

THE OFFICIAL MAGAZINE OF THE OCEANOGRAPHY SOCIETY

Oceanography

CITATION

Cummings, J., L. Bertino, P. Brasseur, I. Fukumori, M. Kamachi, M.J. Martin, K. Mogensen, P. Oke, C.E. Testut, J. Verron, and A. Weaver. 2009. Ocean data assimilation systems for GODAE. *Oceanography* 22(3):96–109, doi:10.5670/oceanog.2009.69.

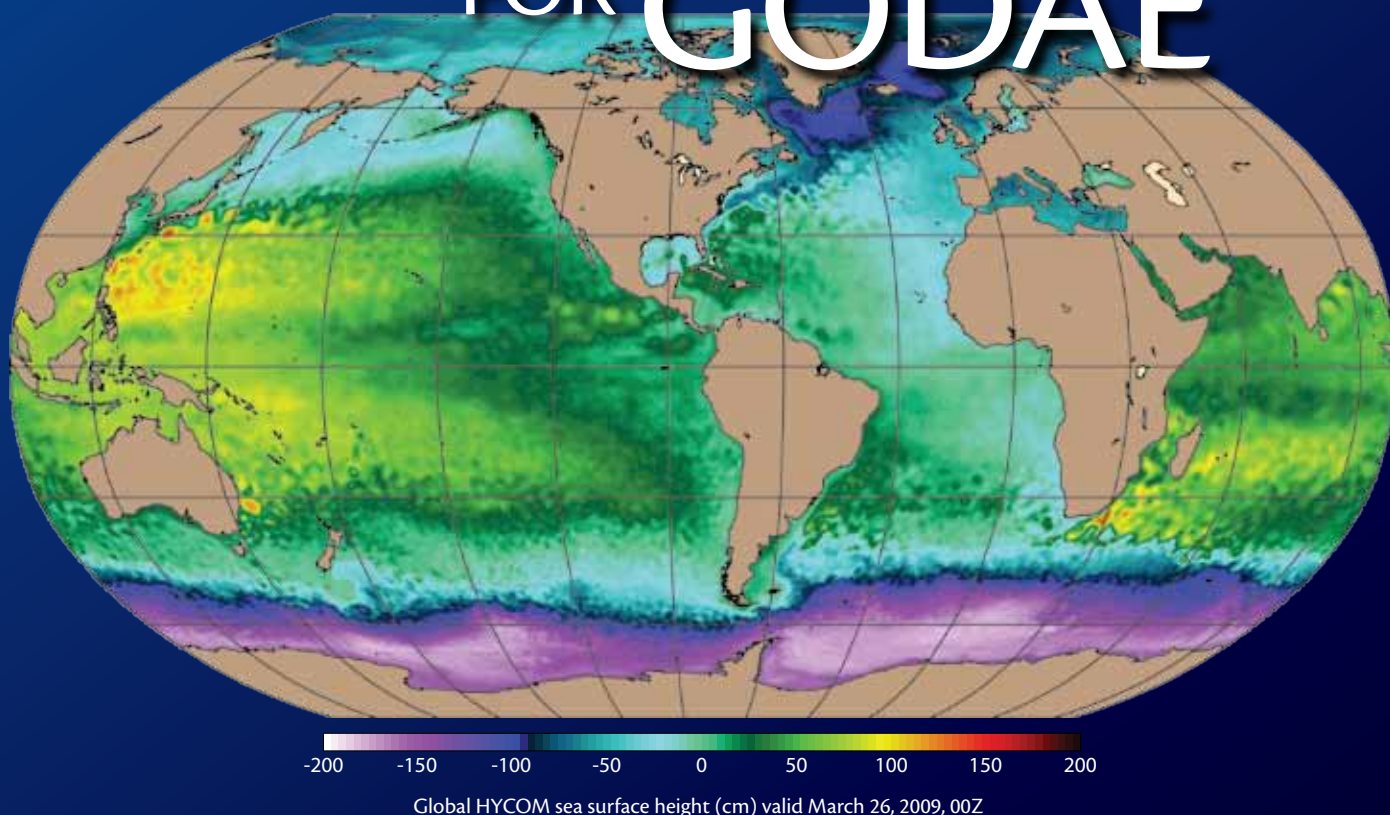
COPYRIGHT

This article has been published in *Oceanography*, Volume 22, Number 3, a quarterly journal of The Oceanography Society. Copyright 2009 by The Oceanography Society. All rights reserved.

USAGE

Permission is granted to copy this article for use in teaching and research. Republication, systematic reproduction, or collective redistribution of any portion of this article by photocopy machine, reposting, or other means is permitted only with the approval of The Oceanography Society. Send all correspondence to: info@tos.org or The Oceanography Society, PO Box 1931, Rockville, MD 20849-1931, USA.

OCEAN DATA ASSIMILATION SYSTEMS FOR GODAE



BY JAMES CUMMINGS, LAURENT BERTINO, PIERRE BRASSEUR, ICHIRO FUKUMORI,
MASAFUMI KAMACHI, MATTHEW J. MARTIN, KRISTIAN MOGENSEN, PETER OKE,
CHARLES EMMANUEL TESTUT, JACQUES VERRON, AND ANTHONY WEAVER

ABSTRACT. Ocean data assimilation has matured to the point that observations are now routinely combined with model forecasts to produce a variety of ocean products. Approaches to ocean data assimilation vary widely both in terms of the sophistication of the method and the observations assimilated, and also in terms of specification of the forecast error covariances, model biases, observation errors, and quality-control procedures. In this paper, we describe some of the ocean data assimilation systems that have been developed within the Global Ocean Data Assimilation Experiment (GODAE) community. We discuss assimilation methods, observations assimilated, and techniques used to specify error covariances. In addition, we describe practical implementation aspects and present analysis performance results for some of the analysis systems. Finally, we describe plans for improving the assimilation systems in the post-GODAE time period beyond 2008.

INTRODUCTION

Ocean data assimilation is a mathematically rigorous process of combining ocean observations and ocean models to extract the most important information from relatively sparse and incomplete observations of time-varying ocean circulation. The main goals of ocean data assimilation are to improve our understanding of ocean circulation and monitor and predict circulation on all relevant temporal and spatial scales. Ocean data assimilation products developed during the Global Ocean Data Assimilation Experiment (GODAE) are used to: (1) initialize ocean models using all available observations through sequential approaches for forecasting, and (2) synthesize observations with ocean models to obtain dynamically

consistent estimates of changing ocean state. Ocean data assimilation is needed because ocean models are likely to have errors due to deficiencies in model physics, grid resolution, lateral boundary conditions, or atmospheric forcing. One important impact of data assimilation is to counter the tendency of ocean models to drift away from reality.

Although major challenges remain, substantial progress has been made during GODAE in developing ocean data assimilation systems. GODAE assimilation systems are now producing ocean forecasts and ocean state estimates on a routine basis, some in near-real time. GODAE systems represent large investments by many national groups that are in the process of transitioning to sustained assimilation activities for both

scientific and operational purposes. This paper summarizes the status of ocean data assimilation systems developed in support of GODAE activities. We focus on the most important aspects of the GODAE assimilation systems without resorting to the mathematical details. For each of the GODAE systems discussed, we briefly describe: (1) the data assimilation method along with practical aspects of implementing the method; (2) the observing systems assimilated and forecast error covariances, including specification of the background, observation, and multivariate aspects of the assimilation; (3) performance results of the assimilation; and (4) future plans in the post-GODAE time period (beyond 2008). Note that the modeling and operational aspects of the assimilation systems are discussed elsewhere in this special issue (see Dombrowsky et al. and Hurlburt et al.), and are not repeated here. We provide an extensive reference list for those readers who seek more detail on one or more of the systems described in this paper.

GODAE systems discussed include: the BLUElink> Ocean Data Assimilation System (BODAS), operational at the Australian Bureau of Meteorology; the near-real-time Estimating the Circulation and Climate of the Ocean

Table 1. Data assimilation methods used by GODAE systems

System Name	Country	Data Assimilation Method	Reference
BODAS	Australia	Ensemble Optimal Interpolation	Oke et al., 2008
ECCO-JPL	USA	Kalman filter and smoother	Fukumori, 2002
FOAM	UK	Analysis Correction	Martin et al., 2007
Mercator	France	Static SEEK filter	Brasseur et al., 2005
MOVE/MRI.COM	Japan	Multivariate 3DVAR	Fujii and Kamachi, 2003
NCODA	USA	Multivariate Optimal Interpolation	Cummings, 2005
NEMOVAR	European Union	Multivariate Incremental 3DVAR	Weaver et al., 2005
TOPAZ	Norway	Ensemble Kalman filter	Evensen, 2006

(ECCO) system, operational at the Jet Propulsion Laboratory; the Forecast Ocean Assimilation Model (FOAM), operational at the UK Met Office; Mercator, operational in France; the Multivariate Ocean Variational Estimation/Meteorological Research Institute Community Ocean Model (MOVE/MRI.COM), operational at the Japan Meteorological Agency; the Navy Coupled Ocean Data Assimilation (NCODA) system, operational at US Navy oceanography centers; Nucleus for European Modeling of the Ocean VARIational data assimilation (NEMOVAR), planned for research and operations at several European institutions; and Towards an Operational Prediction system for the North Atlantic European coastal Zones (TOPAZ), operational exploitation of which has been transitioned to the Norwegian meteorological services.

METHODS

A variety of methods exist for assimilating observations into ocean models, and GODAE systems reflect this

diversity in methodology. GODAE assimilation methods range from relatively simple schemes, such as Analysis Correction and Optimal Interpolation, to more sophisticated schemes, such as variational and ensemble techniques. Table 1 lists the assimilation methods used in GODAE systems. Detailed descriptions of GODAE data assimilation schemes can be found in the references shown in table. (See, for

example, Brasseur (2006) and Wunsch (2006) for general discussions.) Most GODAE systems are run in near-real time for ocean monitoring and forecasting purposes, while some systems are exclusively executed for times in the past in what is referred to as “reanalysis mode.” The basic data inputs into any assimilation system are the innovations. Innovations are the differences between the observations and the model

James Cummings (james.cummings@nrlmry.navy.mil) is Oceanographer, Naval Research Laboratory, Monterey, California, USA. **Laurent Bertino** is Leading Scientist, Modeling and Data Assimilation Group, Nansen Environmental and Remote Sensing Center, Bergen, Norway. **Pierre Brasseur** is Research Scientist, Conseil National de la Recherche Scientifique/Laboratoire des Ecoulements Geophysiques et Industriels (CNRS/LEGI), Grenoble, France. **Ichiro Fukumori** is Group Supervisor, Ocean Circulation Division, Jet Propulsion Laboratory, Pasadena, California, USA. **Masafumi Kamachi** is Head, Second Laboratory, Oceanographic Research Department, Meteorological Research Institute, Tsukuba, Japan. **Matthew J. Martin** is Ocean Data Assimilation Scientist, Met Office, Exeter, UK. **Kristian Mogensen** is Scientist, European Centre for Medium-Range Weather Forecasting, Reading, UK. **Peter Oke** is Research Scientist, Commonwealth Scientific and Industrial Research Organisation, Hobart, Tasmania, Australia. **Charles Emmanuel Testut** is Scientist, Mercator Océan, Toulouse, France. **Jacques Verron** is Directeur de Recherche, Centre National de la Recherche Scientifique, Laboratoire des Ecoulements Géophysiques et Industriels, Grenoble, France. **Anthony Weaver** is Senior Researcher, Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique, Toulouse, France.

Table 2: Observing systems assimilated by each of the GODAE systems

System	Sea Level	Subsurface Temperature and Salinity	Surface Temperature	Sea Ice
BODAS	Along-track data from satellite altimeters, coastal tide gauges	Argo, CTD, XBT, and moorings	Satellite data	
ECCO-JPL	Along-track data from TOPEX/Poseidon and Jason-1	Argo, CTD, XBT, and moorings	Reynolds SST analysis	
FOAM	Along-track data from satellite altimeters	Argo, CTD, XBT, and moorings	In situ and satellite data	OSI-SAF sea ice analysis
Mercator	Along-track data from satellite altimeters	Argo, CTD, XBT, and moorings	NOAA RTG SST analysis	
MOVE/MRI.COM	Along-track data from all satellite altimeters	Argo, CTD, XBT, and moorings	MGDSST SST analysis	MGDSST sea ice analysis
NCODA	Along-track data from satellite altimeters	Argo, CTD, XBT, moorings, drifting buoys, and gliders	In situ and satellite data	SSM/I and SSMIS sea ice concentration
NEMOVAR	Along-track data from satellite altimeters	Argo, CTD, XBT, and moorings	In situ and satellite data	
TOPAZ	Gridded sea level anomaly maps	Argo	Reynolds SST analysis	AMSR sea ice concentration and sea ice drift products from CERSAT

predictions of the observed variables. Innovations are a measure of model error at the update cycle interval. Correspondingly, the basic data outputs of any assimilation system are the residuals. Residuals are the differences between the analyzed fields and the observed variable after the assimilation. Residuals measure the fit of the analysis to the observations.

OBSERVATIONS ASSIMILATED

The main observations assimilated by the GODAE systems are the sea level anomaly (SLA) data provided by satellite altimeters; subsurface temperature and salinity data from Argo floats, moored and drifting buoys, expendable

bathythermograph (XBT) temperature, and conductivity-temperature-depth (CTD) recorders; in situ and satellite sea surface temperature data; and satellite sea ice concentration and drift data. Table 2 lists the observations assimilated by each GODAE system based on analysis variable and observing system. Observations from satellite altimeters and the various in situ subsurface measuring systems are assimilated by all of the systems, with BODAS also assimilating SLA from tide gauge data. A number of the systems assimilate sea surface temperature (SST) analyses from various sources that have already combined the SST data into a gridded product, such as the

US National Oceanic and Atmospheric Administration (NOAA) real-time global (RTG) and the Japanese Merged Global Daily SST (MGDSST) analyses. Other systems directly assimilate orbital satellite SST retrievals along with in situ SST observations from ships and buoys. NCODA assimilates sea ice concentration data from the Special Sensor Microwave Imager (SSM/I) and Special Sensor Microwave Imager/Sounder (SSMIS) satellite series. Two of the systems assimilate gridded sea ice concentration fields from the French Ocean and Sea Ice Satellite Application Facility (OSI-AF) and Japanese MGDSST. TOPAZ assimilates sea ice concentration retrievals from

the Advanced Microwave Scanning Radiometer (AMSR) and sea ice drift data from the Center for Satellite Exploitation and Research (CERSAT).

ERROR COVARIANCES

A crucial aspect of all ocean data assimilation schemes is the way in which the background and observation error covariance matrices are specified, or the way in which the model and/or observations are perturbed in the case of ensemble schemes. The background error covariance matrix determines how information is spread from the observations to the model grid points and model levels. The background error covariances should also ensure that observations of one model variable produce dynamically consistent corrections in other model variables. Observation errors are typically assumed to be uncorrelated and consist of two parts: a measurement error and a representation error. Measurement errors reflect

the background error, determines the relative weight given to the observation in the analysis. The ratio of the observation error to the background error is expected to be close to unity if the assimilation system fits the data to within prescribed observation error limits and the background errors are consistent with the model-data errors.

Gaussian or second-order autoregressive (SOAR) functions are used to model the background error covariances in many of the GODAE systems (FOAM, MOVE/MRI.COM, NCODA, NEMOVAR). FOAM assumes there are two main sources of model forecast error: one due to errors in the forcing of the model by atmospheric fields, and the other due to model dynamical errors. The background error covariance matrix in FOAM, therefore, is specified to be the combination of two SOAR functions, each with their associated variance and correlation length scale (see Martin et al., 2007, for details). The

Fujii and Kamachi, 2003, for details). NCODA uses a SOAR function where the horizontal correlation length scales vary with location and are specified as the first baroclinic Rossby radius of deformation (Chelton et al., 1998) multiplied by a scaling factor. (The scaling is on the order of 1.3 to 2.8, with small latitude dependence). Flow-dependence is introduced in the analysis by adjusting the horizontal correlations with a tensor computed from forecast model SLA gradients. The flow-dependent tensor tends to spread innovations along rather than across the SLA contours, which are used as a proxy for the circulation field. Vertical correlation length scales in NCODA evolve from one analysis cycle to the next and are computed from forecast vertical density gradients using a change in density mixing criterion. In this way, vertical length scales vary with depth and are large (small) when the water column stratification is weak (strong) (see Cummings, 2005, for details). In NEMOVAR, the model state variables in the background error covariance are transformed to new variables whose background errors are approximately uncorrelated using analytical balance relationships (T-S, hydrostatic, geostrophic, and dynamic height relations). By transforming variables in this way, the background-error covariance matrix can be assumed to be univariate with respect to the transformed variables. The univariate correlations are then modeled implicitly using an anisotropic diffusion operator. The correlation functions implied by the diffusion model are approximately Gaussian (see Weaver et al., 2005, for details).

In other GODAE systems (BODAS, Mercator), time series from prior

“OCEAN DATA ASSIMILATION HAS MATURED TO THE POINT THAT OBSERVATIONS ARE NOW ROUTINELY COMBINED WITH MODEL FORECASTS TO PRODUCE A VARIETY OF OCEAN PRODUCTS.”

what is known about the accuracy of the instruments and the ambient conditions in which the instruments operate. Representation errors, on the other hand, are model dependent and poorly known. Note that the magnitude of the observation error, in combination with

MOVE/MRI.COM analysis scheme uses anisotropic and inhomogeneous horizontal Gaussian decorrelation scales and vertically coupled temperature-salinity (T-S) empirical orthogonal function (EOF) modes to form the background error covariances (see

model integrations are used to provide background error covariance information. BODAS uses an ensemble of long model runs without assimilation to form anomaly fields that have scales and variability that resemble mesoscale ocean

each other, the two error sources are empirically estimated in observation space from auto- and cross-covariances among the observations and their model simulation counterpart (see Fukumori, 2002, for details).

of 11 days that yields global coverage for all data sets assimilated (altimetry, SST, and Argo). The Mercator and TOPAZ assimilation systems are based on a seven-day assimilation cycle, while FOAM and NCODA are run on a daily cycle using synoptic time windows for the observations.

“THE MAIN GOALS OF OCEAN DATA ASSIMILATION ARE TO IMPROVE OUR UNDERSTANDING OF OCEAN CIRCULATION AND MONITOR AND PREDICT CIRCULATION ON ALL RELEVANT TEMPORAL AND SPATIAL SCALES.”

circulation. The background error covariances are then derived directly from this ensemble of intraseasonal anomalies (see Oke et al., 2008, for details). Mercator uses a three-dimensional multivariate EOF decomposition of the model variability, followed by a truncation of the EOF series, to represent the background error covariances. Model variability at the time scale of a week is estimated from ensembles of model anomalies using an a priori interannual simulation as a reference (see Testut et al., 2003, for details).

TOPAZ is unique in its use of an ensemble of model and analysis estimates to evolve the covariances in both space and time. The TOPAZ model ensemble provides both multivariate and flow-dependent background error covariances (see Evensen, 2006, for details). ECCO uses the method of covariance matching (Fu et al., 1993) to estimate both model and data errors. By assuming that errors of the data and those of a model simulation are independent of

PRACTICAL ASPECTS OF IMPLEMENTATION

The assimilation of millions of observations into high-resolution, nonlinear numerical models of ocean circulation is far from trivial. One of the major difficulties in ocean data assimilation is finding practical algorithms that make the solution computationally affordable while maintaining the accuracy of the solution in terms of the fit of the analysis to the observations within the specified error characteristics. In this section we describe some practical aspects of implementing GODAE ocean data assimilation systems.

Cycling and Time Windows

The GODAE systems use different strategies for cycling and for specifying the time window of observations used in the assimilation. For example, BODAS runs on a seven-day cycle in reanalysis mode and on a three- to four-day cycle for real-time operational forecasts, but always uses an observation time window

Quality Control

Various ocean data quality-control procedures are used by GODAE systems to ensure that erroneous data are not assimilated. Some systems use externally processed observations (Mercator uses data processed by the Coriolis data center at Ifremer), while other systems have developed their own automatic quality-control procedures, such as those used by NCODA, FOAM (Ingleby and Huddleston, 2007), and MOVE/MRI.COM (Fujii et al., 2005). For systems executed in reanalysis mode, the observational data have often undergone more extensive delayed-mode, scientific quality-control procedures that are not available in near-real time.

Model Biases

One assumption in the data assimilation methods used by the GODAE systems is that the observations and the model are free of systematic error. This assumption is not valid in practice, and systematic errors in both the models and the observations can cause significant problems in the assimilation and in the skill of the subsequent forecast. In the ECCO near-real-time system, temporal anomalies of sea level and in situ temperature profiles relative to their respective time-means are employed in the assimilation and the time-mean errors of the model are not corrected. In FOAM, schemes have been

developed to deal with biases, including pressure corrections (Bell et al., 2004), systematic errors in the altimeter assimilation (Lea et al., 2008), and biases in the various sources of satellite SST data (Stark et al., 2007). NCODA makes use of satellite SST bias corrections from the data providers, and performs a bias correction on salinity measurements from Argo, which are known to drift with time as floats age.

Initialization

All GODAE systems employ multivariate corrections to the model state vector through statistical relationships, dynamical balance equations, or combinations of the two. A majority of GODAE systems (BODAS, FOAM, Mercator, MOVE/MRI.COM, NCODA, and NEMOVAR) initialize the model for the next forecast using the incremental analysis updating scheme (IAU; Bloom et al., 1996). This approach was found not to be necessary in the TOPAZ system.

point. However, while reducing the computational load, localization can also have a negative impact by degrading the dynamical balance of the analysis fields; thus, it must be implemented with care. BODAS and Mercator divide the model domain into subdomains where the analysis is computed independently. In BODAS, observations are assimilated within a subdomain plus a surrounding halo region, which results in seamless analyses from adjoining subdomains. In Mercator, elliptical influence radii are defined a priori from SLA satellite observations with axes of about 200–500 km, which are on the order of several Rossby radii of deformation at mid latitudes. NCODA uses overlapping analysis volumes that are dynamically created based on observation density and background correlation length scales. A total of eight volume solutions are obtained for each analysis grid point with volume size encompassing up to eight correlation length scales. This combination

profiles) within a radius of 700 km for satellite data and 1000 km for Argo data.

ECCO employs several approximations to reduce the estimation's computational requirements pertaining to derivation of the model state error covariance matrix. These approximations include evaluating independent errors separately from one another (partitioning; Fukumori, 2002), estimating only the most dominant modes of the errors (state reduction; Fukumori and Malanotte-Rizzoli, 1995), and deriving and using the asymptotic limit of the time-evolving error covariances (Fukumori et al., 1993).

The FOAM, MOVE/MRI.COM, and NEMOVAR analysis systems do not localize but rather invoke iterative methods to solve the global minimization problem. The Analysis Correction scheme in FOAM uses a fixed number of 10 iterations to reduce the computational burden of solving the optimal interpolation equation. This number of iterations was determined by performing sensitivity experiments and assessing the increase in accuracy obtained by increasing the number of iterations. The three-dimensional variational methods used in MOVE/MRI.COM and NEMOVAR use preconditioning methods to improve the convergence properties of the minimization (for details see Fujii, 2005, and Tshimanga et al., 2008).

“ALTHOUGH MAJOR CHALLENGES REMAIN, SUBSTANTIAL PROGRESS HAS BEEN MADE DURING GODAE IN DEVELOPING OCEAN DATA ASSIMILATION SYSTEMS.”

Efficiency

A number of techniques are used to improve assimilation efficiency. Many GODAE systems use a form of localization to reduce the need to consider distant observations (where the correlation is essentially zero) in determining the analysis at a particular model grid

of overlapping volumes, and a large number of correlation length scales within a volume, produces smooth analysis increments and reduces the departure from geostrophy that occurs when interpolating different analysis solutions. TOPAZ applies localization by selecting the 50 nearest observations (respectively

ASSIMILATION PERFORMANCE

In this section, we provide examples of GODAE system assimilation performance. Note that additional performance aspects of the GODAE systems are discussed elsewhere in this special issue (see Hernandez et al.).

BODAS

Oke et al. (2008) present a comprehensive assessment of BODAS when applied for version 1.5 of the BLUElink> ReANalysis (BRAN; a multiyear, data-assimilating model run). Through a series of comparisons with both assimilated and withheld observations, Oke et al. (2008) show that, around Australia, BRAN fields are typically within 4–10 cm and 0.4–1°C of observed SLA and SST, respectively; within 1°C and

0.15 psu of observed subsurface T and S, respectively; and within about 0.2 m s⁻¹ of observed near-surface currents from surface drifting buoys. Recent assessments of the operational BLUElink> system shows similar performance to BRAN. Figure 1 is an example of BODAS applied to version 2.1 of BRAN showing a series of comparisons between six-day composite Advanced Very High Resolution Radiometer (AVHRR) SST fields and five-day averaged model

SST fields. Overlaid on the BRAN SST fields are five-day Lagrangian trajectories derived from the time-varying surface velocities computed by BRAN. BRAN shows good agreement with the observed features. This result demonstrates that the Ensemble Optimal Interpolation (EnOI) system used by BODAS, including error estimates therein, while not optimal, is capable of constraining an eddy-resolving model in this highly energetic region.

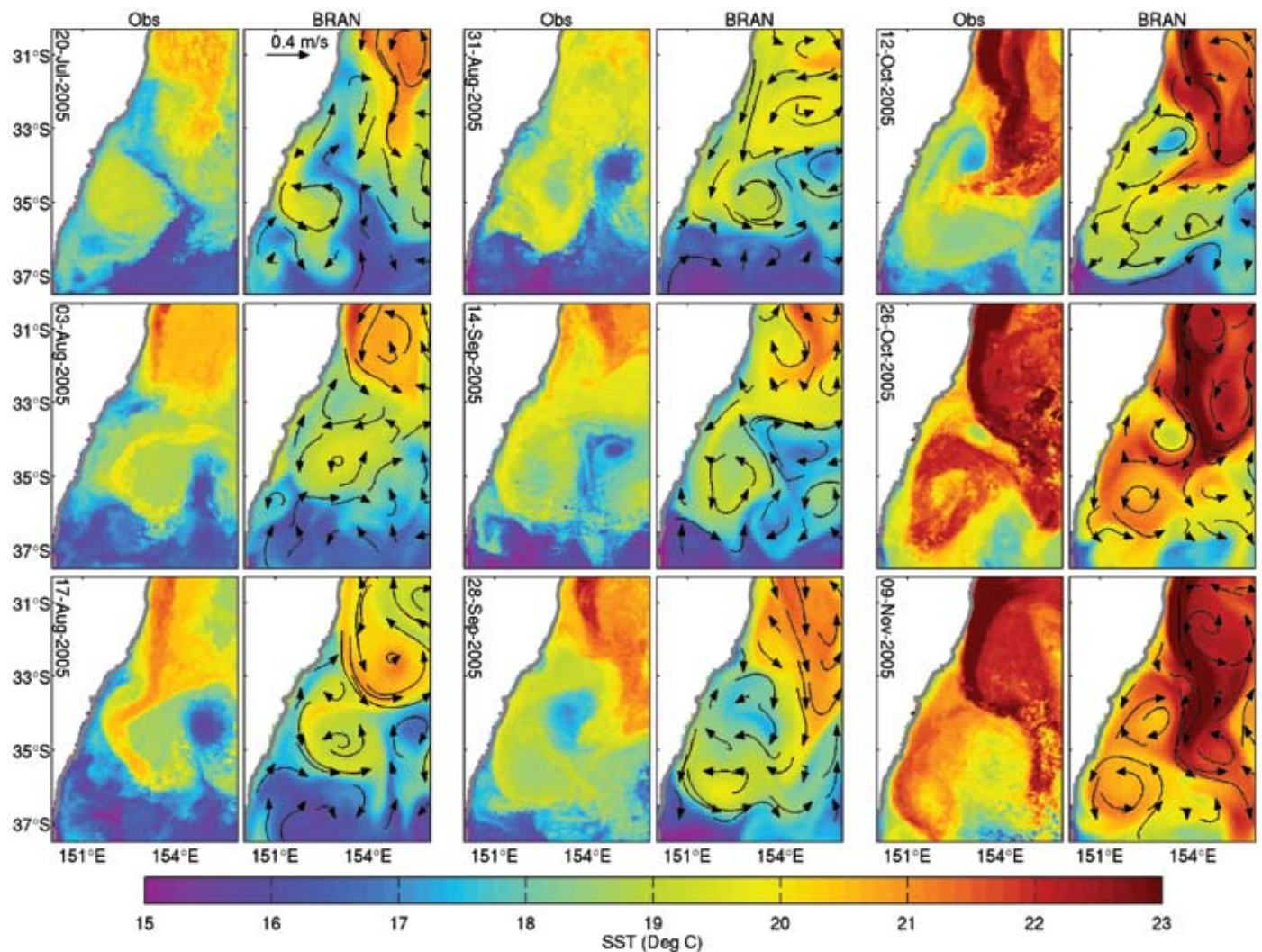


Figure 1. A series of comparisons between six-day composite Advanced Very High Resolution Radiometer sea surface temperature (SST) (columns 1, 3, and 5) and five-day averaged SST in the Tasman Sea, from BRAN2p1 with five-day Lagrangian trajectories from reanalyzed surface velocities overlaid (columns 2, 4, and 6).

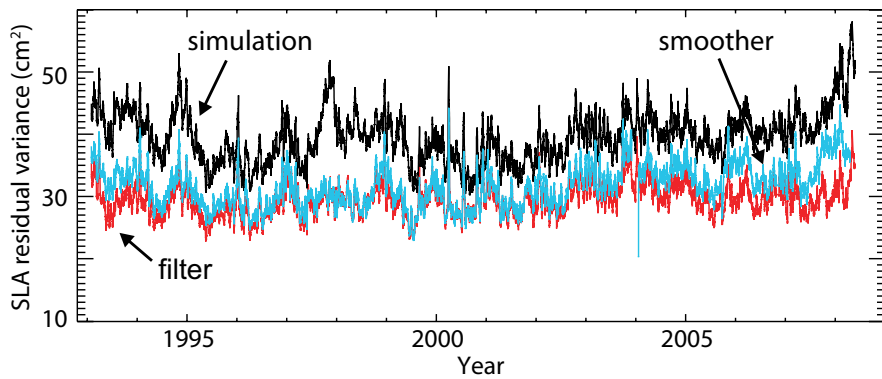


Figure 2. Time series of the variance of ECCO-JPL model-data sea level anomaly differences: simulation (black), Kalman filter (red), and RTS smoother (blue). The smoother variances are nearly indistinguishable from the filter, whereas the simulation residuals are substantially larger.

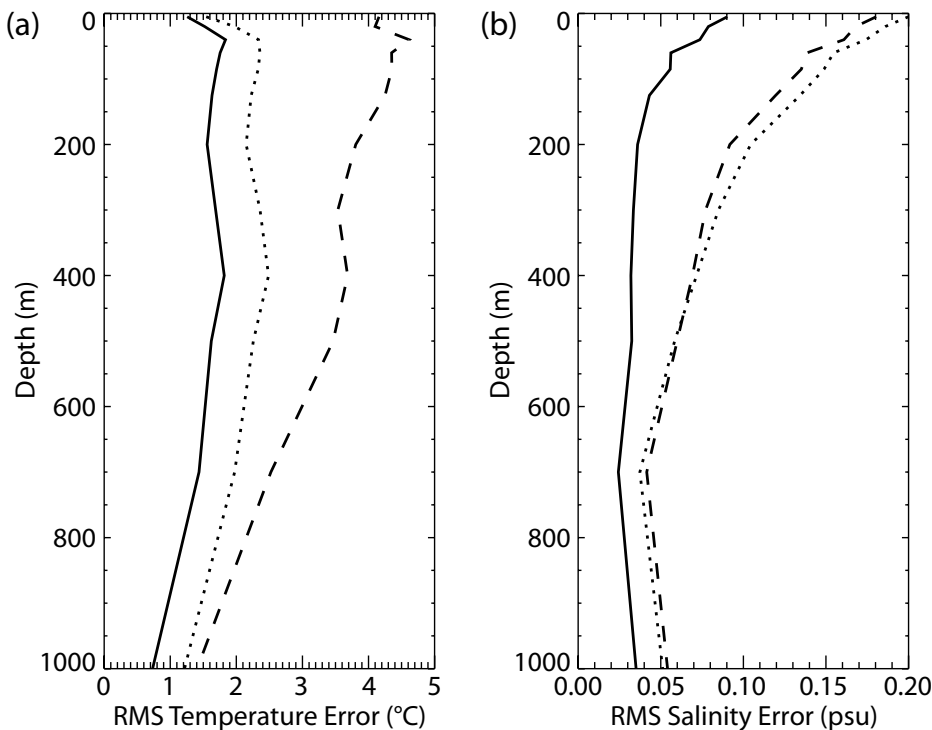


Figure 3. RMS errors in (a) temperature ($^{\circ}\text{C}$) and (b) salinity (PSU) analyses compared to in situ profile observations before they are assimilated, averaged between January 2001 and July 2005 over the North Atlantic model domain. Results are shown for runs that assimilated all in situ data (solid) and all in situ data except Argo (dotted), and for no data assimilation (dashed).

ECCO

Figure 2 illustrates the fidelity of the near-real-time analysis in terms of variance of model-data differences in SLA. The Kalman filter estimate generally has a smaller model-data difference than does a model simulation unconstrained by the observations. Remaining model-data differences are largely due to mesoscale variability not resolved by the model (representation error) and are comparable to theoretical expectations based on formal uncertainty estimates. The smoothed estimate shown is a model simulation forced by winds estimated by the approximate Rauch-Tung-Striebel smoother. On the one hand, due to approximations in the smoother in addition to those in the filter, model-data differences in this smoothed estimate are slightly larger than those in the filtered result. On the other hand, unlike the filtered result, the temporal evolution of this smoothed estimate is physically consistent, owing to the explicit estimation of model error sources (i.e., inaccuracies of winds in this particular example; Fukumori, 2006). This consistency result permits studies of causal mechanisms underlying observed changes in the ocean, such as in mixed-layer temperature and near-surface water mass characteristics (e.g., Kim et al., 2007; Wang et al., 2004).

FOAM

Figure 3 shows the impact of assimilation of in situ profile data in the FOAM $1/9^{\circ}$ North Atlantic model for a set of five-year integrations. The temperature and salinity errors are significantly reduced when assimilating the in situ data. When no Argo data are assimilated, the salinity errors are marginally worse

than with no data assimilated in the top 400 m of the water column, probably due to the temperature-only assimilation disrupting the density structure in the model. These statistics show both the importance of the Argo data and the beneficial impact of data assimilation on the model fields.

Mercator

Figure 4 illustrates the quality of the control for SST in the assimilation from a global $1/4^\circ$ hindcast simulation over the period 2002–2008. Comparison with the zonal temporal evolution of the RTG SST at 3°N shows how the assimilation reproduces well the observed signal in terms of magnitude and phase in regions characterized by Tropical Instability Wave physical processes.

MOVE/MRI

An assimilation experiment (analysis/reanalysis) was conducted from January 1948 to December 2007 for the global and North Pacific systems, and from January 1985 to September 2007 for the western North Pacific system. A total of 138 cases of prediction experiments for Kuroshio path variability south of Japan were also conducted from February 1993 to July 2004. Ninety-day lead time predictions showed realistic predictability (Usui et al., 2008). In the western North Pacific, there are two shallow water masses: the warm, salty Kuroshio water in the subtropical gyre and the cold, fresh Oyashio water in the subpolar gyre. These two water masses merge and produce many mesoscale eddies and additional water masses at the boundary of the subtropical and subpolar gyres in the area east of Tohoku. Figure 5 shows observed

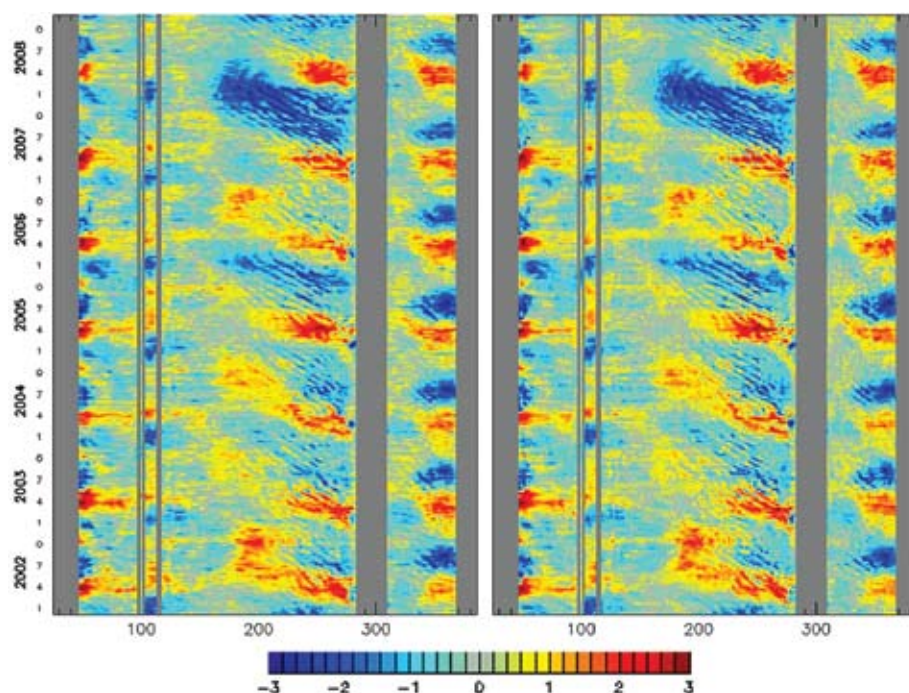


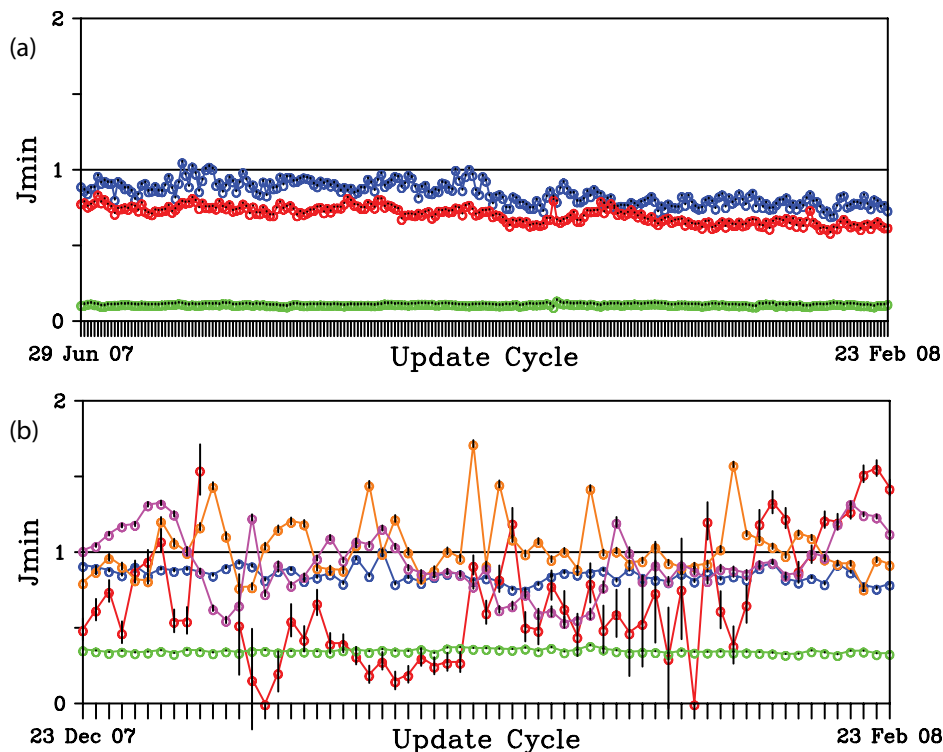
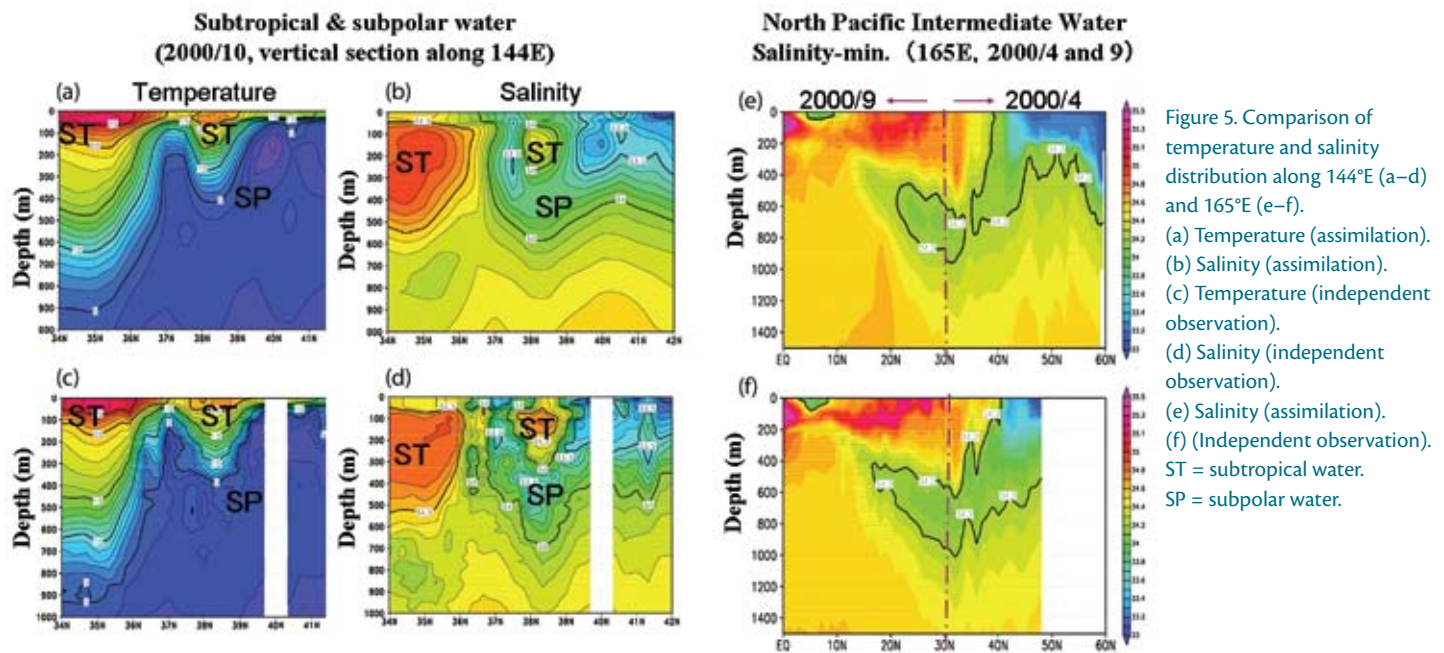
Figure 4. Hovmöller diagrams of sea surface temperature (SST) anomalies at 3°N over the period 2002–2008 for RTG-SST assimilated data (left) and analyzed model SST (right) from a Mercator hindcast simulation on the global $1/4^\circ$ configuration.

temperature and salinity distributions along 144°E from Japan Meteorological Agency line measurements. The warm and saline Kuroshio water and its frontal structure are clearly seen south of 36°N , with a Kuroshio warm water eddy between 38°N and 39°N . The Kuroshio warm water eddy is surrounded by cold Oyashio water. Salinity distributions also highlight the Kuroshio warm eddy structure with salty eddy water surrounded by Oyashio freshwater. The right panel shows observed salinity distributions along 137°E from Japan Meteorological Agency line measurements. It shows typical North Pacific Intermediate Water (NPIW) characterized by a minimum salinity core near 600–800-m depth. The depth distribution and lateral extent of NPIW is important for understanding Pacific

decadal oscillations. The assimilation reproduces the salinity distribution of NPIW quite well.

NCODA

Performance of the NCODA analysis system is routinely assessed by examination of model innovation and analysis residual time series from the cycling assimilation. Residual root mean square (RMS) errors are consistently less than innovation RMS errors, indicating that the analysis is making effective use of the observations, and residual mean errors are indistinguishable from zero for all analysis variables, indicating an unbiased analyzed state. Consistency of the specified background and observation error variances with the innovation vector is monitored each update cycle by the J_{\min} diagnostic (Daley and Barker, 2001).



When normalized by the number of observations assimilated, the expected value of the J_{min} is one. If $J_{min} < 1.0$, either the observation or background errors are specified too large; if $J_{min} > 1.0$, the errors are too small or erroneous data are being assimilated. Figure 6a shows an eight-month time series of daily J_{min} diagnostics computed for selected analysis variables in the Atlantic basin of the global HYCOM/NCODA system. For the three-dimensional temperature and SST analysis variables, J_{min} values are close to one during the first four months of the assimilation, but slowly drift to values less than one in the last half of the assimilation. For altimeter SLA and synthetic temperature profiles (Figure 6b), however, the J_{min} diagnostics indicate that the error variances assigned to these data types are clearly incorrect. For profile observing systems, Figure 6b shows that error variances specified for Argo float temperatures are nearly

correct, while the J_{\min} diagnostics for XBT, CTD, and fixed buoys show a large amount of day-to-day fluctuation. The daily fluctuations of XBT, CTD, and buoy diagnostics are likely due to the fact that the actual forecast background error has significant variations, while the error variances specified in the assimilation may be correct on average but not in specific situations. This problem is most apparent for XBT and CTD observing systems, which sample sporadically in both space and time. Argo, on the other hand, is a global observing array with more than 300 floats surfacing around the globe each day. The consistency of the Argo sampling pattern likely helps maintain the consistency of the temperature background error variances in the HYCOM/NCODA assimilation.

NEMOVAR

A cycled 3DVAR experiment using a global version of NEMOVAR has been conducted for the 20-year period 1987–2006. The assimilated data are T and S profiles from the ENACT/ENSEMBLES quality-controlled database (Ingleby and Huddleston, 2007). The surface forcing fluxes are derived from ERA40. Results from the experiment are summarized in Figure 7, which shows the global mean and RMS of the model fit to the T and S data. The control analysis, produced without assimilating data, is too warm and too salty compared to observations (thin red curves). The assimilation improves the mean and RMS fit to the data both in the analysis and background, the fit for the former being somewhat better, as expected. Note that the fit to the data achieved by the 3DVAR analysis is degraded with the Incremental Analysis Updates (IAU) procedure, with

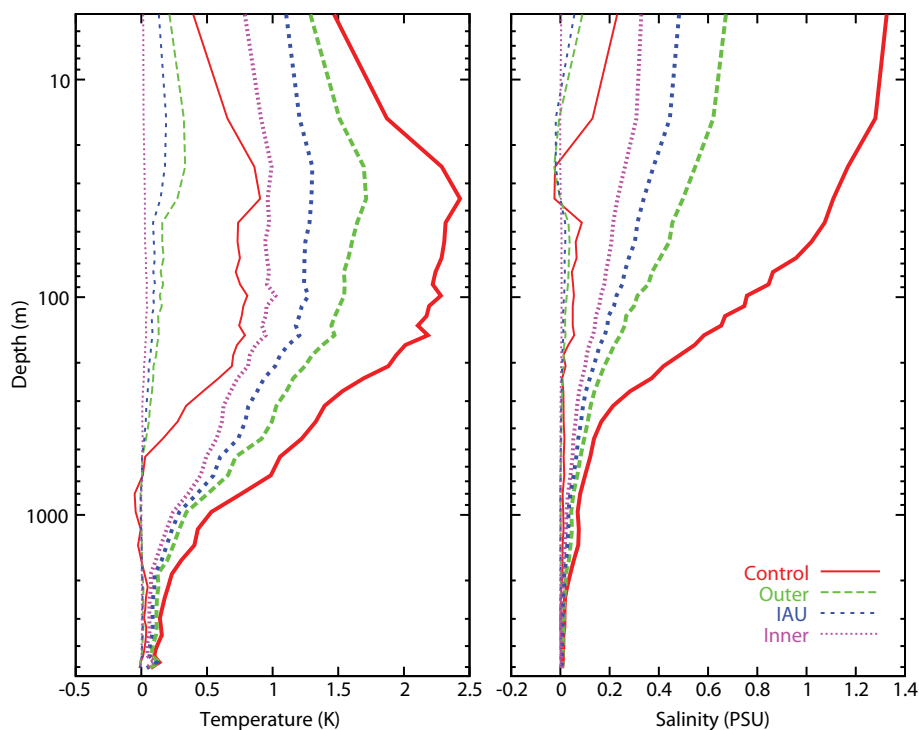


Figure 7. Vertical profiles of the 1988–2006 time-mean (thin curves) and RMS (thick curves) of the globally averaged model-minus-observation differences for temperature (left panel) and salinity (right panel) in a $1^\circ \times 1^\circ$ global NEMO configuration. The display shows statistics from a control experiment in which no profile data are assimilated (red solid curve) and from a 3DVAR (NEMOVAR) experiment. For the 3DVAR experiment, the statistics are shown for the background-minus-observations (green dashed curve labeled “Outer”), the analysis-minus-observations after minimization (pink dotted curve labeled “Inner”), and the analysis-minus-observations after IAU (blue dashed curve).

the model-data fit after IAU lying in between the 3DVAR analysis residual and the innovation. This property of IAU achieves temporal smoothness in the analyses at the expense of degrading the fit to the data, especially near the beginning of the assimilation window.

TOPAZ

Figure 8 shows the forecast skills for sea ice in the Barents Sea; the forecast does slightly better than persistence and a larger reduction is attained by data assimilation, which reduces the error by 20% at each analysis. The residual SLA error in the North Atlantic after

assimilation is between 10 and 15 cm, which is, in part, the result of a seasonal trend. It is unlikely that surface forcing alone is the cause of these errors, because surface forcing variability should have been captured in the ensemble-derived background error covariance matrix. Rather, the SLA bias may indicate a too-coarse vertical resolution of the model and an underestimation of the thermal expansion effects.

FUTURE PLANS

A number of GODAE systems plan to upgrade to more advanced data assimilation schemes in the coming years.

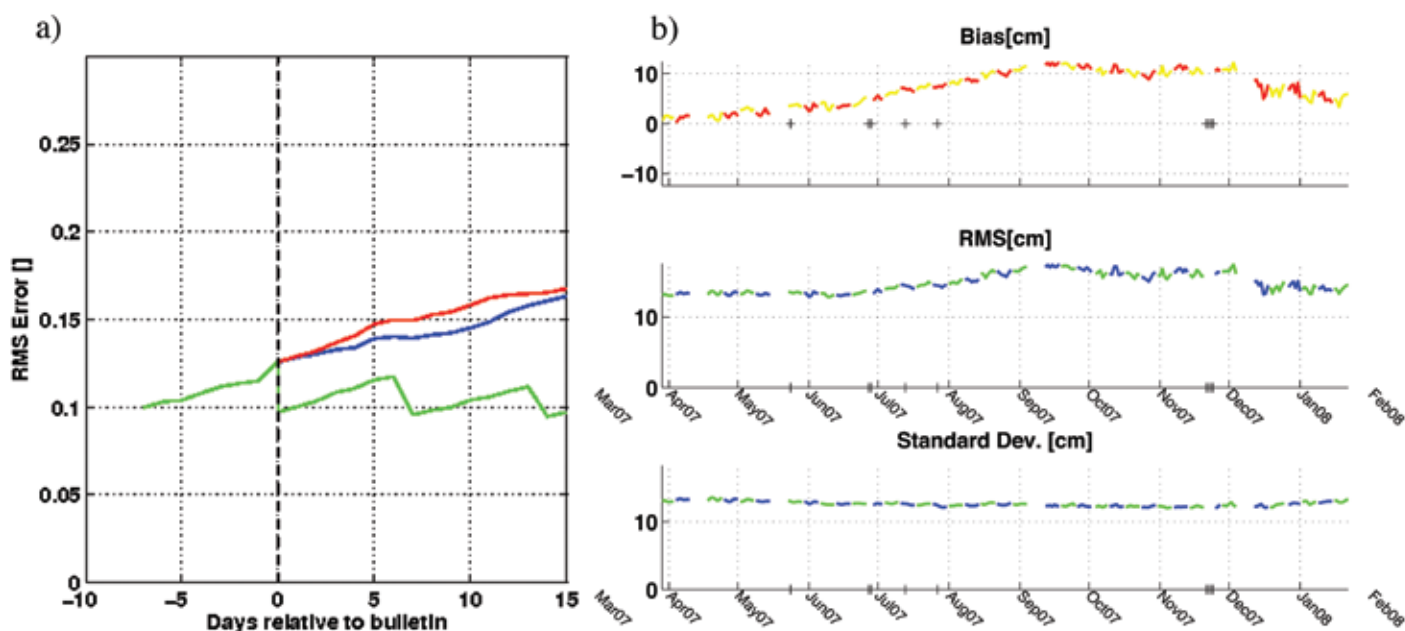



Figure 8. Performance of the TOPAZ data assimilation system. (a) Forecast skills of ice concentrations in the Barents Sea on average for the period April to September 2007; blue is the forecast, red is persistence, and green the analysis. (b) Statistics of residual errors of sea level anomaly in the North Atlantic Drift from April 2007 to February 2008; alternate weeks are drawn in different colors to highlight the assimilation steps.

BODAS will move to a hybrid EnOI/EnKF scheme, FOAM and NCODA will implement 3DVAR schemes, MOVE/MRI.COM and NEMOVAR will both be developed into 4DVAR schemes, and ECCO-JPL plans to combine the filter/smoothing estimates with the ECCO 4DVAR estimates. Significant improvement to the statistical parameterizations used in the systems is also planned. Proper specification of the background and observation error covariances has been identified as critical for the success of GODAE assimilation systems. The error covariances must dynamically connect all model state variables, evolve with time, and accurately reflect the proper balance between observation errors and forward model errors. This problem is an active area of research in ocean data assimilation. New observing systems, both satellite based and in

situ, continue to be deployed at the national and international level. Use of these new observing systems and improved use of existing observing systems is clearly required in the future. Nevertheless, as the GODAE experiment ends and sustained assimilation activities begin, GODAE data assimilation systems will continue to evolve and improve with time.

ACKNOWLEDGEMENTS

The authors wish to thank the reviewers of the paper, Eric Dombrowsky and Pierre De Mey, for their comments and help on improving the manuscript. Pierre De Mey was especially helpful for sharing his unpublished manuscript "Approaches to Data Assimilation within GODAE." The authors give special thanks to Mike Bell, one of the editors of this special issue, for his diligence

on shaping the content of the paper to clearly convey descriptions and results of the GODAE assimilation systems. 

REFERENCES

- Bell, M.J., M.J. Martin, and N.K. Nichols. 2004. Assimilation of data into an ocean model with systematic errors near the equator. *Quarterly Journal of the Royal Meteorological Society* 130:873–893.
- Bloom, S.C., L.L. Takacs, A.M. Da Silva, and D. Ledvina. 1996. Data assimilation using incremental analysis updates. *Monthly Weather Review* 124:1,256–1,271.
- Brasseur, P. 2006. Ocean data assimilation using sequential methods based on the Kalman filter. From theory to practical implementations. Pp. 271–316 in *Ocean Weather Forecasting: An Integrated View of Oceanography*. E.P. Chassignet and J. Verron, eds, Springer.
- Brasseur P., P. Bahurel, L. Bertino, F. Birol, J.-M. Brankart, N. Ferry, S. Losa, E. Rémy, J. Schröter, S. Skachko, and others. 2005. Data assimilation for marine monitoring and prediction: The Mercator operational assimilation systems and the MERSEA developments. *Quarterly Journal of the Royal Meteorological Society* 131:3,561–3,582.

- Chelton, D.B., R.A. DeSzoeke, M.G. Schlax, K.E. Naggar, and N. Siwertz. 1998. Geographical variability of the first baroclinic Rossby radius of deformation. *Journal of Physical Oceanography* 28:433–460.
- Cummings, J.A. 2005. Operational multivariate ocean data assimilation. *Quarterly Journal of the Royal Meteorological Society* 131:3,583–3,604.
- Daley, R., and E. Barker. 2001. NAVDAS formulation and diagnostics. *Monthly Weather Review* 129:869–883.
- Dombrowsky, E., L. Bertino, G.B. Brassington, E.P. Chassignet, F. Davidson, H.E. Hurlburt, M. Kamachi, T. Lee, M.J. Martin, S. Mei, and M. Tonani. 2009. GODAE systems in operation. *Oceanography* 22(3):80–95.
- Evensen, G. 2006. *Data Assimilation: The Ensemble Kalman Filter*. Springer, 280 pp.
- Fu, L.-L., I. Fukumori, and R.N. Miller. 1993. Fitting dynamic models to the Geosat sea level observations in the Tropical Pacific Ocean. Part II: A linear, wind-driven model. *Journal of Physical Oceanography* 23:2,162–2,181.
- Fujii, Y. 2005. Preconditioned Optimizing Utility for Large-dimensional analyses (POpULAR). *Journal of Oceanography* 61:167–181.
- Fujii, Y., and M. Kamachi. 2003. Three-dimensional analysis of temperature and salinity in the equatorial Pacific using a variational method with vertical coupled temperature-salinity EOF modes. *Journal of Geophysical Research* 108(C9), 3297, doi:10.1029/2002JC001745.
- Fujii, Y., S. Ishizaki, and M. Kamachi. 2005. Application of nonlinear constraints in a three-dimensional variational ocean analysis. *Journal of Oceanography* 61:655–662.
- Fukumori, I. 2002. A partitioned Kalman filter and smoother. *Monthly Weather Review* 130:1,370–1,383.
- Fukumori, I. 2006. What is data assimilation really solving, and how is the calculation actually done? Pp. 317–342 in *Ocean Weather Forecasting: An Integrated View of Oceanography*. E.P. Chassignet and J. Verron, eds, Springer.
- Fukumori, I., and P. Malanotte-Rizzoli. 1995. An approximate Kalman filter for ocean data assimilation: An example with an idealized Gulf Stream model. *Journal of Geophysical Research* 100:6,777–6,793.
- Fukumori, I., J. Benveniste, C. Wunsch, and D.B. Haidvogel. 1993. Assimilation of sea surface topography into an ocean circulation model using a steady-state smoother. *Journal of Physical Oceanography* 23:1,831–1,855.
- Hurlburt, H.E., G.B. Brassington, Y. Drillet, M. Kamachi, M. Benkiran, R. Bourdallé-Badie, E.P. Chassignet, G.A. Jacobs, O. Le Galloudec, J.-M. Lellouche, and others. 2009. High-resolution global and basin-scale ocean analyses and forecasts. *Oceanography* 22(3):110–127.
- Ingleby, B., and M. Huddleston. 2007. Quality control of ocean temperature and salinity profiles: Historical and real-time data. *Journal of Marine Systems* 65:158–175.
- Kim, S.-B., T. Lee, and I. Fukumori. 2007. Mechanisms controlling the interannual variation of mixed layer temperature averaged over the NINO3 region. *Journal of Climate* 20:3,822–3,843.
- Lea, D., J.-P. Drecourt, K. Haines, and M.J. Martin. 2008. Ocean altimeter assimilation with observational- and model-bias correction. *Quarterly Journal of the Royal Meteorological Society* 134:1,761–1,774.
- Martin, M.J., A. Hines, and M.J. Bell. 2007. Data assimilation in the FOAM operational short-range ocean forecasting system: A description of the scheme and its impact. *Quarterly Journal of the Royal Meteorological Society* 133:981–995.
- Oke, P.R., G.B. Brassington, D.A. Griffin, and A. Schiller. 2008. The Bluelink Ocean Data Assimilation System (BODAS). *Ocean Modelling* 21:46–70.
- Stark, J.D., C.J. Donlon, M.J. Martin, and M.E. McCulloch. 2007. OSTIA: An operational, high resolution, real time, global sea surface temperature analysis system. Paper presented at Oceans '07 IEEE Conference, "Marine Challenges: Coastline to Deep Sea," June 18–21, 2007, Aberdeen, Scotland.
- Testut C.-E., P. Brasseur, J.M. Brankart, and J. Verron. 2003. Assimilation of sea-surface temperature and altimetric observations during 1992–1993 into an eddy-permitting primitive equation model of the North Atlantic Ocean. *Journal of Marine Systems* 40–41:291–316.
- Tshimanga, J., S. Gratton, A.T. Weaver, and A. Sartenauer. 2008. Limited-memory preconditioners, with application to incremental four-dimensional variational data assimilation. *Quarterly Journal of the Royal Meteorological Society* 134:753–771.
- Usui, N., H. Tsujino, H. Nakano, and Y. Fujii. 2008. Formation process of the Kuroshio large meander in 2004. *Journal of Geophysical Research* 113, C08047, doi:10.1029/2007JC004675.
- Wang, O., I. Fukumori, T. Lee, and B. Cheng. 2004. On the cause of eastern equatorial Pacific Ocean T-S variations associated with El Niño. *Geophysical Research Letters* 31, L15309, doi:10.1029/2004GL020188.
- Weaver, A.T., C. Deltel, E. Machu, S. Ricci, and N. Daget. 2005. A multivariate balance operator for variational ocean data assimilation. *Quarterly Journal of the Royal Meteorological Society* 131:3,605–3,625.
- Wunsch, C. 2006. *Discrete Inverse and State Estimation Problems: With Geophysical Fluid Applications*. Cambridge University Press, New York, NY, 384 pp.