An evaluation of high-resolution ocean reanalyses in the California Current System

Dillon J. Amaya[1], Michael A. Alexander[1], James D. Scott[1,3], and Michael G. Jacox[1,2]

[1]Physical Science Laboratory, Earth System Research Laboratory, National Oceanic and Atmospheric Administration

[2]Environmental Research Division, Southwest Fisheries Science Center, National Oceanic and Atmospheric Administration

[3]Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder

Corresponding author: Dillon J. Amaya, dillon.amaya@noaa.gov, 816-916-8348

Abstract

Sparse and inconsistent coverage of ocean observations makes analysis of climate impacts on ocean physics and marine ecosystems challenging. As a result, ocean reanalyses (i.e., ocean models constrained by observations through data assimilation) were developed to provide historical ocean state estimates that are spatially and temporally uniform. Recent advances in high performance computing and the number and quality of observations have led to the development of high-resolution ocean reanalyses, which offer an opportunity to investigate coastal ocean variability with enhanced fidelity. In this study, we evaluate the ability of three high-resolution ocean reanalyses, including the Global Ocean Reanalysis and Simulations (GLORYS), the Ocean Reanalysis System version 5 (ORAS5), and the California Current System Reanalysis (CCSRA), to accurately represent ocean temperature and salinity (from the surface to the bottom), sea surface height, and mesoscale activity in the California Current Large Marine Ecosystem (CCLME). Specifically, we compare these reanalyses to a variety of assimilated and independent *in situ* and satellite derived observations along the U.S. west coast. We find that the reanalyses generally reproduce large-scale variability in temperature and sea surface height within the CCLME, including effects of major ENSO events and recent marine heatwaves. We also show that GLORYS and CCSRA, with their finer horizontal resolution, have enhanced fidelity in simulating nearshore ocean parameters such as coastal sea level and bottom temperature along the continental shelf. Our results suggest that these tools can be used to study the fine-scale features of the California Current System over the past several decades.

1. Introduction

The spatial inhomogeneity of global ocean observations in the historical record makes rigorous analyses of long-term ocean climate variability and change challenging. Ocean model simulations generate continuous data in both time and space, ideal for ocean climate studies, but are also affected by biases and errors that can lead to an unrealistic representation of key physical processes. To overcome the respective limitations of observations and model simulations, while harnessing the strengths of each, modelling centers blend dynamical models with atmospheric and oceanic observations through data assimilation techniques. These efforts have resulted in observationally constrained estimates of the climate state that: (1) Combine the full spatiotemporal coverage of models with the accuracy of observations, and (2) Are easily accessible for scientific and industrial applications (Balmaseda et al., 2015; Storto et al., 2019).

Despite these advantages, there can still be large uncertainties in ocean reanalysis products. In particular, some parts of the global open ocean are historically under-observed (e.g., the Southern Ocean), the ocean subsurface is much less constrained than the satellite-observable surface, and coastal regions are often dominated by mesoscale ocean features that may not be properly resolved by the relatively coarse resolution of many global ocean reanalyses (Balmaseda et al., 2015; de Souza et al., 2021; Lee et al., 2009). As a result, reanalysis estimates in these regions may be increasingly dependent on the underlying ocean model configuration and its parameterization of subgrid scale processes, leading to large differences between ocean reanalyses from different modeling centers. (Balmaseda et al., 2015; Masina et al., 2011; Storto et al., 2019; Xue et al., 2017, 2012; Zhu et al., 2012).

Recent advances in ocean model physics (Breivik et al., 2015), the number and quality of observations (Legeais et al., 2018), and data assimilation techniques (Sakov et al., 2012; Storto et al., 2018) have led to the development of several high-resolution global and regional ocean reanalyses. With horizontal resolution as fine as 8 km, these new tools offer a unique opportunity to study ocean variability and change, as well as their impacts on marine species distributions and populations, with enhanced fidelity. Given the expanding use of ocean reanalyses for regional studies of oceanography and ecology, it is important to verify the accuracy of these high-resolution state estimates against independent (i.e., unassimilated) *in situ* ocean observations where available (de Souza et al., 2021; Xie et al., 2008). Indeed, increasing model resolution presents its own set of challenges, including (among others) potential errors in simulating mesoscale eddies and sharp gradients in ocean properties, coupling tides and waves, downscaling of atmospheric forcing, and the accuracy of bathymetry (e.g., (Storto et al., 2019). For example, (de Souza et al., 2021) compared several high-resolution ocean reanalyses to a variety of observations in New Zealand coastal waters and showed that some of the datasets inaccurately represented important coastal boundary currents.

Another region that would also benefit from a thorough intercomparison of high-resolution ocean reanalyses is the California Current Large Marine Ecosystem (CCLME; Figure 1). The CCLME is home to a highly productive marine ecosystem with primary productivity and fish catch disproportionately high for its spatial extent (Chavez and Messié, 2009). This elevated productivity can be attributed to upwelling of nutrient rich waters along the U.S. west coast, which is driven by the seasonal intensification of northerly winds in the spring and summer. These seasonal wind changes and the associated impact on ocean properties are sensitive to climate variability on time scales of weeks to decades (Checkley and Barth, 2009). In particular, upwelling in the CCLME is strongly modulated by large-scale climate modes such as the El Niño-Southern Oscillation (ENSO), the Pacific Decadal Oscillation, and the North Pacific Gyre Oscillation (Di Lorenzo et al., 2008; Jacox et al., 2015, 2014). In recent years, there has been rapid growth in the use of ocean reanalyses for oceanographic and ecological research in the CCLME, with reanalyses providing the environmental information used to develop oceanographic indices (Jacox et al., 2018; Santora et al., 2020), and to model species distribution shifts (e.g., (Abrahms et al., 2019; Brodie et al., 2018) and population fluctuations (e.g., (Schroeder et al., 2014; Tolimieri et al., 2018).

There have been some previous comparisons between high-resolution reanalyses and observations in the CCLME. For example, (Schroeder et al., 2014) showed that the high-resolution (1/10˚) data assimilative implementation of the Regional Ocean Model System (ROMS) from the University of California Santa Cruz (hereafter referred to as the California Current System reanalysis or CCSRA) compares well to *in situ* hydrographic measurements of ocean temperature, salinity, and upper ocean stratification during boreal winter and spring in the coastal ocean between Monterey Bay and Pt. Arenas. Additionally, (Neveu et al., 2016) showed that the CCSRA can credibly capture the spatial distribution of Eddy Kinetic Energy (EKE) off California's coast. While these studies indicate that high-resolution ocean reanalyses may provide an accurate and spatiotemporally consistent depiction of ocean properties in the CCLME over the last several decades, they are primarily focused on a single regional ocean model. A careful analysis of the strengths and weaknesses of several different high-resolution ocean reanalyses in this region would benefit marine scientists interested in leveraging these tools for research into ocean climate variability and its impact on marine ecosystems in the CCLME.

In this study, we independently verify the mean and variability of key ocean parameters in the CCLME from three high-resolution ocean reanalyses—the Global Ocean Reanalysis and Simulations (GLORYS), the Ocean Reanalysis System version 5 (ORAS5), and CCSRA. In particular, we focus on the historical representation of temperature and salinity at the ocean surface, within the water column, and at the ocean bottom along the continental shelf, as well as SSH along the U.S. west coast. These variables were chosen due to their importance as leading indicators of marine resource response to climate variability (Ottersen et al., 2010; Pinsky et al., 2013) and due to the availability of *in situ* and satellite observations suitable for model reanalysis evaluation.

2. Data and Methods

2.1 The high-resolution ocean reanalyses

Below are descriptions of the high-resolution ocean reanalyses evaluated in this study. The spatiotemporal availability of each is further summarized in Figure 1 and Table 1.

a) GLORYS

We evaluate data from the Global Ocean Reanalysis and Simulations (GLORYS) version 1 global ocean reanalysis (Lellouche et al., 2021). Available through the Copernicus Marine Environmental Monitoring Service (CMEMS), GLORYS offers daily mean and monthly mean ocean variables at 1/12˚ (~8 km) horizontal resolution with 50 vertical levels. The reanalysis is generated using the Nucleus for European Modelling of the Ocean (NEMO) ocean model, forced at the surface by the European Center for Medium-Range Weather Forecasts (ECMWF) ERA-Interim atmospheric reanalysis. Output is available for 1993-2019, during which the model assimilates along-track satellite altimetry, satellite sea surface temperature (SST), sea ice concentrations, and *in situ* profiles of temperature and salinity from the Coriolis Ocean database ReAnalysis (CORA) dataset (Szekely et al., 2019).

b) ORAS5

The global Ocean Reanalysis System version 5 (ORAS5; (Zuo et al., 2019, 2017), which includes a dynamic-thermodynamic sea ice model and surface wave model, was developed at ECMWF. Like GLORYS, it uses the NEMO ocean model (version 3.4.1) and surface forcing from the ERA-Interim reanalysis. The ORAS5 assimilates in-situ profiles of temperature and salinity from the "EN4" dataset (Good et al., 2013), merged *in situ* and satellite SST from HadISST2, and along track satellite altimeter-derived sea-level anomalies from AVISO (Archiving, Validation and Interpretation of Satellite Oceanographic data). Monthly and daily mean fields are available from 1979 to present, at a horizontal resolution of 0.25˚ (~25 km). There are 75 vertical depth levels, with layer thickness increasing from 1 m near the surface to 200 m in the deep ocean. ORAS5 consists of five ensemble members obtained using perturbing forcing fields and slight sampling differences in the observations that were assimilated. Here we used the ensemble mean of the five members.

c) CCSRA

A series of California Current System (CCS) ocean reanalyses have been developed by the Ocean Modeling group at UC Santa Cruz (https://oceanmodeling.ucsc.edu). Here, we evaluate a historical reanalysis covering 1980-2010 (Neveu et al., 2016), and an extension covering 2011-2018. In both cases, the CCSRA employ the Regional Ocean Modeling System with 4-dimensional variational data assimilation (ROMS 4D-Var; (Moore et al., 2011). The domain covers nearly the entire the U.S. west coast (30-48˚N) and offshore to 134˚W with a horizontal resolution of 0.1˚ (~10 km) and 42 terrain-following vertical levels (Figure 1). Surface forcing for the 1980-2010 reanalysis is derived from a combination of ECMWF atmospheric reanalyses (ERA-40 and ERA-Interim) and cross-calibrated multiplatform (CCMP) winds, while the extension uses higher-resolution forcing from the Naval Research Laboratory's Coupled Ocean Atmosphere Mesoscale

Prediction System (COAMPS). Ocean boundary conditions are provided by the Simple Ocean Data Assimilation (SODA) product for 1980-2010, and from the GLORYS product after 2010. Assimilated data include satellite SST from AVHRR, AMSR-E, and MODIS, satellite SSH from AVISO/CMEMS, and *in situ* temperature and salinity profiles from the EN3 database (Ingleby and Huddleston, 2007) for 1980-2010 and from the EN4 database (Good et al., 2013) post-2010.

Due to changes in the surface forcing and ocean boundary conditions between the historical CCSRA and its extension, care must be taken when combining the two into a continuous record. In general, agreement is good for quantities that are well constrained by observations, including temperature. For SSH, the switch in ocean boundary conditions introduces a small offset (Section 3.6) that could be misinterpreted as a trend or low frequency variability. Past studies have found that some variables (e.g., surface properties, upper ocean stratification) are suitable for concatenation (Brodie et al., 2018) while others (e.g., subsurface currents) are not (Tolimieri et al., 2018). Thus, the prudent approach for any user would be to examine the reanalyses for any obvious inconsistencies that might preclude stitching them together for a specific application.

2.2 Observations used for model evaluation

Below are descriptions of the observations used to validate the three reanalyses described above. The spatiotemporal availability of all observations is further summarized in Figure 1 and Table 2.

a) Sea surface temperature and sea surface salinity

We first compare SST from the three reanalyses to those from the NOAA Optimum Interpolation Sea Surface Temperature version 2.1 (OISSTv2.1; (Huang et al., 2021; Reynolds et al., 2007) dataset, available for 1981 to present. The OISST blends satellite measurements with *in* *situ* data from ship, buoys and Argo floats, using a number of steps to reduce biases and provide the data on a 0.25° grid. However, the processes inherent in creating the OISST smooths the SST field and thus the true or feature resolving resolution is less than 0.25° (Reynolds et al., 2013); the smoothing varies in time and space but can be more pronounced in coastal regions (Reynolds and Chelton, 2010). Therefore, in order to further validate the coastal environment in the reanalyses, we compare with SST measurements from six nearshore stations that span most of the U.S. west coast (see Figure 1 yellow diamonds and Table 2). These *in situ* observations were not assimilated into any of the reanalyses, and thus, provide an independent estimate of SST variability.

Sea surface salinity (SSS) from the reanalyses is compared to the Level-4 SSS data from the Multi-Mission Optimally Interpolated Sea Surface Salinity (OISSS) Global Dataset V1.0 (Melnichenko et al., 2016). This dataset optimally interpolates Level-2 swath measurements of SSS from the Aquarius, Soil Moisture Active Passive (SMAP), and Soil Moisture and Ocean Salinity (SMOS) satellite missions to produce monthly mean SSS estimates from August 2011 to present on a global 0.25˚ grid. We limit our SSS comparisons to 2012-2018, which is the longest overlapping period between the satellite observations and the reanalyses.

b) *Water column temperature and salinity*

Water column temperature and salinity was obtained from two sources. First, temperature profiles between the surface and 500m (or the bottom if shallower) were obtained from the California Underwater Glider Network (CUGN, (Rudnick et al., 2017). The gliders follow three paths originating from the California coast (at Monterey, Pt. Conception and Dana Pt.) extending southwestward ~500 km offshore, with each out-and-back section taking 2-3 weeks to complete (Figure 1; purple lines). Gliders dive between the surface and 500 m with each dive taking approximately three hours and covering three km horizontally. Several statistical methods are applied to the original data, including a least squares fit and objective mapping, to obtain anomalies and a mean seasonal cycle on a rectangular grid as a function of depth, distance offshore, and time (Rudnick et al. 2017). We compare the reanalyses to this gridded data, which has a vertical resolution of 10 m, a horizontal resolution of 5 km, and a temporal resolution of 10 days. Our comparisons are for the period 2007 to 2018, the longest overlapping time period for the glider data and the reanalyses. The CUGN data is not assimilated into ORAS5 or GLORYS, but it is assimilated into the historical CCSRA reanalysis from 1980-2010. The CCSRA extension from 2011-2018 does not assimilate CUGN data.

Each reanalysis is further compared with Argo profile measurements of water column temperature and salinity in the CCLME. Argo is a global network of autonomous profiling floats that measure the temperature and salinity of the ocean's upper 2000 m. Since 1999, the Argo program has collected more than 2 million hydrographic profiles worldwide (Jayne et al., 2017). Here, we evaluate 18,971 quality-controlled Argo profiles in the CCLME covering 2002-2018 (Figure 1a; shading). For our comparisons, we average profiles in three sub-regions within the CCLME: (1) 40˚N-50˚N (referred to as North LME or NLME), (2) 30˚N-40˚N (referred to as Central LME or CLME), and (3) 20˚N-30˚N (referred to as South LME or SLME). See Figure S1 for sub-region boundaries.

Prior to analysis, individual Argo profiles were categorized into one of the three CCLME sub-regions based on their latitude and longitude. Profile measurements were then aggregated in the vertical into 20m bins in order to acquire uniformly spaced temperature profiles in each sub-region as a function of time. Finally, we averaged all profiles within a given month to produce monthly mean water column temperature measurements as a function of depth. See Figure S1 and Figure S2 for the density of Argo measurements in depth and time. While Argo measurements are assimilated into each of the ocean reanalyses discussed in this study, an intercomparison between Argo and the reanalyses may shed light on any potential deficiencies in the individual data assimilation schemes. In particular, differences in model physics, resolution, and the incorporation of other *in situ* datasets may impact the overall assimilation of Argo measurements. When compared to Argo data, all temperature and salinity anomalies are relative to the period 2002-2018.

c) Bottom temperature

Bottom temperature data were obtained from three sources. First, the deepest portion of the glider profiles described above were used as near-bottom temperature measurements. Some comparisons were not practical as: (1) the shelf was too narrow for comparing reanalyses to gliders west of Dana Pt.; (2) the CCSRA grid points were too far away for a reasonable comparison with the CUGN location at 410 m in Monterey Bay as a result of using smoothed bathymetry, and (3) the ORAS5 reanalysis was too coarse for comparisons with the CUGN data to be meaningful. Second, we used temperature observations from conductivity, temperature, depth sensor (CTD) casts at three locations (See Table 2) along the Newport Hydrographic Line (Huyer et al., 2007), which extends west from the Oregon coast at 44.65°N (Figure 1; green line), to estimate the variability of bottom temperature during 2008 to 2018. Third, we use bottom temperature data that are collected as part of the U.S. West Coast Groundfish Bottom Trawl Survey (WCGBTS; (Keller et al., 2017), conducted between May and October of each year by NOAA’s Northwest Fisheries Science Center. Since 2003, the WCGBTS has covered the shelf/slope region of the entire U.S. west coast, sampling bottom depths of 55-1280m using a random stratified sampling design, with ~500-700 total stations per year (Figure 1; red dots). Each tow is ~15 minutes in duration, covering ~0.55 km horizontally. The Newport Line and bottom trawl survey data are not assimilated in any of the reanalyses. As mentioned previously, the CUGN data is only assimilated in the first segment of the CCSRA data (1980-2010). Therefore, many of these observations offer an independent metric by which to make our comparisons

While the stratified random sampling pattern of the trawl data precludes the generation of climatologies and thus an evaluation of bottom temperature variability, it does provide much more thorough spatial coverage from which we can assess the ability of reanalyses to reproduce mean patterns of bottom temperature. Bottom temperature variability will instead be assessed using the measurements from the CUGN and Newport Line, which are more consistent in time and space than the trawl data. Comparing bottom temperature between observations and reanalyses is made further complicated by the relatively narrow shelf and steep bathymetry off the U.S. west coast. Even with ~10 km reanalysis resolution, the true bottom depth can be very different from the reanalysis bottom depth at the nearest grid point, which in turn can produce large differences in the observed and reanalysis bottom temperature. One can account for this discrepancy by matching the bottom temperature measurements with reanalysis output at the same depth (not necessarily at the reanalysis bottom). Specifically, for each bottom temperature observation we first find the closest reanalysis grid cell where the model bottom is deeper than the observed depth, and then linearly interpolate the reanalysis water temperature profile to the depth of the observation (Figure 2). To illustrate the influence of using the reanalysis bottom temperature rather than a depth-matched temperature, we compare the two methods for the bottom trawl data (Section 3.5).

d) Coastal sea surface height

To verify sea level variability along the U.S. west in the ocean reanalyses we compare them to data from nine tide gauges (Figure 1; black dots) maintained by the Joint Archive for Sea Level (JASL), which is a partnership between the University of Hawaii Sea Level Center (UHSLC) and the National Centers for Environmental Information (NCEI). Here, we utilize the Research Quality Data Set (RQDS). These observations are not assimilated in any of the reanalyses discussed here, and therefore offer an independent metric by which to verify sea level changes. For this comparison, monthly SSH anomalies are relative to a long-term climatology of 1993-2018, which is the longest overlapping period between the tide gauges and the reanalyses.

e) Eddy Kinetic Energy

Finally, ocean mesoscale variability is ubiquitous along the U.S. west coast and plays a key role in modulating regional ocean dynamics such as coastal upwelling in the CCLME, which can impact primary productivity through the vertical transport of key nutrients in and out of the euphotic zone (Gruber et al., 2011; Renault et al., 2016). We assess the representation of mesoscale variability in the ocean reanalyses using geostrophic Eddy Kinetic Energy (EKE), calculated as:

$$EKE = \frac{1}{2}\left(U'^{2}_{g} + V'^{2}_{g}\right) \tag{1}$$

Where U'_g and V'_g, respectively, are the zonal and meridional components of the daily mean geostrophic surface current anomalies estimated from daily mean SSH anomalies. The EKE in the ocean reanalyses is compared to the EKE calculated from AVISO satellite altimetry measurements (Ducet et al., 2000), which provides daily mean SSH anomalies from 1993-2012 on a 0.25˚ grid. For consistency with the AVISO data, SSH anomalies from each reanalysis are computed relative to the long-term mean of a 1993-2012 base period. The EKE in each reanalysis was first calculated on the native model grid and then interpolated to the AVISO 0.25˚ grid in order to compare them to the coarser AVISO data.

2.3 Analysis approach

a) Statistical methods

The purpose of this study is to inform potential users of these high-resolution reanalyses whether (and in what context) these different data assimilative models may act as a reliable substitute for the comparably noisy, discontinuous and sporadic raw ocean measurements. To support this goal, we evaluate each ocean parameter using a set of common statistical comparisons that are broadly applicable to a range of potential research applications. These comparisons include assessments of the reanalyses' mean state and variability as measured by the mean bias, root-mean-square error (RMSE), and correlation coefficient relative to observations. Where appropriate statistical significance is evaluated using a Student's t-test with a 95% confidence interval after correcting the degrees of freedom for lag-1 autocorrelation.

b) Observation limitations

Due to the different spatiotemporal coverage and sampling schemes of different observational platforms, each is well suited to some types of analyses but not others. For instance, to assess the degree to which reanalysis data represents the observed variability of an ocean parameter at any given location (as measured by RMSE and/or the correlation coefficient), consistent measurements are required at that location for many years to derive a representative climatology and anomalies. Some observational platforms do provide long records at one location (such as the coastal station data analyzed in Section 3.1b), but these measurements are not useful for assessing the variability of an ocean parameter over large areas. As a result, to validate variability in reanalyses on a broader scale, we either: (1) aggregate sporadic point measurements in time and space to produce a consistent record suitable for calculating anomalies for comparison to the models (e.g., our approach with Argo profiles in Sections 3.3b and 3.4b) or (2) rely on observational products that have been infilled or interpolated onto a grid, such as AVISO, OISST, and OISSS. Gridded observations may feature statistical artifacts or biases introduced during interpolation (Reynolds et al., 2013; Reynolds & Chelton, 2010). However, by also comparing the reanalyses to raw and *in situ* point measurements, such as the Newport line data, shore stations, tide gauges, bottom trawl measurements, and Argo profiles, we hope to assess any sensitivities our results may have to our choice of observations. We return to these topics in more detail in the Discussion section.

3. Results

3.1 Sea surface temperature

a) Comparisons with OISST

The annual mean SST pattern in OISST features a tongue of cool SSTs along the U.S. west coast, which is associated with the southward advection of cold water from high-latitudes by the California Current, as well as broadscale upwelling driven by the climatological northerly winds (Figure 3a). Overall, the reanalyses show modest SST biases relative to the annual mean OISST data (Figure 3b-d). Additionally, each reanalysis exhibits a similar bias pattern: cool along the U.S. west coast and warm offshore. Among the three reanalyses, GLORYS shows the largest warm SST biases (~0.5˚C) from 32˚N-42˚N, while the U.S. west coast SST biases in ORAS5 and CCSRA are generally insignificant.

The greatest SST variability, as indicated by the standard deviation of the monthly means in OISST, is adjacent to the coast with regions of somewhat enhanced variability extending farther offshore of the California coast (Figure 3e). Additionally, there are areas of higher variability along the southern half of Oregon and portions of California, including between Pt. Arena and Pt. Reyes (~38°N) and in the vicinity of Pt. Conception and the Channel Islands (~34°N). While the nearshore peak in the SST standard deviation is likely associated with upwelling driven SST changes, the offshore extension of elevated variability may be associated with enhanced eddy activity in this region (see also Section 3.7). The monthly mean SST anomalies in each reanalysis are highly correlated with the corresponding SST anomalies from OISST (Figure 3f-h), indicating that the reanalyses credibly reproduce the regional structure of the observed variability. This is further supported by the high pattern correlations between the OISST monthly standard deviation pattern and those from the reanalyses (Figure S3e-h). Among the reanalyses, CCSRA most closely resembles OISST, featuring the highest point-by-point correlations with the observed monthly mean SST anomalies and the highest pattern correlation with the observed monthly standard deviation pattern. In comparison, ORAS5 exhibits less variability than OISST over most of the domain and GLORYS SST variability is generally greater than that in OISST, particularly in the southwest portion of the domain (Figure S3f-g). These differences contribute to the overall lower (yet still significant) point-by-point correlations between ORAS5/GLORYS and the OISST monthly mean SST anomalies (Figure 3f-g).

b) Comparisons with nearshore stations

The reanalyses also generally compare well with SST data from six nearshore stations spanning the U.S. west coast (Figures 4 and S4). At the northernmost stations (Stonewall; 44.7˚N and Charleston; 43.3˚N), CCSRA and GLORYS exhibit insignificant annual mean biases, while ORAS5 has a significant warm bias at Charleston. However, all three reanalyses show high monthly mean correlations and relatively small RMSE of ~0.25-0.4˚C when compared to station data at these latitudes. At Trinidad Beach (41.1˚N), GLORYS and ORAS5 exhibit significant cold biases. While CCSRA does not have a significant mean bias at this station, all three reanalyses show lower (yet still significant) monthly correlations, which also correspond to elevated RMSE values of ~0.5-0.6˚C. The reanalyses compare most favorably at the Farallon Islands (37.7˚N), where they have the smallest biases, highest correlations, and lowest RMSE values of any of the stations compared here. The southern stations (Newport Beach; 33.6˚N and Scripps Pier; 32.9˚N) show the largest differences among the reanalyses. For example, at Newport Beach, GLORYS exhibits a significant cold bias of about 1˚C, while ORAS5 and CCSRA show warm biases of ~0.8˚C and 1˚C, respectively. At the Scripps Pier, both ORAS5 and GLORYS show significant cold biases, while CCSRA does not have a significant bias. These potential reanalysis errors are further highlighted by overall lower monthly correlations and higher RMSE values at these southern stations, especially for GLORYS at Newport Beach. The cold biases in GLORYS at these southern latitudes may be due to enhanced upwelling rates. We will explore this possibility in more detail in Section 3.3a.

3.2 Sea surface salinity

a) Comparisons with OISSS

The annual mean SSS pattern in OISSS features salinity values that decrease with latitude, reaching a minimum in the coastal regions of the Pacific Northwest, which is likely associated with increased freshwater fluxes from the Columbia River outflow at ~46˚N (Figure 5a). Despite the high pattern correlations between the annual mean SSS in the reanalyses and the OISSS data (Figure S5b-d), each of the reanalyses have significant mean SSS biases, particularly off the Oregon and Washington coast (Figure 5b-d). The bias patterns in GLORYS and ORAS5 are similar, with mainly salty biases offshore north 45˚N and mainly fresh biases in a horseshoe pattern along the coast and extending offshore from 40˚N-50˚N. South of 40˚N, both GLORYS and ORAS5 have mostly insignificant biases. In contrast, CCSRA has significant salty biases from 40˚N-50˚N and significant negative biases offshore south of 40˚N. The large biases in each of the reanalyses near the Columbia River outflow suggests unrealistic freshwater forcing in this region. GLORYS and ORAS5 have potentially too much freshwater input, while the salty nearshore biases in CCSRA are consistent with its omission of freshwater sources (Neveu et al., 2016).

Monthly SSS variability is relatively weak throughout the CCS, except near the Columbia River outflow where monthly standard deviations reach as high as 0.35 PSU (Figure 5e). The monthly mean SSS anomalies in GLORYS are highly correlated with the OISSS anomalies offshore and south of about 40˚N, but are not significantly correlated with the observations off the coast of Oregon and Washington (Figure 5f). In comparison, ORAS5 has the highest overall correlations with the satellite data; however, it is less correlated along the coast from 30˚N-50˚N (Figure 5g). The SSS anomalies in CCSRA are significantly correlated with OISSS south of 38˚N, but the correlations are insignificant nearly everywhere north of 38˚N (Figure 5h). Overall, ORAS5 has the highest pattern correlation with the observed monthly mean standard deviation pattern ($r = 0.84$), while GLORYS and CCSRA have weaker pattern correlations of 0.77 and 0.60, respectively (Figure S5e-h).

3.3 Water column temperature

a) Comparisons with glider lines

When compared to annual mean temperature data from the CUGN, each of the three ocean reanalyses displays a different annual mean temperature bias pattern that is broadly consistent across the three different glider lines (Figures 6 and S6). For example, GLORYS shows a nearshore significant warm bias centered at ~30m depth, as well as a cold bias along the continental shelf at ~60-180m depth at Monterey and Pt. Conception at the surface at Dana Pt. (Figure 6d-f). The warm bias may be due to GLORYS having a slightly deeper nearshore mean thermocline than observed (Figure S6a-f). Due to the strong vertical temperature gradients found within the thermocline, even subtle shifts in its mean position can result in large model biases. The coastal cold biases are most significant at Pt. Conception and may be related to differences in nearshore vertical transport rates, as indicated by the shallower isotherm slopes below ~90m and within ~40km of the coast in GLORYS compared to CUGN (Figure S6b,e). The cold biases near the surface at Dana Pt. may be due to enhanced upwelling in GLORYS, as indicated by the more vertically tilted annual mean isotherms above 90m within 20km of the coast. Enhanced upwelling at these latitudes may also explain the significant cold SST biases seen at Newport Beach and the Scripps Pier (Figure 4a). At each glider line, ORAS5 shows significant warm temperature biases at all vertical levels within ~150-200km of the coastline (Figure 6g-i). These warm biases are likely the result of weaker overall upwelling rates in ORAS5, as indicated by the generally flatter nearshore mean isotherms when compared to observations (Figure S6g-i). Weaker upwelling in ORAS5 may be due to the coarser resolution of its ocean model. Finally, CCSRA features significant warm biases at each glider line that extend offshore and slope upwards towards to the coast (Figure 6j-l). These biases are related to a systematically deeper mean thermocline in CCSRA when compared to CUGN (Figure S6j-l). Overall, the annual mean temperature in GLORYS has the lowest pattern root mean square error (RMSE) when compared to the CUGN data at each location (ranging from 0.17˚C to 0.21˚C), while ORAS5 tends to have the highest (ranging from 0.30˚C to 0.36˚C).

All three reanalyses show high monthly mean correlations with the CUGN data above ~30m at Monterey, above ~50m at Pt. Conception, and above ~90 m at Dana Pt (Figure 7). The correlations in each reanalysis are lower below these depths, with the exception of GLORYS at Monterey and Pt. Conception where there are high correlations offshore at all depths. Overall, GLORYS has the highest correlations with observations across each CUGN domain, while ORAS5 has the weakest. The different correlation patterns among reanalyses may be related to their different monthly mean standard deviation patterns when compared to observations (Figure S6, shading). Overall, GLORYS and ORAS5 have similar pattern RMSE values with the observed temperature variability maps (ranging from 0.08˚C to 0.13˚C), while CCSRA has slightly higher RMSE values (ranging from 0.13˚C to 0.16˚C).

b) Comparisons with Argo profiles

All three reanalyses generally reproduce the timing and relative magnitude of CCLME subsurface temperature anomalies observed by Argo floats (Figure 8), including major interannual warming events in 2004-2006 and 2014-2016 related to El Niño and Northeast Pacific marine heatwaves (Amaya et al. 2016; Li et al. 2020) and cooling events in 2007-2009 related to La Niña (Okumura and Deser, 2010). The multi-year warming from 2014-2016 seen in observations and the reanalyses is most pronounced and consistent in the CLME and SLME, likely due to the southward shift in large-scale anomalous atmospheric forcing associated with major marine heatwaves in the Northeast Pacific in late 2014 and early 2015 (Amaya et al., 2016). Argo measurements in the SLME also show resurgent warming in 2017-2018, which is generally reproduced by GLORYS and ORAS5 (SLME is outside the CCSRA domain). Finally, there is evidence in the Argo record of downward propagating warm waters from the surface in 2014-2015 to 100-150 m in 2015-2016 (particularly in the CLME and SLME). These downward propagating anomalies are evident in each reanalysis, although the magnitude of the CLME anomalies during this period are overall weaker in CCSRA.

While the ocean reanalyses generally have weaker temperature anomalies then those computed from Argo, the discrepancy is likely due to the relatively small number of Argo profile measurements in both depth and time that go into area averages (Figures S1 and S2). Indeed, if we resample daily mean GLORYS vertical temperature data to the same time and depth locations as the Argo data, we find that agreement between the two is greatly improved (Figure S7). This result highlights an advantage of ocean reanalyses, which provide a uniform dataset in time and space, over *in situ* observations that can give a biased view of the ocean state due to under sampling (see Section 5 for a more detailed discussion).

3.4 Water column salinity

a) Comparisons with glider lines

The reanalyses each exhibit salinity bias patterns that are broadly consistent across the different glider lines (Figure 9). For example, GLORYS has significant fresh biases near the coast, which peak at the surface at Monterey and at ~60m-90m at Pt. Conception and Dana Pt. (Figure 9d-f). In contrast, ORAS5 has significant salty biases mainly below 120m along each glider line, with significant fresh biases near the surface at Monterey and Pt. Conception (Figure 9g-i). Many of the fresh biases seen in GLORYS and ORAS5 may be associated with differences in vertical transport near the coast. For example, the CUGN data at Monterey shows annual mean isohalines of 33.1-33.5 PSU outcropping within ~200km of the coastline, however, neither GLORYS nor ORAS5 show outcropping isohalines beyond 33.2 PSU, resulting in fresh biases nearshore (Figure S8). On the other hand, CCSRA has isohalines at Monterey that have a similar slope to observations, resulting in weaker biases nearshore. At Dana Pt., there is a layer of saltier water seen in the annual mean CUGN data from the surface to ~60m within ~100km of the coast (Figure 9c). This shallow, salty water is not reproduced in any of the reanalyses, contributing to the fresh biases seen here in GLORYS and ORAS5 (Figure S8, bottom row). Additionally, CCSRA has a systematically weaker and shallower halocline at all three glider lines than in the CUGN, resulting in positive significant salty biases that slope upwards from offshore to onshore and which overlie significant fresh biases at deeper depths (Figures 9j-l and S8).

In general, the monthly mean correlations between CUGN and reanalysis water column salinity are overall lower than the corresponding temperature correlations (comparing Figures 7 and 10). The lower overall salinity correlations are consistent with the large differences seen in the monthly mean standard deviation patterns between the CUGN and reanalysis data (Figure S8). Despite these clear deficiencies, the reanalyses do have regions of significant salinity anomaly correlations along each glider line, with the highest correlations for all three reanalyses found at Dana Pt (Figure 10, bottom row). In this region, each reanalysis shows significant correlations throughout the water column, with the highest values near the surface for GLORYS and ORAS5 and from 90-120m for CCSRA. Further, both GLORYS and CCSRA show significant (albeit weaker) correlations with CUGN salinity data throughout the water column at Monterey and Pt. Conception, with the highest correlations in GLORYS found near the surface west of 120˚W and the highest correlations in CCSRA found below ~50m along each line (Figure 10, top and middle rows). In contrast, ORAS5 has noticeably weaker correlations at Monterey and Pt. Conception, with insignificant values throughout much of the water column (Figure 10g-h). However, ORAS5 does have significant correlations near the surface along the glider lines and at depth within ~40km of the coast.

b) Comparisons with Argo profiles

The reanalyses broadly reproduce the timing of major interannual salinity anomalies observed by Argo profiles (Figure 11). For example, in the NMLE and CLME there was a period of fresher than normal conditions from the surface to ~150m from 2003-2006, which is seen in

GLORYS and ORAS5, but less so in CCSRA. This period of fresh anomalies also encompasses the SLME in GLORYS and ORAS5, but there are very few Argo profiles during this time with which to validate the reanalyses (Figure S2c). Additionally, Argo profiles show a period of salty anomalies beginning in ~2016 in each sub-region that is broadly captured by the reanalyses. Despite these similarities, there are some important differences between the Argo salinity data and the reanalyses. For example, Argo shows fresh anomalies in 2009-2010 and 2013-2015 in the NLME, which are mostly absent in the reanalyses. There are also larger and more persistent salty anomalies in GLORYS and ORAS5 in the SLME from 2014-2016 than seen in Argo. While GLORYS and ORAS5 show larger anomalies during this time period, the Argo profiles tend to have larger salinity anomalies overall, which as discussed previously is likely related to the limited number of individual Argo profiles in these regions (Figures S1-S2, S7).

3.5 Bottom Temperature

a) Comparisons with the CUGN and Newport Line

Both GLORYS and CCSRA produce monthly mean bottom temperature data that are significantly correlated with nearby CUGN values at the Monterey Bay (Line 66) and Pt. Conception (Line 80) locations as well as data taken from CTD casts along the Newport Line (Figure 12; note ORAS5 is omitted from this comparison due to its inability to resolve the shelf). For the CUGN lines, both the GLORYS and CCSRA bottom depth correlations decrease with depth, while the reanalyses have peak correlations with the Newport Line data at 55m depth. At CUGN Line 66, GLORYS and CCSRA have similar correlations with observations. However, at CUGN Line 80 and along the Newport Line, GLORYS has noticeably higher correlations.

b) Comparisons with bottom trawl data

While the stratified random sampling pattern of the trawl data precludes the generation of climatologies, it does provide much more thorough spatial coverage from which we can assess the ability of reanalyses to reproduce mean patterns of bottom temperature. As described in the methods (Section 2.2c), there are discrepancies between reanalysis bottom depth and true bottom depth, and consequently there is considerable scatter when comparing the observed and reanalysis bottom temperature at the same location. Despite these discrepancies, there is good agreement between the reanalyses (especially GLORYS) and the observations (Pearson correlation coefficients r = 0.76 and 0.92 for CCSRA and GLORYS, respectively; Figure 13). This effect is less pronounced in GLORYS than in CCSRA, as the terrain-following coordinate system used in the latter requires additional bathymetric smoothing that introduces greater differences between modeled and true bottom depth. However, when bottom depth differences are accounted for (i.e., reanalysis temperature is taken from the depth of the trawl measurement), both GLORYS and CCSRA exhibit strong fidelity to observed bottom temperatures (r = 0.97 and 0.96, respectively). Patterns of mean bias differ between the two reanalyses, with CCSRA tending to be too warm at the coldest temperatures and too cold at the warmest temperatures, while GLORYS is slightly warm at high temperatures (Figure 13).

3.6 Sea surface height

Coastal sea level measurements from tide gauges and reanalyses exhibit large monthly and interannual fluctuations that are likely associated with local wind forcing and propagating coastally trapped waves (Amaya et al., 2022) that may be stochastically forced or driven by major ENSO events, including the 1997-1998 and 2015-2016 El Niño events (Figures S9-S11). The SSH anomalies in each reanalysis are significantly correlated with the tide gauge measurements at every location (Figure 14a). However, GLORYS produces the highest correlation values at every station except the Humboldt Bay tide gauge at ~40˚N. Latitudinal patterns in RMSE roughly mirror those in the correlation, with the lowest RMSEs in the south of the domain and larger values in the north (Figure 14b). GLORYS tends to have the lowest RMSE values, while ORAS5 has the highest values north of 36˚N.

3.7 Eddy Kinetic Energy

The annual mean EKE pattern from AVISO data shows a band of elevated values that roughly follow the North American coastline from 20˚N-45˚N, with peak values centered offshore in the CCLME around 37˚N (Figures 15 and S12). In comparison, GLORYS has a significant positive bias throughout much of the CCLME from 30˚N-40˚N, while ORAS5 has a significant negative bias throughout the entire domain. The EKE bias pattern in CCSRA exhibits significant positive biases near the edges of the regional model domain, likely associated with the lateral boundary conditions used to force the regional model, however, CCSRA tends to have the smallest biases in the main EKE region (Figure 15b-d). When considering monthly EKE variability, we see that both GLORYS and CCSRA are significantly correlated with the observations throughout their respective domains, while ORAS5 generally has insignificant correlations (Figure 15e-h).

The strong negative EKE biases and low correlations seen in ORAS5 may be due, in part, to the model's coarse resolution (0.25˚), which is not eddy resolving. Additionally, although ORAS5 does assimilate satellite altimetry data, these observations are rejected from the assimilation scheme in the nearshore environment (i.e., when the ocean bottom depth is shallower than 500m) and are strongly down weighted within ~800km of the coastline (Zuo et al. 2019). As a result, the mean structure and variability of the mesoscale features are generally not captured in

ORAS5. It is also important to note that our comparisons are somewhat hampered by the resolution of the gridded AVISO data. The higher EKE values in both GLORYS and CCSRA are due to their higher horizontal resolutions, which allow their respective ocean models to simulate finer scale circulation features, thus increasing the level of eddy activity relative to AVISO. Therefore, in reality (where ocean currents are not limited by horizontal grid resolution), it is likely that the magnitude of EKE in the CCS is actually closer to that seen in the high-resolution reanalyses.

4. Summary

In this study, we compared the output from three high-resolution ocean reanalysis products—the 1/4˚ ORAS5, the 1/12˚ GLORYS, and the 1/10˚ CCSRA—to a variety of *in situ* and satellite-derived observations in the CCLME. For surface temperature, we found that all three analyses were generally able to capture the observed mean state and monthly variability as measured by satellite observations and coastal station data over the last several decades. In particular, when comparing to unassimilated SST data from six stations along the U.S. west coast, we showed that the nearest grid cells in each reanalysis were highly correlated with the observations, with typical correlation values exceeding 0.80 and even reaching as high as 0.97 at some stations (Figure 4). Comparing across the reanalyses, we found that CCSRA has the most accurate depiction of monthly SST throughout the CCLME, while GLORYS and ORAS5, respectively, slightly overestimated and underestimated the larger-scale SST variability (Figures 3 and S3). For sea surface salinity, we found significant biases in all three reanalyses near the outflow of the Columbia River at 46˚N (Figures 5 and S5), suggesting an influence of unrealistic (or missing) freshwater forcing in the models at this location.

Comparisons to different measures of water column temperature throughout the CCS saw large differences among the ocean reanalyses. For example, each reanalysis depicted significant warm mean temperature biases relative to CUGN data (Figure 6). These biases were the result of differences in the mean position of the thermocline (as in GLORYS and CCSRA) or differences in vertical transport rates (as in ORAS5). Despite these mean biases, monthly mean temperature data from each reanalysis were significantly correlated with the CUGN data above ~50m. However, GLORYS showed a clear advantage over ORAS5 and CCSRA with the highest correlations throughout the water column (Figure 7). Comparisons to CUGN salinity data were less favorable among the reanalyses (Figures 9 and 10), with each reanalysis showing large biases related to differences in vertical transport rates (as in GLORYS and ORAS5) and differences in the mean position of the halocline (as in CCSRA), as well as weaker overall point-by-point correlations than with the corresponding temperature observations. All three reanalyses credibly reproduce the large-scale subsurface temperature and salinity anomalies measured by Argo profiles, including the downward propagation of recent warm anomalies associated with the 2015-2016 marine heatwave and the recent salty conditions throughout the CCS after 2016 (Figures 8 and 11).

When comparing bottom temperature measurements from the reanalyses to observations, GLORYS was the best performer. In particular, GLORYS consistently had the highest correlations with monthly mean bottom temperature estimates from the CUGN and the Newport Line (Figure 12). Additionally, due to GLORYS having more realistic bathymetry than CCSRA, it also had bottom temperatures that were much more highly correlated with bottom temperature measurements from nearby trawls (Figure 13). However, accounting for differences in bottom depth between the trawl measurements and the reanalysis led to marked improvements for

CCSRA. Overall, the favorable comparisons between the reanalyses and various bottom temperature observations are impressive given the difficulty of comparing pointwise measurements on the steep bathymetry along the west coast.

The reanalyses also showed impressive correlations with monthly mean SSH measurements from nine coastal tide gauges, producing significant correlation values ranging from 0.67 to 0.91 and accurately depicted major El Niño-driven sea level changes along the U.S. west coast (Figures 14 and S9-S11). However, CCSRA and GLORYS continued to stand out, producing lower overall RMSE values at each tide gauge location when compared to ORAS5. Finally, both GLORYS and CCSRA produced patterns of ocean mesoscale activity (i.e., EKE) that compared well to AVISO satellite measurements, while ORAS5 underestimated the intensity of EKE throughout the CCS due to its coarser resolution and the decision to down weight satellite altimetry data near coastlines during the assimilation process (Figure 15).

5. Discussion

5.1 Choosing a reanalysis for California Current ecosystem science

As is often the case when evaluating reanalyses (Balmaseda et al., 2015; Storto et al., 2019), the “best” product to use depends on the application. However, based on the results above, we can offer some general guidelines as well as some more targeted examples.

5.1.1 General considerations

Model resolution is an obvious consideration when choosing a reanalysis. For research into coastal ocean processes along the U.S. west coast, the higher resolution of GLORYS or CCSRA clearly provides enhanced fidelity of the nearshore environment including surface and bottom temperature, sea level variability, mean coastal upwelling, and the representation of coastally trapped waves (Amaya et al., 2022). However, given the small scale of many of the eddies off the U.S. west coast, Neveu et al. (2016) concluded that the CCSRA horizontal resolution was insufficient to fully represent the observed EKE variability. Throughout the broader CCLME, ORAS5 is generally comparable to GLORYS and CCSRA (with the exception of its poorer representation of EKE), so studies interested in larger-scale ocean variability may prefer ORAS5 with its coarser resolution and smaller overall storage requirements.

Another clear consideration is the spatiotemporal coverage of a particular analysis. For example, GLORYS only provides data starting in 1993, so studies requiring output prior to 1993 would be limited to either CCSRA or ORAS5. Similarly, the CCSRA domain limits analysis to the CCS region, whereas the global ocean models used by GLORYS and ORAS5 do not have such geographical restrictions. As a result, studies utilizing GLORYS and ORAS5 are able to investigate the relationship between the CCS and remote regions (e.g., tropical Pacific) within the same reanalysis dataset. In general, the resolution and performance of GLORYS and CCSRA will make them preferable to ORAS5 unless the application requires both global coverage and a historical record extending earlier than 1993, or if the additional computing/storage burden incurred by using a higher resolution reanalysis is prohibitive.

Finally, potential users of these products should note that they may all be quite limited in certain respects. In our analysis, a chief example is the relatively large salinity errors in each reanalyses when compared to observations. In particular, the large mean salinity biases and weaker overall monthly mean correlations may limit the utility of the reanalysis salinity data in the CCS, especially in the nearshore environment and at higher latitudes near the Oregon and Washington border. Indeed, to the best of our knowledge, none of the reanalyses considered here explicitly represent freshwater inputs (i.e., rivers) to the California Current System at all. Thus, they are not well tailored to applications that are very sensitive to that aspect of the oceanography, though the effects of freshwater inputs will be captured indirectly by assimilation of temperature and salinity observations in the coastal ocean. We do note, however, that our salinity comparisons were limited to relatively short periods (2012-2018 for OISSS, 2007-2018 for CUGN, and 2002-2018 for Argo) and that satellite salinity data have known biases and errors (particularly at higher latitudes), which may influence our comparisons (e.g., Melnichenko et al., 2014). Therefore, it is possible that the apparent deficiencies seen in the reanalysis salinity data will improve as satellite salinity measurements increase in number and accuracy.

5.1.2 Targeted applications

The general guidelines offered above can be further tailored in the context of specific applications, which we demonstrate here by drawing on a range of CCS case studies. For example, in recent years ocean reanalyses have been increasingly relied upon to generate ecologically-relevant oceanographic indices. The Temperature Observations to Avoid Loggerheads (TOTAL; Welch et al., 2019)) tool tracks the risk of Loggerhead Turtle bycatch in California's drift gillnet fishery based on SST anomalies in the Southern California Bight. In summers following persistent warm SST anomalies, temporary closures can be enacted. In this case, CCSRA is likely the most attractive reanalysis due its ability to capture the mean and variability of SST anomalies in the region, and its long record that provides a more robust assessment of historical variability, especially related to ENSO events. The Habitat Compression Index (HCI; Schroeder et al., 2022), which tracks the presence of cool-water habitat nearshore, has been related to regional ecosystem shifts and whale entanglement risk. Again, the fidelity of CCSRA for fine-scale nearshore SST variability, as well the relatively high resolution of its wind forcing and associated representation of coastal upwelling, make it well suited to this application.

Moving to more complex ecological models, additional considerations will drive the choice of reanalysis. As described above, CCSRA is a good choice for surface-oriented analyses, and has been successfully applied in species distribution models focused on the near-surface environment (e.g., Becker et al., 2016). In contrast, GLORYS more realistically captures the bathymetry of the relatively narrow shelf off the US west coast and generally does a better job reproducing observed bottom temperature variability; therefore, it is likely a better choice for species distribution models of benthic organisms such as groundfish (e.g., Ward et al., 2022), provided the shorter historical record is adequate. Issues of internal consistency of reanalyses can also have different impacts depending on the nature of ecological model employed. Changes in the configuration of CCSRA, between its 1980-2010 historical run and an extension starting in 2011, introduce inconsistencies in some fields. For aspects of the ocean circulation that are well constrained by observations or surface forcing (e.g., SST, SSH, upwelling), these inconsistences are of less concern. But some ecological models rely on hydrographic properties that are not well constrained by observations and are more sensitive to changes in model configuration. For example, life-stage specific recruitment models have been developed for several groundfish species in the CCS (Haltuch et al., 2020; Tolimieri et al., 2018) based on mixed layer depth, ocean temperature, and alongshore and cross-shore currents in different vertical and horizontal sectors of the water column. For the subsurface alongshore and cross-shore currents in particular, the 1980-2010 and post-2010 versions of CCSRA cannot be combined as a consistent reanalysis. In this case GLORYS would offer a self-consistent alternative, and while data limitations preclude direct assessment of subsurface currents, the subsurface structure (temperature and salinity) in GLORYS is generally very good relative to the other reanalyses.

5.2 Sources of differences between reanalyses

It is important to consider what factors may lead to one reanalysis comparing more favorably to observations than another reanalysis, particularly if those observations are assimilated by each of the reanalyses (e.g., Argo data is assimilated in all three reanalyses analyzed here). As discussed previously, differences in model resolution likely plays a key role in producing different reanalysis solutions, especially in the nearshore coastal region. Additionally, the resolution and fidelity of the atmospheric forcing for each of the ocean models may play a role. Different data assimilation schemes (e.g., 3D-Var versus 4D-Var) and different subgrid parametrizations may also impact how reanalyses compare to observations. For example, the underestimation of sea level variance in ORAS5 is partly due to suboptimal parameter specifications for observation errors and data sampling (Zuo et al. 2019). However, it is difficult to assess the sensitivity of reanalysis errors to these model design choices without delving deeper into the raw model forcing files or (in some cases) the actual model code, neither of which are readily available to the average user of these reanalyses. Therefore, deciding which ocean reanalysis is most appropriate for a given application should be based on which physical processes are of interest and what computational resources are available to the user.

5.3 Considerations for comparing reanalyses and observations

In addition to the CCS-specific results, our analysis further highlights several important considerations that are generally applicable when comparing raw and/or post-processed observations to reanalysis products. A critical factor for why reanalyses may differ from each other and from nature is the change in the number and types of data that are assimilated over time. Inclusion of new sources of data can lead to discontinuities, while the lack of data, especially early in the record, enhance the contribution of model bias to reanalysis errors. For example, Lellouche et al. (2021) noted that salinity coverage by ARGO was insufficient to constrain model error prior to 2014 in GLORYS. Temporal changes in atmospheric reanalyses used as boundary conditions also can influence ocean reanalyses. The horizontal resolution of the SSTs used as boundary conditions in ERA-interim increased in 2002. This increased small-scale variability in the atmospheric reanalysis winds, which were transmitted to the ocean reanalyses that used ERA-interim, including the three reanalyses examined here; ORAS5 and CCSRA also used other atmospheric reanalyses during portions of their record, which likely contributed to discontinuities. Another change in the reanalyses occurred in 2004, with the inclusion of a large number of ARGO profiles. To accommodate the increase in the vertical profiles the time window in which the bias correction was performed was reduced by a third in GLORYS (Lellouche et al., 2021). This led to a rapid increase of EKE in GLORYS, which may have contributed to its excessive SST variability.

In addition to temporal changes in the type and number of observations, some mean biases may depend on the particular observational data used for comparisons, particularly if one is comparing to raw measurements that have been post-processed or smoothed onto a uniform grid (such as OISST, OISSS, and AVISO data analyzed here) since the interpolation process may introduce statistical artifacts or biases (Reynolds et al., 2013; Reynolds & Chelton, 2010). Although, the weaknesses of interpolating raw observations onto a grid may be partially outweighed by the benefits of post-processing bias adjustments to satellite and ship observations to compensate for platform differences and sensor biases over time (e.g., Reynolds 1993; Reynolds et al., 2007; Banzon et al., 2016; Huang et al., 2020). Further, reanalysis data at any given grid cell represents the characteristics of a volume of water, while *in situ* measurements are often from single points and may therefore benefit from similar bias corrections for comparisons with model output (Chang et al., 2021).

Additionally, the raw Argo profile measurements showed stronger temperature changes than those observed in any of the reanalyses, which may be surprising considering each reanalysis assimilates Argo. However, resampling GLORYS to match the Argo data produced much more consistent results (Figure S7), indicating that apparent discrepancies are in fact largely due to sampling differences. Similarly, EKE values derived on the native GLORYS grid were much higher than the coarser AVISO gridded observations (Figures 15 and S12), and it is possible that the real-world intensity of EKE may be closer to the values seen in GLORYS than in pure observations. For benthic conditions over the continental shelf and slope, an important consideration is that the depth of the real ocean bottom can be significantly different than the depth of the nearest reanalysis grid cell. This depth difference leads to discrepancies between the observed and reanalysis bottom temperature, which are worse when the bathymetry is less realistic (i.e., compare CCSRA and GLORYS in Figure 13). A more accurate representation of the bottom conditions can be obtained by using the reanalysis temperature at the depth of the real ocean bottom, even if it is not the bottom in the reanalysis. However, this extra analytical step is not trivial and requires obtaining the full water column temperature data from the reanalysis, rather than a single level. In general, differences between observations and reanalyses are greatly reduced by controlling for the sampling limitations presented by the observations. Indeed, if there were substantially more Argo profiles in the CCS or if the spatial footprint of satellite altimetry measurements was more similar to GLORYS or CCSRA, then the observations may begin to look more like the reanalyses. This suggests that even our best estimates of the "truth" can sometimes be limited by sampling frequency in time and space, and therefore, that the high-resolution reanalyses may provide a more realistic and more uniform representation of undersampled ocean variables in the CCLME.

Acknowledgements

We thank two anonymous reviewers for helpful comments that improved the quality of our analysis. We also thank Andrew Moore and Chris Edwards for their helpful insight on this work. This work was supported in part by the NOAA Cooperative Agreement with CIRES, NA17OAR4320101. GLORYS output is available at: https://resources.marine.copernicus.eu/products. CCSRA output is available at: https://oceanmodeling.ucsc.edu. ORAS5 output is available at: http://icdc.cen.uni-hamburg.de/thredds/catalog/ftpthredds/EASYInit/oras5/catalog.html. Newport line data were provided by Dr. Kym Jacobson (NOAA-NWFSC) and Jennifer Fisher and Samantha Zeman (OSU-CIMRS). Individual Argo float profiles are available at: http://www.coriolis.eu.org/Data-Products/Data-Delivery. CUGN data can be found at: https://spraydata.ucsd.edu/projects/CUGN/. California shore station data can be found at: https://shorestations.ucsd.edu/shore-stations-data/download-all-data/. Stonewall Bank station data is at: https://www.ndbc.noaa.gov/station_history.php?station=46050. Charleston station data is at: https://data.nanoos.org/erddap/tabledap/NOS_9432780_Met.html. Tide gauge data can be found at: http://ilikai.soest.hawaii.edu/UHSLC/jasl.html. OISSTv2.1 data can be found at: https://psl.noaa.gov/data/gridded/data.noaa.oisst.v2.highres.html. Bottom trawl data can be found

at: https://www.webapps.nwfsc.noaa.gov/data/map. CUGN is available at: https://spraydata.ucsd.edu/climCUGN/.

References

Abrahms, B., Welch, H., Brodie, S., Jacox, M.G., Becker, E.A., Bograd, S.J., Irvine, L.M., Palacios, D.M., Mate, B.R., Hazen, E.L., 2019. Dynamic ensemble models to predict distributions and anthropogenic risk exposure for highly mobile species. Diversity and Distributions 25, 1182–1193. https://doi.org/10.1111/ddi.12940

Amaya, D.J., Bond, N.E., Miller, A.J., Deflorio, M.J., 2016. The evolution and known atmospheric forcing mechanisms behind the 2013-2015 North Pacific warm anomalies. US CLIVAR Variations 14, 1–6.

Amaya, D.J., Jacox, M.G., Dias, J., Alexander, M.A., Karnauskas, K.B., Scott, J.D., Gehne, M., 2022. Subseasonal-to-Seasonal Forecast Skill in the California Current System and Its Connection to Coastal Kelvin Waves. Journal of Geophysical Research: Oceans 127, e2021JC017892. https://doi.org/10.1029/2021JC017892

Balmaseda, M.A., Hernandez, F., Storto, A., Palmer, M.D., Alves, O., Shi, L., Smith, G.C., Toyoda, T., Valdivieso, M., Barnier, B., Behringer, D., Boyer, T., Chang, Y.S., Chepurin, G.A., Ferry, N., Forget, G., Fujii, Y., Good, S., Guinehut, S., Haines, K., Ishikawa, Y., Keeley, S., Köhl, A., Lee, T., Martin, M.J., Masina, S., Masuda, S., Meyssignac, B., Mogensen, K., Parent, L., Peterson, K.A., Tang, Y.M., Yin, Y., Vernieres, G., Wang, X., Waters, J., Wedd, R., Wang, O., Xue, Y., Chevallier, M., Lemieux, J.F., Dupont, F., Kuragano, T., Kamachi, M., Awaji, T., Caltabiano, A., Wilmer-Becker, K., Gaillard, F., 2015. The ocean reanalyses intercomparison project (ORA-IP). Journal of Operational Oceanography 8, s80–s97. https://doi.org/10.1080/1755876X.2015.1022329

Becker, E.A., Forney, K.A., Fiedler, P.C., Barlow, J., Chivers, S.J., Edwards, C.A., Moore, A.M., Redfern, J.V., 2016. Moving Towards Dynamic Ocean Management: How Well Do Modeled Ocean Products Predict Species Distributions? Remote Sensing 8, 149. https://doi.org/10.3390/rs8020149

Breivik, Ø., Mogensen, K., Bidlot, J.R., Balmaseda, M.A., Janssen, P.A.E.M., 2015. Surface wave effects in the NEMO ocean model: Forced and coupled experiments. Journal of Geophysical Research: Oceans 120, 2973–2992. https://doi.org/10.1002/2014JC010565

Brodie, S., Jacox, M.G., Bograd, S.J., Welch, H., Dewar, H., Scales, K.L., Maxwell, S.M., Briscoe, D.M., Edwards, C.A., Crowder, L.B., Lewison, R.L., Hazen, E.L., 2018. Integrating dynamic subsurface habitat metrics into species distribution models. Frontiers in Marine Science 5. https://doi.org/10.3389/fmars.2018.00219

Chang, J.H., Hart, D.R., Munroe, D.M., Curchitser, E.N., 2021. Bias Correction of Ocean Bottom Temperature and Salinity Simulations From a Regional Circulation Model Using Regression Kriging. Journal of Geophysical Research: Oceans 126. https://doi.org/10.1029/2020JC017140

Chavez, F.P., Messié, M., 2009. A comparison of Eastern Boundary Upwelling Ecosystems. Progress in Oceanography 83, 80–96. https://doi.org/10.1016/j.pocean.2009.07.032

Checkley, D.M., Barth, J.A., 2009. Patterns and processes in the California Current System. Progress in Oceanography 83, 49–64. https://doi.org/10.1016/j.pocean.2009.07.028

de Souza, J.M.A.C., Couto, P., Soutelino, R., Roughan, M., 2021. Evaluation of four global ocean reanalysis products for New Zealand waters–A guide for regional ocean modelling. New Zealand Journal of Marine and Freshwater Research 55, 132–155. https://doi.org/10.1080/00288330.2020.1713179

Di Lorenzo, E., Schneider, N., Cobb, K.M., Franks, P.J.S., Chhak, K., Miller, A.J., McWilliams, J.C., Bograd, S.J., Arango, H., Curchitser, E., Powell, T.M., Rivière, P., 2008. North Pacific Gyre Oscillation links ocean climate and ecosystem change. Geophysical Research Letters 35. https://doi.org/10.1029/2007GL032838

Ducet, N., Le Traon, P.Y., Reverdin, G., 2000. Global high-resolution mapping of ocean circulation from TOPEX/Poseidon and ERS-1 and -2. Journal of Geophysical Research: Oceans 105, 19477–19498. https://doi.org/10.1029/2000jc900063

Good, S.A., Martin, M.J., Rayner, N.A., 2013. EN4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates. Journal of Geophysical Research: Oceans 118, 6704–6716. https://doi.org/10.1002/2013JC009067

Gruber, N., Lachkar, Z., Frenzel, H., Marchesiello, P., Münnich, M., McWilliams, J.C., Nagai, T., Plattner, G.K., 2011. Eddy-induced reduction of biological production in eastern boundary upwelling systems. Nature Geoscience 4, 787–792. https://doi.org/10.1038/ngeo1273

Haltuch, M.A., Tolimieri, N., Lee, Q., Jacox, M.G., 2020. Oceanographic drivers of petrale sole recruitment in the California Current Ecosystem. Fisheries Oceanography 29, 122–136. https://doi.org/10.1111/fog.12459

Huang, B., Liu, C., Banzon, V., Freeman, E., Graham, G., Hankins, B., Smith, T., Zhang, H.M., 2021. Improvements of the Daily Optimum Interpolation Sea Surface Temperature (DOISST) Version 2.1. Journal of Climate 34, 2923–2939. https://doi.org/10.1175/JCLI-D-20-0166.1

Huyer, A., Wheeler, P.A., Strub, P.T., Smith, R.L., Letelier, R., Kosro, P.M., 2007. The Newport line off Oregon - Studies in the North East Pacific. Progress in Oceanography 75, 126–160. https://doi.org/10.1016/j.pocean.2007.08.003

Ingleby, B., Huddleston, M., 2007. Quality control of ocean temperature and salinity profiles - Historical and real-time data. Journal of Marine Systems 65, 158–175. https://doi.org/10.1016/j.jmarsys.2005.11.019

Jacox, M.G., Edwards, C.A., Hazen, E.L., Bograd, S.J., 2018. Coastal Upwelling Revisited: Ekman, Bakun, and Improved Upwelling Indices for the U.S. West Coast. Journal of Geophysical Research: Oceans 123, 7332–7350. https://doi.org/10.1029/2018JC014187

Jacox, M.G., Fiechter, J., Moore, A.M., Edwards, C.A., 2015. ENSO and the California current coastal upwelling response. Journal of Geophysical Research: Oceans 120, 1691–1702. https://doi.org/10.1002/2014JC010650

Jacox, M.G., Moore, A.M., Edwards, C.A., Fiechter, J., 2014. Spatially resolved upwelling in the California Current System and its connections to climate variability. Geophysical Research Letters 41, 3189–3196. https://doi.org/10.1002/2014GL059589

Jayne, S.R., Roemmich, D., Zilberman, N., Riser, S.C., Johnson, K.S., Johnson, G.C., Piotrowicz, S.R., 2017. The argo program: Present and future. Oceanography 30, 18–28. https://doi.org/10.5670/OCEANOG.2017.213

Keller, A.A., Wallace, J.R., Methot, R.D., 2017. The Northwest Fisheries Science Center’s West Coast Groundfish Bottom Trawl Survey: History, Design, and Description. NOAA Technical Memorandum 47.

Lee, T., Awaji, T., Balmaseda, M.A., Greiner, E., Stammer, D., 2009. Ocean state estimation for climate research. Oceanography 22, 160–167. https://doi.org/10.5670/oceanog.2009.74

Legeais, J.F., Ablain, M., Zawadzki, L., Zuo, H., Johannessen, J.A., Scharffenberg, M.G., Fenoglio-Marc, L., Joana Fernandes, M., Baltazar Andersen, O., Rudenko, S., Cipollini, P., Quartly, G.D., Passaro, M., Cazenave, A., Benveniste, J., 2018. An improved and homogeneous altimeter sea level record from the ESA Climate Change Initiative. Earth System Science Data 10, 281–301. https://doi.org/10.5194/essd-10-281-2018

Lellouche, J.-M., Greiner, E., Bourdallé Badie, R., Garric, G., Melet, A., Drévillon, M., Bricaud, C., Hamon, M., Le Galloudec, O., Regnier, C., Candela, T., Testut, C.-E., Gasparin, F., Ruggiero, G., Benkiran, M., Drillet, Y., Le Traon, P.-Y., 2021. The Copernicus Global 1/12° Oceanic and Sea Ice GLORYS12 Reanalysis. Frontiers in Earth Science 9. https://doi.org/10.3389/feart.2021.698876

Masina, S., Di Pietro, P., Storto, A., Navarra, A., 2011. Global ocean re-analyses for climate applications. Dynamics of Atmospheres and Oceans 52, 341–366. https://doi.org/10.1016/j.dynatmoce.2011.03.006

Melnichenko, O., Hacker, P., Maximenko, N., Lagerloef, G., Potemra, J., 2016. Optimum interpolation analysis of Aquarius sea surface salinity. Journal of Geophysical Research: Oceans 121, 602–616. https://doi.org/10.1002/2015JC011343

Melnichenko, O., Hacker, P., Maximenko, N., Lagerloef, G., Potemra, J., 2014. Spatial Optimal Interpolation of Aquarius Sea Surface Salinity: Algorithms and Implementation in the North Atlantic. Journal of Atmospheric and Oceanic Technology 31, 1583–1600. https://doi.org/10.1175/JTECH-D-13-00241.1

Moore, A.M., Arango, H.G., Broquet, G., Powell, B.S., Weaver, A.T., Zavala-Garay, J., 2011. The Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilation systems. Part I - System overview and formulation. Progress in Oceanography 91, 34–49. https://doi.org/10.1016/j.pocean.2011.05.004

Neveu, E., Moore, A.M., Edwards, C.A., Fiechter, J., Drake, P., Crawford, W.J., Jacox, M.G., Nuss, E., 2016. An historical analysis of the California Current circulation using ROMS 4D-Var: System configuration and diagnostics. Ocean Modelling 99, 133–151. https://doi.org/10.1016/j.ocemod.2015.11.012

Ottersen, G., Kim, S., Huse, G., Polovina, J.J., Stenseth, N.C., 2010. Major pathways by which climate may force marine fish populations. Journal of Marine Systems 79, 343–360. https://doi.org/10.1016/j.jmarsys.2008.12.013

Pinsky, M.L., Worm, B., Fogarty, M.J., Sarmiento, J.L., Levin, S.A., 2013. Marine taxa track local climate velocities. Science 341, 1239–1242. https://doi.org/10.1126/science.1239352

Renault, L., Molemaker, M.J., Mcwilliams, J.C., Shchepetkin, A.F., Lemarié, F., Chelton, D., Illig, S., Hall, A., 2016. Modulation of wind work by oceanic current interaction with the atmosphere. Journal of Physical Oceanography 46, 1685–1704. https://doi.org/10.1175/JPO-D-15-0232.1

Reynolds, R.W., Chelton, D.B., 2010. Comparisons of daily Sea surface temperature analyses for 2007-08. Journal of Climate 23, 3545–3562. https://doi.org/10.1175/2010JCLI3294.1

Reynolds, R.W., Chelton, D.B., Roberts-Jones, J., Martin, M.J., Menemenlis, D., Merchant, C.J., 2013. Objective determination of feature resolution in two sea surface temperature analyses. Journal of Climate 26, 2514–2533. https://doi.org/10.1175/JCLI-D-12-00787.1

Reynolds, R.W., Smith, T.M., Liu, C., Chelton, D.B., Casey, K.S., Schlax, M.G., 2007. Daily high-resolution-blended analyses for sea surface temperature. Journal of Climate. https://doi.org/10.1175/2007JCLI1824.1

Rudnick, D.L., Zaba, K.D., Todd, R.E., Davis, R.E., 2017. A climatology of the California Current System from a network of underwater gliders. Progress in Oceanography 154, 64–106. https://doi.org/10.1016/j.pocean.2017.03.002

Sakov, P., Counillon, F., Bertino, L., Lister, K.A., Oke, P.R., Korablev, A., 2012. TOPAZ4: An ocean-sea ice data assimilation system for the North Atlantic and Arctic. Ocean Science 8, 633–656. https://doi.org/10.5194/os-8-633-2012

Santora, J.A., Mantua, N.J., Schroeder, I.D., Field, J.C., Hazen, E.L., Bograd, S.J., Sydeman, W.J., Wells, B.K., Calambokidis, J., Saez, L., Lawson, D., Forney, K.A., 2020. Habitat compression and ecosystem shifts as potential links between marine heatwave and record whale entanglements. Nature Communications 11. https://doi.org/10.1038/s41467-019-14215-w

Schroeder, I.D., Santora, J.A., Mantua, N., Field, J.C., Wells, B.K., Hazen, E.L., Jacox, M., Bograd, S.J., 2022. Habitat compression indices for monitoring ocean conditions and ecosystem impacts within coastal upwelling systems. Ecological Indicators 144, 109520. https://doi.org/10.1016/j.ecolind.2022.109520

Schroeder, I.D., Santora, J.A., Moore, A.M., Edwards, C.A., Fiechter, J., Hazen, E.L., Bograd, S.J., Field, J.C., Wells, B.K., 2014. Application of a data-assimilative regional ocean modeling system for assessing California Current System ocean conditions, krill, and juvenile rockfish interannual variability. Geophysical Research Letters 41, 5942–5950. https://doi.org/10.1002/2014GL061045

Storto, A., Alvera-Azcárate, A., Balmaseda, M.A., Barth, A., Chevallier, M., Counillon, F., Domingues, C.M., Drévillon, M., Drillet, Y., Forget, G., Garric, G., Haines, K., Hernandez, F., Iovino, D., Jackson, L.C., Lellouche, J.M., Masina, S., Mayer, M., Oke, P.R., Penny, S.G., Peterson, A.K., Yang, C., Zuo, H., 2019. Ocean reanalyses: Recent advances and unsolved challenges. Frontiers in Marine Science. https://doi.org/10.3389/fmars.2019.00418

Storto, A., Oddo, P., Cipollone, A., Mirouze, I., Lemieux-Dudon, B., 2018. Extending an oceanographic variational scheme to allow for affordable hybrid and four-dimensional data assimilation. Ocean Modelling 128, 67–86. https://doi.org/10.1016/j.ocemod.2018.06.005

Szekely, T., Gourrion, J., Pouliquen, S., Reverdin, G., 2019. The CORA 5.2 dataset: global in-situ Temperature and Salinity measurements dataset. Data description and validation. The CORA 5.2 dataset: global in-situ Temperature and Salinity measurements dataset. Data description and validation 1–20. https://doi.org/10.5194/os-2018-144

Tolimieri, N., Haltuch, M.A., Lee, Q., Jacox, M.G., Bograd, S.J., 2018. Oceanographic drivers of sablefish recruitment in the California Current. Fisheries Oceanography 27, 458–474. https://doi.org/10.1111/fog.12266

Ward, E.J., Barnett, L.A.K., Anderson, S.C., Commander, C.J.C., Essington, T.E., 2022. Incorporating non-stationary spatial variability into dynamic species distribution models. ICES Journal of Marine Science 79, 2422–2429. https://doi.org/10.1093/icesjms/fsac179

Welch, H., Hazen, E.L., Briscoe, D.K., Bograd, S.J., Jacox, M.G., Eguchi, T., Benson, S.R., Fahy, C.C., Garfield, T., Robinson, D., Seminoff, J.A., Bailey, H., 2019. Environmental indicators to reduce loggerhead turtle bycatch offshore of Southern California. Ecological Indicators 98, 657–664. https://doi.org/10.1016/j.ecolind.2018.11.001

Xie, J., Zhu, J., Li, Y., 2008. Assessment and inter-comparison of five high-resolution sea surface temperature products in the shelf and coastal seas around China. Continental Shelf Research 28, 1286–1293. https://doi.org/10.1016/j.csr.2008.02.020

Xue, Y., Balmaseda, M.A., Boyer, T., Ferry, N., Good, S., Ishikawa, I., Kumar, A., Rienecker, M., Rosati, A.J., Yin, Y., 2012. A comparative analysis of upper-ocean heat content variability from an ensemble of operational ocean reanalyses. Journal of Climate 25, 6905–6929. https://doi.org/10.1175/JCLI-D-11-00542.1

Xue, Y., Wen, C., Kumar, A., Balmaseda, M., Fujii, Y., Alves, O., Martin, M., Yang, X., Vernieres, G., Desportes, C., Lee, T., Ascione, I., Gudgel, R., Ishikawa, I., 2017. A real-time ocean reanalyses intercomparison project in the context of tropical pacific observing system and ENSO monitoring. Climate Dynamics 49, 3647–3672. https://doi.org/10.1007/s00382-017-3535-y

Zhu, J., Huang, B., Balmaseda, M.A., 2012. An ensemble estimation of the variability of upper-ocean heat content over the tropical Atlantic Ocean with multi-ocean reanalysis products. Climate Dynamics 39, 1001–1020. https://doi.org/10.1007/s00382-011-1189-8

Zuo, H., Balmaseda, M.A., Mogensen, K., 2017. The new eddy-permitting ORAP5 ocean reanalysis: description, evaluation and uncertainties in climate signals. Climate Dynamics 49, 791–811. https://doi.org/10.1007/s00382-015-2675-1

Zuo, H., Balmaseda, M.A., Tietsche, S., Mogensen, K., Mayer, M., 2019. The ECMWF operational ensemble reanalysis-analysis system for ocean and sea ice: A description of the system and assessment. Ocean Science 15, 779–808. https://doi.org/10.5194/os-15-779-2019

Table 1 Attributes of the three reanalyses used in this study. The acronyms in the table are: First Guess at the Appropriate Time (FGAT); Copernicus Marine Environment Monitoring Service (CMEMS); eXpendable BathyThermograph (XBT); Mechanical Bathythermograph (MBT); Advanced Very High-Resolution Radiometer (AVHRR); Coriolis Ocean database ReAnalysis (CORA); Operational Sea Surface Temperature and Ice Analysis (OSTIA).

Reanalysis	**GLORYS**	**ORAS5**	**CCSRA**
Institute	Mercator Ocean International (consortium)	ECMWF	University of California Santa Cruz
Ocean Model	NEMO	OCEAN5	ROMS
Domain	global	global	California Current System
Horizontal resolution	1/12°	1/4°	1/10°
Levels	50	75	42
Vertical Coordinate	Depth (z)	Depth (z)	Terrain following (ρ)
First level	0.5 m	0.5 m	variable
Atmospheric Forcing	ERA-Interim	ERA-Interim (1979-2015), ECMWF-NWP (2015-present)	ERA40 (1982-87), ERA40+CCMP winds (1987-2001), ERA Interim+CCMP winds (2001-present)
Assimilation Scheme	3D-Var Kalman Filter	3D-Var FGAT with 5 day window	4D-Var
Assimilated Observations	AVHRR: SST, CMEMS: SLA, CORA: In-situ T/S profiles	HadISST2 + OSTIA: SST, AVISO: SLA, EN4: *In situ* T/S profiles with XBT and MBT correction	AVHRR, AMSR-E, and MODIS: SST, AVISO/CMEMS: SSH, EN3: *In situ* T/S profiles for 1980-2010 EN4: *In situ* T/S profiles for post-2010
Archive period	daily/monthly	daily	6 hourly
Period	1993-2019	1979-2018	1980-2019
References	Lellouche et al. 2021	Zuo et al. 2017, 2019	Neveu et al. 2016

Table 2 Description of observations used for comparison with reanalyses.

Data Set	Description	Period	References
OISST v2.1	Merged satellite and *in situ* SST data mapped to a 1/4° grid	1982-present	Reynolds et al. 2007; Huang et al. 2021
OISSS v1.0	Merged satellite SSS data mapped to a 1/4˚ grid	2012-2018	Melnichenko et al. 2016
Nearshore stations	SST from Stonewall Bank (44.7°N, 124.5°W), Charleston (43.3°N, 124.3°W), Trinidad Beach (41.1°N, 124.1°W), Farallons: (37.7°N, 123.0°W, Newport Beach (33.6°N, 117.9°W), Scripps Pier (32.9°N, 117.3°W).	1993-present (with gaps)	https://shorestations.ucsd.edu/shore-stations-data/
Gliders	California Underwater Glider Network (CUGN), 3 sections that extend ~500 km offshore. Surface to 500 m (or near bottom)	2008-2018	Rudnick et al. 2017
Newport line	CTD casts at three locations along the Newport Hydrographic Line extends offshore from Oregon coast at 44.7°N for locations: T1: 124.1°W (25 m) T3: 124.1°W (55 m) T4: 124.3°W (75 m) Bottom depth in parentheses	2008-2018	Huyer et al. 2007
Argo	Profiling floats drift at a depth of 1000 meters over 10 days, then dive to 2000 meters and return to the surface. ~4000 currently deployed over the global ocean with 18,97 profiles in the CCLME during 2002-18.	2002-2018	Jayne et al. 2017
Tide Gauges	Sea level from 9 tide gauges spanning the west coast	Varies	
Bottom Trawls	Bottom temperature measurements from U.S. West Coast Groundfish Bottom Trawl Survey (WCGBTS)	Varies	Keller et al., 2017
AVISO	Satellite altimetry measurements on a 0.25˚ grid	1993-2012	Ducett et al. 2000

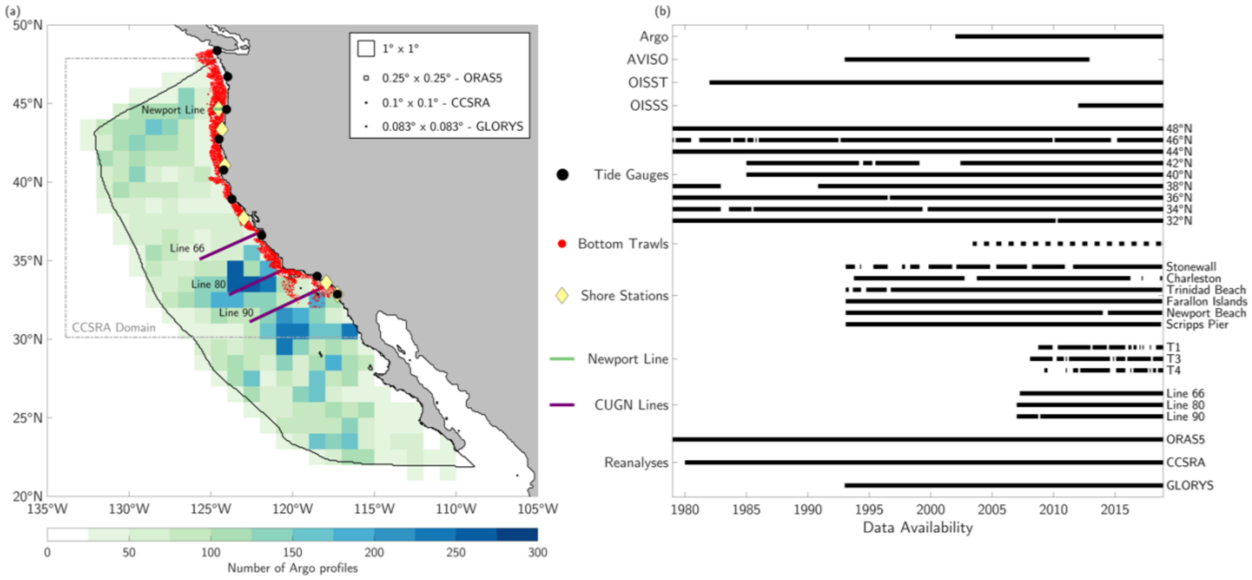

Figure 1 (a) Spatial distribution and (b) temporal availability of all data sets used in this study. Shading in (a) denotes total number of Argo profiles since 2002, binned in 1˚ x 1˚ grid cells. Black line outlines in the CCLME. Dashed gray line denotes the CCSRA regional domain. Black lines for each dataset in (b) indicates that the data was available for the given time step somewhere within the CCLME.

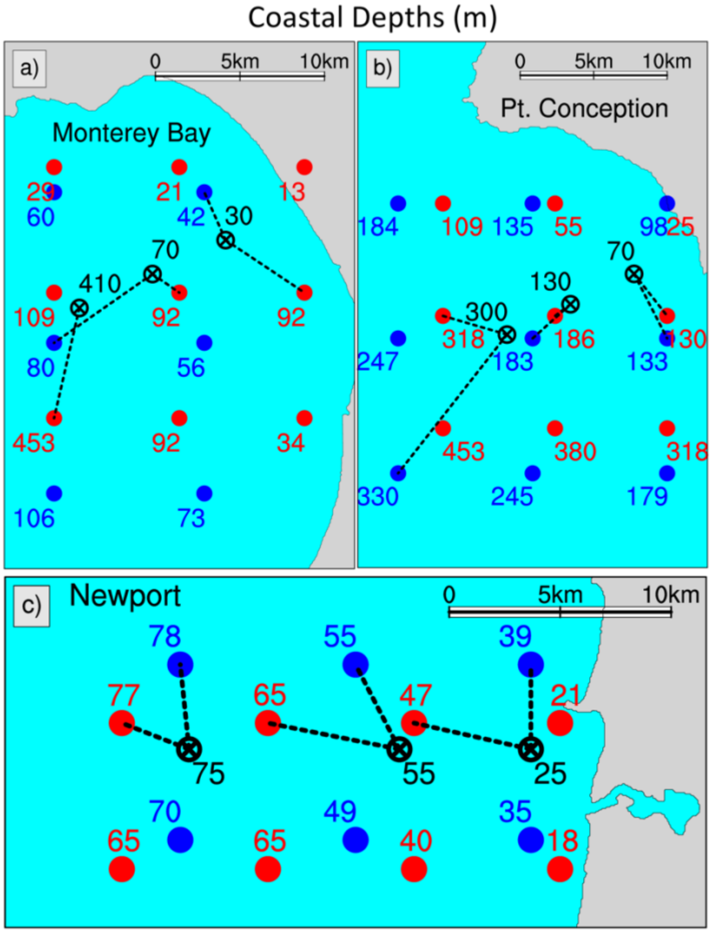

Figure 2 Bottom temperature locations used for comparing the CUGN and Newport Line (black X's) to nearby GLORYS (red dots) and CCSRA (blue dots) grid points in Figure 12. The comparisons are conducted as follows: the nearest reanalysis grid point where the bottom is deeper than at the observed location is selected (dashed line connect the observed and reanalysis data locations), then the temperatures are interpolated from the reanalyses bottom depth to the observed bottom depth. Comparisons are made at depths of 30m, 70m and 410m in Monterey Bay and at 70m, 130m and 300m off Pt. Conception, and at 25m, 55m, 75m on the Newport line. Comparisons between CCSRA and CUGN at 410m in Monterey Bay were omitted since the nearest deeper CCSRA grid points were too far away for a reasonable comparison. The ORAS5 reanalysis was omitted since it is too coarse for this comparison to be meaningful.

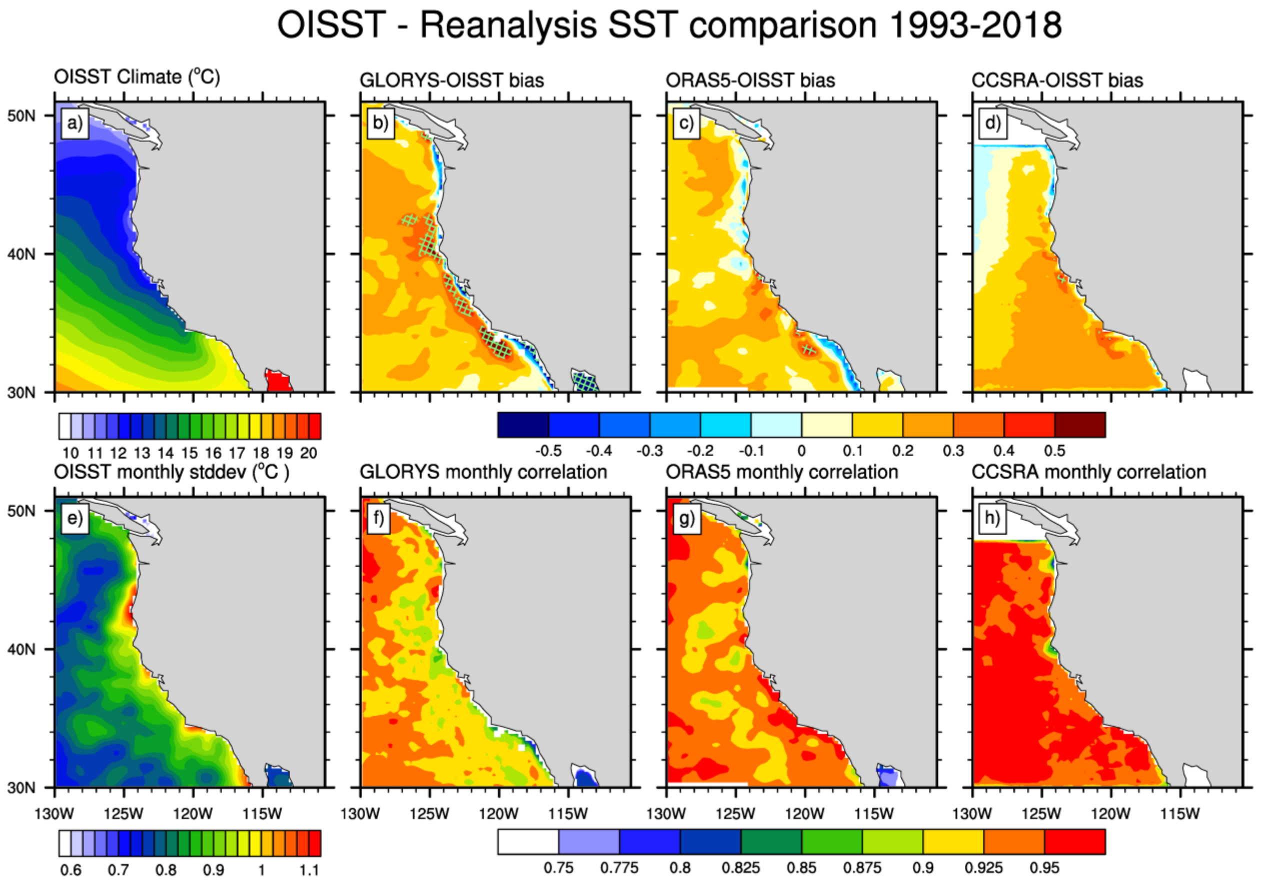

Figure 3 (a) Annual mean SST (˚C) pattern from OISST. (b)-(d) Annual mean SST bias patterns from the GLORYS, ORAS5 and CCSRA reanalyses, respectively. (e) Monthly mean standard deviation pattern (˚C) from OISST. (f)-(h) Anomaly correlation coefficients between monthly mean SST anomalies from OISST and each reanalyses. Significant biases at 95% confidence in the top row are denoted by green hatching. All monthly correlations reported in the bottom row are significant at 95% confidence.

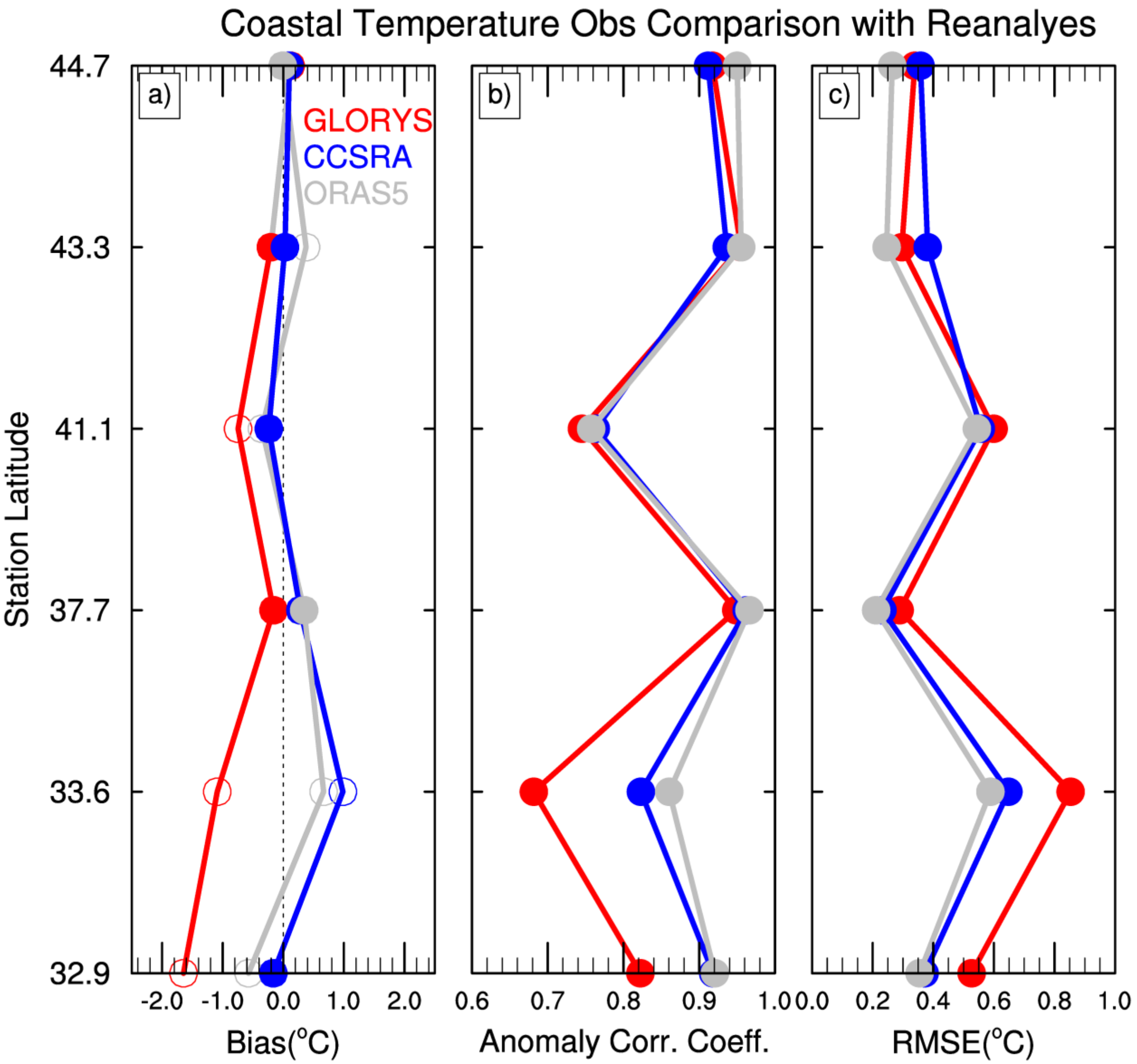

Figure 4 (a) Annual mean SST bias values (˚C) from GLORYS (red), CCSRA (blue), and ORAS5 (grey) relative to the mean SST at each shore station (indicated by latitude). (b) As in (a), but for the anomaly correlation coefficient between monthly mean SST anomalies from each reanalyses and each shore station. (c) As in (a), but for the monthly root mean square error (RMSE; ˚C). Open circles in (a) indicate significant mean biases at 95% confidence. Closed circles in (b) indicate significant correlations with 95% confidence.

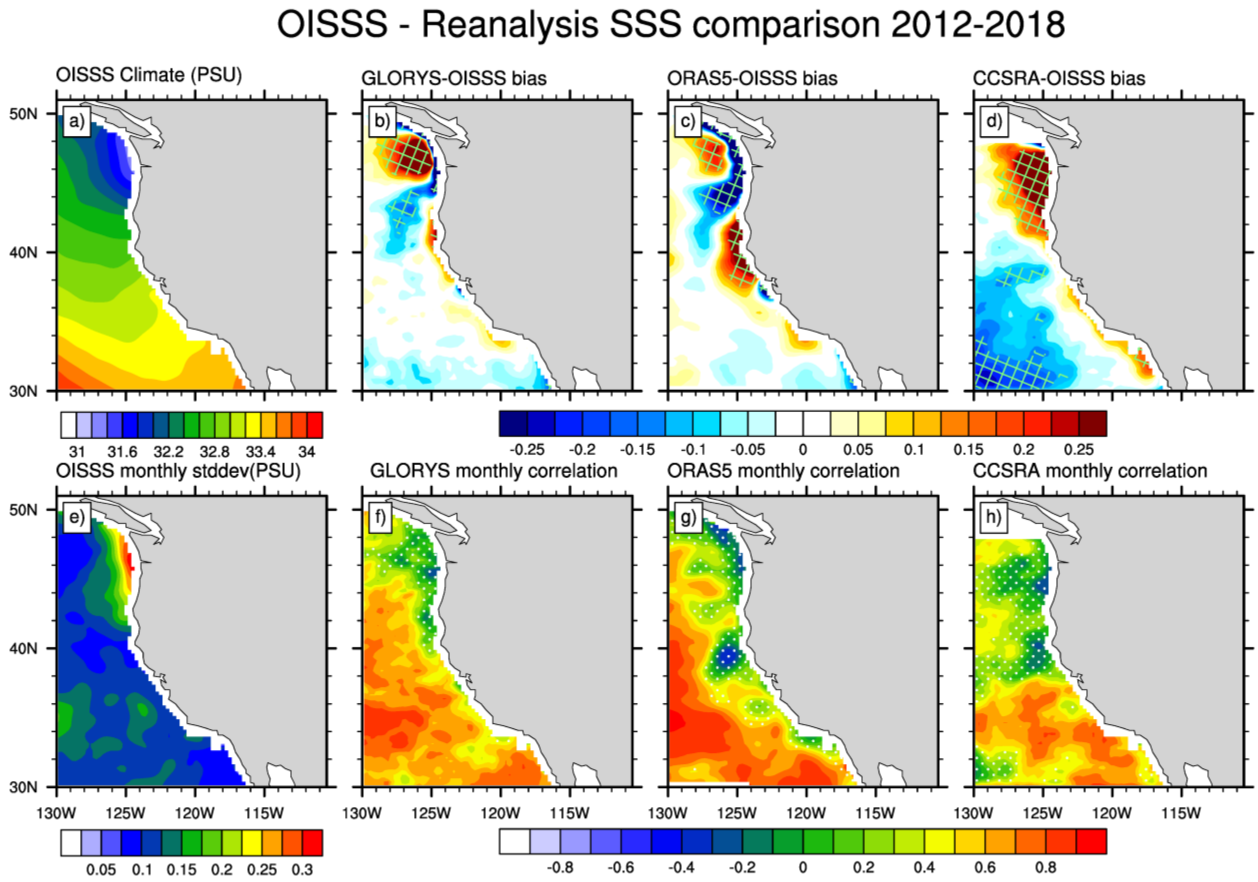

Figure 5 As in Figure 3, but for OISSS and reanalysis sea surface salinity (PSU).

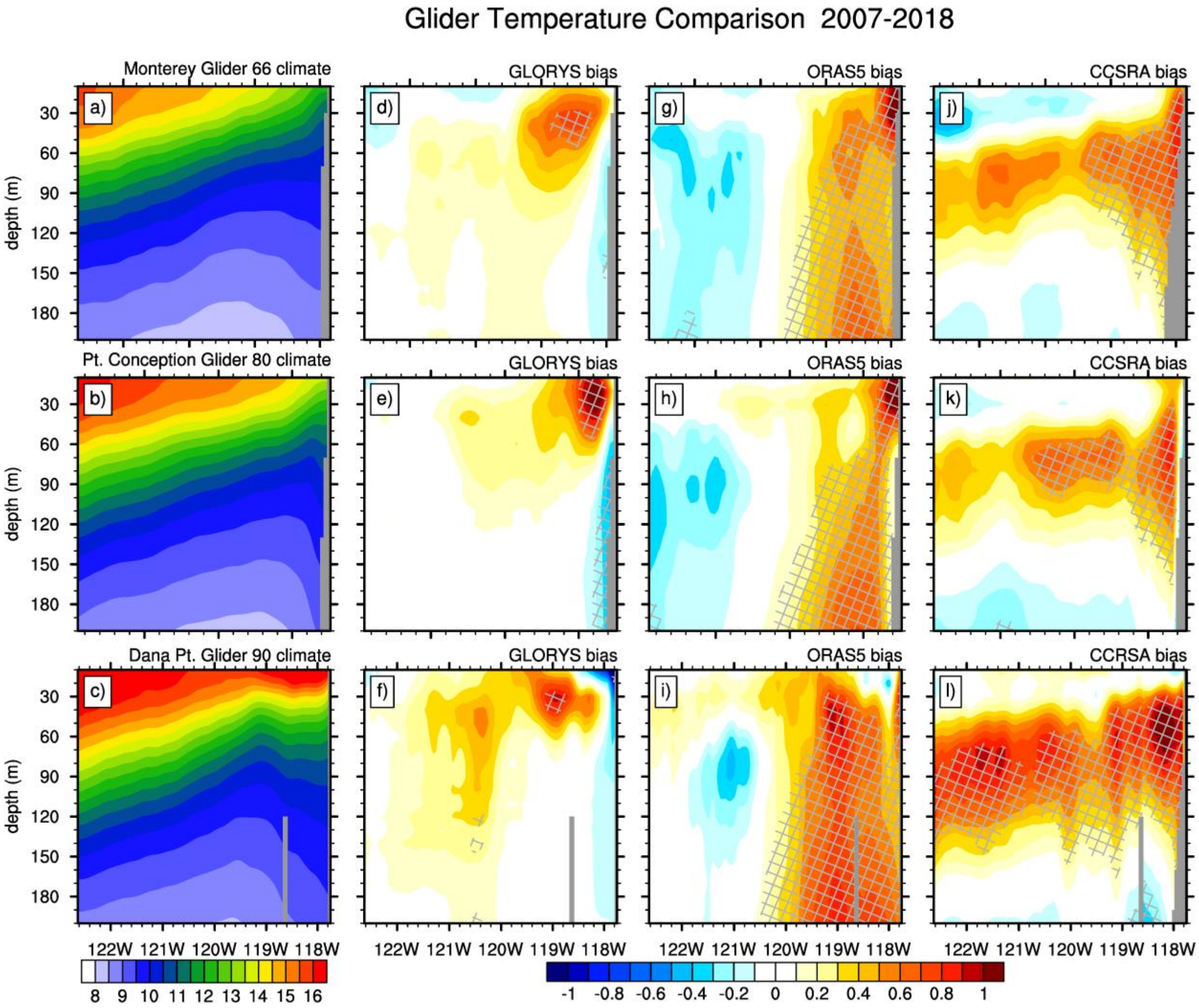

Figure 6 (a)-(c) Annual mean temperature (˚C) along CUGN Line 66 (Monterey), Line 80 (Pt. Conception), and Line 90 (Dana Pt.), respectively. (d)-(l) Annual mean temperature bias (˚C) along each glider line in (d)-(f) GLORYS, (g)-(i) ORAS5, and (j)-(l) CCSRA. Gray hatching indicates a significant mean bias with 95% confidence.

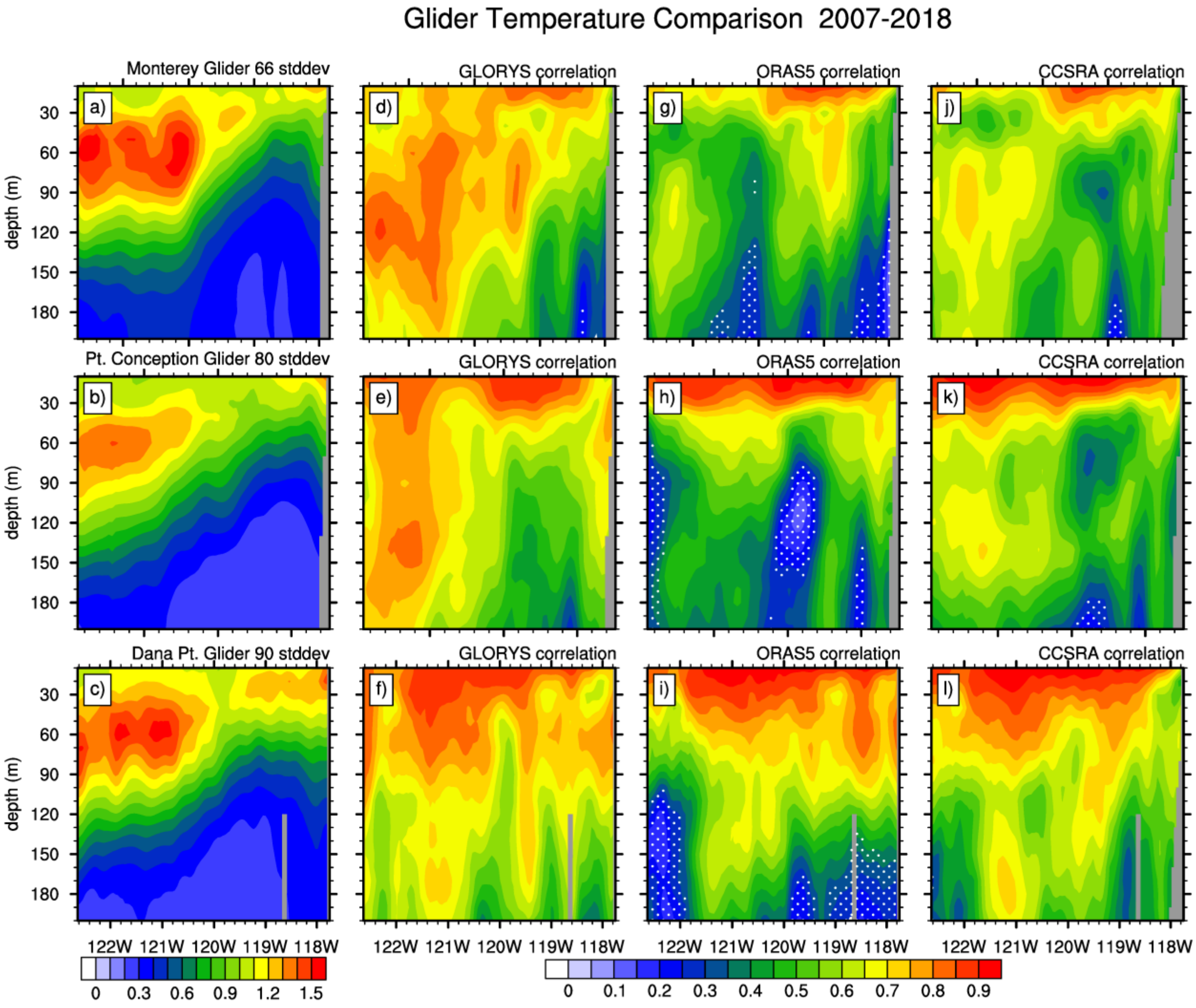

Figure 7 (a)-(c) Monthly mean temperature standard deviation (˚C) along CUGN Line 66 (Monterey), Line 80 (Pt. Conception), and Line 90 (Dana Pt.), respectively. (d)-(l) Anomaly correlation coefficients of monthly mean temperature data from CUGN with (d)-(f) GLORYS, (g)-(i) ORAS5, and (j)-(l) CCSRA. White stipples indicate an insignificant correlation with 95% confidence.

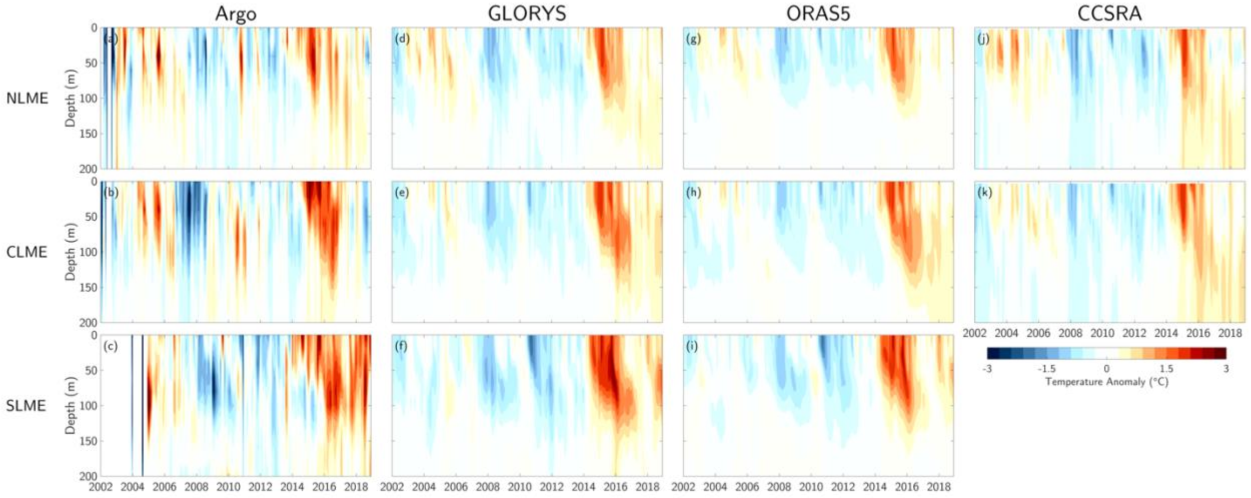

Figure 8 Depth/time cross-sections of monthly mean water temperature anomalies (˚C) averaged in three CCLME sub-regions–the North LME (NLME; top row), Central LME (CLME; middle row), and South LME (SLME; bottom row). Data are from (a)-(c) Argo profiles binned in the vertical in 20 m bins, (d)-(f) GLORYS, (g)-(i) ORAS5, and (j)-(k) CCSRA. Note the reanalysis data feature their native vertical resolution. See Methods for more details.

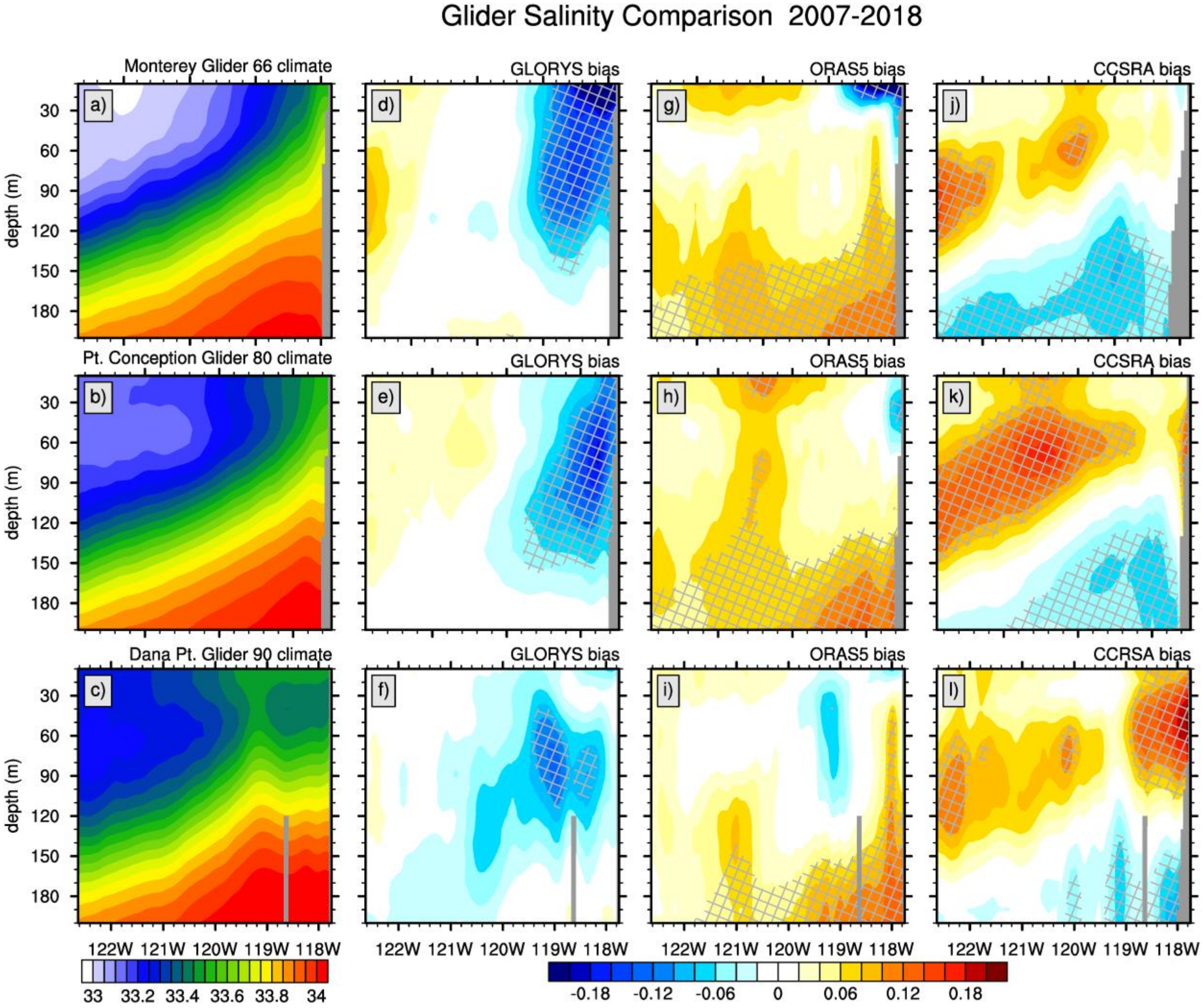

Figure 9 As in Figure 6, but for CUGN and reanalysis water column salinity (PSU).

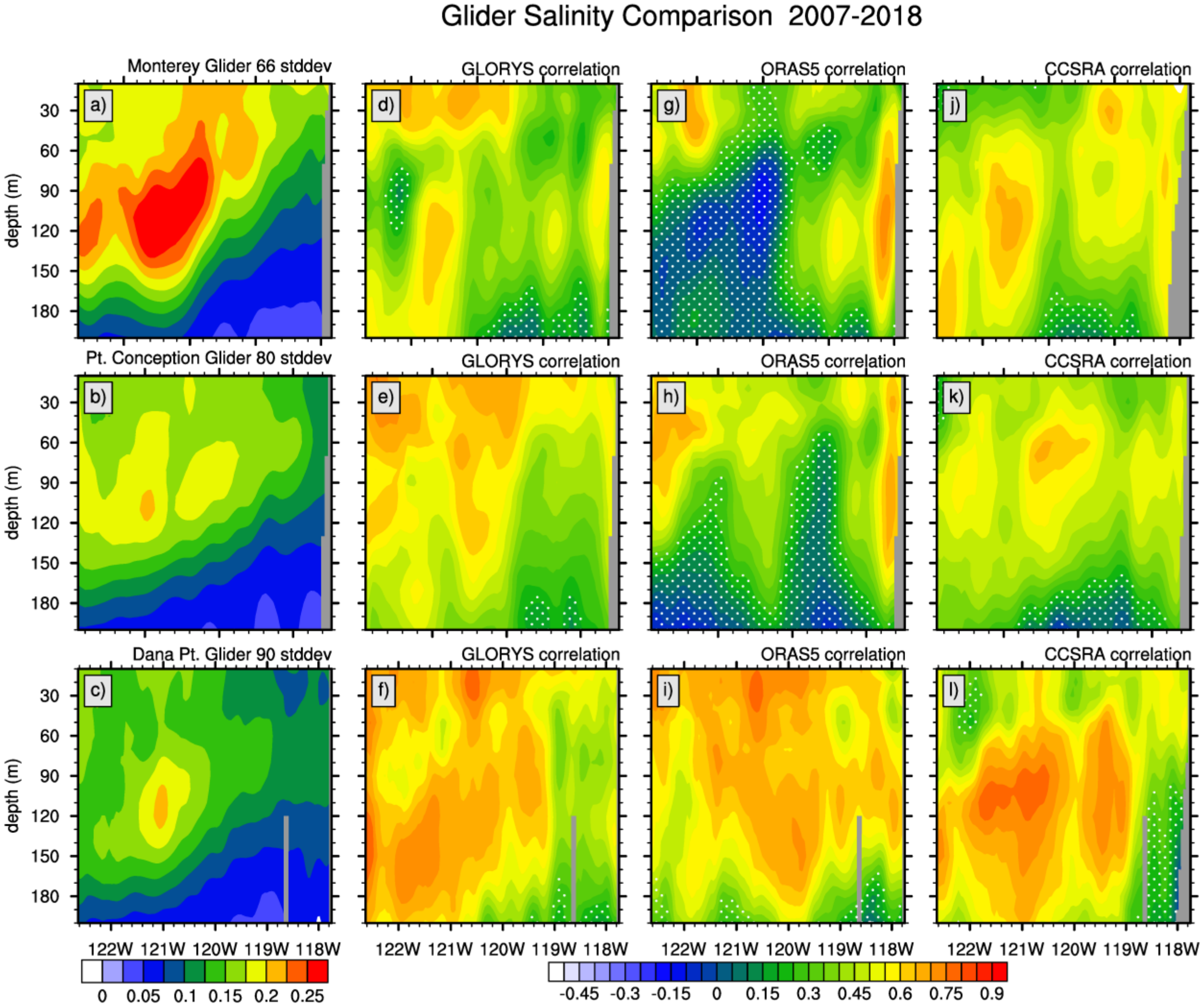

Figure 10 As in Figure 7, but for CUGN and reanalysis water column salinity (PSU).

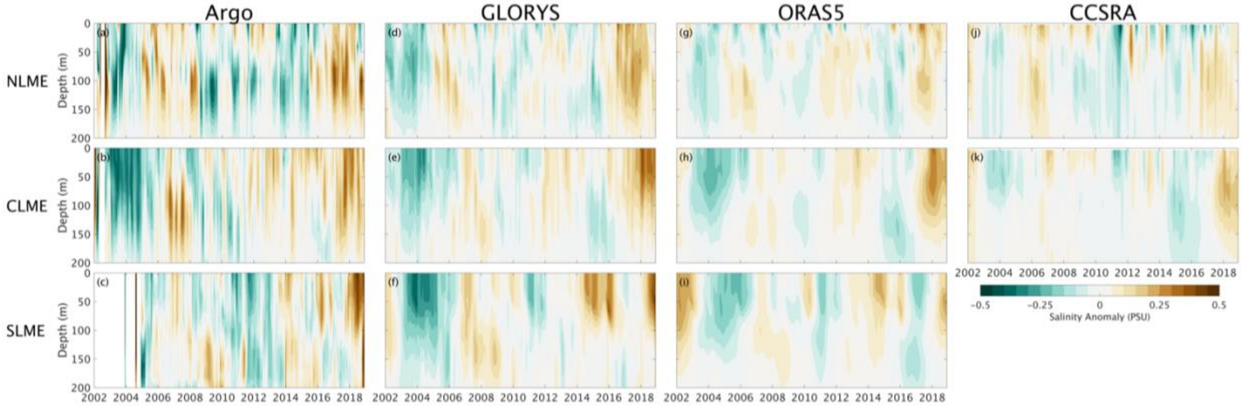

Figure 11 As in Figure 8, but for Argo and reanalysis salinity anomalies (PSU).

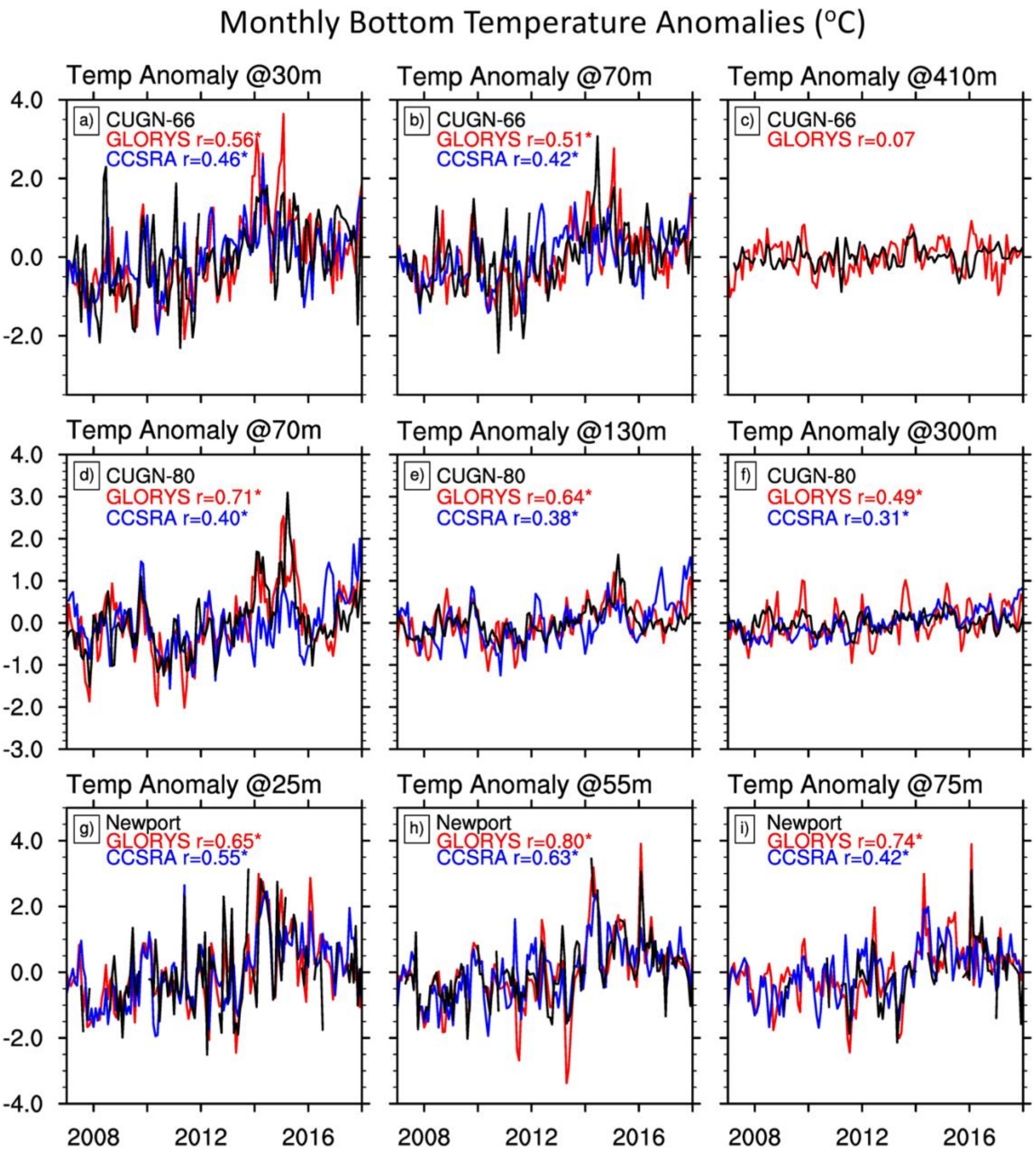

Figure 12 Monthly mean bottom temperature anomaly (˚C) timeseries at select locations along (a)-(c) CUGN Line 66, (d)-(f) CUGN Line 80, and (g)-(i) the Newport Line. In each panel the observational data are in black and data from the nearest GLORYS and CCSRA grid cells are in red and blue, respectively. See Figure 2 for precise data locations. Anomaly correlation coefficients between the observations and each reanalyses are shown in each panel. Asterisks indicate significant correlations with 95% confidence.

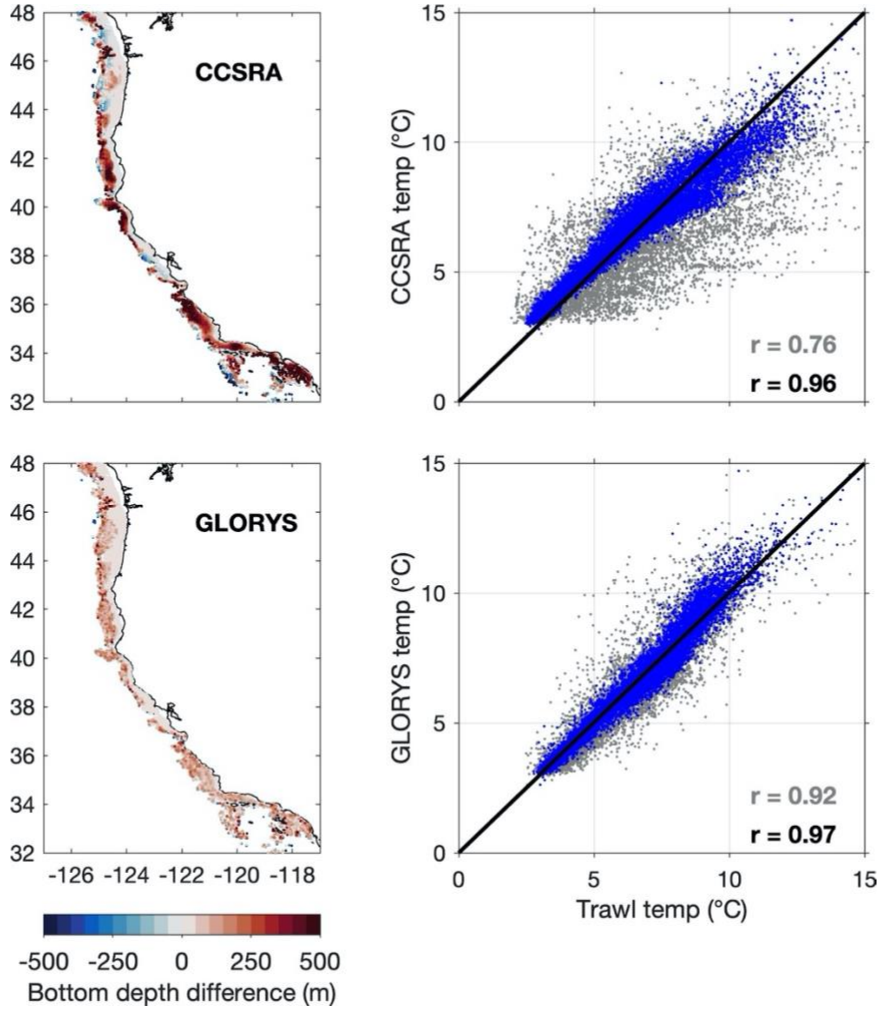

Figure 13 Comparison of CCSRA and GLORYS with bottom temperature data from the west coast groundfish bottom trawl survey. (left) maps indicate differences between trawl depths and model bottom depths at the same location. (right) scatter plots of observed bottom temperatures compared to reanalysis temperatures extracted two ways: (1) at the model bottom (gray), which may be a substantially different depth from the trawl depth, and (2) at the depth of trawl sample (blue). See Methods for more detail.

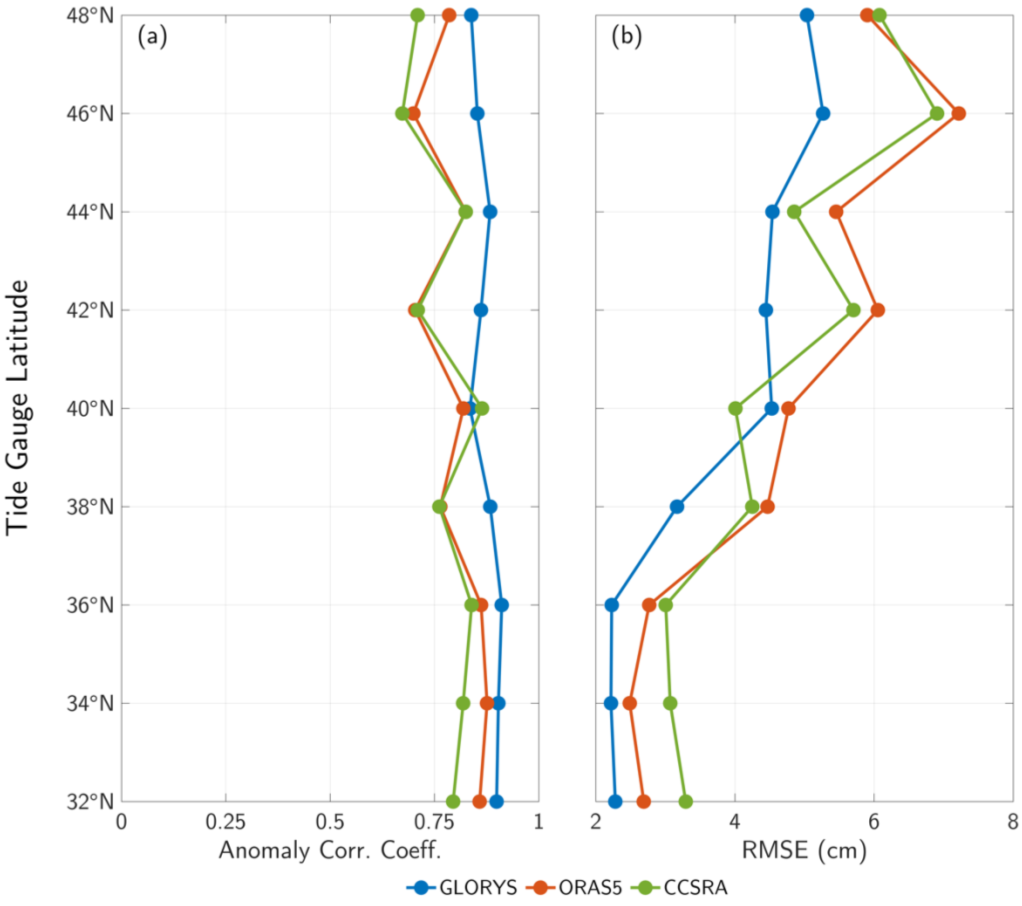

Figure 14 As in Figure 4b and 4c, but for SSH anomaly comparisons between the nine tide gauges and the nearest grid cells in GLORYS (blue), ORAS5 (red), and CCSRA (green).

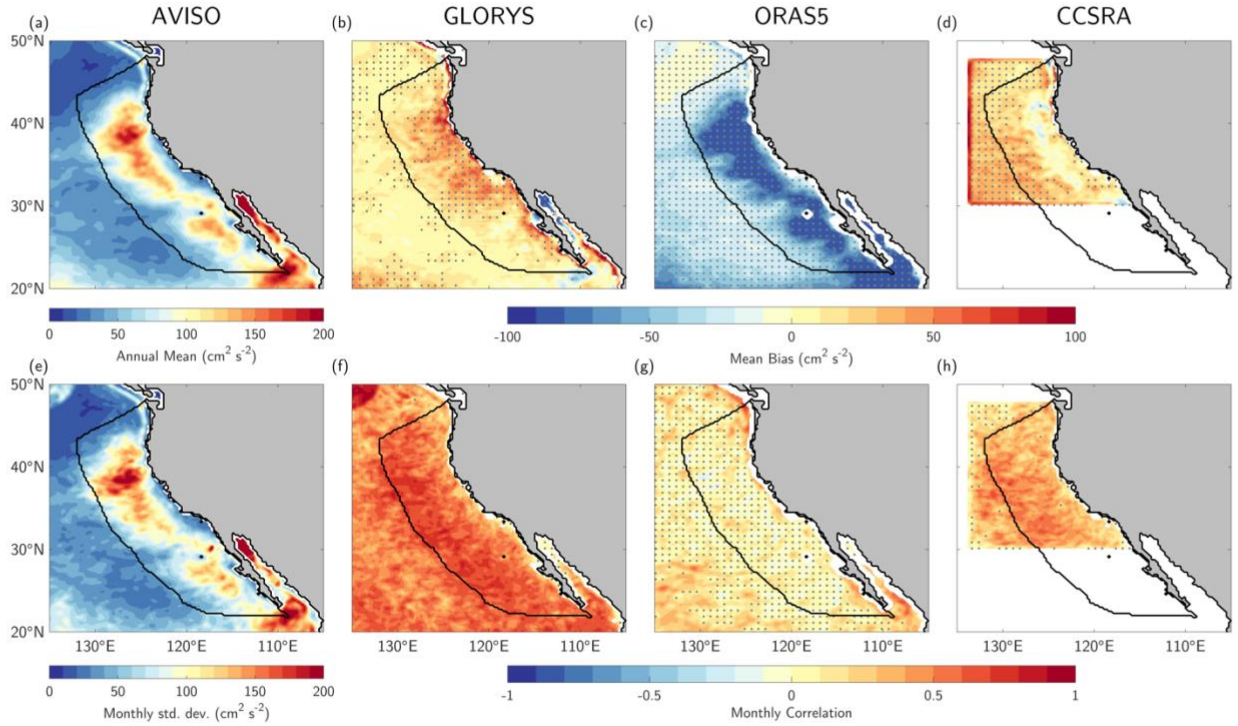

Figure 15 (a) Annual mean geostrophic Eddy Kinetic Energy (EKE; cm^2 s^{-2}) based on SSH anomalies from AVISO satellite observations. (b)-(d) Annual mean EKE biases in GLORYS, ORAS5, and CCSRA, respectively. (e) Monthly mean EKE standard deviation in AVISO. (f)-(h) Anomaly correlation coefficients between monthly mean EKE values from AVISO and each reanalyses. Stipples in (b)-(d) indicate a significant mean bias with 95% confidence. Stipples in (f)-(h) indicate an insignificant correlation with 95% confidence.