

A Hybrid Global Ocean Data Assimilation System at NCEP

STEPHEN G. PENNY

*National Oceanic and Atmospheric Administration/National Centers for Environmental Prediction,
and Department of Atmospheric and Oceanic Science, University of Maryland,
College Park, College Park, Maryland*

DAVID W. BEHRINGER

National Oceanic and Atmospheric Administration/National Centers for Environmental Prediction, College Park, Maryland

JAMES A. CARTON AND EUGENIA KALNAY

*Department of Atmospheric and Oceanic Science, University of Maryland, College Park,
College Park, Maryland*

(Manuscript received 13 November 2014, in final form 14 June 2015)

ABSTRACT

Seasonal forecasting with a coupled model requires accurate initial conditions for the ocean. A hybrid data assimilation has been implemented within the National Centers for Environmental Prediction (NCEP) Global Ocean Data Assimilation System (GODAS) as a future replacement of the operational three-dimensional variational data assimilation (3DVar) method. This Hybrid-GODAS provides improved representation of model uncertainties by using a combination of dynamic and static background error covariances, and by using an ensemble forced by different realizations of atmospheric surface conditions. An observing system simulation experiment (OSSE) is presented spanning January 1991 to January 1999, with a bias imposed on the surface forcing conditions to emulate an imperfect model. The OSSE compares the 3DVar used by the NCEP Climate Forecast System (CFSv2) with the new hybrid, using simulated in situ ocean observations corresponding to those used for the NCEP Climate Forecast System Reanalysis (CFSR).

The Hybrid-GODAS reduces errors for all prognostic model variables over the majority of the experiment duration, both globally and regionally. Compared to an ensemble Kalman filter (EnKF) used alone, the hybrid further reduces errors in the tropical Pacific. The hybrid eliminates growth in biases of temperature and salinity present in the EnKF and 3DVar, respectively. A preliminary reanalysis using real data shows that reductions in errors and biases are qualitatively similar to the results from the OSSE. The Hybrid-GODAS is currently being implemented as the ocean component in a prototype next-generation CFSv3, and will be used in studies by the Climate Prediction Center to evaluate impacts on ENSO prediction.

1. Introduction

The National Centers for Environmental Prediction (NCEP) has used the same 3D variational data assimilation (3DVar) approach to provide initial conditions and verification of the ocean state since its development in the late 1980s (Derber and Rosati 1989). The computationally inexpensive 3DVar was implemented

operationally at NCEP within the Global Ocean Data Assimilation System (GODAS) in 2003 (Behringer and Xue 2004; Behringer 2007) and as part of the Climate Forecast System (CFS) in 2004 (<http://rda.ucar.edu/datasets/ds094.0>; Saha et al. 2006). Increases in computing power and advances in data assimilation algorithms have made it practical to use higher-cost ensemble-based approaches that offer more representative error estimates and improvements in accuracy. For the atmosphere, a hybrid 3DVar–ensemble Kalman filter (EnKF) became operational in the NCEP Global Forecast System (GFS) in 2012 (Kleist 2012; Wang et al. 2013). That hybrid system is also intended as the

Corresponding author address: Stephen G. Penny, Department of Atmospheric and Oceanic Science, University of Maryland, College Park, 4254 Stadium Dr., College Park, MD 20742-2425.
E-mail: steve.penny@noaa.gov

atmospheric component of the next-generation CFS (CFSv3), thus motivating the development of an ensemble-based counterpart for the ocean. Here we present a set of idealized experiments designed to explore the potential gain in accuracy we might expect relative to the current operational approach.

This ocean data assimilation upgrade must address five issues: 1) introduce time-evolving estimates of correlated errors within the ocean model forecast, 2) introduce quantitative error representation of surface forcing conditions at the upper ocean boundary, 3) provide filter stability even where observations are sparse, 4) allow for a wide variety of new observational data types such as time-varying gravity, and 5) present itself as an extension of the current system. Numerous studies have explored ensemble-based data assimilation for the global ocean suggesting that such a filter can address some of these issues (Evensen 1994; Zhang et al. 2007; Keppenne et al. 2008; Zhang and Rosati 2010; Penny 2011; Penny et al. 2013; Karspeck et al. 2013). However, such filters tend to have problems in regions of sparse data (filter stability), and they would have represented a dramatic change relative to the current operational system at NCEP. Other approaches using 3DVar (Balmaseda et al. 2013) or ensemble optimal interpolation (Oke et al. 2013) have provided advances in ocean data assimilation, but do not explicitly represent the dynamical evolution of the forecast error covariance. Martin et al. (2015) give a further review of various operational ocean data assimilation systems.

The “Hybrid-Gain” approach developed by Penny (2014) addresses the issues outlined above. This approach combines the gain matrices of the local ensemble transform Kalman filter (LETKF; Hunt et al. 2007) and the operational 3DVar (Derber and Rosati 1989). The result is a dynamically varying background error covariance estimate, driven by model uncertainties, that simultaneously leverages the climatological background error covariance estimate to which the current operational system has been tuned. As a further benefit, relative to LETKF alone the hybrid reduces sensitivity to tuning parameters such as localization and inflation. A separate study by Hamrud et al. (2014) recently examined the Hybrid-Gain approach within the context of the European Centre for Medium-Range Weather Forecasts (ECMWF) operational hybrid of four-dimensional variational data assimilation (4DVar) and ensemble data assimilation (EDA). They found that their Hybrid-Gain EnKF-Variational system significantly outperformed either their 4DVar or EnKF used alone, and was competitive with the operational 4DVar-EDA.

Here we present a set of idealized observing system simulation experiments (OSSEs) using the surface

forcing and the ocean observing system available during the 8-yr period 1991–98 to explore the comparative benefits of the new hybrid filter relative to the current operational filter in the context of global ocean data assimilation. The period is chosen because it contains dramatic interannual variability, notably the 1997/98 El Niño and the Indian Ocean dipole event (Saji et al. 1999; Webster et al. 1999). We find that in producing spatio-temporally varying estimates of uncertainty, the new hybrid filter provides a more accurate estimate of the true ocean state than the current operational 3DVar, with an overall reduction in temperature and salinity biases.

2. Methodology

A nature run using an ocean general circulation model was commenced in 1985 using the CFSR ocean state estimate as initial conditions. This nature run is considered to be the true evolution of the ocean and is sampled at the time and place where operationally assimilated temperature and salinity profiles are available from 1991 to 1998, with random errors added to these observations as described below. Three sets of experiments are compared, the only differences being the data assimilation algorithm and surface forcing. These data assimilation experiments are conducted driven by imperfect surface forcing. The first set uses the current operational 3DVar as described in Saha et al. (2014), the second set uses LETKF, and the third set uses the new hybrid. A reference “perfect forcing” data assimilation experiment is conducted using perfect initial conditions and perfect surface forcing to evaluate the range of performance of the operational 3DVar configuration in the OSSE. Finally, to give an indication of applicability in an operational setting, a preliminary reanalysis using real observational data over the same time period is performed for both 3DVar and the hybrid.

a. Model

The Geophysical Fluid Dynamics Laboratory (GFDL) Modular Ocean Model (MOM4p1) uses the configuration implemented for the Climate Forecast System Reanalysis (CFSR), as described by Saha et al. (2010). The model is a finite difference version of the ocean primitive equations using the Boussinesq and hydrostatic approximations. We use a tripolar (Murray 1996) staggered Arakawa B grid in the horizontal and the z^* coordinate geometric height in the vertical. The zonal resolution is $1/2^\circ$. The meridional resolution is $1/4^\circ$ between 10°S and 10°N and gradually increases to $1/2^\circ$ poleward of 30°S and 30°N . There are 40 layers in the vertical, with 10-m resolution from 0 to 240 m, gradually increasing to about 511 m in the bottom layer. This results in 27 layers in the

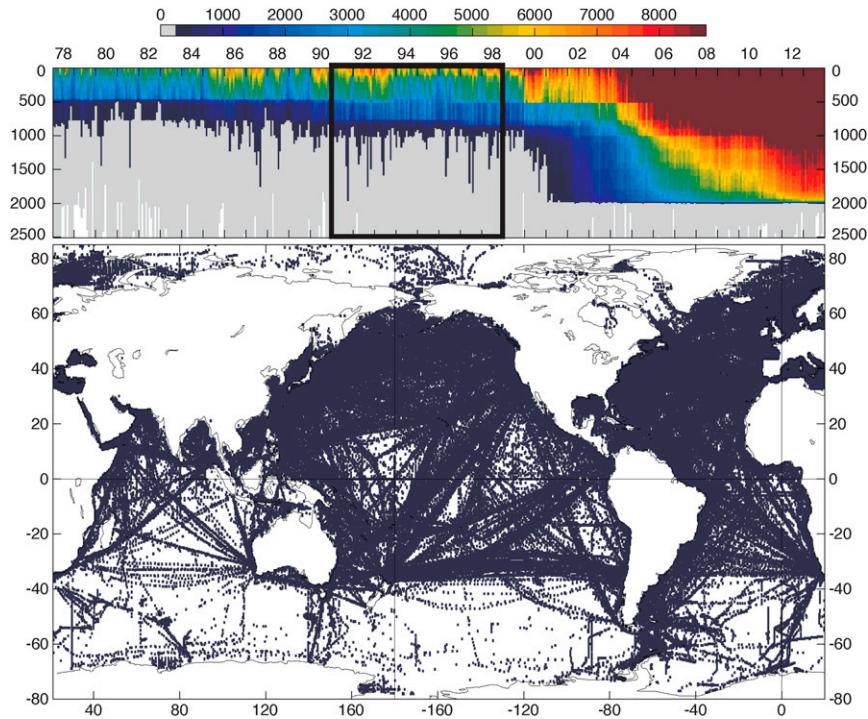


FIG. 1. (top) Global coverage of temperature profiles used for the experiment period by vertical depth (m) from 1978 to 2013 with the experiment period highlighted by the black rectangle and (bottom) the spatial distribution of all observations from 1991 to 1999.

upper 400 m with a bottom depth of 4478 m. The isoneutral method developed by Gent and McWilliams (1990) and the nonlinear scheme of Smagorinsky (Griffies and Hallberg 2000) are used for horizontal mixing of tracers and momentum, respectively. Vertical mixing follows the nonlocal K -profile parameterization of Large et al. (1994). The upper boundary is an explicit free surface, and for the ocean floor we use the Ocean Circulation and Climate Advanced Modeling (OCCAM) 0.2° bathymetry. Variations in model physics, such as those explored by Griffies et al. (2009) and Danabasoglu et al. (2014), are not considered here.

b. Surface forcing and initial ensemble generation

The nature run is forced by surface fluxes from the NCEP–DOE Reanalysis v2 (R2) (Kanamitsu et al. 2002). To degrade surface forcing to reflect uncertainties in the surface meteorology, we generate a 56-member ensemble of surface conditions using ensemble perturbations from the Twentieth Century Reanalysis Project (20CR) daily averaged 3-h forecasts (Compo et al. 2006, 2011; Whitaker et al. 2004). These perturbations are differences from the 20CR ensemble mean for surface momentum stress (N m^{-2}), precipitation rate ($\text{kg m}^{-2} \text{s}^{-1}$), net downward shortwave radiation (W m^{-2}), specific humidity at 2 m (kg kg^{-1}), and air temperature at

2 m ($^\circ\text{C}$). The net downward longwave radiation (W m^{-2}) and mean sea level pressure (Pa) are used unperturbed from the R2. Thus, for the OSSEs, the 20CR ensemble perturbations are used to simulate the uncertainty in the R2 forcing. Members 1–28 (out of 56) are selected to drive the ensemble-method ocean forecasts, thus imposing an intentional bias to the ensemble mean, while two randomly selected members from these 28 drive forecasts are used for two experiments using 3DVar. For the real-data experiment, the full 56-member ensemble of 20CR surface forcing perturbations is used for the hybrid, centered at the R2.

Following the NCEP procedure used in the operational 3DVar, we relax sea surface temperature (SST) for the nature run and all experiments to the daily NOAA Optimal Interpolation (OI) v2 SST product constructed from Pathfinder AVHRR data (Reynolds et al. 2002, 2007) on a 10-day time scale. Sea surface salinity is relaxed to the monthly climatology based on the World Ocean Database 1998 (Conkright et al. 1999) with a 30-day time scale.

c. Observations

Simulated temperature and salinity profiles are sampled from the nature run based on observing networks operational during 1991–98 (Fig. 1). This

historical observational data coverage is sparse and often irregular. In particular, we use the expendable bathythermograph (XBT) profiles acquired from the Global Temperature-Salinity Profile Project (GTSP; <http://www.nodc.noaa.gov/GTSP/>; Sun et al. 2010), conductivity-temperature-depth (CTD) profiles, the Tropical Atmosphere Ocean/Triangle Trans-Ocean Buoy Network (TAO/TRITON) (McPhaden et al. 1998), the Prediction and Research Moored Array in the Tropical Atlantic (PIRATA; Bourlès et al. 2008) and the Research Moored Array for African-Asian-Australian Monsoon Analysis and Prediction (RAMA). We also sample simulated salinity measurements at the locations of the climatologically based synthetic salinity observations used for the CFSR (Saha et al. 2010). Observations poleward of 65°N are excluded because they are not used in the CFSR.

Observational errors are added to these simulated profiles in two parts: a representativeness error ϵ_r is a scalar normal random variable applied to the entire profile with standard deviation given by vector $\boldsymbol{\sigma}^r$; and an instrument error ϵ^n is a vector-valued normal random variable applied to each individual measurement within the profile, with scalar standard deviation σ_n an order of magnitude smaller than $\|\boldsymbol{\sigma}^r\|$:

$$\mathbf{y}_i^o = H(\mathbf{x}^i) + \epsilon_r \boldsymbol{\sigma}_i^r + \sigma_n \epsilon_i^n. \quad (1)$$

Here, the vector \mathbf{y}_i^o is a subset of the total observations representing a single profile i , and $H(\cdot)$ transforms the model state to the observation location (via interpolation). The $\boldsymbol{\sigma}^r$ values vary with depth according to the square root of the vertical temperature gradient and are scaled to have values between 1°C at depth and 1.5°C at the surface for temperature, and between 0.05 and 0.15 psu for salinity.

This setup is meant to emulate the operational environment. It is common practice (although not accurate) to set the off-diagonal elements of the error covariance matrix \mathbf{R} to zero. For all algorithms used here, the observational errors are assumed to be uncorrelated (i.e., we use a diagonal \mathbf{R} matrix). However, these correlations do exist in real-world observations due to representativeness errors relative to the model grid. This is an inaccuracy in the operational system that we have intended to emulate. The observations are binned temporally by day. For the 4D-LETKF component, innovations are generated by comparing daily observations to the model forecasts for the corresponding day. Identical observations, observation errors, and error covariance matrix \mathbf{R} are used for all experiments.

d. Data assimilation

The hybrid method is based on Penny (2014) and utilizes two fully functioning and independent data assimilation systems: the 4D-LETKF ocean assimilation system (Penny 2011; Penny et al. 2013), based on algorithms developed by Ott et al. (2004) and Hunt et al. (2007), and the NCEP operational 3DVar (Behringer et al. 1998; Vossepoel and Behringer 2000; Behringer 2007; Saha et al. 2010). LETKF uses error covariance estimates provided by the ensemble to update the temperature, salinity, zonal velocity, and meridional velocity fields. The NCEP operational 3DVar updates only the temperature and salinity fields. We note that this is not a fundamental limitation of 3DVar, as other operational variational systems have the capability to correct unobserved fields such as zonal and meridional velocity as well (Balmaseda et al. 2013).

We utilize the LETKF algorithm of Hunt et al. (2007) to obtain an analysis by minimizing the objective function:

$$J_P(\mathbf{x}) = (\mathbf{x} - \bar{\mathbf{x}}^b)^T (\mathbf{P}^b)^{-1} (\mathbf{x} - \bar{\mathbf{x}}^b) + [\mathbf{y}^o - H(\mathbf{x})]^T \mathbf{R}^{-1} [\mathbf{y}^o - H(\mathbf{x})]. \quad (2)$$

This objective function expresses our relative confidence between the mean background state $\bar{\mathbf{x}}^b$ (via the background error covariance matrix \mathbf{P}^b) and the observations \mathbf{y}^o (via the observation error covariance matrix \mathbf{R}). In practice, the model space dimension m is typically much larger than the observation space dimension l , which is typically much larger than the ensemble space dimension k . We use subscripts to convey these dimensions in the equations that follow.

Localization is achieved by performing a separate analysis at each grid point, using only observations within a specified geospatial radius. This is called observation-space localization, or R localization (Greybush et al. 2011), to differentiate from localization approaches that eliminate long-distance correlations by operating on the background error covariance matrix. The estimated errors associated with the observations are scaled with a Gaussian function based on their distance from the analyzed grid point. Our choice of horizontal localization radius is roughly a factor of 3 larger than the Rossby radius of deformation (Chelton et al. 1998), decreasing linearly from a σ radius of 720 km at the equator to 200 km at the poles. Vertical inconsistencies caused by a lack of observations in the deep ocean are known to generate spurious currents and vertical velocities in ocean analyses (Zhang and Rosati 2010; Karspeck et al. 2013). To preserve balance within each model ocean column, we do not apply localization in the vertical dimension.

The details for formulating the sequential update from the objective function (2) are given by Hunt et al. (2007). We first compute the analysis error covariance in ensemble space,

$$\tilde{\mathbf{P}}_{k \times k}^a = [(k-1)\mathbf{I}_{k \times k} + (\mathbf{Y}_{l \times k}^b)^T \mathbf{R}_{l \times l}^{-1} \mathbf{Y}_{l \times k}^b]^{-1}, \quad (3)$$

where $\mathbf{Y}^b = \mathbf{H}\mathbf{X}^b$, assuming a linearized observation operator \mathbf{H} , with perturbations from the ensemble mean state forming the columns of \mathbf{X}^b . The symmetric square root of this matrix is computed to obtain the weight matrix \mathbf{W}^a that determines the linear combination of background ensemble members used to form the analysis ensemble:

$$\mathbf{W}_{k \times k}^a = [(\tilde{\mathbf{P}}_{k \times k}^a)^{1/2}]. \quad (4)$$

To transform the background ensemble to the analysis ensemble, we multiply these weights with each of the background ensemble member perturbations:

$$\mathbf{X}_{m \times k}^a = \mathbf{X}_{m \times k}^b \mathbf{W}_{k \times k}^a. \quad (5)$$

These weights are applied identically to each model level, thus maintaining the first-order vertical balance in the analysis. The LETKF analysis perturbations computed in (5) are added to an updated analysis ensemble mean to create the final analysis ensemble:

$$\bar{\mathbf{w}}_{k \times 1}^a = \tilde{\mathbf{P}}_{k \times k}^a (\mathbf{Y}_{l \times k}^b)^T \mathbf{R}_{l \times l}^{-1} (\mathbf{y}^o - \mathbf{H}\bar{\mathbf{x}}^b)_{l \times 1}, \quad (6)$$

$$\bar{\mathbf{x}}_{m \times 1}^a = \mathbf{X}_{m \times k}^b \bar{\mathbf{w}}_{k \times 1}^a + \bar{\mathbf{x}}_{m \times 1}^b, \quad (7)$$

where $\bar{\mathbf{x}}^b$ is the mean background state. This procedure is performed independently at each model grid point. The values at the center of each local analysis are compiled to form the global analysis $\bar{\mathbf{x}}^a$.

While the traditional hybrid covariance approach discards the ensemble mean information $\bar{\mathbf{x}}^a$, the hybrid gain approach utilizes the LETKF analysis ensemble mean as an improved background estimate for 3DVar. The 3DVar component of our hybrid minimizes the following objective function, which is similar to (2):

$$J_B(\mathbf{x}) = (\mathbf{x} - \bar{\mathbf{x}}^a)^T \hat{\mathbf{B}}^{-1} (\mathbf{x} - \bar{\mathbf{x}}^a) + (\mathbf{y}^o - \mathbf{H}\mathbf{x})^T \mathbf{R}^{-1} (\mathbf{y}^o - \mathbf{H}\mathbf{x}). \quad (8)$$

The preferred choice of $\hat{\mathbf{B}}$ is the true analysis error covariance matrix; however, as this is unknown it is instead approximated with the climatological background error covariance matrix used in the operational 3DVar. Following Derber and Rosati (1989), the horizontal component of the error covariance matrix \mathbf{B} is approximated numerically by repeated applications of a Laplacian smoother. Behringer et al. (1998) and Vossepoel and

Behringer (2000) describe the resulting horizontal covariance between any two points as

$$a(z) \exp \left[- \left(\frac{d_1^2}{b_1^2} + \frac{d_2^2}{b_2^2} \right) (\cos \phi)^{-1} \right], \quad (9)$$

where $a(z)$ is proportional to the square root of the local vertical temperature gradient (at depth z) computed from the forecast, d_1 and d_2 are distances between points in the zonal and meridional directions, b_1 and b_2 are parameters defining the horizontal length scale, and $\cos(\phi)$ scales the parameters by latitude ϕ . Within 10° of the equator $b_1 = 2b_2$ (900 and 450 km, respectively). Poleward of 10° the ratio b_1/b_2 is tapered exponentially so that at 30° the ratio $b_1/b_2 = 1.4$, at 50° $b_1/b_2 = 1.1$, and at 60° $b_1/b_2 = 1.0$. The vertical covariance is estimated similarly, with length scales specified as a function of depth. At any level the vertical scale is twice the ocean model layer thickness (Saha et al. 2010).

As in the operational 3DVar, the preconditioned conjugate gradient (PCG) algorithm implemented for the ocean by Derber and Rosati (1989) is applied to minimize the residual $dJ/d(\Delta\mathbf{x}) = \mathbf{r}$ of the cost functional J , with respect to the correction field $\Delta\mathbf{x} = (\mathbf{x} - \bar{\mathbf{x}}^a)$, using preconditioning matrix \mathbf{B} :

$$\mathbf{B}\mathbf{r} = \Delta\mathbf{x} + \mathbf{B}\mathbf{H}^T \mathbf{R}^{-1} (\mathbf{y}^o - \mathbf{H}\mathbf{x}). \quad (10)$$

Upon convergence of PCG we find the correction field $\Delta\mathbf{x}$ that approximates the minimum of the functional J over the global domain within a small tolerance, and consequently the analysis estimate \mathbf{x}^a that corresponds to this minimizer.

As described above, \mathbf{B} is an overestimation of the analysis error covariance matrix. Therefore, a simple parameter α is used to weaken the correction made by the 3DVar minimization. We update the analysis mean as a weighted combination of LETKF and the 3DVar correction and recenter the ensemble to this hybrid analysis:

$$\bar{\mathbf{x}}_{\text{Hybrid}}^a = \alpha \mathbf{x}^a + (1 - \alpha) \bar{\mathbf{x}}^a, \quad (11)$$

$$\mathbf{X}_{\text{Hybrid}}^a = \mathbf{X}^a + \bar{\mathbf{x}}_{\text{Hybrid}}^a \mathbf{v}^T, \quad (12)$$

where $\mathbf{v} = (1, 1, \dots, 1, 1)^T$ is a column of k ones used to add the mean to each column of \mathbf{X}^a .

While this algorithm is algebraically equivalent to forming hybrid gain matrices derived from 3DVar and LETKF (Penny 2014), it has the advantage that a hybrid data assimilation system can be formed from two existing operational-scale data assimilation systems with almost no modification to the individual component

TABLE 1. Parameters for OSSEs in this study. Observations are sampled from the nature run. A reference perfect-forcing case is used to evaluate the impact of observational noise in a perfect system. The remaining experiments use perturbations (pert.) from the Twentieth Century Reanalysis Project (20CR) applied to the NCEP–DOE R2 surface forcing. Identical errors are used for each observation across all experiments.

Experiment	Data assimilation	Surface forcing	Ensemble size	Observations	Figure color
Nature (truth)	None	NCEP–DOE R2	1	None	—
Perfect-forcing	3DVar	NCEP–DOE R2	1	T/S profiles	Gray
3DVar (surface case 01)	3DVar	20CR pert. 01	1	T/S profiles	Red
3DVar (surface case 16)	3DVar	20CR pert. 16	1	T/S profiles	Dark red
LETKF	LETKF	20CR pert. 1–28	28	T/S profiles	Green
Hybrid ($k = 8$)	Hybrid LETKF/3DVar	20CR pert. 1–8	8	T/S profiles	Light blue
Hybrid	Hybrid LETKF/3DVar	20CR pert. 1–28	28	T/S profiles	Blue

systems. The pure 3DVar is recovered by simply bypassing the LETKF procedure, while the pure LETKF is recovered by setting $\alpha = 0$, thus bypassing the 3DVar correction.

We note that the formulation of \mathbf{B} for 3DVar does not contain cross-covariance between temperature and salinity. In the operational system, corrections to salinity are achieved by assimilating a “synthetic” salinity observation derived by applying the climatological relationship between temperature and salinity to the observed temperature at that location. In our OSSE, we instead sample an observed salinity profile from the nature run at every synthetic salinity profile location to ensure this multivariate balance is approximately known by 3DVar.

e. Experiment design and evaluation approach

A single model run forced by the R2 surface forcing was selected to serve as the nature run, or “truth” dataset for all OSSEs. The experiment period 1991–98 was selected to coincide with the introduction of the TOPEX/Poseidon satellite altimetry (September 1992) and the strong ENSO signals present (1997/98). Against this truth dataset, we compare (i) two cases of 3DVar, “(01)” and “(16)”, using model integrations forced by randomly selected surface conditions from surface forcing ensemble members 1 and 16, respectively; (ii) LETKF, using $k = 28$ ensemble members; (iii) the hybrid with $\alpha = 0.5$, using $k = 28$ and $k = 8$ ensemble members; and (iv) a reference “perfect” 3DVar assimilation presented as a baseline for this comparison, using perfect initial conditions and perfect surface forcing (R2). The analysis cycle interval is 5 days. For reference, the details of these experiments are reiterated in Table 1.

We define bias as the difference between an estimator’s expected value and the true value of the parameter being estimated. In this case, the true value of the surface forcing (R2) is the mean of the 56-member ensemble, while the surface forcing bias is the difference

between the mean of members 1–28 and the R2. For example, the time-averaged wind stress curl is shown compared to the root-mean square (RMS) of the wind stress curl bias in Fig. 2. Relative to the climate variability, the largest bias is in the tropics. Ensemble statistics of all surface forcing fields are given as part of the 20CR dataset and can be accessed, for example, via the NOAA/Earth System Research Laboratory (ESRL)/Physical Sciences Division (PSD).

The two cases using 3DVar with surface forcing sampled from the 28 surface forcing conditions used by LETKF may be considered a small “ensemble of 3DVars.” This small sample is used to approximate the bias in the 28-member subset of surface forcing conditions. Such an ensemble could be increased to 28 members to ensure identical biases in the surface forcing. Additional experiments (not shown) were run with 3DVar using a variety of common ocean assimilation practices, such as the use of incremental analysis updates (IAU; Bloom et al. 1996), and extended observation windows that reuse observations across multiple analysis cycles. Results for 3DVar are presented for the configuration using only IAU, which exhibited the lowest root-mean-square errors (RMSEs) among all combinations of these approaches.

For a set of points x_i , for $i = 1$ to n , in a given domain D within the model space, the RMSEs for a given time are calculated between the analysis and truth as a spatial average:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i^a - x_i^t)^2}. \quad (13)$$

We evaluate RMSE for the global ocean and subregions of the major ocean basins to form general statistics that quantify the OSSE results. Because the true state is not known for real-world systems, we also calculate the RMS deviations (RMSDs) between the observations and forecast, computed in the observation space:

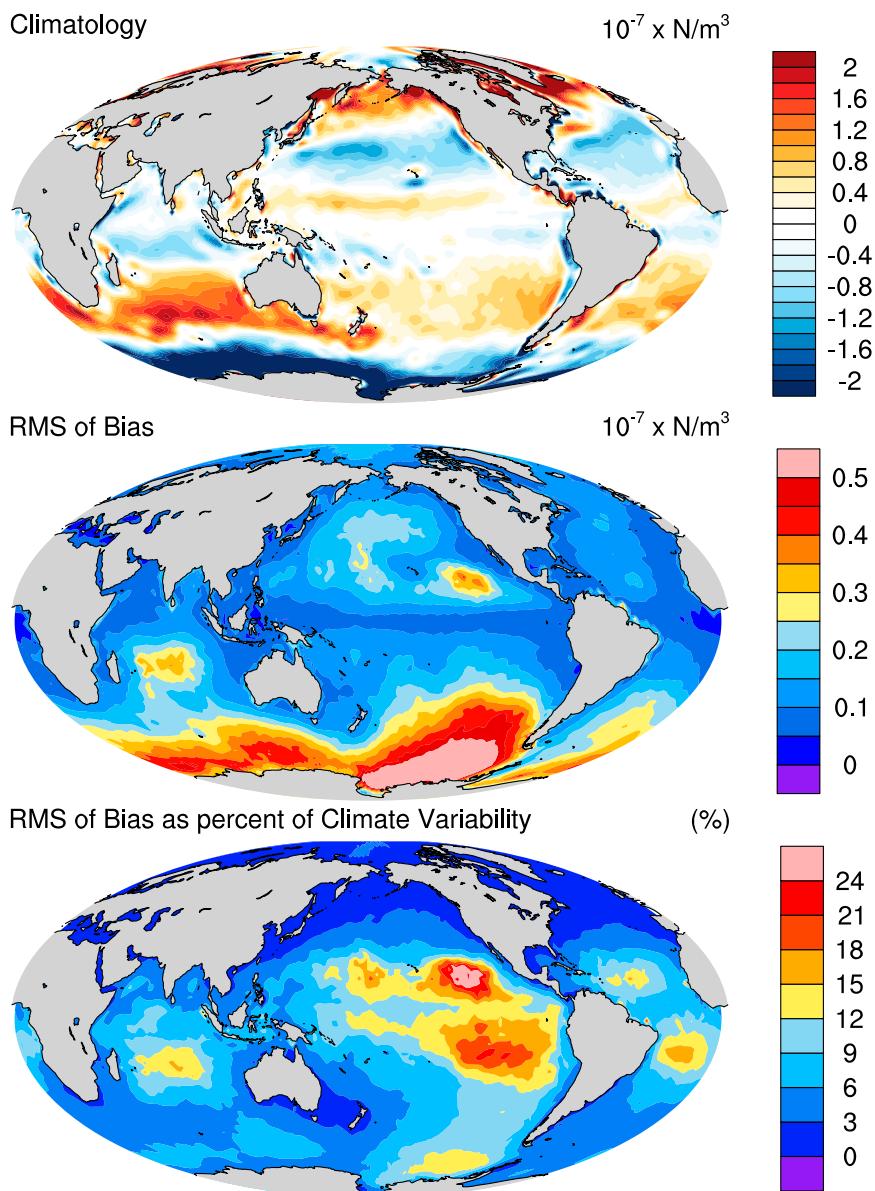


FIG. 2. (top) Climatological average of the wind stress curl for the 56-member mean (R2) from 1991 to 1998. (middle) RMS of daily wind stress curl bias computed as the difference between the 56-member mean and the 28-member subset mean over the same time period. (bottom) RMS of daily wind stress curl bias normalized by the standard deviation of the R2 around the climatological average, also over the same time period.

$$\text{RMSD} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i^o - Hx_i^b)^2}. \quad (14)$$

Both RMSEs and RMSDs are computed pointwise. Bias is computed as the average signed error or deviation. Diagnostic statistics as derived by Desroziers et al. (2005) and applied by Hoffman et al. (2013) and Martin et al. (2015) are used to evaluate the prescribed error covariances within the system. We evaluate the

diagnosed observation error in relation to the prescribed observation error, varying by depth and time. We evaluate the background error in comparison to the RMSEs.

We use the 20°C isotherm as a proxy for the thermocline depth when examining analysis errors in the tropical Pacific. Changes in ocean heat content in the equatorial Pacific have been identified as important indicators of ENSO variability (Wyrtki 1985; Cane et al. 1986; Zebiak 1989; Jin 1997; Meinen and McPhaden

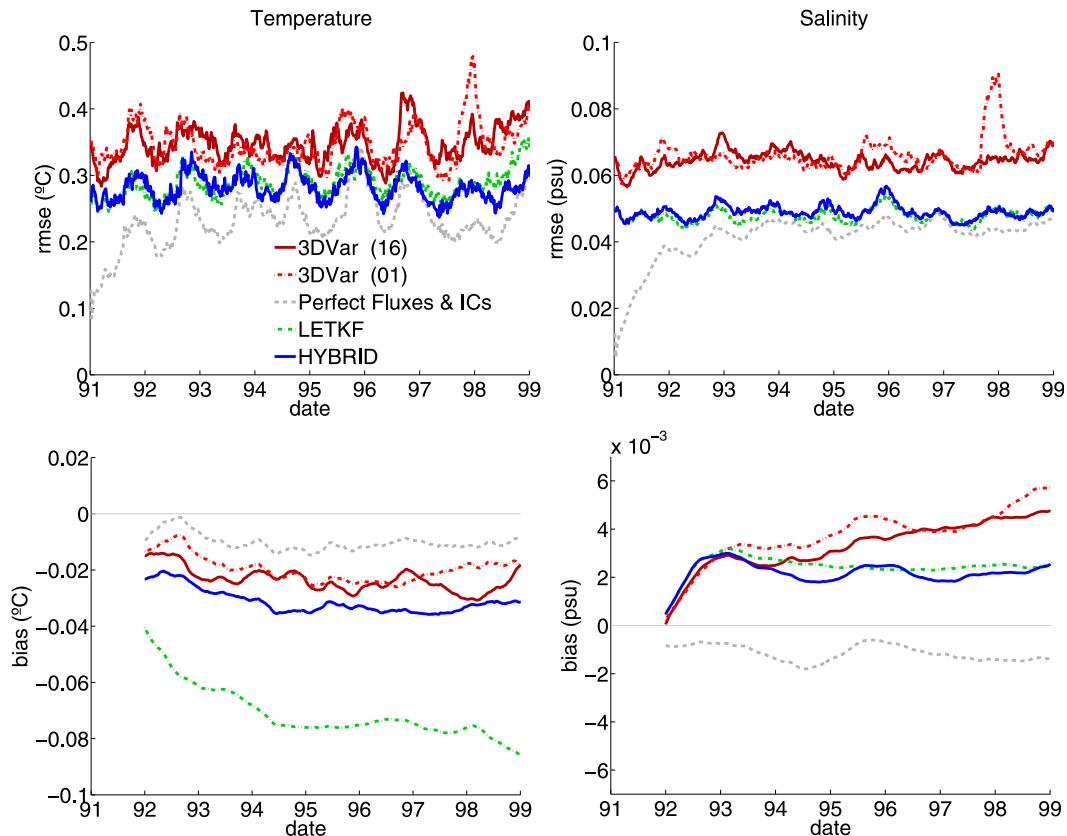


FIG. 3. (top) Global analysis-minus-truth-RMS errors (RMSEs) and (bottom) 12-month moving average of biases in (left) temperature and (right) salinity for the top 700 m. Results are shown for the hybrid (blue), LETKF (green), 3DVar (using surface forcing perturbations: 01 is red, 16 is dark red), and the reference perfect-forcing case (gray). The ensemble methods produce lower RMSEs than 3DVar and approach the RMSE levels of the reference perfect-forcing case. A small cold bias is produced by LETKF and salty bias produced by 3DVar, both gradually increasing throughout the experiment period. The hybrid reduces these biases while maintaining a similar RMSE accuracy with LETKF.

2000; Kessler 2002; Trenberth et al. 2002) and in the prediction of summer monsoon rainfall in India (Rajeevan and McPhaden 2004). McPhaden (2003) noted that unlike for SST there is no spring persistence barrier for integrated upper ocean heat content. It has similarly been found in modeling studies that accurate initialization of upper ocean heat content in the tropical Pacific often reduces the prominence of the spring prediction barrier for SST (Smith et al. 1995; Xue et al. 2000). Chen et al. (1995) demonstrated that reduced noise in the wind stress field prior to initialization also effectively eliminated the spring barrier to prediction in the Zebiak and Cane (1987) model.

Finally, observed-minus-forecast RMSDs are computed for the same time period (1991–98) for a preliminary reanalysis using observational data from the historical record. These experiments use only observed temperature and salinity profiles for LETKF, but include synthetic salinity profiles for the 3DVar component. In

addition, the surface forcing ensemble is expanded to include the full set of 56-member 20CR perturbations, centered at the R2. The model and data assimilation configurations are otherwise the same as for the OSSE described above.

3. Results

We first examine the global upper ocean. We focus on the upper 700 m containing the thermocline at most latitudes, a layer spanned by the XBT profiles. Next, we examine RMSEs in regional subdomains, and examine the North Atlantic in more detail. We then examine the tropics, including temperature and salinity errors along the equator, and errors in thermocline depth.

In the global average, there is an 18% and 24% reduction in RMS temperature and salinity errors, respectively (Fig. 3), for the ensemble-based approaches versus 3DVar. Both the hybrid and LETKF have

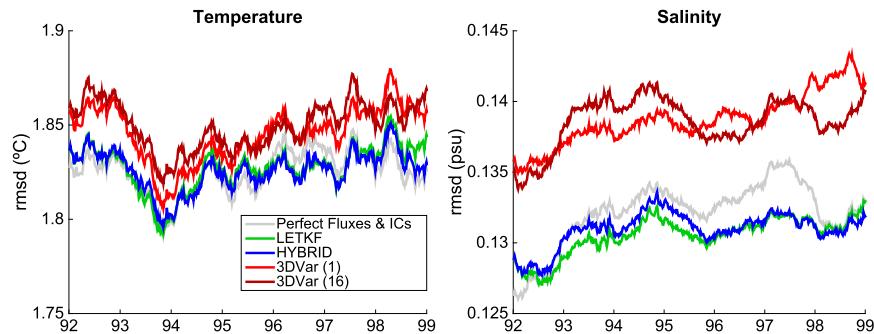


FIG. 4. Global observed-minus-forecast-RMS deviations (RMSD) for 3DVar (casec01 is red and 16 is dark red), the reference perfect-forcing 3DVar (gray), LETKF (green), and the hybrid (blue) for (left) temperature and (right) salinity.

RMSEs close to the levels produced by the reference perfect-forcing system. Similar qualitative results can be ascertained from the observed-minus-forecast RMSD metric (Fig. 4). However, we note that the magnitude of the RMSD is not indicative of the magnitude of the true error, but rather the combined error and observational noise.

The RMSEs of the reference perfect-forcing/perfect-initial-condition system can be interpreted as the amount of observational noise admitted into the system by a perfect-model 3DVar, demonstrating that inaccuracies in the representation of \mathbf{B} exacerbate the impact of observational noise on the analysis. In some regions (e.g., the North Pacific, North Atlantic, and tropical Atlantic) the hybrid and LETKF produced mean errors lower than this baseline level.

LETKF and the hybrid have smaller time-mean RMSEs than 3DVar in all regions, with RMS errors for temperature reduced by: 14% in the tropical Pacific, 29% in the tropical Atlantic, 25% in the tropical Indian Ocean, 22% in the North Pacific, 37% in the North Atlantic, and 13%–17% in the southern oceans. RMSEs for salinity are reduced by 20% in the tropical Pacific, 39% in the tropical Atlantic, 30% in the tropical Indian Ocean, 26% in the North Pacific, 36% in the North Atlantic, and 16%–24% in the southern oceans. There are also large reductions in RMS errors of the ocean currents, particularly in the Southern Hemisphere. RMS errors for zonal velocity are reduced by 16%–31% in the tropical oceans, 29%–39% in the northern oceans, 45% in the South Pacific, 43% in the South Atlantic, and 66% in the Southern Ocean. RMS errors for meridional velocity are reduced by 21%–36% in the tropical oceans, 33%–45% in the northern oceans, 43% in the South Pacific, 45% in the South Atlantic, and 66% in the Southern Ocean.

We show the median regional RMS errors for temperature, the first and third quartiles, confidence limits,

and outliers over the course of the 1991–98 experiments in Fig. 5. The height of each box indicates the interquartile range, the horizontal line inside indicates the median, and whiskers extend to approximately ± 2.7 standard deviations. When comparing two methods, it is typical to consider the difference in medians statistically significant if the notches inside their respective boxes do not overlap. Median RMS errors for the ensemble-based methods are below that of both 3DVar cases and the reference perfect-forcing case in the North Pacific, North Atlantic, and tropical Atlantic. RMSEs for the larger ensemble ($k = 28$) LETKF and hybrid are significantly below those for 3DVar in all regions. The smaller ensemble hybrid ($k = 8$) performs similarly to the larger ensemble hybrid ($k = 28$) case in lower-resolution regions, while the more highly resolved tropics benefit from increased ensemble size.

The ensemble methods produce significantly lower RMSEs than 3DVar for all prognostic model variables in the North Atlantic (Fig. 6). The hybrid is again shown to reduce biases versus LETKF alone. The RMS errors for LETKF are significantly below the reference perfect-forcing 3DVar for all variables, indicating that the background error used by 3DVar is overestimated in this region. Upon examining the temporal evolution between 100- and 200-m depths, we find that errors introduced via noise in assimilated observations are amplified along strong temperature or salinity gradients. Because of the simple structure of its error covariance matrix, 3DVar produces approximately isometric analysis increments that have low fidelity relative to the spatial structure of these gradients. As a result, fast-growing errors may advect for weeks to months around the gyre until new observations are available within a range close enough to correct these transported localized errors. For the ensemble methods however, the ensemble spread naturally increases along strong fronts, thus allowing LETKF to form a dynamically aware error

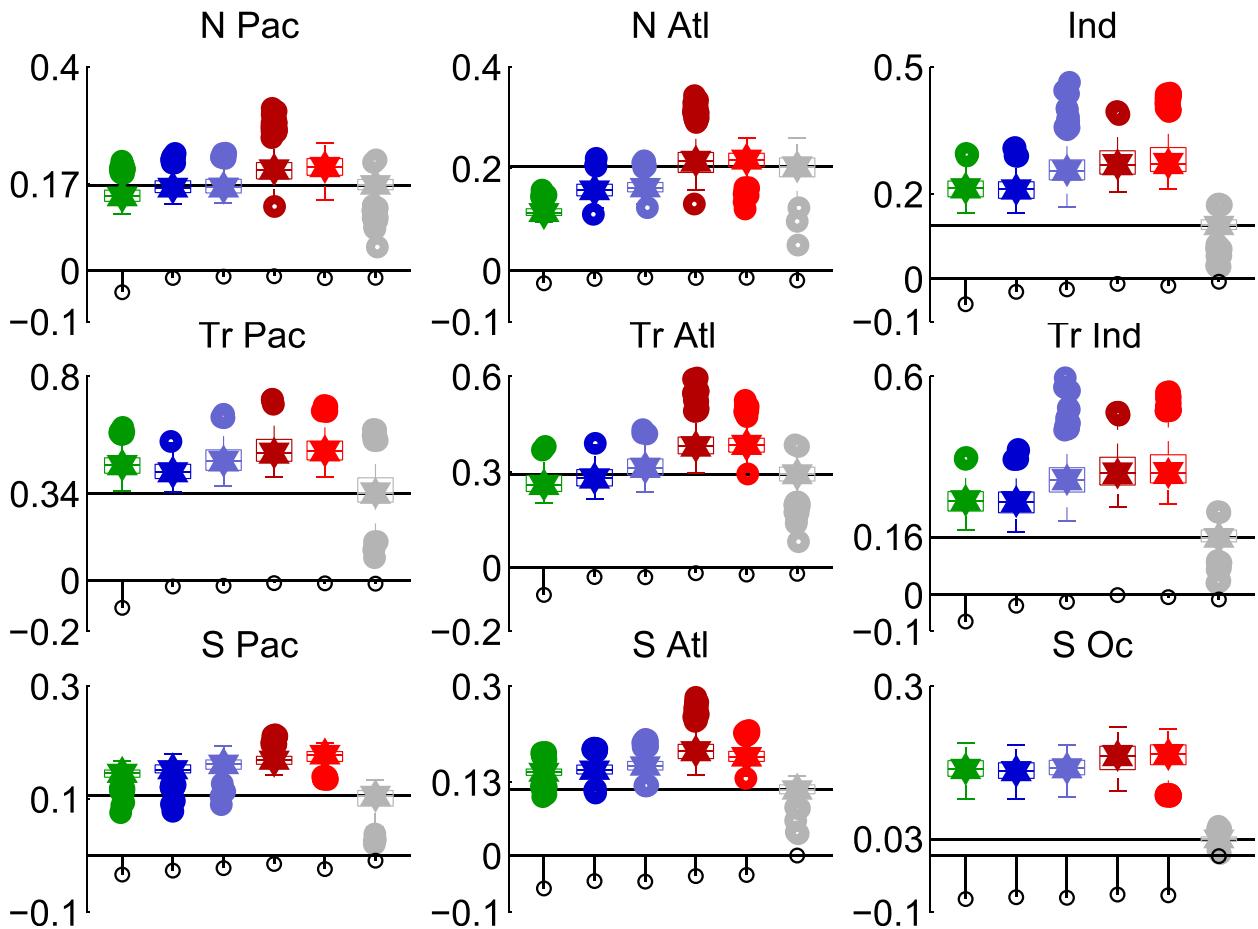


FIG. 5. Box-and-whisker plots of analysis minus truth RMSE (color) and stem plots of bias (black open circles) for temperature ($^{\circ}\text{C}$). Results are shown for LETKF (green) with ensemble size $k = 28$, the hybrid with ensemble size $k = 28$ (blue) and $k = 8$ (light blue), the operational 3DVar with two randomly chosen surface forcings 16 (dark red) and 01 (red), and the reference perfect-forcing case (gray). Outlier values are indicated with open circles in the respective color. The median of the reference perfect-forcing case is further indicated with a horizontal black line to aid visual comparison. (from top left to bottom right) Results are subdivided into regions as defined in Table 1. With statistical significance, RMS errors for the ($k = 28$) ensemble methods are typically below those for 3DVar. Relative to LETKF, the hybrid reduces imposed biases introduced in the surface forcing ensemble.

covariance matrix that reduces the magnitude of fast-growing errors before they are advected.

Figure 7 shows time mean statistics along the equator from the surface to 300-m depth. The distribution of ensemble spread indicates that the largest uncertainty in temperature is at the location of the thermocline depth. The ensemble of surface forcing conditions largely drives this uncertainty. The largest estimated errors in the analysis are also concentrated at the thermocline depth. This is in part due to the fact that observational errors are larger where the observed vertical gradients are larger (due to errors in representativeness), and in part due to the background error covariance being larger at this depth.

Errors in heat content in the tropics are largely due to misplacement of the thermocline depth. We explore

errors in the thermocline by comparing the 20°C isotherm depth of the nature run (truth) to 3DVar and the hybrid in the equatorial Pacific during two years that include the 1997/98 ENSO (Fig. 8). The impacts of surface forcing errors on the 20°C isotherm are illustrated by comparing the reference perfect-forcing 3DVar with the perturbed-forcing 3DVar. In all experiments we observe errors propagating eastward at about 1.9 m s^{-1} , indicative of uncertainties in the timing, intensity and propagation speed of equatorial Kelvin waves (Picaut et al. 2002). We also observe westward propagating errors that are particularly noticeable throughout the second half of 1998, which appear to correspond to mixed Rossby-gravity (Yanai) waves (Shinoda 2010) traveling at about $0.3\text{--}0.4\text{ m s}^{-1}$. As these are present in the reference perfect-forcing case, their genesis can be attributed to observational noise.

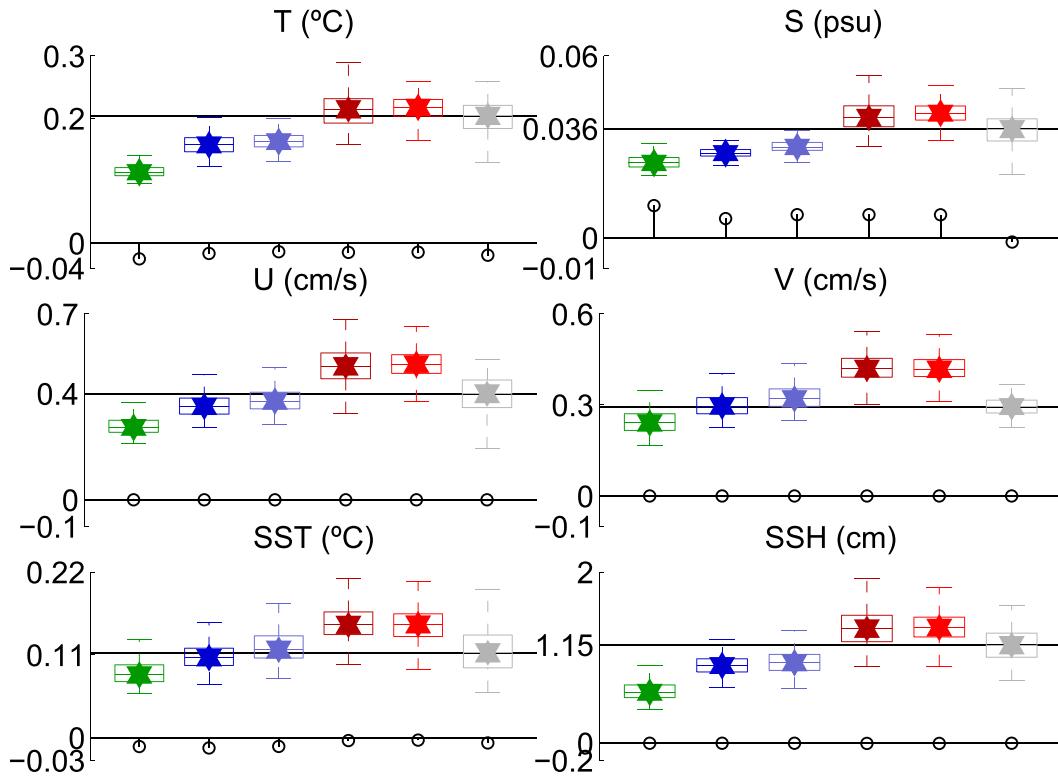


FIG. 6. As in Fig. 5, but only for the North Atlantic and for (from top left to bottom right) T , S , U , V , SST, and SSH. Outliers are not shown. The hybrid reduces imposed biases introduced in the surface forcing ensemble that particularly affect the temperature and salinity fields for LETKF.

Additional uncertainties in surface forcing conditions intensify these errors because the resulting background state is typically farther from observations and consequently requires a larger adjustment.

During the fourth quarter (Q4) of 1997, 3DVar produces a wide region in the eastern Pacific that is at first too deep and then too shallow by over 20m (corresponding to the strong spike in RMSE during this

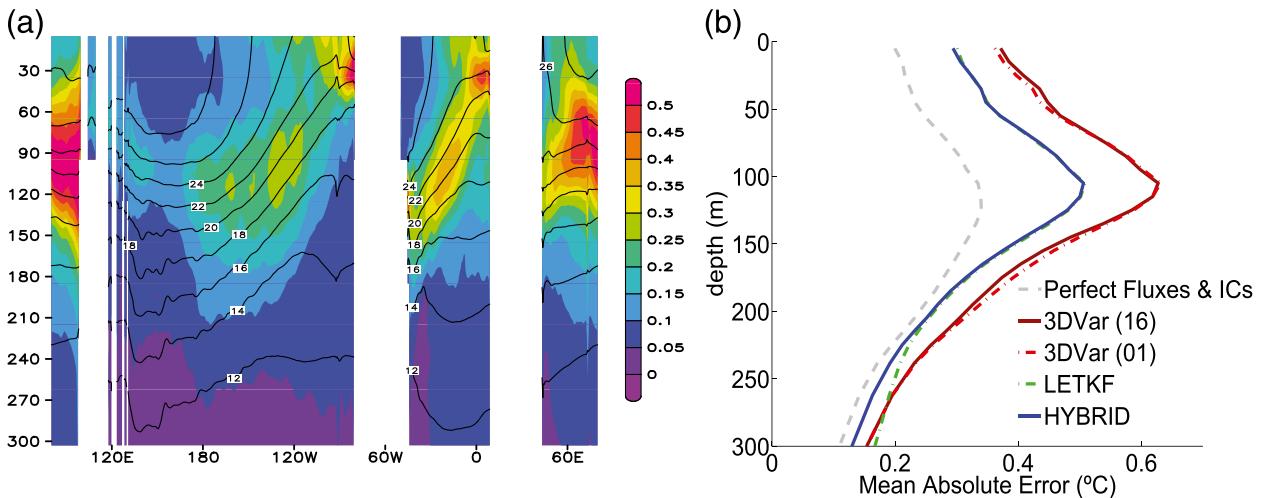


FIG. 7. (a) Time-averaged (1991–98) temperature (contours) and ensemble spread (color shaded) for the top 300 m at the equator in $^{\circ}\text{C}$ for the three longitude regions indicated. Ensemble spread is associated with the thermocline depth. It increases in the top 50 m in the eastern Pacific and Atlantic Oceans, and has a maximum in the Indian Ocean. (b) The zonally averaged time-mean absolute analysis error at the equator for the five different methods. The errors reach their maximum at the thermocline depth. The ensemble methods reduce error relative to the operational 3DVar.

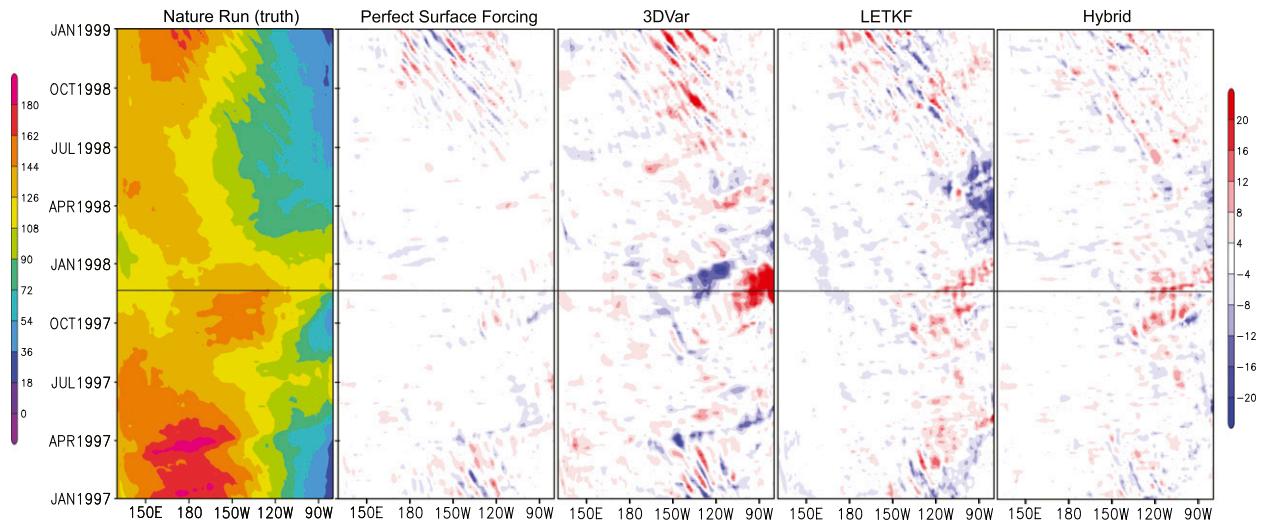


FIG. 8. The 20° isotherm depth (m) averaged 5°S–5°N and 130°E–80°W in the equatorial Pacific during 1997/98. (from left to right) Nature run (truth), errors in the reference perfect-forcing 3DVar, 3DVar (surface case 01), LETKF, and hybrid. The hybrid reduces persistent errors in the eastern Pacific for 3DVar (surface case 01) and LETKF associated with the onset of ENSO, as well as errors caused by observational noise.

period, as seen in Fig. 3). This corresponds to the transition into the peak deepening of the thermocline in the eastern Pacific in the fall of 1997. As the El Niño subsides during the spring of 1998, both 3DVar and LETKF have increased errors in their response to the reversal. The ensemble spread in this region is small, and is indicative of a temporary filter divergence because LETKF overly trusts the model forecast versus the observations during this period. The hybrid is the most accurate in this case, as it compensates for the underestimated ensemble spread with a fixed minimum background error covariance. As shown in Fig. 9, LETKF introduces a shallow bias in the western Pacific that is reduced using the hybrid. The hybrid reduces 1) the majority of the analysis errors exhibited by LETKF throughout the western Pacific, 2) the most extreme errors in the eastern Pacific exhibited by 3DVar in Q4 1997 and LETKF during spring 1998, and 3) the westward propagating errors of the reference perfect-forcing case throughout the second half of 1998.

The three cases for 3DVar, two using randomly selected surface forcing conditions from the 28-member ensemble and one using perfect forcing, indicate a range of possible performance of 3DVar depending on the surface boundary conditions. The perfect initial conditions in the reference perfect-forcing case are forgotten within the first few months. This time scale is largely determined by the number of observations assimilated, and the magnitude of their errors. A large spike in RMSEs is present in the 3DVar (01) case during the 1997/98 ENSO period. Because the only difference

between this case and the reference perfect-forcing case is the surface forcing, and the spike does not appear in the 3DVar (16) case, the cause can be attributed to the surface forcing rather than a fundamental aspect of the 3DVar method.

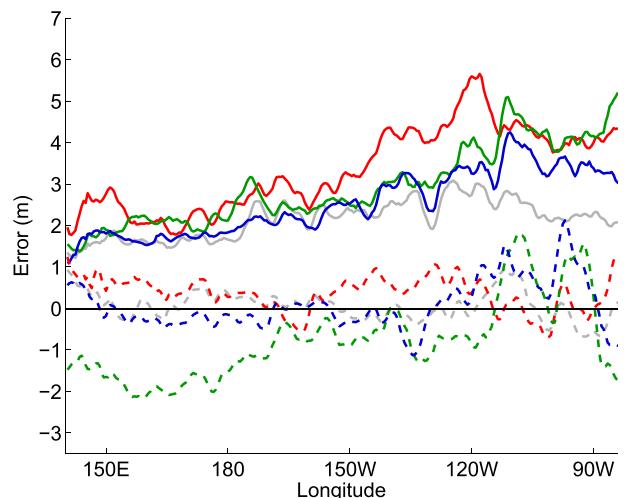


FIG. 9. Time mean absolute errors (solid) and bias (dashed) of the 20° isotherm depth in the equatorial Pacific from 5°S to 5°N, corresponding to Fig. 7 for the reference perfect-forcing 3DVar (gray), 3DVar (surface case 01) (red), LETKF (green), and hybrid (blue). Across the equatorial Pacific, the hybrid reduces average absolute errors below those of 3DVar and LETKF. LETKF introduces a shallow bias in the western Pacific that is reduced using the hybrid, thus reducing the absolute errors for the hybrid to approximately equal with the reference perfect-forcing 3DVar.

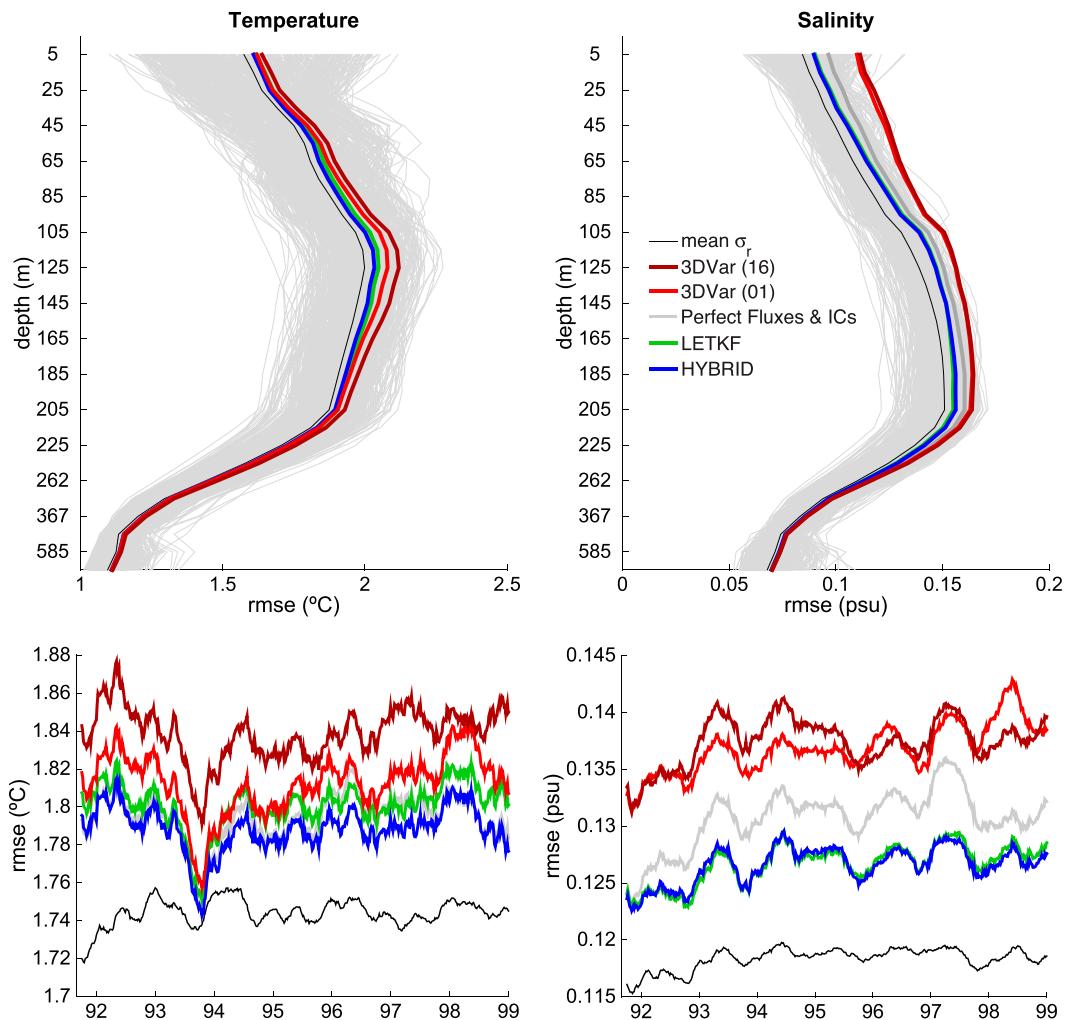


FIG. 10. (top) Daily profiles of mean global observation error (σ_r) for (left) T and (right) S from 1991 to 1998 (light gray), and their average (black) computed in observation space with the diagnosed observation errors shown for 3DVar (red for 01 and dark red for 16), the reference perfect-forcing 3DVar (dark gray), LETKF (green), and the hybrid (blue). (bottom) The 9-month moving average of daily mean global observation error (black) and diagnosed observation error (colors as in top) for (left) T and (right) S .

The Desroziers et al. (2005) approach is used to diagnose the observation error as shown in Fig. 10, in which the diagnosed observation errors for the ensemble methods are closer to the prescribed values than for the 3DVar experiments. The observation error variance is prescribed, and therefore known. The observation error cross-covariance, however, is assumed to be zero even though this assumption is known to be inaccurate. The diagnosed observation error, which is larger than the prescribed values, may be compensating for this inaccuracy.

The RMSE of each experiment, the known time-averaged ensemble spread of LETKF and the hybrid, and the corresponding climatological background error used by 3DVar are all shown in Fig. 11. The time-averaged

ensemble spread is nearly identical between LETKF and the hybrid. The ensemble spread underestimates the RMSEs, on average, while the climatological background error is an overestimate. Using the hybrid largely eliminates the bias implemented through the surface forcing fields on the LETKF solution.

A preliminary application of the hybrid to real historical temperature and salinity profile data indicates a clear reduction in temperature and salinity RMSD using the hybrid versus 3DVar (Fig. 12). Analogous results with the OSSE lead us to believe that this likely translates to a reduction in RMSEs. Biases are generally reduced by the hybrid. In particular, a steady growth in the temperature and salinity biases exhibited by 3DVar is completely eliminated by the hybrid. A gradual

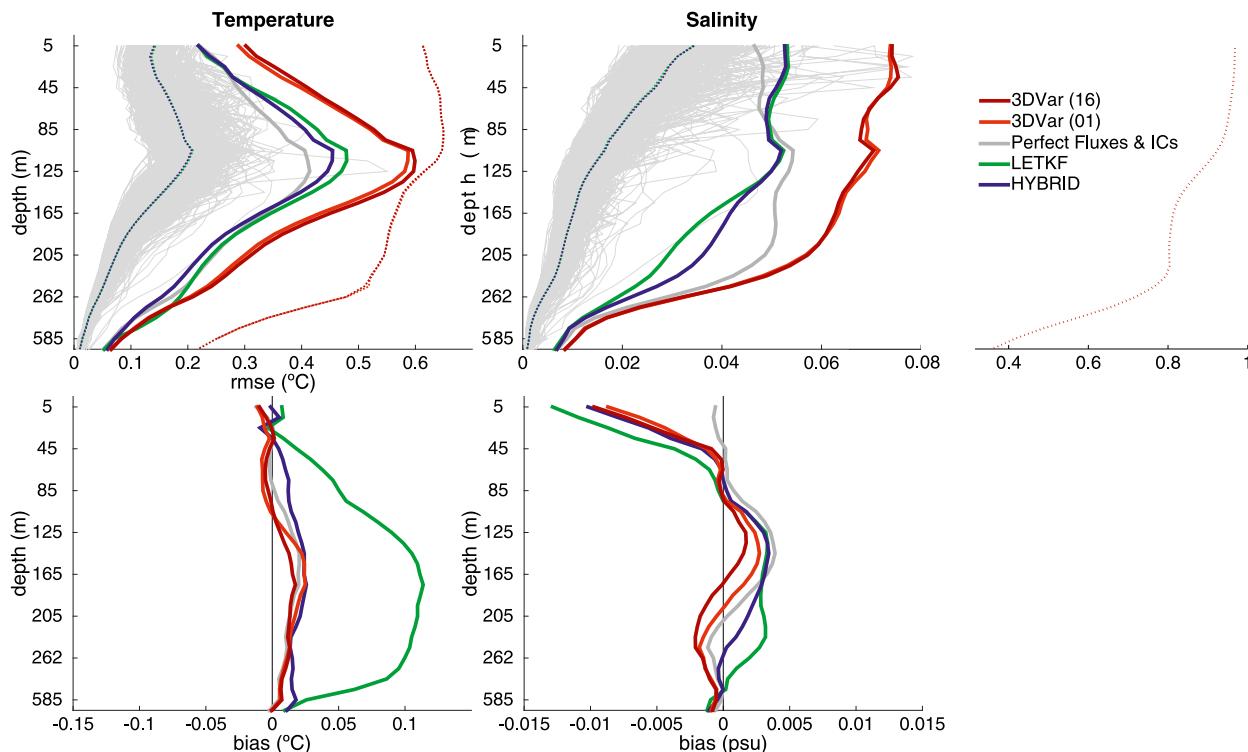


FIG. 11. (top) Global average forecast minus truth RMSE (solid gray and colored lines) compared to global average forecast ensemble spread (dashed gray and red lines) by depth for (left) T and (right) S over the experiment period 1991–98, computed in observation space. The distribution of daily global averages of ensemble spread for LETKF is shown (light gray) for the period 1991–98. (bottom) Corresponding bias for (left) T and (right) S (note the change in scale). Results are shown for LETKF (green), the hybrid (blue), 3DVar (red for 01 and dark red for 16), and the reference perfect-forcing 3DVar (dark gray). The LETKF ensemble typically underpredicts the errors attributable to the small sample size of the ensemble distribution; while 3DVar consistently overpredicts the errors. LETKF improves accuracy vs 3DVar but causes an increase in temperature and salinity biases. The hybrid, utilizing a combination of error estimates coming from both LETKF and 3DVar maintains the accuracy of LETKF while reducing the introduced biases.

decrease in temperature RMSD over time is paired with a gradual increase in salinity RMSD.

4. Conclusions

The Hybrid-GODAS significantly reduces RMS errors versus the NCEP operational 3DVar and reduces growing biases in temperature and salinity produced respectively by LETKF and 3DVar. The improvements in accuracy are found to hold for both observed and unobserved model fields, assessed both globally and regionally. In particular, the hybrid has superior performance in the tropical Pacific when compared to all other methods. Two main features of the Hybrid-GODAS are responsible for these improvements: 1) the improved representation of errors in the surface forcing and interior ocean via an ensemble and 2) the inflationary effects of the hybrid gain.

While 3DVar is limited to a single realization of surface forcing conditions, the LETKF component of the

hybrid can simultaneously use multiple plausible surface forcing scenarios. Further, the observation–localization used by LETKF allows the analysis ensemble to be formed with different linear combinations of the ocean members in different regions around the globe, thus expanding the impacts of the surface forcing ensemble diversity. Of course, any hybrid method is limited by the effectiveness of its climatological component. To improve the hybrid outside of the tropical Pacific, further work must be done to improve the background error covariance matrix used in the operational 3DVar.

Examination of the 20°C isotherm, a proxy for thermocline depth, in the equatorial Pacific reveals the largest errors occurred in the eastern Pacific during ENSO transitions. During the 1997/98 El Niño, surface conditions in one of the randomly chosen forcing perturbations used to drive the 3DVar caused broad errors in the 20°C isotherm depth throughout the latter half of 1997. Errors induced by both observational noise and surface forcing in the equatorial Pacific are generally

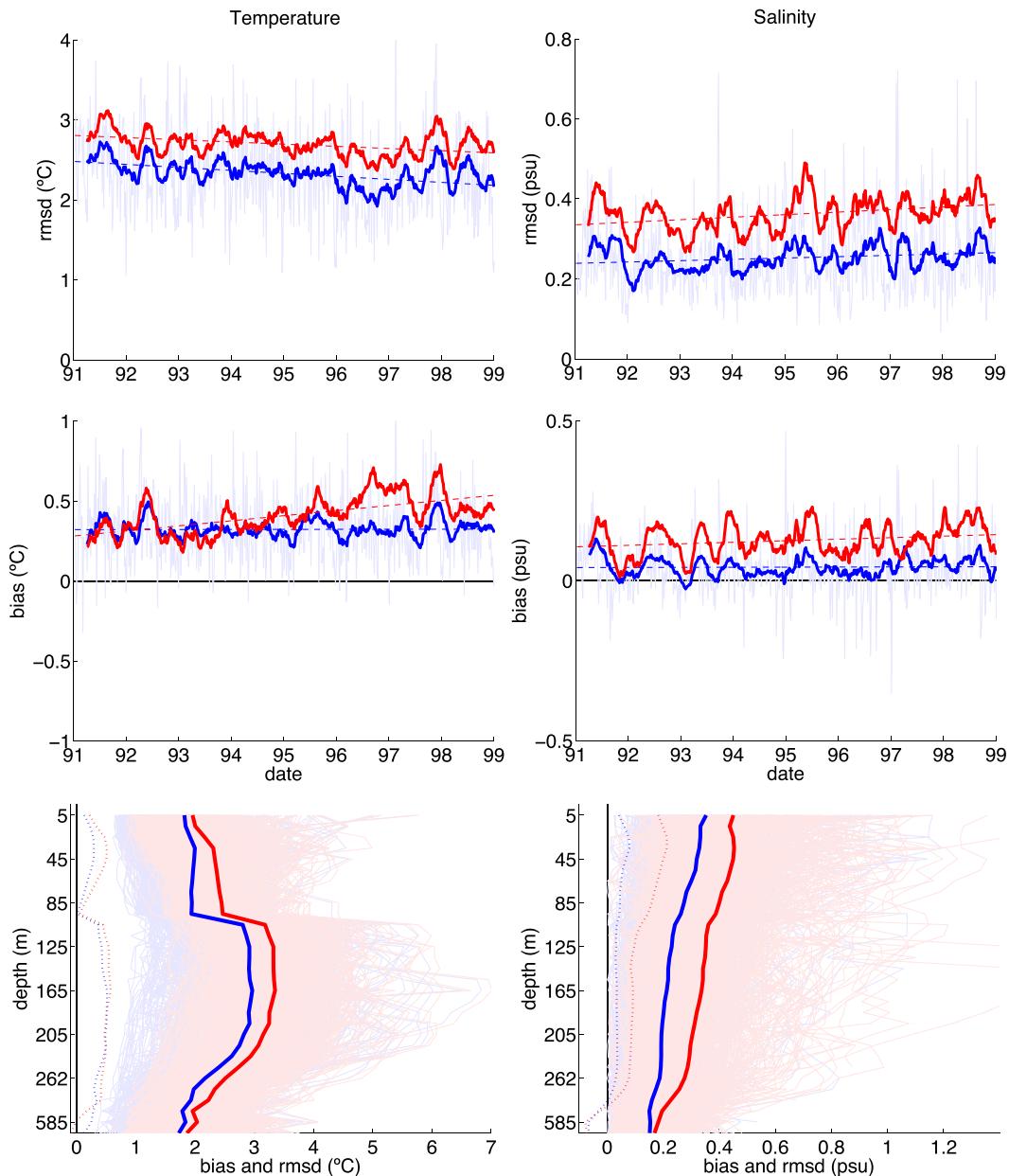


FIG. 12. Observed minus forecast (top) RMSD and (middle) bias for the historical data reanalysis experiment using the hybrid system (blue) and 3DVar (red) for (left) T and (right) S ; included is the 3-month moving average (solid lines), a linear regression (dashed lines), and the daily data for the hybrid (light blue thin lines). (bottom) The RMSD (solid lines) and bias (dotted lines) for each vertical level with daily data for the hybrid (light blue thin lines) and 3DVar (light red thin lines).

reduced by the hybrid. The hybrid analysis of the tropical Pacific simultaneously provides an improved estimate of upper ocean heat content along with quantitative estimates of uncertainty. With the combination of improved upper ocean heat content, improved SST, and consideration of surface wind stress uncertainty, we anticipate the Hybrid-GODAS will lead to increased

skill in ENSO prediction within a coupled model framework.

The results of a preliminary reanalysis using real observational data over the same time period as the OSSE indicate that a similar qualitative improvement is produced with the hybrid versus the operational 3DVar. While errors in the model and observed data increase

the total noise present in the system, the reduction in observed-minus-forecast RMSD in both temperature and salinity is unambiguous. Further, biases are generally reduced and positive trends in the bias of both temperature and salinity errors are completely eliminated by the hybrid. Future work is needed to extend this reanalysis through the Argo era, in which significant changes occurred in the global ocean observing network.

The EnKF methods appear to offer an advantage over a deterministic 3DVar approach. While we have demonstrated quantitatively that the mean states have improved, it is difficult to isolate whether the primary advantage comes from a better estimation of the first or second statistical moments. Extending the set of two perturbed-forcing 3DVar OSSE cases to an exhaustive set of 28 cases using surface forcing corresponding to each of the LETKF and hybrid ensemble members could strengthen the argument for the latter. We have, however, identified reduction in errors by LETKF along evolving temperature and salinity gradients, as compared to the coarse corrections made by the dynamically unaware 3DVar. We note that only errors in surface forcing have been considered for representing model errors in this study. Increased representation of systematic model errors is needed for further improvement of ensemble-based ocean data assimilation.

The Hybrid-GODAS is currently being implemented as the ocean component in the prototype next-generation CFSv3 at NCEP. In addition to its application for seasonal forecasting, the NCEP Climate Prediction Center (CPC) uses the output from GODAS to monitor and understand ENSO in near-real time (Huang et al. 2010). We anticipate that the improvements in state estimation and the introduction of uncertainty estimates, particularly for SST, will significantly enhance the utility of GODAS for this purpose. Further, we expect the hybrid approach to be applied with the Simple Ocean Data Assimilation (SODA) analysis of decadal ocean variability (Carton et al. 2000a,b; Carton and Giese 2008) at the University of Maryland.

Acknowledgments. We acknowledge Dan Barrie and the NOAA Climate Program Office (CPO) Modeling, Analysis, Predictions and Projections (MAPP) program for making this work possible (NA130AR4310108). James Carton is supported by NSF Grant OCE0752209. Eugenia Kalnay is supported by Monsoon Mission Grant MMSERPUnivMaryland. We would like to acknowledge Hendrik Tolman and Suranjana Saha of the Environmental Modeling Center (EMC) for collaboration and for providing additional computing resources. Finally, we acknowledge Bill Lapenta for encouraging

the development of a hybrid data assimilation system for the ocean. The NOAA Research and Development High Performance Computing Program provided computing and storage resources that have contributed to the research results reported within this paper (<http://rdhpcs.noaa.gov>). Support for the Twentieth Century Reanalysis Project dataset is provided by the U.S. Department of Energy, Office of Science Innovative and Novel Computational Impact on Theory and Experiment (DOE INCITE) program, the Office of Biological and Environmental Research (BER), and the NOAA CPO.

REFERENCES

- Balmaseda, M. A., K. Mogensen, and A. T. Weaver, 2013: Evaluation of the ECMWF ocean reanalysis system ORAS4. *Quart. J. Roy. Meteor. Soc.*, **139**, 1132–1161, doi:10.1002/qj.2063.
- Behringer, D. W., 2007: The Global Ocean Data Assimilation System (GODAS) at NCEP. *11th Symp. on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, and Land Surface*, San Antonio, TX, Amer. Meteor. Soc., 3.3. [Available online at <https://ams.confex.com/ams/pdfpapers/119541.pdf>.]
- , and Y. Xue, 2004: Evaluation of the global ocean data assimilation system at NCEP: The Pacific Ocean. *Eighth Symp. on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, and Land Surface*, Seattle, WA, Amer. Meteor. Soc., 2.3. [Available online at <https://ams.confex.com/ams/pdfpapers/70720.pdf>.]
- , M. Ji, and A. Leetmaa, 1998: An improved coupled model for ENSO prediction and implications for ocean initialization. Part I: The ocean data assimilation system. *Mon. Wea. Rev.*, **126**, 1013–1021, doi:10.1175/1520-0493(1998)126<1013:AICMFE>2.0.CO;2.
- 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.
- Bourlès, B., and Coauthors, 2008: The PIRATA Program: History, accomplishments, and future directions. *Bull. Amer. Meteor. Soc.*, **89**, 1111–1125, doi:10.1175/2008BAMS2462.1.
- Cane, M. A., S. E. Zebiak, and S. C. Dolan, 1986: Experimental forecasts of El Niño. *Nature*, **321**, 827–832, doi:10.1038/321827a0.
- Carton, J. A., and B. S. Giese, 2008: A reanalysis of ocean climate using Simple Ocean Data Assimilation (SODA). *Mon. Wea. Rev.*, **136**, 2999–3017, doi:10.1175/2007MWR1978.1.
- , G. Chepurin, X. Cao, and B. Giese, 2000a: A Simple Ocean Data Assimilation analysis of the global upper ocean 1950–95. Part I: Methodology. *J. Phys. Oceanogr.*, **30**, 294–309, doi:10.1175/1520-0485(2000)030<0294:ASODAA>2.0.CO;2.
- , —, and —, 2000b: A Simple Ocean Data Assimilation analysis of the global upper ocean 1950–95. Part II: Results. *J. Phys. Oceanogr.*, **30**, 311–326, doi:10.1175/1520-0485(2000)030<0311:ASODAA>2.0.CO;2.
- Chelton, D., R. A. deSzoeke, M. G. Schlax, K. El Naggar, and N. Siwertz, 1998: Geophysical 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.

- Chen, D., S. E. Zebiak, A. J. Busalacchi, and M. A. Cane, 1995: An improved procedure for El Niño forecasting. *Science*, **269**, 1699–1702, doi:10.1126/science.269.5231.1699.
- Compo, G. P., J. S. Whitaker, and P. D. Sardeshmukh, 2006: Feasibility of a 100-year reanalysis using only surface pressure data. *Bull. Amer. Meteor. Soc.*, **87**, 175–190, doi:10.1175/BAMS-87-2-175.
- , and Coauthors, 2011: The Twentieth Century Reanalysis Project. *Quart. J. Roy. Meteor. Soc.*, **137**, 1–28, doi:10.1002/qj.776.
- Conkright, M. E., and Coauthors, 1999: World Ocean Database 1998, documentation and quality control version 2.0. National Oceanographic Data Center Internal Rep. 14, 127 pp.
- Danabasoglu, G., and Coauthors, 2014: North Atlantic simulations in Coordinated Ocean-ice Reference Experiments phase II (CORE-II). Part I: Mean states. *Ocean Modell.*, **73**, 76–107, doi:10.1016/j.ocemod.2013.10.005.
- Derber, J. D., and A. Rosati, 1989: A global oceanic data assimilation system. *J. Phys. Oceanogr.*, **19**, 1333–1347, doi:10.1175/1520-0485(1989)019<1333:AGODAS>2.0.CO;2.
- Desroziers, G., L. Berre, B. Chapnik, and P. Poli, 2005: Diagnosis of observation, background and analysis-error statistics in observation space. *Quart. J. Roy. Meteor. Soc.*, **131**, 3385–3396, doi:10.1256/qj.05.108.
- Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, 10 143–10 162, doi:10.1029/94JC00572.
- Gent, P. R., and J. C. McWilliams, 1990: Isopycnal mixing in ocean circulation models. *J. Phys. Oceanogr.*, **20**, 150–155, doi:10.1175/1520-0485(1990)020<0150:MIOCM>2.0.CO;2.
- Greybush, S. J., E. Kalnay, T. Miyoshi, K. Ide, and B. Hunt, 2011: Balance and ensemble Kalman filter localization techniques. *Mon. Wea. Rev.*, **139**, 511–522, doi:10.1175/2010MWR3328.1.
- Griffies, S. M., and R. W. Hallberg, 2000: Biharmonic friction with a Smagorinsky-like viscosity for use in large-scale eddy-permitting ocean models. *Mon. Wea. Rev.*, **128**, 2935–2946, doi:10.1175/1520-0493(2000)128<2935:BFWASL>2.0.CO;2.
- , and Coauthors, 2009: Coordinated Ocean-Ice Reference Experiments (COREs). *Ocean Modell.*, **26**, 1–46, doi:10.1016/j.ocemod.2008.08.007.
- Hamrud, M., M. Bonavita, and L. Isaksen, 2014: EnKF and Hybrid Gain Ensemble Data Assimilation. ECMWF Tech. Rep. 733, 34 pp. [Available online at http://old.ecmwf.int/publications/library/ecpublications/_pdf/tm/701-800/tm733.pdf.]
- Hoffman, R. N., J. V. Ardizzone, S. M. Leidner, D. K. Smith, and R. Atlas, 2013: Error estimates for ocean surface winds: Applying Desroziers diagnostics to the cross-calibrated, multiplatform analysis of wind speed. *J. Atmos. Oceanic Technol.*, **30**, 2596–2603, doi:10.1175/JTECH-D-13-00018.1.
- Huang, B., Y. Xue, D. Zhang, A. Kumar, and M. J. McPhaden, 2010: The NCEP GODAS ocean analysis of the tropical Pacific mixed layer heat budget on seasonal to interannual time scales. *J. Climate*, **23**, 4901–4925, doi:10.1175/2010JCLI3373.1.
- Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D*, **230**, 112–126, doi:10.1016/j.physd.2006.11.008.
- Jin, F.-F., 1997: An equatorial recharge paradigm for ENSO. Part I: Conceptual model. *J. Atmos. Sci.*, **54**, 811–829, doi:10.1175/1520-0469(1997)054<0811:AEORPF>2.0.CO;2.
- Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter, 2002: NCEP–DOE AMIP-II Reanalysis (R-2). *Bull. Amer. Meteor. Soc.*, **83**, 1631–1643, doi:10.1175/BAMS-83-11-1631.
- Karspeck, A. R., S. Yeager, G. Danabasoglu, T. Hoar, N. Collins, K. Raeder, J. Anderson, and J. Tribbia, 2013: An ensemble adjustment Kalman filter for the CCSM4 ocean component. *J. Climate*, **26**, 7392–7413, doi:10.1175/JCLI-D-12-00402.1.
- Keppenne, C. L., M. M. Rienecker, J. P. Jacob, and R. Kovach, 2008: Error covariance modeling in the GMAO ocean ensemble Kalman filter. *Mon. Wea. Rev.*, **136**, 2964–2982, doi:10.1175/2007MWR2243.1.
- Kessler, W. S., 2002: Is ENSO a cycle or a series of events? *Geophys. Res. Lett.*, **29**, 2125, doi:10.1029/2002GL015924.
- Kleist, D. T., 2012: An evaluation of hybrid variational-ensemble data assimilation for the NCEP GFS. Ph.D. dissertation, University of Maryland, 149 pp. [Available online at <http://drum.lib.umd.edu/handle/1903/13135>.]
- Large, W. G., J. C. McWilliams, 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.
- Martin, M. J., and Coauthors, 2015: Status and future of data assimilation in operational oceanography. *J. Oper. Oceanogr.*, **8** (Suppl.), s28–s48, doi:10.1080/1755876X.2015.1022055.
- McPhaden, M. J., 2003: Tropical Pacific Ocean heat content variations and ENSO persistence barriers. *Geophys. Res. Lett.*, **30**, 1480, doi:10.1029/2003GL016872.
- , and Coauthors, 1998: The Tropical Ocean-Global Atmosphere observing system: A decade of progress. *J. Geophys. Res.*, **103**, 14 169–14 240, doi:10.1029/97JC02906.
- Meinen, C. S., and M. J. McPhaden, 2000: Observations of warm water volume changes in the equatorial Pacific and their relationship to El Niño and La Niña. *J. Climate*, **13**, 3551–3559, doi:10.1175/1520-0442(2000)013<3551:OOWWVC>2.0.CO;2.
- Murray, R. J., 1996: Explicit generation of orthogonal grids for ocean models. *J. Comput. Phys.*, **126**, 251–273, doi:10.1006/jcph.1996.0136.
- Oke, P. R., and Coauthors, 2013: Towards a dynamically balanced eddy-resolving ocean reanalysis: BRAN3. *Ocean Modell.*, **67**, 52–70, doi:10.1016/j.ocemod.2013.03.008.
- Ott, E., and Coauthors, 2004: A local ensemble Kalman filter for atmospheric data assimilation. *Tellus*, **56A**, 415–428, doi:10.1111/j.1600-0870.2004.00076.x.
- Penny, S. G., 2011: Data assimilation of the Global Ocean using the 4D-local ensemble transform Kalman filter (4D-LETKF) and the Modular Ocean Model (MOM2). Ph.D. dissertation, University of Maryland, 141 pp. [Available online at <http://hdl.handle.net/1903/11716>.]
- , 2014: The hybrid local ensemble transform Kalman filter. *Mon. Wea. Rev.*, **142**, 2139–2149, doi:10.1175/MWR-D-13-00131.1.
- , E. Kalnay, J. A. Carton, B. R. Hunt, K. Ide, T. Miyoshi, and G. A. Chepurin, 2013: The local ensemble transform Kalman filter and the running-in-place algorithm applied to a global ocean general circulation model. *Nonlinear Processes Geophys.*, **20**, 1031–1046, doi:10.5194/npg-20-1031-2013.
- Picaut, J., E. Hackert, A. J. Busalacchi, R. Murtugudde, and G. S. E. Lagerloef, 2002: Mechanisms of the 1997–1998 El Niño–La Niña, as inferred from space-based observations. *J. Geophys. Res.*, **107**, 3037, doi:10.1029/2001JC000850.
- Rajeevan, M., and M. J. McPhaden, 2004: Tropical Pacific upper ocean heat content variations and Indian summer monsoon rainfall. *Geophys. Res. Lett.*, **31**, L18203, doi:10.1029/2004GL020631.
- Reynolds, R. W., N. A. Rayner, T. M. Smith, D. C. Stokes, and W. Wang, 2002: An improved in situ and satellite SST

- analysis for climate. *J. Climate*, **15**, 1609–1625, doi:10.1175/1520-0442(2002)015<1609:AIISAS>2.0.CO;2.
- , T. M. Smith, C. Liu, D. B. Chelton, K. S. Casey, and M. G. Schlax, 2007: Daily high-resolution blended analyses for sea surface temperature. *J. Climate*, **20**, 5473–5496, doi:10.1175/2007JCLI1824.1.
- Saha, S., and Coauthors, 2006: The NCEP Climate Forecast System. *J. Climate*, **19**, 3483–3517, doi:10.1175/JCLI3812.1.
- , and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.*, **91**, 1015–1057, doi:10.1175/2010BAMS3001.1.
- , and Coauthors, 2014: The NCEP Climate Forecast System version 2. *J. Climate*, **27**, 2185–2208, doi:10.1175/JCLI-D-12-00823.1.
- Saji, N. H., and Coauthors, 1999: A dipole mode in the tropical Indian Ocean. *Nature*, **401**, 360–363.
- Shinoda, T., 2010: Observed dispersion relation of Yanai waves and 17-day tropical instability waves in the Pacific Ocean. *SOLA*, **6**, 17–20, doi:10.2151/sola.2010-005.
- Smith, T., A. G. Barnston, M. Ji, and M. Chelliah, 1995: The impact of Pacific Ocean subsurface data on operational prediction of tropical Pacific SST at the NCEP. *Wea. Forecasting*, **10**, 708–714, doi:10.1175/1520-0434(1995)010<0708:TIOPOS>2.0.CO;2.
- Sun, C., and Coauthors, 2010: The data management system for the Global Temperature and Salinity Profile Programme. *Proceedings of OceanObs.09: Sustained Ocean Observations and Information for Society*, Vol. 2, ESA Publication WPP-306, doi:10.5270/OceanObs09.cwp.86.
- Trenberth, K. E., J. M. Caron, D. P. Stepaniak, and S. Worley, 2002: Evolution of El Niño–Southern Oscillation and global atmospheric surface temperatures. *J. Geophys. Res.*, **107**, 4065, doi:10.1029/2000JD000298.
- Vospepoel, F. C., and D. W. Behringer, 2000: Impact of sea level assimilation on salinity variability in the western equatorial Pacific. *J. Phys. Oceanogr.*, **30**, 1706–1721, doi:10.1175/1520-0485(2000)030<1706:IOSLAO>2.0.CO;2.
- Wang, X., D. Parrish, D. Kleist, and J. Whitaker, 2013: GSI 3DVar-based ensemble–variational hybrid data assimilation for NCEP Global Forecast System: Single-resolution experiments. *Mon. Wea. Rev.*, **141**, 4098–4117, doi:10.1175/MWR-D-12-00141.1.
- Webster, P. J., A. M. Moore, J. P. Loschnigg, and R. P. Leben, 1999: Coupled ocean–atmosphere dynamics in the Indian Ocean during 1997–98. *Nature*, **401**, 356–360, doi:10.1038/43848.
- Whitaker, J. S., G. P. Compo, X. Wei, and T. M. Hamill, 2004: Reanalysis without radiosondes using ensemble data assimilation. *Mon. Wea. Rev.*, **132**, 1190–1200, doi:10.1175/1520-0493(2004)132<1190:RWRUED>2.0.CO;2.
- Wyrski, K., 1985: Water displacements in the Pacific and the genesis of El Niño cycles. *J. Geophys. Res.*, **90**, 7129–7132, doi:10.1029/JC090iC04p07129.
- Xue, Y., A. Leetmaa, and M. Ji, 2000: ENSO prediction with Markov models: The impact of sea level. *J. Climate*, **13**, 849–871, doi:10.1175/1520-0442(2000)013<0849:EPWMMT>2.0.CO;2.
- Zebiak, S. E., 1989: Ocean heat content variability and ENSO cycles. *J. Phys. Oceanogr.*, **19**, 475–486, doi:10.1175/1520-0485(1989)019<0475:OHCVAE>2.0.CO;2.
- , and M. A. Cane, 1987: A model El Niño–Southern Oscillation. *Mon. Wea. Rev.*, **115**, 2262–2278, doi:10.1175/1520-0493(1987)115<2262:AMENO>2.0.CO;2.
- Zhang, S., and A. Rosati, 2010: An inflated ensemble filter for ocean data assimilation with a biased coupled GCM. *Mon. Wea. Rev.*, **138**, 3905–3931, doi:10.1175/2010MWR3326.1.
- , M. J. Harrison, A. Rosati, and A. Wittenberg, 2007: System design and evaluation of coupled ensemble data assimilation for global oceanic climate studies. *Mon. Wea. Rev.*, **135**, 3541–3564, doi:10.1175/MWR3466.1.