Ocean Data Impacts in Global HYCOM*

JAMES A. CUMMINGS

Oceanography Division, Naval Research Laboratory, Monterey, California

OLE MARTIN SMEDSTAD

QinetiQ North America, Stennis Space Center, Mississippi

(Manuscript received 14 January 2014, in final form 24 April 2014)

ABSTRACT

The impact of the assimilation of ocean observations on reducing global Hybrid Coordinate Ocean Model (HYCOM) 48-h forecast errors is presented. The assessment uses an adjoint-based data impact procedure that characterizes the forecast impact of every observation assimilated, and it allows the observation impacts to be partitioned by data type, geographic region, and vertical level. The impact cost function is the difference between HYCOM 48- and 72-h forecast errors computed for temperature and salinity at all model levels and grid points. It is shown that routine assimilation of large numbers of observations consistently reduces global HYCOM 48-h forecast errors for both temperature and salinity. The largest error reduction is due to the assimilation of temperature and salinity profiles from the tropical fixed mooring arrays, followed by Argo, expendable bathythermograph (XBT), and animal sensor data. On a per-observation basis, the most important global observing system is Argo. The beneficial impact of assimilating Argo temperature and salinity profiles extends to all depths sampled, with salinity impacts maximum at the surface and temperature impacts showing a subsurface maximum in the 100–200-m-depth range. The reduced impact of near-surface Argo temperature profile levels is due to the vertical covariances in the assimilation that extend the influence of the large number of sea surface temperature (SST) observations to the base of the mixed layer. Application of the adjoint-based data impact system to identify a data quality problem in a geostationary satellite SST observing system is also provided.

1. Introduction

Assessment of the impact of observations on reducing ocean model forecast error from data assimilation is a fundamental aspect of any ocean analysis and forecasting system. The purpose of assimilation is to reduce the model initial condition error. Improved initial conditions should lead to an improved forecast. However, it is likely that not all observations assimilated have equal value in reducing forecasting error. Estimation of which observations are best and the determination of locations where forecast errors are sensitive to the initial conditions are essential for improving the data assimilation system itself and for the design and implementation of future observing systems. Typically, the impact of assimilated observations is assessed using data denial studies in what is referred to as an observing system experiment (OSE). In an OSE a sequence of assimilation runs are performed where one or more components of the observing system are withheld. Impacts are determined by estimating changes in the forecasts compared to a run when all observations are assimilated (Oke and Schiller 2007; Balmaseda and Anderson 2009; Lellouche et al. 2013; Lea et al. 2014). Given the large number of ocean observing systems, an OSE is computationally a very expensive method for determining the value of observations assimilated if it is applied systematically to all observation datasets. An OSE has a further disadvantage in that modifications to the observing systems assimilated from the data denial change the value of the remaining

* Naval Research Laboratory Contribution Number NRL/JA/7320-14-2049.

Corresponding author address: James A. Cummings, Naval Research Laboratory, 7 Grace Hopper Ave., Stop 2, Monterey, CA 93943.
E-mail: james.cummings@nrlmry.navy.mil; ole.smedstad.ctr@nrlssc.navy.mil

DOI: 10.1175/JTECH-D-14-00011.1

observations. Because of the computational expense and difficulty in interpretation of the results, an OSE is performed intermittently and is not a viable method for routine determination of observation data impacts.

An alternative approach has been developed to estimate the impact of the observations on the forecast by using an adjoint sensitivity method (Langland and Baker 2004). The technique computes the variation in forecast error due to the assimilated data. Observation impact is estimated simultaneously for the complete set of observations assimilated. There is no need to selectively add or remove observations in the assimilation to estimate observation value as in an OSE. The procedure is computationally inexpensive and can be used for routine observation monitoring. This aspect of the adjoint method is advantageous, since ocean observing and assimilation/forecast systems are in continuous evolution, requiring an efficient procedure that allows the impact of observations to be regularly assessed. Data impacts can be partitioned for any subset of the data assimilated—instrument type, observed variable, geographic region, or vertical level—with traceability to individual platforms based on station identifiers, or call signs.

The monitoring of impacts of observations on short-range forecasts in numerical weather prediction (NWP) using the adjoint of the data assimilation system has become routine (Cardinali 2009), as evidenced further by the series of World Meteorological Organization (WMO) conferences and technical reports (e.g., Andersson and Sato 2012). Gelaro and Zhu (2009) compare results from a standard OSE with the adjoint-based method and show that the two approaches provide consistent estimates of the overall impacts of the major observing systems assimilated. Moore et al. (2011b) have applied the method in the ocean using the Regional Ocean Modeling System (ROMS) four-dimensional variational data assimilation system (4DVAR; Moore et al. 2011a). They quantified the impact of individual observations and observation platforms on the alongshore and cross-shore transport of the central California Current system associated with wind-induced coastal upwelling. Recently, the data impact method has been adapted for ensemble-based systems by Kalnay et al. (2012) and Ota et al. (2013).

This paper describes application of the adjoint-based procedure to the estimation of the impact of the assimilation of observations on reducing ocean model forecast error in the navy's global Hybrid Coordinate Ocean Model (HYCOM) ocean analysis/forecast system. Forecast error gradients and actual model–data differences are used to estimate the impact of each observation assimilated. Since forecast errors grow and decay at different rates throughout the model domain, a large model–data difference does not necessarily lead to a large data impact. Observations can make small changes to the initial conditions and still have a large data impact if the location of the observation is in a dynamically sensitive region. The results illustrate the types of diagnostics that can be routinely obtained with the adjoint-based method in an operational context.

The paper is organized as follows. Section 2 gives a brief description of the global HYCOM analysis/forecasting system. Section 3 outlines the data impact procedure, including a description of the limitations of the method as applied in this study. Results of the data impact system cycling with global HYCOM are illustrated in section 4. Summary and conclusions are presented in section 5.

2. Global ocean forecasting

The U.S. Navy Global Ocean Forecasting System (GOFS) consists of HYCOM and the Navy Coupled Ocean Data Assimilation (NCODA) components. The global system is described in a series of technical reports (Metzger et al. 2008, 2010) and more recently in Metzger et al. (2014). The NCODA data assimilation is described in detail in Cummings and Smedstad (2013). As configured within GOFS version 3, HYCOM has a horizontal equatorial resolution of 0.08° or ~1/12° (~7-km midlatitude) resolution. This makes HYCOM eddy resolving. Eddy-resolving models can more accurately simulate western boundary currents and the associated mesoscale variability, and they better maintain more accurate and sharper ocean fronts. In particular, an eddy-resolving ocean model allows upper-ocean topographic coupling via flow instabilities, while an eddy-permitting model does not because fine resolution of the flow instabilities is required to obtain sufficient coupling (Hurlburt et al. 2008b). The coupling occurs when flow instabilities drive abyssal currents that in turn steer the pathways of upper-ocean currents. In ocean prediction this coupling is important for ocean model dynamical interpolation skill in data assimilation/nowcasting and in ocean forecasting, which is feasible on time scales up to about a month (Hurlburt et al. 2008a).

The global HYCOM grid is on a Mercator projection from 78.64°S to 47°N and north of this it employs an Arctic dipole patch, where the poles are shifted over land to avoid a singularity at the North Pole. This gives a midlatitude (polar) horizontal resolution of approximately 7 km (3.5 km). This version employs 32 hybrid vertical coordinate surfaces with potential density referenced to 2000 m, and it includes the effects of thermobaricity (Chassignet et al. 2003). Vertical coordinates can be isopycnals (density tracking), often best in the deep

stratified ocean; levels of equal pressure (nearly fixed depths), best used in the mixed layer and unstratified ocean; and sigma levels (terrain following), often the best choice in shallow water. HYCOM combines all three approaches by choosing the optimal distribution at every time step. The model makes a dynamically smooth transition between coordinate types by using the layered continuity equation. The hybrid coordinate extends the geographic range of applicability of traditional isopycnic coordinate circulation models toward shallow coastal seas and unstratified parts of the World Ocean. It maintains the significant advantages of an isopycnal model in stratified regions while allowing more vertical resolution near the surface and in shallow coastal areas, hence providing a better representation of the upper-ocean physics. HYCOM is configured with options for a variety of mixed-layer submodels. This version uses the *K*-profile parameterization (KPP) of Large et al. (1994). In this study, the ocean model used 3-hourly Navy Operational Global Atmospheric Prediction System (NOGAPS) forcing obtained from the Fleet Numerical Meteorology and Oceanography Center (FNMOC), which includes air temperature at 2 m, surface specific humidity, net surface shortwave and longwave radiation, total (large scale plus convective) precipitation, ground/sea temperature, zonal and meridional wind velocities at 10 m, mean sea level pressure, and dewpoint temperature at 2 m. The first six fields are input directly into the ocean model or used in calculating components of the heat and buoyancy fluxes, while the last four fields are used to compute surface wind stress with temperature- and humidity-based stability dependence. NOGAPS forcing is available on the FNMOC 0.5 degree resolution application grid and extends out to 120 h (i.e., the length of the HYCOM/NCODA forecast). The global HYCOM forecast system includes a built-in energy loan thermodynamic ice model. In this nonrheological system, ice grows or melts as a function of SST and heat fluxes.

NCODA is an oceanographic implementation of the Naval Research Laboratory Atmospheric Variational Data Assimilation System (NAVDAS), a three-dimensional variational data assimilation (3DVAR) technique developed for navy NWP systems (Daley and Barker 2001). The NCODA 3DVAR analysis variables are temperature, salinity, geopotential (dynamic height), and u, v vector velocity components. All ocean variables are analyzed simultaneously in three dimensions. The horizontal correlations are multivariate in geopotential and velocity, thereby permitting adjustments to the mass fields to be correlated with adjustments to the flow fields. The velocity adjustments (or increments) are in geostrophic balance with the geopotential increments, which are in hydrostatic agreement with the temperature and salinity increments. The NCODA 3DVAR problem is formulated in observation space as

$$\mathbf{x}_a = \mathbf{x}_f + \mathbf{HB}^{\mathrm{T}}(\mathbf{HBH}^{\mathrm{T}} + \mathbf{R})^{-1}[\mathbf{y} - \mathbf{Hx}_f], \qquad (1)$$

where $\mathbf{x}_a$ is the analysis vector, $\mathbf{x}_f$ is the forecast background vector, $\mathbf{B}$ is the positive-definite background error covariance matrix, $\mathbf{H}$ is the forward operator, $\mathbf{R}$ is the observation error covariance matrix, and $\mathbf{y}$ is the observation vector. At the present time, the forward operator is spatial interpolation performed in three dimensions by fitting a surface to a $4 \times 4 \times 4$ gridpoint target and evaluating the surface at the observation location. Thus, $\mathbf{HBH}^{\mathrm{T}}$ is approximated directly by the background error covariance between observation locations, and $\mathbf{BH}^{\mathrm{T}}$ directly by the error covariance between observation and grid locations. For the purpose of discussion, the quantity $[\mathbf{y} - \mathbf{Hx}_f]$ is referred to as the innovation vector, $[\mathbf{y} - \mathbf{Hx}_a]$ is the residual vector, and $\mathbf{x}_a - \mathbf{x}_b$ is the increment (or correction) vector.

The 3DVAR observation vector contains all of the synoptic temperature, salinity, and velocity observations that are within the geographic and time domains of the forecast model grid and update cycle interval. The analysis makes full use of all sources of the operational ocean observations. Ocean observing systems currently assimilated by NCODA 3DVAR are listed in Table 1, along with typical global data counts per day. After data thinning and preprocessing, NCODA routinely assimilates about 2.2 million observations per day onto the global HYCOM grid, which contains more than 520 million grid points. All ocean observations are subject to data quality control (QC) procedures prior to assimilation. The need for quality control is fundamental to a data assimilation system. Accepting erroneous data can cause an incorrect analysis, while rejecting extreme, but valid, data can miss important events. The NCODA 3DVAR analysis was codeveloped and is tightly coupled to an ocean data QC system. Cummings (2011) provides an overview of the NCODA ocean data quality control procedures.

The global HYCOM 3DVAR analysis is split into seven overlapping regions covering the global ocean. The boundary between the different regions follows the natural boundary of the continents. The Atlantic, Indian, and Pacific Ocean regions cover the Mercator projection part of the model grid. The remaining four regions cover the irregular part of the model domain, one region in the Antarctic, one each in the northern part of the Atlantic and Pacific, and the Arctic Ocean. Multiple analysis solutions are obtained in the overlap regions, which ensure that the model update will be

TABLE 1. List of observation data types assimilated by NCODA 3DVAR in global HYCOM, with typical daily data counts. Note that profile data counts are for the entire profile. Profiles typically contain hundreds of levels that are assimilated as independent observations. NOAA stands for the National Oceanic and Atmospheric Administration satellite. MetOp is for Meteorological Operation. *Suomi-NPP* VIIRS is for the *Suomi–National Polar-Orbiting Partnership* Visible Infrared Imager Radiometer Suite. Meteosat is for the Meteorological Satellite. *COMS-1* denotes the *Communication, Ocean and Meteorological Satellite 1*. C-MAN is for the Coastal Marine Automated Network. *Cryosat-2* is for the *Cyrosphere Satellite-2*. DMSP is for the Defense Meteorological Satellite Program. SSM/I is for the Special Sensor Microwave Imager.

Data type	Data source	Specifications	Number
Satellite SST	*NOAA-18 NOAA-19*	Infrared 2-km day, night retrievals	4 800 000
	NOAA-18 NOAA-19	Infrared 8-km day, night, relaxed day retrievals	800 000
	MetOp-A	Infrared 2-km day, night retrievals	15 000 000
	MetOp-A	Infrared 8-km day, night, relaxed day retrievals	450 000
	MetOp-B	Infrared 2-km day, night retrievals	15 000 000
	MetOp-B	Infrared 8-km day, night, relaxed day retrievals	450 000
	Suomi-NPP VIIRS	Infrared 1-km day, night retrievals	40 000 000
	GOES-13 GOES-15	Infrared 12-km day, night retrievals	2 000 000
	MeteoSat-2	Infrared 8-km day, night retrievals	220 000
	COMS-1	Infrared 8-km day, night retrievals	220 000
In situ SST	Ships	Engine room intake	6500
		Hull contact sensor	1000
		Bucket temperature	100
		CMAN station	100
	Drifting buoy		34 000
	Fixed buoy		7000
Satellite altimeter	*Jason-2*, Altika, *Cryosat-2*	SSHA	150 000
Sea ice concentration	*DMSP-F13*, *-F14*, *-F15*	SSM/I 25-km retrievals	900 000
	DMSP-F16, *-F17*, *-F18*	SSMIS 25-km retrievals	1 200 000
Profiles	Drifting buoy	Temperature	50
	Fixed buoy		1200
	Argo		600
	XBT		100
	TESAC (CTD)		3500
	Drifting buoy	Salinity	50
	Fixed buoy		800
	Argo		600
	TESAC (CTD)		3000

smooth across boundaries that occur in open ocean areas. There are no limitations in the 3DVAR that prevent the analysis from being executed on the full global HYCOM grid. However, at the present time, limitations in the data setup program do not allow the full global grid in memory, which has resulted in the need to use basin-scale grid domains that overlap slightly. This problem is being addressed by implementation of a distributed memory, domain decomposition Message Passing Interface in the data preparation step of the assimilation system.

The assimilative run of the 3DVAR cycling with global HYCOM on a daily basis (24-h update cycle) reported here was initialized on 29 November 2011. It is a real-time run and must deal with data latency issues associated with some of the ocean observing systems. Satellite altimeter and profile observations have the longest time delays before the data are available for assimilation in real time. The delays in the altimeter data are 72–96 h because of the need to provide precise orbit corrections to improve the accuracy of the measurements. Similarly, receipt of profile data can be delayed up to ~72 h for various reasons. Since ocean data are so sparse, it is important to use all of the data in the assimilation. Accordingly, in real-time applications, NCODA 3DVAR has the capability to select data for the assimilation based on receipt time (the time the observation is received at the center) instead of observation time. In this way all data received since the previous analysis are used in the next real-time run of the 3DVAR. However, data selected using receipt time will necessarily contain nonsynoptic measurement times. Comparing observations to model background fields valid at different times is a source of error in the analysis. This error is reduced by comparing observations against time-dependent background fields using the first guess at appropriate time (FGAT) technique. Hourly forecast fields are used in the FGAT for assimilation of SST observations in order to maintain a diurnal cycle in the model. Daily-averaged forecast fields are used in FGAT for profile data types.

The 3D temperature, salinity, and u, v velocity analysis increments are incrementally inserted into the model during the first 6 h of the forecast using the incremental analysis update procedure (Bloom et al. 1996).

Assimilation of altimeter sea surface height anomalies (SSHA) in global HYCOM uses the synthetic profile technique. Sea surface height from satellite altimeters is an integral measurement, and its assimilation requires an estimate of the covariability of dynamic height versus temperature and salinity at depth. These relationships have been derived from an analysis of historical profile observations and are contained in the Modular Ocean Data Assimilation System (MODAS) database (Fox et al. 2002). MODAS produces synthetic profiles of temperature and salinity consistent with the observed, along-track altimeter SSHA. Salinity is computed in a two-step process, where first a synthetic temperature profile is generated from the altimeter SSHA and then a synthetic salinity profile is derived from the synthetic temperature profile using stored temperature–salinity correlations. The observed altimeter time-mean SSHA must exceed a 3-cm prescribed noise level to be used in the MODAS synthetic profile generation algorithm. In addition, projection of the altimeter SSHA signal onto the model subsurface density field can produce unrealistic results when the vertical stratification is weak. In the absence of specific knowledge about how to partition SSHA anomaly into baroclinic and barotropic structures in these weakly stratified regions, MODAS synthetic profiles are rejected for assimilation when the top-to-bottom temperature difference of the HYCOM forecast is less than 5°C. This necessarily limits the use of altimeter SSHA data at high latitudes.

3. Data impact procedure

Adjoint-based observation sensitivity provides a feasible all-at-once approach to estimating observation impact. Observation impact depends on the forecast error metric, the innovations (model–data differences at the update cycle interval), and the observations number. First, a scalar error norm is defined that is a measure of some aspect of the forecast error. Following Langland and Baker (2004), forecast error is defined as the difference between two forecasts of different lengths separated in time by the update cycle interval and valid at the same time. Here, we use forecasts of 48 and 72 h, which correspond to the 24-h update cycle used in global HYCOM. Forecast errors result from inaccuracies in the initial conditions, the atmospheric forcing, and the nonlinear forecast model. However, differences between forecast errors from forecasts of different lengths verifying at the same time are solely due to the assimilation of observations, which makes it an appropriate cost function for data impact studies. For example, if there were no observations assimilated 48 h ago, then the trajectory of the 48- and 72-h forecasts will be the same and their differences will be zero at the verifying analysis time. However, observations are usually assimilated and the two forecast trajectories will differ as a result.

Quadratic estimates of forecast errors at two different times are given by

$$e_{48} = \langle (x_{48} - x_0)(x_{48} - x_0) \rangle \tag{2}$$

$$e_{72} = \langle (x_{72} - x_0)(x_{72} - x_0) \rangle , \tag{3}$$

where $\mathbf{x}_{48}$ and $\mathbf{x}_{72}$ are forecast states of 48- and 72-h length, and $\mathbf{x}_0$ is the verifying analysis. The outer brackets represent a scalar inner product. These estimates of forecast errors are calculated separately for model temperature, salinity, and velocity fields. The difference between the forecast errors is given by

$$\Delta e_{48}^{72} = e_{48} - e_{72} , \tag{4}$$

with the expectation that e_{48} will be less than e_{72}, since the 48-h forecast will likely be closer to the true ocean state than the 72-h forecast.

Normally, calculation of observation sensitivity and observation impact is a two-step process that involves the adjoint of the forecast model and the adjoint of the data assimilation system. In the first step, the model adjoint maps the forecast error differences given by Eqs. (2) and (3) back to the analysis time, providing analysis or initial condition sensitivity gradients in model grid space. The second step is to extend the initial condition sensitivity gradients from model space into observation space using the adjoint of the assimilation system. However, a limitation of the method as applied in this study is that the adjoint of the HYCOM forecast model is not available. Because of this limitation, initial condition sensitivity gradients are approximated directly from the forecast error differences given by Eq. (4). Since these forecast errors are valid at the 48-h forecast period and not at the analysis time, we have to assume the errors are stationary. Advection and the likelihood of nonlocal influences on the forecast error estimates will invalidate this assumption as the forecast period is extended. This necessarily limits the data impact calculations to relatively short forecast periods. The choice of 48-h forecasts in this study is expected to minimize the effect of nonlocal influences on forecast errors at depths below the mixed layer where oceanographic time scales are relatively long, but it is likely forecast errors near the surface are inflated because of the shorter time scales there. Nevertheless, forecast error differences given by

Eq. (4) are exact and their direct use in the data impact system still provides useful information. The lack of the forward model adjoint also means that we cannot use forecast error cost functions that combine model prognostic variables, such as the energy norm used in NWP data impact studies (Rabier et al. 1996), or the alongshore and cross-shore transport of the California Current system investigated by Moore et al. (2011b). Instead, we are limited to simple error norms calculated directly from model prognostic variables. Consequently, we cannot draw any conclusions about whether the assimilation of temperature observations have a greater impact on reducing HYCOM forecast errors than the assimilation of salinity observations.

Forecast error gradients are projected from model space to observation space using the adjoint of the NCODA 3DVAR assimilation procedure according to

$$\partial J/\partial \mathbf{y} = \mathbf{K}^{\mathrm{T}} \Delta e_{48}^{72}, \tag{5}$$

where $\mathbf{K} = \mathbf{B}\mathbf{H}^{\mathrm{T}}[\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}} + \mathbf{R}]^{-1}$ is the Kalman gain matrix of Eq. (1), with the adjoint of $\mathbf{K}$ given by $\mathbf{K}^{\mathrm{T}} = [\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}} + \mathbf{R}]^{-1}\mathbf{H}\mathbf{B}$. The observation sensitivity vector $\partial J/\partial \mathbf{y}$ is the forecast error gradients in observation space; its elements exist at the observation locations. Term J represents the forecast error norm for temperature or salinity. The only difference between the forward and adjoint of the analysis system is the transposition of the post-multiplication of the solution in observation space to grid space. In the adjoint, a premultiplier $\mathbf{H}\mathbf{B}$ from grid space to observation space is applied first followed by a conjugate gradient solver. Note that $[\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}} + \mathbf{R}]^{-1}$ is symmetric or self-adjoint and operates the same way in the forward and adjoint directions (Baker and Daley 2000). The assimilation procedure is essentially linear, so the adjoint calculation for the 3DVAR is not subject to much error, other than the residual error that results from specifying a convergence criterion in the descent algorithm of the solver.

The observation sensitivity gradient vector ($\partial J/\partial \mathbf{y}$) is then used in a modified form of the observation impact equation (Langland and Baker 2004) given by

$$\delta e_{48} = \langle |(\mathbf{y} - \mathbf{H}\mathbf{x}_f)|, \partial J/\partial \mathbf{y}) \rangle, \tag{6}$$

where $(\mathbf{y} - \mathbf{H}\mathbf{x}_f)$ is the innovation vector. The brackets represent a scalar inner product. In the modification of the impact equation used here, the innovation vector appears as absolute values. The use of absolute values means δe_{48} is directly proportional to the sign and the magnitude of the observation sensitivity gradient. This form of the impact equation is used in adaptive, targeted observing applications of the adjoint-based data impact method, where innovations of future hypothetical observations are unknown and require the use of proxy values. Interpretation of the observation impact equation is straightforward. A negative δe_{48} value indicates a beneficial observation in the assimilation of the observation-reduced HYCOM 48-h forecast error, while a positive δe_{48} value indicates a nonbeneficial observation (forecast error actually increased from assimilation of the observation). Nonbeneficial impacts are not expected, since the assimilation is expected to decrease forecast error by producing improved initial conditions. However, if nonbeneficial impacts occur, and they are persistent, then that may indicate problems with data quality or model performance. Thus, the data impact system can be used as an effective observing system monitoring tool for diagnosing data quality issues or identifying areas where the model has significant predictability limits. We show an example of persistent nonbeneficial data impacts from the assimilation of SST retrievals from a geostationary satellite due to a data quality issue.

The background error covariance matrix used in the NCODA 3DVAR controls how information is spread from the observations to the model grid points and model levels. It also ensures that observations of one model variable produce dynamically consistent corrections in other model variables. The NCODA adjoint effectively performs the reverse of this information spreading by taking as input forecast error gradient fields and computing the matrix of weights given to the error gradient grid points in the evaluation of the observation sensitivity problem at the observation locations. Formulation of the NCODA 3DVAR background error covariances is summarized here, with more detail given in Cummings and Smedstad (2013). Error covariances in the 3DVAR are first separated into a background error variance and an error correlation. The error correlations are modeled using second-order autoregressive (SOAR) functions and are separated into horizontal and vertical components. Horizontal correlation length scales are proportional to the first baroclinic Rossby radius of deformation using estimates computed from the historical profile archive by Chelton et al. (1998). Rossby radius length scales qualitatively characterize scales of ocean variability and vary from ~10 km at the poles to ~200 km along the equator. The length scales increase rapidly near the equator, which allows for stretching of the zonal scales in the assimilation of observations in the equatorial waveguide. Flow dependence is introduced in the analysis by modifying the horizontal correlations with a tensor computed from HYCOM sea surface height (SSH) gradients. The flow-dependent tensor spreads innovations along rather than

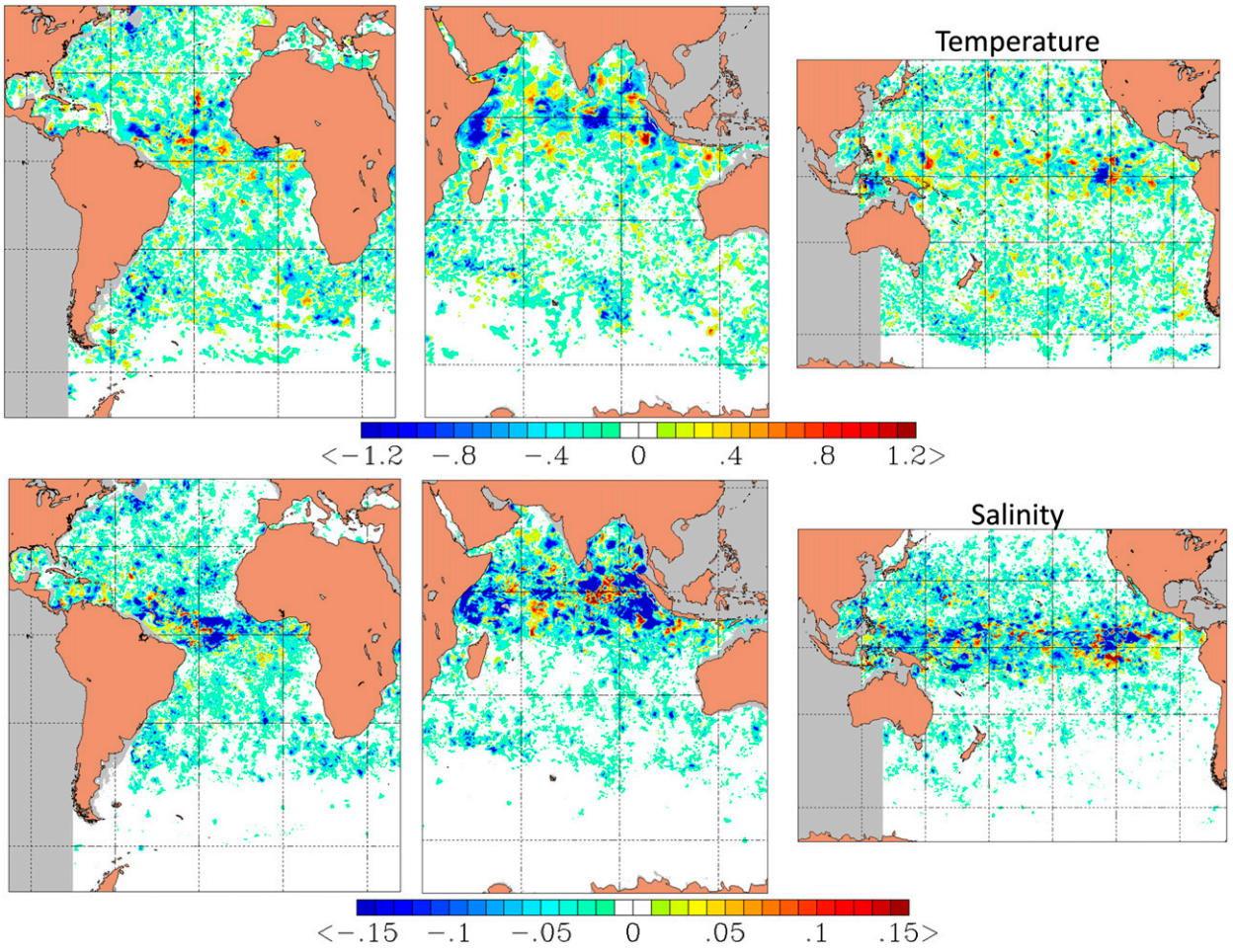

FIG. 1. Instantaneous forecast error gradients [Eq. (4)] for (top) temperature (°C) and (bottom) salinity (PSU) at 100-m depth, valid 1800 UTC 1 Nov 2012. Results are presented for each analysis ocean basin. Positive values (warm colors) indicate forecast error growth; negative values (cool colors) indicate forecast error reduction.

across SSH contours, which are used as a proxy for the circulation field. Flow dependence is necessary in the analysis, since error correlations across an ocean front are expected to be characteristically shorter than error correlations along the front. A similar tensor is used to account for the influence of coastlines in the analysis by rotating and stretching horizontal correlations along the coast while minimizing or removing correlations into the land. Background error correlations close to the coast are expected to be anisotropic because of horizontal advection from alongshore currents. Vertical correlation length scales vary with location and depth and evolve from one analysis cycle to the next. They are defined on the basis of background vertical density gradients calculated using a change in density stability criteria. The method produces vertical correlation length scales that are long when the water column stratification is weak and short when the water column is strongly stratified. Model variability is used as a proxy for the background error variances. Model variability is estimated from an inverse time-weighted history of differences between successive forecasts at the update cycle interval. Since the forecasts are separated in time by an assimilation step, the models are on different trajectories and the variability estimates include the influence of the observations. A time-weighted history of forecast differences is used to improve the estimate due to sampling limitations. The result is background error variances in the 3DVAR that vary with location and depth, and evolve with time, although in practice the error variances tend to evolve to a quasi-steady state. Finally, a multivariate balance operator is used to couple the mass and velocity fields in the analysis. Assimilation of temperature and salinity observations generates geostrophic balanced increments of u, v vector velocity using multivariate correlations with the geopotential, which in turn is computed from temperature

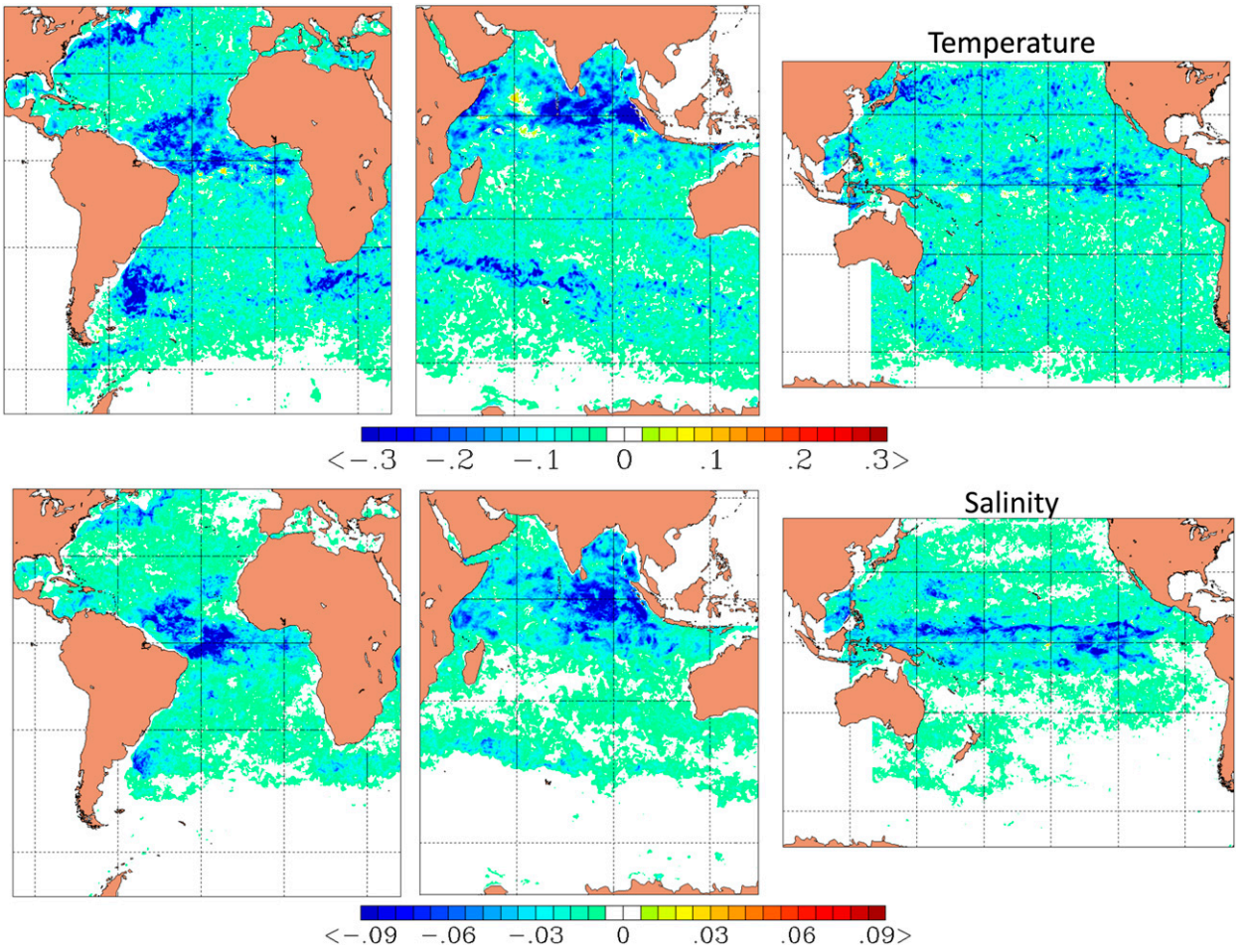

FIG. 2. Averaged forecast error gradients [Eq. (4)] for (top) temperature (°C) and (bottom) salinity (PSU) for the month of November 2012. Results are presented for each analysis ocean basin. Positive values (warm colors) indicate forecast error growth; negative values (cool colors) indicate forecast error reduction.

and salinity. The strength of the geostrophic coupling of the velocity/geopotential correlations is scaled to zero within 1° of latitude from the equator. The multivariate correlations are derived from the first and second derivatives of the SOAR horizontal correlation function as described in Daley (1991).

4. Results

Global HYCOM was cycled sequentially with the NCODA 3DVAR from 16 September to 30 November 2012 using a 24-h update cycle. A summary of the observation data types assimilated in the model with typical daily observation data counts is given in Table 1. Forecast error gradients are computed daily for differences between 48- and 72-h forecasts of temperature, salinity, and u, v vector velocity using Eq. (4). The NCODA adjoint is executed at the end of each analysis update cycle to obtain the observation sensitivities for use in the observation impact equation [Eq. (6)]. Observation impacts are available for each observation assimilated and are partitioned into contributions made by instrument type, geographic domain, and vertical level. The separate NCODA 3DVAR analyses for each of the seven overlapping regions covering the global ocean provide a natural way to partition the data impact results by geographic domain. Results presented here are for the Atlantic, Indian, and Pacific Ocean basins. Specifically, we consider the impact of assimilating temperature observations on reducing HYOCM 48-h forecast temperature error, and the impact of assimilating salinity observations on reducing HYCOM 48-h forecast salinity error. Recall that because of the lack of the HYCOM model adjoint, we are limited to simple forecast error norms calculated from model variable fields.

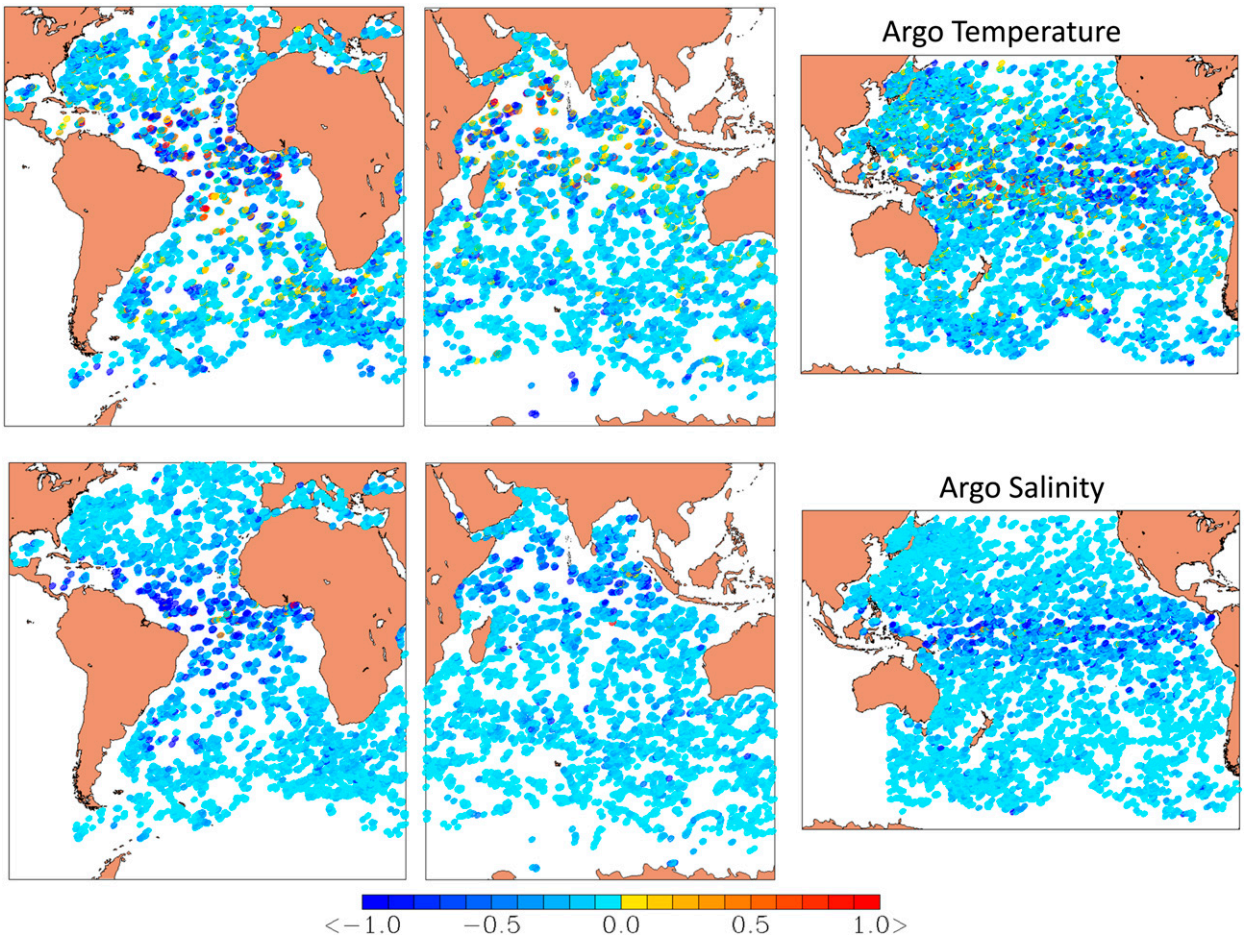

FIG. 3. Impacts of Argo profiles on reducing HYCOM 48-h (top) temperature (°C) and (bottom) salinity (PSU) forecast error. Beneficial impacts are negative values (cool colors); nonbeneficial impacts are positive values (warm colors). Results are pooled over the 16 Sep–30 Nov 2012 time period.

a. Forecast error gradients

The forecast error gradient fields are the basic information input into the NCODA 3DVAR adjoint when calculating observation data impacts. The gradient fields are fully three dimensional and are calculated daily. To illustrate some typical error patterns, Fig. 1 shows temperature and salinity forecast error gradients at 100-m depth in the Atlantic, Indian, and Pacific basins valid 1800 UTC 1 November 2012. Positive and negative areas of forecast errors are seen for each analysis variable in all of the ocean basins, indicating that on any given day HYCOM forecast errors are both increasing (positive values) and decreasing (negative values) in different areas of the model domain. These patterns will vary with depth and evolve over time in accordance with changes in the observing systems assimilated and the variable skill of the HYCOM forecast. Magnitudes of the instantaneous temperature and salinity forecast errors are greatest in the tropics and western boundary current regions.

Time-averaged temperature and salinity forecast error gradient fields at 100-m depth for the month of November are shown in Fig. 2. In general, negative values are found almost everywhere, an indication that, on average, the assimilation is consistently reducing HYCOM 48-h forecast errors. The magnitude of the temperature forecast error reduction is greatest in the tropics and western boundary currents regions, including the Antarctic Circumpolar Current east of the Agulhas region in the Indian Ocean basin. Salinity forecast error reduction is greatest in the tropics for all of the ocean basins, with some additional large salinity forecast error reduction in the Gulf Stream and Brazil/Malvinas Current regions in the Atlantic Ocean basin. The time-averaged fields also show limited areas where temperature forecast error differences are positive, indicating assimilation of temperature observations in those

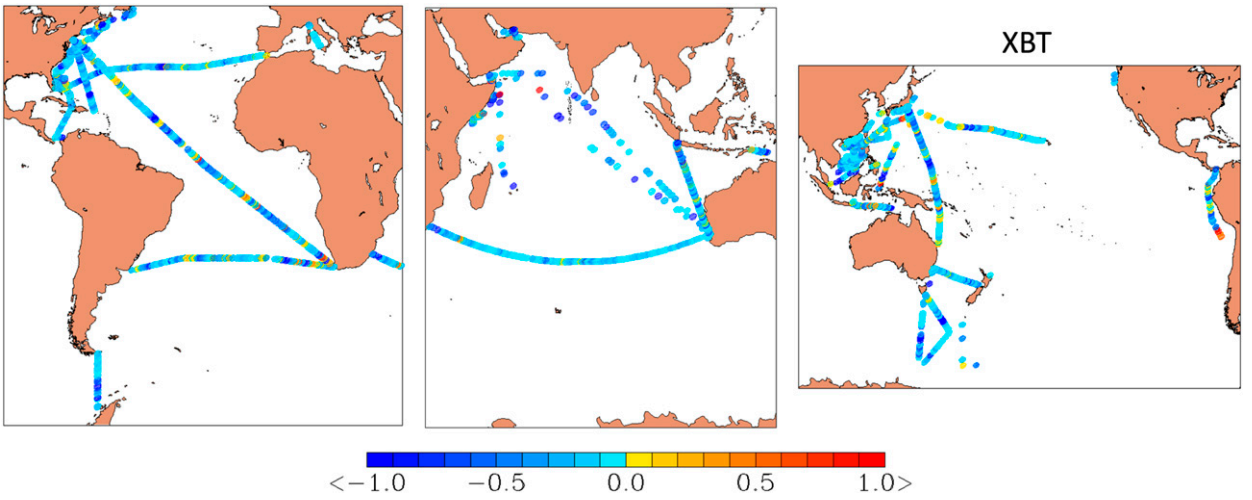

FIG. 4. Impacts of XBT profiles on reducing HYCOM 48-h temperature (°C) forecast error. Beneficial impacts are negative values (cool colors); nonbeneficial impacts are positive values (warm colors). Results are pooled over the 16 Sep–30 Nov 2012 time period.

locations during the month of November actually increased forecast error. These small areas are primarily near the equator in the eastern tropical Atlantic, western tropical Pacific, and Indian Oceans. This increase in forecast error in limited areas could be due to localized, reduced HYCOM predictability arising from instabilities in the system, or due to the assimilation of erroneous data. We suspect the latter in some cases, in particular the increased forecast errors at 8°N, 158°W and 2°S, 164°W in the Pacific, which appear to be associated with the assimilation of specific fixed mooring profiles.

Time-averaged forecast errors that are slightly negative or slightly positive are not color filled in the forecast error maps shown in Fig. 2. These are areas where the data impacts are essentially neutral. A neutral impact does not imply that the model has no forecast error (or skill), nor does it imply that the assimilation system is not efficient in extracting the information content of the observations. It is possible that the forecast is already quite accurate in those areas and that little correction is required. As a result, observations assimilated may improve the quality of the forecast background every day, but when averaged over multiple update cycles, the impacts of observations assimilated on reducing HYCOM forecast error in those areas are quite small. This outcome is likely to be the case for salinity in southeastern Pacific. The area is well sampled by Argo (see Fig. 3), and forecast errors of HYCOM temperature are consistently reduced from the assimilation of Argo temperatures, yet forecast errors of HYCOM salinity show little change. It is concluded that, on average, HYCOM predicts salinity better than temperature in southeastern Pacific.

b. Data impacts

It has been demonstrated that routine assimilation of large numbers of observations work together to consistently reduce global HYCOM 48-h forecast error for both temperature and salinity. An advantage of the observation impact equation is that it allows quantification of the impacts of individual observations assimilated over the global domain. To summarize these results, individual impacts are partitioned by geographic location, data type, latitude band, and vertical level. Because the impact of observations has large variability from day to day, we pool the data impacts over the time period of the experiment: 16 September–30 November 2012. Impact results presented for any group partition is the sum of all individual observation impacts in that group normalized by the number of observations. The normalization is done to facilitate the intercomparison among data types since, for example, temperature observations are dominated numerically by synthetic profiles derived from satellite altimeter SSHA and salinity observations are dominated numerically by Argo. In the figures that follow, beneficial impacts ($\delta e_{48} < 0$) are plotted using cool color shades and nonbeneficial impacts ($\delta e_{48} > 0$) are plotted using warm color shades.

The geographic variation of the impacts of assimilating temperature and salinity observation data types on reducing HYCOM 48-h forecast error are presented for Argo (Fig. 3); expendable bathythermograph (XBT; Fig. 4); fixed buoys (Fig. 5); MODAS synthetics (Fig. 6); SST (Fig. 7); temperature, salinity, and current (TESAC;

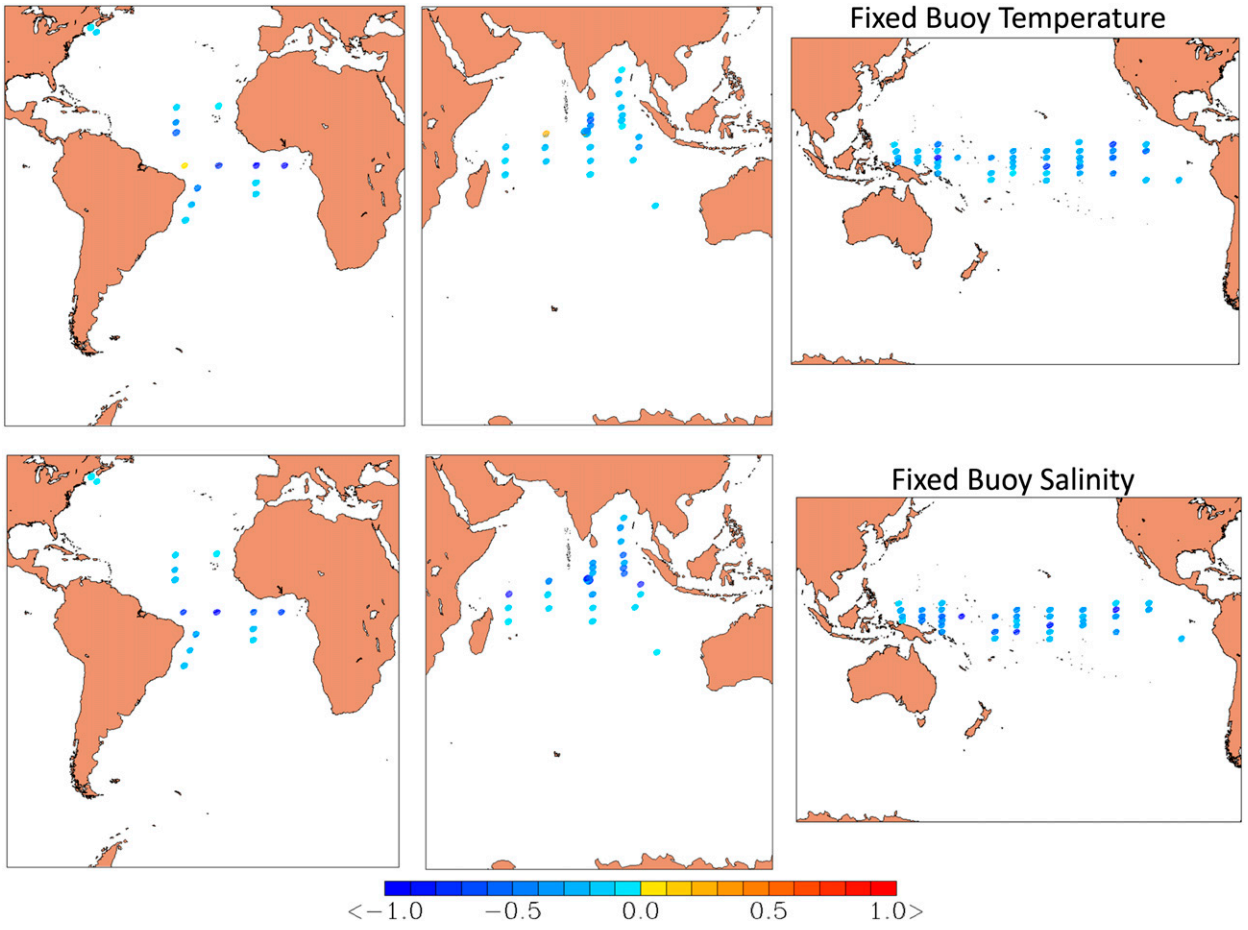

FIG. 5. Impacts of fixed buoy profiles on reducing HYCOM 48-h (top) temperature (°C) and (bottom) salinity (PSU) forecast error. Beneficial impacts are negative values (cool colors); nonbeneficial impacts are positive values (warm colors). Results are pooled over the 16 Sep–30 Nov 2012 time period.

Fig. 8); and animal sensors (Fig. 9). TESAC is the name of a WMO code form for reporting temperature and salinity profiles from ocean gliders and shipboard conductivity–temperature–depth (CTD) measurements. In the case of profile data types (Argo, XBT, TESAC, fixed buoy, animal sensor), measurements from different levels within a profile are treated as independent observations in the assimilation. The results shown here are the summed impacts of the separate depth-level observations in each vertical profile. In addition, each displayed point averages the impacts of multiple profiles if during the two-and-half-month time period of the study more than one profile occurred within a HYCOM grid cell (~7-km midlatitude). This averaging of data impact results on the HYCOM grid is also relevant for 1) fixed buoy arrays assimilated as daily averages of profiles reported almost hourly, 2) MODAS synthetics derived from 10- and 35-day repeat along-track satellite altimeter SSHA measurements, and 3) SST observations from multiple satellites and in situ networks (see Table 1). Results partitioned by observing system show that a majority of temperature observations assimilated have beneficial impacts ($\delta e_{48} < 0$), although nonbeneficial impacts ($\delta e_{48} > 0$) are seen in some Argo, XBT, TESAC, and animal sensor temperature profiles. Assimilation of salinity observations, however, is always beneficial, with relatively large data impacts from assimilation of Argo salinity profiles at low latitudes.

Figure 10 presents the normalized observation impacts for temperature observing systems averaged over the time period of the experiment within the ocean basins. Figure 11 presents similar results for salinity observing systems. The results show that, on average, impacts from assimilation of temperature and salinity measurements from all observing systems are beneficial ($\delta e_{48} < 0$), with the most beneficial data assimilated being temperature and salinity observations from the tropical

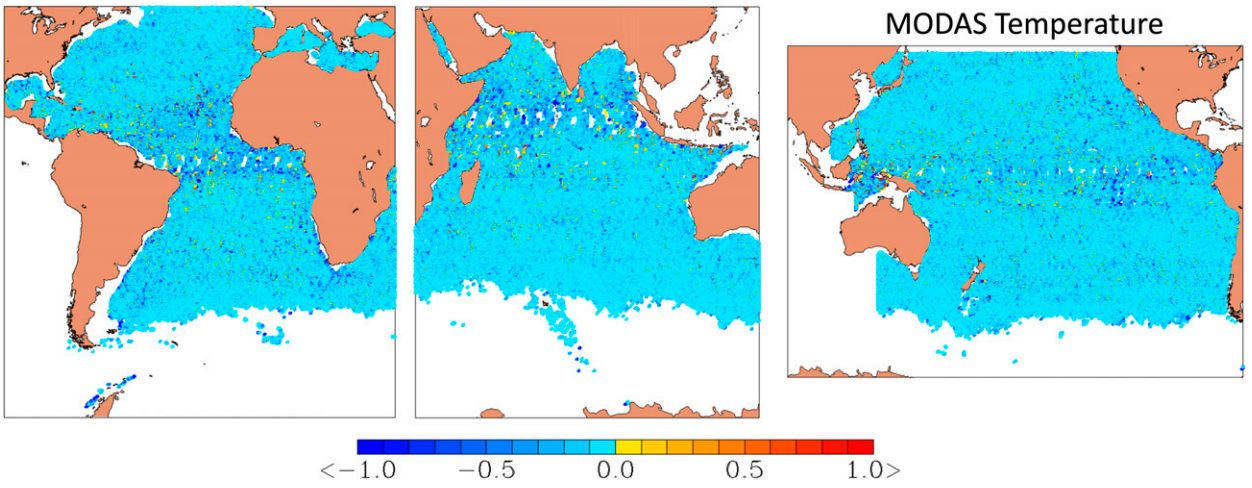

FIG. 6. Impacts of MODAS synthetic profiles on reducing HYCOM 48-h temperature (°C) forecast error. Beneficial impacts are negative values (cool colors); nonbeneficial impacts are positive values (warm colors). Results are pooled over the 16 Sep–30 Nov 2012 time period.

fixed buoy arrays [Tropical Atmosphere Ocean (TAO) and Triangle Trans-Ocean Buoy Network (TRITON) moorings in the Pacific, Prediction and Research Moored Array in the Tropical Atlantic (PIRATA) moorings in the Atlantic, Research Moored Array for African–Asian–Australian Monsoon Analysis and Prediction (RAMA) moorings in the Indian Ocean]. Impacts from the assimilation of Argo, XBT, and animal sensor profiles are also highly beneficial. The large data impacts from the TESAC salinity profiles in the Indian Ocean is not considered to be a robust result, since it is based on few observations from limited sampling locations in the Gulf of Oman and along the Indonesian coast (see Fig. 8).

Synthetic temperature profiles derived from altimeter SSHA measurements using the MODAS methodology have the smallest data impacts, even though synthetics are the most numerous temperature profile data type assimilated. These marginally beneficial impacts are a clear indication that assimilation of altimeter SSHA using MODAS synthetic profiles is suboptimal. MODAS uses climatological relationships to infer temperature at depth from altimeter SSHA. These relationships are well defined in western boundary current areas, where SSHA

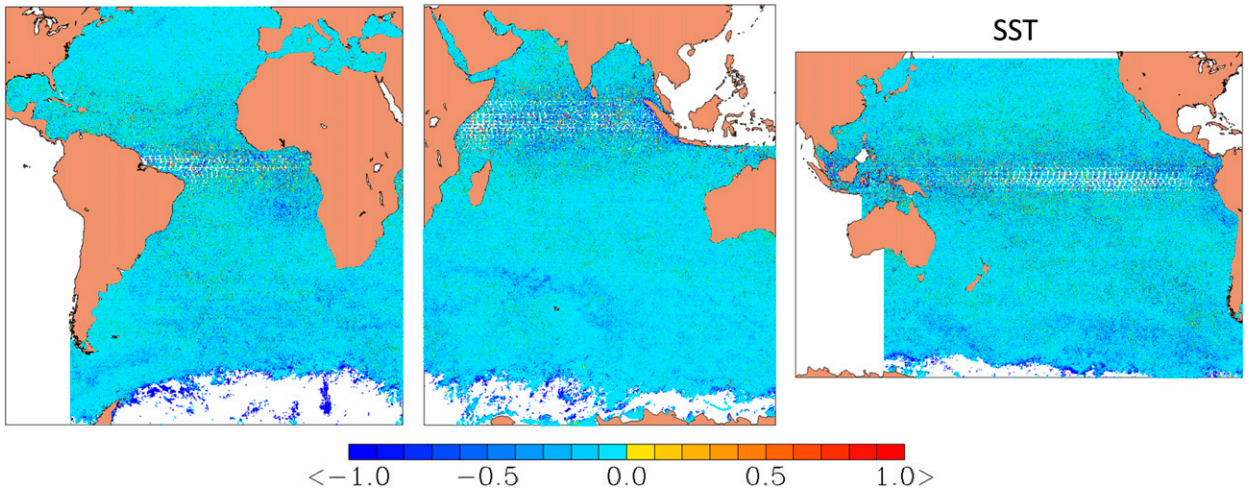

FIG. 7. Impacts of all sources of SST (satellite and in situ) on reducing HYCOM 48-h temperature (°C) forecast error. Beneficial impacts are negative values (cool colors); nonbeneficial impacts are positive values (warm colors). Results are pooled over the 16 Sep–30 Nov 2012 time period.

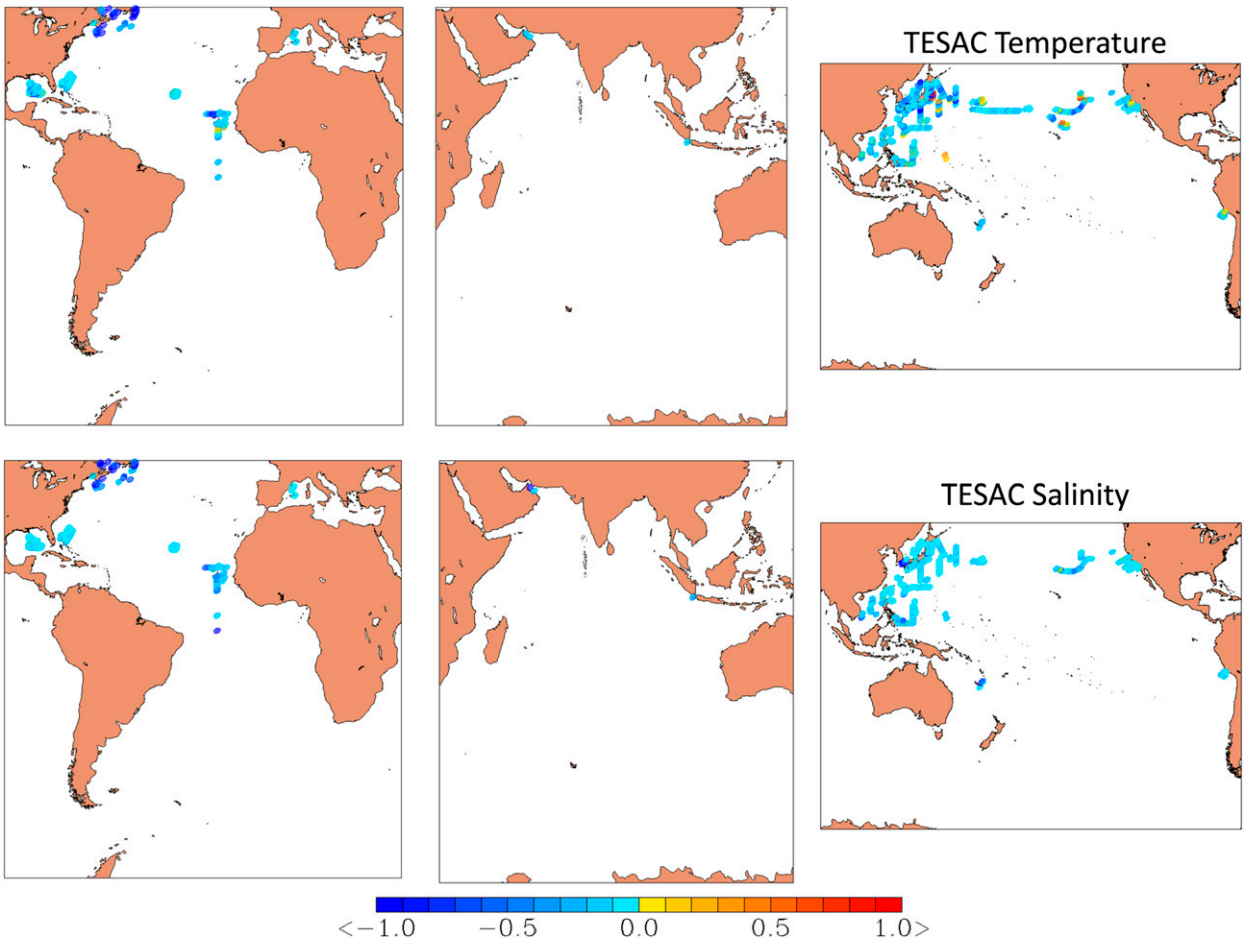

FIG. 8. Impacts of TESAC profiles on reducing HYCOM 48-h (top) temperature (°C) and (bottom) salinity (PSU) forecast error. Beneficial impacts are negative values (cool colors); nonbeneficial impacts are positive values (warm colors). Results are pooled over the 16 Sep–30 Nov 2012 time period.

is a good indicator of thermocline depth, but elsewhere dynamic height cannot be adequately described simply in terms of vertical temperature structure. The altimeter SSHA subsurface inference problem needs to take into account the effects of salinity and nonsteric signals in the altimeter SSHA measurements. Further, conversion of altimeter SSHA to synthetic temperature profiles requires an estimate of a reference mean dynamic topography (MDT). The MDT must match that contained in the time mean altimeter data, which is a nontrivial problem. Synthetic salinity profiles from MODAS, generated in a two-step process described previously, are not directly assimilated in global HYCOM, since it was found that including MODAS salinity profiles in the analysis created a bias in the model and the generation of spurious circulation features when assimilating real salinity data from Argo. MODAS salinity, however, is used in the multivariate balance operator to provide velocity corrections when assimilating MODAS temperatures.

It is interesting to note the relatively large data impacts of the animal sensor profiles given the very low data counts. These are high-quality temperature and salinity profiles obtained from CTD sensors attached to animals, in particular elephant seals. The deep-diving elephant seals obtain environmental data at depths of more than 900 m. The CTD sensors record and transmit ocean structure information to polar-orbiting satellites when the animal surfaces. The foraging behaviors of the animals primarily bring them to ocean frontal zones in search of food in hard-to-reach polar regions. The seals basically serve as targeted observing platforms, providing high-impact data in dynamically sensitive areas.

Profiles from the tropical mooring arrays are found to have the greatest impact on reducing HYCOM 48-h temperature and salinity forecasts. However, in 2011 real-time observations from the TAO array began to decrease. By 2013 distribution of data from TAO

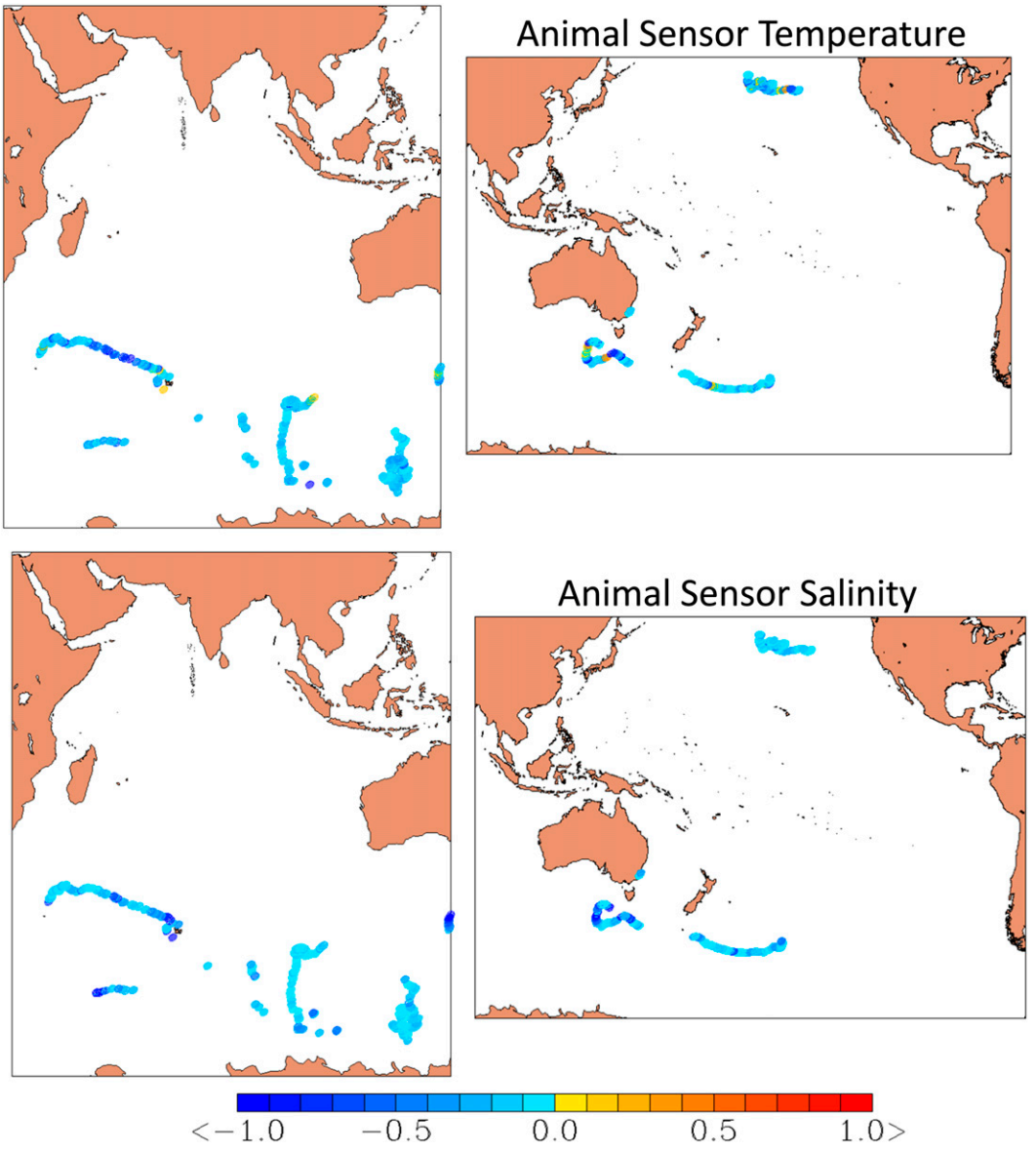

FIG. 9. Impacts of animal sensor profiles on reducing HYCOM 48-h (top) temperature (°C) and (bottom) salinity (PSU) forecast error. Beneficial impacts are negative values (cool colors); nonbeneficial impacts are positive values (warm colors). Results are pooled over the 16 Sep–30 Nov 2012 time period.

became very sporadic in the central and eastern equatorial Pacific. At the same time the density of Argo floats increased in the region. A basic question is whether assimilation of Argo can compensate for the decrease or loss of TAO/TRITON data. Figure 12 compares zonal averages of Argo and fixed buoy data impacts within 5° latitude bands in the ocean basins. Argo observes at nearly all latitudes, but the greatest impact of Argo temperature and salinity observations is in the tropics (±10° latitude). In the Pacific, the magnitude of Argo data impacts is similar to that of the tropical moorings within the same latitude bands, in particular Argo salinity. Argo has an advantage over the tropical moorings by providing deeper and denser vertical sampling, but it is difficult to maintain floats along the equator, although that problem is improving with the implementation of Iridium communication. Argo has a disadvantage in that, solely in terms of ocean observing, profiling floats cannot reproduce the highly temporally resolved data time series that can only be obtained from fixed moorings (e.g., Kessler et al. 1996). It is clear that temperature and salinity observations are essential for

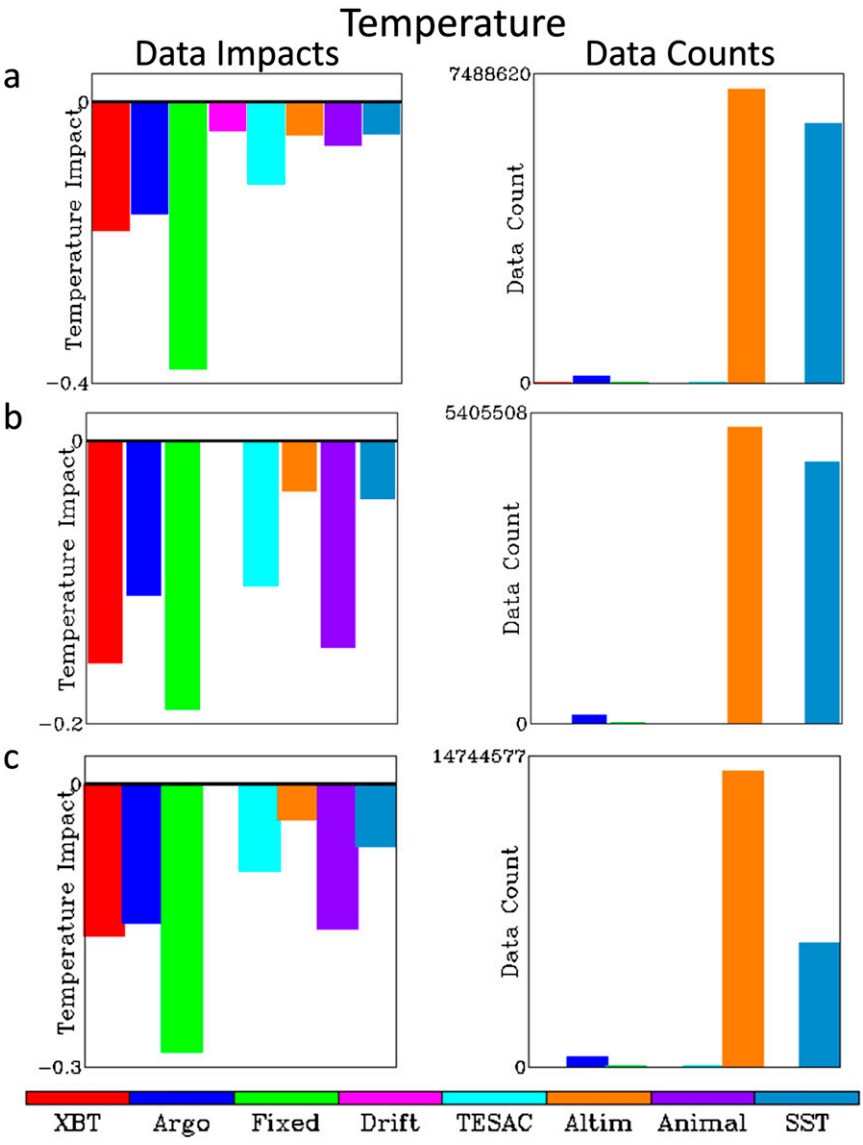

FIG. 10. Mean observation impact for temperature observing systems. Negative values indicate beneficial data impacts. Results are partitioned by ocean basin: (a) Atlantic Ocean, (b) Indian Ocean, and (c) Pacific Ocean. Note that the scale of the data impacts varies with ocean basin. Includes all temperature observations (°C) assimilated over the 16 Sep–30 Nov 2012 time period: Argo—Argo array; fixed—fixed buoys; drift—drifting buoys with thermistor chains; TESAC; Altim—MODAS synthetic profiles; animal—animal sensor profiles; SST—satellite and in situ SST measurements.

constraining forecast errors in HYCOM ocean heat content, stratification, and circulation in the tropics. However, it is unclear if Argo floats can compensate for the loss of TAO/TRITON data. The results presented here show that Argo and TAO/TRITON are highly complementary during the time period of the study, an indication that both observing systems are needed into the future.

The only true global temperature ocean observing systems are Argo, satellite altimeters, and SST, which includes SST data from both satellites and in situ networks (ships and buoys). Other observing systems utilize opportunistic, burst-mode-type sampling regimes restricted to shipping lanes (XBT), fixed locations (moorings), and localized areas (TESAC and animal sensor). Figure 13 shows a comparison of normalized observation impacts for the global temperature observing systems. As previously mentioned, normalization by the number of observations is needed to compare impacts of observing

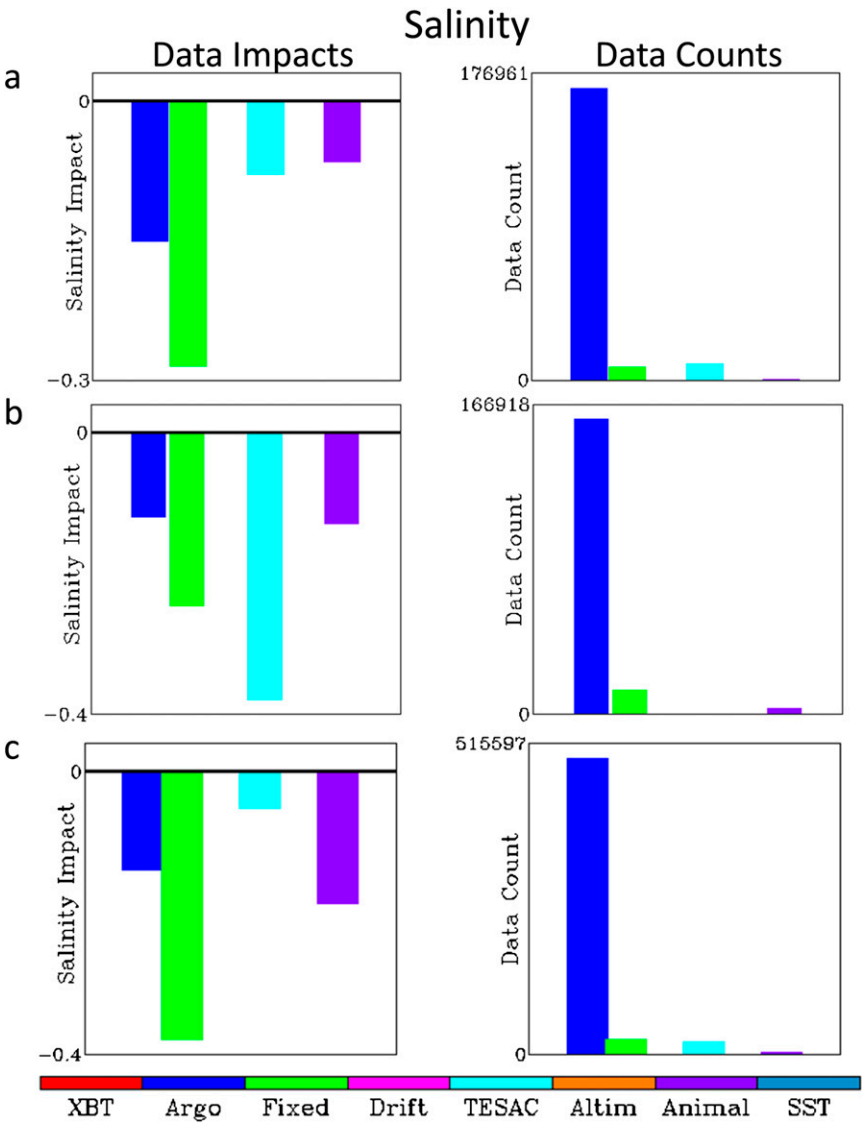

FIG. 11. Mean observation impact for salinity observing systems. Negative values indicate beneficial data impacts. Results are partitioned by ocean basin: (a) Atlantic Ocean, (b) Indian Ocean, and (c) Pacific Ocean. Note that the scale of the data impacts varies with ocean basin. Includes all salinity observations assimilated over the 16 Sep–30 Nov 2012 time period: Argo—Argo array; Fixed—fixed buoys; TESAC; Animal—animal sensor profiles. Salinity units are in PSU.

systems that have large numerical data differences. Normalized data impact comparisons are valid if the data assimilated within an observing system can be considered as independent observations. Prior to assimilation in global HYCOM, all observing systems are processed to remove data redundancies and to minimize correlations among the observations. In particular, satellite altimeter SSHA and SST data are thinned by averaging synoptic observations in spatially varying sized bins based on horizontal correlation length scales. The effect of this adaptive data thinning is seen in Fig. 7, which shows increased decimation of SST observations at low latitudes, where length scales defined using Rossby radius increase rapidly toward the equator. Assuming independent observations after the data thinning processing, it is found that on a per-observation basis, Argo is the most beneficial source of temperature observations assimilated for reducing HYCOM 48-h forecast error. The impact of Argo is more than twice that of altimeter-derived synthetic temperature profiles and SST. However, as shown below, assimilation of SST is complementary to Argo in the upper levels of the water column.

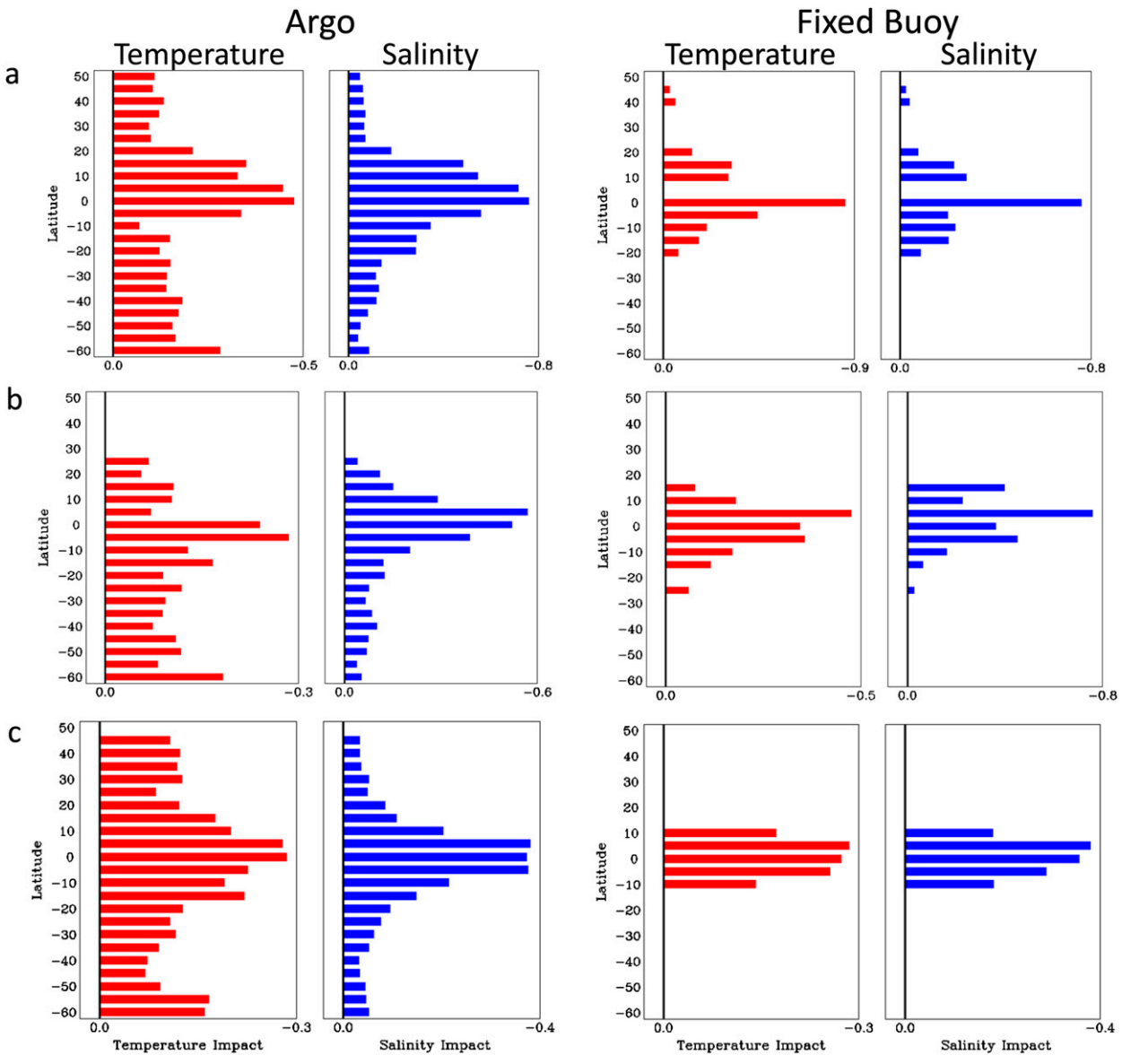

FIG. 12. Mean observation impacts of Argo and fixed buoy temperature (°C) and salinity (PSU) stratified by 5° latitude bands. Results are partitioned by ocean basin: (a) Atlantic Ocean, (b) Indian Ocean, and (c) Pacific Ocean. Note that the scale of the data impacts varies with ocean basin. Negative values indicate beneficial data impacts.

Figure 14 partitions Argo data impacts according to vertical levels of the observations. Data impacts are summed within vertical layers that increase in thickness with depth. Approximately 50% of the impact of assimilating Argo temperature observations is in the upper 300 m of the water column. For Argo salinity, more than half of the data impacts are in the upper 100 m. Argo temperature data impacts show a subsurface maximum in the 100–200-m-depth range. The impact of Argo salinity, however, is greatest at the surface and decreases monotonically with depth. The diminished impact of near-surface Argo temperature profile observations is likely due to the simultaneous assimilation of the large number of satellite-derived SST observations. As described previously, single-level surface observations such as SST are projected downward in the 3DVAR using vertical correlation length scales defined by locally varying vertical density gradients (e.g., mixed layer depth). The vertical covariance extension and the relatively high density of satellite SST measurements combine to effectively constrain HYCOM model temperatures in the upper part of the water column, resulting in a reduced impact of near-surface, less frequent Argo temperature profile observations.

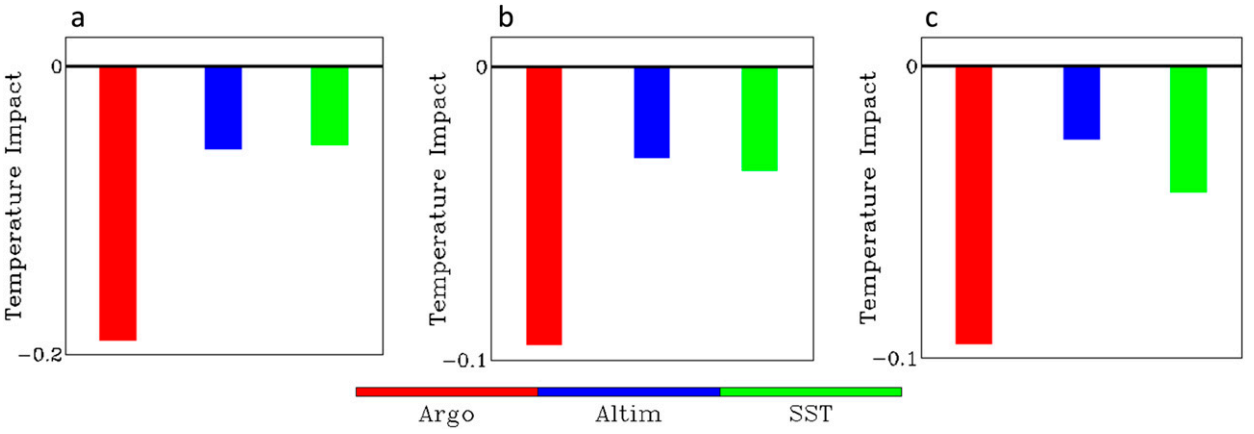

FIG. 13. Mean observation impact for global temperature observing systems (°C). Negative values indicate beneficial data impacts. Results are partitioned by ocean basin: (a) Atlantic Ocean, (b) Indian Ocean, and (c) Pacific Ocean. Note that the scale of the data impacts varies with ocean basin. Includes all Argo, altimeter, and SST observations assimilated over the 16 Sep–30 Nov 2012 time period.

c. *Data quality*

As previously mentioned, the data impact system can be used to identify data quality problems that otherwise are difficult to detect. Persistent nonbeneficial impacts may indicate problems with data quality or model performance. An example of a data quality issue is shown here for SST retrievals from *Geostationary Operational Environmental Satellite-13* (*GOES-13*). SST retrievals from *GOES-13* are generated by the Naval Oceanographic Office (NAVOCEANO) using the empirical split-window formulation. Retrievals are generated only for cloud-free radiances using data with satellite zenith angles up to a maximum of 70°. This value exceeds the zenith angle NAVOCEANO normally uses for retrievals from polar-orbiting satellites, which is restricted to angles of 55° or less. The *GOES-13* SST data are quality controlled and assimilated in the NCODA 3DVAR along with many other sources of satellite SST (Table 1). In general, the impact of *GOES-13* is beneficial in that assimilation of the data during the time period of the study reduce HYCOM surface temperature forecast errors (mean $\delta e_{48} = -0.2$ in the Pacific, mean $\delta e_{48} = -0.1$ in the Atlantic). However, the geographic distribution of nonbeneficial *GOES-13* retrievals ($\delta e_{48} > 0$) shows a distinct pattern that indicates assimilation of GOES retrievals near the edge of the disk are more likely to increase HYCOM forecast error than assimilation of retrievals in the center of the disk (Fig. 15). The longer atmospheric pathlength of the surface-emitted infrared radiances at high zenith angles is likely adding noise to the data from an increase in total column water vapor and the presence of other atmospheric constituents (i.e., aerosols). This atmospheric variation is not adequately modeled or corrected in the empirical NAVOCEANO split-window retrieval algorithm.

5. Summary and conclusions

The NCODA 3DVAR system cycling with global HYCOM assimilates a wide variety of space-based and in situ ocean observations. The 3DVAR combines observations with the HYCOM 24-h forecast of the global ocean state to generate improved initial conditions for the next forecast run of the model. In this study we present results on the impact of the assimilation of various operational observing systems on reducing HYCOM 48-h forecast error. The method uses an estimate of the forecast error in model gridpoint space derived from two model forecast trajectories of different lengths that verify at the same time. The adjoint of the 3DVAR assimilation extends the model space forecast error gradients into observation space. A modified form of the observation impact equation derived by Langland and Baker (2004) is then used to estimate the impact of every observation assimilated. The method is efficient with computational costs roughly equivalent to a single run of the 3DVAR. As such, it can be used for routine observation monitoring in operations. It provides an all-at-once approach to estimating observation impacts. There is no need to selectively remove observing systems to determine impacts as in a data denial experiment. The method automatically adjusts to changes in the observation suite assimilated as new observing systems are introduced and to changes in the forecast model as model resolution increases or as new physics are introduced. The disadvantage of the method as applied here is that due to the lack of HYCOM, adjoint-only

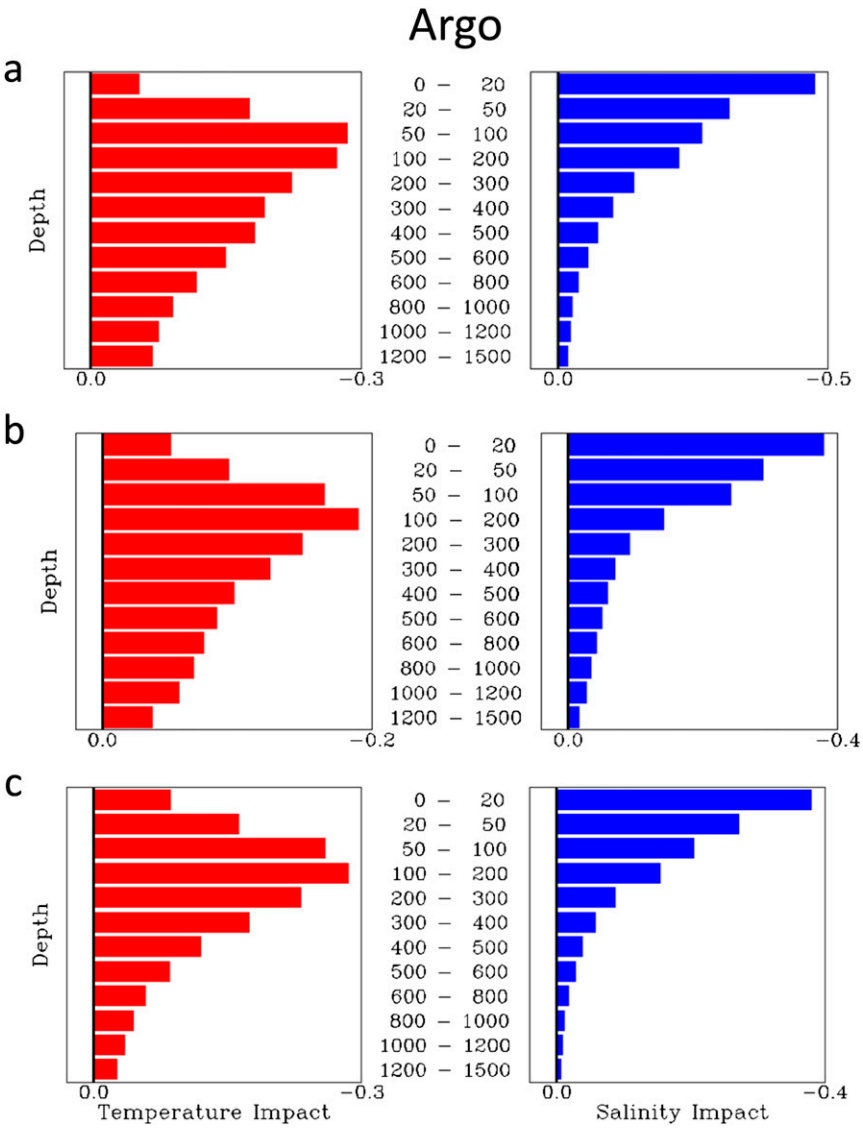

FIG. 14. Mean data impacts of Argo (left) temperature (°C) and (right) salinity partitioned by vertical pressure level. Negative values indicate beneficial data impacts. Note that the thickness of the vertical layers increases with depth. The depth range of the vertical levels for each layer increase with depth and are shown. Results are partitioned by ocean basin: (a) Atlantic Ocean, (b) Indian Ocean, and (c) Pacific Ocean. Note that the scale of the data impacts varies with ocean basin.

simple cost functions and short forecast periods could be used. Thus, only limited aspects of HYCOM forecast error [impact of assimilation of temperature (salinity) observations on temperature (salinity) forecast error] are assessed. Even with these disadvantages and limitations, it is shown that the method provides useful information on data impacts in global HYCOM. It is now possible to efficiently and routinely evaluate the entire global set of oceanographic observations assimilated by HYCOM, determining which data are most valuable and which data are redundant or do not add significant value.

As noted by Langland and Baker (2004) and Lea et al. (2014), the impact results of the adjoint method or an OSE strongly depend on the assimilation system and the forecast model. The issue of whether the data impact results presented here can be generalized to all ocean forecast systems is thus unknown. However, a series of observing system data denial experiments have been proposed by Global Ocean Data Assimilation Experiment (GODAE) OceanView (GOV), the follow-on to the GODAE (Bell et al. 2009), to assess data impacts in operational global ocean forecasting systems.

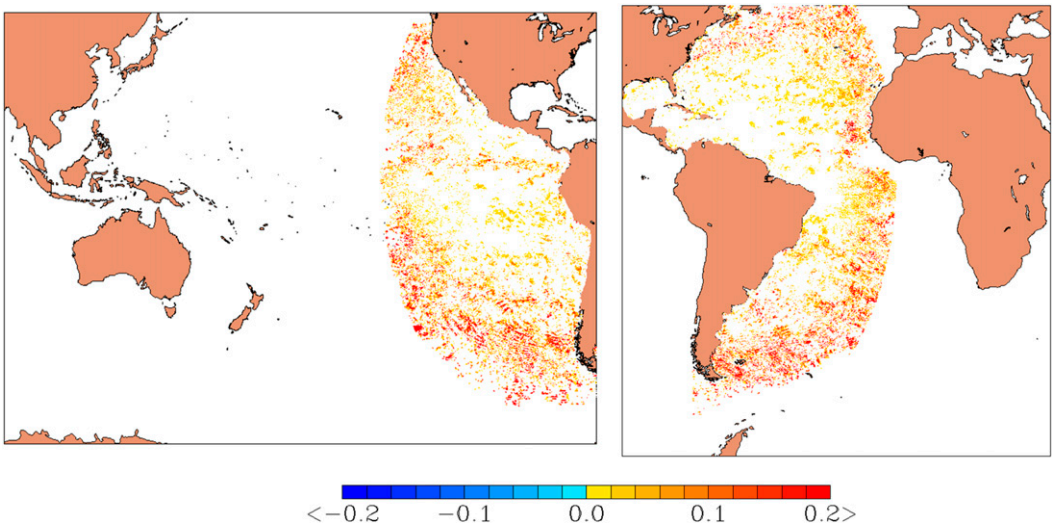

FIG. 15. Geographic distribution of nonbeneficial impacts (positive values, warm colors) from assimilation of *GOES-13* satellite SST retrievals (°C). Beneficial impacts of *GOES-13* data are not displayed. Results are averaged within HYCOM grid locations for the month of November 2012.

Comparison of data impact results from multiple systems will allow for a better assessment of the global ocean observing system in terms of which observations are best and defining locations where forecast errors are sensitive to the initial conditions. The adjoint-based data impact system developed for global HYCOM and described here will contribute to the planned GOV data impact experiments.

It is shown that routine assimilation of large numbers of observations work together to consistently reduce global HYCOM 48-h forecast error for both temperature and salinity. The assimilation of salinity observations is always beneficial, while the assimilation of temperature observations does, on rare occasions, result in nonbeneficial profiles. The largest error reduction in global HYCOM is due to the assimilation of temperature and salinity profiles from the tropical fixed mooring arrays. However, Argo has equivalent data impacts when compared with the tropical moorings in identical latitude ranges. Argo, XBT, and animal sensor data are the next most important observing systems assimilated. The beneficial impact of assimilating Argo temperature and salinity profiles extends to all depths sampled, with salinity impacts maximum at the surface and temperature impacts showing a subsurface maximum in the 100–200-m-depth range. The vertical covariances in the 3DVAR work to extend the influence of the numerous satellite SST observations to the base of the mixed layer, which reduces the impact of near-surface Argo temperature profile levels. Finally, a data quality issue was diagnosed with *GOES-13* retrievals at high satellite zeniths using the data impact system.

The forecast error and data impact results presented here clearly indicate that global HYCOM forecast errors are greatest in the tropics. The assimilation works hard to reduce low-latitude HYCOM forecast errors over time, especially for salinity. The results show that continuous and routine observing is needed at low latitudes to adequately constrain the global HYCOM. The impacts of Argo and the tropical fixed mooring arrays are equivalent in the Pacific. Thus, the two observing systems can be considered complementary when initializing ocean forecast models. Argo samples deeper with improved vertical resolution, but fixed moorings observe more frequently and have the potential for direct measurements of other variables that can be used in ocean model assimilation (ocean currents) and air/sea coupling (meteorological variables such as air temperature and wind speed). Given the recent degradation in the availability of TAO mooring profiles for real-time ocean data assimilation, it is important that the design of a long-term, cost-effective tropical ocean observing system be made and that prioritization be given to its deployment and maintenance.

Acknowledgments. This work was funded in part by the NRL base project ''Observation Impact Using a Variational Adjoint System''. Funding was also received from the National Oceanographic Partnership Program (NOPP) through the project ''U.S. GODAE: Global-Ocean Prediction with the Hybrid Coordinate Ocean Model (HYCOM)'' and the Office of Naval Research (ONR) under Program Element 61153N. The Department of Defense High Performance Computing Modernization

Program provided grants of computer time at Major Shared Resource Centers operated by the Naval Oceanographic Office, Stennis Space Center, Mississippi. We thank the two anonymous reviewers, whose comments and suggestions greatly improved the manuscript.

REFERENCES

Andersson, E., and Y. Sato, 2012: Final report of the Fifth WMO Workshop on the Impact of Various Observing Systems on Numerical Weather Prediction. WMO World Weather Watch Tech. Rep. 2012-1, 25 pp. [Available online at http://www.wmo.int/pages/prog/www/OSY/Meetings/NWP5_Sedona2012/Final_Report.pdf.]

Baker, N., and R. Daley, 2000: Observation and background adjoint sensitivity in the adaptive observation-targeting problem. *Quart. J. Roy. Meteor. Soc.,* **126,** 1431–1454, doi:10.1002/qj.49712656511.

Balmaseda, M. A., and D. L. T. Anderson, 2009: Impact of initialization strategies and observations on seasonal forecast skill. *Geophys. Res. Lett.,* **36,** L01701, doi:10.1029/2008GL035561.

Bell, M. J., M. Lefèbvre, P.-Y. Le Traon, N. Smith, and K. Wilmer-Becker, 2009: GODAE: The Global Ocean Data Assimilation Experiment. *Oceanography,* **22,** 14–21, doi:10.5670/oceanog.2009.62.

Bloom, S. C., L. L. Takacs, A. M. Da Silva, and D. Ledvina, 1996: Data assimilation using incremental analysis updates. *Mon. Wea. Rev.,* **124,** 1256–1271, doi:10.1175/1520-0493(1996)124<1256:DAUIAU>2.0.CO;2.

Cardinali, C., 2009: Monitoring the observation impact on the short-range forecast. *Quart. J. Roy. Meteor. Soc.,* **135,** 239–250, doi:10.1002/qj.366.

Chassignet, E. P., L. T. Smith, G. R. Halliwell, and R. Bleck, 2003: North Atlantic simulations with the Hybrid Coordinate Ocean Model (HYCOM): Impact of the vertical coordinate choice, reference pressure, and thermobaricity. *J. Phys. Oceanogr.,* **33,** 2504–2526, doi:10.1175/1520-0485(2003)033<2504:NASWTH>2.0.CO;2.

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. *J. Phys. Oceanogr.,* **28,** 433–460, doi:10.1175/1520-0485(1998)028<0433:GVOTFB>2.0.CO;2.

Cummings, J. A., 2011: Ocean data quality control. *Operational Oceanography in the 21st Century,* A. Schiller and G. B. Brassington, Eds., Springer, 91–121.

——, and O. M. Smedstad, 2013: Variational data assimilation for the global ocean. *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications,* S. K. Park and L. Xu, Eds., Vol. II, Springer-Verlag, 303–343.

Daley, R., 1991: *Atmospheric Data Analysis.* Cambridge University Press, 457 pp.

——, and E. Barker, 2001: NAVDAS formulation and diagnostics. *Mon. Wea. Rev.,* **129,** 869–883, doi:10.1175/1520-0493(2001)129<0869:NFAD>2.0.CO;2.

Fox, D. N., W. J. Teague, C. N. Barron, M. R. Carnes, and C. M. Lee, 2002: The Modular Ocean Data Assimilation System. *J. Atmos. Oceanic Technol.,* **19,** 240–252, doi:10.1175/1520-0426(2002)019<0240:TMODAS>2.0.CO;2.

Gelaro, R., and Y. Zhu, 2009: Examination of observation impacts derives from observing system experiments (OSEs) and adjoint models. *Tellus,* **61A,** 179–193, doi:10.1111/j.1600-0870.2008.00388.x.

Hurlburt, H. E., and Coauthors, 2008a: Eddy-resolving global ocean prediction. *Ocean Modeling in an Eddying Regime, Geophys. Monogr.,* Vol. 177, Amer. Geophys. Union, 353–381.

——, E. J. Metzger, P. J. Hogan, C. E. Tilburg, and J. F. Shriver, 2008b: Steering of upper ocean currents and fronts by the topographically constrained abyssal circulation. *Dyn. Atmos. Oceans,* **45,** 102–134, doi:10.1016/j.dynatmoce.2008.06.003.

Kalnay, E., Y. Ota, T. Miyoshi, and J. Liu, 2012: A simpler formulation of forecast sensitivity to observations: Application to ensemble Kalman filters. *Tellus,* **64A,** 18462, doi:10.3402/tellusa.v64i0.18462.

Kessler, W. S., M. C. Spillane, M. J. McPhaden, and D. E. Harrison, 1996: Scales of variability in the equatorial Pacific inferred from the TAO buoy array. *J. Climate,* **9,** 2999–3024, doi:10.1175/1520-0442(1996)009<2999:SOVITE>2.0.CO;2.

Langland, R., and N. Baker, 2004: Estimation of observation impact using the NRL atmospheric variational data assimilation adjoint system. *Tellus,* **56A,** 189–201, doi:10.1111/j.1600-0870.2004.00056.x.

Large, W. G., J. C. Mc Williams, and S. C. Doney, 1994: Oceanic vertical mixing: A review and a model with a nonlocal boundary layer parameterization. *Rev. Geophys.,* **32,** 363–403, doi:10.1029/94RG01872.

Lea, D. J., M. J. Martin, and P. R. Oke, 2014: Demonstrating the complementarity of observations in an operational ocean forecasting system. *Quart. J. Roy. Meteor. Soc.,* doi:10.1002/qj.2281, in press.

Lellouche, J.-M., and Coauthors, 2013: Evaluation of global monitoring and forecasting systems at Mercator Océan. *Ocean Sci.,* **9,** 57–81, doi:10.5194/os-9-57-2013.

Metzger, E. J., H. E. Hurlburt, A. J. Wallcraft, J. F. Shriver, L. F. Smedstad, O. M. Smedstad, P. Thoppil, and D. S. Franklin, 2008: Validation test report for the Global Ocean Prediction System V3.0 – 1/12° HYCOM/NCODA: Phase I. NRL Memo. Rep. NRL/MR/7320–08-9148, 82 pp. [Available online at http://www7320.nrlssc.navy.mil/pubs/2008/metzger-2008.pdf.]

——, and Coauthors, 2010: Validation test report for the Global Ocean Forecast System V3.0 – 1/12° HYCOM/NCODA: Phase II. NRL Memo. Rep. NRL/MR/7320–10-9236, 70 pp. [Available online at http://www7320.nrlssc.navy.mil/pubs/2010/metzger1-2010.pdf.]

——, and Coauthors, 2014: U.S. Navy Operational Global Ocean and Arctic Ice Prediction Systems. *Oceanography,* in press.

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

——, and Coauthors, 2011b: The Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilation systems: Part III—Observation impact and observation sensitivity in the California Current system. *Prog. Oceanogr.,* **91,** 74–95, doi:10.1016/j.pocean.2011.05.005.

Oke, P. R., and A. Schiller, 2007: Impact of Argo, SST, and altimeter data on an eddy-resolving ocean reanalysis. *Geophys. Res. Lett.,* **34,** L19601, doi:10.1029/2007GL031549.

Ota, Y., J. C. Derber, E. Kalnay, and T. Miyoshi, 2013: Ensemble-based observation impact estimates using the NCEP GFS. *Tellus,* **65A,** 20038, doi:10.3402/tellusa.v65i0.20038.

Rabier, F., E. Klinker, P. Courtier, and A. Hollingsworth, 1996: Sensitivity of forecast errors to initial conditions. *Quart. J. Roy. Meteor. Soc.,* **122,** 121–150, doi:10.1002/qj.49712252906.