

©SODA3: A New Ocean Climate Reanalysis

JAMES A. CARTON, GENNADY A. CEPURIN, AND LIGANG CHEN

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

(Manuscript received 12 March 2018, in final form 29 April 2018)

ABSTRACT

This paper describes version 3 of the Simple Ocean Data Assimilation (SODA3) ocean reanalysis with enhancements to model resolution, observation, and forcing datasets, and the addition of active sea ice. SODA3 relies on the ocean component of the NOAA/Geophysical Fluid Dynamics Laboratory CM2.5 coupled model with nominal $1/4^\circ$ resolution. A scheme has also been implemented to reduce bias in the surface fluxes. A 37-yr-long ocean reanalysis, SODA3.4.2, created using this new SODA3 system is compared to the previous generation of SODA (SODA2.2.4) as well as to the Hadley Centre EN4.1.1 no-model statistical objective analysis. The comparison is carried out in the tropics, the midlatitudes, and the Arctic and includes examinations of the meridional overturning circulation in the Atlantic. The comparison shows that SODA3.4.2 has reduced systematic errors to a level comparable to those of the no-model statistical objective analysis in the upper ocean. The accuracy of variability has been improved particularly poleward of the tropics, with the greatest improvements seen in the Arctic, accompanying a substantial reduction in surface net heat and freshwater flux bias. These improvements justify increasing use of ocean reanalysis for climate studies including the higher latitudes.

1. Introduction

Here we report on revisions to the Simple Ocean Data Assimilation (SODA) ocean reanalysis in the decade since the publication of the most recent version 2 (Carton and Giese 2008). These revisions include changes to the ocean model, the addition of active sea ice, new forcing datasets, changes to the updating data, and modifications to the data assimilation scheme.

The origins of ocean reanalysis can be traced to efforts by meteorologists in the 1990s to use data assimilation to create uniformly gridded reconstructions of the evolving state of the global atmosphere at time scales ranging from diurnal through decadal (Kalnay et al. 1996). These atmospheric reanalyses provided, as a by-product, all of the

terms required to specify surface heat, mass, and momentum fluxes, which the oceanographers could then use to drive ocean general circulation models (e.g., Carton et al. 2000a,b). Version 2 of SODA (SODA2) was developed as part of an effort to address inadequacies in the initial ocean reanalyses, including ocean models that were not truly global with resolution that did not permit eddies. Several different forcing datasets were used, but for the recent version of SODA2, SODA2.2.4, forcing is provided by the NOAA Twentieth Century Reanalysis version 2 (CR20v2) of Compo et al. (2011) (see Giese and Ray 2011). The data assimilation algorithm used in SODA2 is based on the computationally efficient optimal interpolation.

The design of the SODA3 framework recognizes that future reanalyses are likely to be carried out in a fully coupled atmosphere–ocean–sea ice system. For this reason SODA3 reanalyses are built on the Modular Ocean Model, version 5, ocean component of the Geophysical Fluid Dynamics Laboratory CM2.5 coupled model (Delworth et al. 2012), with fully interactive sea ice and improved $0.25^\circ \times 0.25^\circ$ horizontal and 50-level vertical resolution. The hydrographic dataset now includes the latest release of the World Ocean Database, a 40% increase over the dataset used in SODA2.2.4, as well as upgrades to the SST datasets. Assimilation of additional

© Denotes content that is immediately available upon publication as open access.

© Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/JCLI-D-18-0149.s1>.

Corresponding author: James A. Carton, carton@atmos.umd.edu

TABLE 1. Summary of parameters for SODA2.2.4 and SODA3.4.2. (OI is optimum interpolation. Expansions of acronyms are available online at <http://www.ametsoc.org/PubsAcronymList>.)

Name	Model numerics	Resolution (horizontal, vertical)	Surface forcing	Datasets	Assimilation	Flux bias correction	Time period
SODA2.2.4	POP2	$0.4^\circ \times 0.4^\circ$, 42L	20CRv2	WOD09, COADS, and Pathfinder SST	OI	No	1871–2010
SODA3.4.2	MOM5.1	$0.25^\circ \times 0.25^\circ$, 50L	ERA-Interim	WOD13 and COADS2.1	OI	Yes	1980–present

datasets, including sea level and near-surface currents as well, is being explored, but since each updating dataset has the potential to introduce systematic errors these additional datasets are excluded from the reanalysis described here.

Indeed, earlier generations of ocean reanalysis have contained systematic errors that raised concerns about their application for some problems such as climate change studies (e.g., they are not included in the Fifth Assessment Report of the Intergovernmental Panel on Climate Change; [Rhein et al. 2013](#)). These systematic errors have several sources in addition to changing observing systems, including measurement bias, inaccurate model physics and numerical resolution, and biases in fluxes and initial conditions. In recent years the oceanographic community has made great efforts to remove biases from the observation sets (e.g., [Levitus et al. 2009](#); [Boyer et al. 2016](#)) and to upgrade models to reduce model error ([Balmaseda et al. 2015](#)), leaving bias in surface fluxes ([Brunke et al. 2011](#); [Ren et al. 2014](#)) as a major source of systematic error.

To address the broad issue of systematic errors SODA3 consists of a growing ensemble of reanalyses, some members of which explore the impact of various assimilated datasets, while others explore dependence on surface forcing datasets. To address the issue of bias in surface forcing in SODA3 we adopt the iterative flux correction procedure of [Carton et al. \(2018\)](#) in which flux error is estimated from the misfits obtained from an initial ocean reanalysis to alter fluxes for a revised ocean reanalysis. For the experiment shown here, SODA3.4.2, we restrict the assimilated data to the basic hydrographic data and SST with forcing derived from the European Centre for Medium-Range Weather Forecasts interim reanalysis (ERA-Interim) daily average surface radiative and state variables ([Dee et al. 2011](#)).

The emphasis in the results section of this paper is to provide users with information and encouragement about the usefulness of the new 37-yr SODA3.4.2 ocean reanalysis for climate studies. Following a presentation of data and methods, [section 3](#) begins with a side-by-side comparison to the previous SODA2.2.4 and to available temperature and salinity data to explore the issue of systematic error as well as to evaluate accuracy. We

include in many of these comparisons the Hadley Centre EN4.1.1 analysis of [Good et al. \(2013\)](#) to clarify the changes in systematic error and accuracy that occur when shifting from a no-model statistical objective analysis to a full model-based ocean reanalysis. In the second half of [section 3](#) we discuss the behavior of SODA3.4.2 in the context of a selection of important problems in ocean climate variability. [Section 4](#) includes a discussion of the future development of SODA3.

2. Data and methods

This section describes the data, model, forcing, and data assimilation used in the construction and evaluation of SODA3 (see [Table 1](#)).

a. Data

The main datasets that SODA ingests are the World Ocean Database of historical hydrographic profiles ([Boyer et al. 2013](#)) and in situ and remotely sensed SST. Because of data rescue efforts the World Ocean Database now includes 15.4×10^6 profiles. Of these, 10.8×10^6 profiles were collected during our period of interest, including 1.8×10^6 profiling float profiles. Between 1980 and 2000, after $1^\circ \times 1^\circ$ spatial and 5-day temporal binning, the dataset contains typically 1000–5000 independent profiles per month, a number that increases to in excess of 10 000 profiles per month by 2010 ([Fig. 1](#)). Throughout this study we provide salinity values in practical salinity units based on the Practical Salinity Scale 1978 (PSS 78) and potential, rather than in situ, temperature.

The record prior to the 2000s is dominated by bathythermograph profiles that are subject to important evolving instrument-dependent biases (e.g., [Reverdin et al. 2009](#)). Here we adopt the corrections proposed by [Levitus et al. \(2009\)](#). The random error associated with the bathythermographs is also significant, but still less than the error of representativeness introduced by unresolved processes such as tides and eddies ([Janjić et al. 2018](#)). The bathythermographs have several additional limitations. First, they lack salinity observations. Second, they were shallow, generally less than 700 m. Third, they were mainly collected along ship tracks in the

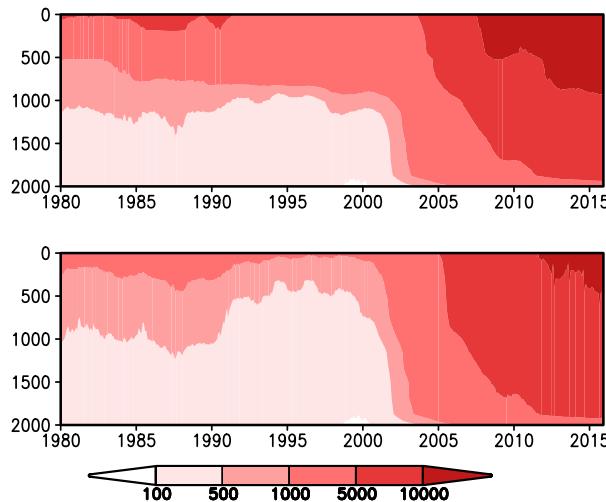


FIG. 1. Number of hydrographic profile observations used by SODA3, after spatial and temporal binning to remove redundancies, with depth and time 1980–2015, for (top) temperature and (bottom) salinity. Units are observations per month.

Northern Hemisphere. All of these limitations have been addressed by the gradual deployment of the Argo profiling floats beginning in the early 2000s.

In addition to profile hydrography data we assimilate both in situ and satellite SST observations. In situ SST observations are obtained from the International Comprehensive Ocean–Atmosphere Data Set (ICOADS) version 5 release 2 SST database (Woodruff et al. 2011). Since 1981 we have also included satellite SST observations. For the period October 1981–December 2002 we use the L3 Pathfinder version 5.2 Advanced Very High Resolution Radiometer SST data (Casey et al. 2010). The L3 designation indicates that the product has been binned into 4-km bins, but bins with no observations are unfilled. We limit the SST data we use to nighttime only to avoid introducing a warm bias associated with what for our model would be an unresolved near-surface diurnal warming. All NOAA satellites used by Pathfinder prior to NOAA-17 have a nighttime equatorial crossing at around 0200 LT. NOAA-17 has an equatorial crossing at around 2200 LT (Ignatov et al. 2004).

The Pathfinder SST data have not been corrected for the reduction in temperature when moving from the near-surface skin layer, whose thickness is measured in microns, down to the depth of the uppermost model level (see Fig. S1 in the supplemental material). For SODA3 we have added an empirical correction for this skin–bulk difference, following Grodsky et al. (2008). Beginning in 2003 we have found that Pathfinder SST has some unphysical features, including a temperature spike in 2004 and a cooling trend (Fig. S1). To avoid

these problems we switch to the NOAA Advanced Clear-Sky Processor for Oceans (ACSPO) Level 2P (L2P) SST product, using VIIRS and MODIS, at the beginning of 2003 (Ignatov et al. 2016). For ACSPO, nighttime is defined for each pixel by the requirement that the solar zenith angle exceeds 90° . The switch in SST datasets seems not to introduce a noticeable jump in the analysis SST (Fig. S2).

Opportunities to make a completely independent assessment of the accuracy of the ocean reanalysis hydrographic fields are limited because so much of the historical observation archive has already been assimilated. Here we track forecast error by collecting statistics of the differences between the assimilated observations ω^o and the forecasts of the same variables mapped onto the observation locations and times $\mathbf{H}(\omega^f)$. We track analysis error by examining the differences between ω^o and the similarly mapped analysis variables $\mathbf{H}(\omega^a)$. In these expressions ω represents either potential temperature or practical salinity.

As mentioned in the introduction a goal of SODA3 is to reduce systematic error. To help identify the presence of systematic error in SST we also compare the ocean reanalyses to the $1/4^\circ \times 1/4^\circ$ gridded statistical objective analysis of bulk SST of Reynolds et al. (2007) and to the $1^\circ \times 1^\circ$ gridded Met Office Hadley Centre EN4.1.1 monthly subsurface temperature and salinity statistical objective analysis of Good et al. (2013). EN4.1.1 relies on the same World Ocean Database 2009 (WOD09) dataset as SODA2.2.4 with the same Levitus et al. (2009) bathythermograph bias correction as used in SODA3.

SODA3 does not assimilate velocity information, which provides the opportunity to carry out an independent check of ocean reanalysis velocity against Tropical Atmosphere Ocean (TAO) TRITON mooring observations along the equator where the flow is ageostrophic. In the subtropical North Atlantic the RAPID–MOCHA measurements of overturning at 26°N (Cunningham et al. 2007) provide another opportunity to evaluate ocean reanalysis transport. This latter calculation is highly sensitive to small errors, which requires that it be carried out on the original grid, which is not available for SODA2.2.4. Additional comparison to satellite sea level is provided in the supplemental material.

The cryospheric component of SODA3 is also unconstrained in the experiments described in this paper. Our examination of the cryospheric component focuses on the Arctic. In the Arctic the summer minimum sea ice extent, defined by the area of the Arctic with at least 15% sea ice coverage, varies widely from year to year. To provide some information about how well SODA3.4.2

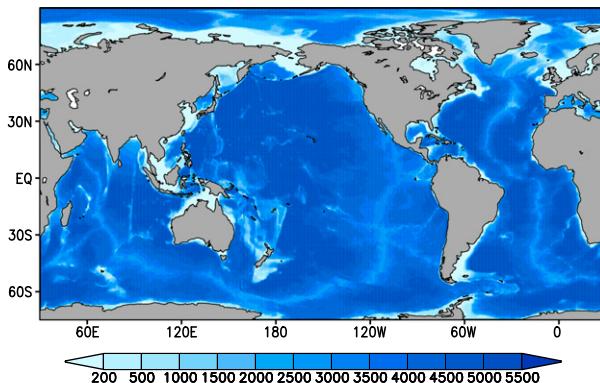


FIG. 2. ETOPO5 topography used in SODA3 (m). Note the inclusion of many marginal seas.

reproduces the observed sea ice cover we include comparison to the microwave-based concentration estimates of the seasonal minimum provided by the National Snow and Ice Data Center (NSIDC). To evaluate the mean stratification below the ice we compare the ocean reanalyses to the $1^\circ \times 1^\circ$ Polar Science Center (PSC) Hydrographic Climatology version 3 (PHC3.0, Steele et al. 2001), bearing in mind that this climatology was constructed primarily from data prior to the 1990s and thus lacks the influence of the more recent Arctic warming.

b. Model

The global ocean–sea ice model is the ocean–sea ice component of the GFDL CM2.5 coupled model (Delworth et al. 2012) with 1440×1070 eddy permitting quasi-isotropic horizontal grid cells that vary in size from 28 km at the equator to 14 km at 60°N . In the Arctic the grid splits into two geographically displaced north poles giving, for example, approximately 5-km horizontal resolution at Fram Strait east of northern Greenland. Bottom topography is based on the NOAA/National Centers for Earth Information 5' grid Earth topography (ETOPO5) (Fig. 2) with a few modifications to narrow passages. The topographic mask has a number of changes from that used for SODA2 including resolved marginal seas such as Hudson Bay, and improved resolution of important passages. In the Indonesian Throughflow this improved resolution allows representation of features such as Timor Strait and Micronesia, which are largely missing from the SODA2 topography (Fig. S3 in the supplemental material). ETOPO5 predates much of the altimeter era and as a consequence there are a few errors in the topography, for example on the Siberian shelves.

The model has 50 levels in the vertical using a z^* vertical coordinate (Griffies et al. 2000). In the vertical the grid telescopes from 10-m resolution in the upper

100 m to much coarser resolution in the deep ocean. The model uses a third-order advection scheme on an Arakawa B grid with no explicit horizontal diffusion and a scale-selective Smagorinsky viscosity, enhanced in the region of the western boundary currents. In addition, enhanced nearshore tidal mixing is parameterized. The sea ice model uses the GFDL Sea Ice Simulator dynamics and thermodynamics of Winton (2000) on the same numerical grid. Snow albedo is fixed to be 0.85, which lies in the middle of observational estimates, while ice albedo has a high value of 0.8. Surface fluxes are calculated in the GFDL Flexible Modeling System coupler, which takes into account the changing distributions of snow and ice as well as changing SST and surface currents. Monthly continental discharge is provided by Dai et al. (2009) with monthly Greenland discharge following Bamber et al. (2012).

c. Data assimilation

SODA3.4.2 uses a linear deterministic sequential filter in which the ocean state ω^a is constructed from a forecast ω^f based on the difference between observations ω^o and ω^f mapped onto the observation variable and its location $\mathbf{H}(\omega^f)$:

$$\omega^a = \omega^f + \mathbf{K}[\omega^o - \mathbf{H}(\omega^f)], \quad (2.1)$$

where the gain matrix \mathbf{K} determines the impact of the observations, depending on the observation error covariance $\mathbf{R}^o \equiv \langle \mathbf{e}^o \mathbf{e}^{o\top} \rangle$ and the model forecast error covariance $\mathbf{P}^f \equiv \langle \mathbf{e}^f \mathbf{e}^{f\top} \rangle$. A direct implementation of (2.1) would introduce shocks and spurious waves. Instead we implement the incremental analysis update procedure of Bloom et al. (1996) using an update cycle of 10 days (a period chosen to be consistent with the data sampling and the time scale of ocean variability). The incremental analysis update procedure is a predictor–corrector digital filter that slowly modifies the model state over an analysis cycle. The form of $\mathbf{K} = \mathbf{P}^f \mathbf{H}^T (\mathbf{H} \mathbf{P}^f \mathbf{H}^T + \mathbf{R}^o)^{-1}$ is determined by minimizing the expected variance of the analysis error subject to some simplifying assumptions, including the assumption that the model forecast and observation errors are unbiased, and thus the analysis errors are unbiased. The presence of biased forecast errors and their relationship to bias in surface forcing is discussed in Carton et al. (2018) and in section 2d. The observation errors included in SODA3.4.2 are assumed spatially uncorrelated so \mathbf{R}^o is diagonal with a signal-to-noise ratio of 1.4 (profile observations) and 1.2 (SST observations). Figure S1 in the supplemental material shows that this assumption is violated for remotely sensed SST. Further discussion of observation error

characteristics is provided in, for example, [Oke and Sakov \(2008\)](#) and [Janjić et al. \(2018\)](#).

For the experiment described here the model forecast error \mathbf{P}^f has the form

$$\begin{aligned} \mathbf{P}^f[\Delta x, \Delta y, \Delta z, \Delta t, \bar{y}, z_1, z_2, \rho(z), \Delta \eta] \\ = \Phi \exp \left[-\left(\frac{\Delta \eta^2}{\gamma_\eta^2} + \left| \frac{\Delta x}{\gamma_x} \right| + \left| \frac{\Delta y}{\gamma_y} \right| + \left| \frac{\Delta z}{\gamma_z} \right| + \left| \frac{\Delta t}{\gamma_t} \right| \right) \right], \quad (2.2) \end{aligned}$$

with prespecified dependence on the zonal, meridional, and vertical spatial differences, Δx , Δy , and Δz , respectively, and time difference Δt between forecast and observation points. The dependence of \mathbf{P}^f on change in sea level height $\Delta \eta$, with $\gamma_\eta = 1\text{m}$, allows the error covariance to depend on the flow field, elongating the error covariance scale along strong currents such as the Gulf Stream. For the other scales (i.e., γ_i) we follow [Carton and Giese \(2008\)](#), who include some anisotropy (zonal scales are larger than meridional scales in the tropics) and some latitude dependence (horizontal scales decrease somewhat with increasing latitude). Experiments have shown a general lack of sensitivity of the ocean reanalysis to the specification of these scales. A significant improvement is obtained by switching to replacing (2.2) with an evolving estimate from an ensemble Kalman filter ([Penny et al. 2015](#)), but this will increase the computational cost by perhaps a factor of 50. Because of the improvement, some SODA3 ensemble members, including SODA3.4.2, will be rerun with the ensemble Kalman filter.

The forecast error covariance between temperature and salinity errors is determined from an empirical table constructed from the hydrographic profile dataset. Temperature forecast errors modify the salinity forecast through the temperature–salinity forecast error covariance so that we maintain climatological water mass properties in the absence of direct observations of changing water masses. [Figure 3](#) illustrates the relationship between these variables at a single location in the western subtropical North Atlantic. At this location the temperature–salinity relationship is quite similar for all three analyses we consider, SODA3.4.2, SODA2.2.4, and EN4.1.1 at levels shallower than about 900 m. At deeper levels, the SODA2.2.4 temperature–salinity relationship has greater scatter indicating that new water masses are being created.

Forecast errors within the mixed layer are observed to have a much higher vertical correlation than forecast errors spanning the base of the mixed layer. To determine the depth at which this change in vertical correlation occurs we first estimate the mixed layer depth prior to each assimilation cycle from the forecast using

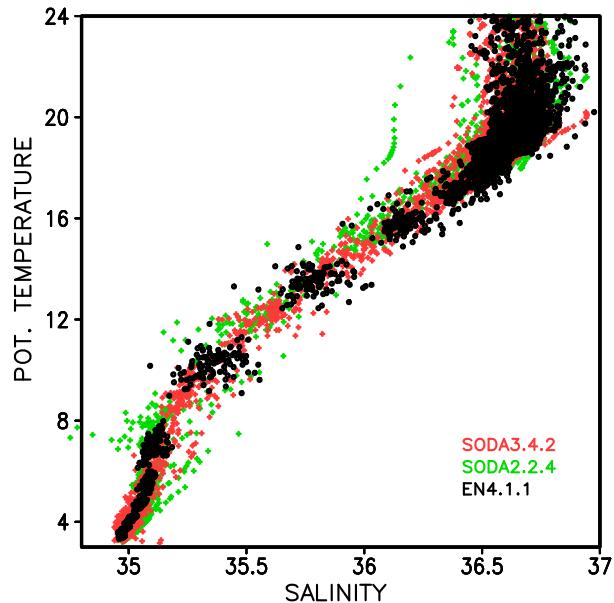


FIG. 3. Potential temperature ($^{\circ}\text{C}$) vs salinity (psu) for SODA2.2.4 (green), SODA3.4.2 (red), and EN4.1.1 (black) at a fixed location (32°N , 62°W) just east of Bermuda in the subtropical North Atlantic for the 9-yr period 2000–08. The temperature–salinity relationship is reasonably consistent among the three except for temperatures below 8°C corresponding to depths below 900 m. At those depths SODA2.2.4 shows greater variability.

0.05 standard deviation (σ), somewhat less stringent than the 0.03σ criterion of [de Boyer Montégut et al. \(2004\)](#). A comparison of the SODA3.4.2 climatological maximum monthly mixed layer depth to the observational estimates of [de Boyer Montégut et al.](#) is shown in Fig. S4 of the supplemental material. Information from surface observations, such as SST, is thus effectively propagated down through the mixed layer, but has less influence on water below the mixed layer.

d. Surface forcing

SODA3.4.2 is forced by ERA-Interim ([Dee et al. 2011](#)) near-surface atmospheric variables (neutral winds at 10-m height, air temperature and humidity at 2-m height, sea level pressure, and liquid and solid precipitation) together with downwelling short and longwave radiative fluxes. These are applied to the ocean–sea ice system through the GFDL Flexible Model System flux coupler. For SODA3.4.2 the coupler includes the Coupled Ocean–Atmosphere Response Experiment ([Fairall et al. 2003](#)), version 4 (COARE4), bulk formulas. Some SODA3 ensemble members not described here use an alternative set of bulk formulas by [Large and Yeager \(2009\)](#).

We evaluate the compatibility of this flux specification with the ocean observations by accumulating the gridded

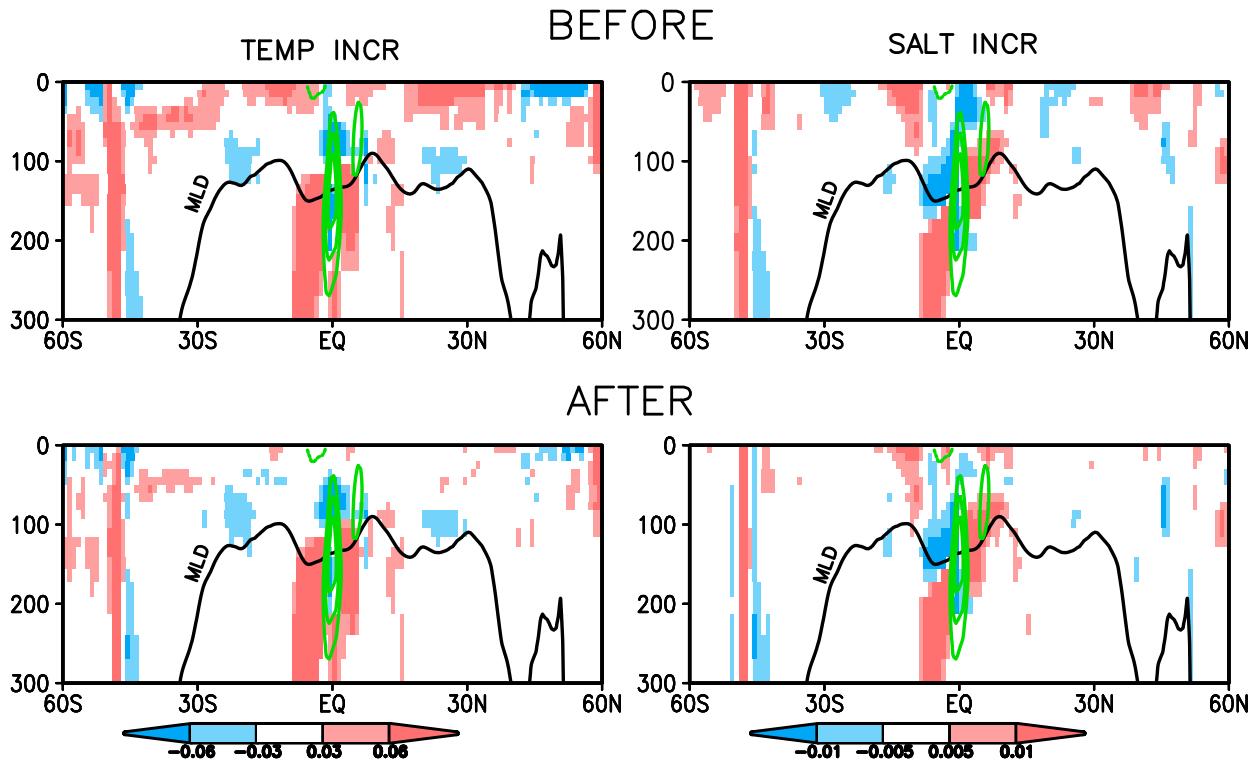


FIG. 4. Mean (left) temperature $\theta^o - \mathbf{H}(\theta^f)$ [K (10 day) $^{-1}$] and (right) salinity increments $S^o - \mathbf{H}(S^f)$ [psu (10 day) $^{-1}$] with latitude and depth (m), averaged in the longitude band 170°W – 180° for experiments using ERA-Interim forcing. The average depth of the density-based mixed layer (MLD) is shown with the black contour and zonal velocity with green contours (0.2 m s^{-1} contour interval). (top) Analysis increments before heat and freshwater fluxes are corrected. (bottom) Analysis increments for SODA3.4.2, after heat and freshwater fluxes have been corrected as described in the text. Note the reduction in misfit except near the strong equatorial currents.

potential temperature and salinity analysis increments $\{\delta\theta = [\theta^o - \mathbf{H}(\theta^f)]\}$ and $\delta S = [S^o - \mathbf{H}(S^f)]\}$ during a preliminary 8-yr ocean reanalysis 2007–14. A latitudinal slice of the time mean of these increments is shown in Fig. 4 (top). It is evident that the time mean increments are largest within the mixed layer except near the equator where the increments correspond to the locations of strong currents. As described in Carton et al. (2018), the presence of these time mean increments in the mixed layer highlights the presence of biases in net surface heat and freshwater flux. As shown in Carton et al. we can apply net surface heat flux and surface freshwater flux increments, δQ and $\delta(P - E + R)$, derived from the incremental heat and freshwater budget equations:

$$\begin{aligned}\delta Q &\approx \frac{\rho C_p}{\Delta t} \int_{z=-D}^0 \delta\theta dz \quad \text{and} \\ \delta(P - E + R) &\approx \frac{1}{S_o \Delta t} \int_{z=-D}^0 \delta S dz,\end{aligned}\quad (2.3)$$

to improve ERA-Interim net surface heat and freshwater flux. Here D is the depth of the mixed layer, S_o is surface salinity, ρ is water density, C_p is water specific

heat, and $\Delta t = 10$ days is the length of an assimilation cycle. Figure S5 in the supplemental material shows the geographic structure of the resulting modifications to time mean ERA-Interim heat and freshwater fluxes while the modified time mean heat and freshwater fluxes are shown in Fig. 5.

To illustrate the impact of the flux bias-correction procedure we present the time mean analysis increments of potential temperature and salinity from SODA3.4.2 (which uses the modified fluxes) in Fig. 4, bottom. Modifying the fluxes reduces the analysis increments and thus increases the compatibility of heat and freshwater exchanges between atmosphere and ocean with the historical ocean observation set. However, in regions of strong currents the model misfits remain even after heat and freshwater flux are corrected. Also, the reader is reminded that small changes to the model state made each assimilation cycle, even if they do not lead to a time mean bias, will inflate quadratic terms like vertical eddy temperature transport $w'\theta'$. This concern about estimation of quadratic terms applies to any atmosphere or ocean reanalyses using short assimilation windows.

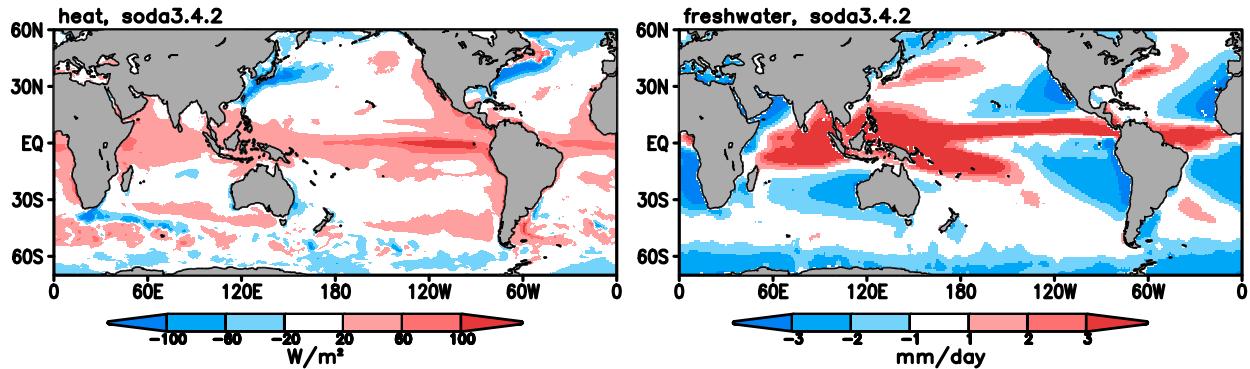


FIG. 5. Time mean net surface (left) heat (W m^{-2}) and (right) freshwater flux (mm day^{-1}).

e. Output datasets

Ocean reanalysis output is saved separately for the ocean and sea ice variables on the native grid at 5-day resolution. The ocean files contain the ocean state (temperature, salinity, and velocity) as well as a number of diagnostic variables such as mixed layer depth and entrainment, individual components of net surface heat flux, sea level, and so on. Volume transport estimates in and out of each model cell are saved on the original grid in a separate file at 10-day resolution.

A subset of the variables contained in the original files are regressed onto a uniform $0.5^\circ \times 0.5^\circ$ horizontal grid using a conservative mapping scheme (retaining the original vertical grid), and then are also monthly averaged (both 5-day and monthly average regressed datasets are available). A 1-yr-long file containing 12 monthly regressed ocean datasets is 3.3 Gb in size, while the corresponding regressed file containing sea ice variables (thickness categories and concentration) is 88 Mb in size. Finally, some ancillary fields are created from the regressed files. These include depth, potential temperature, and salinity on 16 isopotential density surfaces (referenced to 0 m) ranging from 1028 to 1024 kg m^{-3} .

3. Results

a. SST and SSH

We begin by comparing potential temperature and salinity fields at 5-m depth on monthly time scales during 1980–2008, using the gridded Optimum Interpolation Sea Surface Temperature (OISST) analysis of Reynolds et al. (2007) as a reference for this comparison (Fig. 6). The time mean differences show that both SODA3.4.2 and SODA2.2.4 have a small 0.25°C warmer bulk mixed layer than OISST throughout much of the northern subtropics. As shown in Fig. S1, a difference of this size can be introduced simply by the choice of observation

sets. In contrast, south of 60°S SODA2.2.4 has a 1°C warm bias, which we think is due to the lack of sea ice. SODA2.2.4 monthly variability is synchronized with OISST only in the tropics while SODA3.4.2 remains similar to OISST at all latitudes (Fig. 6, right).

Sea level observations are not assimilated, allowing us to use this variable to provide independent comparison. For the period 1993–2015 when continuous satellite sea level is available, SODA3.4.2 explains over 40% of the variance of annual average satellite sea level with correlations between the two in excess of 0.8 and rms differences below 3 cm throughout the tropical Pacific and Indian Oceans (Fig. S6 in the supplemental material). Correlations are lower in the low-variability tropical Atlantic, and also in the Southern Ocean due to the presence of more complicated physics including strong barotropic motions. No attempt has been made, in this comparison, to account for the limited global ocean mass changes in SODA3.4.2.

b. Upper and middepth layers

We next examine subsurface potential temperature and salinity beginning with the upper 300 m. For this comparison we map the ocean reanalyses onto the hydrographic profile observation locations, compute the analysis increments [$\mathbf{H}(\omega^a) - \omega^o$], and then accumulate the time average statistics (mean and root-mean-square after removal of the mean). The time mean of the analysis increments reveals that EN4.1.1 is nearly unbiased, SODA3.4.2 has a warm bias confined to the Pacific and Atlantic equator, and SODA2.2.4 is systematically cool and generally fresh (Figs. 7 and 8, left). The cause of the SODA3.4.2 equatorial warm bias has been traced to the slight weakness of the ERA-Interim equatorial trade winds (Dussin et al. 2016), thus providing insufficient equatorial upwelling. The root-mean-square of the analysis minus observation misfits (Figs. 7 and 8, right) shows that outside of the tropics the SODA2.2.4 analysis errors are within a factor of 2 of the

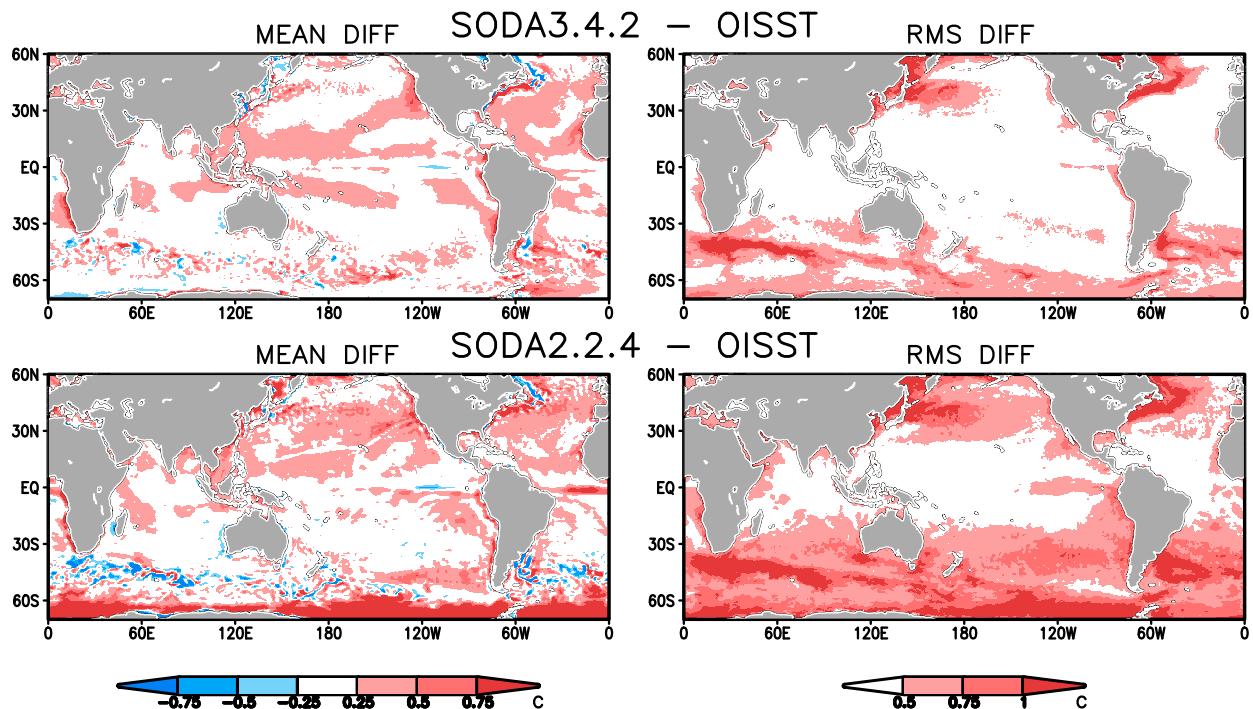


FIG. 6. (left) Mean and (right) root-mean-square difference of monthly SST ($^{\circ}$ C) from the OISST analysis of Reynolds et al. (2007) after removal of the mean difference during 1980–2008, for (top) SODA3.4.2 and (bottom) SODA2.2.4.

monthly RMS variability (Figs. 7 and 8, top right). In contrast, both SODA3.4.2 and EN4.1.1 show lower RMS differences between the analyses and the observation set in the 0–300-m layer.

One component of the 0–300-m monthly potential temperature variability is a weak 0.008 K yr^{-1} (equivalent to 0.25 W m^{-2}) global warming trend, which represents part of the ocean's uptake of the global heat flux imbalance introduced by changing greenhouse gas levels in the atmosphere. All three analyses show quite similar vertically averaged temperature trends, as well as similar details such as the cooling of the upper ocean in the early 1980s and again following the 1991 eruption of Mount Pinatubo (Fig. 9, left). The analyses show the widely discussed hiatus in warming extending only from 1998 to 2004 (e.g., Meehl et al. 2011; Kosaka and Xie 2013; Chen and Tung 2014).

The warming is concentrated in several geographic regions: the western side of the Pacific, the south subtropical Indian Ocean, and the North Atlantic (Fig. S7 in the supplemental material). The 300–1000-m layer warms at an average rate one-quarter that of the upper 300 m (and thus storing 60% as much of the excess planetary heat as the upper layer). The warming in this layer occurs throughout the Atlantic and in the Southern Ocean as well (Fig. S7). In the 1000–2000-m layer EN4.1.1 and SODA3.4.2 are also warming; however, we do not know

the extent to which the increase with time of the observation count in this depth range may contribute to this result. Note that SODA2.4 included a weak relaxation to climatology below 1000 m, which explains the absence of a trend in the 1000–2000-m temperature.

Global ocean salt is nearly constant due to the lack of salt sources. However, shifts in the strength of the atmospheric hydrologic cycle seem to cause basin-scale changes in sea surface salinity (Durack and Wijffels 2010). A recent study by Friedman et al. (2017) suggests that the most dramatic change in recent decades has been the salinification of the surface Atlantic 40°S – 40°N , a result which motivates our examination of Atlantic salinity (Fig. 9, right). While quite noisy, all three analyses do show increasing 0–300-m salinity during 1980–2008. However, as with 1000–2000-m temperatures, we do not know the extent to which the observed positive subsurface salinity trend is caused by the increasing salinity observation count after 2000.

Closely tied to the issue of changing Atlantic salinity is the issue of the changing strength of the meridional overturning circulation. In a previous study using SODA2 Zheng and Giese (2009) estimate a mean overturning transport a 26°N of approximately $16 \times 10^6 \text{ m}^3 \text{ s}^{-1}$ and also document indications of an increasing transport trend over multiple decades. Since 2004 direct observations have been collected by the RAPID-MOCHA

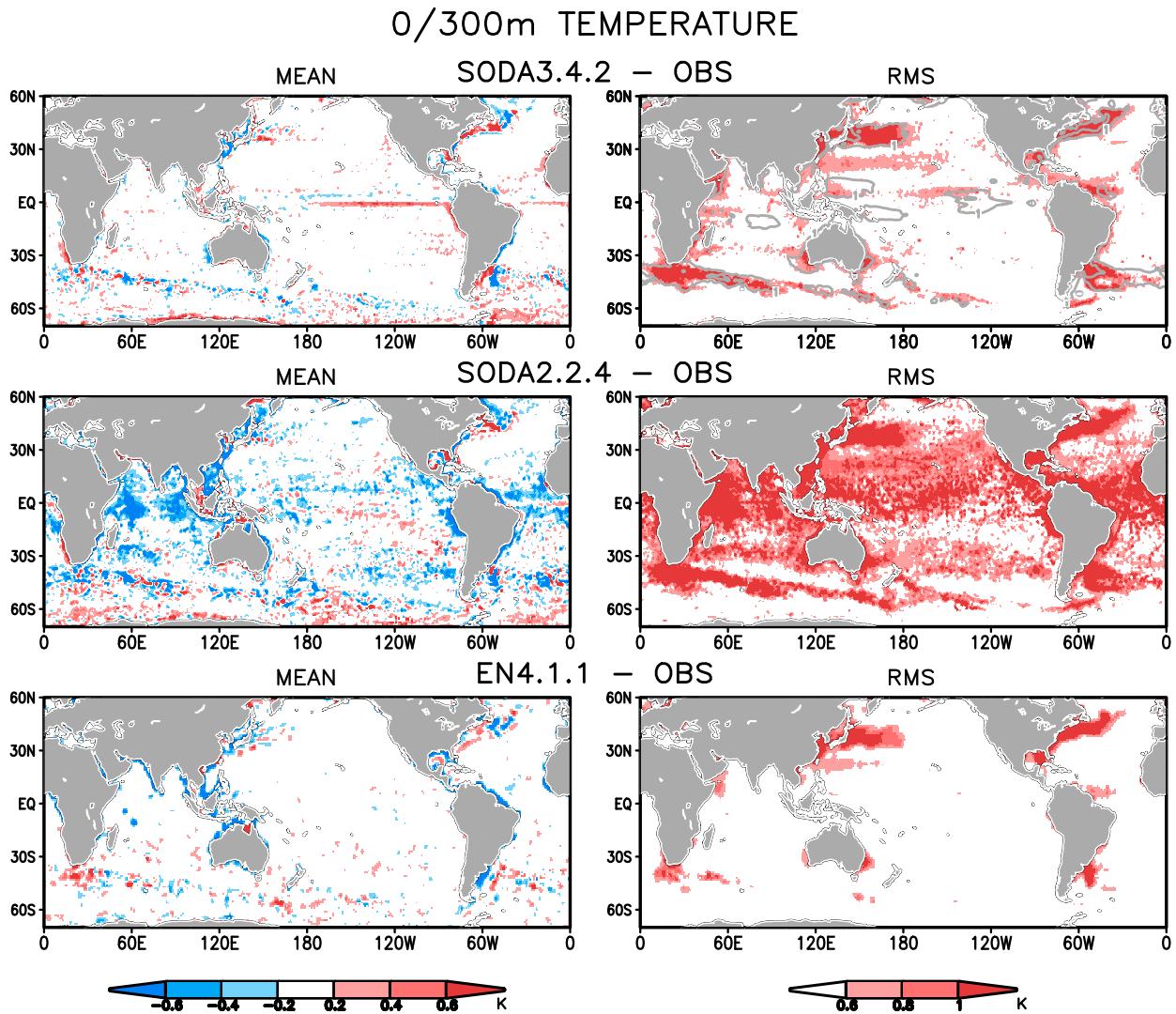


FIG. 7. Statistics of the binned analysis minus observation (OBS) potential temperature differences $H(\theta^a) - \theta^o$ (K) averaged in time and then averaged over 0–300 m during the 29-yr period 1980–2008, for (top) SODA3.4.2 – OBS, (middle) SODA2.2.4 – OBS, and (bottom) EN4.1.1 – OBS, showing the (left) mean and (right) RMS. Contours of the RMS variability of 0–300-m temperature are superposed on the top-right panel for comparison [contour interval (CI) = 1 K].

array at this latitude (Cunningham et al. 2007), which can now be directly compared to SODA3.4.2 monthly transport estimates (Fig. 10). The comparison shows a remarkably close correspondence at seasonal time scales (with a correlation of 0.73), with similar year-to-year variations and indications of a decline in transport by $1\text{--}2 \times 10^6 \text{ m}^3 \text{ s}^{-1}$ since 2004.

c. Arctic

The Arctic is a salinity-controlled, highly stratified ocean, with a cold and fresh shallow surface layer overlying a warmer, saltier upper layer. On the Atlantic side of the Arctic this warm/salty layer is supplied by Atlantic water flowing northward through Fram Strait

and eastward across the Barents Sea Opening. On the Pacific side of the Arctic somewhat cooler, fresher Pacific water flows north through the Bering Strait and eastward into the Canada Basin. In Fig. 11 we examine the time mean stratification at two locations on opposite sides of the Arctic that sample these water masses.

On the Atlantic side, in the Nansen Basin northeast of Svalbard, observations (Anderson and Jones 1992) and the PHC3.0 climatology both show a pronounced warm ($>1^\circ\text{C}$) and salty ($>34.6 \text{ psu}$) Atlantic water layer at depths between 200 and 500 m (Fig. 11, top). SODA3.4.2 has similar salinity stratification, but with Atlantic water temperatures that are up to 0.5°C warmer than the PHC3.0 climatology, perhaps reflecting the impact of

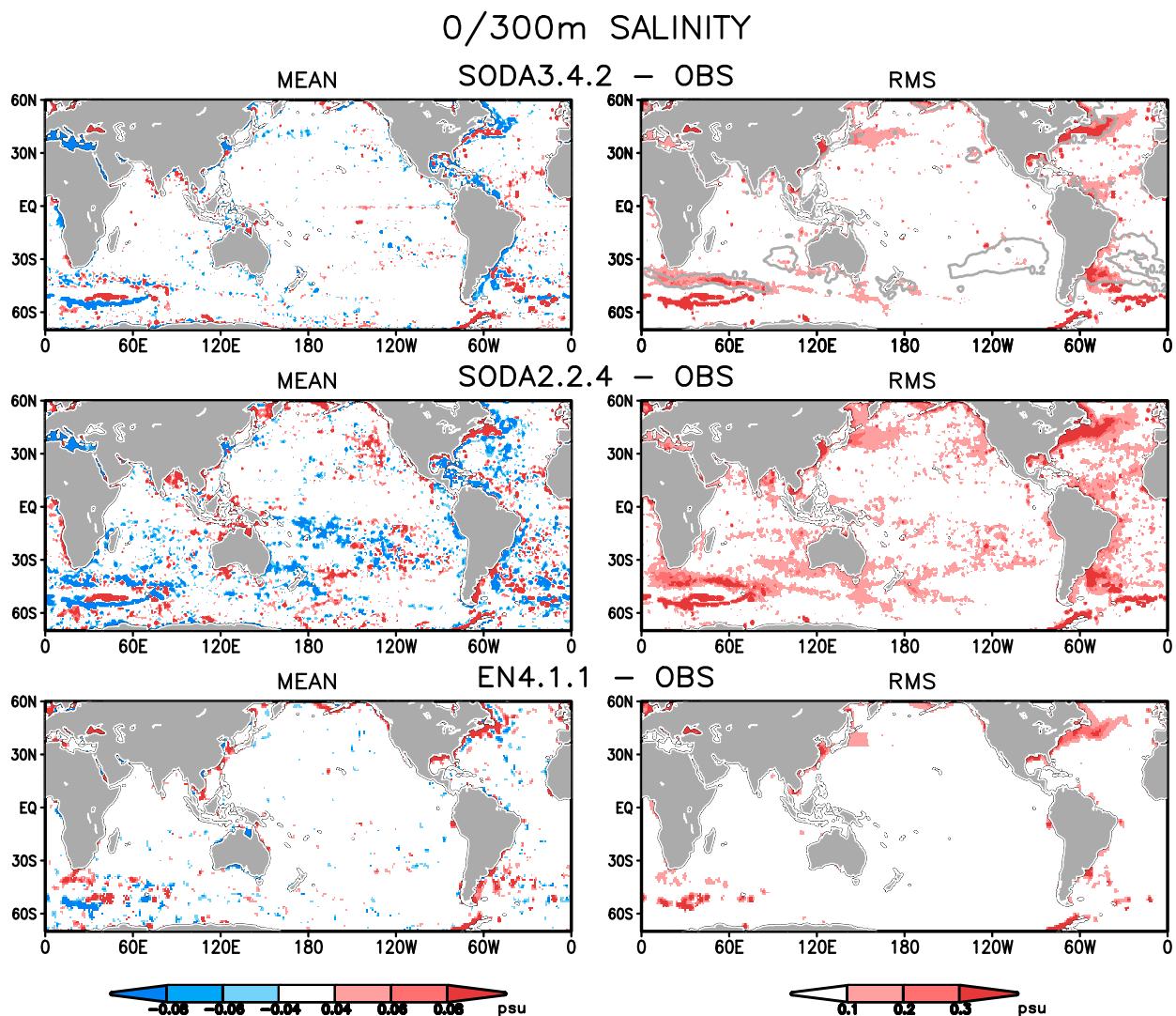


FIG. 8. Statistics of the binned analysis minus observation salinity differences $\mathbf{H}(S^a) - S^o$ (psu) averaged in time and then averaged over 0–300 m during the 29-yr period 1980–2008, for (top) SODA3.4.2 – OBS, (middle) SODA2.2.4 – OBS, and (bottom) EN4.1.1 – OBS, for (left) mean and (right) RMS. Contours of the RMS variability of 0–300-m salinity are superposed on top-right panel for comparison (CI = 0.1 psu).

the recent warming of this water mass (Polyakov et al. 2005; Beszczynska-Möller et al. 2012). In SODA2.2.4, in contrast, the Atlantic water layer rises close to the surface layer. The surface layer in PHC3.0 and SODA3.4.2 is low salinity (33 psu) and near-freezing, consistent with the presence of sea ice. SODA2.2.4, in contrast, has an erroneously warm (0°C) surface layer, much as it does in the Antarctic as well (Fig. 6).

On the Pacific side of the Arctic, east of Bering Strait and within the Beaufort Gyre, PHC3.0 and SODA3.4.2 (Fig. 11, bottom) have a deeper and cooler halocline that freshens to <30 psu at the surface. PHC3.0 also has a thermostad of cool $<-1^\circ\text{C}$ water of Pacific origin lying between 50 and 200 m, which is absent from SODA3.4.2 despite the fact that this reanalysis has reasonable $1.1 \times 10^6 \text{ m}^3 \text{s}^{-1}$

transport through the Bering Strait. Observational studies show that the upper 50 m has complex properties (Shimada et al. 2001; Shimada et al. 2005). Subzero surface temperatures are separated from the thermostad by a thin near-surface temperature maximum, the existence of which is indicated by a slight temperature peak in PHC3.0 at 75-m depth. This shallow temperature maximum is missing from SODA3.4.2. In contrast, SODA2.2.4 has near-surface properties that differ from PHC3.0 and SODA3.4.2, being both very fresh and warming dramatically toward the surface.

SEA ICE

We evaluate the performance of the SODA3.4.2 sea ice by comparing the seasonal minimum sea ice extent to

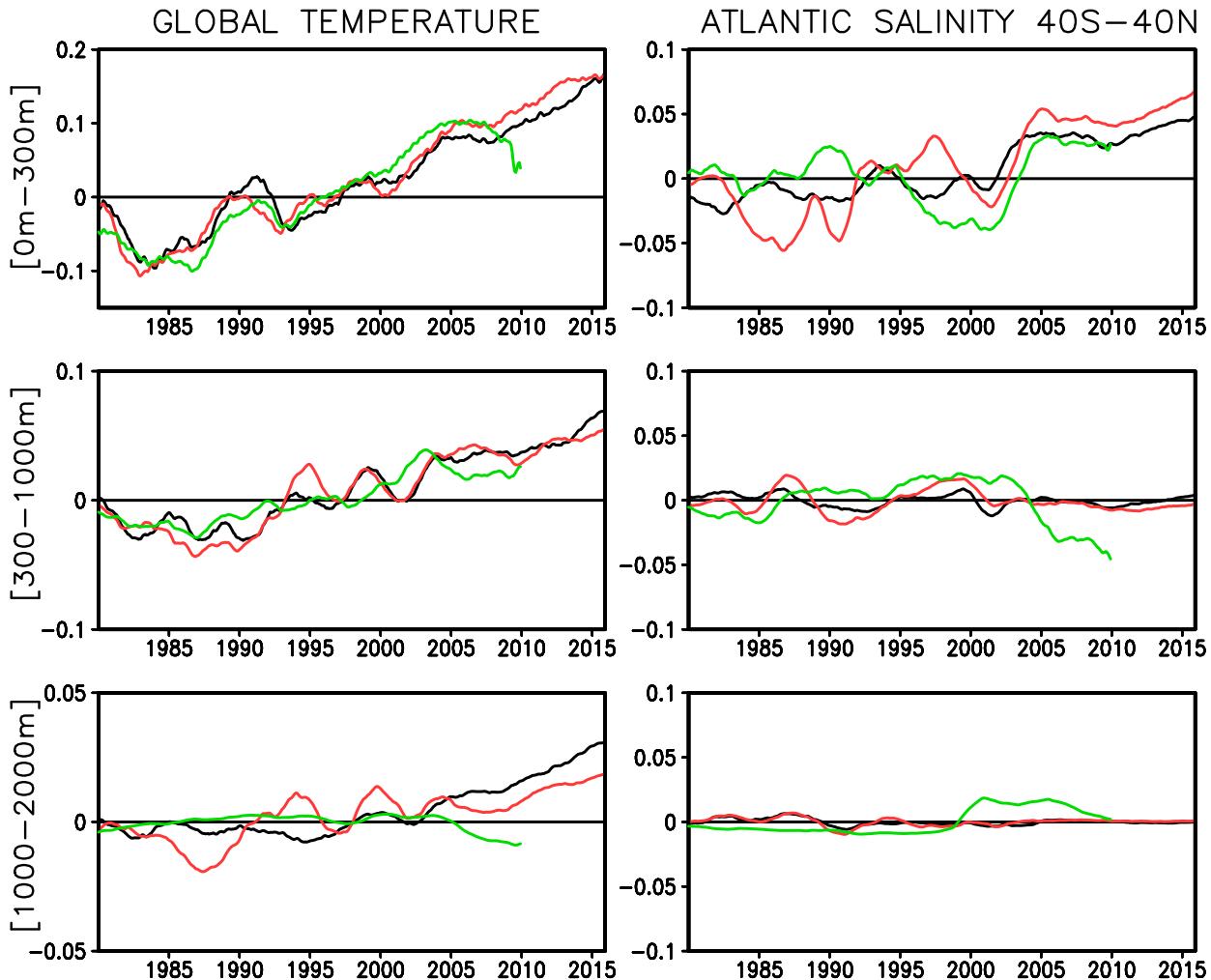


FIG. 9. Spatial average of (left) global (70°S – 60°N) potential temperature (K) and (right) Atlantic (40°S – 40°N) salinity anomalies (psu) from the 1980–2008 climatology in three depth ranges for EN4.1.1 (black), SODA3.4.2 (red), and SODA2.2.4 (green), for (top) 0–300, (middle) 300–1000, and (bottom) 1000–2000 m.

the observed record whose most dramatic observed feature is the gradual decline in seasonal minimum extent. The observations and SODA3.4.2 both show a gradual $0.7\% \text{ yr}^{-1}$ decline in extent, along with similar

year-to-year variability (Fig. 12). Studies of the causes of the two summers of extreme sea ice loss, 2007 and 2012, highlight the importance of anomalous solar insolation and wind-driven sea ice export from the Arctic

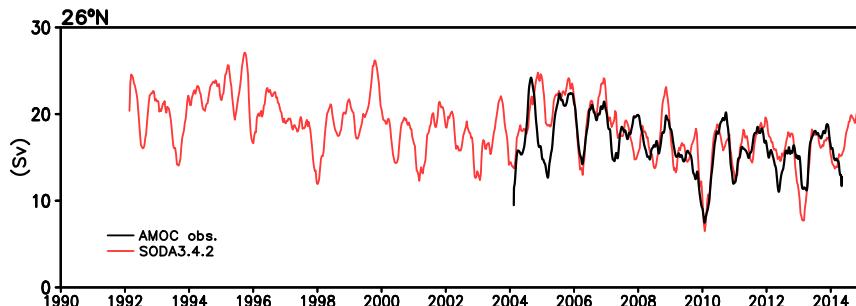


FIG. 10. Monthly average overturning circulation (Sv ; $1 \text{ Sv} \equiv 10^6 \text{ m}^3 \text{ s}^{-1}$) across 26°N in the subtropical North Atlantic. Correlation between SODA3.4.2 and observations is 0.73 with a 99% confidence interval of ± 0.02 .

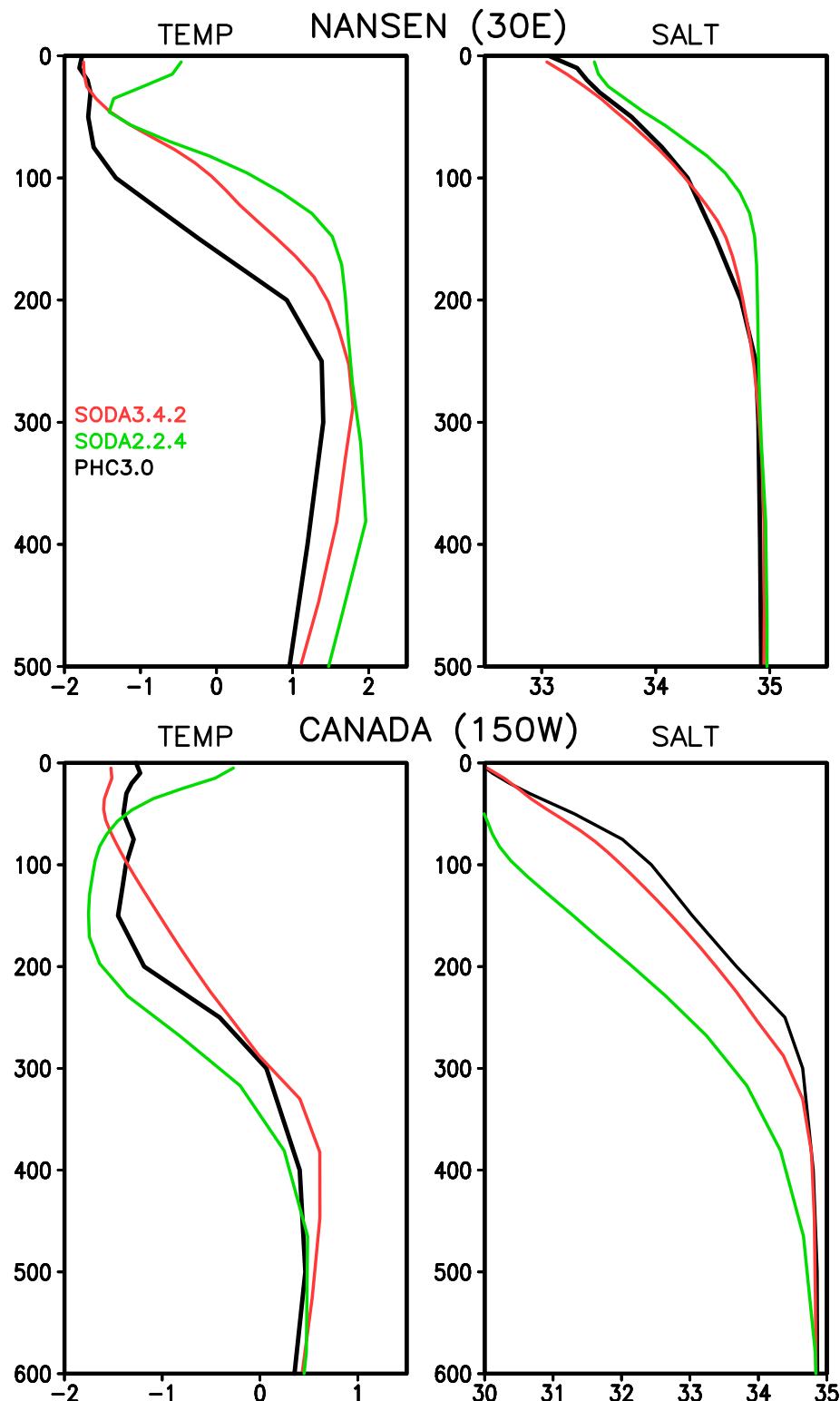


FIG. 11. Time mean stratification 2000–08 (left) temperature ($^{\circ}$ C) and salinity (psu) with depth for PHC3.0 (black), SODA3.4.2 (red), and SODA2.2.4 (green) at two locations on opposite sides of the Arctic: (top) 84 $^{\circ}$ N, 30 $^{\circ}$ E on the edge of the Nansen Basin near the Atlantic inflow to the Arctic, and (bottom) 72.5 $^{\circ}$ N, 150 $^{\circ}$ W in the Canada Basin.

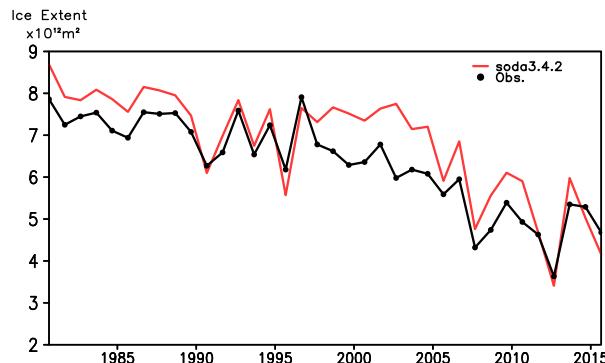


FIG. 12. SODA3.4.2 September Arctic sea ice extent (area with concentration in excess of 15%; 10^{12} m^2). Black line indicates observed sea ice extent.

(Parkinson and Comiso 2013). The appearance of these features in SODA3.4.2 may suggest that the processes controlling the anomalous melting of sea ice, including absorption of anomalous heat flux and anomalous export of sea ice, are being handled correctly. Alternatively, the good agreement in Fig. 12 could be due partly to the imprint of the observed sea ice extent on the distribution of SST observations (no observations where there is sea ice) and surface flux forcing (surface air temperature and air humidity are lower over sea ice).

d. Tropics

Finally we consider the behavior of the ocean reanalyses along the equator, focusing on the Pacific sector. The good agreement with observed temperature and salinity evident in Figs. 7 and 8 is confirmed in a comparison to the time mean profiles of temperature and salinity at two mooring locations in the central-western (165°E) and eastern (140°W) equatorial Pacific (Fig. 13). The central-western Pacific has a rain-freshened surface layer that is slightly fresher than in the ocean reanalyses. From 165°E to 140°W the observed depth of the peak speed of the observed eastward Equatorial Undercurrent shallows from 200 to 110 m and that peak speed strengthens to $>1 \text{ m s}^{-1}$. Both SODA3.4.2 and SODA2.2.4 have equatorial undercurrents that are too weak and shallow at 165°E , while they strengthen to match the observations by 140°W .

In the central-western basin the observations and SODA3.4.2 both have surface currents that vary in response to changing meteorological conditions such as those associated with El Niño, but have a weak time mean. A striking feature of the 1997/98 El Niño, for example, is the appearance of multiple westerly wind bursts in the first half of 1997 (van Oldenborgh 2000), driving eastward surface currents in the central basin that persist for much of 1997. Such strong eastward

currents are evident in SODA3.4.2 throughout 1997 but are confined to the boreal spring season in SODA2.2.4 (Fig. 14).

4. Summary and discussion

This article describes construction of version 3 of the Simple Ocean Data Assimilation (SODA3) ocean reanalysis framework and its application to studies of climate variability, which follows the previous update of SODA (Carton and Giese 2008) a decade ago. SODA3 is built around the eddy-permitting Modular Ocean Model version 5 component of the GFDL CM2.5 coupled model and includes Sea Ice Simulator (SIS) interactive sea ice at high latitudes. Other changes include updates to the datasets being assimilated and changes to the surface forcing fields. Attention is focused on reducing systematic errors with particular attention to reducing biases in surface heat and freshwater fluxes.

In the 10 years since SODA2 was released there have been many improvements to the ocean observing system, advances in assimilation methodology, and a new generation of atmospheric reanalyses. To understand the impact of each of these changes we have given up on the idea of producing a single best ocean reanalysis. Instead SODA3 is designed as a framework for a growing ensemble of ocean reanalyses that allow exploration of the impact of a variety of system parameters including the assimilated datasets, surface forcing datasets, as well as assimilation methodology.

The 37-yr-long SODA3.4.2 ocean reanalysis discussed in this paper is forced by ERA-Interim surface forcings using the COARE4 bulk formulas to compute turbulent fluxes. Other ensemble members we also make available, but are not discussed here, use surface forcings from other NOAA, NASA, and the Japan Meteorological Agency atmospheric reanalyses. SODA3.4.2 assimilates only the basic hydrographic dataset. Future ensemble members will also include sea level, surface salinity, and surface current observations. SODA3.4.2 relies on a simple but efficient implementation of optimal interpolation with modifications to account for mixed layer dynamics. Future ensemble members will also explore the improvements offered by a hybrid ensemble Kalman filter, whose implementation is described separately in Penny et al. (2015), and whose advantage is that it produces objectively estimated inhomogeneous error statistics.

Our evaluation of SODA3.4.2 is organized around comparison to a version of the earlier SODA2 (SODA2.2.4) and to the no-model EN4.1.1 statistical objective analysis during their 28-yr period of overlap 1980–2008. We begin by comparing the two ocean

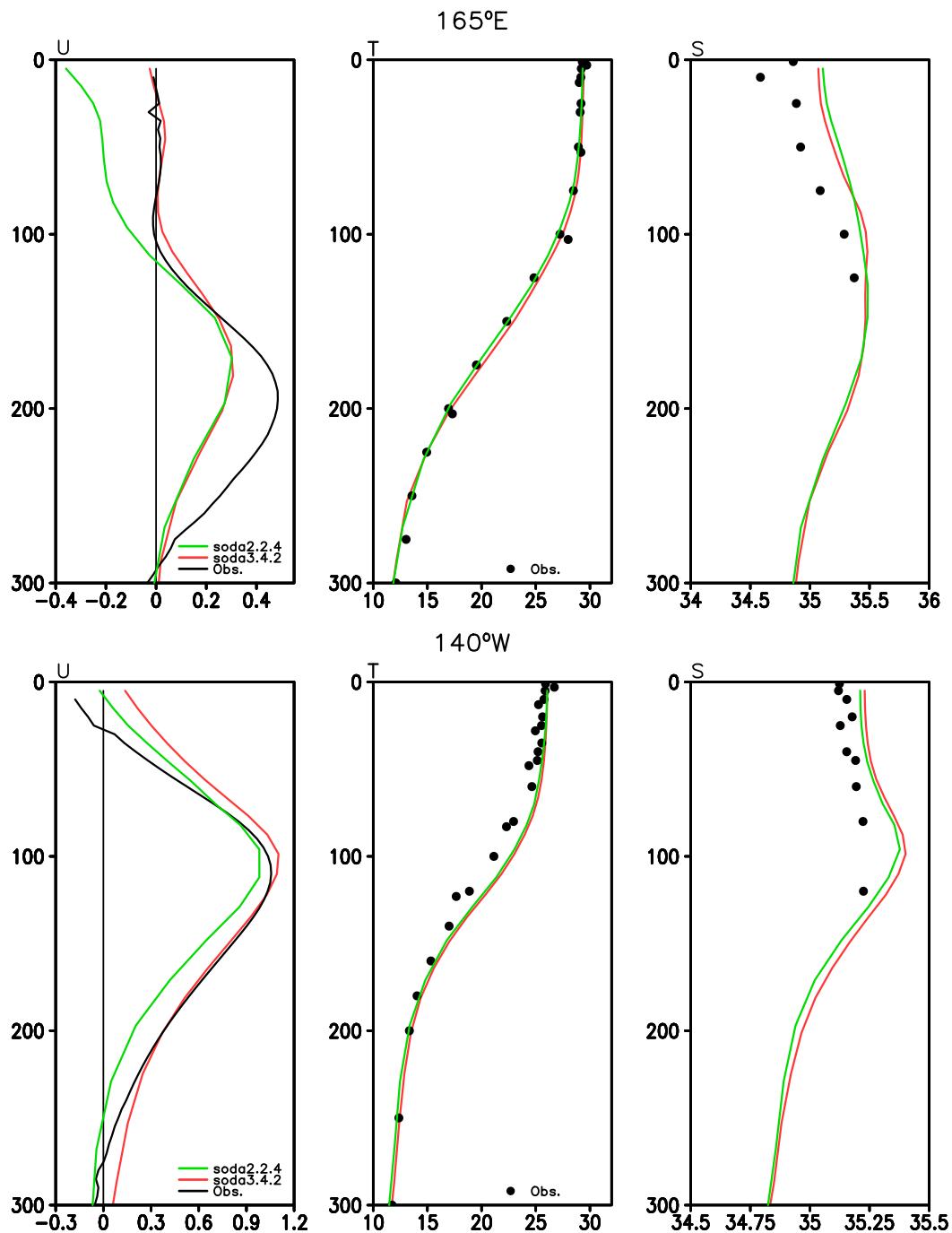


FIG. 13. Time mean vertical profiles of (left) zonal velocity (m s^{-1}), (middle) potential temperature ($^{\circ}\text{C}$), and (right) salinity (psu) with depth at two longitudes along the Pacific equator at (top) 165°E and (bottom) 140°W .

reanalyses to the Reynolds et al. (2007) OISST statistical objective analysis of nominally bulk SST. This comparison reveals that SODA2.2.4 SST is too cool and has a noise level, determined by the RMS difference from OISST, of 0.6° to 1°C . In contrast, for EN4.1.1 and SODA3.4.2 the only systematic difference from OISST

is an equatorial upwelling-related warm bias in the Atlantic and Pacific in SODA3.4.2. EN4.1.1 and SODA3.4.2 have SST noise levels of less than 0.6°C .

We next examine monthly temperature and salinity in the 0–300-m layer in comparison to the historical hydrographic profile dataset (Figs. 7 and 8). Again,

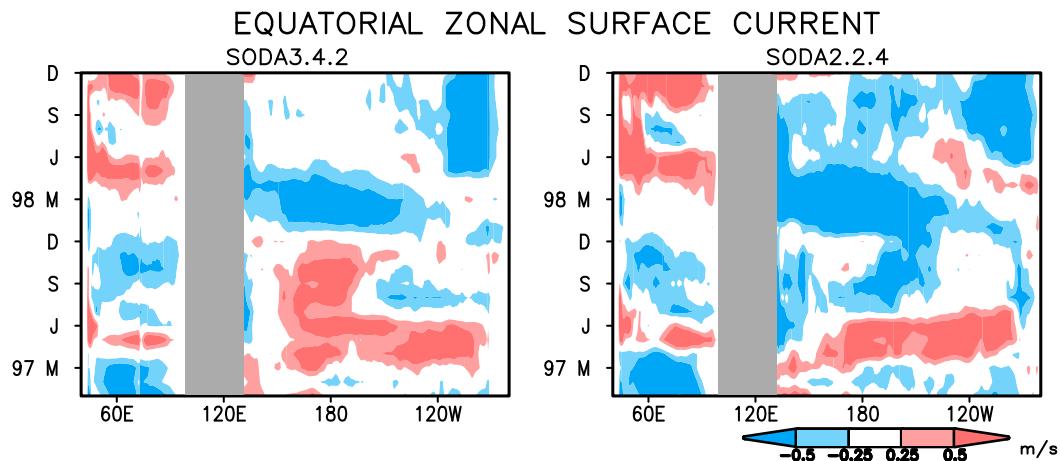


FIG. 14. Monthly average SODA3.4.2 and SODA2.2.4 zonal currents (m s^{-1}) at 5-m depth along the equator in the Indian and Pacific Oceans during the 2-yr period from January 1997 through December 1998.

SODA2.2.4 is systematically cooler than the observations with a 0–300-m temperature noise level of 0.1°C (Fig. 7). EN4.1.1 and SODA3.4.2 have little indication of systematic error and low noise levels. All three analyses show a similar rate of warming of the upper 1000 m of approximately 0.4 W m^{-2} , interrupted by 1–2 cool years following the 1991 eruption of Mount Pinatubo (Fig. 9, left). They show little evidence of a hiatus in the rate of ocean warming extending after 2004.

SODA2.2.4 0–300-m monthly salinity is systematically low in the tropics and subtropics while the other two analyses have less evidence of systematic error and lower noise levels (Fig. 8). All three analyses show that the upper layers of the Atlantic at 40°S – 40°N are becoming more saline. However, the timing of this salinification also corresponds to the introduction of the Argo salinity observations in the early 2000s, raising concerns that the apparent increase in salinity reflects the changing observing system. Finally, we show an encouragingly close correspondence between the SODA3.4.2 overturning circulation and the decade-long observed record at 26°N in the Atlantic.

In the Arctic we compare the stratification at two locations on either side of the Arctic to the PHC3.0 climatology. The correspondence between PHC3.0 and SODA3.4.2 is generally good with two exceptions. SODA3.4.2 has 0.5–1-K warmer temperatures for the layers containing water of Atlantic and Pacific origin than PHC3.0 (possibly reflecting the warming of these layers in recent years). SODA2.2.4 shows more dramatic differences, including vertical displacements of water masses and excess freshening of the Pacific side of the Arctic. The lack of sea ice in SODA2.2.4 seems to lead to excess heating of the surface layer, raising SST by as much as 2 K. In contrast, the SODA3.4.2 sea ice

extent compares well with the observed record. Finally, we examine the behavior of the reanalyses in the tropics. In the tropics SODA3.4.2 and SODA2.2.4 are similar and both are consistent with observations. The main differences we have identified are the near-surface velocity fields, which may reflect differences in specified surface momentum fluxes.

Use of ocean reanalyses for studies of decadal climate variability and change places strong constraints on accuracy and bias. While we still need to clarify the errors introduced by the remaining surface flux errors, the changing ocean observing system, model numerics, and limitations of the data assimilation schemes, we believe the improvements described here support much greater use of this new generation of ocean reanalyses even for decadal climate variability, including variability at high latitudes.

Acknowledgments. Computing resources were provided by the Computational and Information Systems Laboratory of NCAR. The complete hydrographic dataset was obtained from the NOAA/National Centers for Environmental Information World Ocean Database (www.nodc.noaa.gov). In situ SST observations were obtained from the ICOADS archive, which is hosted by the U.S. National Oceanic and Atmospheric Administration (icoads.noaa.gov/e-doc). Satellite AVHRR Pathfinder version 5.2 (PFV5.2) SSTs were obtained from the U.S. National Oceanographic Data Center and GHRSSST (pathfinder.nodc.noaa.gov). PFV5.2 is an updated version of the Pathfinder version 5.0 and 5.1 collections described in Casey et al. (2010). ACSPO MODIS SST data are provided by NOAA STAR. We strongly recommend contacting NOAA SST team led by A. Ignatov before the data are used for any publication

or presentation. The AVISO SSALTO/DUACS altimeter gridded sea level is produced and distributed by the Copernicus Marine and Environment Monitoring Service (marine.copernicus.eu). EN4.1.1 is a product of the Met Office Hadley Centre (September 2017). The Polar Science Center (PSC) Hydrographic Climatology (PHC) PHC3.0 climatology was obtained from the Applied Physics Laboratory (psc.apl.washington.edu/nonwp_projects/PHC/Data3.html). The National Snow and Ice Data Center (NSIDC) Arctic sea ice extent was calculated from the updated daily sea ice extent version 2.1 (September 2017). The TAO TRITON time series were obtained from the NOAA Ocean Climate Stations Project Office of the NOAA/PMEL (www.pmel.noaa.gov/tao/drupal/disdel/; September 2017). RAPID-MOCHA data from the RAPID-WATCH MOC monitoring project are funded by the Natural Environment Research Council and were obtained from www.rapid.ac.uk/rapidmoc (January 2016; <https://doi.org/10.5285/5acfd143-1104-7b58-e053-6c86abc0d94b>). Most of all, we are grateful to Eric Itsweire and the Physical Oceanography Program of the National Science Foundation (OCE-1233942) for providing financial support for this work. The SODA3.4.2 reanalysis datasets along with other reanalysis ensemble members are available through www.soda.umd.edu.

REFERENCES

- Anderson, L. G., and E. P. Jones, 1992: Tracing upper waters of the Nansen Basin in the Arctic Ocean. *Deep-Sea Res.*, **39A** (Suppl. 2), S425–S433, [https://doi.org/10.1016/S0198-0149\(06\)80014-2](https://doi.org/10.1016/S0198-0149(06)80014-2).
- Balmaseda, M. A., and Coauthors, 2015: The Ocean Reanalyses Intercomparison Project (ORA-IP). *J. Oper. Oceanogr.*, **8**, S80–S97, <https://doi.org/10.1080/1755876X.2015.1022329>.
- Bamber, J., M. van den Broeke, J. Ettema, J. Lenaerts, and E. Rignot, 2012: Recent large increases in freshwater fluxes from Greenland into the North Atlantic. *Geophys. Res. Lett.*, **39**, L19501, <https://doi.org/10.1029/2012GL052552>.
- Beszczynska-Möller, A., E. Fahrbach, U. Schauer, and E. Hansen, 2012: Variability in Atlantic water temperature and transport at the entrance to the Arctic Ocean, 1997–2010. *ICES J. Mar. Sci.*, **69**, 852–863, <https://doi.org/10.1093/icesjms/fss056>.
- 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, [https://doi.org/10.1175/1520-0493\(1996\)124<1256:DAUIAU>2.0.CO;2](https://doi.org/10.1175/1520-0493(1996)124<1256:DAUIAU>2.0.CO;2).
- Boyer, T. P., and Coauthors, 2013: World Ocean Database 2013. NOAA Atlas NESDIS 72, S. Levitus, Ed., 209 pp.
- , and Coauthors, 2016: Sensitivity of global upper-ocean heat content estimates to mapping methods, XBT bias corrections, and baseline climatologies. *J. Climate*, **29**, 4817–4842, <https://doi.org/10.1175/JCLI-D-15-0801.1>.
- Brunke, M. A., Z. Wang, X. B. Zeng, M. Bosilovich, and C. L. Shie, 2011: An assessment of the uncertainties in ocean surface turbulent fluxes in 11 reanalysis, satellite-derived, and combined global datasets. *J. Climate*, **24**, 5469–5493, <https://doi.org/10.1175/2011JCLI4223.1>.
- 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, <https://doi.org/10.1175/2007MWR1978.1>.
- , G. Chepurin, X. Cao, and B. S. Giese, 2000a: A Simple Ocean Data Assimilation analysis of the global upper ocean 1950–95. Part I: Methodology. *J. Phys. Oceanogr.*, **30**, 294–309, [https://doi.org/10.1175/1520-0485\(2000\)030<0294:ASODAA>2.0.CO;2](https://doi.org/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, [https://doi.org/10.1175/1520-0485\(2000\)030<0311:ASODAA>2.0.CO;2](https://doi.org/10.1175/1520-0485(2000)030<0311:ASODAA>2.0.CO;2).
- , —, L. Chen, and S. A. Grodsky, 2018: Improved global net surface heat flux. *J. Geophys. Res. Oceans*, **123**, 3144–3163, <https://doi.org/10.1002/2017JC013137>.
- Casey, K. S., T. B. Brandon, P. Cornillon, and R. Evans, 2010: The past, present and future of the AVHRR Pathfinder SST Program. *Oceanography from Space: Revisited*, V. Barale, J. F. R. Gower, and L. Alberotanza, Eds., Springer, 273–287, https://doi.org/10.1007/978-90-481-8681-5_16.
- Chen, X., and K.-K. Tung, 2014: Varying planetary heat sink led to global-warming slowdown and acceleration. *Science*, **345**, 897–903, <https://doi.org/10.1126/science.1254937>.
- Compo, G. P., and Coauthors, 2011: The Twentieth Century Reanalysis Project. *Quart. J. Roy. Meteor. Soc.*, **137**, 1–28, <https://doi.org/10.1002/qj.776>.
- Cunningham, S. A., and Coauthors, 2007: Temporal variability of the Atlantic meridional overturning circulation at 26.5°N. *Science*, **317**, 935–938, <https://doi.org/10.1126/science.1141304>.
- Dai, A., T. Qian, and K. E. Trenberth, 2009: Changes in continental freshwater discharge from 1948 to 2004. *J. Climate*, **22**, 2773–2792, <https://doi.org/10.1175/2008JCLI2592.1>.
- de Boyer Montégut, C., G. Madec, A. S. Fischer, A. Lazar, and D. Iudicone, 2004: Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. *J. Geophys. Res.*, **109**, C12003, <https://doi.org/10.1029/2004JC002378>.
- Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, <https://doi.org/10.1002/qj.828>.
- Delworth, T. L., and Coauthors, 2012: Simulated climate and climate change in the GFDL CM2.5 high-resolution coupled climate model. *J. Climate*, **25**, 2755–2781, <https://doi.org/10.1175/JCLI-D-11-00316.1>.
- Durack, P. J., and S. E. Wijffels, 2010: Fifty-year trends in global ocean salinities and their relationship to broad-scale warming. *J. Climate*, **23**, 4342–4362, <https://doi.org/10.1175/2010JCLI3377.1>.
- Dussin, R., B. Barneir, L. Brodeau, and J. M. Molines, 2016: The making of the DRAKKAR forcing set DFS5 DRAKKAR/MyOcean Report 01-04-16, 34 pp., https://www.draakkar-ocean.eu/publications/reports/report_DFS5v3_April2016.pdf.
- Fairall, C. W., E. F. Bradley, J. E. Hare, A. A. Grachev, and J. B. Edson, 2003: Bulk parameterization of air-sea fluxes: Updates and verification for the COARE algorithm. *J. Climate*, **16**, 571–591, [https://doi.org/10.1175/1520-0442\(2003\)016<0571:BPOASF>2.0.CO;2](https://doi.org/10.1175/1520-0442(2003)016<0571:BPOASF>2.0.CO;2).
- Friedman, A. R., G. Reverdin, M. Khodri, and G. Gastineau, 2017: A new record of Atlantic sea surface salinity from 1896 to 2013 reveals the signatures of climate variability and long-term

- trends. *Geophys. Res. Lett.*, **44**, 1866–1876, <https://doi.org/10.1002/2017GL072582>.
- Giese, B. S., and S. Ray, 2011: El Niño variability in simple ocean data assimilation (SODA), 1871–2008. *J. Geophys. Res.*, **116**, C02024, <https://doi.org/10.1029/2010JC006695>.
- Good, S. A., M. J. Martin, and N. A. Rayner, 2013: EN4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates. *J. Geophys. Res.*, **118**, 6704–6716, <https://doi.org/10.1002/2013JC009067>.
- Griffies, S. M., R. C. Pacanowski, and R. W. Hallberg, 2000: Spurious diapycnal mixing associated with advection in a z -coordinate ocean model. *Mon. Wea. Rev.*, **128**, 538–564, [https://doi.org/10.1175/1520-0493\(2000\)128<0538:SDMAWA>2.0.CO;2](https://doi.org/10.1175/1520-0493(2000)128<0538:SDMAWA>2.0.CO;2).
- Grodsky, S. A., J. A. Carton, and H. Liu, 2008: Comparison of bulk sea surface and mixed layer temperatures. *J. Geophys. Res.*, **113**, C10026, <https://doi.org/10.1029/2008JC004871>.
- Ignatov, A., I. Lasloe, E. D. Harrod, K. B. Kidwell, and G. P. Goodrum, 2004: Equator crossing times for NOAA, ERS and EOS sun-synchronous satellites. *Int. J. Remote Sens.*, **25**, 5255–5266, <https://doi.org/10.1080/01431160410001712981>.
- , and Coauthors, 2016: AVHRR GAC SST Reanalysis Version 1 (RAN1). *Remote Sens.*, **8**, 315, <https://doi.org/10.3390/rs8040315>.
- Janjić, T., and Coauthors, 2018: On the representation error in data assimilation. *Quart. J. Roy. Meteor. Soc.*, <https://doi.org/10.1002/qj.3130>, in press.
- Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471, [https://doi.org/10.1175/1520-0477\(1996\)077<0437:TNYRP>2.0.CO;2](https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2).
- Kosaka, Y., and S.-P. Xie, 2013: Recent global-warming hiatus tied to equatorial Pacific surface cooling. *Nature*, **501**, 403–407, <https://doi.org/10.1038/nature12534>.
- Large, W. G., and S. G. Yeager, 2009: The global climatology of an interannually varying air-sea flux data set. *Climate Dyn.*, **33**, 341–364, <https://doi.org/10.1007/s00382-008-0441-3>.
- Levitus, S., J. I. Antonov, T. Boyer, R. A. Locarnini, H. E. Garcia, and A. V. Mishonov, 2009: Global ocean heat content 1955–2008 in light of recently revealed instrumentation problems. *Geophys. Res. Lett.*, **36**, L07608, <https://doi.org/10.1029/2008GL037155>.
- Meehl, G. A., J. M. Arblaster, J. T. Fasullo, A. Hu, and K. E. Trenberth, 2011: Model-based evidence of deep-ocean heat uptake during surface-temperature hiatus periods. *Nat. Climate Change*, **1**, 360–364, <https://doi.org/10.1038/nclimate1229>.
- Oke, P. R., and P. Sakov, 2008: Representation error of oceanic observations for data assimilation. *J. Atmos. Oceanic Technol.*, **25**, 1004–1017, <https://doi.org/10.1175/2007JTECHO558.1>.
- Parkinson, C. L., and J. C. Comiso, 2013: On the 2012 record low Arctic sea ice cover: Combined impact of preconditioning and an August storm. *Geophys. Res. Lett.*, **40**, 1356–1361, <https://doi.org/10.1002/grl.50349>.
- Penny, S. G., D. W. Behringer, J. A. Carton, and E. Kalnay, 2015: A hybrid global ocean data assimilation system at NCEP. *Mon. Wea. Rev.*, **143**, 4660–4677, <https://doi.org/10.1175/MWR-D-14-00376.1>.
- Polyakov, I. V., and Coauthors, 2005: One more step toward a warmer Arctic. *Geophys. Res. Lett.*, **32**, L17605, <https://doi.org/10.1029/2005GL023740>.
- Ren, L., E. Hackert, P. Arkin, and A. J. Busalacchi, 2014: Estimating the global oceanic net freshwater flux from Argo and comparing it with satellite-based freshwater flux products. *J. Geophys. Res.*, **119**, 7869–7881, <https://doi.org/10.1002/2013JC009620>.
- Reverdin, G., F. Marin, B. Bourlès, and P. L'Herminier, 2009: XBT temperature errors during French research cruises (1999–2009). *J. Atmos. Oceanic Technol.*, **26**, 2462–2473, <https://doi.org/10.1175/2009JTECHO655.1>.
- Reynolds, R. W., 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, <https://doi.org/10.1175/2007JCLI1824.1>.
- Rhein, M., and Coauthors, 2013: Observations: Ocean. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 255–315.
- Shimada, K., E. C. Carmack, K. Hatakeyama, and T. Takizawa, 2001: Varieties of shallow temperature maximum waters in the western Canadian Basin of the Arctic Ocean. *Geophys. Res. Lett.*, **28**, 3441–3444, <https://doi.org/10.1029/2001GL013168>.
- , M. Itoh, S. Nishino, F. McLaughlin, E. Carmack, and A. Proshutinsky, 2005: Halocline structure in the Canada Basin of the Arctic Ocean. *Geophys. Res. Lett.*, **32**, L03605, <https://doi.org/10.1029/2004GL021358>.
- Steele, M., R. Morley, and W. Ermold, 2001: PHC: A global ocean hydrography with a high-quality Arctic Ocean. *J. Climate*, **14**, 2079–2087, [https://doi.org/10.1175/1520-0442\(2001\)014<2079:PAGOHW>2.0.CO;2](https://doi.org/10.1175/1520-0442(2001)014<2079:PAGOHW>2.0.CO;2).
- van Oldenborgh, G. J., 2000: What caused the onset of the 1997–98 El Niño? *Mon. Wea. Rev.*, **128**, 2601–2607, [https://doi.org/10.1175/1520-0493\(2000\)128<2601:WCTOOT>2.0.CO;2](https://doi.org/10.1175/1520-0493(2000)128<2601:WCTOOT>2.0.CO;2).
- Winton, M., 2000: A reformulated three-layer sea ice model. *J. Atmos. Oceanic Technol.*, **17**, 525–531, [https://doi.org/10.1175/1520-0426\(2000\)017<0525:ARTLSI>2.0.CO;2](https://doi.org/10.1175/1520-0426(2000)017<0525:ARTLSI>2.0.CO;2).
- Woodruff, S. D., and Coauthors, 2011: ICOADS Release 2.5: Extensions and enhancements to the surface marine meteorological archive. *Int. J. Climatol.*, **31**, 951–967, <https://doi.org/10.1002/joc.2103>.
- Zheng, Y., and B. S. Giese, 2009: Ocean heat transport in Simple Ocean Data Assimilation: Structure and mechanisms. *J. Geophys. Res.*, **114**, C11009, <https://doi.org/10.1029/2008JC005190>.