

Some Challenges and Advances in Regional Ocean Data Assimilation

Andrew M. Moore

*Department of Ocean Sciences
University of California at Santa Cruz
1156 High Street
Santa Cruz CA 95064 U.S.A.*

1. Introduction

Since the early attempts at ocean data assimilation in the 1980's (e.g. Robinson and Leslie, 1985; Moore et al, 1987) the field has matured considerably, and a diverse array of ocean data assimilation systems now exist. Global ocean data assimilation efforts have been propelled in part by the coordinated efforts of the Global Ocean Data Assimilation Experiment (GODAE), the growing need for seasonal forecasts, as well as growing concerns over climate change. There are a growing number of groups producing global ocean analyses of the circulation of the past (see for example <http://www.godae.org/Ocean-products.html>). In addition, some operational centres routinely generate analyses for current conditions. Global ocean data assimilation, however, presents a considerable challenge because of the size of the inverse problem involved, so the resolution (horizontal and vertical) of global assimilation products is often limited. The highest resolution products currently available are performed on grids with horizontal grid-spacing typically $\sim 1/4$ - $1/6$ degree, so the effective resolutions are 2-3 times lower than this, which is marginal for resolving much of the important mesoscale variability in the open ocean, and certainly inadequate for capturing important circulation features of the coastal regions. For this reason, there has also been a push to develop regional ocean data assimilation systems which utilize higher resolution grids. Two approaches to regional ocean data assimilation are typically used: either the regional model is nested within a global data assimilating model, or the regional model is run stand-alone and boundary condition information is provided by a global assimilating model. While there are clear advantages and disadvantages to both approaches, the stand-alone approach offers greater flexibility since the analyses can be produced using a variety of global circulation estimates as boundary conditions, thus providing a range of uncertainty estimates. Some current regional ocean data assimilation efforts can also be found at <http://www.godae.org/Ocean-products.html>.

While ocean data assimilation efforts have typically lagged developments in meteorology and numerical weather prediction, the methods used in the ocean span a wide level of complexity, ranging from simple nudging, multivariate optimal interpolation, various flavours of Kalman filters and smoothers, variational methods (3D-Var and 4D-Var), and more recently ensemble variational approaches.

The field has now reached a level of maturity where ocean data assimilation circulation estimates are used for a variety of scientific and practical applications. On the scientific side, data assimilation

products have been used to yield insights into the dynamics of the ocean circulation, and shed light on oceanic aspects of climate variability and climate change, particularly in relation to the El Niño Southern Oscillation (ENSO) and the North Atlantic meridional overturning circulation. On the more practical side, data assimilative models have enjoyed wide-spread use in oil spill modeling, particularly in relation to the recent Deep Water Horizon disaster in the Gulf of Mexico, as well as search and rescue efforts, contaminant dispersal, and ocean forecasting.

2. Some Challenges of Ocean Data Assimilation

Some of the challenges currently faced in ocean data are reviewed and discussed in this section, and as we shall see, these are intimately linked to fundamental physical properties of the ocean circulation. To provide a context and a motivation for these and later topics, the focus here will be on a specific region of the ocean circulation, namely the California Current system. However, while the focus will be on regional ocean data assimilation, many of the challenges discussed apply equally to global ocean data assimilation as well.

2.1. The California Current Large Marine Ecosystem (CCLME)

The California Current Large Marine Ecosystem (CCLME) is one of 65 LMEs that have been identified by NOAA and the United Nations Environment Program (UNEP) which collectively account for ~95% of global fisheries biomass (see <http://www.lme.noaa.gov>). The CCLME is one of five LMEs that are subject to seasonal variations in coastal upwelling in which cold, nutrient rich water is brought to the surface, creating conditions that are favourable for high levels of primary productivity, although primary productivity is not as high in the CCLME as it is in other regions.

The subtropical gyre spans the entire Pacific basin from east to west, and the California Current (CC) forms its eastern boundary. The signature of the seasonal coastal upwelling circulation in the CC is illustrated in Fig. 1 which shows satellite images of sea surface temperature (SST) and ocean colour (a proxy for upper ocean Chlorophyll due to photosynthesizing marine organisms) during August, 2005.

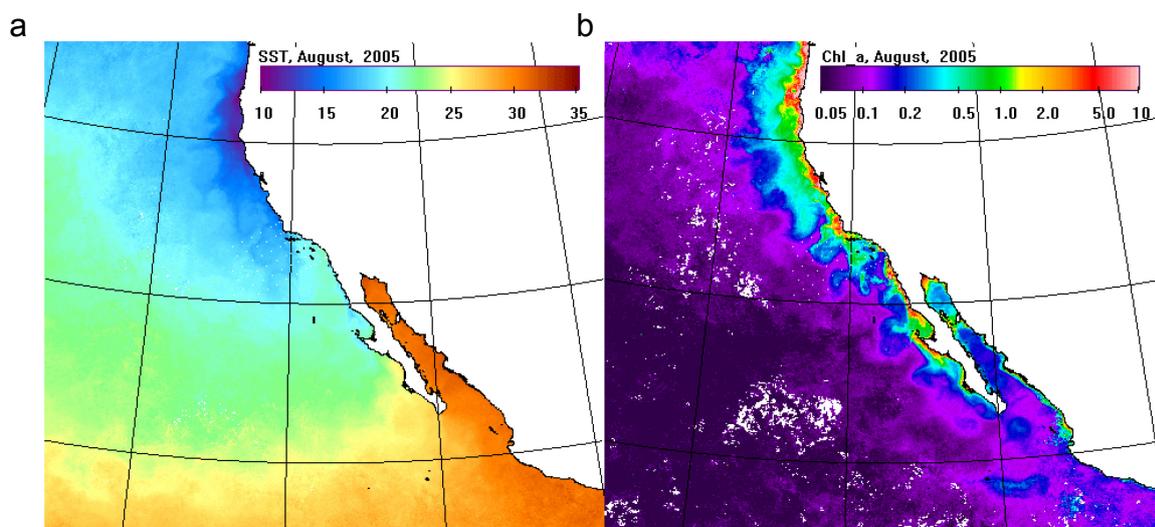


Figure 1: Satellite images of (a) SST and (b) ocean colour from the North East Pacific Ocean during August, 2005.

Despite the prevailing late summer-time conditions, the surface ocean along much of the coast of central California and Oregon is cold (10 C or less, Fig. 1a), and high levels of ocean productivity are evident (Fig. 1b). The cause of these oceanic conditions is well known, and has been extensively studied for several decades (see Hickey (1998) for an excellent review). During the northern hemisphere summer, an atmospheric high pressure system covers much of the North Pacific Ocean, producing equatorward, upwelling favourable winds along most of the west coast of North America. In the surface ocean, the net Ekman transport is offshore which near the coast creates an offshore pressure gradient that in turn enhances the equatorward flowing branch of the subtropical gyre, creating the CC (see Checkley and Barth (2009) for an excellent recent review). As the CC flows toward the equator, it is strongly influenced by the coastal topography and bathymetry such as coastal capes and sea floor escarpments. The CC is also subject to baroclinic instability along parts of its journey, and south of Cape Mendocino (~45N) the circulation is dominated by mesoscale eddies, filaments and jets. The signatures of these circulation features are clearly evident in the ocean colour image of Fig. 1b. The mesoscale features are in turn subject to secondary instabilities as illustrated in Fig. 2 which shows a high resolution satellite SST image in the vicinity of San Francisco Bay and Monterey Bay. The formation of sub-mesoscale circulation features on the mesoscale eddies and filaments is evident. The broad range of space- (and time-) scales that encompass the North Pacific Ocean (from the basin-scale to the sub-mesoscale) has profound implications not only for the circulation but also for the marine ecosystems that inhabit the region.

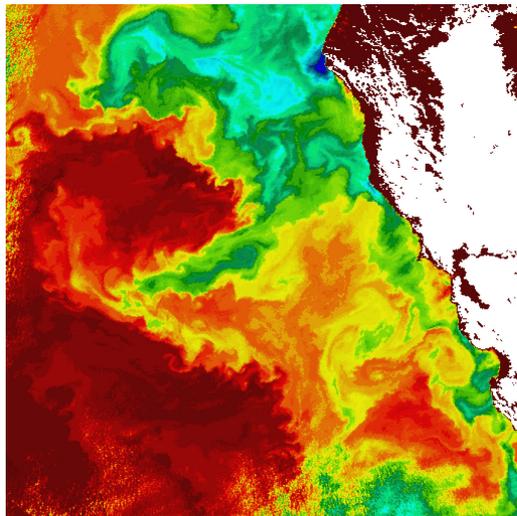


Figure 2: A detailed satellite SST image in the vicinity of San Francisco Bay and Monterey Bay, California.

Coastal upwelling within the CCLME occurs via two different mechanisms, and is illustrated schematically in Fig. 3 from Rykaczewski and Checkley (2007). Nearshore, the divergence of the surface Ekman flow at the coast creates a narrow region of intense upwelling where vertical velocities can be quite large ($O(1-10 \text{ m day}^{-1})$). The width of this upwelling region is governed by the baroclinic Rossby radius of deformation which is ~25-30 km along the California coast (e.g. Chelton et al, 1998). The relatively large vertical velocities in proximity to the rich sediment deposits of the continental shelf yield substantial nutrient fluxes which support a community of large phytoplankton, primarily diatoms. The diatoms are grazed by relatively large species of zooplankton which in turn are the

favoured diet of anchovies. Farther offshore, the strength of the equatorward winds increases creating a region of positive wind stress curl which in turn produces a broad region of upwelling. The wind stress curl-induced upwelling velocities, however, are about an order of magnitude smaller than those close to shore. The weaker upwelling velocities and absence of sedimentary sources of micronutrients in the offshore regions yield significantly lower nutrient fluxes, which in turn support phytoplankton and grazing zooplankton communities comprised of smaller species compared to the more nutrient rich coastal communities. The small zooplankton found offshore are the preferred diet of sardines. Thus, not only is coastal upwelling an important component of the ocean circulation in this region, it also has a profound influence on the distribution and type of lower trophic level marine organisms, and the populations of forage fish that they support. The complex meso- and sub-mesoscale circulations (cf Figs. 1b and Fig. 2) also profoundly influence the marine ecosystems of the CCLME by, for example, exporting nutrients and plankton offshore in the case of coastal jets, and by creating regions of localized upwelling and downwelling at the boundaries of interacting eddies and along frontal regions.

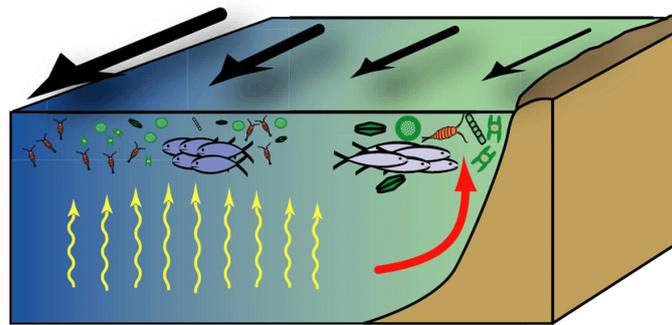


Figure 3: A schematic showing upwelling driven by divergence at the California coast due to offshore Ekman transport by the alongshore equatorward winds (red arrow), and upwelling due to wind stress curl (yellow arrows) farther offshore. The region of strong nearshore upwelling is favored by large phytoplankton, large zooplankton and anchovies (grey fish) while the region of weaker curl driven upwelling is favored by smaller species of phytoplankton and zooplankton which are the preferred diet of sardines (blue fish). (From Rykaczewski and Checkley, 2007).

At higher trophic levels, the sardines and anchovies form a critical part of the diet for apex predators, such as predatory fish, sea birds, and marine mammals. The low species diversity of the forage fish population of the food web comprised of anchovies and sardines is an example of what is commonly referred to as a wasp-waisted food web. Such food webs are particularly sensitive and vulnerable to variability and changes in the ocean circulation, and display other self-limiting behaviours (Bakun, 2006). For example, the circulation changes within the CCLME during El Niño events can have catastrophic effects on the marine ecosystems of the California region. Developing an understanding of the complex interrelationship between the ocean circulation and the marine ecosystems at all trophic levels is of critical importance for effective management and protection of marine communities and resources. Clearly, detailed estimates of the ocean circulation over the broad range of space- and time-scales that characterize the Pacific-wide circulation are required, so ocean observing systems and data assimilation are critical. With this in mind, some of the challenges presented by ocean data assimilation will be considered.

2.2. Observing the Oceans

These days, observations of the ocean are made from a variety of platforms. The largest source of observations are earth orbiting satellites that provide measurements of SST and ocean colour from radiometers, sea surface height (SSH) from altimeters, and more recently sea surface salinity from an instrument aboard the Aquarius satellite. However, probably the single most important addition to the ocean observing network during the last decade is the network of Argo profiling floats. The Argo floats drift freely with the ocean circulation, and via a clever design are able to change their density by increasing their volume thus enabling them to descend through the water column to depths of upto 2000m. As they rise and descend they record profiles of ocean temperature and salinity which are broadcast to ground receiving stations via satellite when the floats are at the surface. The time and distance between broadcast locations also provides information about the average ocean currents. Each Argo float measures a profile about once every 10 days, and each float is designed to make about 150 profiles during its operational lifetime. At the present time, there are approximately 3000 Argo floats spanning the world ocean (see Fig. 4).

More traditional ocean observing platforms such as moorings, buoys and ships, the mainstay of ocean observing networks for many decades, also form an important component of the global ocean observing network. For example, the TAO/TRITON array in the tropical Pacific, the PIRATA array in the tropical Atlantic, and the more recent RAMA array in the Indian Ocean, form a critical component of tropical ocean circulation monitoring, and in the case of the tropical Pacific provide crucial early indicators of approaching ENSO events. During the 1970s and 1980s, the oceanographic community relied heavily on the volunteer observing ship program, and in many remote parts of the global ocean volunteer observing ships continue to provide valuable data.

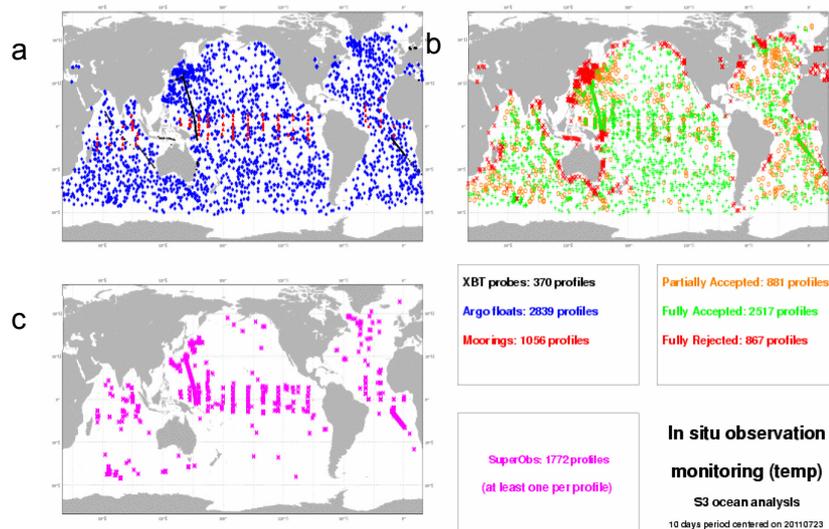


Figure 4: A typical 10 day sample of sub-surface hydrographic observations that are used in the ECMWF global ocean analysis system. All available observations are shown in (a), while those observations that are accepted and rejected after quality control and screening are indicated in (b). The locations of super observations are shown in (c).

In coastal regions, new and more exotic technologies are now being used to routinely observe the ocean. For example, high frequency (HF) radars positioned along the coast can be used to estimate near shore surface ocean currents via the radar backscatter from the rough sea surface. The resolution of the measurements and maximum range offshore are a function of the radar frequency. A comprehensive array of coastal HF radars are now in place along much of the west coast of the U.S. (Kim et al, 2011). In addition, advances in technology have also led to a revolution in the use of gliders and other autonomous under water vehicles (AUVs) for making a wide range of ocean measurements. Gliders and AUVs are extremely versatile, and can make measurements in locations (e.g. in shallow water or close to hazards) and conditions (e.g. under sea ice) not possible from ships, and at a fraction of the cost. AUVs can also be preprogrammed for specific missions, and recently one such vehicle become the first to successfully cross the North Atlantic Ocean (<http://smithsonianscience.org/2010/12/rutgers-glider-to-be-added-to-the-collections-of-the-natural-history-museum/>).

Tagging of marine mammals (e.g. seals, sea lions and whales), large predatory fish (e.g. tuna and sharks), and sea birds (e.g. albatross) also provides a wealth of information about the state and health of the ocean, both directly and indirectly. For example, as part of the Tagging of Pacific Predators (TOPP) program, adult elephant seals are fitted with tags that measure ocean temperature. Adult elephant seals spend much of the year at sea, diving upto 1000m in search of food. Each dive of a tagged individual therefore provides a vertical profile of ocean temperature which is broadcast to a receiving station onshore via satellite (<http://topp.org>). The TOPP program has been very successful, and has provided a wealth of useful oceanographic data, in some cases in areas of the world ocean that have never been observed before.

Figure 4 shows a typical 10 day sample of hydrographic observations that are used in the ECMWF 3D-Var global ocean analysis system. The extent of the Argo profiling floats is clearly evident in Fig. 4a along with the PIRATA, TAO/TRITON and RAMA mooring arrays mentioned earlier. Several expendable bathythermograph (XBT) lines from volunteer observing ships are also present during this period. It is important to realize that while Fig. 4a gives the impression that the global oceans are well sampled on 10 day timescales, there are still relatively large areas devoid of subsurface observations, and the spatial resolution of the array is fairly coarse and certainly does not resolve mesoscale or sub-mesoscale circulation features like those illustrated in Figs. 1 and 2. In addition, the hydrographic data illustrated in Fig. 4a extends only to a maximum depth of 2000 m. The deep ocean below 2000 m is essentially unobserved. Figure 4b gives an indication of the observations that are accepted and rejected based on quality control checks, while Fig. 4c shows which profiles were used to form super-observations in cases where the vertical resolution of the observations is higher than that of the model.

As noted earlier, the lion's share of ocean observations are made by Earth orbiting satellites. However, while satellites provide a wealth of information about the surface conditions of the ocean, they provide no direct information about the subsurface circulation. A great deal of effort has therefore gone into exploring ways in which surface information from satellites can be projected in the vertical to provide dynamically consistent information about the subsurface density structure. One method that has been very influential is that proposed by Cooper and Haines (1996), and is widely used in many global and regional ocean data assimilation systems. Away from the ocean surface, fluid columns conserve potential vorticity and, as such, are largely constrained to follow isopycnal surfaces. Using the hydrostatic balance relation, Cooper and Haines astutely observed that this important property of the subsurface ocean circulation can be respected if each isopycnal is displaced

by an amount $\Delta h = \Delta p_s / (g\Delta\rho)$ where Δp_s is the change in surface pressure deduced from measurements of SSH, and $\Delta\rho$ is the density difference between the top and bottom of the water column. Information from SST observations can also be incorporated to account for the presence of the ocean mixed layer.

Other methods for projecting satellite surface observations into the vertical water column based on EOF and SVD analyses are also commonly used. In this case, historical data from the model and observations are used to compute the EOFs of temperature and salinity in the vertical. The principal component time series of the leading EOFs are then regressed onto the surface observations (SST and SSH). The regression coefficients are then used to compute pseudo-subsurface observations that are assimilated during subsequent analysis cycles. Some representative examples of this approach are describe by Dobricic et al (2005) for the Mediterranean Forecasting System (MFS) and by Fox et al (2002) for the U.S. Navy Modular Ocean Data Assimilation System (MODAS).

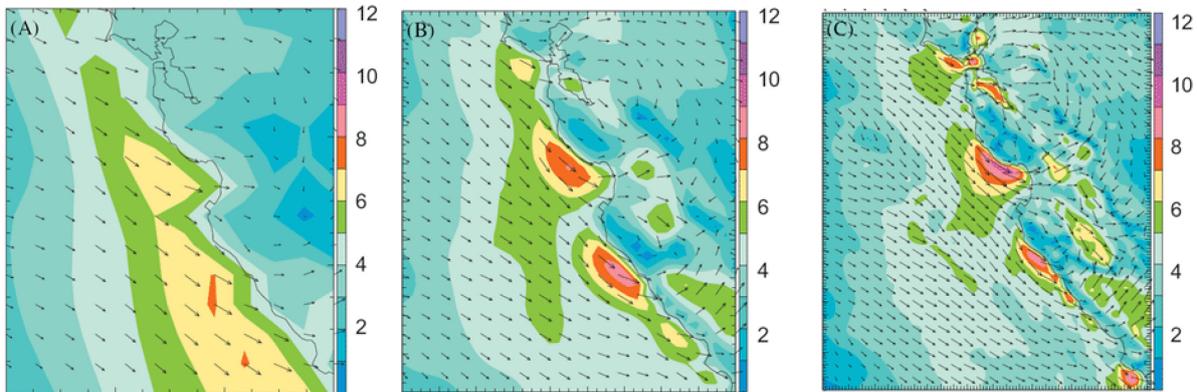


Figure 5: Vectors of surface wind direction and wind speed from the Naval Research Laboratory's Coupled Ocean Atmosphere Mesoscale Prediction System (COAMPS) at (a) 27 km, (b) 9 km, and (c) 3 km horizontal resolution. (From Doyle et al, 2009).

2.3. The Ocean Control Vector

It goes without saying that the ocean circulation is strongly controlled by the surface forcing. This is particularly exemplified in coastal regions where coastal orography can significantly influence the speed and direction of the ocean surface winds. This is graphically illustrated in Fig.5 for part of the central California coast in the vicinity of San Francisco Bay and Monterey Bay. Figure 5, from Doyle et al. (2009) shows the wind speed and direction from the Naval Research Laboratory's Coupled Ocean Atmosphere Mesoscale Prediction System (COAMPS) run at three different horizontal resolutions. At the coarsest resolution (27 km), the upwelling favorable along shore winds are fairly weak and occupy a broad swath along the coast (Fig. 5a). As the resolution is increased to 9 km, the coastal orography of the adjacent land boundary is better resolved, and localized maxima in the alongshore, upwelling winds develop in the lee of Point Reyes, Año Nuevo, and Point Sur (Fig. 5b). At higher resolution still, 3 km in this case, the orographically controlled wind maxima are further enhanced (Fig. 5c). In addition to producing stronger localized winds, the local wind stress curl is also enhanced which contributes further to coastal upwelling. Ocean colour images from satellites confirm that the regions of orographically enhanced winds in Fig. 5c are associated with enhanced primary production, as one would expect. Therefore, in order to capture the mesoscale regions of coastal

upwelling and so explore their impact on local ecosystems, it is important to have high resolution surface forcing with which to drive ocean models. Furthermore, the ocean and atmosphere are significantly coupled, even at the ocean mesoscale (Chelton et al., 2007) which further highlights the need for high resolution coupled models over the ocean. In general, however, surface forcing fields with adequate horizontal resolution are not available, which can seriously hamper efforts to assimilate ocean observations. Therefore, in order to provide additional degrees of freedom for fitting an ocean model to observations, the surface boundary conditions are often included as a control variable in variational data assimilation systems.

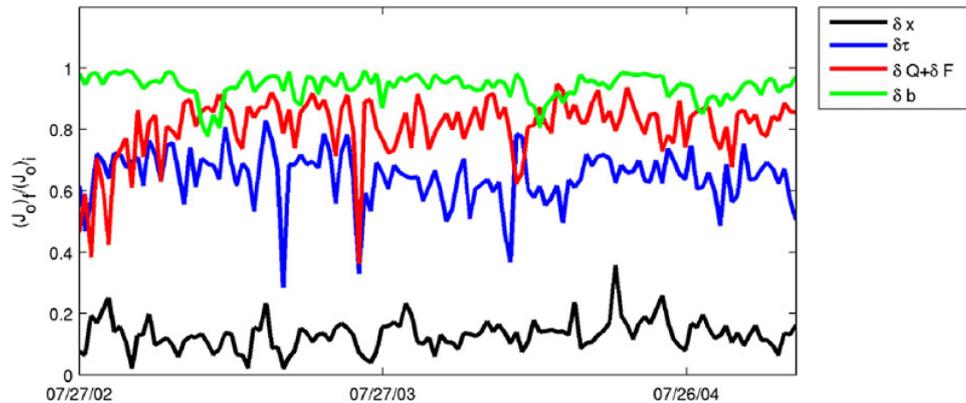


Figure 6: Time series of ratio of the final to the initial value of the observation cost function term, J_o , for four different sequences of 4D-Var calculations. In one case (black curve) the control vector was comprised only of the initial conditions. In a second experiment (blue curve) the control vector was comprised only of the surface wind stress components. In the last two cases, the control vector was comprised of the surface heat flux and freshwater flux (red curve), and only the open boundary conditions (green curve). (From Moore et al, 2011b).

To illustrate the importance of augmenting the data assimilation control vector with the surface forcing, Fig. 6 shows the results from a series of experiments by Moore et al (2011b) using the Regional Ocean Modeling System (ROMS) configured for the CC. More complete details of the data assimilation system used in this case will be presented later in Section 3. At this juncture it is sufficient to note that the data assimilation system used is a 4D-Var scheme run sequentially using 7 day assimilation cycles spanning the period July 2002 – Dec 2004. As described elsewhere in this volume, the 4D-Var cost function can be expressed as $J=J_b+J_o$ where J_b measures departures of the control vector from the *prior* (background) and J_o measures the departure of the model state-vector from the observations. The inverse *prior* (or background) error and observation error covariance matrices are used as the respective weights in J_b and J_o . Figure 6 shows the ratio of the final to the initial value of J_o for each 7 day data assimilation cycle for the CC from four sequences of experiments. In one case, only the model initial conditions were used as elements of the 4D-Var control vector, and in this case J_o is reduced by ~80-90% when data are assimilated. In a second experiment, only the time evolving fields of surface wind stress forcing were used as the elements of the control vector. In this case, the reduction in J_o when data are assimilated is much less dramatic than in the previous case, but nonetheless J_o is reduced by ~30-50%. This illustrates the significant role that surface wind stress plays in the CC on weekly timescales in controlling the ability of the model to fit the observation. The other two experiments shown in Fig. 6 are for the cases when the

4D-Var control vector is comprised solely of either the surface heat and freshwater fluxes or the lateral open boundary conditions. In these cases, the ability of the model to fit the observations is severely limited.

2.4. Prior Error Covariance Modeling of the Ocean

As in meteorology, the specification and modeling of *prior* (or background) error covariances represents a significant challenge in ocean data assimilation. There are several reasons for this, and only a few of them will be discussed here. One primary difficulty, however, arises from the complicated density structure of the ocean which is illustrated in Fig.7. Figure 7 reveals that isopycnals in contact with the atmosphere at middle and high latitudes plunge toward the deep ocean, before shallowing again at subtropical and tropical latitudes. This structure is a direct consequence of potential vorticity conservation (Pedlosky, 1996) in which fluid columns in contact with the atmosphere acquire vorticity through the process of “ventilation” and are then constrained to move equatorward. When the fluid columns encounter fluid of a lower density, they sink (the process of “subduction”) carrying with them potential vorticity which they conserve as they move away from the surface, moving equatorward along isopycnals. This process is so fundamental to shaping the large-scale ocean circulation, that it should be rigorously respected by any data assimilation procedure.

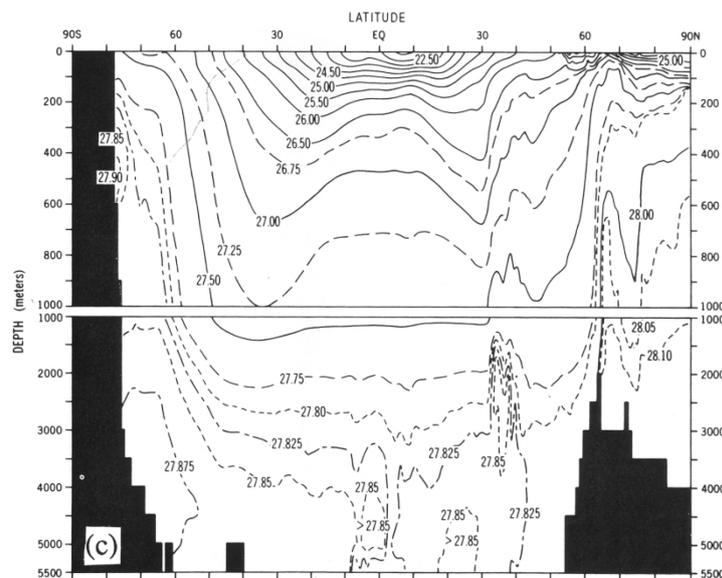


Figure 7: The zonal, global, annual average ocean density structure from Levitus (1982).

The lack of any significant cross-isopycnal flow over much of the ocean complicates any data assimilation procedure, since error correlation functions are expected to largely follow the structure of the isopycnals. This presents a significant challenge, but nonetheless it can be done as illustrated in Fig. 8 adapted from Weaver and Courtier (2001). Figure 8a shows a cross-equatorial temperature section from the central Pacific Ocean, and reveals the complex thermocline structure that is associated with the predominant ocean circulation. Near 10N there is a pronounced dome in the thermocline associated with a region of positive wind stress curl in the vicinity of the Intertropical Convergence Zone. Associated with this dome is the geostrophic, eastward flowing, near-surface North Equatorial Counter Current (NECC). By using a diffusion equation in conjunction with a

diffusion tensor, Weaver and Courtier were able to construct correlation functions that respect the complicated thermocline and circulation structure. An example is shown in Fig. 8b for the case of a single observations located at around 70m depth on the equatorward side of the NECC thermocline dome. These correlation functions were used for both 3D-Var and 4D-Var. However, such calculations are complicated and numerically intensive and, as such, have not yet seen wide spread use by the ocean data assimilation community.

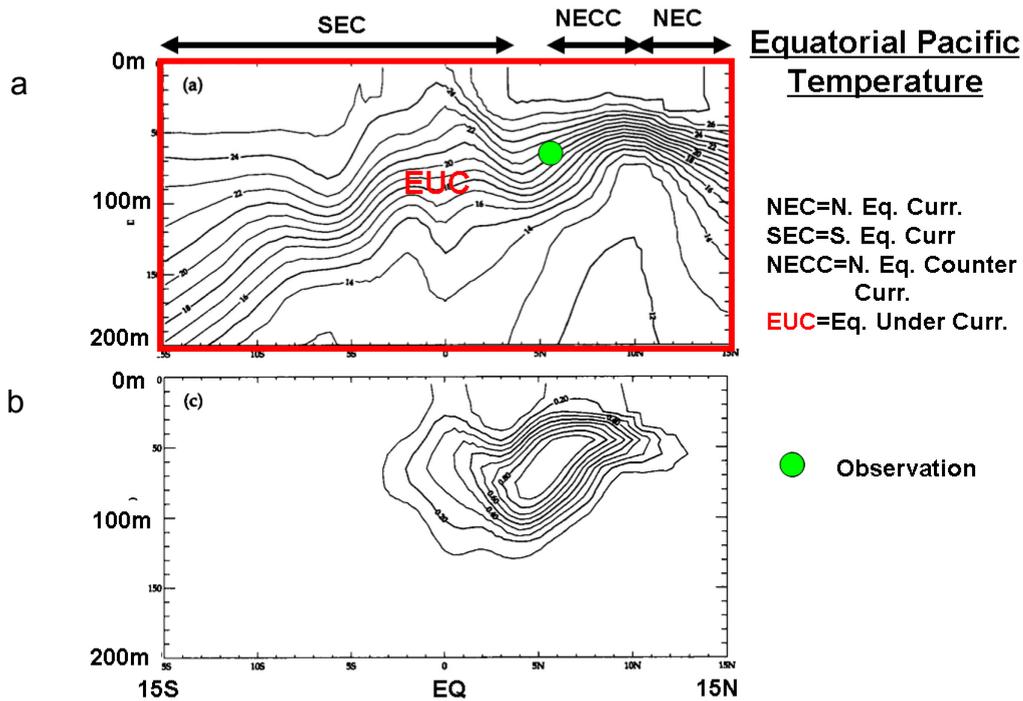


Figure 8: (a) An equatorial cross-section of upper-ocean temperature from the central tropical Pacific showing the dominant circulation features. (b) The correlation function for a single observation at the location indicated, and computed using a diffusion equation with a diffusion tensor. Figure is adapted from Weaver and Courtier (2001).

Another aspect of the ocean circulation that complicates error covariance models in coastal regions is the presence of complex coastlines and bathymetry. The land-sea boundary can support a wide range of coastally trapped ocean waves (LeBlond and Mysak, 1980), and the coastline acts as a waveguide, channeling energy away from the equator along eastern boundaries and toward the equator along western boundaries. This will clearly introduce a significant degree of inhomogeneity into error covariance functions when approaching the coast from the open ocean. One recent attempt to model this transition in a smooth and seamless fashion is illustrated in Fig. 9 from Li et al (2008). In this case, the circulation field was decomposed into a stream function and velocity potential, and a Kronecker product formulation used to locally rotate and stretch the error correlation functions so that they are aligned with the coastline. Figure 9 shows the error correlation function for two single observations: one located in deep water away from the coast, and another on the continental shelf closer to shore. This approach is currently used in a near-real time 3D-Var system.

Ocean tracers can also present challenges in ocean data assimilation since they are bounded by zero and required to be positive quantities. This is similar to the problem encountered in numerical weather prediction in connection with moisture-related variables such as specific humidity and liquid water content. Most currently used data assimilation methods are predicated on the assumption that errors in the state variables and observations that are normally distributed. This condition is violated for strictly positive tracer variables, so it is common to transform the original variable into a new variable that belongs to an approximate normal distribution. In the case of atmospheric humidity, a log-normal distribution is often assumed (e.g. Fletcher, 2010). However, since the variable transformation involved is generally nonlinear (certainly the case for a log-normal transformation), this can complicate the data assimilation procedure, especially in the case of incremental approaches.

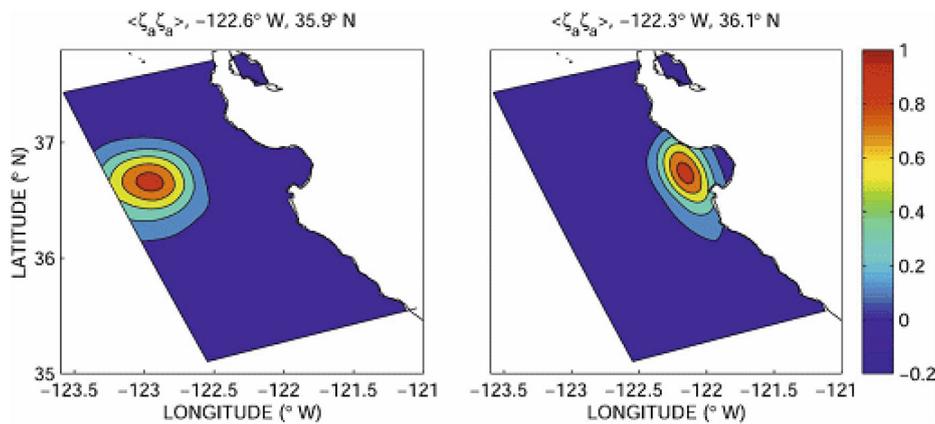


Figure 9: Two realizations of correlation functions for single observations of sea surface height in deep water (left) and near the coast (right) from Li et al (2008).

A very important and dynamically active tracer in the ocean is salinity. The average salinity of the ocean is $\sim 35\text{‰}$, and over much of the ocean the salinity is sufficiently far from zero that there is little chance of negative values developing if a normal distribution is assumed. However, in the coastal regions where there is significant freshwater discharge from rivers and streams, the salinity can be very low, and a different treatment for salinity in data assimilation may be advisable. Other examples of important passive ocean tracers include nutrients, phytoplankton, and zooplankton (cf Fig. 1b), and data assimilation of biological ocean data has been an active area of research for some time (e.g. Friedrichs, 2002). For some tracers, the assumption of a log-normal distribution may be appropriate as suggested by Fig. 10 from Campbell (1995) which shows a log-linear frequency diagram of many years of *in situ* ocean Chlorophyll observations from the northern hemisphere. In other cases, more complex distributions are apparent.

To illustrate, Figs. 11a and 11b from Simon and Bertino (2009) show frequency plots of diatom counts and silicate concentration, the important limiting macro-nutrient for diatoms, from a coupled physical-biological model of the North Atlantic. Diatom counts and silicate are clearly both positive and described by highly non-normal distributions. However, using a technique referred to as interpolated anamorphosis, Simon and Bertino were able to identify empirical non-linear transformations which yield new variables with the distributions shown in Figs. 11c and 11d which are close to Gaussian. These new distributions were used successfully to assimilate data using an ensemble Kalman filter approach.

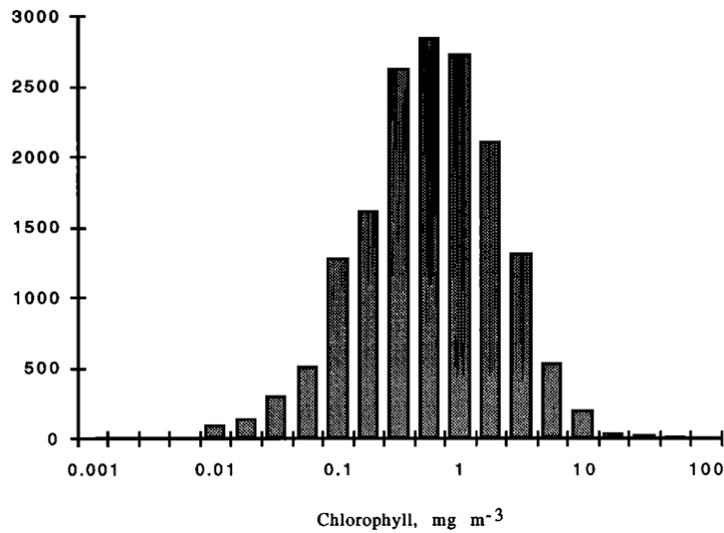


Figure 10: The distribution of in situ chlorophyll observations from the northern hemisphere where the concentration is plotted on a log-scale, from Campbell (1995).

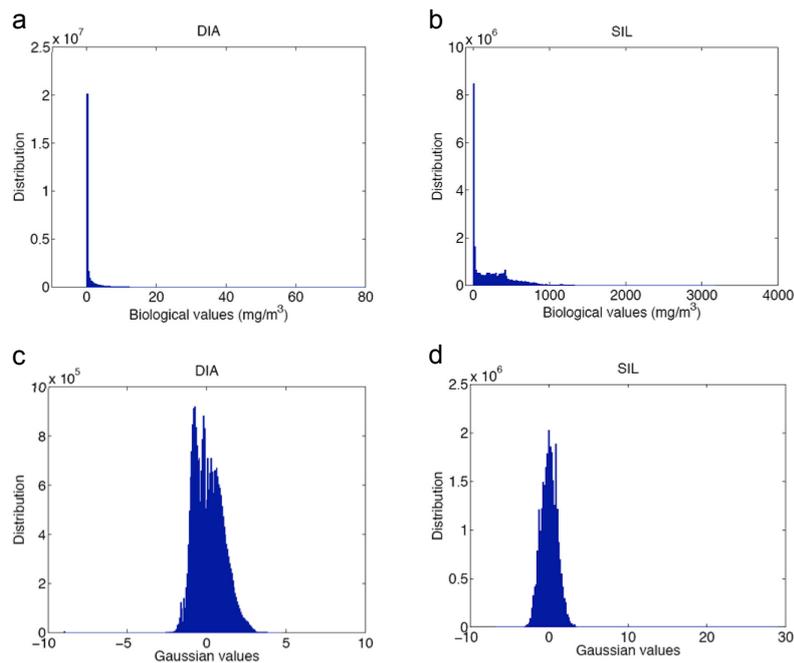


Figure 11: Probability distributions of (a) diatoms, and (b) silicate from a numerical model of the North Atlantic, from Simon and Bertino (2009). The probability distributions for the same variables are plotted in (c) and (d) after an empirically derived interpolated anamorphosis transformation has been applied to the original model data.

2.5. Initialization Shock and Dynamical Filtering

Initialization shock is seldom discussed in relation to ocean data assimilation, probably because the inertia-gravity waves that form part of the dynamical adjustment process are not thought to have an overly detrimental impact on the large-scale ocean circulation. In the tropical oceans and coastal regions, however, the impacts of initialization shock can be more severe because the waveguides in

these regions can channel energy great distances. At the mesoscale and sub-mesoscale, gravity waves can also be an important part of the circulation signal, so differentiating between legitimate and spurious wave modes may be challenge. One aspect of ocean modeling where initialization shock is a serious issue, however, is modeling of ocean ecosystems. As reliable estimates of the ocean circulation become available from data assimilating models, ecosystem modelers have become interested in exploring the impact that these more accurate representations of the physical environment have on the distribution of marine organisms. To illustrate, Fig. 12 from Raghukumar (2011) shows an example of the detrimental influence of inertia-gravity wave shocks if left unchecked on simulations of the lower trophic CC ecosystem.

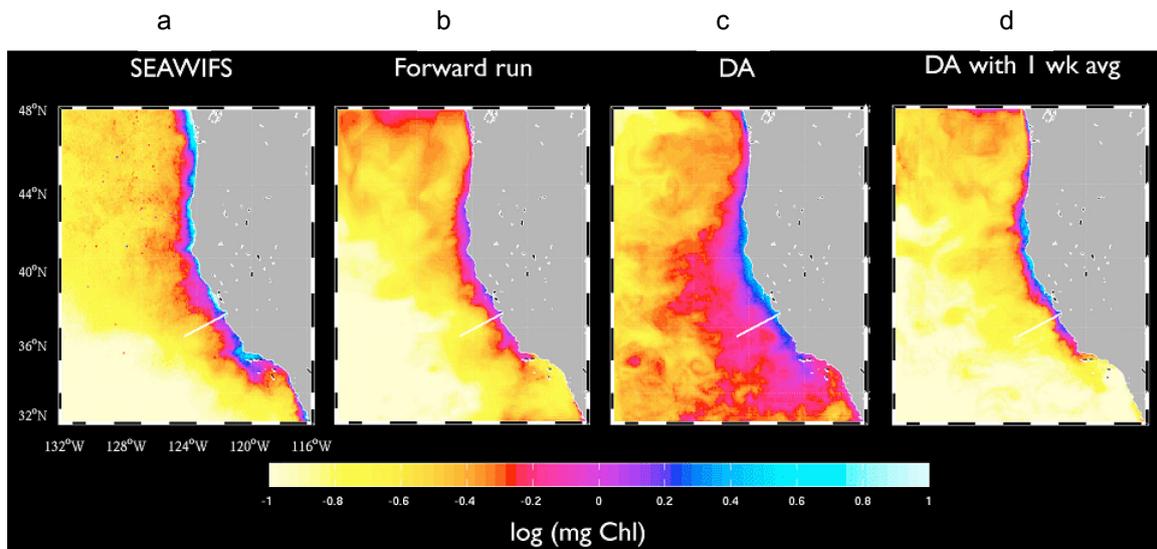


Figure 12: Maps of surface ocean chlorophyll in the California Current system from SeaWiFS satellite observations (a), and from ROMS coupled to a model of the upper trophic ecosystem (b-d). In (b) ROMS is run without data assimilation. In (c) 4D-Var data assimilation is used to reconstruct a best estimate of the physical environment, while in (d) the vertical velocity from data assimilation is low-passed filtered before being used in the ecosystem model. Courtesy of Kaustubha Raghukumar, UC Santa Cruz.

Figure 12a shows observations of surface chlorophyll from the SeaWiFS instrument aboard the SeaStar spacecraft. Elevated levels of chlorophyll pigment are clearly evident along much of the U.S. west coast, indicative of phytoplankton blooms. A simulation of the chlorophyll concentration at the same time from a coupled physical-biological model without data assimilation is shown in Fig. 12b. While the model reproduces some of the qualitative features in the observations, many of the details are misrepresented, and there are clearly issues along the open northern boundary. However, as discussed in section 2.1, the lower trophic marine ecosystems of the CC are intimately linked to the physical environment, so the level of disagreement between Figs. 12a and 12b is perhaps to be expected. By assimilating satellite and hydrographic observations though we would expect to recover some aspects of the observed physical circulation and perhaps more reliably simulate the ecosystem behavior. Figure 12c shows the result of using a 4D-Var analysis of the ocean circulation to drive a new simulation of the CC ecosystem. However, the new simulation is now substantially degraded compared to the original simulation without data assimilation, and Fig. 12c shows unrealistically high levels of surface chlorophyll where none is observed. The culprit here turns out to be fast inertia-

gravity waves that are generated by dynamical imbalances that exist in the model initial conditions resulting from data assimilation. The highly divergent nature of these waves yields large vertical velocities that inject nutrient into the euphotic zone, creating wide spread favorable conditions for phytoplankton blooms throughout the model domain. Clearly then something must be done to mitigate the effects of the initialization shock on the ecosystem model. Figure 12d shows the result of a third model simulation in which a low-pass filter was applied to the vertical velocity field used in the ecosystem model, thus effectively suppressing the influence of the inertia-gravity waves on the biology. Clearly the ecosystem model simulation is much improved as a result of this action.

The example of Fig. 12 suggests that some kind of normal mode initialization or digital filter initialization is warranted in the 4D-Var assimilation scheme used there. However, it must be remembered that the coastal waveguide supports a wide range of coastally trapped waves that are a legitimate part of the ocean response to changes in local and remote forcing, but which may be seriously impacted by the application of any dynamical filter that penalizes the divergent component of the circulation.

2.6. Summary

While the topics covered in this section are by no means exhaustive, they do give a sense of some of the major challenges encountered in ocean data assimilation at all scales. There are, of course, other significant challenges that are not discussed here. These include the assimilation of biogeochemical data into ocean models; data assimilation in the presence of internal tides; data assimilation in the presence of strong air-sea coupling; model errors; quality control of ocean observations; and bias correction.

3. Examples of Some Recent Advances in Regional Ocean Data Assimilation

The impact and sensitivity of observations on atmospheric analyses has received a great deal of attention in numerical weather prediction, and provides powerful tools for quantifying the value of the observing array (see Cardinali, 2011, this volume). Rienecker (2011, this volume) provides a broad overview of the global ocean observing system. Many of the observation platforms described there, however, are under the constant threat of being discontinued, due to budget cuts, so tools similar to those developed in NWP are also required in oceanography to demonstrate the impact and value of the various ocean observing platforms. In this section, some recent advances in regional ocean data assimilation will be presented that explore the impact of the ocean observing array on ocean circulation estimates.

3.1. The Regional Ocean Modeling System (ROMS) and 4D-Var

ROMS is a one example of a regional ocean model and data assimilation system, but is somewhat unique since a suite of observation impact and observation sensitivity tools have been developed as an integral component of the ROMS 4D-Var system. ROMS is primitive equation, hydrostatic, ocean model that employs orthogonal curvilinear coordinates in the horizontal and terrain following coordinates in the vertical (Shchepetkin and McWilliams, 2005). ROMS also includes a broad suite of user-controlled numerical schemes, physical parameterizations, and open boundary options, making it extremely versatile (see <http://www.myroms.org>). In addition, a suite of 4D-Var algorithms have also

been developed for ROMS, one based on the primal formulation, and two based on the dual formulation. All three ROMS 4D-Var algorithms are written in incremental form (Courtier et al, 1994), and the primal form is similar to that described by Weaver et al (2003). The dual formulations of 4D-Var are based on the Physical-space Statistical Analysis System (PSAS, Courtier, 1997) and the indirect representer method of Egbert et al (1994). The entire ROMS 4D-Var system is described in detail by Moore et al (2011a).

In keeping with section 2, examples will be presented here from ROMS configured for the CC. Specifically, Fig. 13 shows the bathymetry and extent of the model domain that was employed in the experiments presented here. Results will be presented from two different configurations of the model: one using 10 km horizontal resolution and 42 σ -levels (here after referred to as CC10), and another using 30 km resolution and 30 σ -levels (here after CC30). Both configurations of the model have been used to generate a sequence of 4D-Var circulation analyses employing 7 day assimilation windows and spanning the period July 2002 – Dec 2004. In each case, the *prior* (background) surface forcing fields were computed from standard 10 m atmospheric variables from the Naval Research Laboratory COAMPS model (Doyle et al, 2009) using the bulk formulation of Liu et al (1979) and Fairall et al (1996a,b). The *prior* (background) open boundary conditions were taken from the global ocean data assimilation product of ECCO described by Wunsch and Heimbach (2007). The *prior* (background) initial conditions for each 7 day assimilation cycle are the *posterior* (analysis) at the end of the previous assimilation cycle. Full details of the model, observations, and configuration of the data assimilation system are given in Veneziani et al (2009) and Broquet et al (2009, 2011). In all experiments presented here, the 4D-Var control vector is comprised of the initial conditions, the ocean surface forcing, and the open boundary conditions.

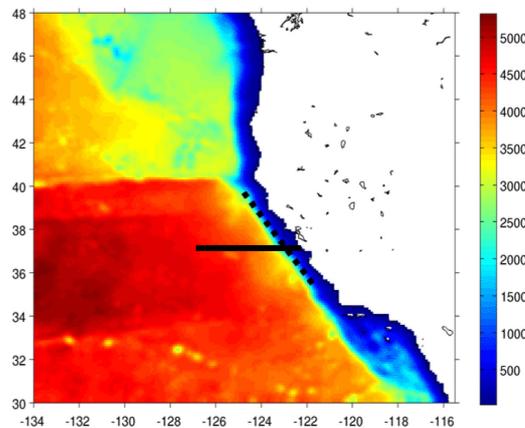


Figure 13: The latitudinal and longitudinal extent of the ROMS domain used to model the California Current. The coloured contours show the ocean bathymetry on a 10 km horizontal grid. Also shown are the sections along which the alongshore and cross-shore transport were defined for the observation impact and observation sensitivity calculations.

The observations assimilated into the models were collected from a variety of different platforms. The majority of the observations (~90%) are in the form of satellite measurements of SST and SSH, although most measurements are of SST. For convenience the gridded SSH product released by Aviso (<http://www.aviso.oceanobs.com>) was used in the experiments described here. The remaining

observations (~10%) were in the form of subsurface hydrographic measurements of temperature and salinity. These data were collected from various different platforms, including Argo drifting floats, ships, moorings, and tagged marine mammals. Figure 14 shows the locations of two routine ocean monitoring networks that are in place along the U.S. west coast, that provide valuable observations of seasonal changes in the upper ocean circulation. Most of the hydrographic observations used in the experiments presented here were taken from the quality controlled EN3 data product available at the U.K. Met Office (URL) and described by Ingleby and Huddleston (2007), except for the tagged marine mammal observations which were kindly provided by Dan Costa at the University of California, Santa Cruz.

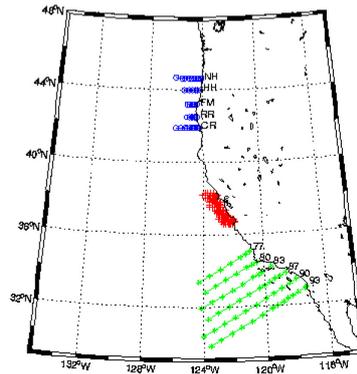


Figure 14: The repeat sampling hydrographic array in place along the U.S. west coast as part of the California Cooperative Fisheries Investigation (CalCOFI) (green), the Global Ecosystem Dynamics Experiment (GLOBEC) (blue), and routine surveys made by the U.S. National Marine Fisheries Service (red).

3.2. Observation Impact vs Observation Sensitivity

The definitions of observation impact and observation sensitivity used here differ subtly from those typically used by the NWP community, so a brief summary is worthwhile. Denoting the ocean state-vector as \mathbf{x} , then it is well known that the *posterior* (analysis) state \mathbf{x}_a can be written as:

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{Kd} \quad (1)$$

where \mathbf{x}_b denotes the *prior* (background) state, \mathbf{K} is the gain matrix, $\mathbf{d} = (\mathbf{y} - H(\mathbf{x}_b))$ is the innovation vector, \mathbf{y} the vector of observation values, and H is the observation operator that samples the *prior* (background) at the observation points. In the case of 4D-Var, H also includes the non-linear model dynamics. As discussed in earlier chapters of this volume (e.g. Lorenc, Barker), the *posterior* in (1) is identified in 4D-Var using iterative techniques via a sequence of linear least-squares minimizations involving the tangent linearization of the model and its adjoint. Since the iteration scheme is never run to complete convergence, the true gain matrix in (1) is never identified, but instead an approximation, $\tilde{\mathbf{K}}$, is found. Hereafter, $\tilde{\mathbf{K}}$ will be referred to as the practical gain matrix, and the *posterior* state estimate obtained after 4D-Var is given by:

$$\mathbf{x}_a = \mathbf{x}_b + \tilde{\mathbf{K}}\mathbf{d}. \quad (2)$$

3.2.1. Observation impact

In observation impact calculations, the focus of attention is usually a scalar function of the state \mathbf{x} , denoted here as $\mathcal{J}(\mathbf{x})$, not to be confused with the cost function. It is customary to consider $\mathcal{J}(\mathbf{x})$ defined over the forecast interval, usually as a measure of the forecast error (e.g. Langland and Baker, 2007; Errico, 2007; Gelaro et al, 2007) in which case $\mathcal{J}(\mathbf{x}) = \mathcal{J}(\mathbf{x}_f)$ where $\mathbf{x}_f = M(\mathbf{x}_a)$ denotes the forecast state vector arising from the analysis \mathbf{x}_a , and M is the forecast model. In the case of 4D-Var, M would evolve the analysis \mathbf{x}_a through both the analysis and forecast periods. However, from (2) $\mathbf{x}_f = M(\mathbf{x}_b + \tilde{\mathbf{K}}\mathbf{d})$ which to first-order in the increment $\delta\mathbf{x} = \tilde{\mathbf{K}}\mathbf{d}$ is given by $\mathbf{x}_f \approx M(\mathbf{x}_b) + \mathbf{M}_b(\tilde{\mathbf{K}}\mathbf{d})$ where \mathbf{M}_b is the tangent linearization of the forecast model linearized about \mathbf{x}_b . Thus the function $\mathcal{J}(\mathbf{x}_f)$ can be similarly approximated as $\mathcal{J}(\mathbf{x}_f) \approx \mathcal{J}(M(\mathbf{x}_b)) + \mathbf{d}^T \tilde{\mathbf{K}}^T \mathbf{M}_b^T (\partial\mathcal{J}/\partial\mathbf{x})|_{\mathbf{x}_b}$ where $(\partial\mathcal{J}/\partial\mathbf{x})|_{\mathbf{x}_b}$ denotes the derivative of \mathcal{J} evaluated at \mathbf{x}_b , and \mathbf{M}_b^T is the adjoint of the tangent linear model. The term $\mathcal{J}(M(\mathbf{x}_b))$ is the value of the function associated with the *prior* (background), while the term:

$$\delta\mathcal{J} = \mathbf{d}^T \tilde{\mathbf{K}}^T \mathbf{M}_b^T (\partial\mathcal{J}/\partial\mathbf{x})|_{\mathbf{x}_b} \quad (3)$$

represents the change in \mathcal{J} due to assimilating the observations. More specifically $\delta\mathcal{J}$ is given by the vector-product $\delta\mathcal{J} = \mathbf{d}^T \mathbf{g}$, where $\mathbf{g} = \tilde{\mathbf{K}}^T \mathbf{M}_b^T (\partial\mathcal{J}/\partial\mathbf{x})|_{\mathbf{x}_b}$. Since each observation is uniquely associated with each element of the innovation vector \mathbf{d} and the vector \mathbf{g} , the contribution or impact of observation j to $\delta\mathcal{J}$ is given by the elements $d_j g_j$ of the dot-product $\mathbf{d}^T \mathbf{g} = \sum_{i=1}^{N_{obs}} d_i g_i$. Equation (3) shows that $\tilde{\mathbf{K}}^T$, the transpose or adjoint of the practical gain matrix, provides the linear map from state-space to observation space that determines the specific contribution of each observation to $\delta\mathcal{J}$.

3.2.2. Observation sensitivity

As noted above, 4D-Var proceeds as a sequence of linear least-squares minimization problems. Therefore, the entire 4D-Var system can be considered more generally as a function, that is to say:

$$\mathbf{x}_a = \mathbf{x}_b + \mathcal{K}(\mathbf{d}, p_a, p_m) \quad (4)$$

where $\mathcal{K}(\mathbf{d}, p_a, p_m)$ denotes the 4D-Var algorithm, p_a are the data assimilation parameters, and p_m are the model parameters. In general $\mathcal{K}(\mathbf{d}, p_a, p_m)$ will be a non-linear function since the minimization algorithms typically employed in 4D-Var involve dot-products of \mathbf{d} with itself. To first order, any small change in the *prior* (background), observations, or parameters will lead to a change in the *posterior* (analysis) given by $\Delta\mathbf{x} \approx (\partial\mathcal{K}/\partial\mathbf{d})\Delta\mathbf{d} + (\partial\mathcal{K}/\partial p_a)\Delta p_a + (\partial\mathcal{K}/\partial p_m)\Delta p_m$ where the derivatives of \mathcal{K} represent the tangent linearization of the entire 4D-Var algorithm, and are evaluated using the unperturbed innovations and parameters. In the remainder of this section, only changes $\Delta\mathbf{d}$ in the innovation vector will be considered (i.e. $\Delta p_a = \Delta p_m = 0$). Consider again the scalar function $\mathcal{J}(\mathbf{x}_f)$

introduced in section 3.2.1. A small change in the innovations $\Delta \mathbf{d}$ will yield a change $\Delta \mathbf{x}_f \approx M(\mathbf{x}_a + \Delta \mathbf{x}) - M(\mathbf{x}_a) \approx \mathbf{M}_f \Delta \mathbf{x}$ where \mathbf{M}_f denotes the tangent linear model linearized about the unperturbed analysis-forecast cycle. The corresponding change $\Delta \mathcal{J}$ in $\mathcal{J}(\mathbf{x}_f)$ is given, to first order by:

$$\Delta \mathcal{J} = \mathcal{J}(\mathbf{x}_f + \Delta \mathbf{x}_f) - \mathcal{J}(\mathbf{x}_f) \approx \Delta \mathbf{x}_f^T (\partial \mathcal{J} / \partial \mathbf{x})|_{\mathbf{x}_f}. \quad (5)$$

where $(\partial \mathcal{J} / \partial \mathbf{x})|_{\mathbf{x}_f}$ represents the derivative of \mathcal{J} evaluated relative to the unperturbed analysis-forecast cycle. But from the preceding statements:

$$\Delta \mathcal{J} \approx \Delta \mathbf{x}_f^T (\partial \mathcal{J} / \partial \mathbf{x})|_{\mathbf{x}_f} \approx \Delta \mathbf{d}^T (\partial \mathcal{K} / \partial \mathbf{d})^T \mathbf{M}_f^T (\partial \mathcal{J} / \partial \mathbf{x})|_{\mathbf{x}_f} \quad (6)$$

where $(\partial \mathcal{K} / \partial \mathbf{d})^T$ is the adjoint of the tangent linearization of the entire 4D-Var algorithm. By analogy with (3), the adjoint operator $(\partial \mathcal{K} / \partial \mathbf{d})^T$ in (6) is the linear map from state-space to observation space and yields the sensitivity of \mathcal{J} to changes in the innovations.

3.3. Observation Impacts in the California Current

As discussed in section 2.1, the CC is an upwelling dominated circulation environment, and undergoes dramatic changes from season-to-season associated with changes in the alongshore wind speed and direction. The associated changes in circulation lead to variations in the alongshore and cross-shore transports which will form the basis of the calculations presented here. Sections along which these transports are computed are indicated in Fig. 13. Specifically, the alongshore transport will be defined as the 7 day average transport along 37 N from the coast to 127 W over the upper 500 m of the water column. This represents an average of three different components of the circulation in this region: the equatorward California Current, the poleward California Undercurrent, and the coastal jet. Conversely, the cross-shore transport will be defined as the 7 day average transport in the upper 15 m of the water column as it passes over the shelf-break (defined here as the 500 m isobath) between 35 N and 40.5 N. The alongshore transport and cross-shore so defined represent convenient scalar measures of the circulation and will be denoted as \mathcal{J}_{37N} and \mathcal{J}_{500m} respectively. For the purpose of illustration, the impact of the observations on \mathcal{J}_{37N} and \mathcal{J}_{500m} will be considered during the analysis cycle only. Forecast cycle impacts are presented in Moore et al (2011c).

Figure 15 shows time series of the *prior* (background) transports $\mathcal{J}_{37N}(\mathbf{x}_b)$ and $\mathcal{J}_{500m}(\mathbf{x}_b)$ for each 7 day assimilation cycle using the model CC10. Each cycle employs 1 outer-loop and 15 inner-loops. The seasonal variations in the alongshore and cross-shore movement of water in response to the seasonally varying alongshore winds is clearly evident. Figure 16 shows time series of the transport increments $\delta \mathcal{J}_{37N}$ and $\delta \mathcal{J}_{500m}$ from each cycle that result from 4D-Var assimilation of all available observations. The contribution of each type of observing platform to $\delta \mathcal{J}$ is also indicated by the different colored bars computed according to (3). The first point note is that data assimilation yields a significant change in both alongshore and cross-shore transport. Secondly, it is clear that satellite SST

observations exert by far the largest impact on the circulation changes. This is of course to be expected since, as noted earlier, $\sim 90\%$ of the available observations take the form of satellite measurements, and most of these are from satellite radiometers. However, a closer inspection of Fig. 16 reveals that during many cycles, the subsurface hydrographic data often exert an equal or greater impact on the transport increments than the satellite observations. For example, salinity observations from Argo drifting floats and the repeat sampling networks of Fig. 14 often times contribute significantly to $\delta\mathcal{J}_{37N}$ and $\delta\mathcal{J}_{500m}$.

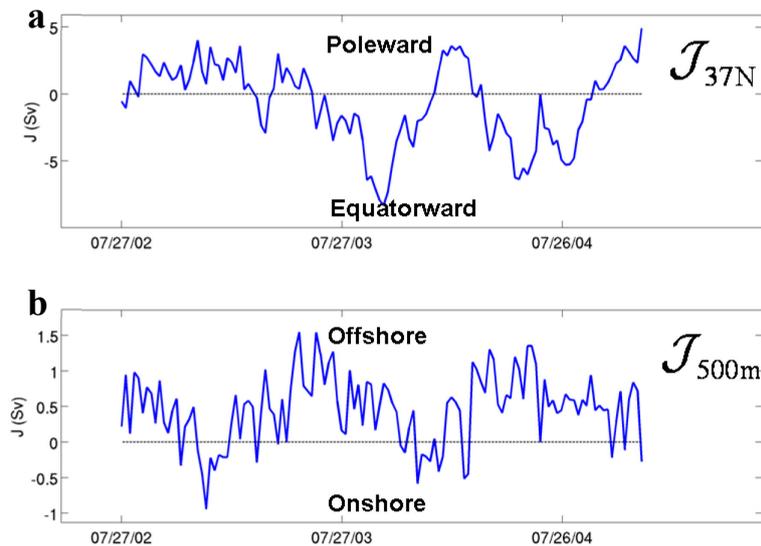


Figure 15: Time series of the prior (background) alongshore transport (a) and cross-shore transport (b) for each 4D-Var data assimilation cycle during the period July 2002-Dec 2004 using CC10. The units of transport are Sverdrups (Sv), where $1 \text{ Sv} = 10^6 \text{ m}^3 \text{ s}^{-1}$.

The root mean square (rms) average over all assimilation cycles of the impact of each observing platform on the transport increments is also indicated in Fig. 16 and reveals that despite the relatively low abundance of *in situ* subsurface hydrographic observations ($\sim 10\%$ of the total number of data), on average they account for $\sim 40\text{-}50\%$ of the impact on $\delta\mathcal{J}_{37N}$ and $\delta\mathcal{J}_{500m}$. Thus the relatively small number of subsurface observations are critical for the analysis of the circulation.

The total number of degrees of freedom (dof) of the observing network can be quantified according to $\text{tr}\{\mathbf{KH}\}$ (Desroziers et al, 2005) where the tangent linear observation operator \mathbf{H} also includes the tangent linear model dynamics in the case of 4D-Var. Moore et al (2011b) have shown that for the same ROMS CC 4D-Var system the $\text{dof} \leq 0.1N_{obs}$, where N_{obs} is the total number of observations, which is suggestive of a high level of redundancy in the satellite observations assimilated into the model. Therefore, it may be possible to increase the impact of the subsurface observations further by decimating in some way the number of satellite data assimilated into the model.

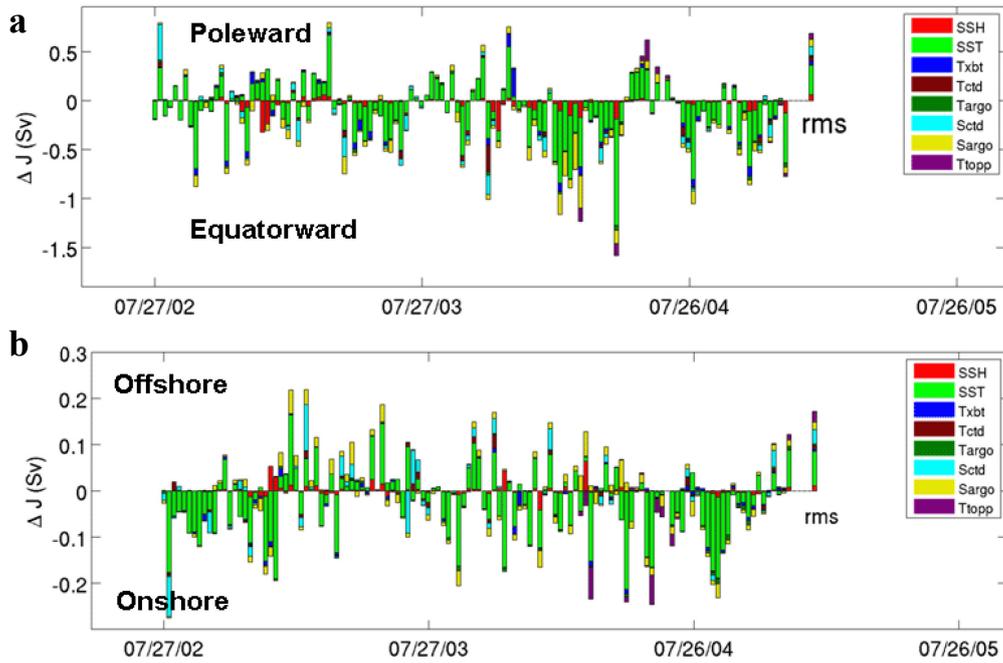


Figure 16: Time series of (a) alongshore transport increments, and (b) cross-shore transport increments from each 4D-Var cycle using CC10. The contribution to the transport increments of all the observations from the different observing platforms are indicated by the different coloured bars. Key: SSH, satellite sea surface height; SST, satellite sea surface temperature; Txbt, temperature from expendable bathy thermographs (XBT) from ships; T,Sctd, temperature or salinity from conductivity, temperature, depth (CTD) profilers aboard ships; T,Sargo, temperature or salinity from Argo profiling drifting floats; Ttopp, temperature from tagged elephant seals. (From Moore et al, 2011c).

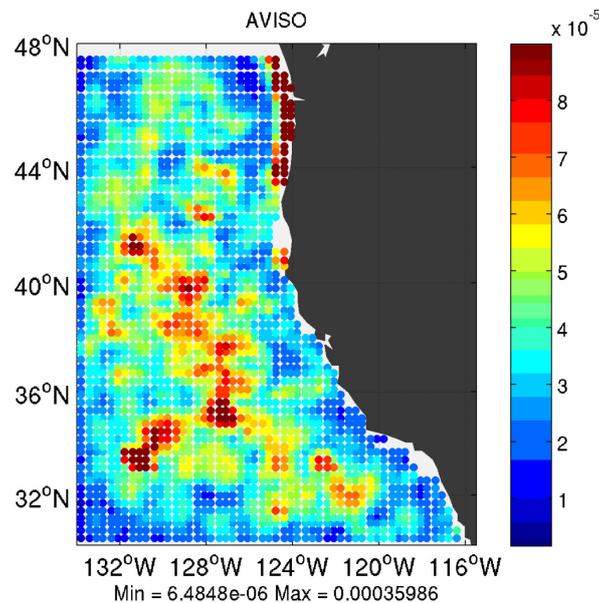


Figure 17: The rms average over all 4D-Var cycles of the impact of satellite SSH observations on alongshore transport, \mathcal{J}_{37N} , at each observation location. Units are Sv. (From Moore et al, 2011c).

The geographical variations of the observation impacts yield clues about the dynamical processes that are most likely at work in redistributing information from the observations. For example, Fig. 17 shows the rms average impact of SSH observations on $\delta\mathcal{J}_{37N}$ at each observation point, averaged over all assimilation cycles. The elevated impact of SSH observations along the coast of Oregon and Washington is suggestive of the role played by coastally trapped waves in transmitting information within the coastal waveguide of the adjoint model. Some of the deep water impacts will be associated with fast, non-dispersive baroclinic inertia-gravity waves. As discussed in section 2.5 some of these waves will be a legitimate part of the ocean response, while others will be the result of dynamical imbalances in the 4D-Var initial conditions. Finally, the relatively broad, somewhat coherent region of elevated sensitivity between 132W-124W, 32N-42N is most likely the result of large-scale changes in the slope of the sea-surface associated with the south eastern portion of the subtropical gyre. Changes in the cross-shore pressure associated with the gyre circulation will result directly in changes in the alongshore geostrophic transport.

3.4. Observation Sensitivity in the California Current

Observation sensitivity as described by (6) is fundamentally different to the observation impact calculations described in the previous section. As demonstrated in section 3.3, observation impact described by (3) quantifies the contribution of each observation to the *actual* increment $\delta\mathcal{J}$ that results in a scalar function due to assimilating observations. This arises from the linear map $\tilde{\mathbf{K}}^T$. Conversely, observation sensitivity given by (6) *predicts* the change $\Delta\mathcal{J}$ that will occur in the scalar function due to a change in the observations or more generally a change in the innovations. This information is derived from the linear map $(\partial\mathcal{K}/\partial\mathbf{d})^T$. Only in the case where the 4D-Var calculation is run to complete convergence will the linear maps of observation impact and observation sensitivity be the same, in which case $\partial\mathcal{K}/\partial\mathbf{d} = \mathbf{K}$, the true gain matrix.

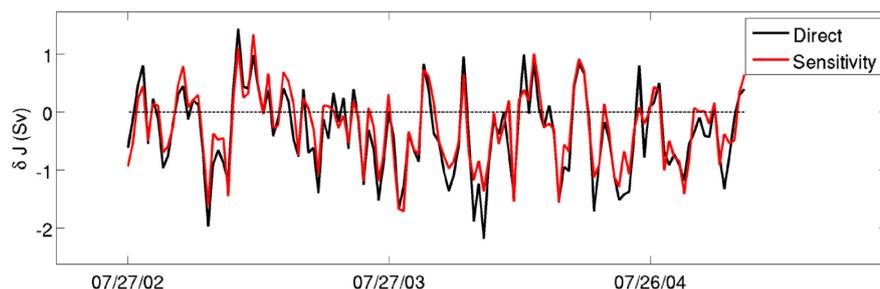


Figure 18: Time series of the change, $\Delta\mathcal{J}_{37N}$, in the alongshore transport when SSH observations are withheld **independently** during each 4D-Var cycle. That is to say, all observations are assimilated during all cycles prior to the cycle under consideration. The black curve shows $\Delta\mathcal{J}_{37N}$ when observations are withheld and each 4D-Var cycle is repeated. The red curve, on the other hand, shows the prediction of $\Delta\mathcal{J}_{37N}$ computed using $(\partial\mathcal{K}/\partial\mathbf{d})^T$ according to (6) and (7). (From Moore et al, 2011c).

The power and utility of the observation sensitivity embodied in (6) is demonstrated here in relation to the alongshore transport, \mathcal{J}_{37N} . In this case, results are shown using the CC30 model configuration since the adjoint operation $(\partial\mathcal{K}/\partial\mathbf{d})^T$ involves the same computational effort as 4D-Var itself for each \mathcal{J} identified. Specifically, the case is considered in which perturbations $\Delta\mathbf{y}$ were added to the observations \mathbf{y} , in which case $(\partial\mathcal{K}/\partial\mathbf{d})^T \equiv (\partial\mathcal{K}/\partial\mathbf{y})^T$. A specific form for the perturbations $\Delta\mathbf{y}$ was chosen, namely:

$$\Delta\mathbf{y} = -\mathbf{W}\mathbf{d} = -\mathbf{W}(\mathbf{y} - H(\mathbf{x}_b)) \quad (7)$$

where \mathbf{W} is a diagonal matrix. Choosing the diagonal elements of \mathbf{W} to be 0 or 1 simulates the effect of withholding specific observations from the analysis. If the number of observations is N_{obs} , then \mathbf{W} is an $N_{obs} \times N_{obs}$ matrix, with leading diagonal $w_{i,i}$, where $i=1, \dots, N_{obs}$. If $w_{j,j}=0$ this leaves the j^{th} observation y_j unchanged, whereas $w_{j,j}=1$ introduces a perturbation to y_j that yields a zero innovation, and corresponds to the situation where the background and the perturbed observation are in perfect agreement at the j^{th} observation location. By simulating the removal of specific observations in this way, equation (6) yields a powerful tool for predicting the effect of withholding observations (i.e. changing the observing network) on the analysis increments.

To demonstrate, Fig. 18 shows two time series of the change in alongshore transport, $\Delta\mathcal{J}_{37N}$, that arises when all of the altimeter measurements of SSH were withheld *independently* from each 4D-Var analysis cycle. That is to say, *all* observations were assimilated during all previous cycles except during the cycle under consideration. In one case the SSH observations were withheld, and $\Delta\mathcal{J}_{37N}$ was computed by repeating each 4D-Var analysis. In the second case, however, the $\Delta\mathcal{J}_{37N}$ were predicted using (6) and observation perturbations of the form (7). Figure 18 indicates that the degree of agreement between the prediction of $\Delta\mathcal{J}_{37N}$ based on (6) and direct computation of $\Delta\mathcal{J}_{37N}$ by repeating the 4D-var analyses is remarkably good. The advantage of using (6), however, is that for a given \mathcal{J} , the 4D-Var adjoint $(\partial\mathcal{K}/\partial\mathbf{d})^T$ need only be run once to yield the sensitivity of $\Delta\mathcal{J}$ to any combination of observations, provided the number of observations withheld does not exceed a modest fraction of the total number of observations, otherwise the approximations employed break down. The results of Fig. 18 demonstrate considerable utility, and provide an alternative and complimentary approach to Observing System Experiments (OSEs).

3.5. 4D-Var Sensitivity and Posterior Error Estimates

The sensitivity operator $(\partial\mathcal{K}/\partial\mathbf{d})^T$ has even wider utility than discussed in section 3.4, and here a demonstration of its use in computing the expected *posterior*/analysis errors in \mathcal{J} is presented. It is well known (e.g. Daley, 1991) that the expected *posterior*/analysis error covariance, \mathbf{A} , can be expressed as:

$$\mathbf{A} = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{B}(\mathbf{I} - \mathbf{K}\mathbf{H})^T + \mathbf{K}\mathbf{R}\mathbf{K}^T. \quad (8)$$

However, direct computation of \mathbf{A} using (8) (or equivalent forms) is difficult in practice because of the large dimension of the problem, and is particularly challenging in 3D-Var and 4D-Var because the information required to estimate \mathbf{A} is not readily available in a convenient form. However, there has been growing interest recently in ensemble 4D-Var methods (Belo Pereira and Berre, 2006) where each member of the ensemble is generated by repeating the 4D-Var analysis using perturbed observations and perturbed *priors*. If the perturbations are drawn from distributions with the *prior* error covariances \mathbf{R} and \mathbf{B} , then Belo Pereira and Berre (2006) have shown the covariance of the perturbed analyses about the unperturbed analysis mimics the expected analysis error covariance. This provides a convenient and practical method for estimating \mathbf{A} which can be used to provide flow dependent background covariance information for the next data assimilation cycle. Inspired by this approach, Moore et al (2011d) have recently demonstrated that if the analysis perturbations are instead expressed in terms of $\partial\mathcal{K}/\partial\mathbf{d}$, the expected *posterior*/analysis error covariance can also be expressed as:

$$\mathbf{A} = (\mathbf{I} - (\partial\mathcal{K}/\partial\mathbf{d})\mathbf{H})\mathbf{B}(\mathbf{I} - (\partial\mathcal{K}/\partial\mathbf{d})\mathbf{H})^T + (\partial\mathcal{K}/\partial\mathbf{d})\mathbf{R}(\partial\mathcal{K}/\partial\mathbf{d})^T. \quad (9)$$

Equation (9) has a similar form to (8) where \mathbf{K} in (8) is replaced by $\partial\mathcal{K}/\partial\mathbf{d}$. This should not come as too much of a surprise since, as noted above, when 4D-Var is run to complete convergence, $\partial\mathcal{K}/\partial\mathbf{d} = \mathbf{K}$. In practice though, 4D-Var calculations are terminated before complete convergence is achieved, in which case \mathbf{K} in (8) is replaced by $\tilde{\mathbf{K}}$, the practical gain matrix. While the form for \mathbf{A} in (9) is still seemingly large and unwieldy, its convergence properties to the true \mathbf{A} are superior to those of (8) using $\tilde{\mathbf{K}}$, and (9) it is particularly convenient for computing the expected errors of functions \mathcal{J} .

To illustrate the utility of (9), the expected error in 4D-Var estimates of the alongshore transport \mathcal{J}_{37N} will be presented. Specifically, the alongshore transport can be represented as $\mathcal{J}_{37N} = \sum_{k=1}^N \mathbf{h}_k^T \mathbf{x}_k$ where $\mathbf{x}_k \equiv \mathbf{x}(t) = \mathbf{x}(k\Delta t)$ is from the model with time step Δt , and $\mathbf{h}_k \equiv \mathbf{h}(t) = \mathbf{h}(k\Delta t)$ is an appropriately defined vector that yields the 7 day average transport along the 37 N section over the desired range of longitude and depth. Using this construct, the *prior* (or background) error variance in \mathcal{J}_{37N} can be expressed as:

$$\left(\sigma_{\mathcal{J}}^b\right)^2 = \left(\sum_{k=1}^N \mathbf{h}_k^T \mathbf{M}_k\right) \mathbf{B} \left(\sum_{i=1}^N \mathbf{M}_i^T \mathbf{h}_i\right) \quad (10)$$

where $\mathbf{M}_k \equiv \mathbf{M}(t, 0) = \mathbf{M}(k\Delta t, 0)$ is the tangent linear model that propagates solutions from time 0 to time t . Similarly, the *posterior* (or analysis) error variance in \mathcal{J}_{37N} can be expressed as:

$$\left(\sigma_{\mathcal{J}}^a\right)^2 = \left(\sum_{k=1}^N \mathbf{h}_k^T \mathbf{M}_k\right) \mathbf{A} \left(\sum_{i=1}^N \mathbf{M}_i^T \mathbf{h}_i\right) \quad (11)$$

where \mathbf{A} is given by (9). A rearrangement of the combination of (9) and (11) reveals that $(\sigma_{\mathcal{J}}^a)^2$ can be computed from a single application of the sensitivity operator $(\partial\mathcal{K}/\partial\mathbf{d})^T$ to $\sum_{i=1}^N \mathbf{M}_i^T \mathbf{h}_i$.

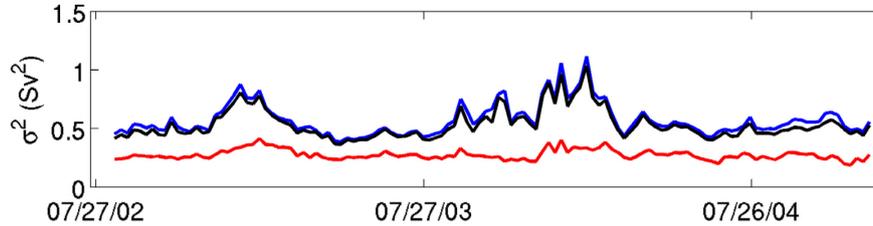


Figure 19: Time series of prior error variance $(\sigma_{\mathcal{J}}^b)^2$ (blue curve) and posterior error variance $(\sigma_{\mathcal{J}}^a)^2$ (red curve) for the alongshore transport, \mathcal{J}_{37N} , computed using (10) and (11) and where \mathbf{A} is given by (9). In addition, posterior error variance estimates are also shown based on (8) and using the practical gain matrix $\tilde{\mathbf{K}}$.

Figure 19 shows time series of $(\sigma_{\mathcal{J}}^b)^2$ and $(\sigma_{\mathcal{J}}^a)^2$ for the *prior* and *posterior* estimates of alongshore transport respectively for each 7 day 4D-Var data assimilation cycle using the CC30 configuration of ROMS. Figure 19 reveals that the error variance in the *posterior* transport estimates is $\sim 50\%$ or less than the expected error in the *prior*. Also shown in Fig. 19 are estimates of the *posterior* error variance using instead \mathbf{A} from (8) in (11) but with \mathbf{K} replaced by $\tilde{\mathbf{K}}$. The *posterior* error estimates in this case are only marginally smaller than the *prior* errors and represent a gross over-estimate of the expected analysis error. This is a well-known symptom of the reduced rank nature of the (8) using $\tilde{\mathbf{K}}$, and is associated with the slow convergence of \mathbf{A} to the true \mathbf{A} as discussed by Moore et al (2011d).

As discussed in section 2.1 the CC circulation has a profound influence on the local marine ecosystems, and one important component of this interaction is the transport of nutrients and marine organisms along the coast (and offshore transport). Thus alongshore transport estimates such as those in Figs. 15 and 16 provide a basis for estimating alongshore biological fluxes. However, direct measurements of transport and flux are seldom made, so the error estimates in Fig. 19 are crucially important for ascertaining the degree of confidence in transport and flux calculations derived from the 4D-Var ocean analyses.

3.6. Summary

Observation impact and observation sensitivity calculations are fast gaining attention in oceanography, and as data assimilation systems evolve and mature, we are likely to see more wide spread use of these important diagnostic tools. The ROMS calculations presented here only scratch the surface, and clearly there is more analysis to be done in order to fully understand the implications of these results for ocean observing systems and ocean state estimation.

4. References

- Bakun, A., 2006: Wasp-waisted populations and marine ecosystem dynamics: Navigating the “predator pit” topographies. *Progress in Oceanography*, **68**, 271-288.
- Belo Pereira, M. and L. Berre, 2006: The use of an ensemble approach to study the background error covariances in a global nwp model. *Mon. Wea. Rev.*, **134**, 2466–2498.
- Broquet, G., C.A. Edwards, A.M. Moore, B.S. Powell, M. Veneziani and J.D. Doyle, 2009: Application of 4D-variational data assimilation to the California Current system. *Dyn. Atmos. Oceans*, **48**, 69-91.
- Broquet, G., A.M. Moore, H.G. Arango and C.A. Edwards, 2011: Corrections to ocean surface forcing in the California Current system using 4D-variational data assimilation. *Ocean Modelling*, **36**, 116-132.
- Campbell, J.W., 1995: The lognormal distribution as a model for bio-optical variability in the sea. *J. Geophys. Res.*, **100**, 13,237-13,254.
- Checkley, D.M. and J.A. Barth, 2009: Patterns and processes in the California Current system. *Progress in Oceanography*, **83**, 49-64.
- Chelton, D.B., R.A. DeSzoeki, M.G. Schlax, K. El Naggar, and N. Siwertz, 1998: Geographical variability of the first baroclinic Rossby radius of deformation. *J. Phys. Oceanogr.*, **28**, 433-460.
- Chelton, D.B., M.G. Schlax and R.M. Samelson, 2007: Summertime coupling between sea surface temperature and wind stress in the California Current system. *J. Phys. Oceanogr.*, **37**, 495-517.
- Courtier, P., 1997: Dual formulation of four-dimensional variational assimilation. *Q. J. R. Meteorol. Soc.*, **123**, 2449-2461.
- Courtier, P., J.-N. Thépaut and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var using an incremental approach. *Q. J. R. Meteorol. Soc.*, **120**, 1367-1388.
- Cooper, M. and K. Haines, 1996: Altimetric assimilation with water property conservation. *J. Geophys. Res.*, **101**, 1059-1077.
- Daley, R., 1991: Atmospheric Data Analysis. Cambridge University Press, Cambridge.
- Desroziers, G., L. Berre, B. Chapnik and P. Poli, 2005: Diagnosis of observation, background and analysis-error statistics in observation space. *Q. J. R. Meteorol. Soc.*, **131**, 3385-3396.
- Dobricic, S., N. Pinardi, M. Adani, A. Bonazzi, C. Fratianni and M. Tonani, 2005: Mediterranean forecasting system: a new assimilation scheme for sea level anomaly and its validation. *Q. J. R. Meteorol. Soc.*, **131**, 3627-3642.
- Doyle, J., Q. Jiang, Y. Chao, and J. Farrara, 2009: High-resolution atmospheric modeling over the Monterey Bay during AOSN. *Deep Sea Res. II*, **56**, 87–99.
- Egbert, G.D., A.F. Bennett and M.C.G. Foreman, 1994: TOPEX/POSEIDON tides estimated using a global inverse model. *J. Geophys. Res.*, **99**, 821-852.

- Errico, R.M., 2007: Interpretations of an adjoint-derived observational impact measure. *Tellus*, **59A**, 273-276.
- Fairall, C.W., E.F. Bradley, J.S. Godfrey, G.A. Wick, J.B. Edson and G.S. Young, 1996a: Cool-skin and warm layer effects on the sea surface temperature. *J. Geophys. Res.*, **101**, 1295-1301.
- Fairall, C.W., E.F. Bradley, D.P. Rogers, J.B. Edson and G.S. Young, 1996b: Bulk parameterization of air-sea fluxes for tropical ocean global atmosphere coupled ocean atmosphere response experiment. *J. Geophys. Res.*, **101**, 3747-3764.
- Fletcher, S.J., 2010: Mixed log-normal-Gaussian four-dimensional data assimilation. *Tellus*, **62A**, 266-287.
- Friedrichs, M.A.M, 2002: Assimilation of JGOFS EqPac and SeaWiFS data into a marine ecosystem model of the central equatorial Pacific Ocean. *Deep Sea Res. II*, **49**, 289-319.
- Fox, D.N., W.J. Teague, C.N. Barron, M.R. Carnes and C.M. Lee, 2002: The modular ocean data assimilation system (MODAS). *J. Atmos. Ocean Tech.*, **19**, 240-252.
- Gelaro, R., Y. Zhu and R.M. Errico, 2007: Examination of various order adjoint-based approximations of observation impact. *Meteorol. Z.*, **16**, 685-692.
- Hickey, B.M., 1998: Coastal oceanography of western north America from the tip of Baja California to Vancouver Island. *The Sea*, **11**, 345-393.
- Ingleby, B. and M. Huddleston, 2007: Quality control of ocean temperature and salinity profile – historical and real-time data. *J. Mar. Res.*, **65**, 158-175.
- Kim, S.Y., E. Terrill, B. Cornuelle, B. Jones, L. Washburn, M. Moline, J. Paduan, N. Garfield, J.L. Largier, G. Crawford and P.M. Kosro, 2011: Observations of high-resolution coastal surface circulation on the U.S. West Coast. *J. Geophys. Res.*, **116**, doi:10.1029/2010JC006669.
- Langland, R.H. and N.L. Baker, 2004: Estimation of observation impact using the NRL atmospheric variational data assimilation adjoint system. *Tellus*, **56A**, 189-201.
- LeBlond, P.H. and L.A. Mysak, 1980: *Waves in the Ocean*. Elsevier, New York.
- Levitus, S., 1982: *Climatological atlas of the world ocean*. NOAA prof. paper **13**, U.S. Govt. Printing Office, Washington D.C.
- Li, Z., Y. Chao, J.C. McWilliams and K. Ide, 2008: A three-dimensional variational data assimilation scheme for the Regional Ocean Modeling System. *J. Atmos. Ocean. Tech.*, **25**, 2074-2090.
- Liu, W.T., K.B. Katsaros and J.A. Businger, 1979: Bulk parameterization of the air-sea exchange of heat and water vapor including molecular constraints at the interface. *J. Atmos. Sci.*, **36**, 1722-1735.
- Moore, A.M., N.S. Cooper and D.L.T. Anderson, 1987: Initialization and data assimilation in models of the Indian Ocean. *J. Phys. Oceanogr.*, **17**, 1965-1977.

- Moore, A. M., H. Arango, G. Broquet, B. Powell, J. Zavala-Garay, and A. Weaver, 2011a: The Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilation systems. I: System overview and formulation. *Progress in Oceanography*, **91**, 34-49.
- Moore, A.M., H.G. Arango G. Broquet, C. Edwards, M. Veneziani, B. Powell, D. Foley, J.D. Doyle, D. Costa and P. Robinson, 2011b: The Regional Ocean Modeling System (ROMS) 4 dimensional variational data assimilation systems, II: Performance and application to the California Current system. *Progress in Oceanography*, **91**, 50-73.
- Moore, A.M., H.G. Arango G. Broquet, C. Edwards, M. Veneziani, B. Powell, D. Foley, J.D. Doyle, D. Costa and P. Robinson, 2011c: The Regional Ocean Modeling System (ROMS) 4 dimensional variational data assimilation systems, III: Observation impact and observation sensitivity in the California Current system. *Progress in Oceanography*, **91**, 74-94.
- Moore, A.M., H.G. Arango and G. Broquet, 2011d: Estimates of analysis and forecast errors derived from the adjoint of 4D-Var. *Mon. Wea. Rev.*, Submitted.
- Pedlosky, J., 1996: *Ocean Circulation Theory*. Springer.
- Raghukumar, K., 2011: Impact of assimilating physical oceanographic information on a complex ecosystem model in the California Current system. In preparation.
- Robinson, A.R. and W.G. Leslie, 1985: Estimation and prediction of oceanic eddy fields *Progress in Oceanography*, **14**, 385-510.
- Rykaczewski, R.R. and D.M. Checkley, 2007: Influence of ocean winds of pelagic ecosystems in upwelling regions. *Proceedings of the National Academy of Sciences*, **105**, 1965-1970.
- Shchepetkin, A.F. and J.C. McWilliams, 2005: The Regional Oceanic Modeling System (ROMS): a split explicit, free-surface, topography-following-coordinate oceanic model. *Ocean Modelling*, **9**, 347-404.
- Simon, E. and L. Bertino, 2009: Application of the Gaussian anamorphosis to assimilation in a 3-D coupled physical-ecosystem model of the North Atlantic with the EnKF: a twin experiment. *Ocean Sci.*, **5**, 495-510.
- Veneziani, M., C.A. Edwards, J.D. Doyle and D. Foley, 2009: A central California coastal ocean modeling study: 1. Forward model and the influence of realistic versus climatological forcing. *J. Geophys. Res.*, **114**, C04015. doi:10.1029/2008JC004774.
- Weaver, A. and P. Courtier, 2001: Correlation modeling on the sphere using a generalized diffusion equation. *Q. J. R. Meteorol. Soc.*, **127**, 1815-1846.
- Weaver, A.T., J. Vialard and D.L.T. Anderson, 2003: Three- and four-dimensional variational data assimilation with a general circulation model of the tropical Pacific Ocean. Part I: Formulation, internal dynamics and consistency checks. *Mon. Wea. Rev.*, **131**, 3605-3625.
- Wunsch, C. and P. Heimbach, 2007: Practical global ocean state estimation. *Physica D*, **230**, 197-208.

