### Validation and Operational Implementation of the Navy Coastal Ocean Model Four Dimensional Variational Data Assimilation System (NCOM 4DVAR) in the Okinawa Trough

#### Scott Smith, Hans Ngodock, Matthew Carrier, Jay Shriver, Philip Muscarella and Innocent Souopgui

Abstract The Navy Coastal Ocean Model Four-Dimensional Variational Assimilation (NCOM 4DVAR) system is an analysis software package that is designed to supplement the current capability of the operational analysis/prediction system known as the Relocatable Navy Coupled Ocean Model (Relo NCOM) system. The present assimilation component of Relo NCOM employs the Navy Coupled Ocean Data Assimilation Three-Dimensional Variational Assimilation (NCODA 3DVAR) system to process and assimilate observations. The NCOM 4DVAR, on the other hand, uses a representer based 4DVAR method and has been found to improve the forecast-skill for several regional applications. This chapter presents the results of validation experiments performed in the Okinawa Trough. The analysis and resulting forecast skill of the two assimilation methods within Relo NCOM (NCOM 4DVAR and NCODA 3DVAR) are compared, and the operational implementation of NCOM 4DVAR is examined to verify that it satisfies operational constraints. The metrics used to validate the NCOM 4DVAR system include: computational efficiency, scalability, robustness, and the prediction accuracy of temperature, sea surface height, and sonic layer depth through NCOM 4DVAR and NCODA 3DVAR analyses. Forecast skill metrics are computed using surface observations of temperature, salinity and sea surface height, and profile observations from gliders and AXBTs (aerial expendable bathythermograph). Overall, the validation reveals that NCOM 4DVAR has lower root mean square errors for both analyses and forecasts than the operational NCODA 3DVAR system.

I. Souopgui

Stennis Space Center, University of Southern Mississippi, Hattiesburg, MS, USA

S. Smith  $(\boxtimes)$  · H. Ngodock · M. Carrier · J. Shriver · P. Muscarella Stennis Space Center, Naval Research Laboratory, Bay St. Louis, MS, USA e-mail: Scott.Smith@nrlssc.navy.mil

<sup>©</sup> Springer International Publishing Switzerland 2017 S.K. Park and L. Xu (eds.), *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications (Vol. III)*, DOI 10.1007/978-3-319-43415-5\_18

#### 1 Introduction

The Navy Coastal Ocean Model four-dimensional variational (NCOM 4DVAR) system is a data assimilative nowcast/forecast ocean modeling and prediction system developed at the Naval Research Laboratory (NRL) for use at the Naval Oceanographic Office (NAVOCEANO) (Smith et al. 2015). This system is built to be used within the same framework as the Relocatable NCOM (Relo NCOM), which is the present operational ocean analysis and prediction tool that the Navy uses for non-global applications. The current data assimilation component of Relo NCOM uses the three-dimensional variational data assimilation method, performed by the Navy Coupled Ocean Data Assimilation system (NCODA 3DVAR, Smith et al. 2012). The newly developed NCOM 4DVAR system is designed to supplement NCODA 3DVAR allowing the user to select either system depending on the application at hand. Regardless of which assimilation option is selected, Relo NCOM uses the same forcing, initial and boundary conditions, and the same ocean model (NCOM) for its forecasting component.

Most ocean models have reduced accuracy and prediction skill at regional and coastal scales where the prediction of tracers, currents, and acoustic properties are important for search and rescue operations, hydrocarbon/chemical spill simulations, environmental prediction, and other Navy operations. While the currently operational NCODA 3DVAR may be ideal for global and large basin scales due to its computational efficiency, NCOM 4DVAR has improved analysis/forecasting capabilities and has shown that it can be operated at sufficiently high resolution in coastal and/or regional areas in a reasonable amount of time (Ngodock and Carrier 2014b). The NCOM 4DVAR is able to provide an improved analysis by accounting for observations at their actual collection times, rather than assuming the observations occur at the same time as in 3DVAR. Also in 4DVAR, observation corrections are temporally correlated and their influence is propagated throughout the entire assimilation window via the model dynamics. This allows more information to be extracted and utilized from sparse observations, thereby producing a more accurate and dynamically consistent analysis, which in turn increases the forecasting predictability skill. Another advantage that NCOM 4DVAR has is the capability to directly assimilate velocity (Carrier et al. 2014) and sea surface height (SSH) (Ngodock et al. 2015) observations without having to use synthetic observations. Synthetic observations consist of temperature and salinity profiles that are derived from SSH observations (Fox et al. 2002). The generation of synthetic observations is required in the NCODA 3DVAR assimilation system because there are no model dynamics or cross-covariances to correlate SSH observations to the other variables.

In this study, the resulting analyses and forecasts from three experiments are analyzed and compared. All experiments are performed using the operational implementation of Relo NCOM for the Okinawa Trough. The only difference between the experiments is that the first uses NCODA 3DVAR and the second and third experiments use NCOM 4DVAR. Two separate NCOM 4DVAR experiments are performed and presented that use different methods of assimilating sea surface height (SSH) observations (described in Sect. 2.3). Every effort is made to keep the forcing, parameters, data, and data processing as similar as possible between the experiments, so that the primary aspect being compared is the assimilation method.

Section 2 describes the Relo NCOM system along with its major components that are relevant to the validation experiments in this study. Then in Sect. 3, the setup of the validation testing experiments for the Okinawa Trough are discussed, followed by the results in Sect. 4. Section 5 goes over some of the implications of applying the NCOM 4DVAR system operationally. Finally, some conclusions are provided in Sect. 6.

#### 2 Components of Relo NCOM

The Relo NCOM system is a flexible data assimilation/forecasting system (Rowley 2010), with most model configuration parameters available for the user to define. The Relo NCOM system consists of a suite of scripts that efficiently handle the input and output data streams, NCODA data processing, the data assimilation, and NCOM forecasts. It also performs the preparation of a new domain, which includes interpolating and setting up the initial and boundary conditions and surface forcing. The initial and boundary conditions are extracted from a larger model, such as the global Hybrid Coordinate Ocean Model (HYCOM) (Metzger et al. 2014). The surface forcing fields can come from the Navy Operational Global Atmospheric Prediction System (NOGAPS, Rosmond 1992; Rosmond et al. 2002); Coupled Ocean Atmosphere Mesoscale Prediction System (COAMPS; Hodur 1997) products generated at the Fleet Numerical Meteorology and Oceanography Command (FNMOC); from COAMPS raw output; or now from the Navy Global Environmental Model (NAVGEM, Hogan et al. 2014). In most cases, atmospheric model wind stresses, radiation fluxes, atmospheric pressure, temperature, and humidity are prepared for the NCOM model, and bulk flux formulae are used in NCOM to calculate surface heat fluxes (Rowley 2010).

#### 2.1 Navy Coastal Ocean Model (NCOM)

The Navy Coastal Ocean Model (NCOM, Martin 2000) is the ocean forecasting component of Relo NCOM. . NCOM was developed primarily from two existing ocean circulation models, the Princeton Ocean Model (POM) (Blumberg and Mellor 1983; 1987) and the Sigma/Z-level Model (SZM) (Martin et al. 1998). NCOM has a free-surface and is based on the primitive equations and hydrostatic, Boussinesq, and incompressible approximations. Turbulent mixing is parameterized by the Mellor-Yamada Level-2.5 (MYL2.5) turbulence closure parameterization (Mellor and Yamada 1982) for vertical diffusion and the Smagorinsky scheme (Smagorinsky 1963) for horizontal diffusion (Carrier et al. 2014). The vertical

mixing enhancement scheme of Large et al. (1994) is used for parameterization of unresolved mixing processes occurring at near-critical Richardson numbers. A source term is included in the model equations to allow for river input and runoff inflows (Barron et al. 2007).

As in the POM, NCOM employs a staggered Arakawa C grid with an orthogonal-curvilinear horizontal grid orientation. Spatial finite differences are mostly second-order centered, but higher-order spatial differences are optional. NCOM features a leapfrog temporal scheme with an Asselin filter to suppress time splitting. Most terms are handled explicitly in time, but surface wave propagation and vertical diffusion are solved implicitly (Martin 2000). In the vertical, NCOM can be configured with terrain-following free-sigma or fixed sigma, or constant *z*-level surfaces or their combination (Barron et al. 2006). Typically, one of two types of combinations is used: the first is a hybrid sigma and *z*-level combination with sigma coordinates applied from the surface down to a designated depth (100–200 m depending on where the shelf break is located), and *z*-levels below this specified depth. The second vertical grid choice typically used is the general vertical coordinate (GVC) grid consisting of a three-tiered structure: (1) a near-surface "free" sigma grid that expands and contracts with the movement of the free surface, (2) a "fixed" sigma, and (3) a *z*-level grid allowing for "partial" bottom (Martin et al. 2008).

#### 2.2 Navy Coupled Ocean Data Assimilation 3D Variational Analysis (NCODA 3DVAR) System

NRL developed and implemented an ocean data analysis component of COAMPS called the Navy Coupled Ocean Data Assimilation System (NCODA; Cummings 2005). The version of NCODA used operationally and in this study employs the 3DVAR method and is capable of processing observations from a large number of different platforms. These include, but are not limited to: satellite sea surface temperature (SST), SSH/altimetry, satellite microwave-derived sea ice concentration, and in situ surface and profile data from ships, drifters, fixed buoys, profiling floats, XBTs (expendable bathythermographs), AXBTs (aerial expendable bathythermographs), CTDs (conductivity, temperature, and depth), and gliders. The observational data are prepared and processed through the NCODA automated data quality control system (NCODA-QC) which identifies spurious observations compared against climatological or model fields and associated variability information (Cummings 2011). Observations that satisfy the quality control are then passed into another NCODA module called NCODA-PREP where they are combined with the previous forecast fields to produce the initial innovations. Observation and forecast errors, and correlation scales are also computed in NCODA-PREP.

The NCODA 3DVAR module reads in the innovations and error covariance information and uses a conjugate gradient routine to minimize a 3D variational cost function to determine the analysis increments in observation space. These increments are then mapped back to the model space using the background error covariances resulting in a set of corrections corresponding to the NCOM forecast fields (Smith et al. 2012). The NCODA 3DVAR system is currently being used operationally at NAVOCEANO in the Relo NCOM, global HYCOM, and COAMPS.

#### 2.3 Navy Coastal Ocean Model Four Dimensional Variational System (NCOM 4DVAR)

The NCOM 4DVAR system operates within the framework of Relo NCOM. The same scripts that are used to operate Relo NCOM with NCODA 3DVAR are used to operate the NCOM 4DVAR, with a few additional parameters for the NCOM model adjoint and the specification of the assimilation window. NCOM 4DVAR uses the same data that is processed by NCODA-QC and it also uses NCODA-PREP to process these observations for the specified domain. NCODA-PREP had to be slightly modified to account for the temporal distribution of the observations and to create time dependent innovations that are required for the NCOM 4DVAR. It should be noted that an observation density-reduction option has been added to the NCOM 4DVAR to ensure that no two observations fall within a correlation scale distance of one another, as too many correlated observations can adversely affect the conditioning of the minimization.

The analysis component of NCOM 4DVAR is a variational assimilation system based on the indirect representer method as described by Bennett (1992, 2002) and Chua and Bennett (2001) and uses the tangent linearization (TL) of the NCOM code and its adjoint. The NCOM 4DVAR system is described in detail by Ngodock and Carrier (2014a), and a full derivation of the representer method can be found in Chua and Bennett (2001). Therefore, only an overview is provided here.

The representer method aims to find an optimal analysis solution as the linear combination of a first guess (i.e., prior model solution) and a finite number of representer functions, one per datum,

$$\hat{u}(x,t) = u_F(x,t) + \sum_{m=1}^{M} \hat{\beta}_m r_m(x,t),$$
(1)

where  $\hat{u}(x, t)$  is the analysis solution,  $u_F(x, t)$  is the prior forecast,  $r_m(x, t)$  is the representer function for the *m*th observation, and  $\hat{\beta}_m$  is the *m*th representer coefficient. The representer coefficients are found by solving the linear system,

$$(\mathbf{R} + \mathbf{O})\boldsymbol{\beta} = \mathbf{y} - \mathbf{H}\mathbf{u}_F,\tag{2}$$

where O is the observation error covariance, y is the observation vector, and H is the linear observation operator that maps the model fields to the observation locations. R is the representer matrix defined as,

$$\mathbf{R} = \mathbf{H}\mathbf{M}\mathbf{B}\mathbf{M}^T\mathbf{H}^T,\tag{3}$$

where **M** is the TL of NCOM,  $\mathbf{M}^T$  is the adjoint of NCOM, and **B** is the model error covariance. Since the matrix  $\mathbf{R} + \mathbf{O}$  is symmetric and positive definite, Eq. (2) can be solved for  $\beta$  iteratively using a linear solver, such as the conjugate gradient method. From Eqs. (2) and (3),  $\beta_m$  can be found for each representer by integrating the adjoint and TL models over some number of minimization steps until convergence.

In the NCOM 4DVAR,  $\beta_m$  is found with a pre-conditioned conjugate gradient solver. The preconditioner here follows from Courtier (1997), where  $\beta$  is redefined as  $\mathbf{u} = \sqrt{\mathbf{O}}\beta$  in the minimization step such that Eq. (2) can be expressed as,

$$\left(\sqrt{\mathbf{O}^{-1}}\mathbf{R}\sqrt{\mathbf{O}^{-1}}+I\right)\mathbf{u}=\sqrt{\mathbf{O}^{-1}}(\mathbf{y}-\mathbf{H}\mathbf{u}_F)$$
(4)

This transformation ensures that there is a lower bound of 1 for the eigenvalues, which insures that the condition number will remain reasonably small and allow the conjugate gradient solver to converge relatively quickly. Once  $\beta$  is determined, Eq. (1) is then used to compute the analysis.

The background and model error covariance in NCOM 4DVAR is univariate and follows the work of Weaver and Courtier (2001) and Carrier and Ngodock (2010). This is deemed acceptable as the application of the TL and adjoint models in the minimization and final sweep provide multivariate balance constraints through the linearized dynamics. It has been shown (Yu et al. 2012) that omitting linear balance constraints does not lead to a significant degradation of the final solution in terms of the fit to observations. The univariate error covariance can be decomposed into a correlation matrix and associated error variance such that,

$$\mathbf{B} = \Sigma \mathbf{C} \Sigma, \tag{5}$$

where  $\Sigma$  is a diagonal matrix consisting of the standard deviations of the background error and **C** is a symmetric matrix of background error correlations. In NCOM 4DVAR, the error standard deviations of the background are used at the initialization of the TL model only, whereas the model error (also contained in the matrix  $\Sigma$ ) is used when the adjoint forces the TL model during integration (i.e., as the TL model integrates forward in time). This allows the weak constraint method to correct for the initial condition error while also adjusting the forward model trajectory based on the specification of the model error. The error correlation, for both the model and the background errors, is not directly calculated and stored in NCOM 4DVAR; rather, the effect of the correlation matrix acting on an input vector is modeled by the solution of a diffusion equation (Weaver and Courtier 2001; Yaremchuk et al. 2013; Carrier and Ngodock 2010; Ngodock 2005).

The NCOM 4DVAR includes two methods for assimilating SSH. The first is by creating synthetic profiles of temperature and salinity (T and S) in the same way as

NCODA 3DVAR (Smith et al. 2012). The second option is to assimilate SSH observations directly. Direct assimilation of SSH is not feasible with the NCODA 3DVAR system because it creates gravity waves. A method has been devised for NCOM 4DVAR to overcome this issue by assimilating SSH observations only into the baroclinic mode of the model. NCOM 4DVAR has an internal routine that checks and adjusts the barotropic mode to ensure that it is in balance with the baroclinic mode, it is in dynamic balance with the model and does not produce gravity waves. A more detailed description of this method is provided in Ngodock et al. (2015).

#### **3** Validation Test Description: Okinawa Trough

The study region encompasses both the Okinawa Trough and Ryukyu Islands of Japan, from 17°N to 34°N and 118°E to 134°E (Fig. 1). The Okinawa Trough region is highly dynamic in nature; it has a complex geometry, sharp bathymetry gradient, a strong Kuroshio current, large barotropic and internal tides, significant river input, and frequent typhoon passage. All of these features provide an excellent testing ground to evaluate the predictive capability of the NCOM 4DVAR assimilation system. The Okinawa Trough is located between Taiwan and southern Japan



Fig. 1 The Okinawa Trough model domain, with 3 km horizontal resolution. The study region encompassed both the Okinawa Trough and Ryukyu Islands of Japan, from 17°N to 34°N and 118°E to 134°E (*dashed lines*) and is a seabed feature of the East China Sea; it is an active, initial back-arc rifting basin which formed behind the Ryukyu arc-trench system in the western Pacific Ocean. A large portion of the domain is more than 1,000 m deep with a maximum depth of 2,716 m.

All of the Relo NCOM Okinawa Trough experiments that are compared in this study are 12 months long (Jan 1, 2007–Dec 31, 2007). Each of these experiments uses surface boundary conditions from the global 0.5° NOGAPS (Rosemond et al. 2002) and lateral boundary conditions from a 6 km Relo NCOM Western Pacific domain. This Western Pacific Relo NCOM is performed operationally at NAVO-CEANO and receives its lateral boundary conditions from the global NCOM (Barron et al. 2007). Each experiment employs NCOM configured with 50 layers in the vertical including 25 free-sigma layers extending to a depth of 116 m with constant z-levels extending down to a maximum of 5500 m.

The following three experiments will be used in this comparison: (1) a standard Relo NCOM using the operational implementation of NCODA 3DVAR (EXP1); (2) the NCOM 4DVAR where the SSH observations are assimilated via synthetic profiles of temperature and salinity generated by the Modular Ocean Data Assimilation System (MODAS, Fox et al. 2002) (EXP2); and (3) NCOM 4DVAR assimilating SSH observations through direct assimilation of the along-track measurements (EXP3). The standard implementation of NCODA 3DVAR (EXP1) uses the MODAS synthetic profiles to assimilate SSH observations, as is the case with EXP2.

The 4DVAR assimilation of along-track SSH (EXP3) is included in this comparison because it is a relatively new technique and has outperformed the other two methods in previous tests. In order to run EXP3, an estimated mean SSH field is needed to transform the observations from height anomalies into the SSH form of the ocean model. Since a sufficiently long enough time period of Relo NCOM does not exist for this domain, a 5-year mean SSH field from the global HYbrid Coordinate Ocean Model (HYCOM) is interpolated to the observation locations and added prior to the inclusion of the data within the assimilation. Another reason to include the direct assimilation of SSH in this study is to identify any model-drift that may be present in the cycling forecast from the 4DVAR analysis. There is a possibility that the assimilation of along-track SSH may produce unrealistic corrections to the thermodynamic state of the model. This is not a concern when synthetic profiles are assimilated, as the generation of these profiles uses climatology to constrain the temperature and salinity profiles. This climatological constraint does not exist when SSH observations are assimilated directly. On the other hand, the 4DVAR does constrain the adjustments to the temperature and salinity by the background around which the adjoint and TL models are linearized. The objective is to determine if this constraint is sufficient to prevent unrealistic adjustments to the thermodynamic structure and, therefore, prevent the model solution from drifting far from reality.

The observational data used in these experiments come from several sources: subsurface in situ and profile temperature (T) and salinity (S) observations were collected from XBTs and Argo Floats, SST observations are collected from

NOAA's GAC and LAC satellites, and SSH observations are from altimeter data obtained from the ENVISAT, GFO, and Jason-1 satellites. The SSH data are processed through the ALtimeter Processing System (ALPS; Jacobs et al. 2002), which is available from the Altimetry Data Fusion Center (ADFC) at the Navy Oceanographic Office (NAVOCEANO). A collection of additional glider and AXBT observations are also provided and used in this study.

At 3 kilometers (km) resolution, the Okinawa Trough domain has a spatial size of 535 by 628 grid points and 50 layers; this corresponds to a total of 16,799,000 grid points. Due to the computational cost of NCOM 4DVAR, which involves solving the adjoint and TL models several times within the minimization driver, the total time to run the assimilation for a model grid of this size is operationally prohibitive.

To reduce the computational time it is necessary to run the NCOM 4DVAR assimilation on a reduced resolution grid. For the 4DVAR experiments (EXP2 and EXP3), the model grid is coarsened by interpolating the 3 km model background to a 6 km analysis grid that covers the same region and vertical structure as the original configuration. This is deemed acceptable as the static spatial covariance scales employed by the NCOM 4DVAR are based on the Rossby radius of deformation, which is approximately 40 km for this region. Once the assimilation is complete on the reduced-grid, the analysis increments are projected back to the original 3 km resolution and added to the full-resolution background state to produce the analysis. A series of experiments conducted during the early testing phase for the NCOM 4DVAR in the Okinawa Trough confirmed that a forecast run at 3 km initialized by a 6 km analysis yields a nearly identical solution as one run from a 3 km analysis. This result, coupled with the fact that the computational cost of the analysis is greatly reduced by the use of the coarse-resolution analysis, justifies this method.

#### 4 Validation of NCOM 4DVAR

The results of the validation testing are broken up into 4 subsections. The first is the analysis and 24-h forecast errors of the different experiments as a function of time throughout the 12-month time period. The remaining 3 subsections focus on subsurface predictability. These statistics are only computed over a 3-month time period (Aug–Oct, 2007), because the vast majority of profile observations are collected during this time. It is also important to examine the predictability and persistence of extended forecasts out to 96 h, and it is prohibitive to perform these extended forecasts for the entire year. In the second and third subsections, the time average predictability of the subsurface temperature and salinity is compared with all assimilated profile data (Sect. 4.2); and then with non-assimilated AXBT and glider profile data for independent verification (Sect. 4.3). Finally, in Sect. 4.4, the predictability of sonic layer depth (SLD) is analyzed.

#### 4.1 Time Distribution of Errors

The first comparison performed on these experiments is to examine the seasonality and errors of the analysis and the corresponding 24-h forecast that is generated by NCOM. To do this, a normalized error metric is computed as a function of time over the 12 month time-period of the experiments. This evaluation can also help determine if any model drift is present in the solution; this would manifest itself as an increasing 24-h forecast error with time as the model solution slowly drifts from reality. The error metric employed for this is a normalized mean absolute error that will be referred to as the  $J_{fit}$  measure,

$$J_{fit} = \frac{1}{M} \sum_{m=1}^{M} \frac{|y_m - H_m x|}{\sigma_m},$$
 (6)

where *M* is the total number of observations;  $y_m$ ,  $H_m$ , and  $\sigma_m$  are the observation, observation operator, and observation error, respectively, associated with the *m*th observation; and **x** is the model state (either the forecast or analysis). Equation 6 indicates that if the forecast or analysis fits the collective observations within their corresponding prescribed observation errors, the  $J_{fit}$  value will be at or below one. If the  $J_{fit}$  value is well below the value of one, then this may indicate that the solution is over-fitting the observations, and the prescribed model errors may need to be reduced.

Figure 2 shows the  $J_{fit}$  normalized error of the analysis (red) and the 24-h forecast (blue) for both 4DAR NCOM experiments. In these figures, the dashed black line represents the overall observation error. The normalized errors ( $J_{fit}$ ) are computed relative to all of the observations that are assimilated; or in the case of the 24-h forecast it is all of the observations that will be assimilated in the next cycle. These observations include temperature, salinity and in the case of EXP3, SSH. For both 4DVAR experiments (Fig. 2a, b) the analysis fits within the observation error for the majority of the time period and the 24-h forecasts are generally within 2 standard deviations of the observation error.

The results in this figure also show that the 4DVAR assimilation of synthetic observations (EXP2) is outperforming the 4DVAR assimilating SSH directly (EXP3). This, however, does not necessarily imply that EXP2 is better, because the normalized error metric ( $J_{fit}$ ) computed for these two experiments uses different observations and observation errors. In EXP2 (Fig. 2a) SSH observations are converted to synthetic temperature and salinity observations, and it is these synthetic observations that are assimilated and used in the  $J_{fit}$  value, along with the remaining in situ temperature and salinity observations. The synthetic observations have a relatively high observation error, higher than the real observations of temperature and salinity that are assimilated directly in EXP3 (Fig. 2b). Therefore, the solution in EXP2 will fit within a lower percentage of the observation error.

Figure 3 displays comparisons of SSH normalized error of the 24-h forecasts generated from NCODA 3DVAR (EXP1) and the two 4DVAR NCOM



**Fig. 2** Normalized error of the analysis (*red*) and 24-h forecast (*blue*) from the 4DVAR experiment assimilating **a** SSH via MODAS synthetics (EXP2) and **b** SSH directly (EXP3). The normalized errors are computed using Eq. 6 and are relative to all of the assimilated observations during the year-long Okinawa Trough experiments

experiments (EXP2 and EXP3). In these error metrics, the SSH forecasts are compared to a SSH map product created by the ALtimeter Processing System (ALPS) (Jacobs et al. 2002). Along-track SSH observations are high resolution in the along-track direction, but sparse in the cross-track direction. This makes comparisons with models difficult as the structure of mesoscale eddies cannot be entirely resolved using instantaneous SSH observations. The ALPS SSH product is a 2D optimal interpolation of sea surface height anomalies (SSHA) from multiple altimetry sources using characteristic covariance information regarding the scale of typical ocean eddies, propagation speeds, and time scales. A 5-year HYCOM mean SSH field is added to the ALPS SSH, in the same manner as the along-track SSH observations.

Figure 3a displays a comparison of the 24-h SSH forecast error between EXP1 (black) and EXP2 (red) for the 12 months of the experiments using the  $J_{fit}$  error metric in Eq. 6. In this comparison, EXP2 exhibits similar SSH forecast error as EXP1, albeit lower from January through May, and higher from May through September. This result is not surprising, as both forecasts are generated from



**Fig. 3** Comparison of the 24-h SSH forecast error resulting from the year-long 3DVAR experiment (EXP1, *black*) and the two 4DVAR experiments: **a** assimilation of SSH via MODAS synthetics (EXP2, *red*) and **b** assimilation of SSH directly (EXP3, *red*). The normalized errors are computed using Eq. 6 and are relative to the SSH maps from ALPS

analyses that use synthetic profiles from MODAS to constrain the mass field. Since MODAS synthetics are based on climatology, they are generally more isotropic and slowly varying; therefore, limiting the advantage of 4DVAR over 3DVAR.

Figure 3b compares EXP3 and EXP1 24-h SSH forecast errors. In the case of EXP3, the 4DVAR analysis assimilates the along-track SSH observations directly. As such, the EXP3 24-h SSH forecast exhibits lower error than EXP1 generally throughout the entire 12-month experiment. This indicates that directly assimilating SSH, rather than through derived synthetic profiles of temperature and salinity, yields a superior SSH forecast. This is consistent with theory, as the observation errors for synthetic profiles are relatively high (Ngodock et al. 2015).

#### 4.2 **Profile Distribution Errors**

It is important for the subsurface thermodynamic characteristics to be captured by the model, thus the first comparison presented is the RMS errors computed as a function of depth, rather than time. This error metric, calculated for the forecasts generated from NCODA 3DVAR and NCOM 4DVAR, presents a comparison of the model layer-by-layer error relative to available profile observations.

Figures 4 and 5 show the 24-h forecast laver-by-laver RMS error value comparison between EXP1 and the 2 4DVAR experiments, EXP2 and EXP3, respectively. These error statistics are calculated for temperature (left panel) and salinity (right panel) relative to profile observations from three out of the 12 months of the experiments (August–October). In these figures, the value (N) in the left panel is the total number of profiles used to compute these statistics during the 3-month time period. Each profile consisted of both temperature and salinity observations down to a particular depth, so the total number of temperature and salinity observations used in these comparisons is the same. NCODA-OC calculates synthetic salinity profiles using MODAS for profile observations of just temperature (such as AXBTs). Then, layer-by-layer, RMS error values are computed for each experiment using forecast-observation comparisons during the entire 3-month period. It should be noted that not all of the profiles used for these comparisons went below 1400 m and many were confined to the upper 100 m. The results shown in these figures reveal that both EXP2 and EXP3 outperform EXP1 in predicting temperature, especially within the depth range of 100–600 m. Whereas, the systems are pretty similar at predicting salinity, except that EXP2 does not have the increased error near 350 m that is in EXP1.

Figure 6 is an overlay of all the error profiles in Figs. 4 and 5 for comparison, including their corresponding 96-h forecasts (dashed lines). As expected, the error



Fig. 4 Comparison of 24-h forecast RMS profile errors between EXP2 (*red*) and EXP1 (*black*) for temperature profiles (*left panel*) and salinity profiles (*right panel*). These are from 3-months (August–October). The value N is the total number of profile observations used in these statistics Fig. 5 Comparison of 24-h forecast RMS profile errors between EXP3 (*red*) and EXP1 (*black*) for temperature profiles (*left panel*) and salinity profiles (*right panel*). These are from 3-months (August–October). The value N is the total number of profile observations used in these statistics



**Fig. 6** Comparison of 24-h (*solid*) and 96-h (*dashed*) forecast RMS profile errors between EXP1 (*black*), EXP2 (*red*), and EXP3 SSH (*blue*) for temperature profiles (*left panel*) and salinity profiles (*right panel*). These statistics are computed over 3-months (August–October). The value N is the total number of profile observations used in these statistics

characteristics grow from the 24-h forecast to the 96-h for both 4DVAR systems. However, the gains provided by the 4DVAR analyses do not degrade much over the period of 96-h and the forecasts generated from the 4DVAR analyses continue to demonstrate skill over the forecasts generated by NCODA 3DVAR. It is interesting to point out that the 96-h forecasts of EXP2 and EXP3 have the same, or better, skill than the 24-h forecast of EXP1.

#### 4.3 Profile Errors Relative to Independent Data

In addition to the 12-month experiments, a series of smaller 3-month runs using the NCOM 4DVAR (with direct SSH and synthetic assimilation) and the NCODA 3DVAR experiments were performed with some data types withheld for independent forecast evaluation. For these comparisons, EXP1 and EXP3 are compared with (1) all withheld glider data (Fig. 7) and (2) all withheld AXBT data (Fig. 8).

Figure 7 illustrates the layer-by-layer  $J_{fit}$  values (Eq. 6) for the EXP3 24-h forecast (red) versus the EXP1 24-h forecast (black) for temperature (left panel) and salinity (middle panel) computed against withheld glider observations. It should be noted that the withheld glider observations were also processed through NCODA-PREP. Therefore, the observation counts in the right panel of this figure are the processed glider observations binned into the NCODA analysis layers in time increments of 3 h. Clearly, the forecast using EXP3 outperformed EXP1 according to this independent data comparison, for both temperature and salinity through all model layers. Figure 8 shows the same comparison, but using withheld AXBT data. There is no salinity AXBT data, therefore there is no panel for salinity. Just as in the glider comparison, EXP3 outperforms EXP1 when compared to this independent data set.

Overall, the results from these experiments indicate that the NCOM 4DVAR analysis system, when assimilating SSH observations directly or through synthetic profiles of temperature and salinity, fits the assimilated observations within the



**Fig. 7** The layer-by-layer  $J_{fit}$  error values (Eq. 6) for the EXP3 24-h forecast (*red*) versus the EXP1 24-h forecast (*black*) for temperature (*left panel*) and salinity (*middle panel*) computed against withheld glider observations (*right panel*)





prescribed observation error. Further, the resulting forecasts generated from the NCOM 4DVAR analyses perform equally or better than the forecasts generated from the NCODA 3DVAR analyses, for both subsurface temperature and salinity, and also for model sea surface height.

#### 4.4 Sonic Layer Depth Prediction

Sonic Layer Depth (SLD) is the depth at which sound speed is maximum in the upper water column. SLD is an important quantity to the Navy, because it is the upper boundary of the SOFAR (sound fixing and ranging) channel in which acoustic signals at certain frequencies can become trapped. Therefore, the prediction skill of SLD is one of the more important metrics that the Navy uses in determining the quality of a prediction system. For the comparison in this section, SLD was calculated using NRL's ProfParam software (Helber et al. 2008) for all of the glider and AXBT profile data (collected during 1 August 2007 through 31 October 2007). The analyses, and 24, 48, 72, and 96-h forecasts of EXP1, EXP2, and EXP3 were interpolated to these observation locations and times.

Table 1 provides the overall statistics of SLD prediction skill of each of the three experiments over the 3-month time period. In this table, N is the total number of SLD observations computed from glider and AXBT profiles. The mean difference

**Table 1** Sonic Layer Depth (SLD) prediction errors of the NCODA 3DVAR and NCOM 4DVAR analyses, along with their ensuing 24, 48, 72, and 96-h forecasts. Errors are relative to the SLD computed from all AXBT and glider profile observations during Aug–Oct 2007. The experiments with the best correlation are in bold

	N	RMS error (m)	Correlation coefficient	Mean diff (m)
Analysis				
EXP1	5579	22.59	0.46	-9.07
EXP2	5579	18.07	0.65	-1.79
EXP3	5579	17.85	0.65	-2.02
NCOM 24 h forecast				
EXP1	5600	21.28	0.52	-7.96
EXP2	5600	19.14	0.61	-2.68
EXP3	5600	18.68	0.63	-2.84
NCOM 48 h forecast				
EXP1	5602	20.41	0.55	-7.14
EXP2	5602	19.71	0.59	-3.43
EXP3	5602	19.27	0.60	-3.58
NCOM 72 h forecast				
EXP1	5531	20.25	0.55	-6.75
EXP2	5531	19.82	0.58	-3.51
EXP3	5531	19.54	0.58	-3.76
NCOM 96 h forecast				
EXP1	5469	20.02	0.55	-6.18
EXP2	5469	19.83	0.57	-3.53
EXP3	5469	19.60	0.58	-4.05

(bias) between the SLDs calculated from the prediction system and data (Model SLD–Data SLD) reveal that the analysis and forecast systems are consistently predicting a shallower SLD than the data in all experiments. The RMS errors from these differences, along with their correlation coefficient, demonstrate that both versions of 4DVAR perform better than the NCODA 3DVAR at predicting SLD for the analysis and all forecast lengths.

Figure 9 displays 2D histograms of occurrence counts between matching observation (x-axis) and model (y-axis) SLDs. Here, the model SLD values are interpolated to each observation location and compared to the observed SLD. The SLD matchups are binned in 5 m resolution cells and the number of occurrences of each matchup within each cell is indicated by the colorbar. This is done for the analysis (top row), 24-h forecast (middle row), and 96-h forecast (bottom row) for EXP1 (left-most column), EXP2 (middle column), and EXP3 (right-most column). Note that the color bar is in log scale. Also, there are no observation counts shallower than 10 m (for both data and model), because the ProfParam software used to calculate SLDs does not allow for a SLD below 10 m. Therefore, more occurrence counts concentrated near the diagonal black line signifies that the model is doing well at predicting SLD.



Fig. 9 Okinawa Trough 2D histograms of SLD (m) of NCODA 3DVAR (*left*), NCOM 4DVAR with synthetic SSH assimilation (*middle*), and NCOM 4DVAR with direct SSH assimilation (*right*) analyses (*top*), 24-h forecasts (*middle*) and 96-h forecasts (*bottom*) relative to SLD computed from profile observations during the 3-month time period of 1 August to 31 October 2007. The *diagonal black line* denotes the locations on each histogram where the modelled SLD matches the observed and the *color bar* denotes the number of counts on a log scale

In the EXP1 analysis histogram, there is an unusual band of modeled SLD counts between 10-15 m depth, and there is an overall significant bias towards the model under-predicting SLD relative to the observations (there are more red squares below and to the right of the diagonal black line). In the EXP3 analysis this bias is significantly reduced; and in the EXP2 analysis, one can barely notice the bias. As the forecast proceeds from 24 to 96-h, there is a clear trend of the shallow modelled SLD bias becoming more pronounced in the 4DVAR and the overall SLD prediction capability of 4DVAR moving towards that of EXP1. However, it can be



**Fig. 10** 2D histograms showing the difference in SLD counts between EXP1 and EXP2 analyses (*left*) and 96-h forecasts (right). *Blue (red) squares* signify that the NCOM 4DVAR has more (less) SLD combination counts than NCODA 3DVAR

seen that EXP2 does better than EXP3 at predicting SLD, and both 4DVAR systems perform significantly better than EXP1 (even after 96-h of forecasts).

To better visualize this improvement, Fig. 10 shows the difference in counts between EXP1 and EXP2 for both the analysis and the 96-h forecast. In this figure, a blue box signifies that EXP2 has more counts at that particular SLD comparison. The analysis histogram (left panel of Fig. 10) has mostly blue boxes along and near the diagonal, and red boxes below and to the right; it is clear that the 4DVAR is doing better and that the NCODA 3DVAR has a significant shallow SLD bias. This improvement persists throughout the 96 h of forecast (right panel of Fig. 10).

#### 5 Operational Implementation of the NCOM-4DVAR

The majority of the NCOM 4DVAR experiments were performed at the DoD Supercomputing Resource Center (DSRC) where the average wall clock time for an analysis/forecast cycle was 70 min. Occasionally, if there were a significant number of observations during a cycle, the conjugate gradient solver would take up to ten iterations to converge, and hence take up to 1.5 h to complete a cycle. Most of the cycles, however, required an hour or less. This puts the time it takes to perform the NCOM 4DVAR for a relatively large domain within the operational constraints of NAVOCEANO.

Although there is room for improvement, much effort was put into the parallelization of the NCOM 4DVAR software such that it can operate on as many processors as efficiently possible. The software is parallelized by splitting the horizontal domain into separate tiles that are each assigned to a processor (CPU). Periodically, throughout the operation of the code, information has to be transferred amongst the tiles. Therefore, as the number of CPUs is increased, so does the amount of data that needs to be transferred amongst the tiles, and there comes a limit when adding more CPUs for a particular domain only marginally decreases the total wallclock time (this is the software's scalability). The core of the NCOM 4DVAR software scales relatively efficiently. Through a number of tests, it was determined that the optimal CPU tile size for the NCOM 4DVAR is about 20  $\times$  20 grid points, therefore, 192 CPUs was the optimal number for the Okinawa Trough domain.

To achieve the goal of performing an analysis/forecast cycle in about an hour, a software module was created to allow the analysis and the forecast to operate at different grid resolutions. As discussed in Sect. 3, this module is used to interpolate the high resolution forecast to a coarser resolution to be used as the background for the 4DVAR analyses. After the analysis, the module is used to interpolate the coarse analysis to the high resolution grid for the initial conditions of the forecast. In the Okinawa Trough experiments, the 4DVAR analysis is run at a 6 km resolution and the forecast is run at 3 km resolution. A number of experiments were performed testing the impact of reducing the resolution of the analysis component of the system and it was determined that the impact on forecast skill was negligible, but the reduction in computation time was tremendous.

Even though the NCOM 4DVAR takes significantly longer and requires more resources to operate than the 3DVAR (the average wall clock time to perform an analysis/forecast cycle with the 3DVAR is 5 min on 12 CPUs) the ability to correlate observations with the dynamics over multiple days significantly improves the analysis and ensuing forecast skill. Also, innovations are computed and applied in NCOM 4DVAR at the actual observation time. Whereas, the assimilation window for 3DVAR is only one day and regardless of when the observations are recorded, their innovations are all applied at a single analysis time.

The optimal assimilation window length for the NCOM 4DVAR can vary depending on the region, grid resolution, and the observations being assimilated. A longer assimilation window allows the observations more time to propagate throughout the domain via the model dynamics, which therefore should improve the model covariance and produce a better analysis. However, increasing the assimilation window increases the computational time and allows more time for small errors that may arise from the TL approximation to potentially grow. Figure 11 shows the impact the length of the assimilation window has on the predictability of the 24-h forecast. In this figure, experiments were performed on the Okinawa Trough domain during August 2007 for assimilation windows ranging from one to five days. The RMS errors of the 24-h forecast of temperature and salinity follow the correct pattern and generally decrease as the assimilation window is increased.



The improvement resulting from increasing the assimilation window, however, is minor and doesn't warrant the extra computational cost. This is why just a 3-day assimilation window was used in the experiments presented in this chapter.

#### 6 Conclusions

In this chapter, the accuracy of NCOM 4DVAR was compared to the operational 3DVAR-based Relo NCOM analysis/prediction system. Year-long experiments were performed for the Okinawa Trough domain. For three of the months in these experiments, the forecasts were extended out to 96 h so that the long-term predictability can be examined. A number of different types of observations were used in both the assimilation and validation: SST observations from satellites, subsurface temperature and salinity profile observations from ARGO floats, AXBTs and gliders, and SSH observations from altimeters. In some of the experiments, portions of the profile data were removed from the assimilation and used as an independent validation data set. The overall results from these experiments indicate that the NCOM 4DVAR analysis system, when assimilating SSH observations directly or through synthetic profiles of temperature and salinity, fits the assimilated observations within the prescribed observation error. Further, the resulting forecasts generated from the NCOM 4DVAR perform equally or better than the forecasts generated from the NCODA 3DVAR for subsurface temperature and salinity, model sea surface height, and sonic layer depth. Finally, it is demonstrated that despite the computational requirements of the NCOM 4DVAR exceeding those of NCODA 3DVAR, they are within the operational constraints.

Acknowledgements The authors were supported by the NRL 6.4 NCOM 4DVAR Rapid Transition Project (projects 4727-04 and 4727-14), which was managed by both Space and Naval Warfare Systems Command under program element 063207N and the Office of Naval Research under program element 0602435N. Numerical simulations were performed at the DoD Supercomputing Resource Center (DSRC) with grants of computer time from the HPCMP Variational Assimilation High Performance Computing (HPC) subproject.

#### References

- Barron CN, Kara AB, Martin PJ, Rhodes RC, Smedstad LF (2006) Formulation, implementation, and examination of vertical coordinate choices in the global Navy Coastal Ocean Model (NCOM). Ocean Modell 11:347–375. doi:10.1016/j.ocemod.2005.01.004
- Barron CN, Kara AB, Rhodes RC, Rowley C, Shriver JF (2007) Validation test report for the 1/8° Global Navy Coastal Ocean Model Nowcast/Forecast System. NRL Tech Report NRL/MR/7320–07-9019. Naval Research Laboratory, Stennis Space Center, MS
- Bennet AF (1992) Inverse methods in physical oceanography. Cambridge University Press, New York 347 pp
- Bennett AF (2002) Inverse modeling of the ocean and atmosphere. Cambridge University Press, New York 234 pp
- Blumberg AF, Mellor GL (1983) Diagnostic and prognostic numerical circulation studies of the South Atlantic Bight. J Geophys Res 88:4579–4592
- Blumberg AF, Mellor GL (1987) A description of a three-dimensional coastal ocean circulation model. In: Heaps N (ed) Three-dimensional coastal ocean models. American Union, New York, N.Y., 208 pp
- Carrier MJ, Ngodock H (2010) Background-error correlation model based on implicit solution of a diffusion equation. Ocean Modell 35:45–53
- Carrier M, Ngodock H, Smith S, Jacobs G, Muscarella P, Ozgokmen T, Haus B, Lipphardt B (2014) Impact of assimilating ocean velocity observations inferred from Lagrangian drifter data using the NCOM 4DVAR. Mon Weather Rev 142:1509–1524
- Chua BS, Bennett AF (2001) An inverse ocean modeling system. Ocean Modell 3:137-165
- Courtier P (1997) Dual formulation of four-dimensional variational assimilation. Q J R Meteorol Soc 123:2449–2461
- Cummings JA (2005) Operational multivariate ocean data assimilation. Q J R Meteorol Soc 131:3583–3604
- Cummings JA (2011) Ocean data quality control. In: Schiller A, Brassington G (eds) Operational oceanography in the 21st century. Springer, pp 91–122. doi:10.1007/978-94-007-0332-2\_4
- Fox DN, Barron CN, Carnes MR, Booda M, Peggion G, Gurley JV (2002) The modular ocean data assimilation system. Oceanography 15:22–28
- Helber RW, Barron CN, Carnes MR, Zingarelli RA (2008) Evaluating the sonic layer depth relative to the mixed layer depth. J Geophys Res 113. doi:10.1029/2007JC004595
- Hodur RM (1997) The naval research laboratory's coupled ocean/atmosphere mesoscale prediction system (COAMPS). Mon Weather Rev 125:1414–1430
- Hogan TF, Liu M, Ridout JA, Peng MS, Whitcomb TR, Ruston BC, Reynolds CA, Eckermann SD, Moskaitis JR, Baker NL, McCormack JP, Viner KC, McLay JG, Flatau MK, Xu L, Chen C, Chang SW (2014) The navy global environmental model. Oceanography 27(3):116–125
- Jacobs GA, Barron CN, Fox DN, Whitmer KR, Klingenberger S, May D, Blaha JP (2002) Operational altimeter sea level products. Oceanography 15(1):13–21

- Large WG, McWilliams JC, Doney SC (1994) Oceanic vertical mixing: a review and a model with a nonlocal boundary layer parameterization. Rev Geophys 32:363–403
- Martin PJ, Peggion G, Yip KJ (1998) A comparison of several coastal ocean models. NRL Report NRL/FR/7322-97-9692, Naval Research Laboratory, Stennis Space Center, MS
- Martin P (2000) Description of the Navy Coastal Ocean Model Version 1.0. NRL report NRL/FR/7322-00-9961, Naval Research Laboratory, Stennis Space Center, MS
- Martin PJ, Barron CN, Smedstad LF, Wallcraft AJ, Rhodes RC, Campbell TJ, Rowley C, Carroll SN (2008) Software design description for the Navy Coastal Ocean Model Version 4.0. NRL Report NRL/MR/7320-08-9149, Naval Research Laboratory, Stennis Space Center, MS
- Mellor GL, Yamada T (1982) Development of a turbulence closure model for geophysical fluid problems. Rev Geophys Space Phys 20:851–875
- Metzger EJ, Smedstad OM, Thoppil PG, Hurlburt HE, Cummings JA, Wallcraft AJ, Zamudio L, Franklin DS, Posey PG, Phelps MW, Hogan PJ, Bub FL, Dehaan CJ (2014) US Navy operational global ocean and arctic ice prediction systems. Oceanography 27(3). http://dx.doi. org/10.5670/oceanog.2014.66
- Ngodock HE (2005) Efficient implementation of covariance multiplication for data assimilation with the representer method. Ocean Modell 8(3):237–251
- Ngodock HE, Carrier M (2014a) A 4DVAR system for the Navy Coastal Ocean Model Part I: System description and assimilation of synthetic observations in Monterey Bay. Mon Weather Rev 142. doi:10.1175/MWR-D-13-00221
- Ngodock HE, Carrier M (2014b) A 4DVAR system for the Navy Coastal Ocean Model Part II: strong and weak constraints assimilation experiments with real observations in the Monterey Bay. Mon Weather Rev 142. doi:10.1175/MWR-D-13-00220
- Ngodock HE, Carrier M, Souopgui I, Smith S, Martin P, Muscarella P, Jacobs G (2015) On the direct assimilation of along-track sea surface height observations into a free-surface ocean model using a weak constraints four dimensional variational (4DVAR) method. Q J R Meotorol Soc (in press). doi:10.1002/qj.2721
- Rosmond TE (1992) The design and testing of the Navy Operational Global Atmospheric Prediction System. Weather Forecast 7:262–272
- Rosmond TE, Teixeria J, Peng M, Hogan TF, Pauley R (2002) Navy Operational Global Atmospheric Prediction System (NOGAPS): forcing for ocean models. Oceanography 15:99–106
- Rowley C (2010) Validation test report for the Relo system. NRL Report NRL/MR/7320-10-9216, Naval Research Laboratory, Stennis Space Center, MS
- Smagorinsky J (1963) General circulation experiments with the primitive equations. I: The basic experiment. Mon Weather Rev 91:99–164
- Smith SR, Cummings JA, Rowley C, Chu P, Shriver J, Helber R, Spence P, Carroll S, Smedstad OM (2012) Validation test report for the Navy Coupled Ocean Data Assimilation 3D Variational Analysis (NCODA-VAR) System, Version 3.43. NRL Memorandum Report NRL/MR/7320-11-9363, Naval Research Laboratory, Stennis Space Center, MS
- Smith SR, Carrier MJ, Ngodock HE, Shriver J, Muscarella P (2015) Validation testing report for the Navy Coastal Ocean Model Four-Dimensional Variational Assimilation (NCOM 4DVAR) System, Version 1.0. NRL Memorandum Report NRL/MR/732-14-9574, Naval Research Laboratory, Stennis Space Center, MS
- Weaver A, Courtier P (2001) Correlation modeling on the sphere using a generalized diffusion Equation. Q J R Meteorol Soc 127:1815–1846
- Yaremchuk M, Carrier M, Smith S, Jacobs G (2013) Background error correlation modeling with diffusion operators. In: Park SK, Xu L (eds) Data assimilation for atmospheric, oceanic and hydrologic applications, vol 2. Springer, Berlin, Heidelberg. doi:10.1007/978-3-642-35088-7\_ 15
- Yu P, Kurapov AL, Egbert GD, Allen JS, Kosro AP (2012) Variational assimilation of HF radar surface currents in a coastal ocean model off Oregon. Ocean Modell 49–50:86–104

## Seon Ki Park · Liang Xu Editors

# Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications (Vol. III)

