Development, Implementation, and Validation of a California Coastal Ocean Modeling, Data Assimilation, and Forecasting System

Yi Chao[1,2], John D. Farrara[2], Hongchun Zhang[1], Kevin J. Armenta[3], Luca Centurioni[4], Francisco Chavez[5], James B. Girton[6], Dan Rudnick[4], Ryan K. Walter[3]

[1] Joint Institute for Regional Earth System Science and Engineering, University of California, Los Angeles, CA 90095, USA

[2] Remote Sensing Solutions, Monrovia, CA 91016, USA

[3] Physics Department, California Polytechnic State University, San Luis Obispo, CA 93407, USA

[4] Scripps Institution of Oceanography, University of California, San Diego, CA 92037, USA

[5] Monterey Bay Aquarium Research Institute, Moss Landing, CA 95039, USA

[6] Applied Physics Laboratory, University of Washington, Seattle, WA 98105, USA

Deep-Sea Research II: CenCal Observations
Submitted, March 7, 2016
Revised October 10, 2016

Abstract

A three-dimensional, near real-time data-assimilative modeling system for the California coastal ocean is presented. The system consists of a Regional Ocean Modeling System (ROMS) forced by the North American Mesoscale Forecast System (NAM). The ocean model has a horizontal resolution of approximately three kilometers and utilizes a multi-scale three-dimensional variational (3DVAR) data assimilation methodology. The system is run in near real-time to produce a nowcast every six hours and a 72-hour forecast every day. The performance of this nowcast system is presented using results from a six-year period of 2009-2015.

The ROMS results are first compared with the assimilated data as a consistency check. RMS differences in observed satellite infrared sea surface temperatures (SST) and vertical profiles of temperature between observations and ROMS nowcasts were found to be mostly less than 0.5^{o}C, while the RMS differences in vertical profiles of salinity between observations and ROMS nowcasts were found to be 0.09 or less. The RMS differences in SST show a distinct seasonal cycle that mirrors the number of observations available: the nowcast is less skillful with larger RMS differences during the summer months when there are less infrared SST observations due to the presence of low-level clouds. The larger differences during summer were found primarily along the northern and central coasts in upwelling regions where strong gradients exist between colder upwelled waters nearshore and warmer offshore waters. RMS differences between HF radar surface current observations and ROMS nowcasts were approximately 7-8 cm s^{-1}, which is about 30% of the time mean current speeds in this region. The RMS differences in sea surface height (SSH) between the AVISO (Archiving, Validation and Interpretation of Satellite Oceanographic) altimetric satellite observations and ROMS nowcasts were about 2 cm.

In addition, the system realistically reproduces the interannual variability in temperatures at the M1 mooring (122.03°W, 36.75°N) in Monterey Bay, including the strong warming of the California coastal ocean during 2014.

The ROMS nowcasts were then validated against independent observations. A comparison of the ROMS nowcast with independent profile observations of temperature and salinity shows RMS differences of 0.7 to 0.92°C and 0.13 to 0.17, which are larger (by up to a factor of 2) than the differences found in the comparisons with assimilated data. Validation of the depth-averaged currents derived from Spray gliders shows that the flow patterns associated with California Current and California Undercurrent/Davidson current systems and their seasonal variations are qualitatively reproduced by the ROMS modeling system.

Lastly, the impact of two recent upgrades to the system is quantified. Switching the lateral boundary conditions from a U.S. west coast regional model to the global HYCOM (HYbrid Coordinate Ocean Model) model results in an improvement in the simulation of the seasonal and interannual variations in the SSH, especially south of Pt. Conception (120.47°W, 34.45°N). The assimilation of altimetric satellite SSH data also results in an improvement in the model surface currents when compared to independent surface drifter observations.

Keywords: coastal ocean modeling, model validation, Californa Current system

1. Introduction

The California coastal ocean is one of the United States' most important resources, both economically for its fisheries and ecologically for its diversity. It is home to several National Marine Sanctuaries, including the Monterey Bay National Marine Sanctuary (*montereybay.noaa.gov*), a federally protected marine area offshore of California's central coast. Both fisheries and ecosystems are quite sensitive to changes in environmental conditions such as temperature, salinity, and currents as evidenced by the impacts of the major warming event of 2014-15 [*Whitney,* 2015; *Opar*, 2015; *NOAA*, 2016a, b]. The need to understand and predict changes in these variables has been underscored recently by this unprecedented warming event in the region [*Di Lorenzo and Mantua*, 2016]. Beginning in 2014, exceptionally warm temperatures developed across a wide area off the California coast [*Bond et al.* 2015; *Zaba and Rudnick,* 2016]. Near-surface positive temperature anomalies exceeded 4^{o}C in certain regions and persisted for much of the year and into 2015 [*Zaba and Rudnick,* 2016].

With the goal of characterizing and predicting environmental conditions in California's coastal ocean in near real-time, we have developed a three-dimensional data-assimilative modeling system based on the Regional Ocean Modeling System (ROMS) code. This system has been producing nowcasts (analyses of the current ocean state) four times per day and a daily 72-hour forecast in near real-time since 2009 to the present time. By near real-time we mean that model nowcasts and forecasts are generally available about 9 hours behind real-time; for example, the 03 UTC nowcast would be available at around 12 UTC. The ROMS modeling system is an integral part of the Central and Northern California Ocean Observing System (CeNCOOS, *http://www.cencoos.org*) and the Southern California Coastal Ocean Observing System (SCCOOS, *http://www.sccoos.org*), two regional associations of the national Integrated

116 Ocean Observing System (https://ioos.noaa.gov/).

117 To place this current modeling work in the context of some of our previous California
118 coastal ocean modeling work, we review here the model configuration used in *Chao et al.* [2009]
119 and discuss some of the differences between that system and the one presented here. The most
120 important thing to note is that these two modeling systems were designed to achieve different
121 goals. The *Chao et al.* [2009] Monterey Bay (MB) system was designed with the aim of
122 simulating (and forecasting) as realistically as possible the summer circulation within and around
123 Monterey Bay (MB) with an emphasis on coastal upwelling and downwelling events and the
124 transitions between them. In order to do this, a relatively high horizontal resolution ocean
125 model and accurate, high-resolution wind forcing were necessary. A nested modeling approach
126 was chosen that was focused on an innermost nest covering a relatively small area (MB and
127 surroundings) at a relatively high horizontal (1 km) resolution. The CA system described in the
128 current paper was designed to be much more comprehensive in terms of the region covered and
129 phenomena we aim to reproduce (see Section 2.1 for a complete description of these
130 phenomena), as we attempt to realistically simulate the environment further offshore – for
131 example, the California current system and its associated mesoscale eddies - and tides (and tidal
132 currents), while still aiming to do reasonably well in simulating nearshore phenomena such as
133 upwelling.

134 These differences in goals are the reason behind many of the differences between the two
135 systems. For example, in the current CA system we have chosen a uniform intermediate
136 horizontal resolution (3 km) and a few more vertical levels (40 versus 32) applied to a much
137 larger single domain compared to the 1 km resolution for the much smaller innermost domain of
138 the MB system. Also, while tides are not essential for simulating coastal upwelling events and

thus were not included in the MB system, they are included in the CA system as they are essential if we want to reproduce tidal phenomena. With the MB system's focus on upwelling, high-resolution coastal winds were essential and thus the COAMPS atmospheric model wind fields were used for the MB system since they were the highest horizontal resolution (3 km for the innermost nest) winds available during the period simulated (summer 2003), while we use the NAM (either 12 or 5 km) winds for the CA system because the high resolution COAMPS model domain does not cover our entire expanded CA domain. Similar considerations apply to the data assimilation, we have switched to a more comprehensive multi-scale 3DVAR methodology that assimilates temperature, salinity, sea surface height, zonal and meridional current data rather than only temperature and salinity (T, S) since while assimilating (T, S) only was adequate (assuming accurate wind forcing) for simulating the upwelling/downwelling events the MB system focused on, the more comprehensive CA system clearly benefits from assimilating all available data types. The advantages of multi-scale 3DVAR used here compared to the single-scale 3DVAR used in MB system will be outlined in Section 2.5 and the advantages of choosing HYCOM output for the lateral boundary forcing in the CA system as opposed to the Levitus climatology used for the outermost nest in the MB system are discussed in Section 4.

We present here the first comprehensive documentation of the model, the data assimilation method, and the performance of six hourly nowcasts. The performance of the ROMS forecasts will be reported in a separate study. We begin with a detailed description of the ROMS-based modeling system, including the external forcing and data assimilation methodology used, in

Section 2. In Section 3, we present a comparison of the ROMS nowcasts with observations that are assimilated and a validation of the ROMS results by comparing them with independent (non-assimilated) observations. In Section 4, we explore the impact of several recent upgrades to the system. Finally, a summary and some concluding remarks are given in Section 5.

2. The ROMS-based California Coastal Modeling, Data Assimilation and Forecast System

2.1 Ocean Model

The California (CA) coastal ocean modeling system is based on the Regional Ocean Modeling System (ROMS) [*Haidvogel et al.,* 2000]. The ROMS configuration used consists of a single domain covering the entire California coastal ocean from Ensenada, Mexico to north of Crescent City, CA and extending approximately 1000 km offshore at a horizontal resolution of 3.3 km (see Figure 1). This particular model configuration was chosen to achieve several objectives: 1) to simulate the major flow features that characterize the California coastal ocean which include the near-surface equatorward California Current system (CCS) that lies several hundred kilometers offshore, the poleward California Undercurrent (CU) that peaks in strength between 100 and 300 meters below the surface and the wintertime inshore Davidson current; 2) to resolve the mesoscale eddies associated with the CCS that are typically the flow type with the largest kinetic energy in this region [*Capet et al.,* 2008a, b]; 3) to cover the entire area observed by the California HF radar surface current mapping network as well as all regions of interest for the Southern California and Central and Northern California coastal ocean observing systems (SCCOOS and CeNCOOS). Note that the system will not resolve the less energetic submesoscale fronts and eddies in the region discussed by *Capet et al.* [2008a, b] and *McWilliams* [2016], nor the very small-scale circulations often generated very near shore (for

example, rip currents). Eddies associated with the California Undercurrent (also known as Cuddies) are submesoscale coherent vortices of horizontal scale less than 10 km and so are not represented in our system. Lastly, analyses are produced by the system every 6 hours (see Section 2.5) because we aim to describe not only sub-tidal variabilities but also diurnal and semi-diurnal tidal currents and near-coastal wind-driven diurnal variability.

ROMS is a free-surface, hydrostatic, three-dimensional primitive equation regional ocean model [*Shchepetkin and McWilliams,* 2005, 2006; *Marchesiello et al*., 2001]. The horizontal discretization uses a boundary-fitted, orthogonal curvilinear formulation. Coastal boundaries are specified as a finite-discretized grid via land/sea masking. Lateral boundary conditions for the CA domain are provided from a separate data-assimilating ROMS system that covers the entire U.S. west coast and all of Baja California at a resolution of 15 km or, as in a recent upgrade, global HYCOM [*http://hycom.org*]. 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. ROMS uses a sigma-type vertical coordinate characterized by coordinate surfaces that follow the bottom topography. In the CA configuration, there are 40 unevenly-spaced sigma surfaces used with the majority of these clustered near the surface to better resolve processes in the mixed layer. Similar, though nested, ROMS configurations have been successfully applied in Monterey Bay, California [*Chao et al.* 2009] and Prince William Sound, Alaska [*Farrara et al.* 2013]. We comment on the differences between the results presented here and those presented for Monterey Bay by *Chao et al.* [2009] in the Summary section.

2.2 Tidal Forcing

The tidal forcing is added through lateral boundary conditions that are obtained from a global barotropic tidal model (TPXO.6) [*Egbert and Erofeeva,* 2002; *Egbert et al.*, 1994], that has a horizontal resolution of 0.25 degrees and uses an inverse modeling technique to assimilate satellite altimetry cross-over observations. Eight major tide constituents at the diurnal and semidiurnal frequencies (M2, K1, O1, S2, N2, P1, K2 and Q1) are used. The barotropic transport from the TPXO.6 solution is adjusted using the ROMS bottom topography because this topography differs from the TPXO.6 bottom topography. The ROMS bottom topography is based on the ETOPO1 dataset [*Amante and Ekins,* 2009] with some selected smoothing of the raw ETOPO1 data applied in regions where the horizontal gradients in depth are very strong. Specifically, the primary requirement we imposed when constructing the bottom topography for our 3 km grid was that the Haney number, $rx_1(h)$, be less than 5 [see *Sikirić et al.* 2009].

One motivation for using a tide-permitting circulation model is that the temperature, salinity, and current data collected by moving platforms (*e.g.*, gliders or other autonomous underwater vehicles) contain both the tidal signal and non-tidal variability. Assimilating these data into a non-tidal model could introduce additional errors.

2.3 Atmospheric Forcing

The atmospheric forcing required by the ROMS model is derived from hourly output of operational forecasts performed with the National Centers for Environmental Prediction (NCEP) North American (NAM) 5 km (the 12 km resolution NAM is used prior to June 2013) model [*http://www.emc.ncep.noaa.gov/index.php?branch=NAM*]. Daily 00 UTC forecasts are used. The surface latent and sensible heat fluxes, as well as surface evaporation rates, are derived from surface air temperatures, surface relative humidity, 10 m winds and ROMS sea surface temperatures (SST) using the bulk formula proposed by [*Kondo,* 1975]. The fresh water flux is

computed as the calculated evaporation rate minus the NAM precipitation rate (E-P). The wind stress is derived from the 10 m winds using the formula of *Large and Pond* [1982]. Net surface solar and terrestrial radiative fluxes are taken directly from the NAM output. The variables used for computing the ocean model forcing have been evaluated against buoy data (not shown). The surface winds are satisfactorily accurate, with RMS errors of 1-2 m s^{-1} being the norm. The surface air temperatures and relative humidity values also show quite good accuracy with errors of less than 1°C and 10%, respectively. On the other hand, there are significant uncertainties in the solar and terrestrial radiation primarily related to deficiencies in the NAM simulation of stratus clouds in the region [*http://www.jcsda.noaa.gov/documents/meetings/wkshp2009/Session-3/3.07.Poster-Zhanqing%20Li.pdf*, *see Figure 4*.]

2.4 Multi-Scale 3DVAR Data Assimilation Scheme

An essential component of this system is the data assimilation (DA) scheme used to generate the nowcast estimates of the three-dimensional ocean state. Data assimilation is a mathematical methodology for optimally synthesizing different types of observations with model first guesses (that is, forecasts). A new two-step multi-scale (MS) three-dimensional variational (3DVAR) data assimilation algorithm is used here. This MS-3DVAR scheme is a generalization of the 3DVAR methodology of *Li et al.* [2008a, b] and is described in detail in *Li et al.* [2013, 2015a, 2015b]. The MS-3DVAR data assimilation methodology was selected because of its ability to propagate observational information, which is often sporadically and irregularly distributed, in both the horizontal and vertical directions through an advanced error covariance formulation, as well as its computational efficiency that enables real-time operational forecasting. The ROMS MS-3DVAR is designed to assimilate all types of observations simultaneously and reliably, while incorporating both the large-scale and small-scale impacts of the observations on the

model fields, a distinct advantage over single-scale 3DVAR systems [*Muscarella et al.,* 2014]. This advantage is realized through the use of background error covariances of multi-decorrelation length scales and by reducing inherent observational representativeness errors. In the implementation used here, the cost function set consists of two components for large and small scales, each using its own set of error covariances with different decorrelation length scales. The scheme is implemented sequentially from large to small scales. Of particular importance for applications such as this one, MS-3DVAR is effective in assimilating two of the most common types of ocean observations in the CA coastal region - sparse vertical profiles and high-resolution surface measurements - simultaneously.

A crucial component of 3DVAR-type assimilation schemes is the specification of the error covariances. For MS-3DVAR, the background error covariance matrices are constructed using the Kronecker product formulation given in *Li et al.* [2008a]. In this formulation, a three-dimensional error covariance matrix is generated using one- and/or two-dimensional error covariance matrices, which are estimated from perturbations derived from an ensemble of multiple long-time simulations of the model by removing the monthly means and ensemble means [*Li et al.,* 2008b]. To apply this method to MS-3DVAR, an additional partitioning of the perturbations into large and small components is performed. The large and small-scale background error covariances are then estimated from the corresponding perturbation fields.

2.5 Operational Implementation of the System

The ROMS nowcast-forecast system is run daily in near real-time. As mentioned above, this means we aim to deliver the nowcasts and daily forecast within nine hours after the time it is valid (that is, the 03 UTC nowcast is to be delivered by 12 UTC). The system incorporates all available real-time streams of data, gathered *in-situ* or remotely sensed (see Section 2.6 for a

283 complete list of the data assimilated). The system is executed following the procedures of
284 numerical weather prediction at operational meteorological centers. An assimilation step is
285 carried out every 6 hours. As mentioned above, a 6-hour window is used because MS-3DVAR
286 aims to describe not only sub-tidal variabilities but also tidal currents and other strong high
287 frequency variability. Specifically, the cycle can be summarized as follows. A 6-hour ROMS
288 forecast is performed, using the analysis valid at 03 UTC as an initial condition. Once the valid
289 09 UTC ROMS forecast is completed and all observations in the time-window between 03 and
290 09 UTC are collected, MS-3DVAR is executed and an analysis valid at 09 UTC is obtained. This
291 completes one cycle. The next cycle begins with a 6-hour ROMS forecast that uses the 06 UTC
292 analysis as its initial condition. The cycle is repeated four times a day and produces analyses
293 valid at 03, 09, 18 and 21 UTC. Once each day at 03 UTC, a 72-hour forecast is performed
294 using the analysis (nowcast) as the initial condition. Archived nowcasts are available since
295 January 2009. In addition to the real-time nowcasts and forecasts, a hindcast with data
296 assimilation (and including the addition of the assimilation of gridded altimetric satellite SSH
297 data and use of a global model output to generate the lateral boundary conditions) covering the
298 years 2009 to the present has recently been completed and provides the primary output used in
299 the analyses presented here.
300
301 ***2.6 Assimilated and Independent Observational Datasets***
302 The following observational data were routinely available in near real-time and were
303 assimilated by the ROMS MS-3DVAR system: 1) the gridded 0.25^{o} Archiving, Validation and
304 Interpretation of Satellite Oceanographic (AVISO) sea surface height data
305 [*ftp.aviso.oceanobs.com*] (in the upgraded version of the system); 2) gridded high frequency

306 (HF) radar surface current data [*http://hfrnet.ucsd.edu*], at 2 km and 6 km horizontal resolutions;
307 3) vertical profiles of temperature and salinity from three Spray gliders traversing California
308 Cooperative Oceanic Fisheries Investigations (CalCOFI) lines 67, 80 and 90, respectively, and
309 the Monterey Bay Aquarium Research Institute (MBARI) mooring (122.03°W, 36.75°N, location
310 shown by the red triangle in Figure 1;
311 *http://www.cencoos.org/data/buoys/mbari/m1tc_skyrocket*); 4) Advanced Very High Resolution
312 Radiometer (AVHRR), Moderate Resolution Infrared Spectroradiometer (MODIS) and
313 Geostationary Operational Environmental Satellite (GOES) satellite SST and ship SST
314 [*ftp://usgodae.org*]. Note that the AVISO sea surface height data from locations where the
315 ocean depth (as determined from the ROMS model bathymetry) is less than 1000 meters are not
316 used in the assimilation or the comparisons presented below, mostly because of the increase in
317 retrieval errors when approaching the coast in shallow waters.

318 The following observational datasets were not assimilated and are used here as independent
319 data for validation of the ROMS system performance: 1) a large set of vertical profiles of
320 temperature and salinity from ship surveys funded by the CalCOFI program
321 [*http://calcofi.org/data.html*] that were primarily in southern California; 2) a smaller set of
322 vertical profiles of temperature and salinity taken off the central California coast by EM-Apex
323 floats during an April 2015 field experiment; 3) vertical profiles of temperature collected
324 intermittently over more than a decade (from 2005 to 2015) in the nearshore coastal environment
325 of San Luis Obispo Bay located along the central coast of California. Temperature measurements
326 were obtained at the end of the Cal Poly Pier (120.76°W, 35.15°N, location shown by the black
327 triangle in Figure 1) using an automated profiling conductivity-temperature-depth (CTD) system;
328 4) surface velocities derived from a number of Lagrangian Surface Velocity Program (SVP)

329 drifters drogued at 15 m depth [*Niiler,* 2001; *Maximenko et al.,* 2013] obtained during the three-
330 year period 2010-2012 and compiled at the University of California, San Diego; 5) depth-
331 averaged velocities derived from the three Spray gliders [*Todd et al.* 2011]. Note that all
332 salinities are reported here in the practical salinity scale and thus have no units.
333
334 **3. Modeling System Performance Assessment**
335
336 ***a. Comparison of the ROMS Reanalysis with Assimilated Datasets −Surface***
337 We begin our assessment of the system performance with a comparison of the 6-hourly
338 analyses (nowcasts) produced with observational datasets that are assimilated. The primary
339 purpose of these comparisons is to provide a consistency check on the data assimilation
340 component of the system. The first category of data we compare is the satellite infrared (IR)
341 SST (AVHRR and MODIS combined). Figure 2a shows the RMS differences (blue solid line)
342 and spatial correlations (purple solid line) between the monthly mean observed IR SST and
343 ROMS SST during the entire 6-year hindcast period (2009-2014). Also shown as green bars is
344 the number of IR SST observations each month (right vertical scale). The RMS differences (blue
345 solid line) range from approximately 0.25^{o}C to 0.75^{o}C and show a distinct seasonal cycle that
346 mirrors the seasonal cycle in the number of observations. There is a clear minimum in the
347 number of observations (and a maximum in RMS differences) during the summer months. The
348 spatial correlations also show the effects of having fewer observations in summer as they dip
349 from values very close to 1.0 to around 0.9. Also shown in Figure 2a (as dashed lines) are the
350 RMS and spatial correlations for the GOES IR satellite data (which were included in the
351 assimilation beginning in 2011). Note that the GOES data is somewhat lower resolution (5 km)

than the MODIS and AVHRR data. The RMS and correlations for the GOES data show a similar evolution and seasonal cycle, though the RMS values occasionally become somewhat larger than those of the AVHRR/MODIS data during winter and spring. Figure 2b clearly displays the seasonal cycle by showing the 6-year mean seasonal cycle (for the AVHRR/MODIS data) in the RMS, spatial correlation and number of observations. The number of observations varies from approximately 0.5 million in July to between 1.5 and 2.5 million in the fall, winter and spring. The corresponding values for the RMS (and spatial correlation) range from approximately 0.75^{o}C (0.95) in July to 0.25-0.4^{o}C (0.99) during fall, winter and spring.

Two example spatial maps of the model minus observed SST differences, for March 2013 and July 2012, are shown in Figure 3 (note that only regions where the differences are larger than 0.25^{o}C are colored). In March 2013, the few regions with differences greater than 0.5^{o}C are more or less randomly distributed, while in July 2012 the regions with large differences are found primarily along the northern and central coasts in the near-coastal upwelling regions where strong gradients exist between colder upwelled waters near shore and warmer offshore waters (see Figure 1 for a September example of the typical summertime SST pattern). The upwelling regions are often covered with low-level stratus clouds during summer and thus few IR satellites SST observations are obtained in this region. Overall, given that standard measures of the accuracy of monthly means of AVHRR satellite SST data [*May et al.* 1998] against *in-situ* observations indicate RMS differences of approximately 0.6^{o}C, we consider the above model RMS values to be low enough to indicate that these data are being successfully assimilated by the system. A validation of the modeled daily mean SST (vs. observed satellite IR SST) using images similar to those in Figure 2a is available in near real-time [see *http://west.rssoffice.com/ca_roms_valid_other?variable=IRsst*]. The RMS differences for these

375 daily means typically range from 0.75 to 1°C. Note that typical RMS values for SST from
376 model runs without data assimilation (not shown) are at least twice as large as those presented
377 here from the run with data assimilation.

378 Another major assimilated observational dataset used for comparison is the gridded HF radar
379 surface current data. The RMS differences (ROMS vs. observed) are shown in Figure 4 for the
380 zonal and meridional components separately for the two different horizontal observational
381 resolutions (two and six kilometers). Note that the two kilometer data coverage is limited to
382 several nearshore (< 100 km) regions that are not contiguous, while the six kilometer data extend
383 further offshore (up to 200 km) and cover nearly the entire California coastal ocean. Figure 4
384 presents the monthly mean RMS differences for the period 2009-2014 between model and
385 observed surface currents, separately for the zonal (blue) and meridional (purple) components
386 and for the 6 km (Figure 4a) and 2 km (Figure 4b) data. Except for the first few months of 2009
387 when values are lower, the 6 km RMS values vary from 6 to 10 cm/s and the 2 km RMS values
388 from 5 to 9 cm/s. Unlike the IR SST RMS values, there is no discernible seasonal cycle in the
389 RMS differences for HF radar surface current data. The number of observations is also shown as
390 green bars. Note that there are significant increases in the number of observations towards the
391 end of 2009 and a more modest increase (in the 6 km data only) at the beginning of 2013.
392 Neither of the changes is accompanied by any meaningful changes in the RMS values.
393 Concerning the accuracy of HF radar surface current data, a summary of a number of studies
394 presented in *Emery et al.* [2004] suggests an envelope of RMS U, V differences between HF
395 radar currents and near-surface *in-situ* observed currents of 7–15 cm/s. Thus, we consider the
396 results presented in Figure 4 to indicate that the HF radar data are being successfully assimilated
397 by the system. A near real-time validation of the modeled daily mean surface currents (vs.

observed HF radar surface currents) using images similar to those in Figure 4 is available at [*http://west.rssoffice.com/ca_roms_valid_radar?variable=6*]. The RMS differences for these daily means typically range from 10 to 15 cm/s. Note that typical RMS values for surface currents from model runs without data assimilation (not shown) are about twice as large as those presented here from the run with data assimilation.

The final surface dataset we examine is the AVISO gridded 0.25° sea surface heights (SSH). Figure 5 presents the monthly mean RMS differences and spatial correlations for the period 2012-2014 between model and observed surface heights (solid lines). As noted above, we only assimilate and compare AVISO SSHs to ROMS output in regions where the ocean depth is greater than 1000 meters. The RMS values and correlations are remarkably stable throughout the period with values staying close to 2 cm and 0.98, respectively. As in the HF radar surface current comparison, we find no discernible seasonal cycle in the RMS differences or correlations for SSH. For this dataset the number of observations is not shown since these gridded fields are produced in such a way that there are no gaps and thus the number of observations does not change from month to month. Park *et al.* [2012] suggest errors in AVISO gridded SSH of approximately 7 cm when compared to values obtained from *in-situ* measurements by pressure-recording inverted echo sounders. Therefore, we consider the results presented in Figure 5 to indicate that the AVISO SSH data are being successfully assimilated by the system. Note that the typical RMS values for SSH from model runs without data assimilation (see the dashed lines in Figure 5, which show RMS and spatial correlations from a run that does not include the AVISO SSH data in the assimilation) are typically substantially larger, ranging from 6 to 10 cm, than those from runs with data assimilation.

b. Comparison of the ROMS Reanalysis with Assimilated Datasets – Subsurface

421 We next examine two of the primary subsurface datasets that are assimilated: vertical profiles
422 of temperature and salinity from the three Spray gliders and the MBARI M1 mooring. As an
423 example, Figure 6 shows a longitude-time section of the temperature along a transect by the
424 glider undertaken in July 2014 along line 80 (see purple line in Figure 1). The observed (Figure
425 6a) and co-located (in time and space) ROMS values (Figure 6b) as well as the difference
426 (Figure 6c) are shown. Aside from a slightly shallower upper ocean mixed layer, the ROMS
427 temperatures are very similar to the observed glider temperatures. Figure 6c shows that the
428 differences are greater than 0.5^{o}C only at some locations near the base of the mixed layer.
429 Overall, for this transect the mean temperature bias and RMS differences are quite small (-
430 0.05^{o}C and 0.44^{o}C). Figure 7 shows the corresponding comparison for the salinity fields. In this
431 case, the largest differences appear near the top of the halocline at the offshore end of the line
432 (Figure 7c), where the differences exceed 0.15. This reflects the fact that the top of the ROMS
433 halocline is located at a shallower depth than the observed. Overall, for this transect the mean
434 salinity bias is +0.005 and the RMS difference is 0.05.
435 A comparison of the ROMS profiles to the Spray glider observations for the entire six-year
436 period (2009-2014) is shown using the scatter diagrams in Figure 8. Note that the points on the
437 diagrams are color-coded according to their depth (see Figure 8 color bar). Overall, the
438 temperature and salinity bias are very small (0.01^{o}C and +0.0001) and the RMS values (0.51^{o}C
439 and 0.086) are comparable to those shown for the July 2014 transect. Note that the model-
440 observation differences tend to be larger near the surface (dark blue and purple points) where the
441 variability in both temperature and salinity is larger. The validation of vertical profiles of
442 temperature and salinity between the ROMS nowcasts against the Spray gliders is now being
443 done in near real-time with images similar to Figure 8 posted online daily

[*http://west.rssoffice.com/ca_roms_vaild_prof?variable=tscat*;

http://west.rssoffice.com/ca_roms_vaild_prof?variable=sscat].

For the M1 mooring, we find that the overall modeled versus observed statistics in vertical profiles of temperature and salinity (not shown) are very similar to those presented for the Spray gliders and so here we take advantage of the fixed location and relatively complete time series of observations to focus on the interannual variability. For further information on the observed ocean conditions at the M1 site, please refer to *Pennington and Chavez* [2017, this volume]. Figure 9 presents a time-depth section of the temperature anomalies (anomalies computed as deviations from the 2009-2014 six-year mean season cycle) as observed (Figure 9a) and the co-located ROMS values (Figure 9b). There are notable periods with cool anomalies in 2010 and 2012 as well as notable warm periods very early in 2010 and especially the latter half of 2014 when the anomalies approach 4ºC. For the most part the model realistically reproduces the timing and vertical extent of these events, including, for example, the deepening of the initially relatively shallow (< 50 m) warming during late 2014. On the other hand, the model overestimates the cool anomalies in late 2009 and underestimates a brief warming event at the end of 2012.

c. Validation of the ROMS Reanalysis with Independent Datasets

In this subsection, we validate the model nowcasts using independent data, that is, data that were not assimilated by the system. Similar to Figure 8, we present in Figure 10 scatter diagrams of ROMS versus observed CalCOFI ship CTD profiles of temperature and salinity. Note that most of these CalCOFI observations were taken off the southern California coast. The points on the diagrams are again color-coded according to their depth. It can be seen that while

467 the overall temperature bias is small (+0.03°C), there is a modest fresh bias in the salinity (-0.05).
468 In addition, the RMS values for both temperature and salinity are approximately twice as large as
469 the values for the assimilated profiles (Spray gliders and M1 mooring), with values of 0.92°C
470 and 0.17, respectively.

471 Though a much smaller dataset, the sampling region for the EM Apex float data shown in the
472 Figure 11 scatter diagram is further north off the central California coast. This comparison
473 shows again a small temperature bias (-0.06°C) and a modest fresh bias (-0.03). The RMS
474 differences are both somewhat smaller than in the CalCOFI comparison (0.72°C for the
475 temperature and 0.13 for the salinity) but still larger than the values for the assimilated profiles.
476 As for the comparisons with Spray glider profiles, in both the CalCOFI and EM Apex float
477 comparisons the model-observation differences tend to be larger near the surface than at depth.

478 For a nearshore location in central California, we use independent observations consisting of
479 vertical profiles of temperature that were collected at the end of the Cal Poly Pier (for location
480 see the black triangle in Figure 1). This comparison is particularly challenging since the ROMS
481 co-located values used are from the 3 km grid box closest to the model coastal boundary and are
482 thus representative of a much larger region of the ocean than the single point observations very
483 close to the coast. Therefore, we chose to compare the average seasonal signal at this location to
484 the model results. The CTD was connected to an automated winch station and recorded vertical
485 profiles of temperature to a depth of approximately 10 meters roughly every 30 minutes. During
486 each vertical profile, the CTD sampled at one second intervals and was lowered at a rate of
487 approximately 0.05 m/s. Temperature data for all of the downcast profiles were bin-averaged
488 into 0.5 m vertical bins. An average annual time series for the downcast temperature data was
489 calculated by averaging all of the data over the period 2009-2015 (to match the period when

490 ROMS output is available) within the same 1-day time window. The climatological mean
491 seasonal cycle obtained is compared to the ROMS seasonal cycle in Figure 12. As can be seen
492 in Figure 12, the seasonal cycle in temperatures at this location features a peak in the late
493 summer/early fall and a minimum during the spring, consistent with wind-driven coastal
494 upwelling. The vertical stratification is relatively weak except during late spring and summer as
495 temperatures warm first in the upper layers. During this time, temperature differences between
496 the top and bottom of the water column approach 5^{o}C.

497 The ROMS seasonal cycle in temperature is broadly similar showing a peak in late summer
498 and fall, minimum during spring, and vertical stratification that is relatively weak except during
499 late spring and summer. However, the ROMS seasonal cycle tends to be somewhat weaker than
500 the observed (summer temperatures tend to be a bit lower and winter temperatures a bit higher).
501 The summer vertical stratification is also weaker than observed. All these shortcomings may be
502 attributed (at least in part) to the relatively coarse spatial resolution of 3 km in ROMS that is
503 representative of a much larger region of the ocean than the single point observation very close
504 to the coast. Nonetheless, it is encouraging to see that the model is able to capture similar
505 seasonal variability at this shallow coastal embayment location.

506 We next compare depth-average currents derived from the Spray gliders with the co-located
507 ROMS depth-average current vectors (Figure 13). The depth range for the averaging is the
508 surface to the ocean bottom or 500 m, whichever is shallower. We present six-year seasonal
509 means of the data in order to focus on the large-scale annual cycle. With this in mind, we
510 examine the vectors (observed are red, ROMS are black) along the three Spray lines shown in
511 Figure 1. During all seasons the observed flow at the offshore ends of the lines tends to be
512 southward reflecting the broad equatorward California current (CC) that dominates the flow in

513 this region for most of the year. The southward currents are strongest during summer (JJA,
514 Figure 13c) when they reach speeds greater than 15 cm/s at some locations. These observations
515 are consistent with the findings of previous observational [*Strub and James* 2000] and modeling
516 [*Di Lorenzo* 2003; *Marchesiello et al.* 2003] studies that have suggested seasonal variations in
517 the California current are closely linked to those in surface winds with the strongest flows during
518 the warmer months. Closer to the coast, the observed depth averaged vectors predominantly
519 reflect the poleward California Undercurrent (CU) that peaks in strength between 100 and 300
520 meters below the surface [*Chelton* 1984; *Garfield et al.* 1999; *Pierce et al.* 2000]. Observations
521 indicate a semiannual seasonal cycle is the dominant temporal variability of the CU at most
522 locations, with a maximum in poleward flow in summer–fall and a secondary peak in winter
523 [*Hickey* 1979; *Chelton* 1984; *Lynn and Simpson* 1987; *Bray et al.* 1999]. In Figure 13, the
524 strongest poleward currents are found during JJA (Figure 13c) and DJF (Figure 13a) with
525 average currents of up to 15 cm/s. During DJF, there is also likely a substantial contribution to
526 the depth-average vectors from the surface seasonal poleward flow that occurs near the coast
527 during winter, known as the Inshore Countercurrent (ICC) or the Davidson Current [*Hickey*
528 1979; *Chelton* 1984; *Lynn and Simpson* 1987]. Generally speaking, the ROMS depth averaged
529 vectors (black) paint a similar picture of the CC and the CU and their seasonal variations. There
530 is some tendency for the ROMS currents to be stronger than the observed. Occasionally there
531 are significant differences between the two that require further investigation. One such
532 difference is seen in the DJF panel (Figure 13a) where the ROMS nearshore currents along Line
533 80 (extending offshore from Pt. Conception) are more westward than northward as observed.
534 Another difference appears during SON (Figure 13d) at a couple of places along Line 90 where

535 the ROMS currents are decidedly northward while the observed currents are weak or even
536 southward.
537
538 **4. Impact of Recent Upgrades to the System**
539
540 We next examine the impact of two recent upgrades to the modeling system: 1) use of
541 global HYCOM output to generate the lateral boundary conditions (LBC), instead of using
542 output from a separate data-assimilating ROMS system that covered the U.S. west coast and Baja
543 California at a horizontal resolution of 15 km, and 2) assimilation of the gridded 0.25^{o} AVISO
544 SSH data. Use of the HYCOM output for the LBC resulted in a substantial positive impact on
545 the seasonal and interannual variability in SSH, especially in the southern portion of the model
546 domain. This impact is demonstrated clearly in Figure 14, which shows the monthly mean SSH
547 anomalies for 2009-2014 at the La Jolla, CA tide gauge (117.26^{o}W, 32.87^{o}N) as observed (thick
548 solid line), from the ROMS run using HYCOM output for the lateral boundary (dashed line) and
549 from the ROMS run using the 15 km U. S. west coast ROMS (thin solid line). Note that the tide
550 gauge data is independent data in that it is not assimilated in either of the ROMS runs. The
551 observed values (thick solid line) show a clear seasonal cycle with peak values during the fall.
552 In addition, there are clear interannual variations superimposed on this seasonal cycle. For
553 example, the fall seasonal peak in 2010 is much smaller than that in 2009, and the 2014 peak is
554 the largest of all, consistent with the unusually warm conditions along the entire CA coast at that
555 time. The run using LBCs from the 15 km ROMS model (thin solid line) shows a much smaller
556 amplitude seasonal cycle and smaller interannual variations as well, having a standard deviation
557 of only 2.2 cm compared to the observed standard deviation of 7.1 cm. The run using LBCs
558 from HYCOM (dashed line) on the other hand, shows more realistic amplitude for both the
559 seasonal cycle and interannual variability. The standard deviation for this run is 5.8 cm. In

addition, the interannual variability shows broad agreement with the observations. For example, the fall peak in 2010 is smaller than that in 2009, and the 2014 peak is the largest.

To assess the impact of assimilating AVISO SSH data, we use the three years of SVP drifter velocities (a total of 45,543 of the observations in this dataset were inside our ROMS model domain). Note that this drifter velocity data is independent data in that it is not assimilated in any of the ROMS runs. Figure 15 summarizes the statistical comparison of these drifter velocities with the co-located ROMS surface velocities in a Taylor diagram. Zonal and meridional velocity statistics are shown separately by the circles and triangles, respectively. The red symbols show the values from the ROMS run without the assimilation of AVISO SSH data and the cyan symbols the values from the run with the assimilation of AVISO SSH data. In this diagram, it is clear that the ROMS run with the assimilation of AVISO SSH data produces surface velocities that compare better with these observed drifter velocities as the red symbols lie in the region of the diagram representing relatively low correlations (near zero) and high RMS values (20-25 cm/s), while the cyan symbols reside closer to the reference 'observed' point (purple square) on the x-axis in a region characterized by correlations near 0.5 and RMS values of about 15 cm/s. These two comparisons indicate the clear beneficial impacts of these two modeling system upgrades.

5. Summary and Discussions

A comprehensive validation of a three-dimensional, near real-time data-assimilative ocean modeling system for the California coastal ocean was presented. The system consists of a data-assimilating regional ocean model forced by the regional NAM operational weather forecast output. The ocean model has a horizontal resolution of approximately 3 km and utilizes a multi-scale 3DVAR data assimilation methodology. The system is run in near real-time to produce a

585 nowcast every six hours and 72-hour forecast every day. An analysis of the performance of the
586 system was presented using results from a seven-year period of 2009-2015. Model performance
587 was evaluated against assimilated and independent (non-assimilated) datasets.

588 The ROMS results were first compared with the assimilated data as a consistency check.
589 RMS differences in satellite infrared sea surface temperatures (SST) and vertical temperature
590 profiles between observations and ROMS nowcasts were found to be mostly less than 0.5^{o}C,
591 while the RMS differences in vertical salinity profiles between observations and ROMS
592 nowcasts were found to be 0.09 or less. The RMS differences in SST show a distinct seasonal
593 cycle that mirrors that in the number of observations available: the nowcast is less skillful with
594 larger RMS differences during the summer months when there are less infrared SST observations
595 due to the presence of low-level clouds. The larger differences during summer are found
596 primarily along the northern and central coasts in upwelling regions where strong gradients exist
597 between colder upwelled waters near shore and warmer offshore waters. RMS differences
598 between HF radar surface current observations and the ROMS nowcasts are approximately 7-8
599 cm s^{-1}, which is about 30% of the time mean current speeds in the region. The RMS differences
600 of SSH between the AVISO altimetric satellite observations and ROMS nowcasts are about 2
601 cm. The system is also found to realistically reproduce the interannual variability in temperature
602 at the M1 mooring in Monterey Bay, including the very strong warming of the California coastal
603 ocean during 2014 including, for example, the deepening of the initially relatively shallow (< 50
604 m) warming during late 2014.

605 The ROMS nowcasts were then validated against independent observations. A
606 comparison of the ROMS nowcast with independent profile observations of temperature and
607 salinity showed RMS differences of 0.7^{o}C to 0.92^{o}C and 0.13 to 0.17, about twice as large as

608 those in the comparisons with assimilated data. Validation versus depth-averaged currents
609 derived from Spray gliders showed that the flow patterns associated with the California Current
610 and California Undercurrent/Davidson Current systems and their seasonal variations are
611 qualitatively reproduced by the ROMS modeling system.

612 Lastly, the impact of several recent upgrades to the system was quantified. Switching
613 the lateral boundary conditions from a U.S. west coast regional model to the global HYCOM
614 model resulted in an improvement in the simulation of the seasonal and interannual variations in
615 the SSH, especially south of Pt. Conception. The assimilation of altimetric satellite SSH data
616 also resulted in an improvement in the model surface currents when compared to independent
617 surface drifter observations.

618 It may be illuminating to compare the results presented here with those of the MB system
619 described above and presented in *Chao et al.* [2009]. In making these comparisons, it is
620 important to keep in mind that *Chao et al*. [2009] compared their model results with a fairly
621 limited amount of observational data gathered within a relatively small area in and near
622 Monterey Bay during a single month-long (August 2003) summer field experiment focused on
623 the simulation of upwelling / downwelling events, while here we use a much larger set of
624 observations that cover a much larger portion of the CA coast and represent all seasons of the
625 year for six years. With this in mind, we note that the RMS differences in assimilated
626 temperature and salinity profiles presented in *Chao et al*. [2009] are comparable to those
627 obtained here (approximately 0.6^{o}C and 0.1). There was not enough independent temperature
628 and salinity data analyzed in *Chao et al*. [2009] to allow us to compare with the RMS differences
629 versus independent temperature and salinity data presented here. For the other variables, such as
630 SSH and velocity, the *Chao et al.* [2009] differences are similar to those obtained in runs without

any data assimilation as these variables were not assimilated at that time.

The ROMS nowcasts described in this paper are posted on the web in near real-time for broader distributions to the public. Two particular users worth mentioning here are the US Coast Guard (USCG) and NOAA Office of Response and Restoration who are using the ROMS nowcasts and forecasts on a routine basis to guide their decision-making in search and rescue and oil spill response operations, respectively. The ROMS nowcasts are also used as an input to a statistical model to detect toxic blooms and HABs [*Anderson et al.,* 2011]. While this paper demonstrates a significant skill in the ROMS nowcasts, the performance of the ROMS forecasts up to 72 hours into the future still remains to be quantified. Coupling the coastal ROMS nowcast/forecast system with an unstructured grid and higher resolution model so as to extend the modeling domain into the San Francisco Bay and Estuary, as well as a Carbon, Silicate, and Nitrogen Ecosystem (CoSiNE) biogeochemical/ecosystem model to enable ecological forecasting, are both currently underway.

Acknowledgements.

The research of Y. Chao and H. Zhang at UCLA JIFRESSE is supported by the NOAA Integrated Ocean Observing System (IOOS) through Southern California Coastal Ocean Observing System (SCCOOS) and Central and Northern California Ocean Observing System (CeNCOOS). The research of Y. Chao and J. Farrara at Remote Sensing Solutions is supported by the NOAA Integrated Ocean Observing System (IOOS) through Alaska Ocean Observing System (AOOS). Portions of this study were also supported by the NOAA ECOHAB program through award NA11NOS4780030 and by the Orange County Sanitation District. Support from the UCLA Joint Institute for Regional Earth System Science and Engineering (JIFRESSE) is gratefully acknowledged. We thank R. Davis for providing the Spray glider data and the CeNCOOS/SCCOOS HF radar team for providing the HF radar derived surface current data. The Cal Poly Pier measurements were supported by the NOAA IOOS program through CeNCOOS (Cal Poly, R. Walter and D. Wendt). Technical discussions and support from the former ROMS team members (Z. Li, X. Wang, P. Li, and Q. Vu) at Jet Propulsion Laboratory are acknowledged. Discussions and encouragement throughout this project period from C. Dong, J.C. McWilliams, and G. Robertson are also acknowledged. We recognize the State of California for their sponsorship of the HF radar network infrastructure under the Coastal Ocean Currents Monitoring Program (COCMP) and the National Oceanic and Atmospheric Administration (NOAA) Integrated Ocean Observing System (IOOS) program for continued operational support.

669
670 **References**
671

672 Amante, C. and B.W. Eakins (2009), ETOPO1 1 Arc-Minute Global Relief Model: Procedures,
673 Data Sources and Analysis. *NOAA Technical Memorandum NESDIS NGDC-24*. National
674 Geophysical Data Center, NOAA. doi:10.7289/V5C8276M.

675 Anderson, C. R., R. M. Kudela, C. Benitez, K. Nelson, E. Sekula, Wood, C.T. Burrell, Y. Chao,
676 G. Langlois, J. Goodman, and D.A. Siegel (2011), Detecting toxic diatom blooms from
677 ocean color and a regional ocean model, *Geophysical. Research Letters,* 38, L04603,
678 doi:10.1029/2010GL045858.

679 Bray, N. A., A. Keyes, W. M. L. Morawitz (1999), The California Current system in the Southern
680 California Bight and the Santa Barbara Channel. *J. Geophys. Res.*, **104**, 7695–7714,
681 doi:10.1029/1998JC900038.

682 Bond, N. A., M. F. Cronin, H. Freeland and N. Mantua (2015), Causes and impacts of the 2014
683 warm anomaly in the northeast Pacific, *Geophysical Research Letters,* doi:
684 10.1002/2015GL063306.

685 Capet, X., J. C. McWilliams, M. J. Molemaker and A. F. Shchepetkin (2008a), Mesoscale to
686 Submesoscale transition in the California current system. Part I: Flow structure, eddy flux,
687 and observational tests. *Journal of Physical Oceanography,* **38**, doi: 10.1175/2007JPO3671.1.

688 Capet, X., J. C. McWilliams, M. J. Molemaker and A. F. Shchepetkin (2008b), Mesoscale to
689 Submesoscale transition in the California current system. Part II: Frontal processes. *Journal of*
690 *Physical Oceanography,* **38**, doi: 10.1175/2007JPO3672.1.

691 Chao, Y., Li, Z., Farrara, J. D., McWilliams, J. C., Bellingham, J., Capet, X., Chavez, F., Choi,
692 J. K., Davis, R., Doyle, J. Frantaoni, D., Li, P. P., Marchesiello, P., Moline, M. A., Paduan,
693 J., Ramp, S. (2009), Development, implementation and evaluation of a data-assimilative
694 ocean forecasting system off the central California coast. *Deep-Sea Research II*, 56, 100-
695 126, doi:10.1016/j.dsr2.2008.08.011.

696 Chelton, D. B. (1984), Seasonal variability of alongshore geostrophic velocity off Central
697 California, *J. Geophys. Res.,* **89**, 3473–3486, doi:10.1029/JC089iC03p03473.

698 Chelton, D. B., R. A. deSzoeke, and M. G. Schlax (1998), Geographical variability of the first
699 baroclinic rossby radius of deformation. *J. Phys. Oceanogr.*, **28**, 433–460.

700 Di Lorenzo, E. (2003), Seasonal dynamics of the surface circulation in the Southern California
701 Current System. *Deep Sea Research Part II: Topical Studies in Oceanography* **50**, 2371–
702 2388, doi:10.1016/S0967-0645(03)00125-5.

703 Di Lorenzo, E., and N. Mantua (2016), Multi-year persistence of the 2014/15 North Pacific
704 marine heatwave, Nature Climate Change, PUBLISHED ONLINE: 11 JULY 2016 | DOI:
705 10.1038/NCLIMATE3082.

706 Doyle, J. D., Q. Jiang, Y. Chao, J. Farrara (2009), High-resolution real-time modeling of the
707 marine atmospheric boundary layer in support of the AOSNII field campaign. *Deep-Sea*
708 *Research II*, 56, 87-99, doi:10.1016/j.dsr2.2008.08.009.

709 Egbert, G. D., Erofeeva, S. Y. (2002), Efficient inverse modeling of barotropic ocean tides. *J.*
710 *Atmos. Oceanic. Technol.*, **19**, 183-204.

711 Egbert, G. D., Bennett, A. F., Foreman, M. G. G. (1994), TOPEX/POSEIDON tides estimated
712 using a global inverse model. *J. Geophys. Res.*, **99**, 24821-24852.

713 Emery, M., L. Washburn, and J. A. Harlan (2004), Evaluating radial current measurements from
714 CODAR high-frequency radars with moored current meters. *J. Atmos. Ocean Tech.*, **21**,
715 1259 – 1271.

716 Farrara, J. D., Y. Chao, Z. Li, X. Wang, X. Jin, H. Zhang, P. Li, Q. Vu, P. Q. Olsson, G. C.
717 Schoch, M. Halverson, M. A. Moline, J. C. McWilliams, F. A. Colas (2013), A data-
718 assimilative ocean forecasting system for the Prince William Sound and an evaluation of its
719 performance during Sound Predictions 2009. *Cont. Shelf Res.*, **63**, S193-S208,
720 doi:10.1016/j.csr:2012.11.008.

721 Garfield, N., C. A. Collins, R. G. Paquette, E. Carter (1999), Lagrangian exploration of the
722 California Undercurrent, 1992–95. *J. Phys. Oceanogr.*, **29**, 560–583, doi:10.1175/1520-
723 0485[1999]029<0560:LEOTCU>2.0.CO2.

724 Haidvogel, D. B., H. G. Arango, K. Hedstrom, A. Beckmann, P. Malanotte-Rizzoli, and A. F.
725 Shchepetkin (2000), Model evaluation experiments in the North Atlantic Basin:
726 Simulations in nonlinear terrain-following coordinates, *Dyn. Atmos. Oceans*, **32**, 239-281.

727 Hickey, B. M. (1979), The California Current System - Hypotheses and facts. *Prog. Oceanogr.*,

728 **8**, 191–279, doi:10.1016/0079-6611(79)90002-8.
729 Kondo, J. (1975), Air-sea bulk transfer coefficients in diabatic conditions. *Bound. Layer Meteor.*,
730 **9**, 91-112.
731 Large, W. G., and S. Pond (1982), Sensible and latent heat flux measurements over the ocean. *J.*
732 *Phys. Oceanogr.*, **12,** 464–482.
733 Li, Z., Chao, Y., McWilliams, J. C., Ide, K. (2008a), A Three-Dimensional Variational Data
734 Assimilation Scheme for the Regional Ocean Modeling System. *Journal of Atmospheric*
735 *and Oceanic Technology*, **25**, 2074-2090.
736 Li, Z., Chao, Y., McWilliams, J. C., Ide, K. (2008b), A three-dimensional variational data
737 assimilation scheme for the Regional Ocean Modeling System: Implementation and basic
738 experiments, *J. Geophys. Res.*, 113, C05002, doi:10.1029/2006JC004042.
739 Li, Z., Y. Chao, J. D. Farrara, J. C. McWilliams (2013), Impacts of distinct observations during
740 the 2009 Prince William Sound field experiment: A data assimilation study, *Cont. Shelf*
741 *Res.*, 63, Suppl. S209–S222, doi:10:1016/j.csr:2012.06.018.
742 Li, Z., J. C. McWilliams, K. Ide, J. D. Farrara (2015a), A Multiscale Variational Data
743 Assimilation Scheme: Formulation and Illustration. *Mon. Wea. Rev.*, **143**, 3804–3822. doi:
744 http://dx.doi.org/10.1175/MWR-D-14-00384.1.
745 Li, Z., J. C. McWilliams, K. Ide and J. D. Farrara (2015b), Coastal Ocean data assimilation using
746 a multi-scale three-dimensional variational scheme. *Ocean Dynamics*, **65**, 1001-1015. doi:
747 10.1007/s10236-015-0850-x.
748 Lynn, R. J., J. J. Simpson (1987), The California Current system: The seasonal variability of its
749 physical characteristics. *J. Geophys. Res.,* **92**, 12,947–12,966, doi:
750 10.1029/JC092iC12p12947.
751 Marchesiello, P., J. C. McWilliams and A. F. Shchepetkin (2001), Open boundary conditions for
752 long-term integration of regional ocean models, *Ocean Modelling*, **3**, 1-20.
753 Marchesiello, P., McWilliams, J. C., Shchepetkin, A. (2003), Equilibrium structure and
754 dynamics of the California Current System. *Journal of Physical Oceanography*, **33**, 753–
755 783, doi:10.1175/1520-0485.
756 Maximenko, N., R. Lumpkin and L. Centurioni (2013), Chapter 12 - Ocean Surface Circulation.
757 In G. Siedler, S.M. Griffies, J. Gould, and J.A. Church (Eds.), Ocean Circulation and
758 Climate, International Geophysics Series, Academic Press, Vol. 103, 283-304.

759 May, D.A., M. M. Parmenter et al. (1998), Operational processing of satellite sea surface
760 temperature retrievals at the Naval Oceanographic Office, *Bull. Amer. Meteor. Soc.,* **79**,
761 397-407.

762 McWilliams, J. C. (2016), Submesoscale currents in the ocean. Proc. R. Soc. A, **472**,
763 20160117.http://dx.doi.org/10.1098/rspa.2016.0117.

764 Muscarella, P. A., M.J. Carrier, H.E. Ngodock (2014), An examination of a multi-scale three-
765 dimensional variational data assimilation scheme in the Kuroshio Extension using the naval
766 coastal ocean model. *Cont. Shelf Res.,***73**, 41-48, doi:10.1016/j.csr:2013.11.009.

767 Niiler, P. P. (2001). The world ocean surface circulation. Ocean Circulation and Climate. G.
768 Siedler, J. Church and J. Gould, Academic Press. **77,** 193-204.

769 Park, J. H., D. R.Watts, K. A.Donohue, and K. L.Tracey (2012), Comparisons of sea surface
770 height variability observed by pressure-recording inverted echo sounders and satellite
771 altimetry in the Kuroshio Extension, *J. Oceanogr.,* **68**, 401–416, doi:10.1007/s10872-012-
772 0108-x.

773 Pennington and Chavez (2017), Decade-scale oceanographic fluctuation in Monterey Bay,
774 California, 1989-2011, this volume.

775 Pierce, S. D., R. L. Smith, P. M. Kosro, J. A. Barth, C. D. Wilson (2000), Continuity of the
776 Poleward Undercurrent along the eastern boundary of the mid-latitude North Pacific, *Deep*
777 *Sea Res.*, Part II, **47**, 811–829, doi:10.1016/S0967-0645(99)00128-9.

778 Shchepetkin, A. F., and J. C. McWilliams (2005), The Regional Ocean Modeling System: A split-
779 explicit, free-surface, topography-following-coordinate ocean model, *Ocean Modeling*, **9**,
780 347-404.

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

785 Sikirić, M.D., Janeković, I. and Kuzmić, M. (2009), A new approach to bathymetry smoothing in
786 sigma-coordinate ocean models. *Ocean Modelling*, doi:10.1016/j.ocemod.2009.03.009.

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

789 Strub, P. T., James, C. (2000), Altimeter-derived variability of surface velocities in the California
790 Current System: 2. Seasonal circulation and eddy statistics. *Deep Sea Research Part II*
791 *Topical Studies in Oceanography,* **47**, 831–870, doi:10.1016/S0967-0645(99)00129-0.

792 Todd, R. E., D. L. Rudnick, M. R. Mazloff, R. E. Davis and B. D. Cornuelle (2011), Poleward
793 flows in the southern California Current System: Glider observations and numerical
794 simulation. *J. Geophys. Res*., **116**, C02026, doi:10.1029/2010JC006536.

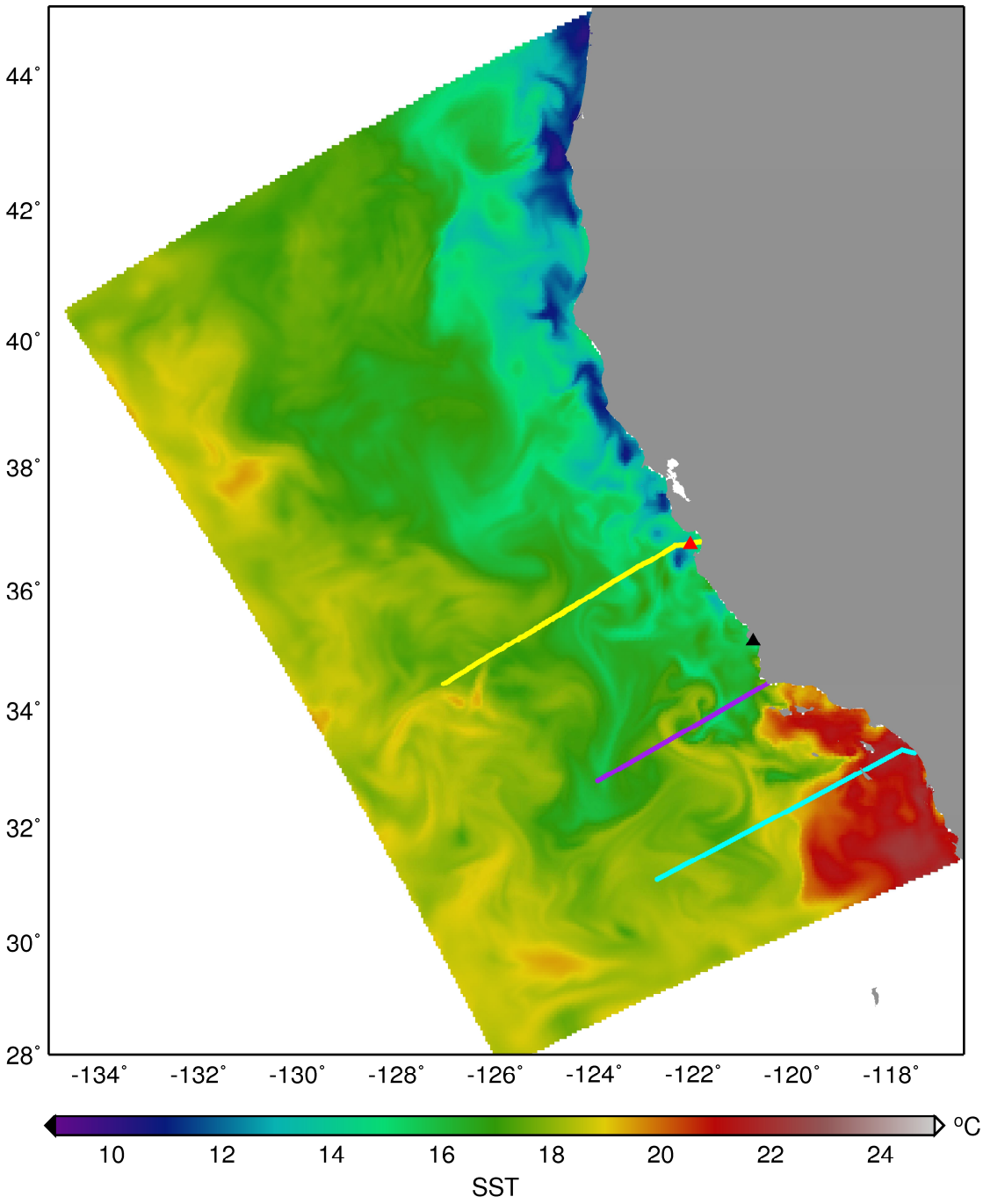
795 Wang, X., Y. Chao, C. Dong, J. Farrara, Z. Li, J. C. McWilliams, J. D. Paduan, L. K. Rosenfeld
796 (2009), Modeling tides in Monterey Bay, California. *Deep-Sea Research II*, 56, 219-231,
797 doi:10.1016/j.dsr2.2008.08.012.

798 Whitney, F. A. (2015), Anomalous winter winds decrease 2014 transition zone productivity in
799 the NE Pacific. *Geophys. Res. Lett.* **42**, 428–431.

800 Zaba, K. and D. L. Rudnick (2016), The 2014-2015 warming anomaly in the Southern California
801 current system observed by underwater gliders, *Geophysical Research Letters,* doi:
802 10.1002/2015GL067550.

803

804



805
806
807 *Figure 1. Daily mean ROMS sea surface temperatures for 28 September 2012 on the ROMS CA-*
808 *3 km model domain (colored region). The yellow, purple and cyan lines show the locations of*
809 *the SIO glider lines 67, 80 and 90, respectively and the red and black triangles the location of*
810 *the M1 mooring and the Cal Poly San Luis Obispo Pier, respectively.*

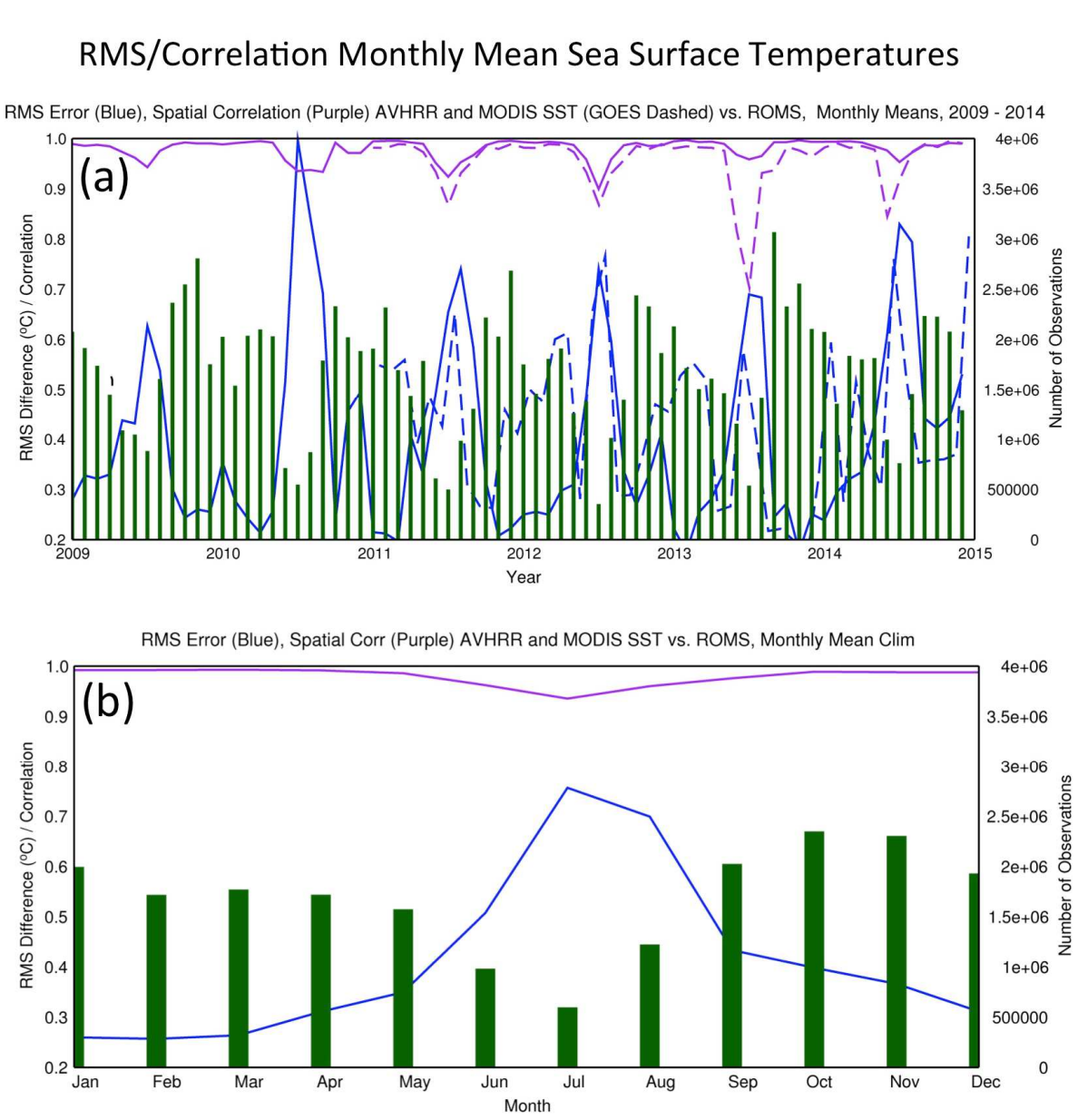


Figure 2. (a) Root-mean square (RMS) differences (°C, blue) and spatial correlations (purple) between monthly mean sea surface temperatures observed by AVHRR and MODIS satellites and the ROMS nowcast analyses. The RMS differences and correlations for the GOES IR satellite data are shown as dashed lines. The green bars indicate the number of observations available each month (scale on right). (b) As above, except for the six-year climatological monthly time series for AVHRR plus MODIS versus ROMS.

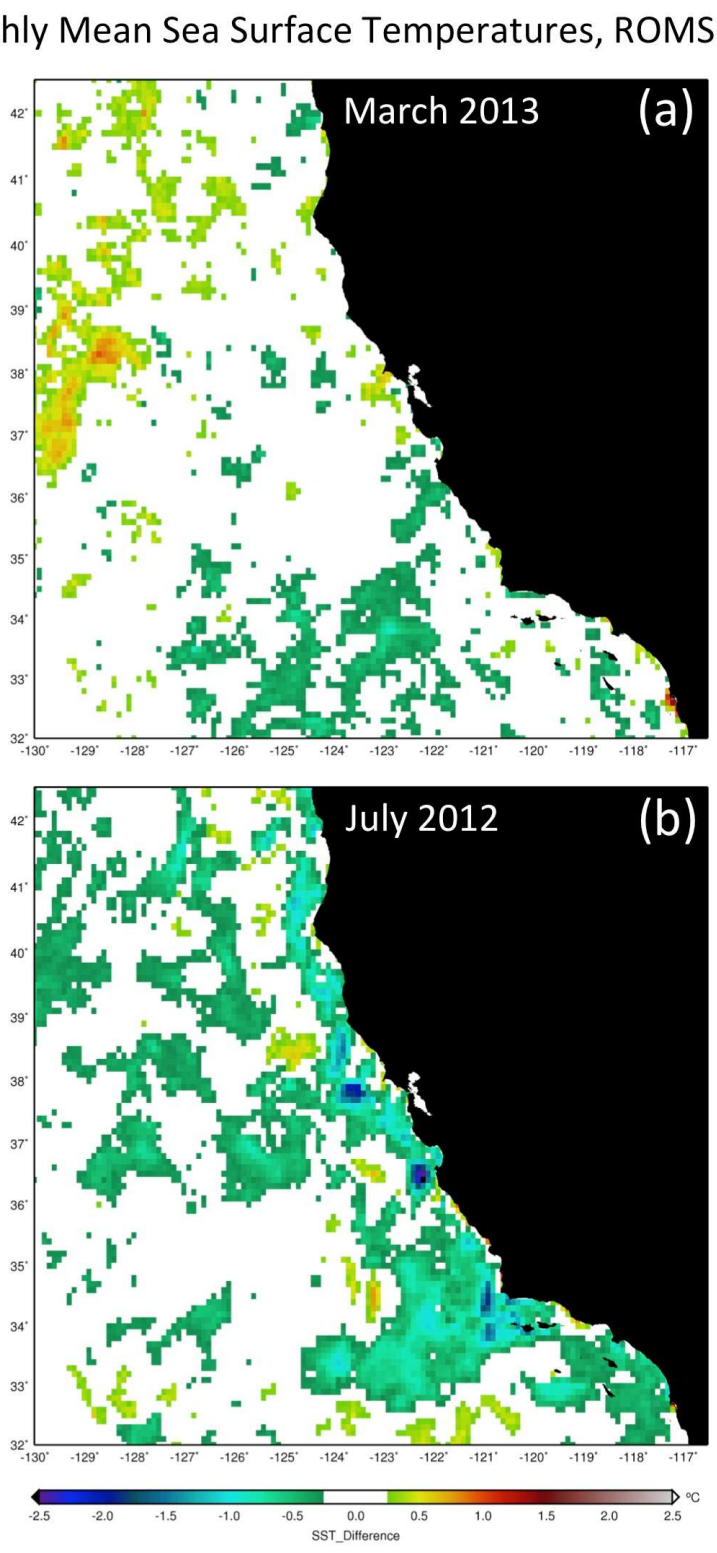


Figure 3. Monthly mean SST differences (°C, ROMS minus observed MODIS) for March 2013 (a) and July 2012 (b). Note that only regions with differences greater than 0.25°C are colored.

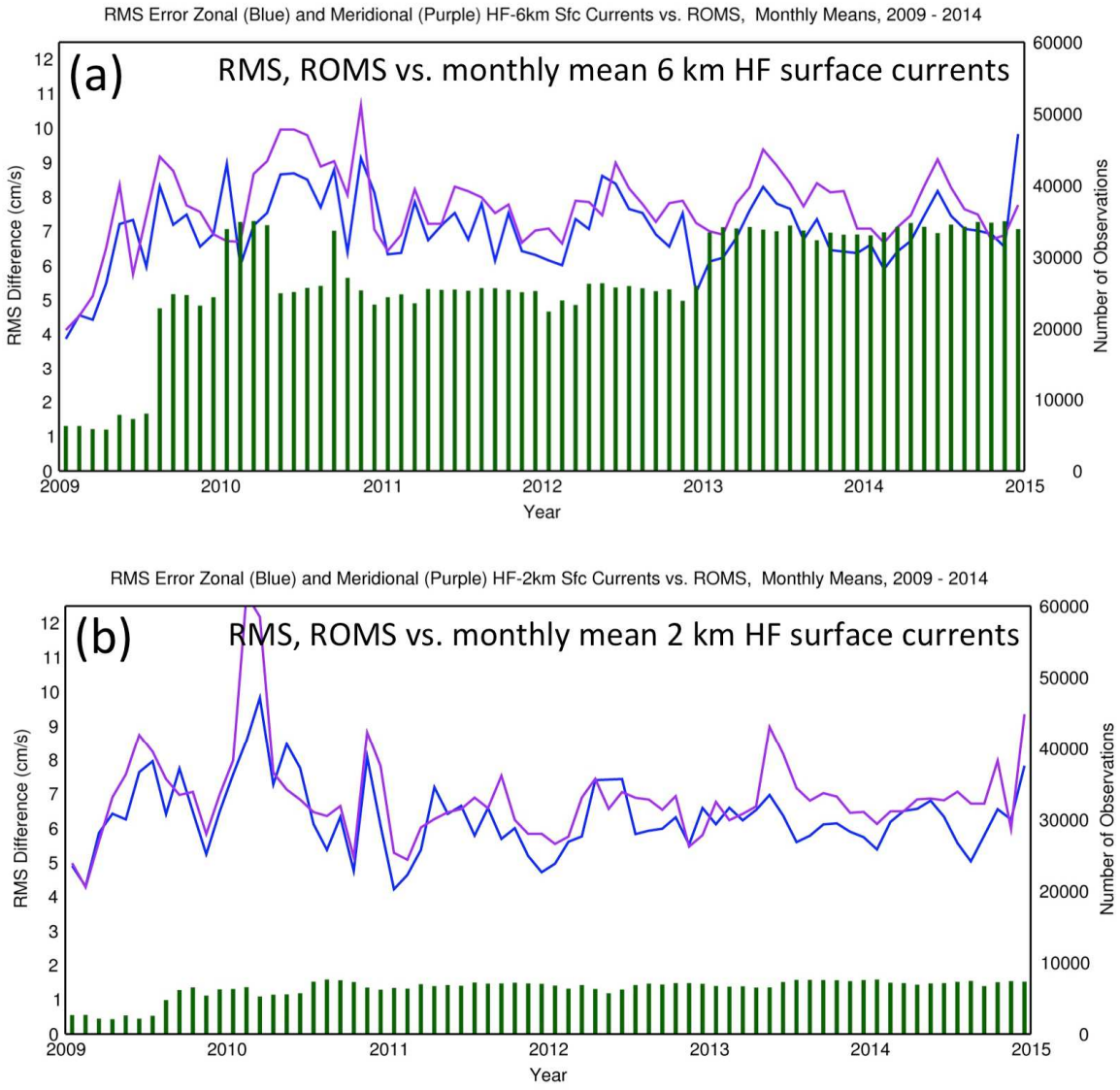


Figure 4. (a) Root-mean square (RMS) differences in the monthly-mean zonal (blue) and meridional (purple) components of the surface velocity (cm/s) between the 6 km HF radar observations and co-located ROMS nowcast analyses. The green bars indicate the number of observations available each month (scale on right). (b) As in (a), except for the 2 km HF radar observations versus ROMS.

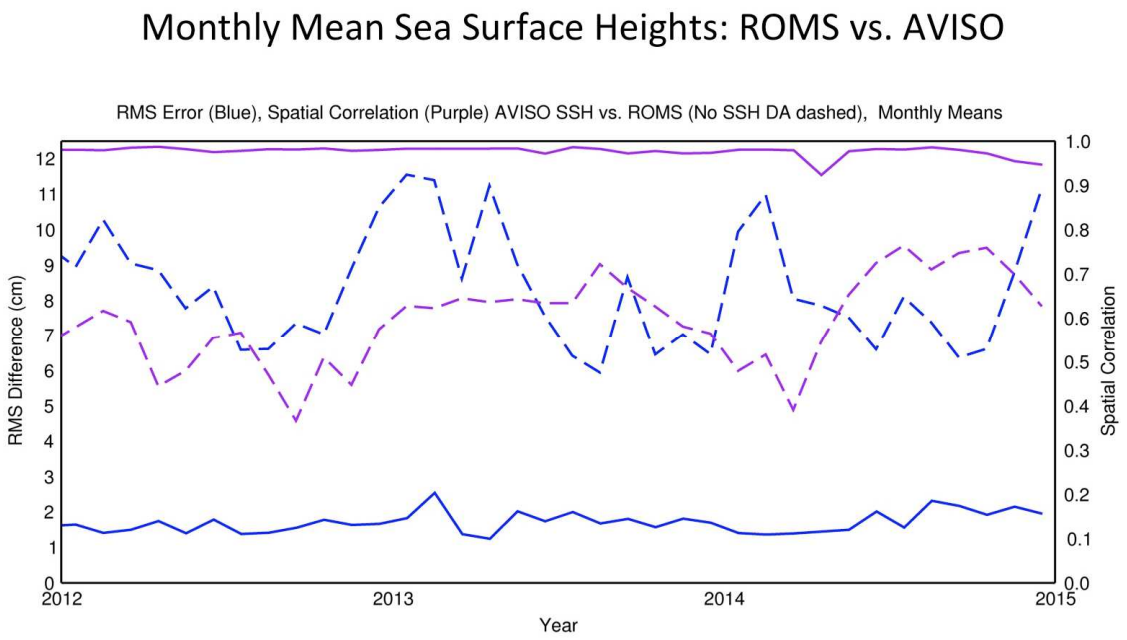


Figure 5. Root-mean square (RMS) differences (cm, blue solid) and spatial correlations (purple solid) between monthly mean AVISO sea surface heights and the ROMS nowcast analyses (the number of observations is not shown as it does not vary). The dashed blue (purple) line depicts the corresponding RMS differences (correlations) obtained from a ROMS run that does not include the assimilation of the AVISO SSH data.

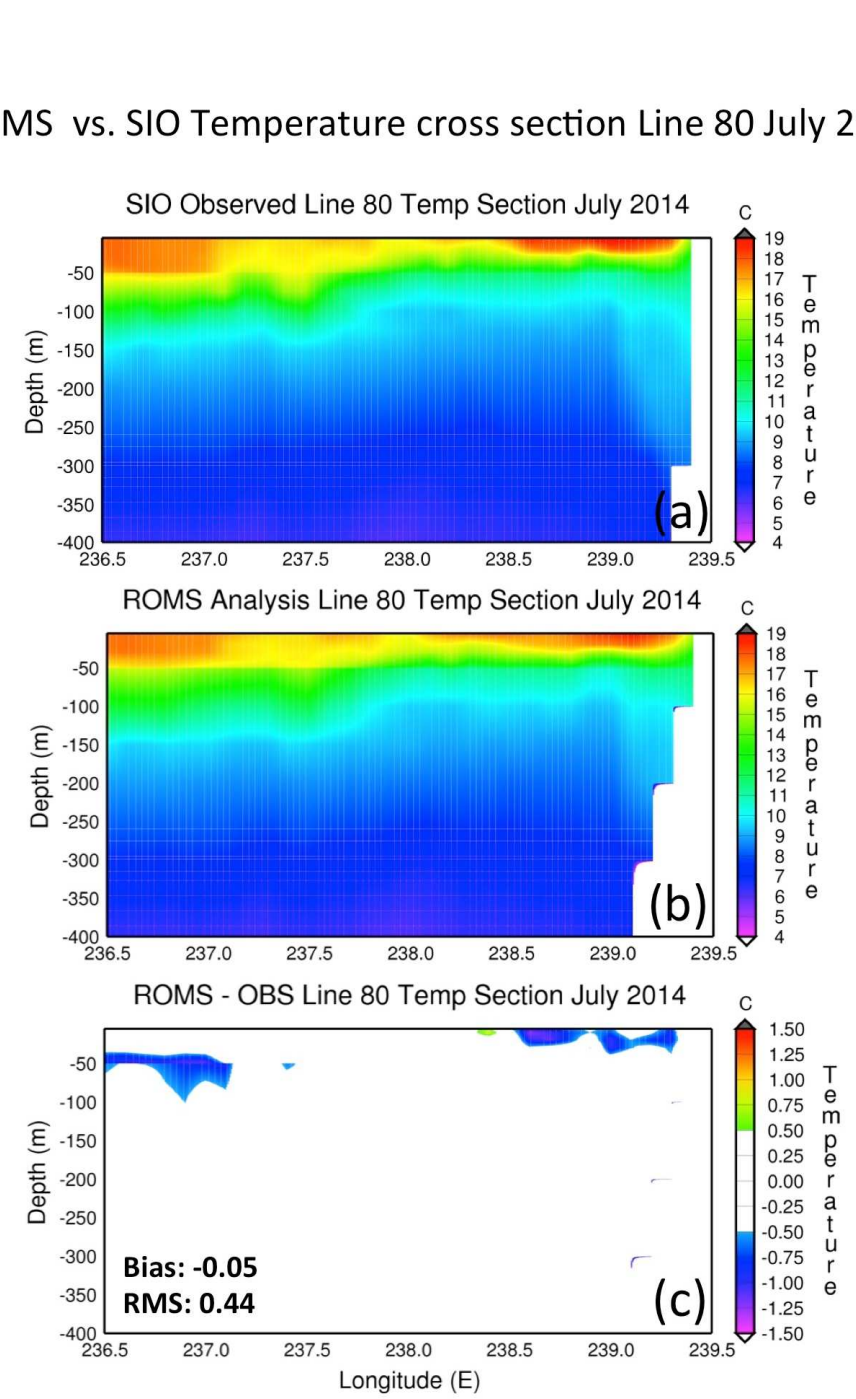


Figure 6. Longitude-depth section of the temperatures (°C) for a transect during July 2014 along Spray glider line 80 as observed (a), from the ROMS modeling system (b) and (c) the difference between the two (ROMS minus observed). Note that the location of line 80 is shown by the purple line on Figure 1.

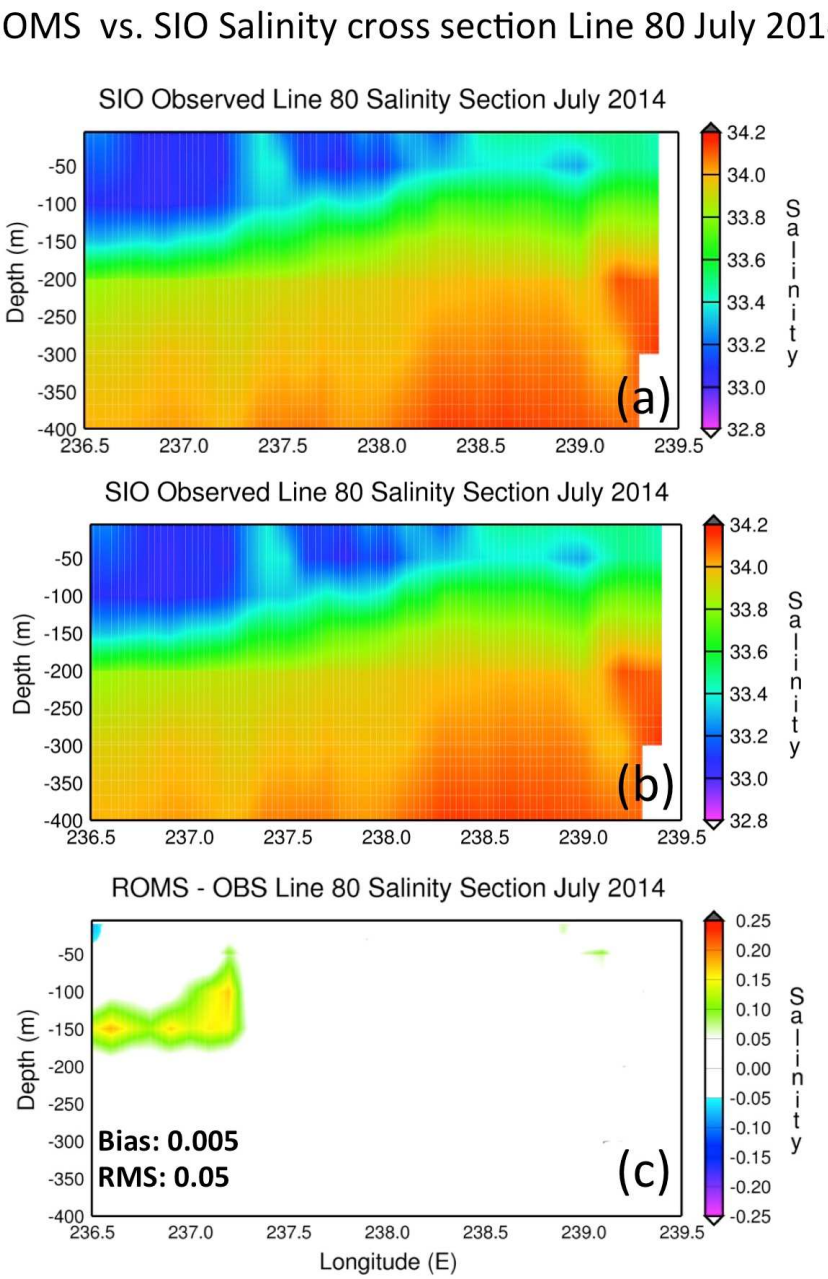


Figure 7. Longitude-depth section of the salinities for a transect during July 2014 along Spray glider line 80 as observed (a), from the ROMS modeling system (b) and (c) the difference between the two (ROMS minus observed). Note that the location of line 80 is shown by the purple line on Figure 1.

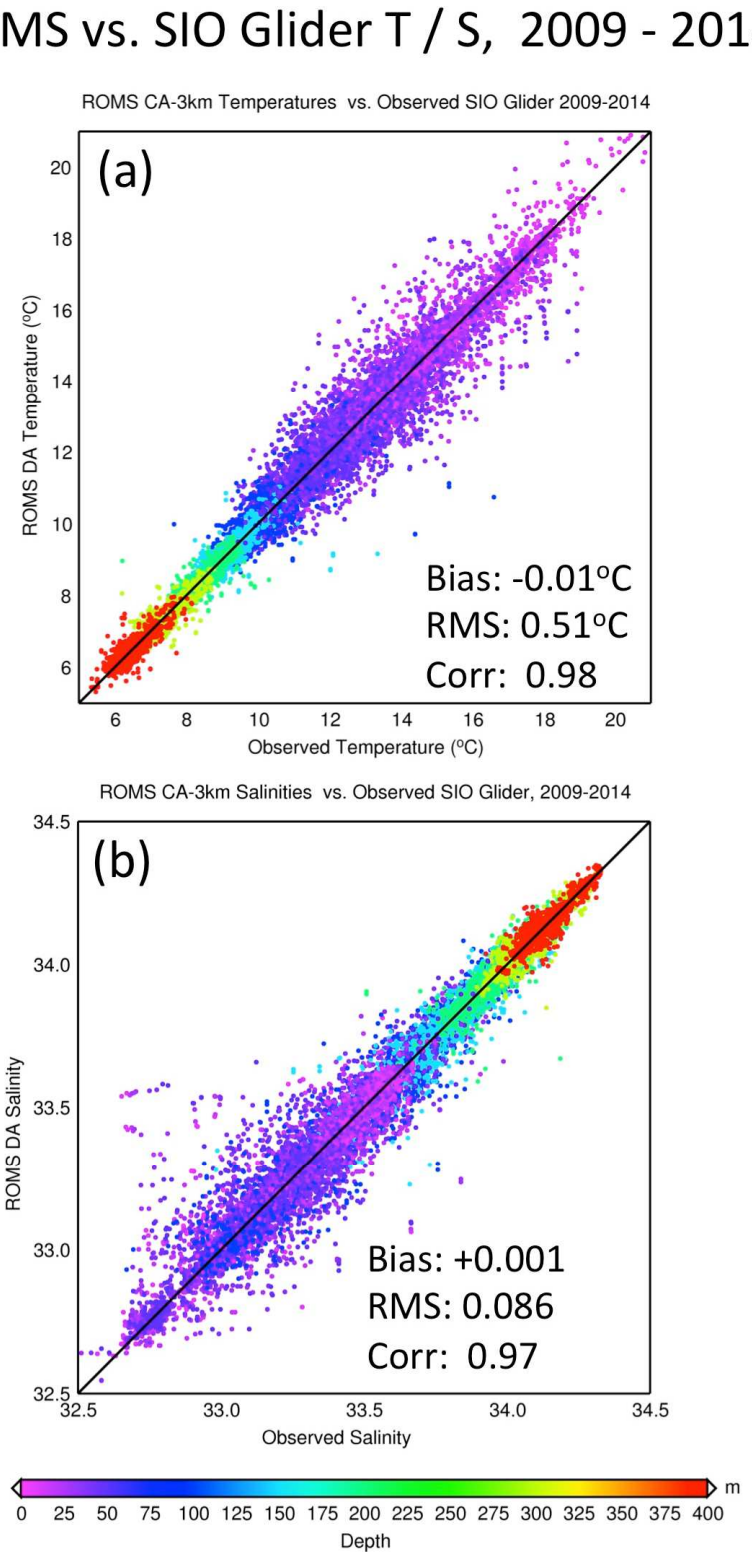


Figure 8. Scatter diagram of ROMS versus Spray glider temperatures (a, °C) and salinities (b).

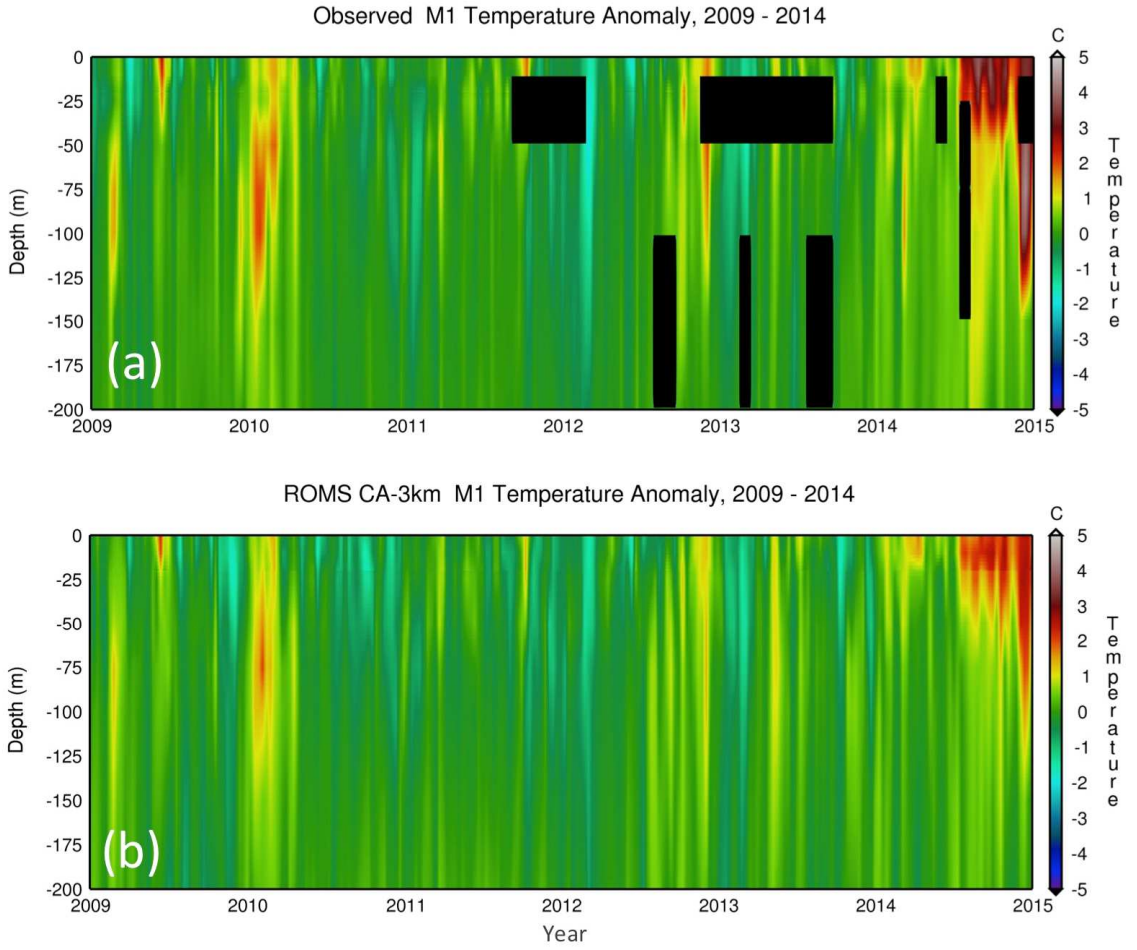


Figure 9. Time-depth section of M1 temperature anomalies (°C) as observed (a) and from the ROMS nowcasts (b). Anomalies are computed as deviations from the 2009-2014 six-year mean season cycle. Note that the location of the M1 mooring is shown by the red triangle on Figure 1.

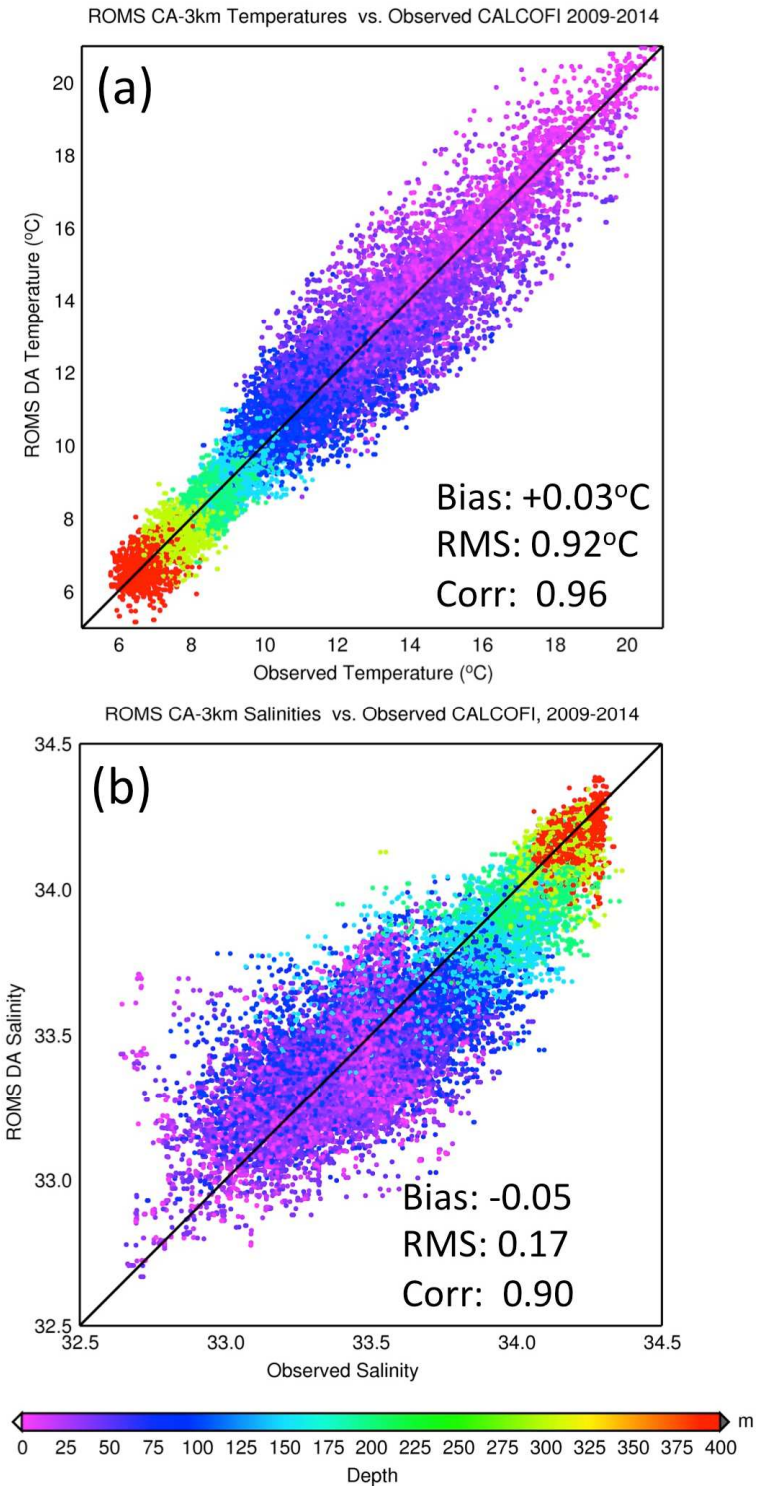


Figure 10. Scatter diagram of ROMS versus CalCOFI ship CTD temperatures (a, °C) and salinities (b).

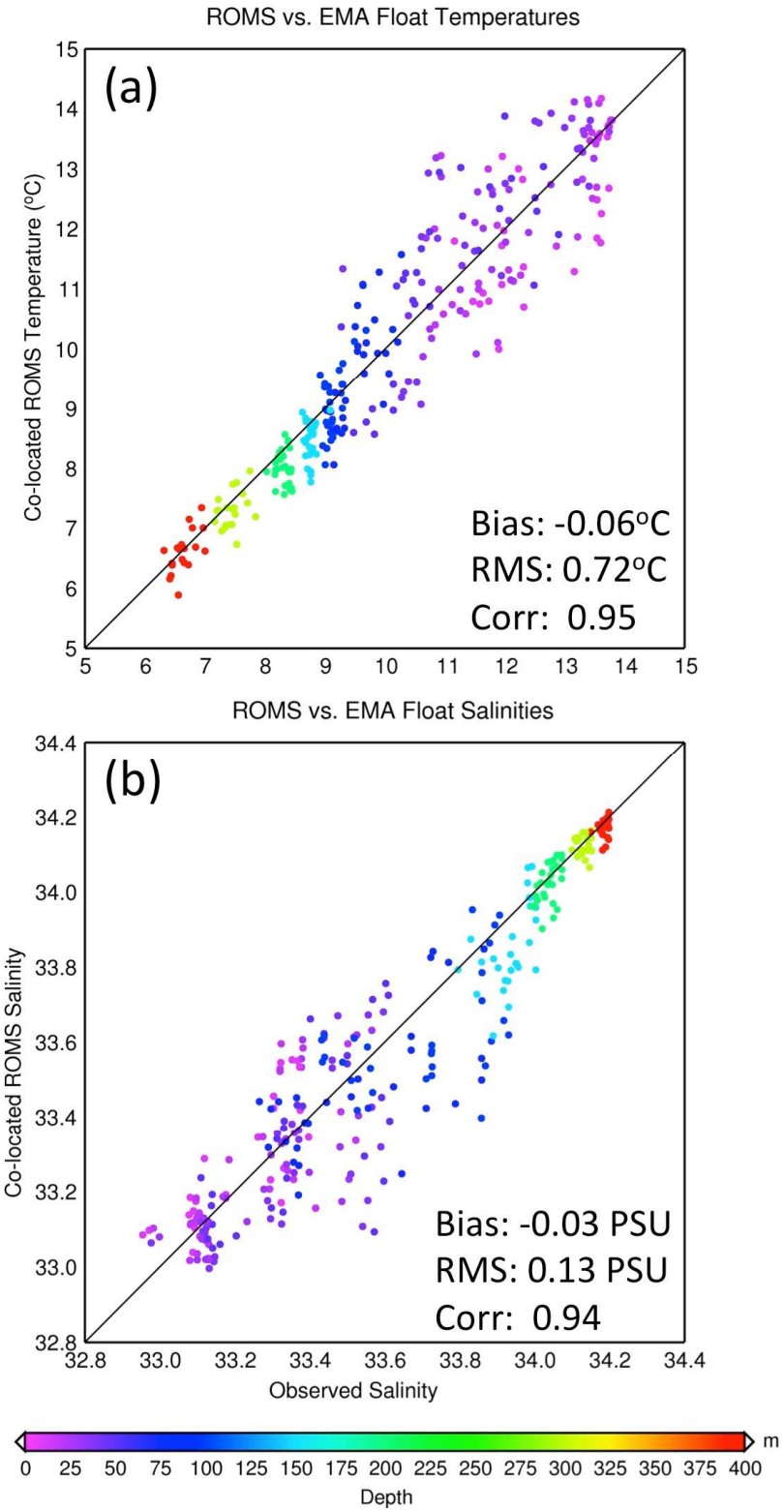
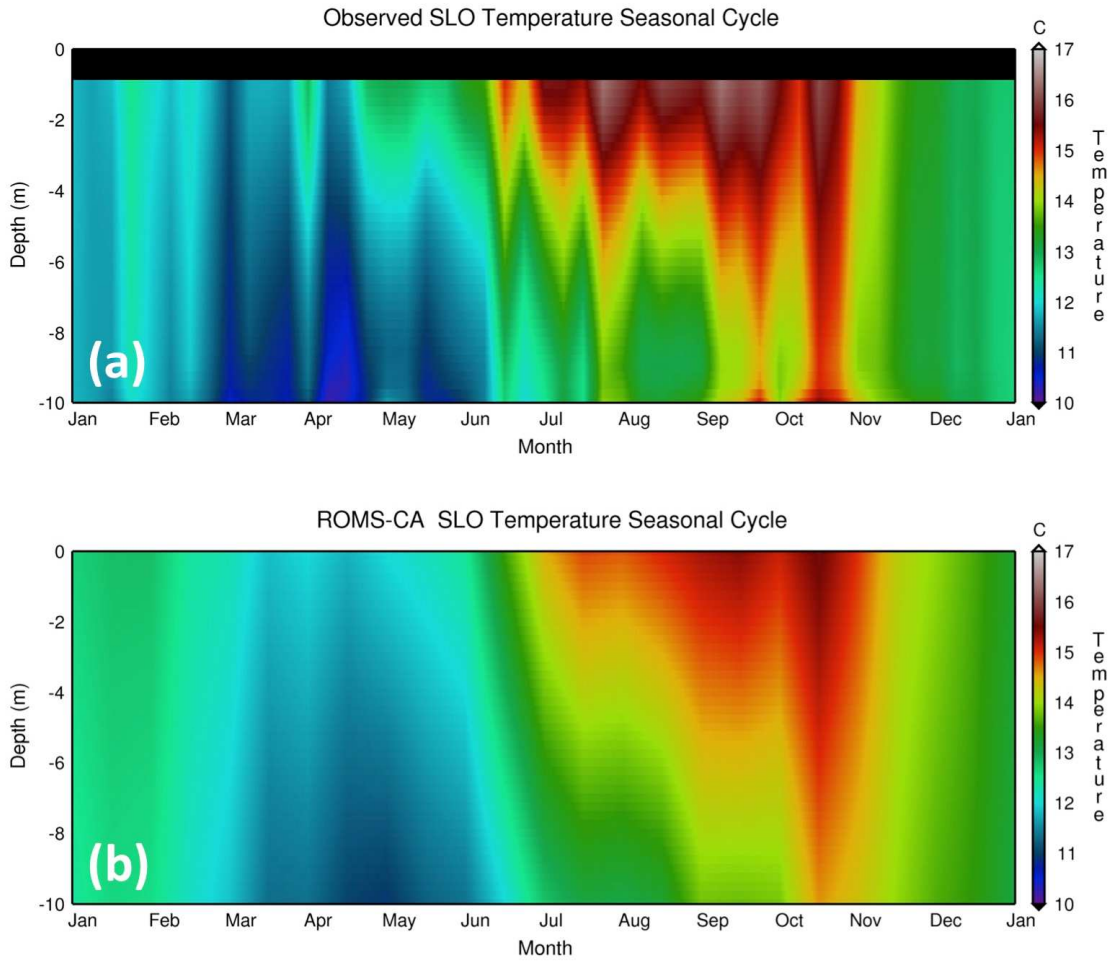


Figure 11. Scatter diagram of ROMS versus EMA float temperatures (a, °C) and salinities (b).



875
876
877
878 *Figure 12. Time-depth section of the climatological seasonal cycle of temperature (°C) at the Cal*
879 *Poly San Luis Obispo Pier as observed (a) and (b) co-located ROMS climatological*
880 *temperatures.*
881

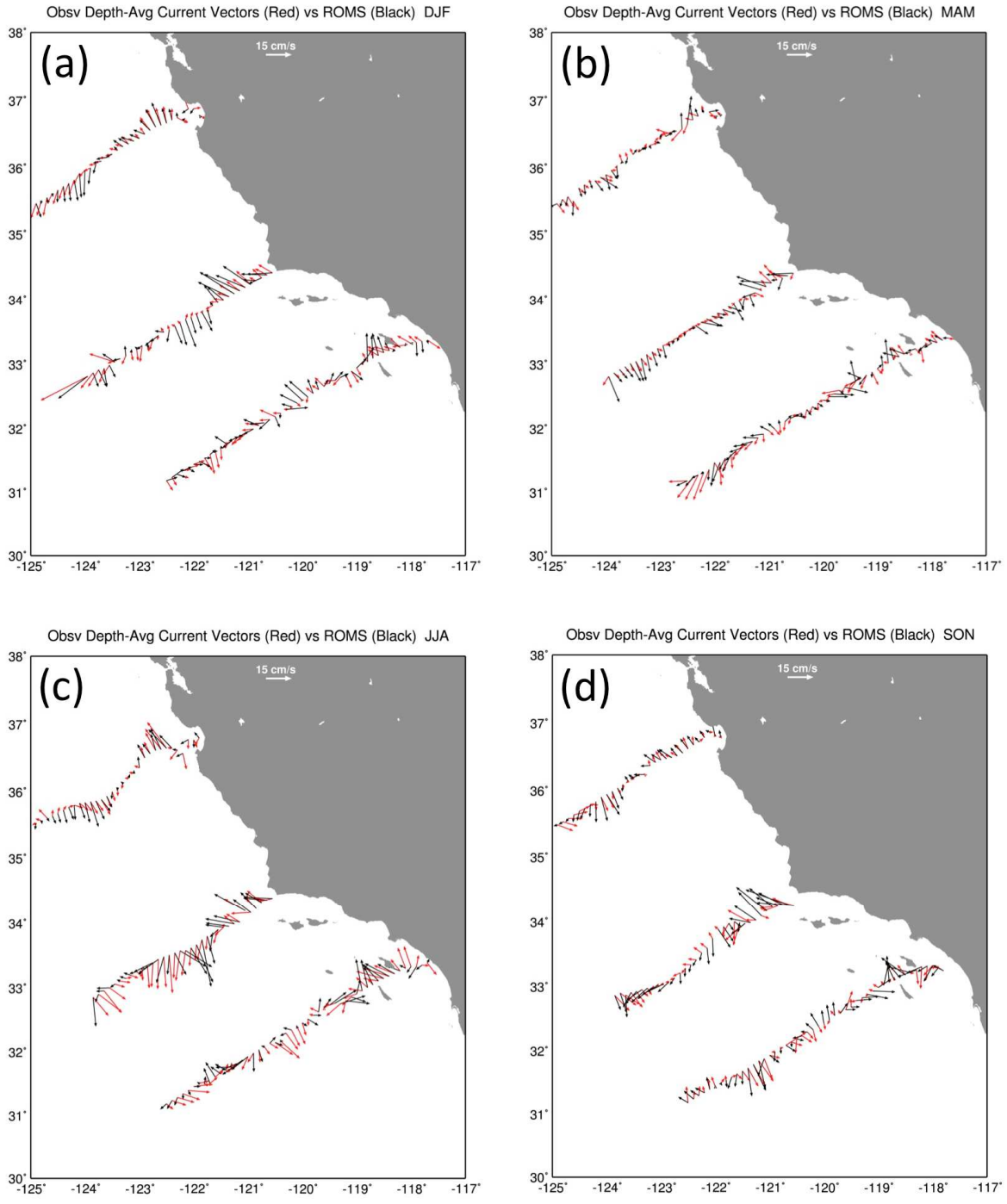


Figure 13. Seasonal mean maps of ROMS (black) versus Spray glider derived (red) depth-averaged (the depth varies with location, please see text for details) velocities (cm/s) for (a) DJF, (b) MAM, (c) JJA and (d) SON.

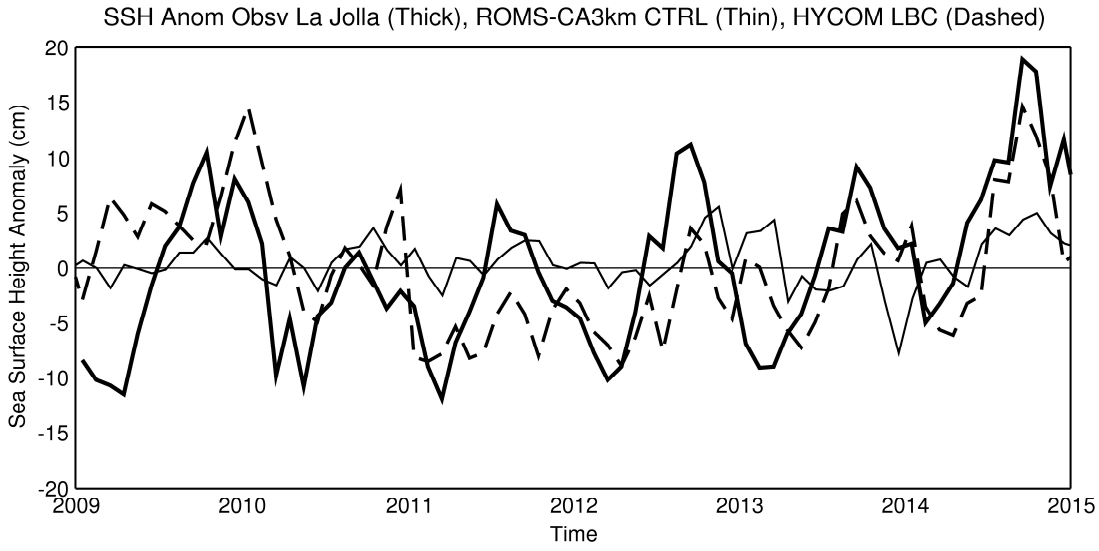


Figure 14. Monthly mean SSH anomalies (cm) as observed at the La Jolla, CA tide gauge station (thick solid line) and the co-located ROMS anomalies from a run using the global HYCOM model output for the lateral boundary conditions (dashed line) and a run using output from a separate data-assimilating ROMS system that covers the entire U.S. west coast at a resolution of 15 km (thin solid line).

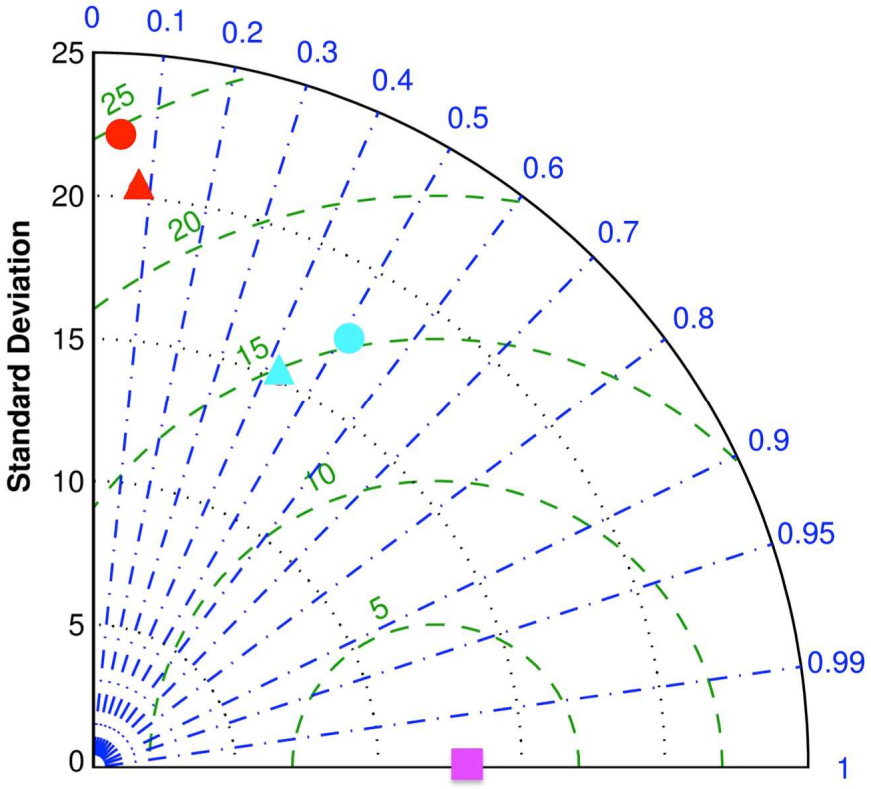


Figure 15. Taylor diagram comparing observed SVP drifter velocities (cm/s) to co-located ROMS surface velocities. The blue lines and labels are for the correlation coefficients and the green dashed lines and labels for the RMS differences (cm/s). Zonal (triangles) and meridional (circles) velocities are shown separately. The red symbols show the values from a run of the ROMS-CA without the assimilation of AVISO SSH data and the cyan symbols the values from a run with the assimilation of AVISO SSH data. The purple square shows the reference (observational) value.