa Cruz, Moss Landing, Point Pinos, and Naval Postgraduate School. The HF radar surface current measurements are available hourly and have a spatial resolution in the range of 1–3 km. Both the SIO and WHOI gliders directly estimated the vertical-average current from a combination of positions gathered when the gliders surfaced and the vehicle's horizontal velocity through the water deduced from compass heading and vertical vehicle velocity. This yields the time-averaged water velocity over the time the glider is diving and ascending but, because the vertical velocity is reasonably constant, this is a good estimate of the vertical average of current. Compass calibration errors contribute an error, primarily to the cross-track component, that is typically 1 cm/s or less. Vertical ocean velocities confound the deduction of horizontal speed through the water from the vehicle's vertical velocity, contributing errors of $O(1 \, cm/s)$ to the estimate of along-track velocity. Most of the SIO-glider velocities are averages to 400 m taken over about 3 h while the WHOI gliders average from the surface to the shallower of the bottom or 200-m depth; because these gliders did not get a position after every dive, the reported averages often apply to several sequential dives taken over a few hours.

3. Atmospheric forcing

The local wind forcing plays an important role in driving the coastal circulation and variability (Hickey, 1998). Alongshore

upwelling-favorable winds drive Ekman transport that moves surface water offshore. The transport divergence at the coastline draws deeper water towards the surface. The spatial gradients in wind or wind curl also can drive a divergence of currents. The resultant "Ekman pumping" is a fundamental forcing agent for coastal circulation and variability.

For a broader region off the US west coast, we use a blended coastal-wind product from atmospheric models and satellite scatterometers at a spatial resolution of 25 km (Chao et al.,
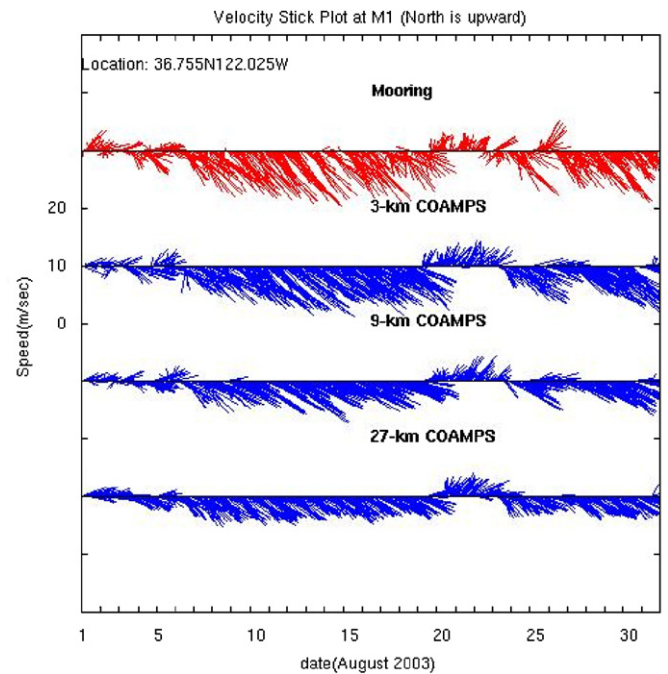

Fig. 3. Hourly wind vector stick plots at the M1 mooring location during August 2003 are shown for the observed (top panel in red) and simulated COAMPS at three spatial resolutions of 27, 9, and 3 km (bottom panels), respectively. North is towards the top of page.
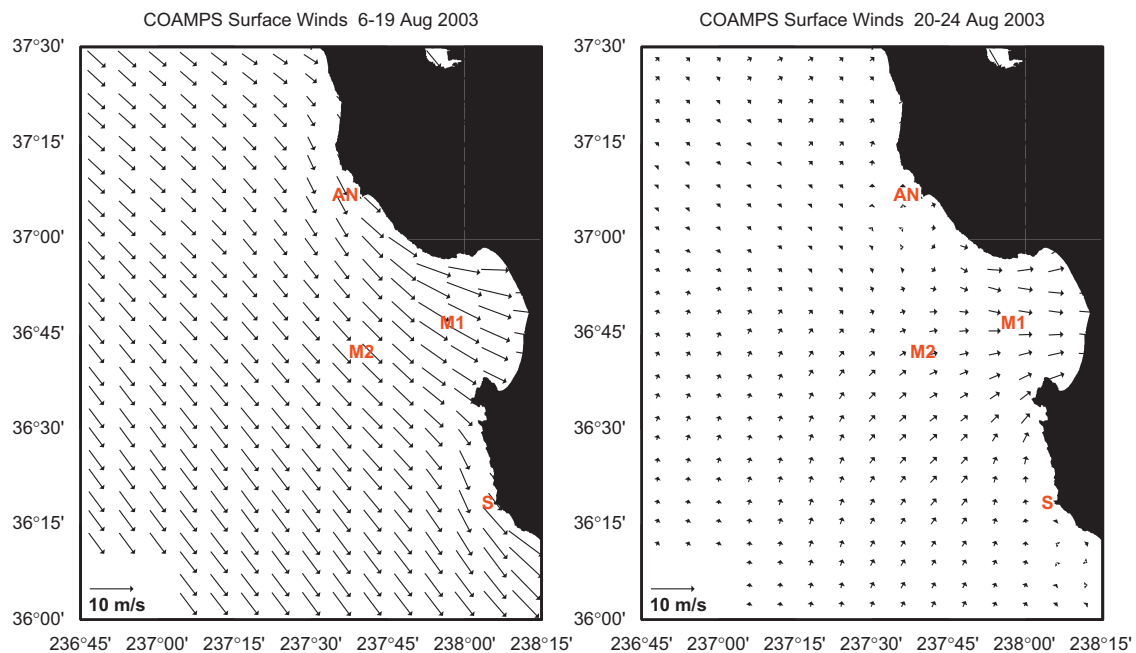

Fig. 2. The 3-km COAMPS wind map during the upwelling (left) and relaxation phases (right). The locations of the M1 and M2 moorings are shown by the symbols in red. Point Año Nuevo (AN) and Point Sur (S) are shown by the symbols in red.

2003). The atmospheric winds are obtained from the 27 km, hourly operational wind products from Fleet Numerical Meteorology and Oceanography Center (FNMOC) produced by Coupled Ocean-Atmosphere Mesoscale Prediction System (COAMPS®) Hodur, 1997). The satellite wind measurements from QuikScat are derived from the Level 3 data product with a resolution of 25 km every 12 h provided by Remote Sensing Systems (http://www.remss.com). The blending algorithm is described in detail by Chao et al. (2003). Both the COAMPS and QuikSCAT data were obtained in real time within 24 h during the AOSN experiment, and combined into a blended wind product to drive ocean models.

The atmospheric circulation over the ocean is strongly influenced by coastal topography, creating small-scale variability in the wind field (Burk et al., 1999; Haack et al., 2001; Capet et al., 2004). For the smaller domain ocean model around the MB, a higher resolution atmospheric model operated in real time during the AOSN experiment by the Naval Research Laboratory (NRL) was used. This high-resolution COAMPS analysis uses four nested grid meshes centered over the MB with horizontal grid increments of 81, 27, 9, and 3 km. The model top is at 30 km with 40 vertical levels, with a vertical grid increment of 10 m at the lowest level. The initial fields for the model are created using an incremental data assimilation procedure based on multi-variate optimum interpolation analyses of upper-air sounding, surface, aircraft and satellite data that are quality controlled and blended with the 12-h COAMPS forecast fields. The initial SST analysis is created through the multi-variate optimum interpolation of available ship, buoy, and satellite observations performed on each model grid mesh. Real-data lateral boundary conditions make use of Navy Operational Global Analysis and Prediction System forecast fields. The topographic data for the real-data simulations are based on the US Defense Mapping Agency's 100-m resolution data set. The COAMPS atmospheric model was executed twice daily during the experiment for a forecast duration of 72 h. The performance of the real-time COAMPS modeling systems as applied in support of the AOSN activities can be found in a companion paper in this special issue (Doyle et al., 2008).

During the spring–summer season, the MB circulation is characterized by two alternate states that occur repeatedly. The upwelling state, which prevails most of the time, is characterized by two upwelling centers in the north (near Pt Año Nuevo; see Fig. 2) and south (near Pt Sur; see Fig. 2) of the MB, respectively. These upwelling centers are presumably generated by the upwelling favorable winds blowing from north to south. During strong upwelling events, a band of cold water is often seen extending across the mouth of MB (Rosenfeld et al., 1994). The relaxation state refers to the relatively shorter periods when upwelling-favorable (or equatorward) winds are
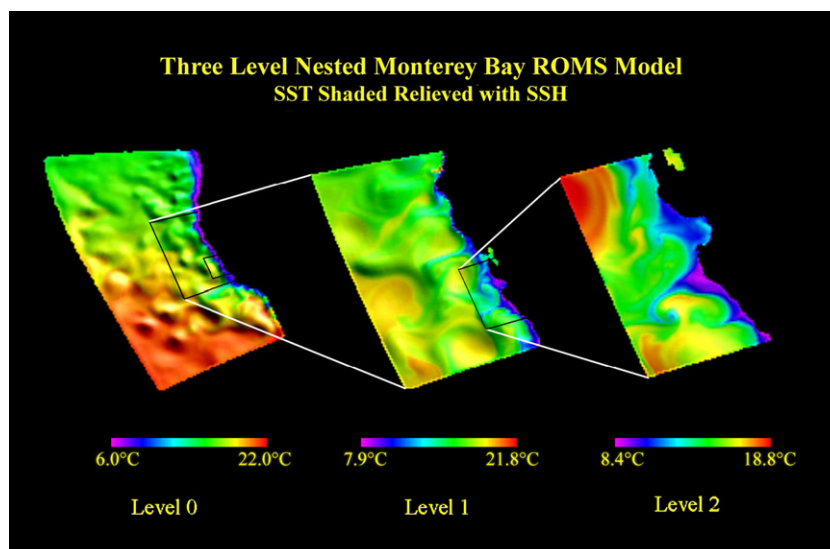

Fig. 4. A nested ROMS configuration with the US West coast ROMS at 15 km, the central California coastal ROMS at 5 km, and the Monterey Bay ROMS at 1.5 km.
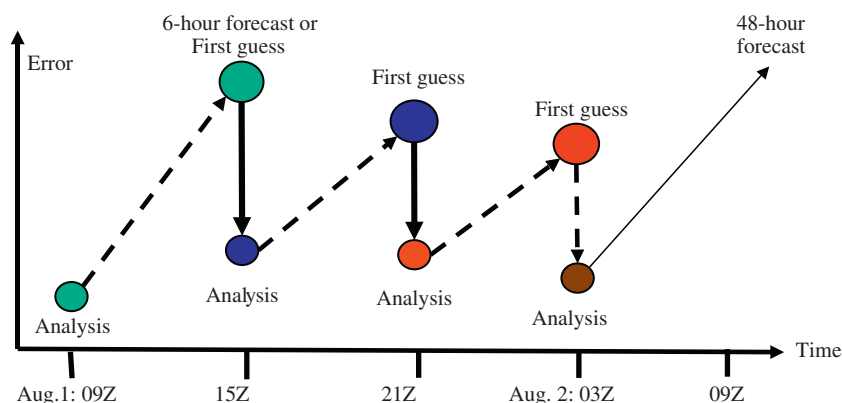

Fig. 5. A schematic diagram illustrating the incremental 3DVAR data assimilation cycle with a window of 6 h.
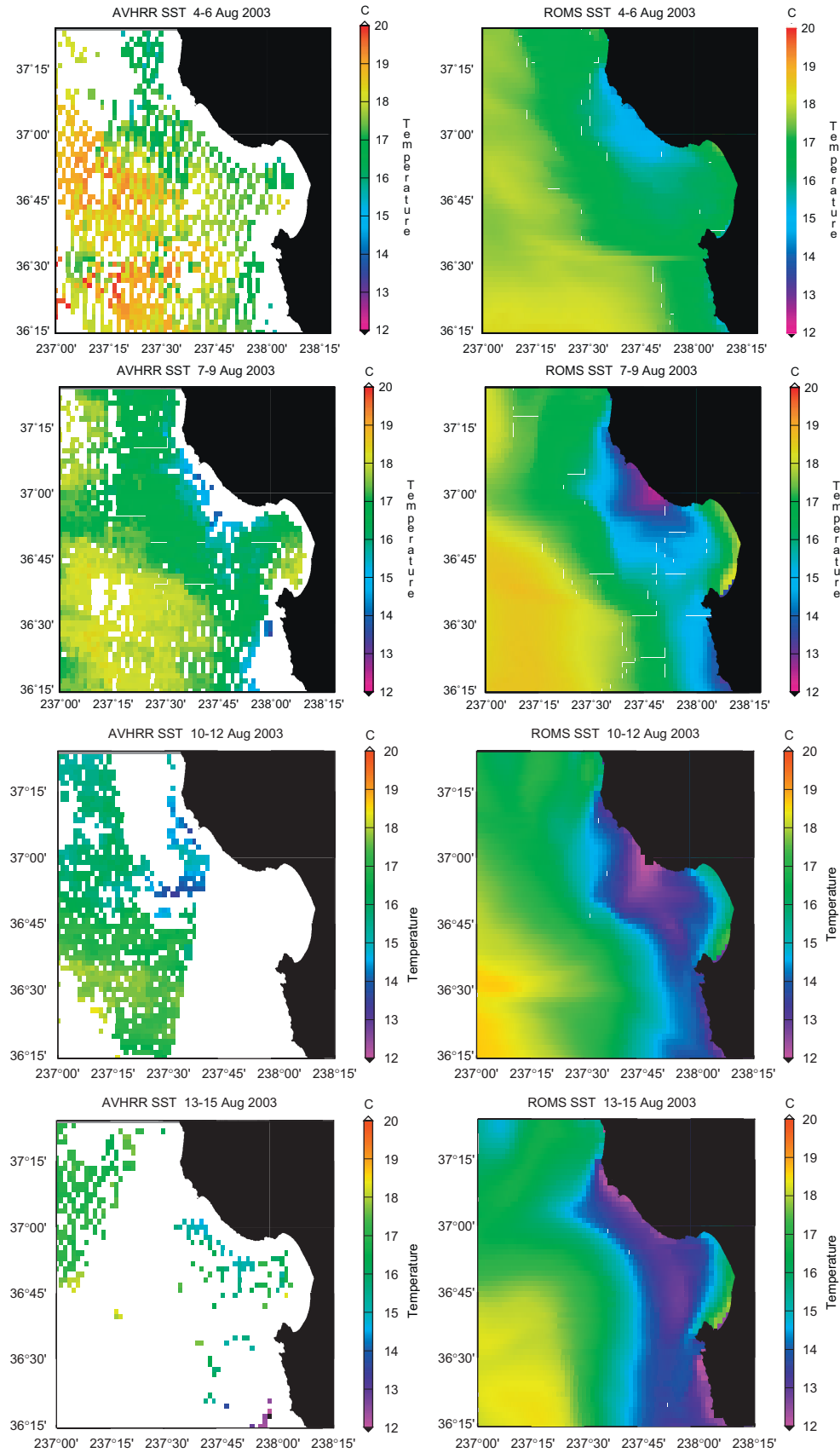
Fig. 6. Horizontal maps of sea-surface temperature as derived from AVHRR observation (left) and the ROMS reanalysis (right) for 4–6 August and 7–9 August, 2003. Note that AVHRR measurements are assimilated in the ROMS reanalysis. Horizontal maps of sea-surface temperature as derived from AVHRR observation (left) and the ROMS reanalysis (right) for 10–12 August and 13–15 August, 2003.

significantly reduced or even reversed, the upwelling is shut down, and the MB is characterized by somewhat more horizontally uniform temperature and salinity distributions. During August 2003, there was a well-defined upwelling state during August 6–19 followed by a brief relaxation state (August 20–24) and another upwelling state toward the end of August (Fig. 2).

Because of the limited wind observations by *in situ* sensors (e.g., MB moorings and NOAA NDBC buoys) in the coastal oceans and the relatively coarse spatial resolution (on the order of 25 km) for satellite (e.g., QuikSCAT) measured winds, coastal ocean models will continue to rely on mesoscale atmospheric models to provide the needed atmospheric forcing. The excellent agreement between the mooring measured winds and the model simulated winds at the finest resolution grid of 3 km (Fig. 3) provide us confidence to use the COAMPS atmospheric forcing for real-time nowcast/forecast and reanalysis.

The model surface boundary conditions use the wind stress computed from the COAMPS 10-m wind using the Large and Pond (1981) drag coefficient formulation that depends upon wind speed. For the heat flux, we directly use the short-wave and long-wave radiative fluxes produced by COAMPS and compute the latent and sensible heat fluxes using the Kondo (1975) bulk formula with prescribed 2-m air temperature and humidity and 10-m wind also provided by COAMPS. In addition, the precipitation produced by COAMPS is used for the fresh-water flux into ROMS. Though typically used by many global and basin-scale climate models, in this study no salinity restoring term toward the climatological surface salinity is used.

4. A real-time ocean modeling, data assimilation and forecast system

4.1. Ocean circulation modeling

The ocean circulation and variability are simulated by the Regional Ocean Modeling System (ROMS). ROMS solves the primitive equations in an Earth-centered rotated Cartesian system of coordinates. The Boussinesq approximation (i.e. where density variations are neglected everywhere except in the gravitational force) is used. ROMS is discretized in a terrain-following curvilinear coordinate. ROMS is a split-explicit, free-surface ocean model, where short time steps are used to advance the surface elevation and barotropic momentum, with a much larger time step used for temperature, salinity, and baroclinic momentum. Detailed descriptions of ROMS and its implementation can be found in Shchepetkin and McWilliams (2005).

The boundary conditions for the model are appropriate for an irregular solid bottom and coastline, free upper surface, and open-ocean sides away from the coastline. These include the forcing influences of surface wind stress and heat and water fluxes, coastal river inflow, bottom drag, and open-ocean outgoing wave radiation and nudging towards the specified basin-scale circulation. In order to obtain well-behaved, long-term solutions for coastal configurations with open boundaries, ROMS uses a radiation boundary scheme that estimates a two-component, horizontal phase velocity near the open boundaries (Raymond and Kuo, 1984; Barnier et al., 1998). This method escapes the usual difficulties associated with large phase speeds and small-scale noise. Thus, it allows large time steps without loss of stability. No enhanced dissipation is needed near the boundary (i.e. "sponge layers") to maintain numerical stability. Large-scale influences are implemented with nudging bands along the open boundaries, where all fields are relaxed toward specified data (which come from observations or a larger domain, coarse resolution model). The nudging rate is adjusted dynamically depending on inward/

outward propagation of information as determined by the radiation algorithm. This method prevents long term drifts (a typical consequence of unforced radiation boundary conditions, which are numerically stable, but fail to supply correct large-scale information) and an over-specification problem (typically manifested by spurious currents along open boundaries in the case of excessive nudging on outflow side, Miyakoda and Rosati, 1977). The ROMS boundary condition has been documented in Marchesiello et al. (2001). For the vertical mixing, ROMS uses a non-local, K-profile PBL scheme (Large et al., 1994) that performs well in both data comparisons and large-domain model solutions (Li et al., 2001).

ROMS is written in FORTRAN as a parallel code, which uses 2D partitioning of the model grid into subdomains that may be assigned to different processors. ROMS can be run on both shared-memory (OpenMP) and distributed memory (MPI) machines. Validation of ROMS parallel performances show that both methods result in viable overall scaling (Wang et al., 2005; Shchepetkin and McWilliams, 2005).
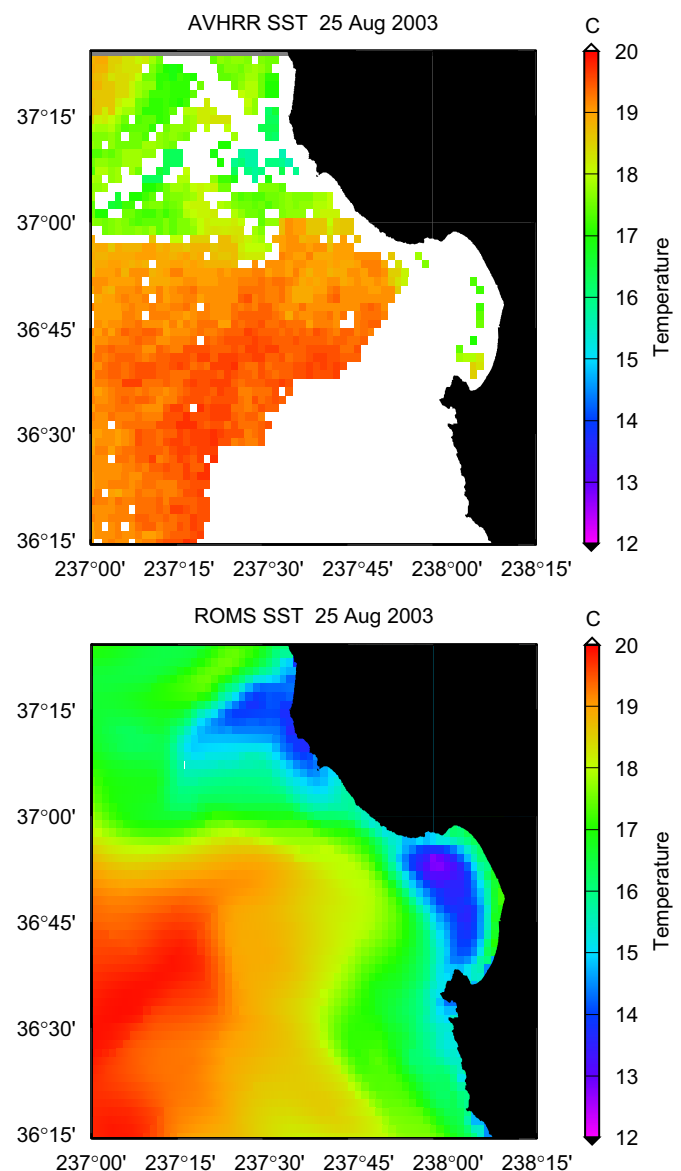

Fig. 7. Horizontal maps of sea-surface temperature as derived from AVHRR observation (top) and the ROMS reanalysis (bottom) for 25 August, 2003. Note that AVHRR measurements are assimilated in the ROMS reanalysis.

4.2. Ocean nested modeling

In addition to the local wind forcing and air–sea heat exchange, the side boundary conditions play a very important role in influencing the fine-scale variability. Presumably, these boundary conditions come from the larger-scale California Current System.

This paper uses a hierarchy of ROMS systems with different model domains and different spatial resolutions. The boundary conditions from the smaller ROMS domain are obtained from the larger ROMS domain. For the AOSN experiment, a three-tier nested ROMS configuration is used including the US West coastal ocean at 15-km resolution, the central California coastal ocean at 5-km resolution, and the MB at 1.5-km resolution (Fig. 4). All three ROMS configurations have 32 vertical sigma layers. The 15-, 5-, and 1.5-km regional ROMS are nested on-line as a single system and run simultaneously exchanging boundary conditions at every time step of the parent grid.

The advantage of using the same ROMS code for both regional and coastal modeling will enable us to understand and to build a predictive capability for the physical circulation and ecosystem response in the MB region. High (~1-km) spatial resolution is required in the MB in order to correctly represent dynamical features and processes such as upwelling plumes, current-topography interactions, and fronts. On a larger scale, mesoscale eddies and filaments associated with the California Current system are important for setting the background environment for the MB circulation and variability. In particular, getting the strength and location of the poleward undercurrent right requires a regional scale model and, in turn, baroclinic instability associated with the vertically sheared meridional boundary currents leads directly to jet meandering and eddy formation. Our approach is to integrate a hierarchy of local, regional and basin-scale models to accurately describe these processes and the interactions between scales.
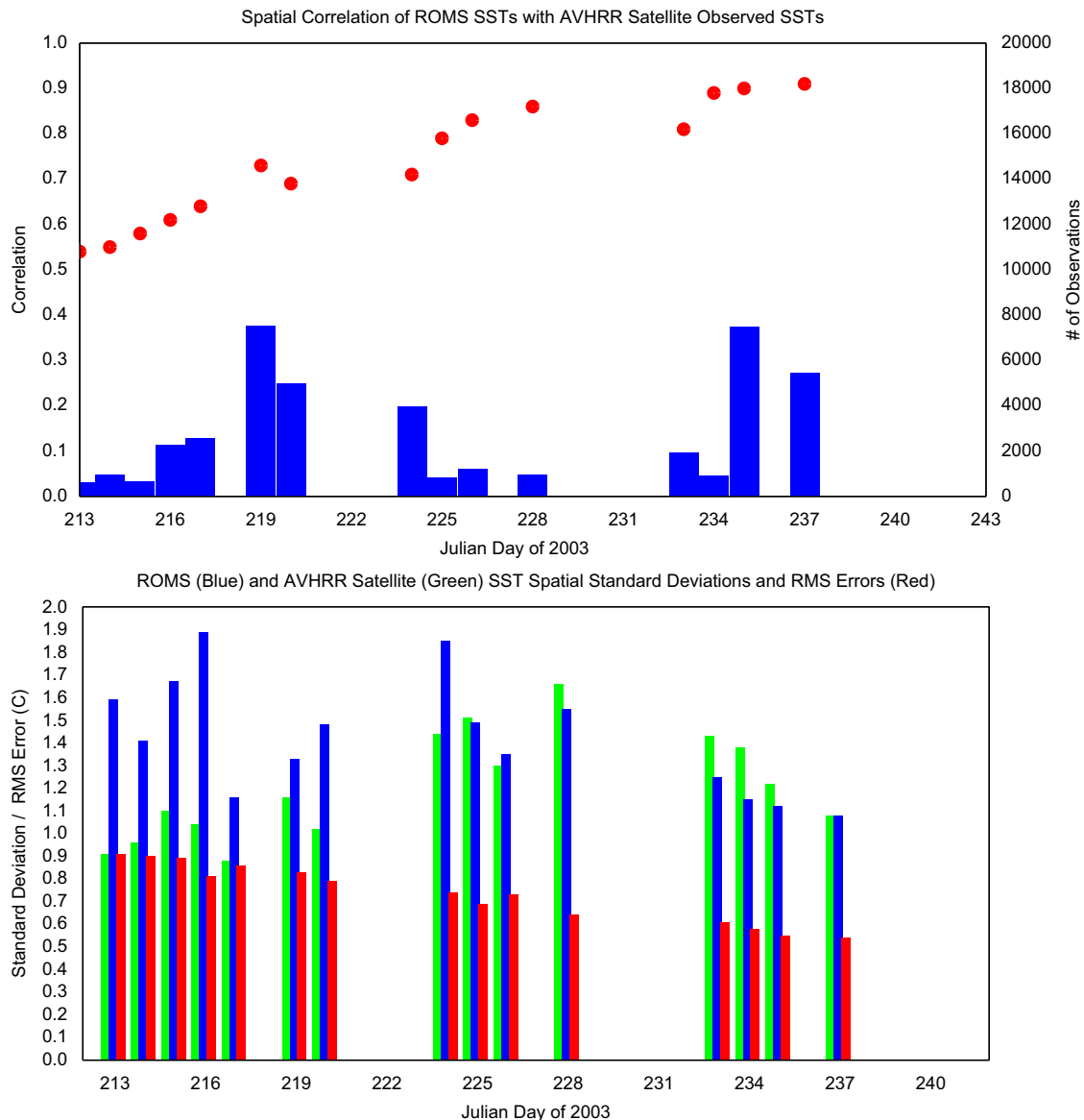

Fig. 8. Top: Time series of the spatial correlation (red dots) between the satellite (AVHRR) measured and ROMS analyzed SST. The number of AVHRR measurements is also shown in blue bars. Bottom: Time series of the standard deviation of the spatial SST patterns as derived from the satellite-measured (green) and the co-located ROMS analyzed (blue) SST and the root-mean-square (RMS) difference (red). Note that AVHRR measurements are assimilated in the ROMS reanalysis.

ROMS is discretized on a structured grid, so local refinement can be performed via nested grids (i.e. fixed high-resolution local models embedded in larger coarse-grid models). The interactions between the two components are twofold: the lateral boundary conditions for the fine grid are supplied by the coarse-grid solution, while the latter is updated from the fine grid solution in the area covered by both grids (Blayo and Debreu, 1999). In our implementation, only one-way nesting is used: coarser grid solutions are used to influence the finer-grid boundary conditions. Long-term simulations have been made to obtain the equilibrium solution (Marchesiello et al., 2003; Penven et al., 2006). The embedded solution shows no discontinuities at the nested domain boundary and a valid representation of the upwelling structure, at a CPU cost only slightly greater than for the inner region alone.

4.3. Ocean data assimilation system

A hierarchy of data assimilation methods has been used in meteorology and oceanography with different levels of sophistication ranging from the relatively simple uni- and multi-variate statistical (optimal) interpolations, variational data assimilation methods implemented for both three-dimensional (3DVAR) and four-dimensional (4DVAR) problems, Kalman Filter and smoother (see Daley, 1991; Li and Navon, 2001 for more detailed reviews). More sophisticated algorithms usually generate better analyses and forecasts, although with a higher computational demand. There must be a compromise between the sophistication of the data assimilation algorithm, computational limitations and time constraints. For a retrospective analysis, one can afford the more sophisticated methods as long as one can fit the problem into the computer available. For the real-time nowcast and forecast problem described here, the computing time for the end-to-end system has to be less than a day. This time requirement eliminates the possibility of using 4DVAR and the Kalman Filter/smoother. Thus, 3DVAR is used in this study.

3DVAR was introduced at the major meteorological centers in the late 1980s and early 1990s and is still widely used operationally in many meteorological centers such as the National Center of Environmental Prediction (NCEP) and the Navy Fleet Numerical Meteorology and Oceanography Center (FNMOC). Because of the extensive experience with it, 3DVAR has been recognized as a reliable data assimilation algorithm. Computationally, 3DVAR is efficient enough to use in a real-time operational system using a small-to-medium size computer. Further, 3DVAR is flexible enough to incorporate various standard and non-standard measurements simultaneously, and thus allows us to assimilate observations from both *in situ* and remotely sensed (aircraft and satellite) data as well as fixed and mobile sensors.

To implement 3DVAR for the coastal ocean, we not only adopted a number of strategies developed by the meteorological data assimilation community, but also proposed particular strategies that are unique for oceanographic applications. These strategies include the implementation of 3D anisotropic and inhomogeneous 3D error correlations based on a Kronecker product, application of particular weak dynamic constraints, and implementation of efficient and reliable algorithms for minimizing the cost function. The algorithm has been detailed in two companion papers (Li et al., 2008a, b). In order to be self-contained, this section provides a brief summary of the 3DVAR data assimilation method implemented in ROMS.

The goal of 3DVAR is to generate the best estimate of the ocean state (x^a) by using the model forecast (x^f) and all the available observations (y). The 3DVAR method is to minimize the defined cost function with respect to the model state (x):

$$J = 0.5(x - x^f)^T B^{-1}(x - x^f) + 0.5(hx - y)^T R^{-1}(hx - y) \quad (1)$$

where h is an observation operator that maps model state variables to the observed quantity, i.e. $y = hx + v$, where v is the observational error. When the model state variables are directly observed, h may simply be a linear interpolation operator from the model grid to the observational point. B is the model forecast error covariance matrix defined as

$$B = \langle (x^f - x^t)(x^f - x^t)^T \rangle \quad (2)$$

where x^t stands for the true ocean state under investigation, the upper script $(\ )^T$ for matrix transpose, and $\langle \cdot \rangle$ for mathematical expectation. R is the error covariance matrix of y defined as

$$R = \langle (y^o - y^t)(y^o - y^t)^T \rangle \quad (3)$$

where y^o corresponds to the observational values and y^t stands for the true value of the observed quantity. In this study, we will only use the diagonal elements of R, which corresponds to the error
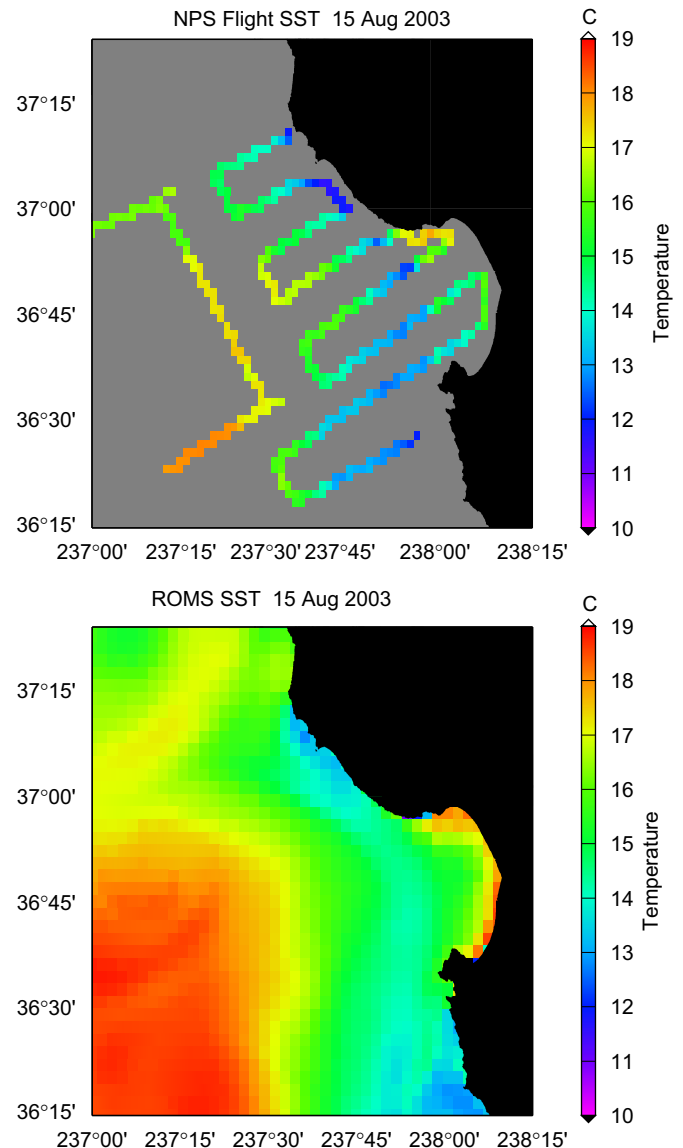

Fig. 9. The composite image of SST on August 15, 2003 as derived from aircraft observations (upper panel) and the ROMS reanalysis (lower panel). Note that aircraft SST measurements are assimilated in the ROMS reanalysis.

variance provided by the observational data sets. The method to estimate the model forecast error covariances will be briefly addressed later.

In this study, we first convert the total fields to the incremental fields, i.e. $\delta x = x - x^f$, and $\delta y = y - hx^f$. δy is also known as an innovation vector. The motivation to work with the incremental fields rather than the total fields is to retain the original forecast fields (x^f) and apply all the operations to the incremental fields. With these two transformations, the cost function of Eq. (1) can be rewritten

$$J = 0.5\delta x^T B^{-1}\delta x + 0.5(h\delta x - \delta y)R^{-1}(h\delta x - \delta y) \tag{4}$$

If δx^a corresponds to the minimum of the incremental cost function (4), the final maximum likelihood estimate would be

$$x^a = x^f + \delta x^a \tag{5}$$

The minimum of a cost function corresponds to the maximum likelihood (Bayesian) estimate, and also to the minimum variance estimate (e.g., Jazwinski, 1970; Cohn et al., 1998), assuming that the observation operator (h) is linear and both the forecast and observation errors follow white Gaussian distributions.

From the cost function (1), the final analysis hinges on the specification of B and R. B represents our trust in the model forecast, while R represents our trust in the observations. A smaller B implies smaller forecast error, and the analysis is more dependent on the model forecast. A systematic method for modeling the model forecast error covariances has been detailed in Li et al. (2008b). One may think that we should always put more trust in the observations. However, the observational errors also can be very complicated. For example, in the coastal ocean small eddies on the scale of a couple of kilometers often exist, and there are very HF oscillations of time scales of minutes to hours. If such small eddies and HF oscillations are not resolved by the
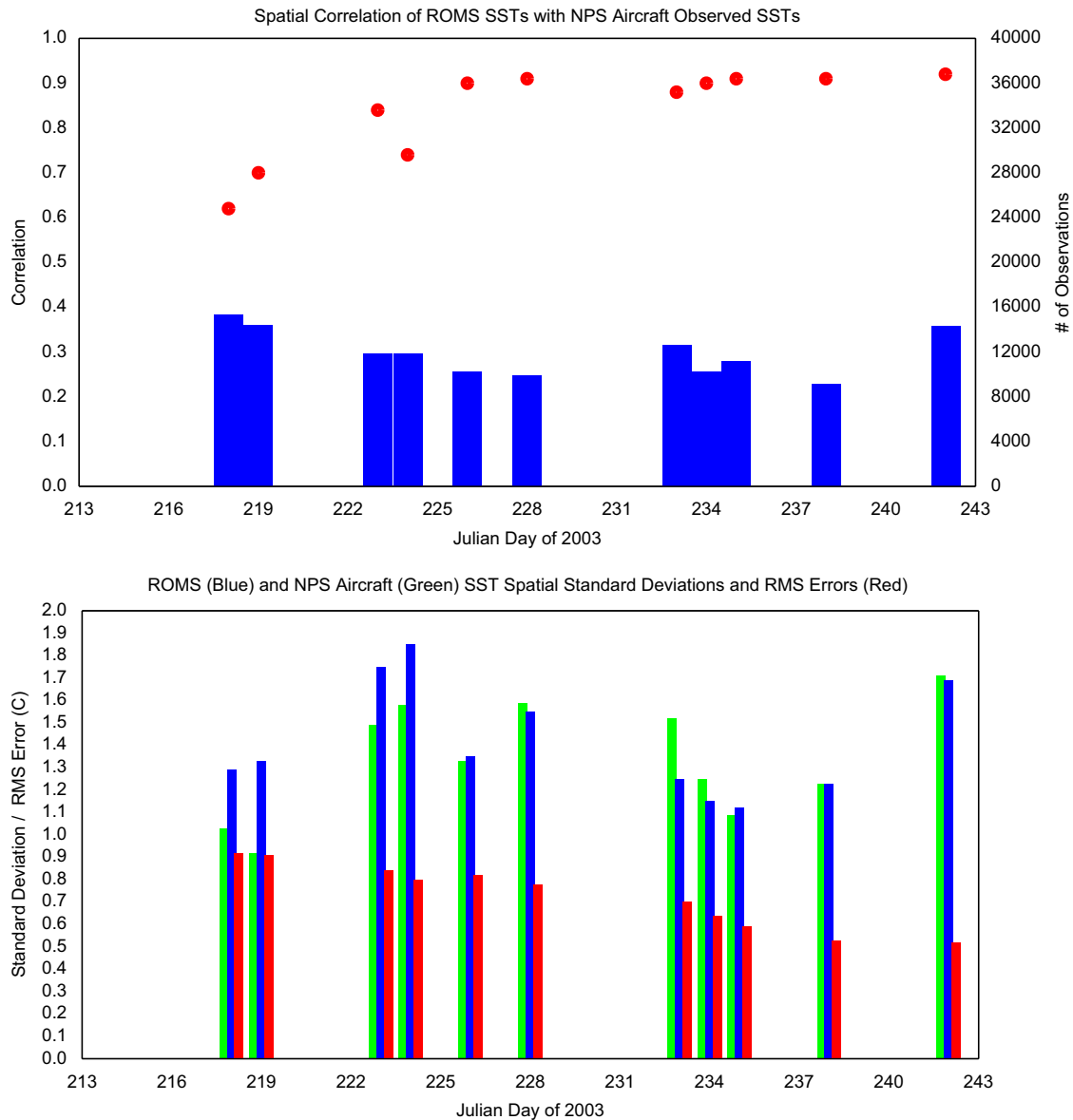

Fig. 10. Top: Time series of the spatial correlation (red dots) between the aircraft (NPS) measured and ROMS analyzed SST. The number of aircraft (NPS) measurements is also shown in blue bars. Bottom: Time series of the standard deviation of the spatial SST patterns as derived from the aircraft (NPS) measured (green) and the co-located ROMS analyzed (blue) SST and the root-mean-square (RMS) difference (red). Note that aircraft SST measurements are assimilated in the ROMS reanalysis.

model, then their effect on the observations must be accounted for in the observational errors. This type of observational error is known as the representation error.

The model forecast error covariance plays another important role in 3DVAR. After the error covariance matrix is normalized by the root mean square (RMS), it becomes a correlation matrix. The correlation allows 3DVAR to spread out observational information horizontally and vertically from the observed locations to the surrounding area. For example, the Jason-1 SSH has a high spatial resolution of about 6 km along the satellite track, and about 220 km across the two adjacent satellite tracks at middle latitudes. The area between two tracks is SSH data void. However, SSHs in this void area may be correlated with those along the tracks. The correlation determines the spreading out from the high-resolution along-track information into the areas between the tracks. During AOSN, the major observations are from moving gliders. The glider profiles can be considered sporadic and heterogeneous in space. The spreading of observational information is crucial for successfully assimilating glider observations.

There is a great challenge in constructing the correlation matrix, however, due to its large size. ROMS with a grid size of 1.5 km has about 6×10^5 grid points. Considering four 3D control variables, the dimension of the correlation matrix is as large as 6×10^{12}. This size is prohibitively large for directly using the full size correlation matrix in terms of both the computing and memory requirements. A strategy is necessary to simplify B, so

that it becomes manageable in size while retaining the major features of the original matrix. Using the Kronecker product, we proposed a formulation to reconstruct B approximately (Li et al., 2008a, b). The reconstructed B requires a manageable amount of memory and renders our 3DVAR computationally efficient. Nevertheless, the reconstructed B is able to incorporate most aspects of inhomogeneity and anisotropy for coastal oceans.

Once the cost function is formulated, we need to find a minimization algorithm. Among the many minimization algorithms available, we select the limited-memory quasi-Newton method (Liu and Nocedal, 1989), mostly because of its extensive use in solving nonlinear problems and its computational efficiency and reliability (Li and Navon, 2001).

4.4. Model initialization, data assimilation and forecast cycle

Starting from the climatological conditions of temperature and salinity with zero current, the outermost ROMS domain with a resolution of 15 km was integrated for multiple years forced with climatological air–sea fluxes in order to reach quasi-equilibrium. This 15-km ROMS solution is used to initialize the 5- and 1.5-km ROMS. The resulting three-tier nested ROMS was then integrated together for multiple years. The one-way coupling in this study used the coarser resolution models to provide the boundary conditions for the finer-resolution models. The exchange of
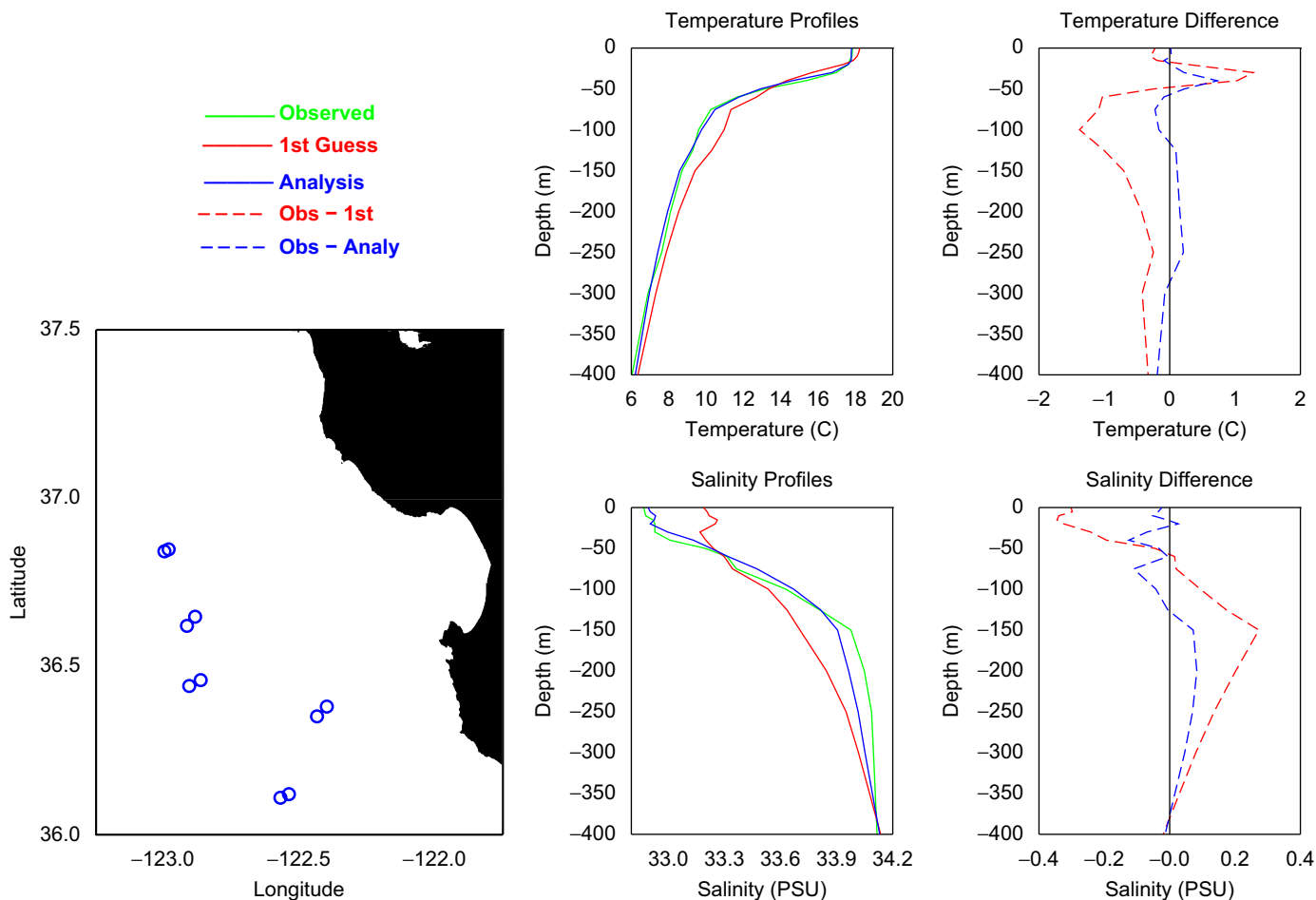

Fig. 11. Comparison of co-located ROMS first guess and analysis fields with glider temperature and salinities averaged over 1 day (15 August 2003). The blue circles represent the locations occupied by the five SIO gliders during this particular day. The observation is plotted in green; the ROMS first guess, which is the 6-h forecast from the previous analysis/nowcast, is plotted in red; the ROMS analysis/nowcast is plotted in blue. The differences between the observations and the ROMS first guess and analysis/nowcast are plotted in red and blue dashed lines, respectively. Note that glider measurements are assimilated in the ROMS reanalysis.

boundary conditions is conducted online every time step of the coarser resolution model. Using the climatological spinup solutions, we start the real-time simulation using air–sea fluxes derived from COAMPS on July 21, 2003, about a week before the AOSN experiment. Although satellite data are assimilated into ROMS from the start of the real-time simulation, it is not until the second week of August when there are significant amounts of *in situ* measurements available for data assimilation (Fig. 1).

During the real-time AOSN experiment, we used a 24-h assimilation window mostly because of the computational constraint. For the reanalysis presented here, we have reduced the assimilation window from 24 to 6 h (Fig. 5). Every 6 h, we produce a nowcast (or analysis). The nowcast (x^a as shown in Eq. (5)) is obtained as the sum of the model first guess (x^f in Eq. (5)) and the assimilation correction term (or increment, δx^a in Eq. (5)). The model first guess is basically a 6-h forecast initialized with the nowcast from 6 h previous. The assimilation correction is obtained by minimizing the cost function taking into account all the data available during a 6-h window (i.e. 3 h before and after the nowcast time). From any nowcast field, a 48-h

forecast can be issued forced by the COAMPS atmospheric forecast.

5. Evaluation of ROMS reanalysis with assimilated data

Since the end of real-time operations at the end of August 2003, a number of modifications have been made in order to produce the most realistic ROMS reanalysis fields that can be used for further diagnostic analysis. During the real-time operations, the assimilation window was set to 24 h, simply because of the computational time required to accomplish the data processing, modeling and data assimilation in real time. Significant efforts have been put into the quality control of various observational data sets. As a result, the ROMS reanalysis presented here is significantly better than the real-time performance. The ROMS reanalysis will be first compared with the assimilated data sets for consistency. In the next section, the ROMS reanalysis is evaluated against the independent measurements that are not assimilated.
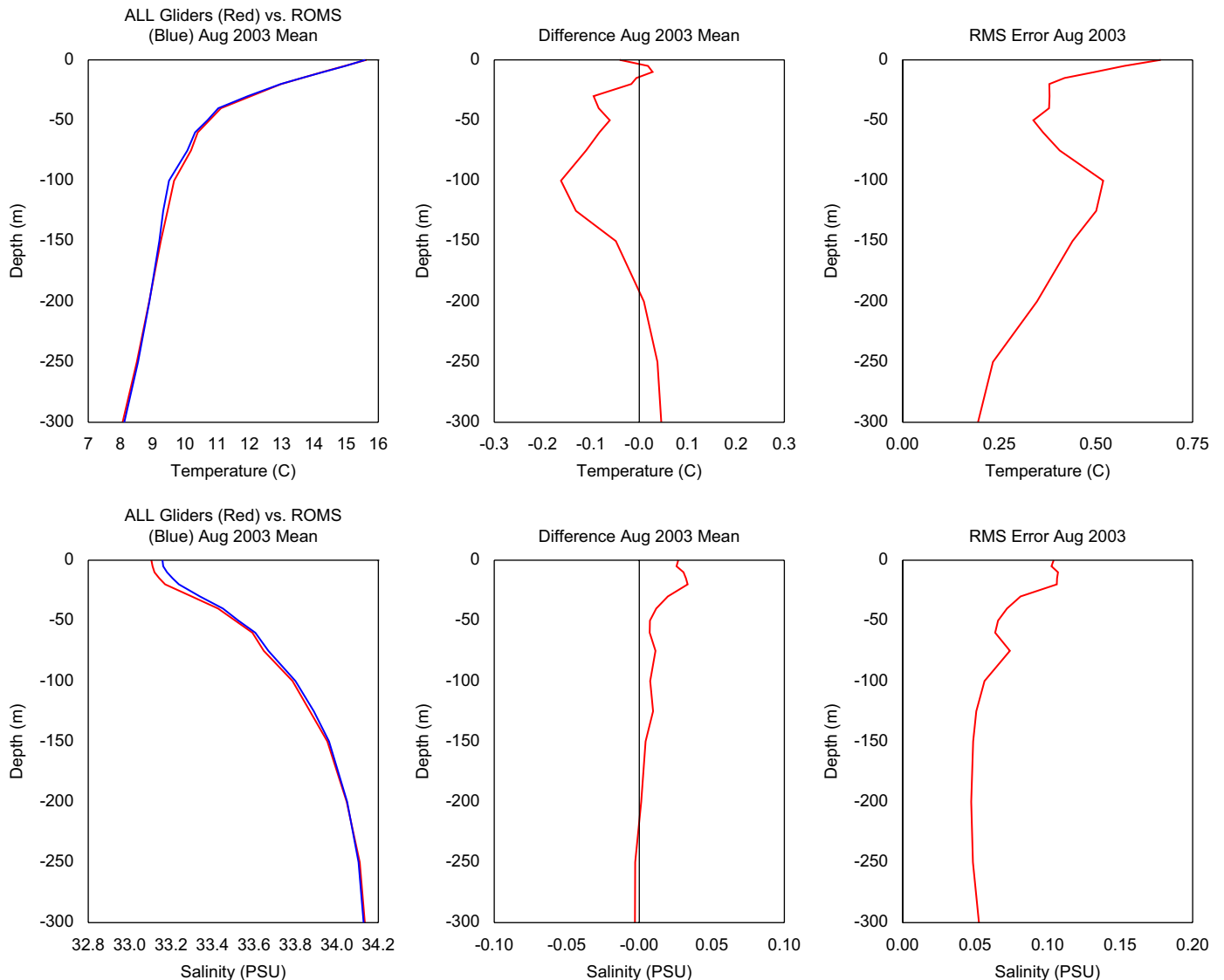

Fig. 12. Vertical profiles of co-located glider and ROMS analyzed temperature (upper panel) and salinity (lower panel) for the mean of the entire month of August 2003 (left), mean differences (ROMS-glider, middle) and RMS differences (right). Note that glider measurements are assimilated in the ROMS reanalysis.

5.1. Sea-surface temperature from satellite

The SST data measured by both satellites (AVHRR) and aircraft are assimilated in ROMS, so they are not considered as independent observations and are only used here for consistency tests. For the AVHRR SST data, relatively cloud-free images (as defined by the majority of the MB area) are available during the development phase of the upwelling state (August 3–18 and August 23–27) as well as during the brief relaxation state (August 20–21).

The first cloud-free AVHRR image was available on August 3 (not shown) when SST is relatively uniform around the MB in the range of 16–17 °C. During August 4–6 (Fig. 6), two upwelling centers with the coldest SST reaching 13 °C are seen in the vicinity of Pt Año Nuevo and Pt Sur, respectively. The cold water around Pt Sur is colder than that at Pt Año Nuevo and occupies a larger area. This description is consistent with previous findings (e.g., Rosenfeld et al., 1994) suggesting this early August 2003 upwelling event is quite typical, representing a repeated upwelling pattern in the spring–summer season.
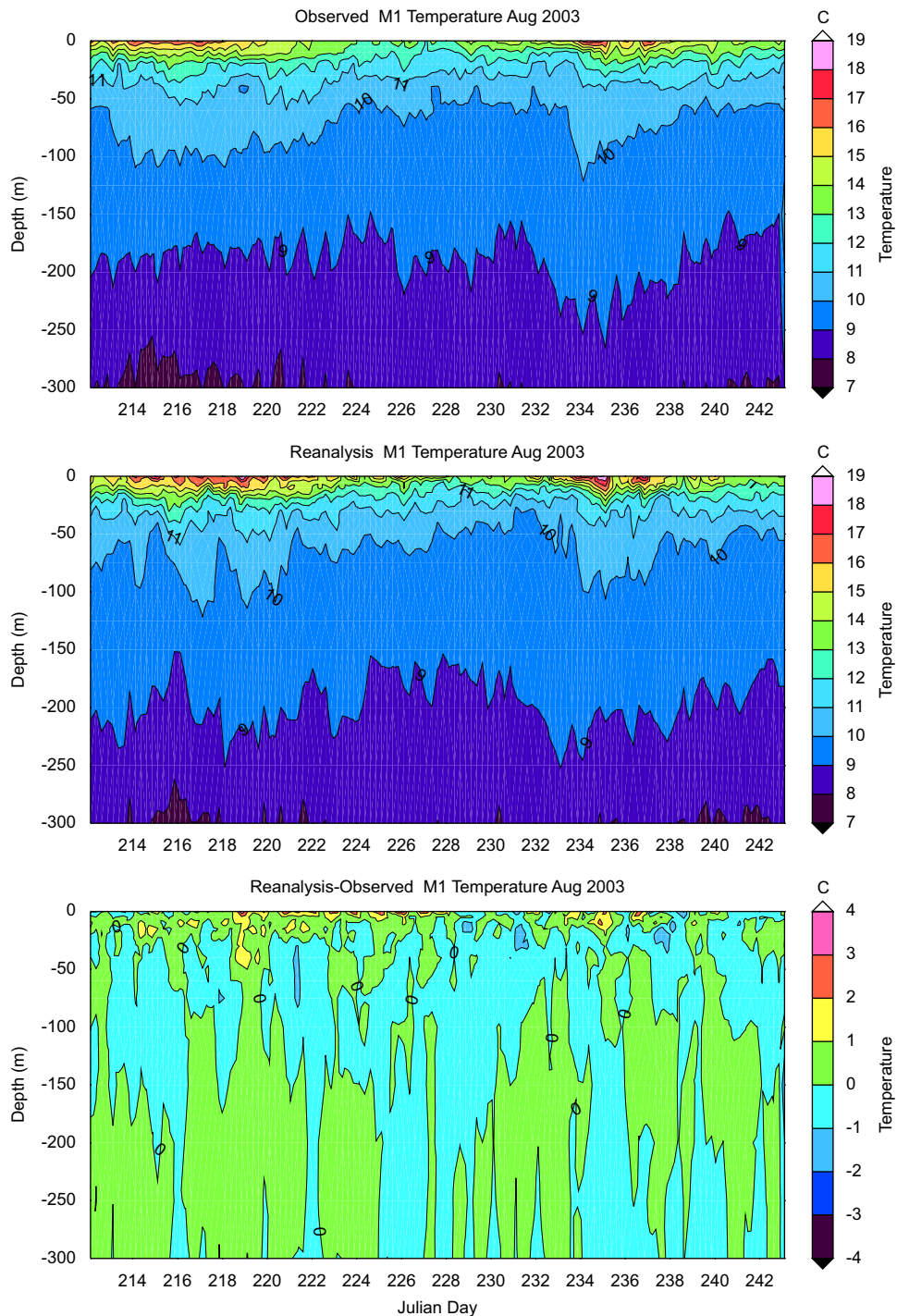

Fig. 13. Time series of temperature as a function of depth as derived from mooring observations at the M1 location (top), the ROMS reanalysis (middle) and the difference between them (bottom). Note that mooring temperature and salinity measurements are assimilated in the ROMS reanalysis.

In the following days (August 7–9 and August 10–12 in Fig. 6), a bifurcated pattern as described in Rosenfeld et al. (1994) is seen in both upwelling centers. Originating at the upwelling centers at Pt Año Nuevo and Pt Sur, cold waters start to extend offshore as well as equatorward. By August 13–15 (Fig. 6), a significant portion of the cold waters originating from Pt Año Nuevo have spread south into MB connecting to cold waters originating from Pt Sur. In the meantime, a warm-water center has developed in the shelf area of the MB with SSTs warmer than 18 °C. It should be pointed out that the coldest water seen in MB during the

upwelling state is crossing the Monterey Canyon (the head of the canyon is approximately co-located with the M1 mooring, see Fig. 2 for M1 mooring location). This suggests that the cold waters seen around the Monterey Canyon are transported into the Bay from outside, rather than directly from the deep Monterey Canyon. For this particular upwelling event in early August 2003, these cold waters originated from the north around or north of Pt Año Nuevo.

From Fig. 6, we conclude that the SST sequence associated with the development of this upwelling state is realistically reproduced
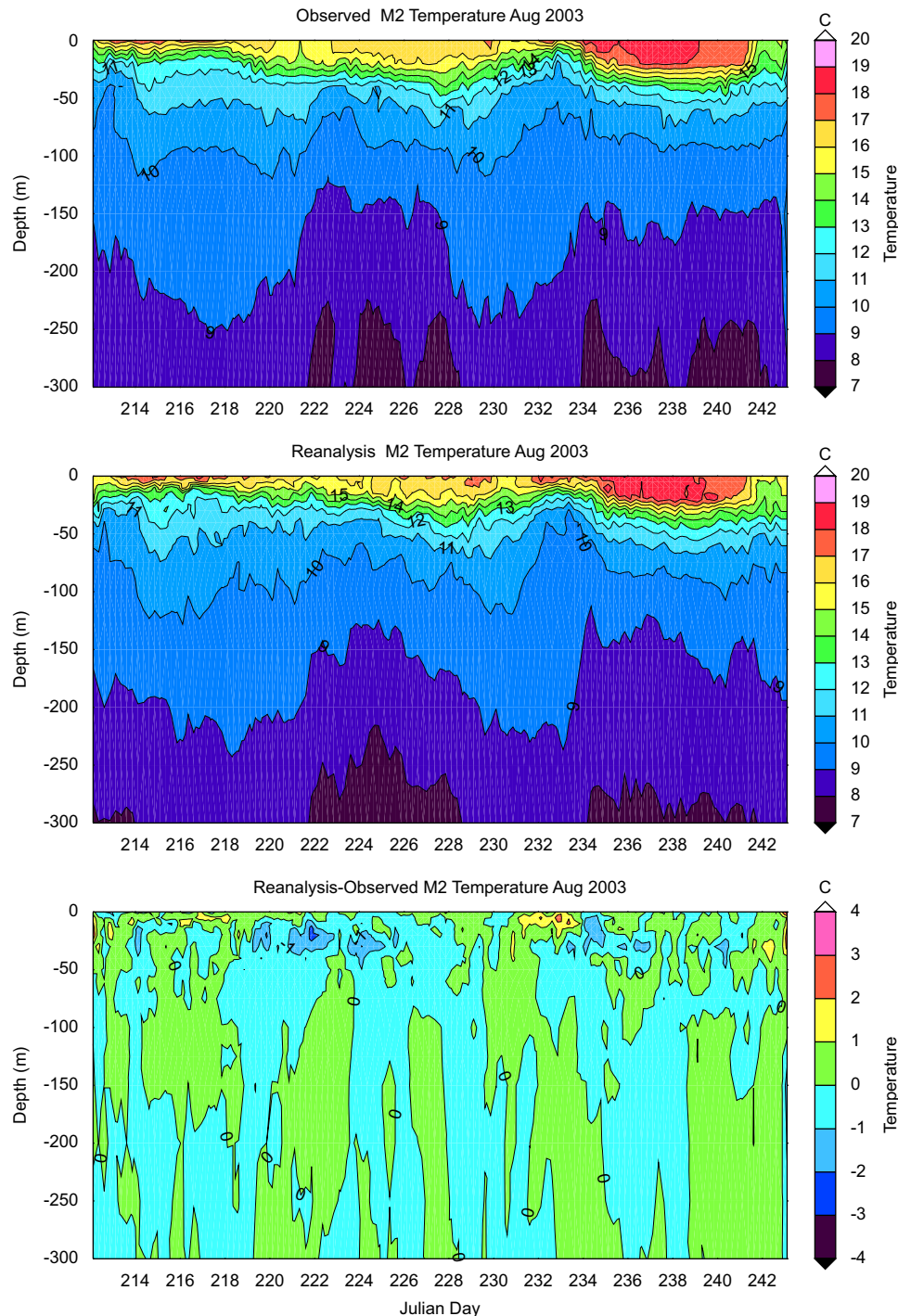

Fig. 14. The same as Fig. 13 except at the M2 location.

by the ROMS reanalysis. This is not surprising because the AVHRR SST data shown in Fig. 6 are assimilated into our ocean model to generate the reanalysis. However, it is always reassuring to see that the ROMS data assimilation system has the ability to digest the AVHRR data and yield a consistent analysis.

Another example of an offshore extension of cold waters is seen during the second upwelling state around August 25 (Fig. 7). The cold waters originally associated with the upwelling center at Pt Año Nuevo have extended offshore. At this time, the ROMS analysis fields also show the extension offshore of this colder water.

A quantitative analysis between the satellite-measured and ROMS reanalyzed SST is shown in Fig. 8. The spatial correlation (all correlations cited are standard Pearson's correlation coefficient) between the satellite and ROMS SST shows a gradual increase suggesting that the positive impact as more data are assimilated into ROMS. The ROMS analyzed standard deviation of the spatial SST patterns also shows better agreement with the observations toward the later part of August. It should be pointed out that there are many days when there are few AVHRR measurements because of cloudy conditions in this area.
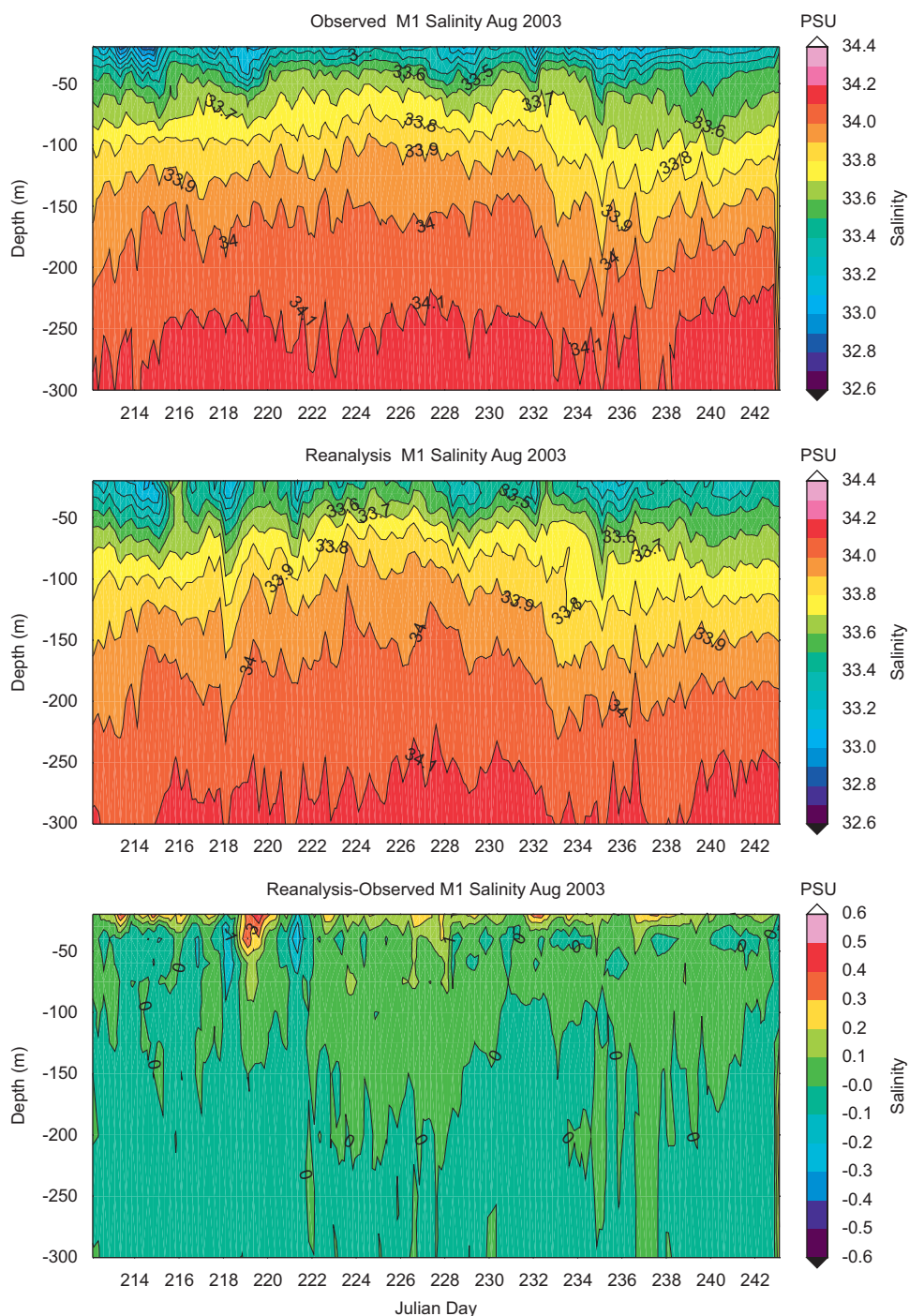

Fig. 15. The same as Fig. 13 except for salinity.

5.2. Sea-surface temperature from aircraft

Because of the lack of SST measurements from satellites due to clouds, a total of 12 aircraft flights were made during the August 2003 field experiment. Each flight took approximately 4–6 h to map the MB. Fig. 9 shows a typical image of the aircraft measured SST on August 15 during the peak of the upwelling. Consistent with the AVHRR observations (see Fig. 6) and aircraft measurements, ROMS reanalysis also shows two cold water or upwelling centers at Pt Año Nuevo and Pt Sur, respectively, while warm waters are seen both offshore and on the MB shelf area. Because the aircraft can fly under the cloud deck (approximately 1000 ft),

this type of SST measurement can be made on a regular basis and therefore has a large impact on model performance during cloudy days when AVHRR data are unavailable.

The quantitative comparison shows similar conclusions as the satellite SST comparison: it takes at least a week before a high correlation is achieved between the aircraft measured and ROMS analyzed SST patterns (Fig. 10). This is consistent with the fact that there is a steady increase of the various *in situ* data during the first week of the field experiment. The standard deviations of SST spatial patterns (Fig. 10) also show better agreement during the latter half of the experiment. It should be pointed out that the number of measurements from the aircraft is significantly larger
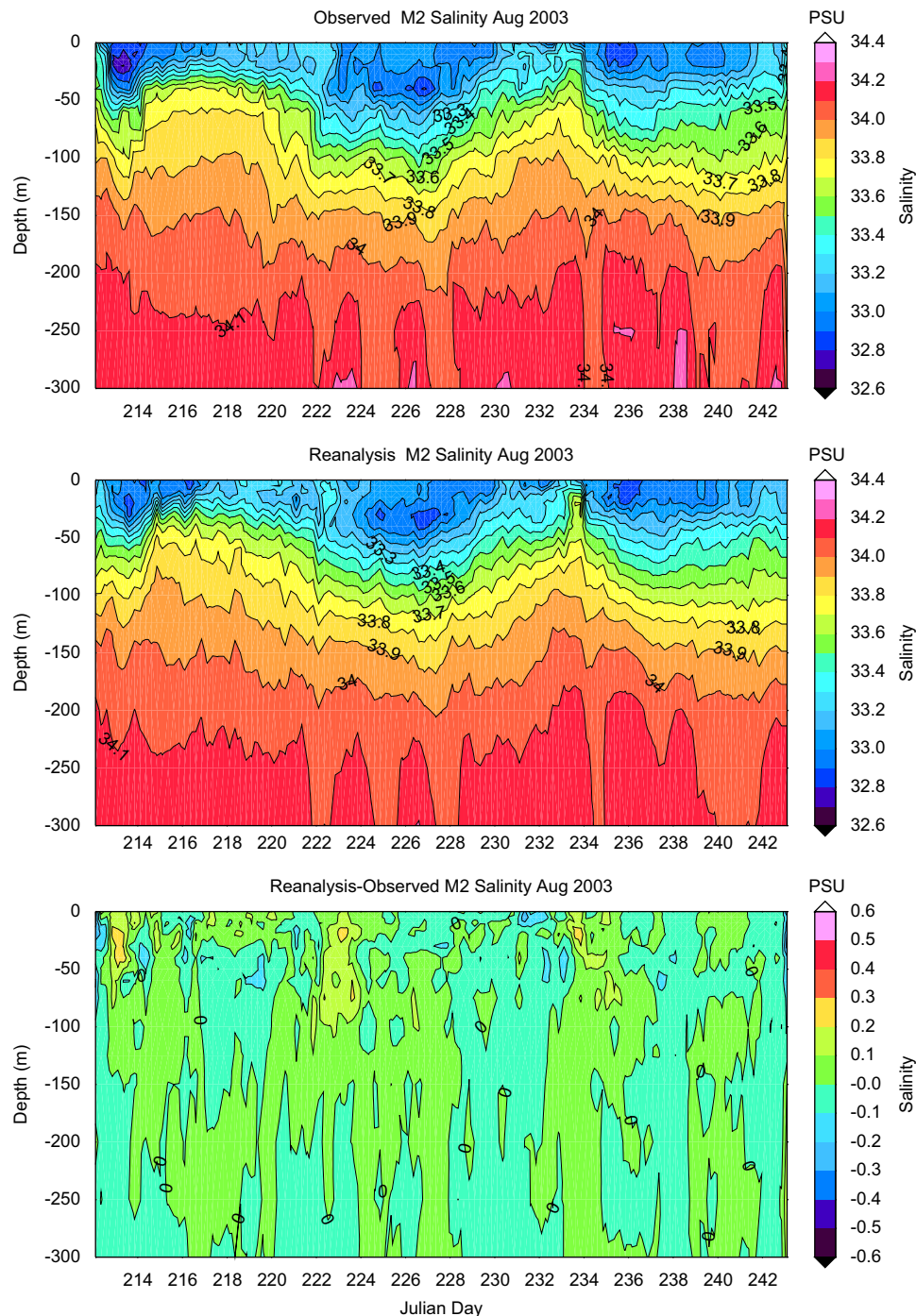

Fig. 16. The same as Fig. 14 except for salinity.

than the satellite measurements, although they have quite different spatial distributions.

5.3. Glider measurements of vertical temperature and salinity profiles

In this subsection, we compare the SIO and WHOI glider temperature and salinity profile data with our ROMS analysis fields. Since these data are assimilated in ROMS, they are not independent observations and, as with the previous AVHRR and aircraft SST data, are used to confirm the consistency of our results. Fig. 11 shows as an example the comparison for a single day (15 August 2003) of the glider profiles and the ROMS first guess (a 6 h forecast) and analysis (after data assimilation) fields. For the temperature, the glider observed (green), first guess (red) and analysis (blue) profiles are virtually indistinguishable. The difference profiles reveal a ROMS cold bias between 40 and 100 m and a warm bias below 120 m in the first guess fields that are reduced, but not eliminated in the analysis fields. These biases are

typical of those found throughout the experiment, both in magnitude and sign. For the salinity, the pattern is a bit more complex, especially near the surface, where variations in the vertical are large; however, between about 50 and 250 m, there is a tendency for the ROMS first guess to be too fresh that is reduced in the analysis fields.

The monthly summary and statistics are shown in Fig. 12 with monthly mean profiles, monthly mean differences (ROMS minus Glider) and RMS differences for all the glider profiles and co-located ROMS analysis profiles. The RMS is calculated as

$$RMS = \left[1/n \sum (x_i - y_i)^2 \right]^{1/2} \tag{6}$$

For the temperature, mean differences near the surface are small, but there is a distinct mean cold bias between 20 and 150 m, though it must be noted that it is quite small (less than 0.2 °C). RMS differences in temperature, on the other hand, are largest near the surface, where they approach 0.75 °C. At all other
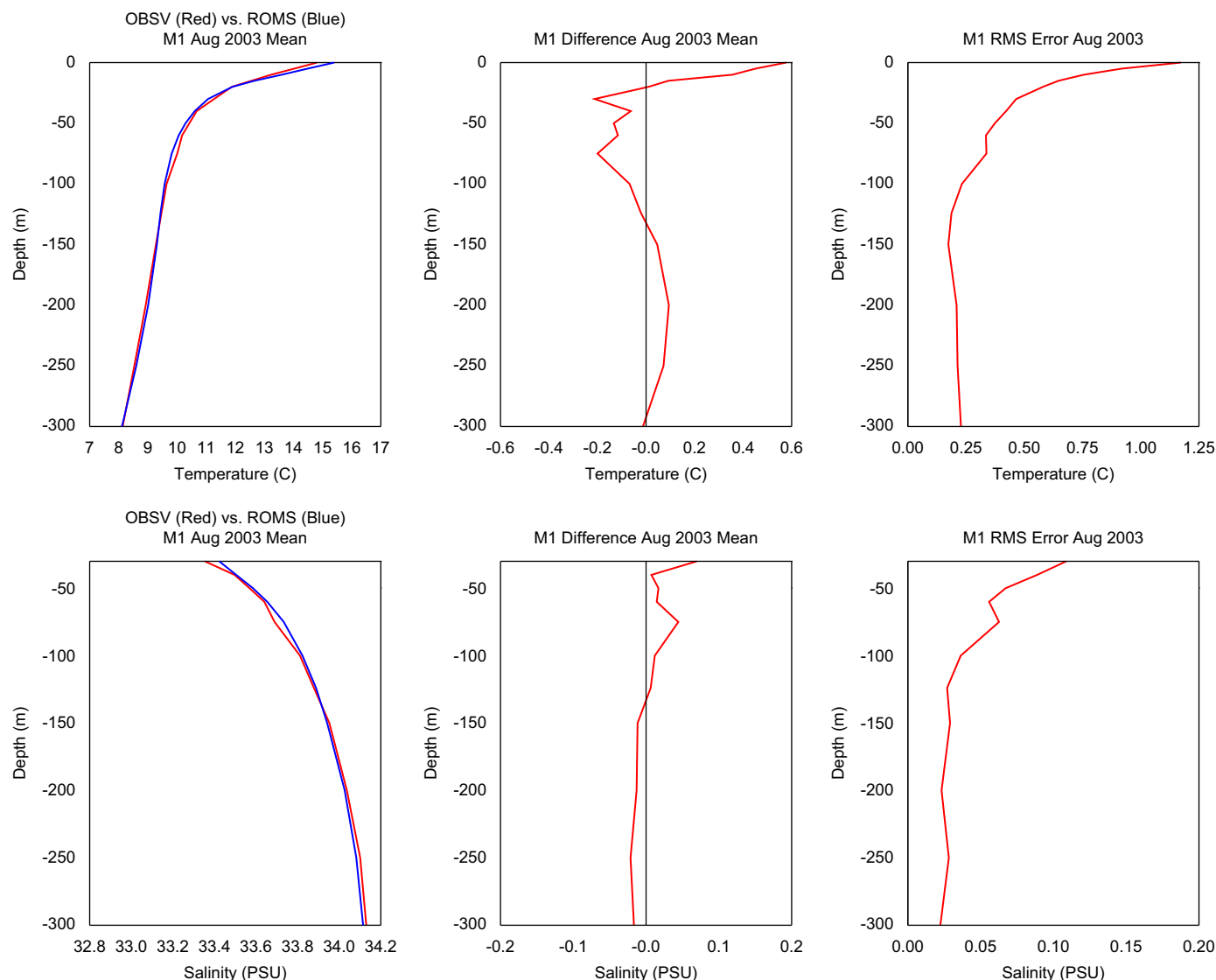

Fig. 17. Comparison profiles of temperature (top) and salinity (bottom) at the M1 mooring location. The left panels show the mean over the August 2003 for mooring observations (red) and ROMS DAS reanalysis (blue). The middle panels show the mean difference and the right panels show the RMS difference. Note that mooring temperature and salinity measurements are assimilated in the ROMS reanalysis.

depths they are generally much smaller, though there is a secondary maximum between 100 and 150 m that approaches 0.5 °C. For the salinity, the mean bias is nearly zero below 150 m, it becomes positive (salty bias) and increases toward a maximum of around 0.05 psu (practical salinity unit). The RMS differences are largest near the surface with values around 0.1 psu. RMS differences then decrease gradually through 150 m before leveling out at a value of approximately 0.05 psu at that level and below.

5.4. Mooring measurements of vertical temperature and salinity profiles

The MBARI moorings, M1 and M2 (locations shown in Fig. 2), collected data throughout the experiment and we use the temperature and salinity measurements, which were assimilated, to verify the consistency of the analysis fields. Fig. 13 shows the temperature as a function of depth and time at the M1 mooring,
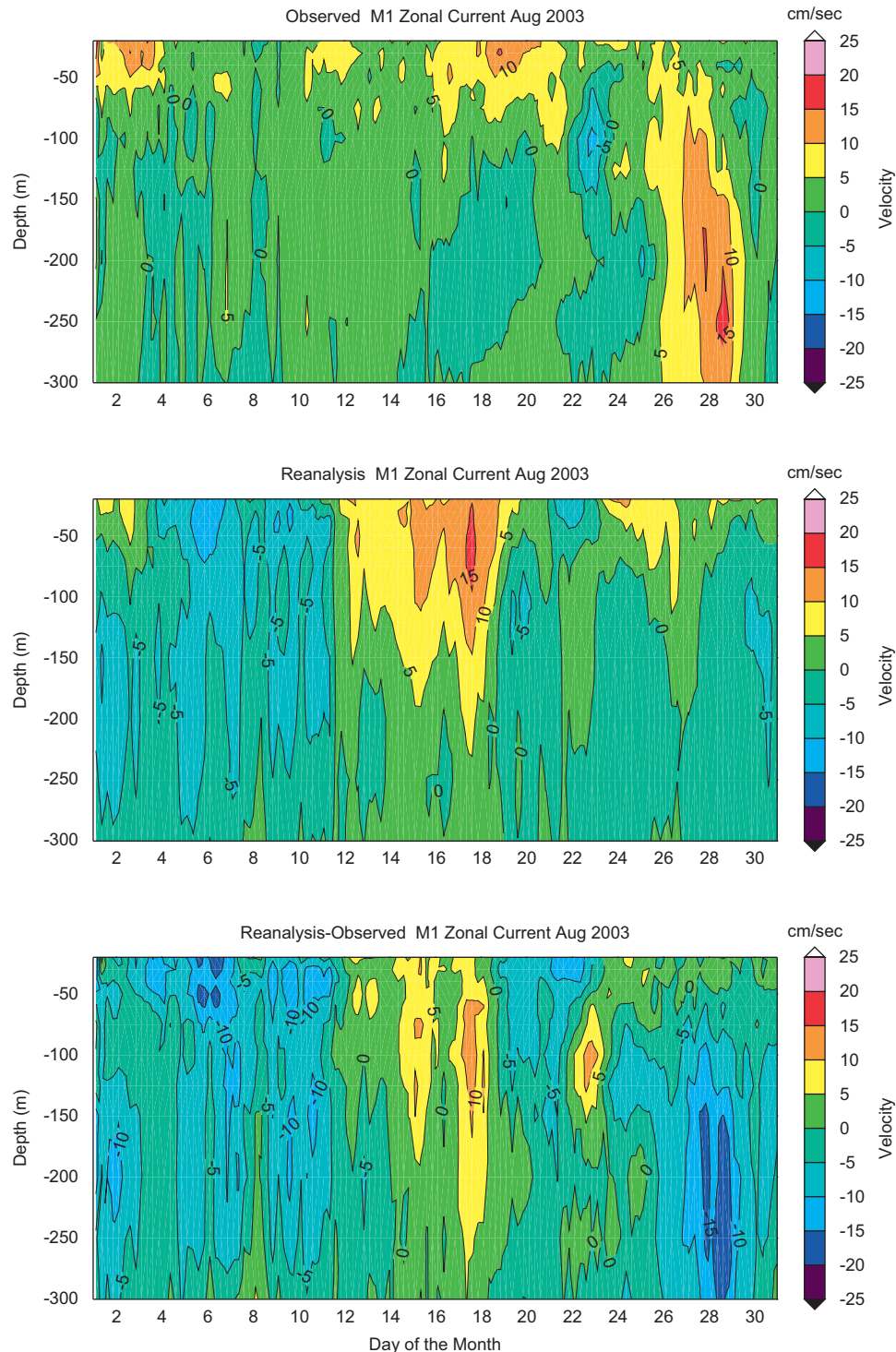

Fig. 18. Time series of zonal current as a function of depth as derived from mooring observations at the M1 location (top), the ROMS reanalysis (middle), and the difference between the two (bottom). A 24-h running mean has been applied with an attempt to filter out the dominant semi-diurnal tidal fluctuations. Note that mooring current measurements are not assimilated in the ROMS reanalysis.

which is within MB. Also shown is the corresponding evolution of the ROMS analysis temperature at this point and the difference between the two. Near the surface waters are warmest during the beginning of the month (before the cold waters generated to the north during the first upwelling event spread into MB) and during the relaxation event later in the month (20–24th), with cooler temperatures in-between and toward the end of the month. The ROMS fields reproduce this evolution quite faithfully. This is confirmed by the difference plot (lowermost panel) that shows differences that are nearly always less than 2 °C and which only occasionally become larger than 1 °C. Near-surface differences are mostly positive (ROMS too warm), while tending to be negative (ROMS too cold) below. Further offshore at M2 (Fig. 14), the near-surface temperatures are warm and tend to increase (though not uniformly) through the month, reaching a maximum just after the relaxation period, before dropping off sharply during the last few days of the month. As at M1, the ROMS analysis reproduces this evolution quite well, including the sharp drop at the end of the month. The largest differences (lowermost panel) of slightly more than 2 °C occur during a small decrease in near-surface
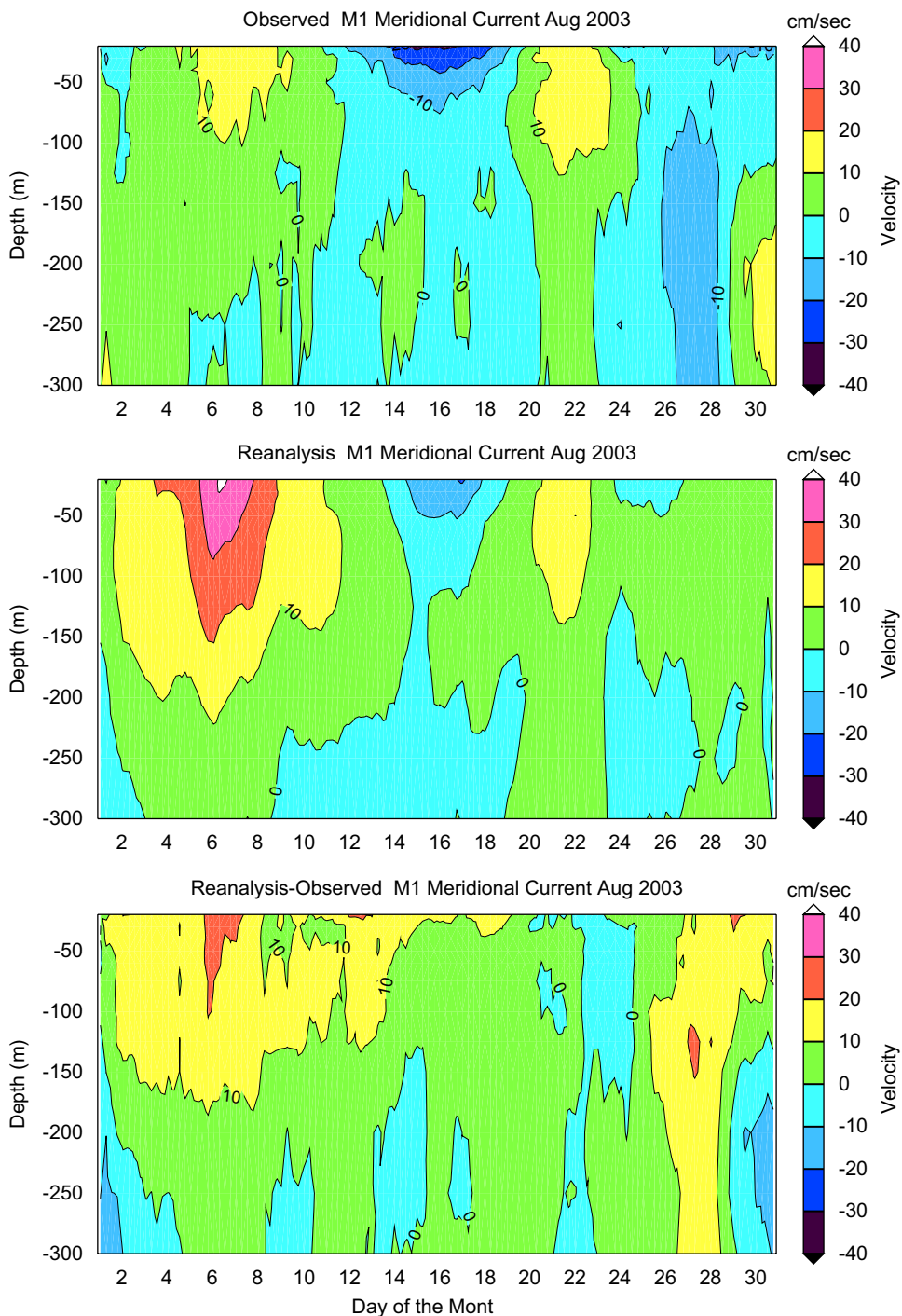

Fig. 19. The same as Fig. 18, except for meridional current.

temperatures at the beginning of the relaxation period. Similar to M1, there is also a tendency for positive differences (ROMS too warm) near the surface and negative differences (ROMS too cold) below.

Comparing mooring salinities as shown in Figs. 15 and 16, we see that at M1 there is a more complicated (than temperature) evolution. Note that we do not compare salinities at M1 at depths shallower than 20 m, as the data at levels above this was of lower quality (F. Chavez, 2006, personal communication). Though the evolution is not a simple one, we see that the ROMS analysis is

able to reproduce most of the observed variability. The exception is an incursion of high salinity water around August 8 in ROMS that has no counterpart in the observations. At this time, differences between ROMS and the mooring exceed 0.4 psu. At all other times, they are less than 0.3 psu, tend to be positive (ROMS too salty) and tend to be largest near 20 m (the shallowest depth compared). There also seems to be a slight tendency for the differences to become smaller with time. At M2, near-surface salinities tend to be low, except for an increase during the relaxation period. The ROMS fields are able to reproduce quite
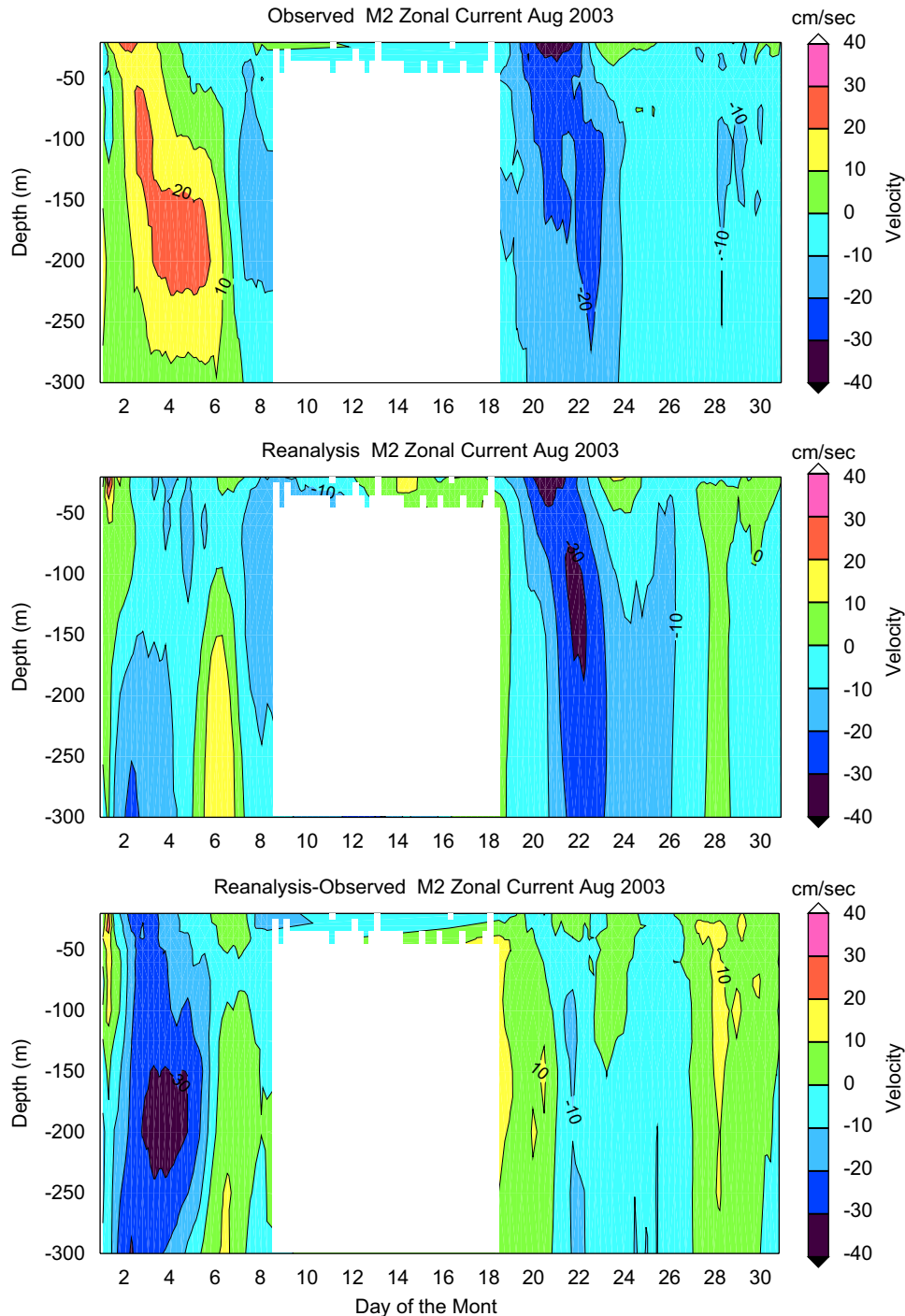

Fig. 20. The same as Fig. 10 except at the M2 location.

well this increase and much of the rest of the variability. Overall, the differences between ROMS and M2 are slightly less than those at M1, with values smaller than 0.2 psu at almost all times. As at M1, the differences tend to be positive (ROMS too salty).

Fig. 17 summaries the comparison of T and S at the M1 mooring showing the monthly mean profiles, differences, and RMS differences. Except near the surface, the temperature difference profile is very similar to the monthly mean all-glider difference profile (Fig. 12) showing negative differences (ROMS too cold) of up to 0.2 °C between 20 and 150 m. Near the surface, there are large positive differences (ROMS too warm) of up to 0.6 °C where

the all-glider plot showed near zero differences. The RMS differences are largest (nearly 1 °C) near the surface and drop off rapidly toward a minimum of 0.25 °C, broadly similar to the all-glider temperature comparison. For the salinity, the differences are only slightly larger than those obtained in the all-glider comparison (generally less than 0.05 psu) and show the same ROMS salty bias near the surface. The profile of RMS differences shows an increase from less than 0.05 psu at 300 m to approximately 0.1 psu near the surface, also very similar to that of the all-glider RMS salinity differences. The summary for M2 (not shown) tells a similar story. Mean temperature differences
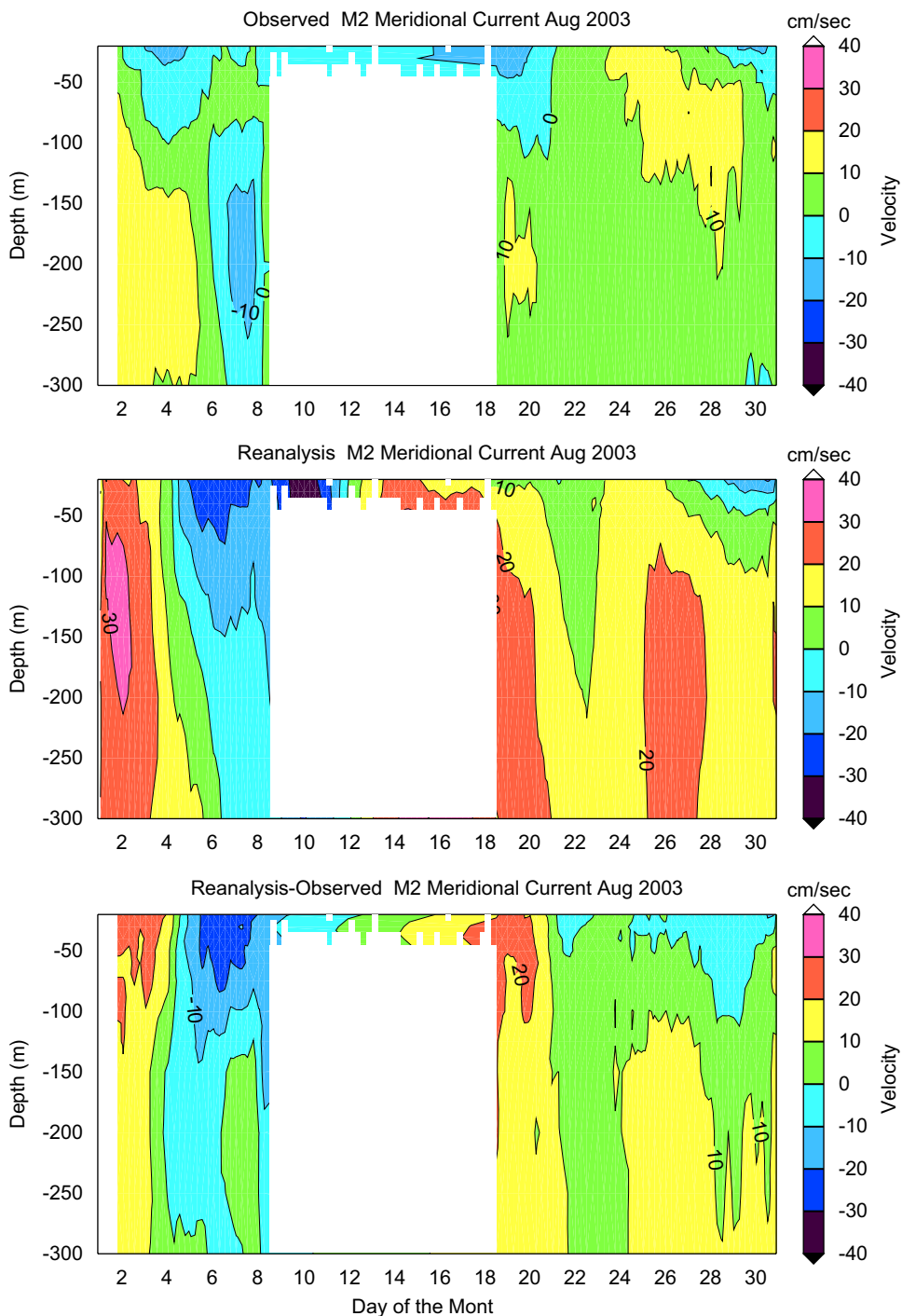

Fig. 21. The same as Fig. 19 except at the M2 location.

are positive (though less so than at M1) near the surface and then negative immediately below, with values of less than 0.5 °C. The RMS differences are largest near the surface and approach 1 °C. For the salinity, the mean and RMS differences are smaller than at M1 and for the all-glider, with the mean differences being less than 0.05 psu everywhere and the RMS differences reaching a maximum of about 0.075 psu near the surface. The mean bias is positive in agreement with the results at M1 and the all-glider.

6. Evaluation of ROMS reanalysis with independent observations

In this section, we subject the analysis to a more challenging test, comparing the ocean current data from moorings, gliders and HF radars with those produced by ROMS.

6.1. Measurements of vertical current profiles from mooring ADCP

Fig. 18 shows the zonal current observed at M1 (within MB) and the ROMS analysis at this location zonal current as a function of depth and time (note that a 24-h running mean

has been applied in an attempt to filter out the fluctuations associated with the semi-diurnal tide). Overall, we note that while there are similarities at some times, there are some large differences as well. For example, significant differences are found below 50 m and during and after the relaxation period. For the M1 meridional current (Fig. 19), there is no corresponding underestimate and overall, the evolution shows a rough similarity to the observed in that there is poleward flow during the first 10 days of the month and during the relaxation period, with weaker equatorward flow in-between and afterwards. Nevertheless, there are clearly substantial differences within this evolution, including a large overestimate in the magnitude of the poleward flow during the early part of the month.

For the zonal current at M2 (Fig. 20), there is a period of eastward currents early in the month, followed by a 10-day period of missing data, followed by westward currents during the relaxation period, and then weak flow for the remainder of the month. The ROMS analysis does well during the relaxation period, but not as well at other times when differences can exceed 20 cm/s. The analysis is better for the meridional current at M2 (Fig. 21) with agreement in the broad features and no glaring
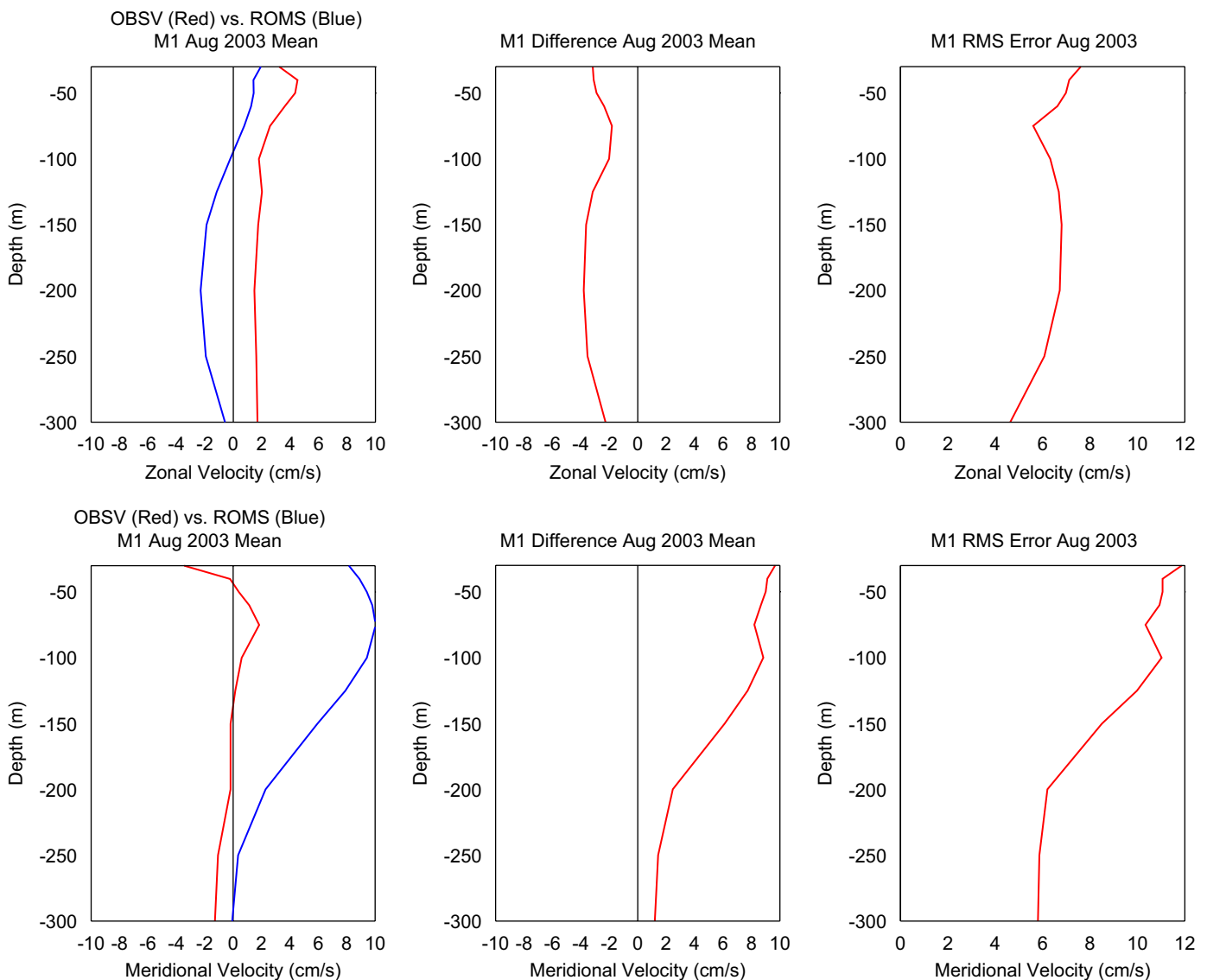

Fig. 22. As in Fig. 17, except for zonal and meridional velocities.

differences in magnitude. The agreement is particularly good for the last half of the month when differences are mostly less than 15 cm/s. Fig. 22 contains a summary for the M1 mooring location (as in Fig. 17 for temperature and salinity) showing the monthly mean profiles, differences, and RMS differences between the reanalysis and observed values of the zonal and meridional velocities. The differences are substantially larger than those found in our comparisons with assimilated data such as T and S. For the zonal velocity, a mean bias of 2–4 cm/s exists throughout most of the vertical range between 300 and 20 m, with the reanalysis showing weakly negative values, while weakly positive velocities were observed. The RMS differences range from 4 to 8 cm/s and are generally larger nearer the surface. For the meridional velocity, the reanalysis shows larger positive values throughout the column, with mean differences that increase from less than 2 cm/s below to 200 m to about 10 cm/s near the surface. The RMS differences are somewhat larger than those for the zonal velocity ranging from approximately 6 cm/s at 300 m to nearly 12 cm/s near the surface.

6.2. Surface current measurements by HF radar

Surface currents were measured by several HF radars during the experiment. A 33-h filter was applied to the hourly HF radar data to remove the tidal current on the diurnal and semi-diurnal time scales (Paduan and Rosenfeld, 1996). Because the 33-h filter also removes a significant portion of the diurnal cycle that is primarily driven by the air–sea fluxes, the following description will only focus on the synoptic fluctuations associated with the upwelling and relaxation period. Maps of the mean August 2003 surface currents are shown in Fig. 23 for the region covered by the HF radar data. The HF radar data reveal an equatorward moving mean surface current with a core in the outer portion of MB with a magnitude of 15–20 cm/s. Mean currents are much weaker inshore of this area and slightly weaker further offshore. The ROMS mean shows a similar pattern with weaker currents inshore and a core of stronger, equatorward flow offshore. The ROMS current is centered further offshore and is stronger (20–25 cm/s) than the HF radar observed one.

Fig. 24 shows the corresponding maps during the upwelling period (6–19 August) and reveals a pattern of observed currents that is quite similar to the monthly mean one, but with stronger flow (note the change in color bar) with core values exceeding 30 cm/s. The ROMS analysis is consist with this feature, showing its offshore-displaced current has also increased in magnitude to greater than 35 cm/s. During the relaxation period (20–24 August) as shown in Fig. 25, the observed currents weaken substantially in the region where they were strong, leaving only a small area of equatorward moving water (10 cm/s) inside MB. In the ROMS analysis, on the other hand, there has been a shift of the equatorward moving current from outside the Bay to inside and a weakening such that the current is now moving at 10–15 cm/s. We do not have a complete explanation for the differences between ROMS and the HF radar measured currents during this period, but note that they could be related to the errors in the ROMS salinity field at M2 (Fig. 16) at this time. A possible source of these salinity errors is the salinity fields used as lateral boundary conditions for the outermost model domain. The temperature, salinity and velocity fields specified there are derived from a monthly Levitus climatology (Levitus et al., 1998) and thus may substantially deviate from the values for a given individual day and year. We are currently in the process of replacing these with fields taken from a daily real-time analysis of the global ocean being produced by the US Navy's Fleet Numerical Meteorological and Oceanographic
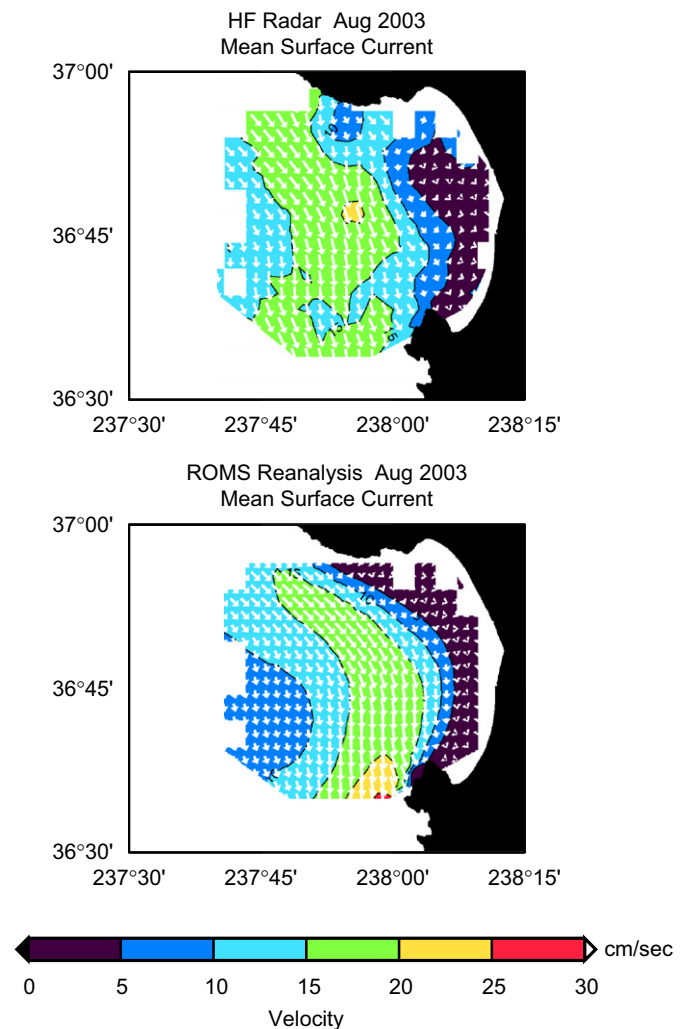

Fig. 23. Horizontal maps of surface current (in arrows with length as well as color representing the current amplitude) averaged over August 2003, as derived from HF radar observations (top) and ROMS reanalysis (bottom). The observed surface current is derived from the 33-h filtered HF radar measurements. Note that HF radar current measurements are not assimilated in the ROMS reanalysis.

Center as part of the GODAE program. We plan to use the most recently available 3-day mean fields from this analysis and expect that these fields will better represent the large-scale ocean boundary conditions. It is also important to keep in mind that no current data at all were assimilated into ROMS during the experiment and so the deficiencies we see in the comparisons with mooring and HF radar currents suggest that the assimilation of such data will be necessary to improve the analysis and forecasts.

6.3. Vertically integrated current derived from gliders

The SIO glider data also were processed to extract a depth-averaged current for each dive. Note that this data was not assimilated into ROMS and thus represents a truly independent dataset for comparison. Fig. 26 shows the depth-averaged SIO glider current vectors (black) along with the co-located ROMS depth-averaged vectors (red) for the 4 weeks of August 2003. The depths are different for each dive and ranges from 50 (or less) to 800 m. Note that not all the glider current vectors have a
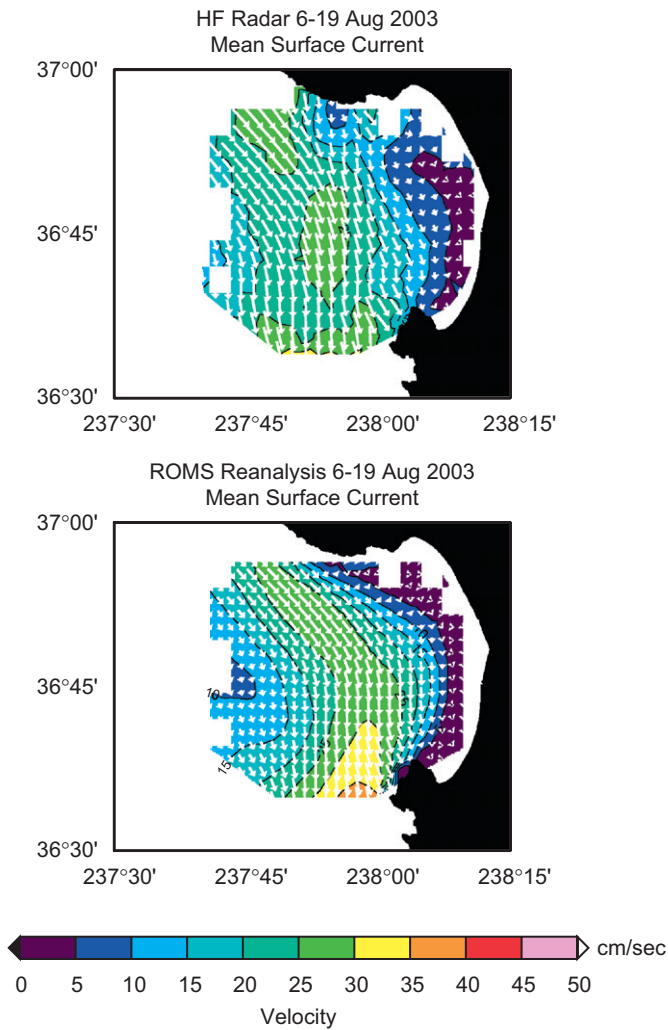
Fig. 24. The same as Fig. 23 except during the upwelling phase: August 6–19, 2003.
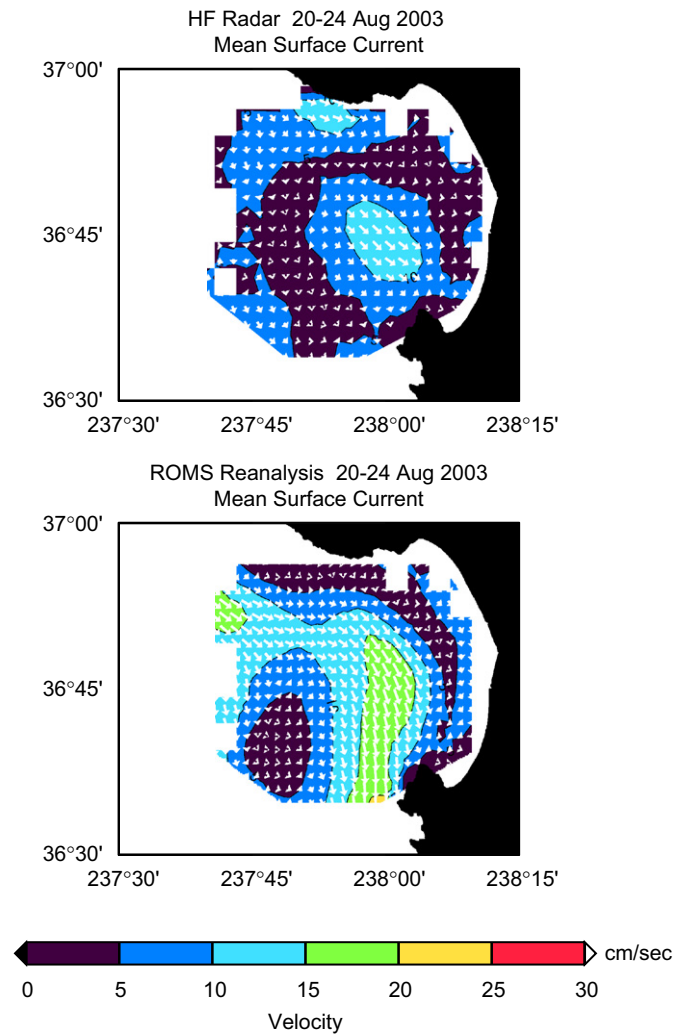
Fig. 25. The same as Fig. 23 except during the relaxation phase: August 20–24, 2003.

corresponding ROMS vector plotted. In these cases we were unable to produce a depth-average from ROMS corresponding to the glider depth average due to differences between the ROMS model bottom topography and the real topography. The primary feature revealed in these data is a poleward flowing undercurrent that tends to be strongest offshore, outside of MB. The ROMS current vectors are in good qualitative agreement with the glider observed ones; the same poleward moving undercurrent can be seen. The RMS differences averaged over all vectors are 3.6 cm/s for the meridional component of the current and 3.3 cm/s for the zonal component.

7. ROMS forecast and predictability

One of the major motivations to implement a coastal ocean observing system and assimilate real-time measurements into numerical models is to issue operational forecasts of the 3D ocean fields similar to what synoptic meteorologists have been doing for the past few decades. Fig. 27 shows the first attempt to quantify the prediction skill by showing the evolution of the spatial correlation (upper panel) and RMS differences (lower panel) of the zonal (red) and meridional(blue) components

between the glider measured and ROMS analyzed depth-averaged velocities as a function of forecast hour. These show a steady decrease in correlation (from the hour 0 values of 0.61 and 0.66 for zonal and meridional components, respectively) and increase in RMS differences (from the values of 3.6 and 3.3 cm/s, respectively) to values nearly one-half (correlation) and double (RMS) the initial ones by the end of the 48-h forecasts. In fact, in the RMS differences, there seems to be an acceleration of the increase after 36 h. This relatively rapid error-growth underscores the need to assimilate current data.

8. Summary and concluding remarks

Based on the Regional Ocean Modeling System (ROMS), we have developed and implemented a real-time forecast system for the central California coastal ocean centered at the MB, California. There are three nested ROMS domains including the US West coastal ocean at 15-km resolution, the central California coastal ocean at 5 km, and the MB region at 1.5 km. All three nested ROMS models have 32 vertical sigma layers, used atmospheric forcing produced by the Navy's COAMPS model and were integrated in
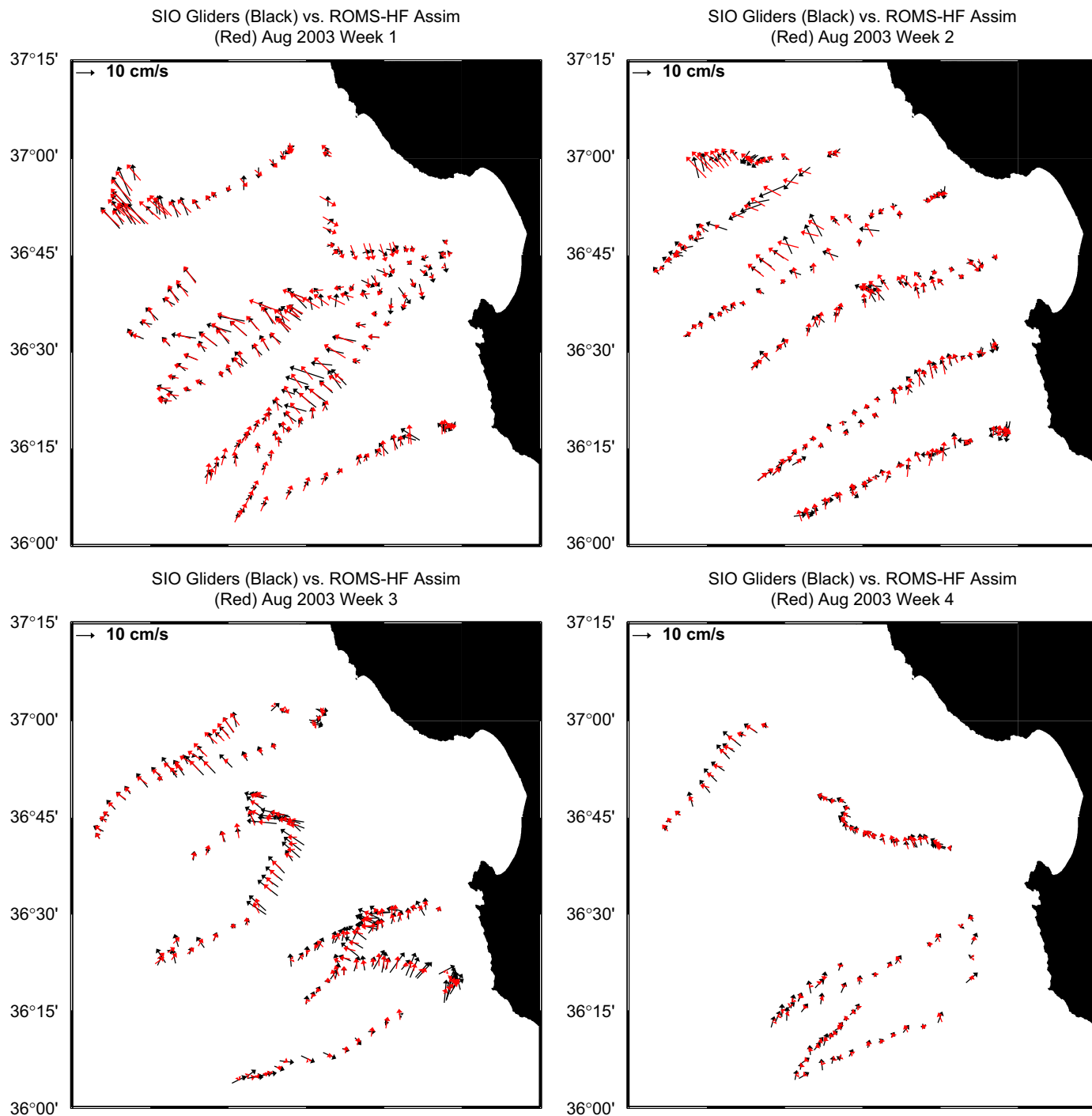
Fig. 26. Weekly maps of the vertically integrated current as derived from the SIO glider measurements (black) and ROMS reanalysis (red) during the first (a), second (b), third (c) and forth (d) week of August 2003. Note that glider derived current data are not assimilated in the ROMS reanalysis.

conjunction with a 3D variational data assimilation algorithm (3DVAR) to produce snapshots of the ocean state every 6 h (the reanalysis) and 48-h forecasts once a day.

Results produced by the ocean forecast system during the AOSN field experiment in August 2003 were evaluated using both assimilated and independent data sets. The ROMS reanalysis temperature and salinity showed excellent quantitative agreement with both assimilated and independent data with mean differences found to be less than 0.6 °C and 0.1 psu, respectively, and RMS differences of less than 1 °C and 0.1 psu. In addition, the

reanalysis showed the ability to capture both the upwelling and relaxation conditions plus a synoptic event—shown in SST data—involving the offshore extension of cold waters during the second upwelling period at the end of August 2003. It should be pointed out that the real-time model performance during the field experiment is somewhat worse than the ROMS reanalysis skill presented here mostly due to improvements in the model and data assimilation schemes as well as the assimilation of additional data that became available during the post-experiment reanalysis phase.
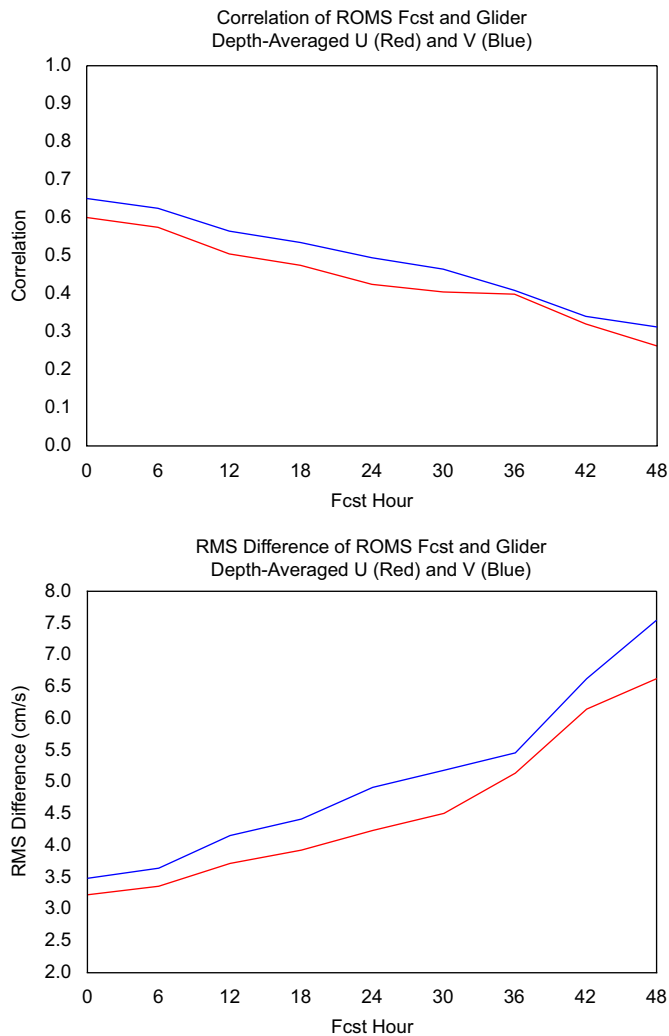
Correlation of ROMS Fcst and Glider Depth-Averaged U (Red) and V (Blue)

RMS Difference of ROMS Fcst and Glider Depth-Averaged U (Red) and V (Blue)

Fig. 27. Spatial correlation (top) and RMS difference (bottom) between the glider derived currents and ROMS forecast currents as a function of forecast lead time (hours). Colors represent the zonal (red) and meridional (blue) currents, respectively. Note that glider-derived current data are not assimilated in the ROMS reanalysis.

Because the current ROMS data assimilation system only assimilates the temperature and salinity data, all the current measurements can be viewed as independent observations. Qualitative agreement was found between the independent current measurements and the ROMS reanalysis. The agreement is particularly good for the vertically integrated current along the SIO glider tracks: the ROMS reanalysis can realistically reproduce the poleward California Undercurrent. Reasonably good agreement is also found in the spatial patterns of the surface current as measured by HF radars, although the amplitudes and exact location of current features can be quite different. This poses the challenge in comparing the ROMS reanalyzed current against the ADCP current measurement at single mooring locations. Our preliminary results suggest the urgent need to include velocity data in our assimilation system.

In summary, we have met most of the AOSN field experiment goals to deliver an operational ocean forecast in real time. During the field experiment, we were able to produce a nowcast and 3-day forecast every day. With the increasing computer power during the reanalysis phase, we are able to reduce the data assimilation window from 24 to 6 h. During the follow-up 2006 MB field experiment, we were able to sustain a 6-h data assimilation window. Except the mooring temperature and salinity measurements, we were able to assimilate all the temperature and salinity data into ROMS from CTDs collected from a variety of *in situ* (ships, gliders, AUVs) and remote sensing (aircraft, satellite) platforms and the ROMS nowcast and forecast fields were provided to the entire AOSN team in real time. Although the ROMS nowcast and forecast results (along with the HOPS products) were presented daily or twice-daily in the AOSN "control" room during the field experiment, they were not routinely used to guide the adaptive sampling to close the feedback loop from the modelers to instrument operators, a major objective of the AOSN field experiment. The reasons are many including at least the following factors. During the 2003 field experiment, we had relatively little experience in distributing and visualizing the 3D model output. The ROMS and HOPS nowcast and forecast fields were sometimes presented in different format and styles. The instrument operators and adaptive sampling planners had a difficult time reconciling the two model results, which sometimes showed significant differences, and using them to make decisions for adjusting the sampling strategy. In addition, neither modeling group was prepared to deliver the large 3D model output in a way transparent to the entire AOSN team. This situation was significantly improved during the 2006 experiment.

In addition to the real-time operational requirements, we have also identified several issues for future improvements. The result that ROMS analyzed offshore currents (e.g., glider derived current) significantly better than those near-shore (e.g., M1 mooring) suggests a challenge in reproducing the complex velocity structure near the coast. The model topography used in the current configuration is overly smoothed and can be improved for the next generation model configuration. During the AOSN experiment, only temperature and salinity data were assimilated into ROMS. Our results suggest the immediate need to assimilate current measurements. The assimilation of velocity observations leads to another challenge due to baroclinic tidal velocities. Currently, we are developing a tidal resolving model, and the data assimilation can be carried out with a model version that explicitly resolves tides.

Another model deficiency is the large-scale boundary condition for the US west coast ROMS domain at 15 km that is currently based on the Levitus et al. (1998) climatology, which contains only the seasonal changes and lacks the year-to-year fluctuations. We are currently looking into several real-time global ocean analysis systems (e.g., the global Navy NCOM model) that can be potentially used to provide real-time large-scale boundary conditions. The 3DVAR data assimilation also can be improved by including a more realistic error covariance. During the 2003 AOSN field experiment, only one deterministic forecast was made for every nowcast initial condition. We are developing methods to make ensemble forecasts using slightly different initial conditions or even different surface forcing or boundary conditions. Collaborations with the adaptive sampling planner to close the loop from models to adaptive sampling are also underway.

Compared to atmospheric observing systems, ocean measurements are relatively sparse in space and often sporadic in time. Numerical ocean models and the associated data assimilation schemes are also less mature. Because of the smaller spatial scales involved in the coastal ocean, there are also computational challenges to model the coastal ocean over a larger region, particularly including the interaction and feedback between different coastal oceans and more importantly between the coastal and regional/global scale processes. As the community deploys more *in situ* and remote sensing sensors in the ocean and improves the modeling and data assimilation methods, our ultimate goal is to make routine nowcasts and forecasts for the

coastal oceans on an operational 24/7 basis and deliver the data and model derived products to users in real time in a user-friendly format.

Acknowledgements

References

Barnier, B., Marchesiello, P., de Miranda, A.P., Molines, J.M., Coulibaly, M., 1998. A sigma-coordinate primitive equation model for studying the circulation in the South Atlantic. Part I: model configuration with error estimates. Deep-Sea Research I 45, 543–572.

Blayo, E., Debreu, L., 1999. Adaptative mesh refinement for finite-difference ocean models: first experiments. Journal of Physical Oceanography 29, 1239–1250.

Burk, S.D., Haack, T., Samelson, R.M., 1999. Mesoscale simulation of supercritical, subcritical, and transcritical flow along coastal topography. Journal of the Atmospheric Sciences 56, 2780–2795.

Capet, X., Marchesiello, P., McWilliams, J.C., 2004. Upwelling response to coastal wind profiles. Geophysical Research Letters 31, L13311.

Chao, Yi, Li, Z., Kindle, J.C., Paduan, J.D., Chavez, F.P., 2003. A high-resolution surface vector wind product for Coastal Oceans: blending satellite Scatterometer measurements with regional Mesoscale atmospheric model simulations. Geophysical Research Letters 30 (1), 1013.

Cohn, S.E., da Silva, A., Guo, D.M., Sienkiewicz, J.M., Lamich, D., 1998. Assessing the effects of data selection with DAO physical-space statistical analysis system. Monthly Weather Review 126, 2913–2926.

Daley, R., 1991. Atmospheric Data Assimilation. Cambridge Atmospheric and Space Science Series, Cambridge University Press, Cambridge, UK, 457pp.

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 AOSNII field campaign. Deep-Sea Research II, this issue [doi:10.1016/j.dsr2.2008.08.009].

Eriksen, C.C., Osse, T.J., Light, R.D., Wen, T., Lehman, T.W., Sabin, P.L., Ballard, J.W., Chiodi, A.M., 2001. Seaglider: a long range autonomous underwater vehicle for oceanographic research. IEEE Journal of Oceanic Engineering 26 (4), 424–436.

Haack, T., Burk, S.D., Dorman, C., Rogers, D., 2001. Supercritical flow interaction within the Cape Blanco–Cape Mendocino orographic complex. Monthly Weather Reviews 129, 688–708.

Hickey, B.M., 1998. Coastal oceanography of western north America from the tip of Baja California to Vancouver Island. In: Robinson, A.R., Brink, K.H. (Eds.), The Sea—The Global Coastal Ocean, Regional Studies and Synthesis. Wiley, New York, pp. 345–393.

Hodur, R.M., 1997. The Naval Research Laboratory's Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS). Monthly Weather Reviews 125, 1414–1430.

Jazwinski, A.H., 1970. Stochastic Processes and Filtering Theory. Academic Press, New York.

Kondo, J., 1975. Air–sea bulk transfer coefficients in diabatic conditions. Boundary-Layer Meteorology 9, 91–112.

Large, W.G., Pond, S., 1981. Open ocean momentum flux measurements in moderate to strong winds. Journal of Physical Oceanography 11, 324–481.

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

Levitus, S., Boyer, T.P., Conkright, M.E., O' Brien, T., Antonov, J., Stephens, C., Stathoplos, L., Johnson, D., Gelfeld, R., 1998. NOAA Atlas NESDIS 18, World Ocean Database 1998: vol. 1: Introduction. US Government Printing Office, Washington, DC, 346pp.

Li, X., Chao, Yi, McWilliams, J.C., Fu, L.-L., 2001. A comparison of two vertical mixing schemes in a Pacific OGCM. Journal of Climate 14, 1377–1398.

Li, Z., Navon, I.M., 2001. Optimality of variational data assimilation and its relationship with the Kalman filter and smoother. Quarterly Journal of the Royal Meteorological Society 127, 661–683.

Li, Z., Chao, Y., McWilliams, J.C., Ide, K., 2008a. A three-dimensional variational data assimilation scheme for the Regional Ocean Modeling System, Part I: formulation. Journal of Geophysical Research, submitted for publication.

Li, Z., Chao, Y., McWilliams, J.C., Ide, K., 2008b. A three-dimensional variational data assimilation scheme for the Regional Ocean Modeling System: implementation and basic experiments. Journal of Geophysical Research 113, C05002, doi:10.1029/2006JC004042.

Liu, D.C., Nocedal, J., 1989. On the limited memory BFGS method for large scale optimization. Mathematical Programming 45, 503–528.

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

Marchesiello, P., McWilliams, J.C., Shchepetkin, A., 2003. Equilibrium structure and dynamics of the California Current System. Journal of Physical Oceanography 33, 753–783.

Miyakoda, K., Rosati, A., 1977. One way nested grid models: the interface condition and numerical accuracy. Monthly Weather Reviews 105, 1108–1118.

Paduan, J.D., Rosenfeld, L.K., 1996. Remotely sensed surface currents in Monterey Bay from shore-based HF radar (CODAR). Journal of Geophysics Research 101, 20669–20686.

Pennington, J., Chavez, F., 2000. Seasonal fluctuations of temperature, salinity, nitrate, chlorophyll and primary production at station H3/M1 over 1989–1996. Deep-Sea Research II 47, 947–973.

Penven, P., Debreu, L., Marchesiello, P., McWilliams, J.C., 2006. Application of the ROMS embedding procedure in the California Current upwelling system. Ocean Modelling 12, 157–187.

Raymond, W.H., Kuo, H.L., 1984. A radiation boundary condition for multi-dimensional flows. Quarterly Journal of the Royal Meteorological Society 110, 535–551.

Rosenfeld, L.K., Schwing, F., Garfield, N., Tracy, D.E., 1994. Bifurcated flow from an upwelling center: a cold water source for Monterey Bay. Continental Shelf Research 14, 931–964.

Shchepetkin, A.F., McWilliams, J.C., 2005. The Regional Oceanic Modeling System: a split-explicit, free-surface, topography-following-coordinate ocean model. Ocean Modelling 9, 347–404.

Sherman, Davis, R.E., Owens, W.B., Valdes, J., 2001. The autonomous underwater glider spray. IEEE Journal of Oceanic Engineering 26 (4), 437–446.

Wang, P., Song, Y.T., Chao, Y., Zhang, H., 2005. Parallel computation of the Regional Ocean Modeling System. International Journal of High Performance Computing Applications 19 (4), 375–385.