SODA4: A Mesoscale Ocean-Sea Ice Reanalysis

James A. Carton[1] and Gennady A. Chepurin[1]

[1] Department of Atmospheric and Oceanic Science
University of Maryland
College Park, MD 20742

Corresponding author: James Carton (carton@atmos.umd.edu)

Key Points:

- A global eddy resolving ocean-sea ice reanalysis is introduced
- The contribution of eddies to mass and heat transport is explored
- Modification of shallow circulation due to summer solar stratification identified

Abstract

This paper introduces the new eddy-resolving global Simple Ocean Data Assimilation version 4 (SODA4) ocean/sea ice reanalysis. As with its predecessor SODA3, sequential data assimilation is used to constrain the evolving temperature and salinity fields using World Ocean Database profiles, in-situ and satellite sea surface temperature observations, and sea ice thickness estimates as constraints. The numerical model is based on NOAA/Geophysical Fluid Dynamics Laboratory MOM5.1/SIS1 numerics with nominal l/10° horizontal resolution, and 75 levels in the vertical. Surface forcing is provided by a bias-corrected version of the European Center for Medium Range Weather Forecasts ERA5 forcing, while continental discharge is provided by a separate monthly observation-based discharge dataset. A 13-year (2010-2022) reanalysis experiment (SODA4.15.2) is compared to the coarser resolution, but otherwise similar SODA3.15.2 and to assimilated and independent observations. These comparisons show that the greatest reduction in time mean bias occurs near strong fronts, which become narrower and stronger. Eddy variability is increased, increasing eddy heat transport. Improved vertical resolution produces shallow summer temperature and salinity stratification which is most noticeable in midlatitudes and the eastern tropical Pacific and Atlantic in Northern Hemisphere winter.

Plain Language Summary

Reanalyses of the atmosphere, ocean, land, and cryosphere provide historical reference for weather and climate studies as well as initial conditions and verification for forecasting studies. This paper describes the new SODA4 eddy-resolving ocean-sea ice reanalysis of variables such as temperature, salinity, sea level, currents, and sea ice. A thirteen-year reanalysis experiment (2010-2022) is compared to a second experiment using a less accurate forecast model. The comparison highlights the changes that result from the use of the improved forecast model. These changes include reduction in mean error and improved representation of the eddy contributions to the circulation and transports.

1 Introduction

Data assimilation is most accurate when it begins with an accurate model forecast. *Griffies et al* (2015), Chang et al. (2020), *Hewitt et al.* (2020), *Jackson et al.* (2020) and others provide compelling evidence that ocean/sea ice as well as coupled forecast models improve substantially when resolution is sufficient to resolve (10-200 km) ocean eddy processes. Yet few global ocean data assimilation systems currently use such fine-resolution forecast models. Here we introduce a new eddy-resolving version of the Simple Ocean Data Assimilation (SODA) ocean/sea reanalysis, SODA4, and explore the impact of the improved forecast model by comparing the results to a reanalysis that is otherwise similar but uses a lower resolution eddy-permitting forecast model during a 12-year period of overlap 2011-2022.

Ocean reanalysis developed from efforts by meteorologists in the 1990s to use data assimilation to combine dynamics- and data-constraints, to create uniformly gridded reconstructions (reanalyses) of the evolving state of the atmosphere at time scales ranging from diurnal through decadal (*Kalnay et al*. 1996). One consequence is that these atmospheric reanalyses also provided the uniformly gridded surface heat, mass, and momentum flux boundary conditions needed to force a physical model of the ocean, which then allowed development of ocean reanalyses (e.g. *Behringer et al.,* 1998; *Carton et al.,* 2000a,b). The early ocean reanalyses suffered from a variety of errors including: i) forcing errors, ii) inaccurate initial conditions, iii) inaccurate ocean model physics, iv) limited ocean observing systems, and v) limitations in the data assimilation procedures themselves. Recent decades have seen reductions of many of these error sources. High quality ocean observing systems have been deployed, such as surface moorings and Argo profilers (*Roemmich et al*., 2019), while calibration projects such as the International Quality-controlled Ocean Database (e.g. *Boyer et al*., 2016; *Cowley et al.,* 2021) and the Group for High Resolution Sea Surface Temperature (*GHRSST Project Office*, 2012) have been correcting biases in historical observations. Data assimilation algorithms are also undergoing rapid development, (*Storto et al.,* 2019), importantly now including adjustments to the meteorological fluxes based on the ocean observations (e.g. *Stammer,* 2002; *Carton et al*., 2018b). Here we explore the impact of an improved forecast model that is better able to resolve the oceanic mesoscale.

The oceanic mesoscale contains steady and transient baroclinic eddies whose horizontal scales range from 230 km in the tropics to 10 km or less at high latitudes (*Eden*, 2007). Because of correlations among variables modeling studies suggest that transient eddies contribute at least 10% of the net poleward heat transport in both the deep tropics and across the subtropical/midlatitude gyre boundaries (*Hecht and Hasumi*, 2008; Volkov et al., 2008; *Griffies et al.,* 2015). In the deep tropics an active mesoscale allows poleward eddy heat fluxes to regulate the temperature of the tongue of cool surface water in the eastern Pacific and Atlantic. Across the gyre boundaries eddies can transport heat down-gradient, warming the higher latitudes and this reducing available potential energy. Stationary eddies can also play important roles in regulating flow past topographic features such as coastal irregularities.

SODA4 is the fourth generation of the SODA reanalysis system, which began with the goal of providing a uniformly gridded historical reference for the evolving physical state of the ocean from monthly to decadal timescales. In SODA4 the latest improvement is the use of a forecast model with resolution increased by a factor of eight and with an improved topographic mapping. The impact of the improved eddy-resolving forecast model was first explored at polar latitudes in the Regional Arctic Ocean/sea ice Reanalysis (RARE1, *Carton et al*., 2023). RARE1 showed many changes to the representation of temperature, salinity, and currents in the Arctic Ocean and neighboring seas relative to a coarser resolution reanalysis or to a data-driven objective analysis. SODA4 extends the developments introduced in RARE1 to the global ocean. Thus, while SODA4 also includes a complete representation of the polar latitudes (similar to RARE1) here we focus on changes observed in the ice-free ocean 60°S-60°N. In addition to presenting summary statistics we highlight examples to illustrate some climate-relevant applications.

This paper is organized as follows. Section 2 describes construction of SODA4: the model, data sets, and data assimilation, the thirteen-year long SODA4.15.2 experiment and the SODA3.15.2 control experiment. Section 3 begins by examining systematic error and then explores the impacts of improved horizontal resolution, which leads to more vigorous eddy production and vertical resolution, which alters near surface stratification.

2 Materials, Methods, and Data

This section describes the data, model, forcing, and data assimilation used in the construction of SODA4 (summarized in **Table 1**).

Table 1 Summary of parameters for SODA3.15.2 and SODA4.15.2 (OI is optimum interpolation, KF is Kalman filter, WOD is World Ocean Database).

Name	Model numerics	Resolution (XxYxZ)	Surface forcing	Datasets	Assimilation	Time Period
SODA3.15.2[1]	MOM5.1 SIS1	0.25°x0.25°x50L	ERA5	WOD18 T/S, COADS3.0, AVHRR SST, GIOMAS	OI	1980-2022
SODA4.15.2	MOM5.1 SIS1	0.1°x0.1°x75L	ERA5	WOD18 T/S, COADS3.0, AVHRR SST, GIOMAS	OI	2010-2022

[1]*Carton et al. (2018b)*

2.1 Model

The global ocean/sea ice model, which is similar to the ocean–sea ice component of the GFDL CM2.6 coupled model (*Winton*, 2014; *Griffies et al.,* 2015), uses MOM5.1/SIS1 numerics (*Winton,* 2000; *Griffies*, 2012). Velocity advection uses a third order Adams–Bashforth scheme on an Arakawa B-grid, while tracer advection uses a third order upwind biased scheme (*Hundsdorfer and Trompert*, 1994) with a predictor-corrector time-filter for sea level (*Griffies* 2004). Vertical turbulent viscosity varies from 1×10^{-4} to 2.5×10^{-3} m^2s^{-1}, while vertical diffusivity varies from 1×10^{-5} to 5×10^{-3} m^2s^{-1}. Additional vertical mixing is added to simulate the impact of tidal mixing following *Lee et al.* (2006).

The ocean model has 3600x2700 eddy resolving quasi-isotropic horizontal grid cells that vary in size from 11 km at the equator to 5.5 km at ±60° (the model resolution required to be considered 'eddy resolving' is discussed in *Hallberg*, 2013). Further northward the north polar cap splits into two geographically displaced poles located on the Eurasian and North American continents. Bottom topography is interpolated from the 30 arcsecond GEBCO 2014 topography (https://www.gebco.net/) with modifications to eliminate orphan points and bays spanned by a single grid point, and to impose a minimum depth of 10m. In the vertical the model has 75 z* vertical levels that expand from fine ~1.1 m resolution near surface to coarse 200m resolution in the deep ocean. Partial bottom cells are used to improve fidelity of the model topography. Monthly continental discharge is provided by a compilation of station gauge estimates including Greenland discharge (*Bamber et al.,* 2018), with the combined dataset adjusted to match estimates of global annual evaporation rates. An examination of the global water budget suggests that our total continental discharge is similar to the 1.3×10^6 m^3/s estimate of *Schmitt* (2008). However, discharge from some of the major river systems such as the Amazon, Orinoco, and Mississippi are too large. This error is apparent in the low salinity of fresh pools in front of these rivers. Gridded Global Ice-Ocean Modeling and Assimilation System sea ice thickness estimates of *Zhang* et al. (2003) are used to constrain sea ice. Relying on PIOMAS/GIOMAS gridded thickness estimates rather than assimilating raw satellite observations allows us to leverage the specialty knowledge of sea ice dynamics and observations of the sea ice group. *Schweiger et al.* (2011) estimate the error in Arctic thickness to be less than 10 cm.

Sea ice has horizontal resolution matching the ocean model and has five snow and ice categories. 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, which was selected in order to reduce the rate of summer sea ice melt. Initial conditions were interpolated from SODA3.15.2 on 29 December, 2009.

2.2 Surface Forcing

Surface forcing is derived from the ERA5 reanalysis of *Hersbach, et al.* (2020) by combining three-hourly estimates of downwelling short and longwave radiative fluxes, six-hourly estimates of neutral winds at 10m height, 2m air temperature and humidity, sea level pressure, and daily liquid and solid precipitation. The '15' in the name SODA4.15.2 refers to the use of this ERA5 forcing. Variables are converted to net thermodynamic and radiative fluxes within the GFDL Flexible Modelling System coupler, which has been modified to implement the Coupled Ocean-Atmosphere Response Experiment (*Fairall et al. 2003*) bulk formulas. The '2' in the name SODA4.15.2 refers to the use of the COARE bulk formulas.

If the ERA5-derived fluxes were directly applied to the ocean, they would produce systematic errors in the ocean and sea ice. We reduce those systematic errors by carrying out an initial experiment to estimate the analysis increments (discussed below), and then use the incremental near-surface heat and freshwater budgets to estimate seasonal corrections to surface heat and freshwater fluxes following *Carton et al.* (2018b). These corrections are applied to modify surface fluxes and thus to reduce the observation-model misfit in the experiments presented here. Because of this 'predictor-corrector' approach each day's analysis is influenced by past, present, and future observations and model forecasts.

2.3 Constraining Data

The main data sets that SODA ingests are the World Ocean Database of historical hydrographic profiles (*Boyer et al.,* 2018) with updates through year 2022, (**Table 1**). During 2010-2022 this dataset consists of more than 6.6 million profiles of which 2.1 million (or 13,400/month) come from Argo drifting profilers and another 2.8 million from ocean gliders. Collated L3 remotely sensed nighttime Infrared SST observations from the NOAA Center for Satellite Applications and Research (*Jonasson et al., 2022*) span the full period of interest, with additional calibrating in situ observations obtained from International Comprehensive Ocean–Atmosphere Data Set (ICOADS) release 3.0 SST database (*Freeman et al., 2016*). Sea level observations are withheld for independent comparison.

2.4 Data assimilation

SODA 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 $\mathbf{H}(\omega^f)$:

$$\omega^a = \omega^f + \mathbf{K}[\omega^o - \mathbf{H}(\omega^f)] \quad (1)$$

where the gain matrix **K** determining the impact of the observations, depends on the observation error covariance $\mathbf{R}^o \equiv \langle \boldsymbol{\varepsilon}^o \boldsymbol{\varepsilon}^{oT} \rangle$ and the model forecast error covariance $\mathbf{P}^\mathbf{f} \equiv \langle \boldsymbol{\varepsilon}^f \boldsymbol{\varepsilon}^{fT} \rangle$. A direct implementation of (1) would introduce shocks and spurious waves. To avoid this, 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 available data and the timescale of ocean variability. The form of $\mathbf{K} = \mathbf{P}^f \mathbf{H}^T \left(\mathbf{H}\mathbf{P}^f \mathbf{H}^T + \mathbf{R}^o\right)^{-1}$ is the consequence of minimizing the expected variance of the analysis error subject to some simplifying assumptions, including the assumptions that the model forecast, observation, and analysis errors are unbiased. Observation errors are also assumed to be unbiased and uncorrelated and as a result $\mathbf{R}^o$ is a diagonal matrix. The *analysis increments* $\mathbf{K}[\omega^o - \mathbf{H}\omega^f]$, which are the gridded corrections to the forecast at each assimilation cycle, are also saved and are used in *Section 3* to evaluate the degree of misfit between the forecasts and the observations. For example, negative time mean temperature analysis increments imply that the model forecasts are biased warm while positive temperature analysis increments imply that the model forecasts are biased cold.

2.5 Comparison datasets

We compare SODA4.15.2 to SODA3.15.2 during a twelve-year period of overlap, 2011-2022, skipping the first year of SODA4.15.2 (**Table 1** second row). SODA3.15.2 shares similar model numerics, forcing, data assimilation, and constraining datasets, but has coarser 0.25°x0.25° horizontal and vertical resolution (SODA3.15.2 has 13 fewer levels the top 50m and 9 fewer in the lowest 4.5 km) and an older, pre-satellite, bottom topography map. To evaluate surface eddy energy levels at 5-dy timescale we compare SODA4.15.2 and SODA3.15.2, to observed sea level, which is not assimilated in either reanalysis. The observed sea level is the 0.25°x0.25° Copernicus Marine Service multi-satellite Data Unification and Altimeter Combination System sea level (*Taburet et al*., 2019).

3 Results

We begin with an examination of the SODA3.15.2 and SODA4.15.2 analysis increments produced by the data assimilation system, and with a comparison to observed sea level, which is an independent observation set. Next, we examine differences in seasonal near surface stratification resulting from improved near surface resolution. Finally, we compare volume and heat transports.

3.1 Comparisons to Observations

The time mean temperature and salinity analysis increments integrated 0-300m, are converted to units of net surface heat (W/m^2) and freshwater (mm/10dy) fluxes in order to make it easier to understand the surface flux errors that would be required to produce these increments (**Fig. 1**). The largest time mean temperature analysis increments occur close to the equator and are similar among SODA3.15.2 and SODA4.15.2. These large analysis increments are not the result of surface heat flux errors, but rather can be traced to a ~4% difference between the zonal wind stress imposed on the ocean models and the value that the ocean models expect based on the observed thermal structure. The consequence of this difference is that the forecast model's equatorial thermocline has a slightly weaker eastward tilt than actually observed. The data assimilation is continually acting to strengthen this tilt by cooling the eastern equatorial Pacific and warming the western equatorial Pacific.

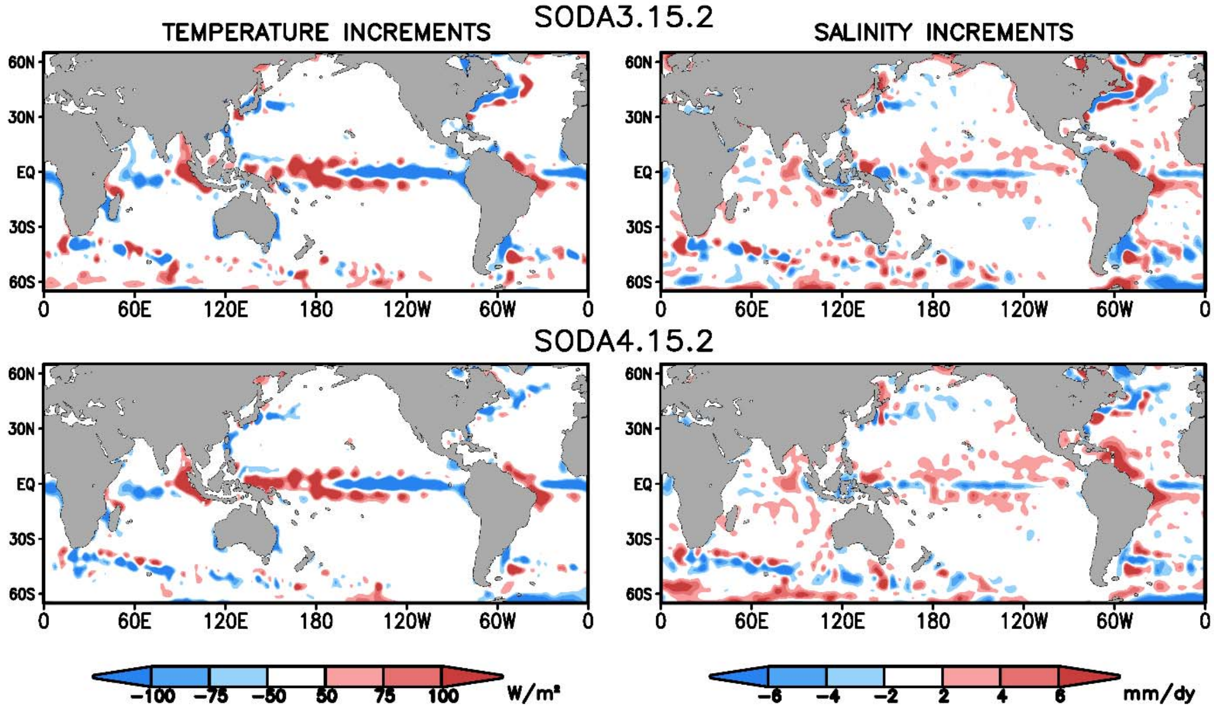

Figure 1 (mean_global_increments.pdf) Time mean (2011-2022) (left) temperature and (right) salinity analysis increments averaged 0-300m. (upper) SODA3.15.2, (lower) SODA3.4.2.

Next, we note that western boundary currents such as the Gulf Stream and Kuroshio show large time mean temperature analysis increments in SODA3.15.2 with negative values near the coast and positive values further to the east (**Fig. 1**). The Gulf Stream region is shown in detail in **Fig. 2**. Negative values near the coast in SODA3.15.2 show that the data assimilation is acting to cool the already cold coastal water, while positive values show the data assimilation is acting to warm the already warm subtropical gyre (blue and red contours, **left Fig. 2**). Thus, the data assimilation is acting to intensify and steepen the Gulf Stream temperature front and thus to narrow and strengthen the Gulf Stream current. The small time-mean temperature analysis increments in SODA4.15.2 show that this forecast model already produces a Gulf Stream front that is narrow and steep and a current that is as strong as observed. The monthly variability of the analysis increments is similar in the two models despite the fact that the model used in SODA4.15.2 is so much more variable (Supporting Information **Fig. S1**)

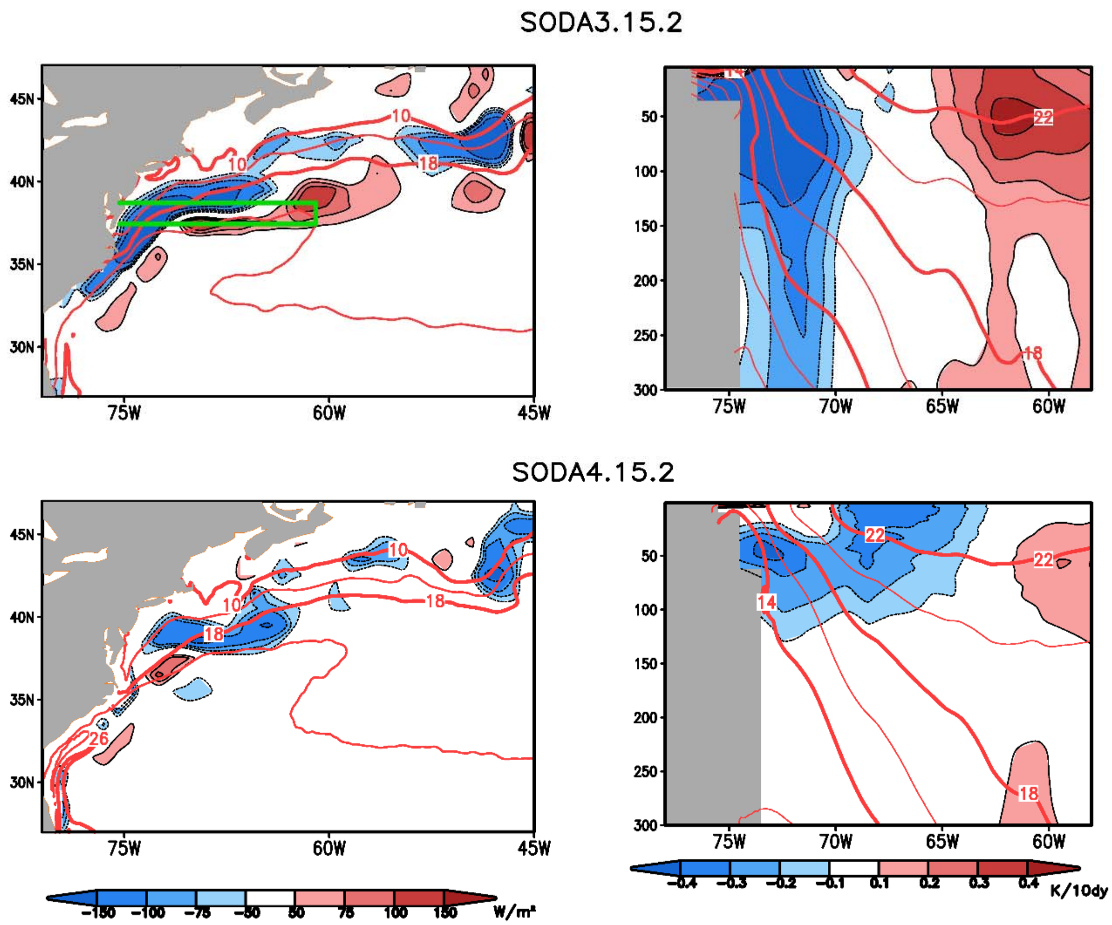

Figure 2 (increment-map-2) (colors) (left) Time mean temperature analysis increments in the upper 300m of the western North Atlantic with latitude and longitude. Contours show time mean temperature averaged 0-100m. (right) Time mean temperature analysis increments with depth in the latitude band 37-38°N. Again, contours show time mean temperature. (upper) SODA3.15.2, (lower) SODA4.15.2.

We next compare surface eddy potential energy in SODA3.15.2 and SODA4.15.2 to observed surface eddy potential energy (**Fig. 3**). The comparison shows that the locations of peaks in eddy potential energy such as the Northwest Atlantic are similar in SODA3.15.2 and SODA4.15.2. But the peak values of SODA4.15.2 eddy potential energy are roughly 50% higher and the region of high eddy potential energy is more extensive, both more closely resembling the observations (lower panels, **Fig. 3**).

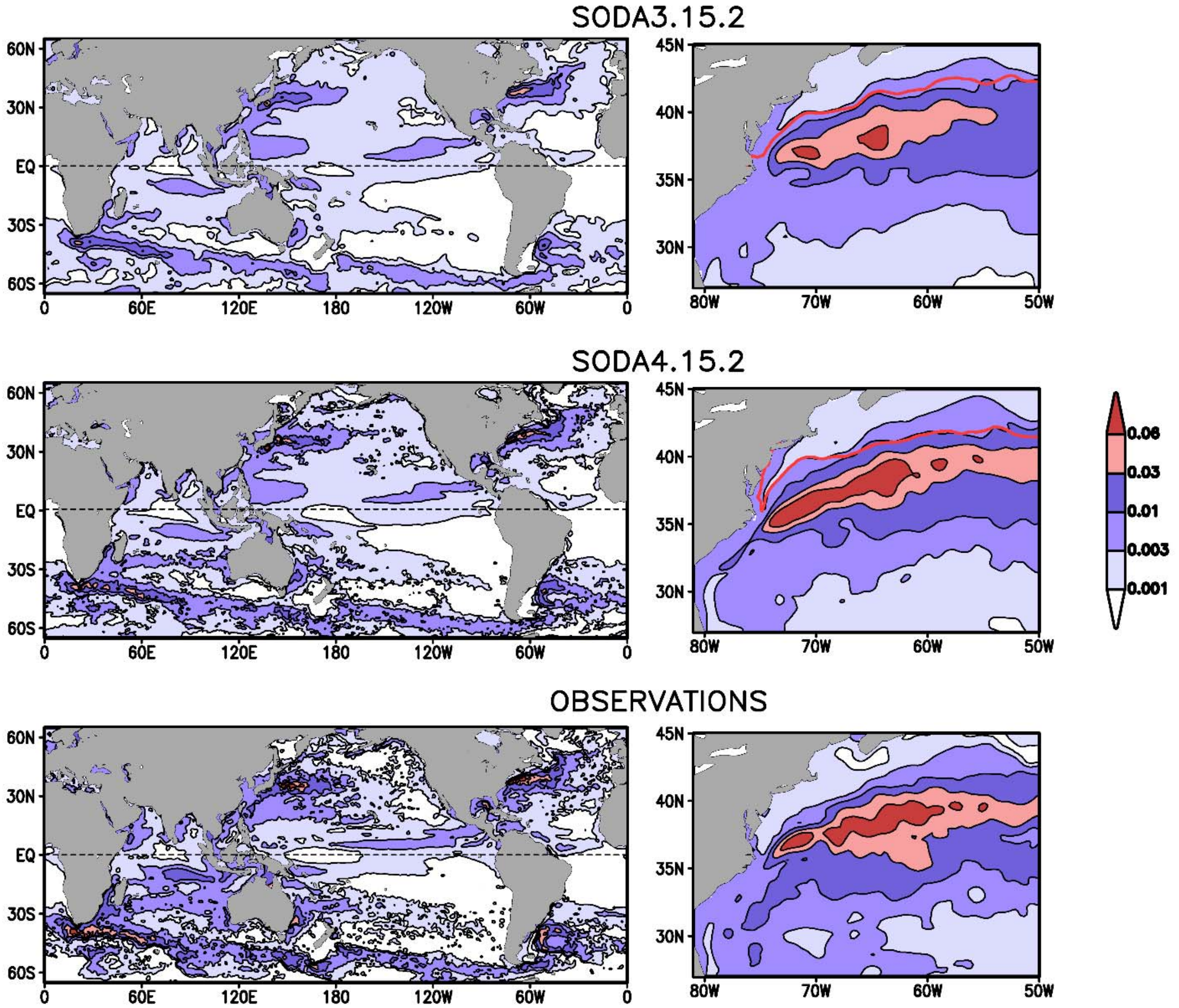

Figure 3 (epe-compare) Time mean (2011-2022) surface eddy potential energy ($\overline{(\eta')^2/2}$). (left) Global ocean, (right) northwest Atlantic. (upper) SODA3.15.2, (middle) SODA3.4.2, and (bottom) observations. Units are m^2/s^2.

3.2 Near surface stratification

The extra vertical resolution used in SODA4.15.2 alters the near-surface penetration of solar radiation and wind-driven turbulence. As a result, SODA4.15.2 has 2-4°C near surface temperature and -0.1 to -0.2psu salinity stratification (measured between the surface and 30m depth in midlatitudes) in summer (**Figs. 4** and **5**). This increase in near-surface temperature and salinity stratification increases near surface density stratification by ~ -0.4 kg/m^3. Much of this additional buoyancy stratification in is sufficiently close to the surface so that it cannot easily be sampled by in situ profiling instruments such as ARGO and may have a diurnal component (**Fig. S3** compares temperature analysis increments at 5m and 30m depth). Particularly large near surface changes in buoyancy also occur along the continental shelves of eastern North Asia and

North America. In the eastern equatorial Pacific the northern hemisphere winter vertical temperature stratification increases by approximately 1°C.

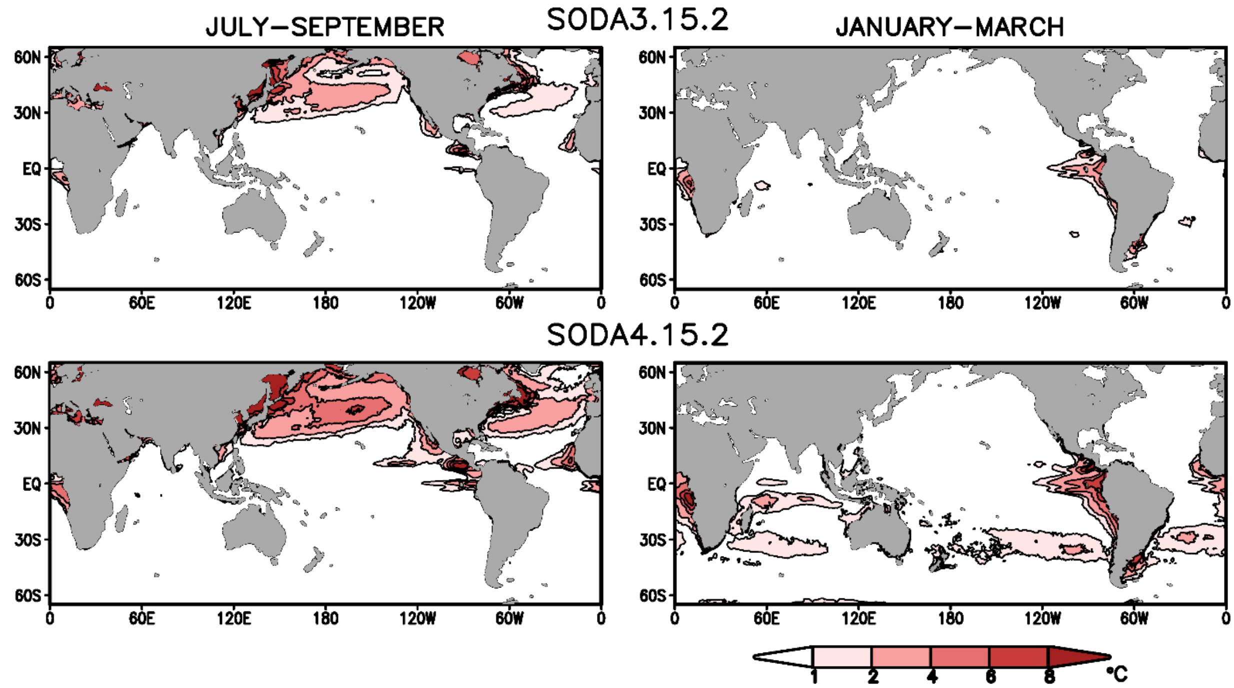

Figure 4 (t0-t30) Climatological seasonal 0-30m temperature difference. (left) July-September, (right) January-March.

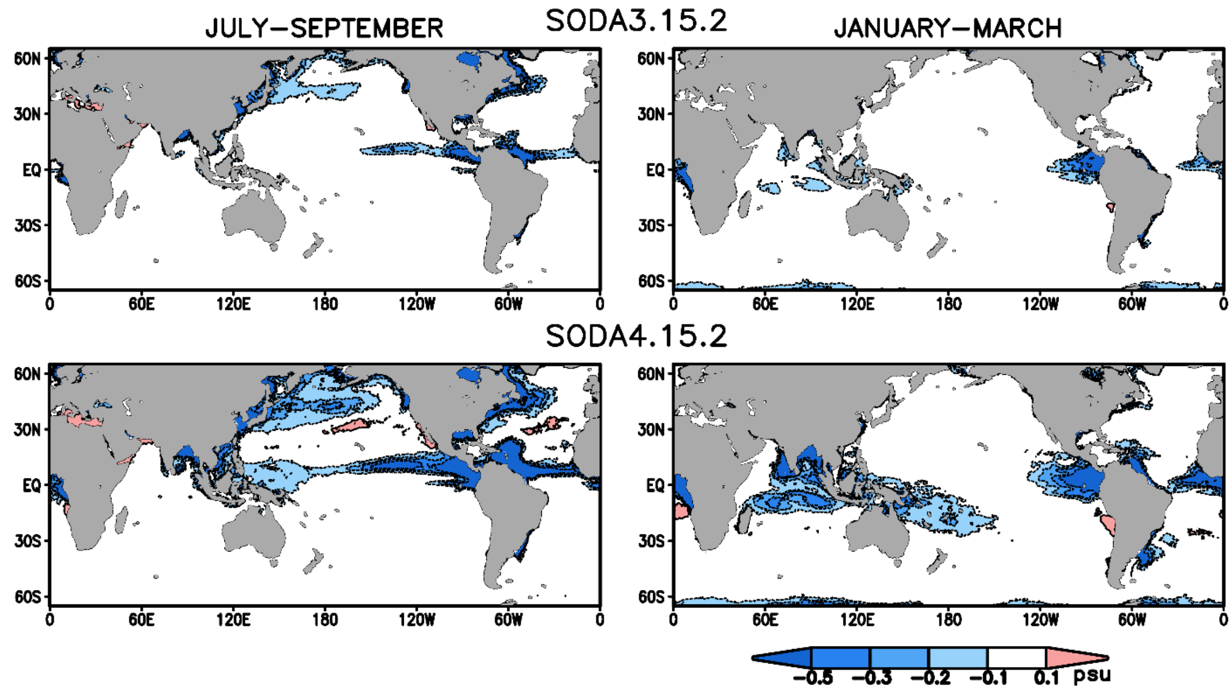

Figure 5 (s0-30) Climatological seasonal 0-30m salinity difference. (left) July-September, (right) January-March.

On the continental shelf of the western North Atlantic previous studies (e.g. *Flagg et al., 2006*; *Lentz*, 2008) have documented a variety of processes influencing along- and cross-shelf exchanges, including seasonal and storm-driven wind drift, solar heating, and interactions with Gulf Stream rings. In **Fig. 6** we show that both SODA3.15.2 and SODA4.15.2 experience a 10°C summer near surface temperature stratification that isolates the surface from the water below 30m. This stratification confines the wind driven cross-shelf exchange to a shallow surface layer and allows deeper water to flow southward along shelf break isobaths. But while SODA3.15.2 shows cool < 8°C water at the shelf break SODA4.15.2 shows even cooler < 6°C water as well as a more extensive southward flow along the shelf break.

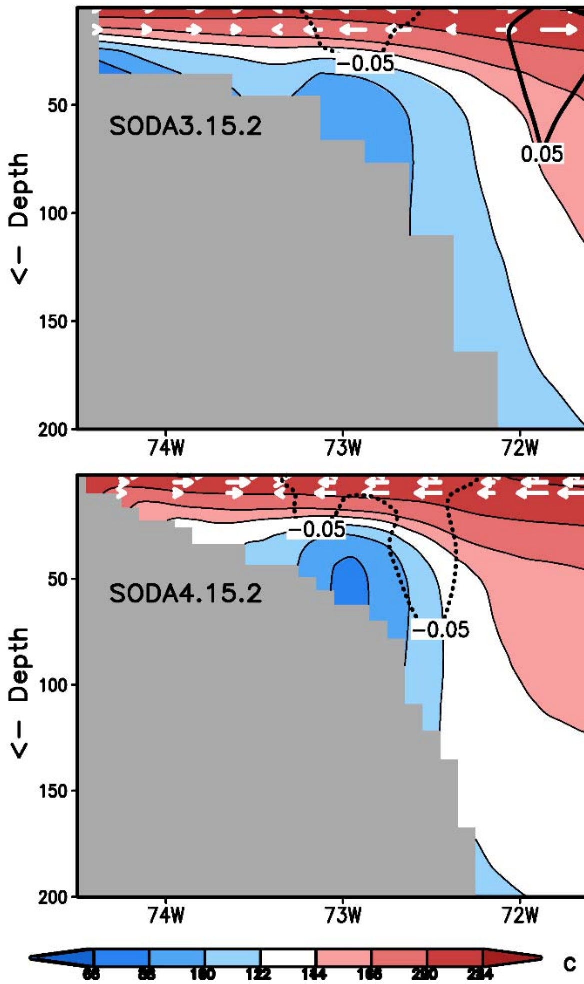

Fig. 6 (39.33-section.pdf) (colors) Summer (JAS) temperature with depth and longitude along 39.33N (averaged 2011-2021). (White arrows) Zonal velocity in the upper 10m. (Black contours) Meridional current.

3.3 Volume Transport

We begin by examining time-mean volume transports. Transports through 1) Drake Passage in the Southern Ocean, measuring the strength of the Circumpolar Current; 2) the Indonesian Throughflow, measuring the Western tropical Pacific to eastern tropical Indian Ocean exchange; and 3) Bering Strait, measuring the North Pacific Water to Arctic Ocean transport agree to within 5% (**Table 2**). Indonesion throughflow transport is carried through three separate straits: Lombok, Ombai, and Timor. In SODA3.15.2 the transport is shared equally among the three while in SODA4.15.2, as in the observations of *Susanto et al.* (2016), more than half of the transport is carried through Timor Strait alone (Supporting Information **Fig. S2**). This difference in mean transport alters the magnitude and temperature of the seasonal and interannual fluctuations.

Table 2 Mean volume transport estimates (in units of 10^6 m^3/s)

Passage	SODA3.15.2	SODA4.15.2	Observations
Drake Passage	160 Eastward	159 Eastward	~157[1]
Indonesian Throughflow	15.4 Westward	14.7 Westward	15[2]
Bering Strait	0.9 Northward	0.9 Northward	0.8[3]

[1]Xu et al (2020); [2]Susanto et al. (2016); [3]Woodgate et al. (2005)

The Atlantic Ocean has instrumented transport arrays across the 59°N, 26.5°N, and 34°S meridians (*Lozier et al.,* 2019; *Frajka-Williams et al.,* 2019; *Kersalé et al*., 2020), spanning the subpolar and subtropical gyres (Supporting Information, **Table S1**). Across 59°N in the section east of Greenland *Lozier et al.* report time mean northward overturning transport in the upper ocean (depths shallower than the 27.66 sigma potential density surface) of 15.6x10^6 m^3/s. For SODA3.15.2 and SODA4.15.2 the corresponding numbers are slightly lower: 14.1x10^6 and 13.8x10^6 m^3/s, but the differences may not be significant. Across the 26.5°N meridian *Frajka-Williams et al*. report an average northward overturning transport above 1000m depth of 17±4x10^6 m^3/s for the period 2004-2017. For this meridian SODA3.15.2 also slightly lower overturning transport of 14.7x10^6 m^3/s while SODA4.15.2 has somewhat higher transport of 18.6x10^6 m^3/s due to a more vigorous western boundary current system.

3.4 Poleward Heat Transport

The oceans absorb heat in the tropics and release that heat poleward of ±20-30° latitude. Between those two latitudes heat is transported by poleward movement of warmer water balanced by equatorward movement of cooler water. *Trenberth, et al*. (2019) asserts that the most accurate way to quantify this transport is through ocean reanalyses such as presented here. In **Fig. 7** we compare heat transport estimates for SODA3.15.2 and SODA4.15.2 in each ocean basin. In the Atlantic SODA3.15.2 and SODA4.15.2 have a maximum northward transport of approximately 1x10^{15} W. However, the two differ in the fact that the latitude of that maximum shifts poleward from 15-20°N in SODA3.15.2 to 20-30°N in SODA4.15.2.

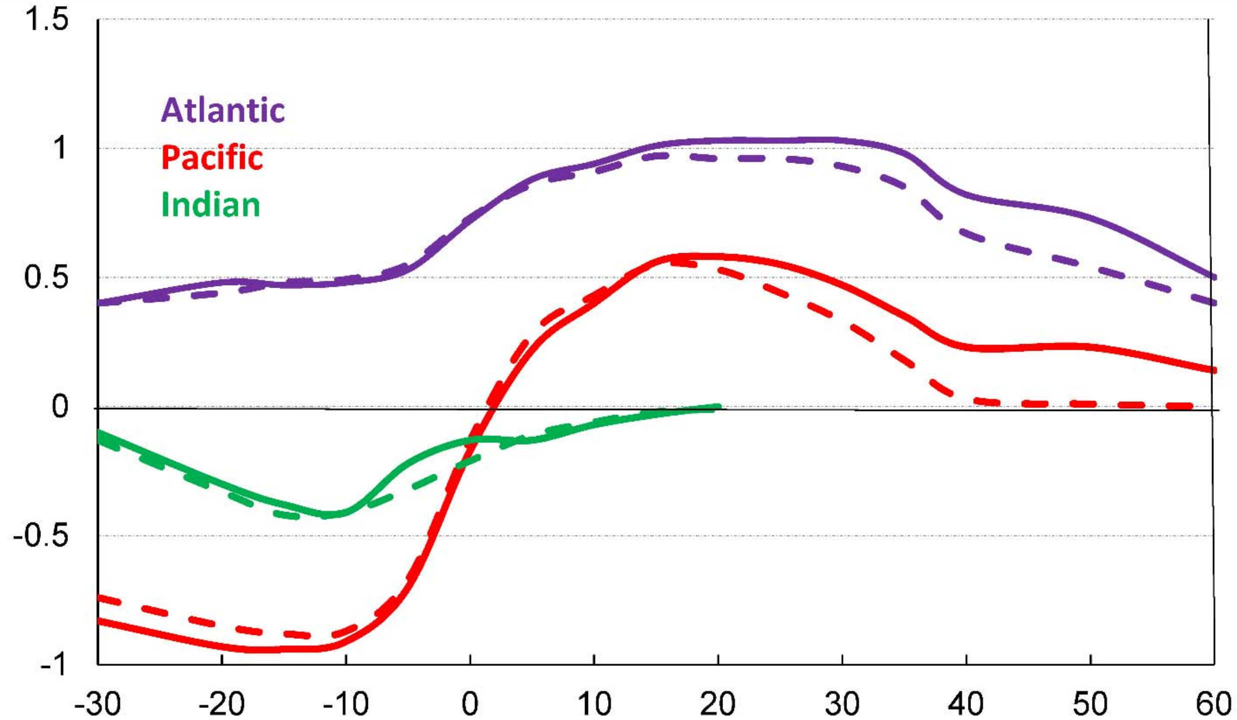

Fig. 7 Mean northward heat transport in the Atlantic, Pacific, and Indian basins with latitude. (dashed) SODA3.15.2, (solid) SODA4.15.2. Units are 10^{15} W.

The cause of this shift is likely due to the increasing contribution of eddy heat transport between 20-30°N in SODA4.15.2 (**Fig. 8**). In the South Atlantic heat transport is toward the equator in both SODA3.15.2 and SODA4.15.2. In the Pacific SODA3.15.2 and SODA4.15.2 also agree in estimating the maximum northward transport to be 0.6×10^{15} W and the minimum to be -0.9×10^{15}W. But as in the case of the Atlantic the latitude of the northward maximum shifts poleward from 15-20°N for SODA3.15.2 to 15-25°N for SODA4.15.2. In the Indian Ocean heat transport is southward, reaching a maximum of -0.4×10^{15} W.

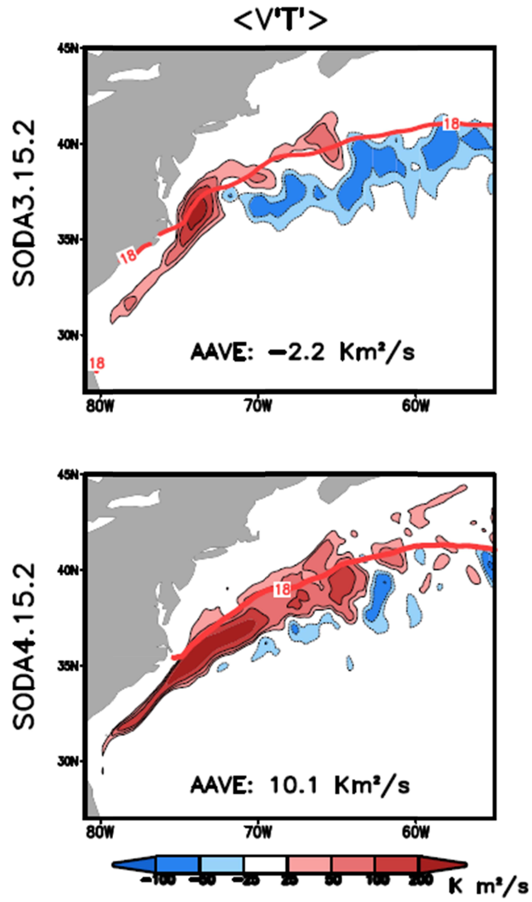

Fig. 8 Mean northward eddy temperature transport integrated over the upper 0-500m. Units are Km^2/s.

4 Discussion

This paper has two goals. The first is to describe an updated version of the Simple Ocean Data Assimilation (SODA4) ocean/sea ice reanalysis that includes an improved eddy-resolving forecast model. The second is to explore the impact of that improved (but noisier) forecast model on the resulting analysis. As part of that exploration, we compare two reanalysis experiments during a twelve-year period, the first, SODA4.15.2, using the new model, and the second, SODA3.15.2, using a forecast model which has 1/8th of the numerical resolution. The results of the comparison highlight areas where the improved forecast model reduces analysis error, but also identifies areas where the improved model does not reduce analysis error because other errors are more significant.

The comparison begins by examining the temperature and salinity increments that are produced as a byproduct of the data assimilation. The comparison then examines differences from an independent satellite sea level data set. The results show that the errors are similar in the deep tropics because here the major error sources error is due to surface wind error. In contrast, systematic errors are reduced in western boundary and frontal regions where the additional horizontal resolution allows fronts to become sharper and steeper, and the flow to produce more eddies. The additional near-surface vertical resolution also increases summer temperature stratification in the uppermost 30 m by as much as 2°-4°C, with corresponding increases in salinity stratification. Much of this additional stratification occurs quite close to the surface. Improved horizontal and vertical stratification also combine to alter the representation of the coastal circulation in summer. In the North Atlantic mid-Atlantic Bight, for example, the additional stratification enhances the isolation of the deeper shelf water from the near surface wind-driven cross-shelf flow. This change reduces cross-shelf exchange and alters the along-shelf transport, a change that has important implications for the usefulness of global ocean reanalyses in the coastal zone.

Time mean, depth integrated volume transports through major passages are broadly unaffected by the additional resolution (partly the result of topographic tuning), but the details of the transports are altered. For example, in the Indonesian Throughflow, the time mean transport is similar, but the fraction carried through individual straits: Lombok, Ombai, and Timor, is altered as is the time evolution of the throughflow transport. In the Atlantic increased resolution alters the overturning transport across the 26.5°N meridian and shifts the latitude of maximum heat transport poleward by 5-10°. Further north across the 59°N meridian in the subpolar gyre the impact of the change in resolution is much smaller. These results confirm previous recommendation based on free-running numerical simulations, e.g. *Griffies* (2015), to use sufficient resolution resolve the oceanic mesoscale and near-surface processes in order to improve the accuracy of ocean/sea ice reanalysis.

Acknowledgments

We gratefully acknowledge our data collectors and aggregators, listed in the Data Availability Statement below. Their efforts have made this work possible. Computing resources were provided by the Computational and Information Systems Laboratory of NCAR. Most of all, we are grateful to Baris M. Uz and the Physical Oceanography Program of the National Science Foundation (OCE1948952) for providing financial support for this work.

Data Availability

The SODA3.15.2 and SODA4.15.2 datasets described in Section 2 are available for download from soda.umd.edu in CF-compliant netCDF4 classic format. SODA4.15.2 ocean and sea ice state estimates are saved separately at the full native grid at 5-day resolution. Each 5.7 Gb 5-dy compressed 32bit floating point NetCDF5 ocean file contains potential temperature, practical salinity, velocity, and a set of diagnostic variables such as mixed layer depth, entrainment, net surface heat flux, sea level, and so on. Each 0.9 Gb 5-dy sea ice file contains five ice categories, total ice thickness and velocity. Each 2.3 Gb 10-day volume transport file contains transport in and out of the horizontal walls of each cell. Many ocean and ice variables have also been regridded onto a convenient regular 0.1° × 0.1° horizontal grid at a subset of 38 vertical levels using a conservative mapping scheme at both 5-dy and monthly resolution. The size of each individual uncompressed regridded ocean file is 2.1 Gb.

Hydrographic profile data was obtained from the NOAA National Center for Environmental Information World Ocean Database (www.nodc.noaa.gov) [last accessed June 2023], and the Unified Database for Arctic and Subarctic Hydrography (*Behrendt et al.,* 2018) from (doi.org/10.5194/essd-10-1119-2018) [accessed, 15 February 2022]. These data sets includes data from the ice tethered profilers, which is available directly at: https://www2.whoi.edu/site/itp. ERA5 hourly surface meteorological variables were obtained through the ECMWF (www.ecmwf.int) [most recently accessed, 15 February 2023]. In situ SST observations were obtained from the ICOADS version 3 archive, hosted by NOAA (icoads.noaa.gov/e-doc) [last accessed 1 July 2023]. Satellite AVHRR Pathfinder Version 5.2 (PFV5.2) SSTs were obtained from (pathfinder.nodc.noaa.gov). ACSPO SST data (*Jonasson, et al.,* 2022) are provided by NOAA. We strongly recommend contacting the NOAA SST team led by A. Ignatov before the data are used for any publication or presentation. *Dai* (2017) (DOI: 10.5065/D6V69H1T) [accessed 7 July 2020], the Arctic Great Rivers Observatory Discharge Dataset (*Shiklomanov et al*., 2021, www.arcticrivers.org/data) [accessed 7 July 2020], and the R-ArcticNET (russia-arcticnet.sr.unh.edu) [accessed 7 July 2020]. Daily combined sea level is available from the CMEMS web portal (http://marine.copernicus.eu/services-portfolio/access-to-products/ [last access: July, 2023]. CMEMS and C3S services are funded by the European Union.

References

Bamber, J., A. Tedstone, M. King, I. Howat, E. Enderlin, M. van den Broeke, & B. Noel (2018), Land ice freshwater budget of the Arctic and North Atlantic Oceans. Part I: Data, methods and results, *J. Geophys. Res*., 123, 1827– 1837, DOI:10.1002/2017JC013605.

Behrendt, A., Sumata, H., Rabe, B., and Schauer, U., (2018), UDASH – Unified Database for Arctic and Subarctic Hydrography, Earth Syst. Sci. Data, 10, 1119–1138, https://doi.org/10.5194/essd-10-1119-2018.

Behringer, D. W., 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, https://doi.org/10.1175/1520-0493(1998),126<1013:AICMFE>2.0.CO;2.

Bloom, S. C., L. L. Takacs, A. M. da Silva, & 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

Boyer, T., C.M. Domingues, A. Good, et al. (2016), Sensitivity of global upper-ocean heat content estimates to mapping methods, XBT Bias Corrections, and Baseline Climatologies. *J. Climate*, **29**, 4817–4842, doi.org/10.1175/JCLI-D-15-0801.1.

Boyer, T.P., O.K. Baranova, C. Coleman, H.E. Garcia, A. Grodsky, R.A. Locarnini, A.V. Mishonov, C.R. Paver, J.R. Reagan, D. Seidov, I.V. Smolyar, K. Weathers, & M.M. Zweng, (2018), World Ocean Database 2018. A.V. Mishonov, Technical Ed., NOAA Atlas NESDIS 87.

Carton, J. A., G. A. Chepurin, X. Cao, & 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.

Carton, J. A., G. A. Chepurin, & X. Cao (2000b), A Simple Ocean Data Assimilation analysis of the global upper ocean 1950–95. Part II: Results. *J. Phys. Oceanogr*., 30, 311–326.

Carton, J.A., G.A. Chepurin, & L. Chen (2018a), SODA3: a new ocean climate reanalysis, *J. Clim*., 31, 6967-6983, DOI:10.1175/JCLI-D-18-0149.1

Carton, J.A., G.A. Chepurin, L. Chen, & S.A. Grodsky (2018b), Improved global net surface heat flux, *J. Geophys. Res*., 123, 3144-3163, DOI:10.1002/2017JC013137

Carton, J. A., and G. A. Chepurin, (2023), RARE: The Regional Arctic Reanalysis. *J. Climate*, 36, 2333–2348. https://doi.org/10.1175/JCLI-D-22-0340.1

Chang, P., Zhang, S., Danabasoglu, G.,Yeager, S. G., Fu, H., Wang, H., et al. (2020). An unprecedented set of high-resolution earth system simulations for understanding multiscale interactions in climate variability and change. *Journal of Advances in Modeling Earth Systems*,12, e2020MS002298. https://doi.org/10.1029/2020MS002298

Cowley, R., R.E. Killick, T. Boyer, V. Gouretski, et al. (2021), International Quality-Controlled Ocean Database (IQuOD), v0.1: The Temperature Uncertainty Specification, *Frontiers in Marine Science*, 8, DOI=10.3389/fmars.2021.689695.

Dai, A. (2017), Dai and Trenberth Global River Flow and Continental Discharge Dataset. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. DOI: 10.5065/D6V69H1T. [Accessed 07 July 2020].

Eden, C. (2007), Eddy length scales in the North Atlantic Ocean, *J. Geophys. Res*., 112, C06004, doi:10.1029/2006JC003901.

Fairall, C.W., E.F. Bradley, J.E. Hare, A.A. Grachev, & J.B. Edson (2003), Bulk Parameterization of Air–Sea Fluxes: Updates and Verification for the COARE Algorithm. *J. Clim*., 16, 571–591, DOI: 10.1175/1520-0442(2003),016<0571:BPOASF>2.0.CO;2

Flagg, C.N., Dunn, M., Wang, D.-P., Rossby, H. T., and Benway, R. L. (2006), A study of the currents of the outer shelf and upper slope from a decade of shipboard ADCP observations in the Middle Atlantic Bight, *J. Geophys. Res*., 111, C06003, doi:10.1029/2005JC003116.

Frajka-Williams, E., I.J. Ansorge, J. Baehr, et al., (2019), Atlantic Meridional Overturning Circulation: Observed Transport and Variability, *Front. Mar. Sci*., 6, DOI: 10.3389/fmars.2019.00260.

Freeman, E., S.D. Woodruff, S.J. Worley, et al. (2016), International Comprehensive Ocean-Atmosphere Data Set (ICOADS), Release 3.0 - Monthly Summary Groups (MSG), Revision 3.0. NOAA National Centers for Environmental Information (NCEI),. [last accessed February, 2023].

GHRSST Project Office (2012), GHRSST Strategy and Implementation Plan (GDIP), [GPO-D02-GDIPv1 Version 1.1 (2012-2022)] Zenodo. https://doi.org/10.5281/zenodo.4700521.

Griffies, S. (2004), Fundamentals of Ocean Climate Models, Princeton University Press, 528pp. ISBN: 9780691118925.

Griffies, S. M. (2012), Elements of the modular ocean model (MOM),: 2012 release (GFDL Ocean Group Technical Report No. 7. Princeton, USA: NOAA/Geophysical Fluid Dynamics Laboratory. 614 + xiii pages.

Griffies, S. M. et al. (2015), Impacts on Ocean Heat from Transient Mesoscale Eddies in a Hierarchy of Climate Models. *J. Clim.* 28, 952–977, https://doi.org/10.1175/jcli-d-14-00353.1.

Hewitt, H.T., Roberts, M., Mathiot, P. et al. Resolving and Parameterizing the Ocean Mesoscale in Earth System Models. Curr Clim Change Rep 6, 137–152 (2020). https://doi.org/10.1007/s40641-020-00164-w

Jackson, L. C., Roberts, M. J., Hewitt, H. T., Iovino, D., Koenigk, T., Meccia, V. L., et al. (2020). Impact of ocean resolution and mean state on the rate of AMOC weakening. Climate Dynamics, 55(7–8), 1711–1732. https://doi.org/10.1007/s00382-020-05345-9

Jonasson, O. A. Ignatov, V. Pryamitsyn, B. Petrenko, & Y. Kihai, (2022), JPSS VIIRS SST Reanalysis Version 3. *Remote Sens.*, 14, 3476. https://doi.org/10.3390/rs14143476.

Hallberg, R. W. (2013), Using a resolution function to regulate parameterizations of oceanic mesoscale eddy effects. *Ocean Modell.*, 72, 92–103, doi: 10.1016/j.ocemod.2013.08.007.

Hecht, M.W., & H. Hasumi, Editors, 2008: Ocean modeling in an eddying regime, Wiley, 409pp, ISBN:9780875904429

Hersbach, H, B. Bell, P. Berrisford, et al. (2020), The ERA5 global reanalysis. Q. J. R. Meteorol Soc., 146, 1999–2049. DOI: 10.1002/qj.3803.

Hundsdorfer, W, & R.A. Trompert, (1994), Method of lines and direct discretization: a comparison for linear advection. *Applied Numerical Mathematics*, 13, 469–490.

Jonasson, O., I Gladkova, & A. Ignatov (2022), A. Towards global daily gridded super-collated SST product from low earth orbiting satellites (L3S-LEO-Daily), at NOAA. Proc. SPIE, 12118, 1211805, doi:10.1117/12.2620103

Kalnay, E., M. Kanamitsu, R. Kistler, et al. (1996), The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, 77, 437–472, https://doi.org/10.1175/1520-0477(1996),077<0437:TNYRP>2.0.CO;2.

Kersalé, M. et al. (2020), Highly variable upper and abyssal overturning cells in the South Atlantic.Sci. Adv.6,eaba7573.DOI:10.1126/sciadv.aba7573

Lee, H.-C., A. Rosati, & M. Spelman (2006), Barotropic tidal mixing effects in a coupled climate model: ocean conditions in the northern Atlantic, *Ocean Model.*, 11, 464–470. DOI: 10.1016/j.ocemod.2005.03.003.

Lentz, S. J. (2008), Observations and a Model of the Mean Circulation over the Middle Atlantic Bight Continental Shelf. *J. Phys. Oceanogr.*, 38, 1203–1221, https://doi.org/10.1175/2007JPO3768.1.

Lozier, M. S., et al. (2019), A sea change in our view of overturning in the subpolar North Atlantic. *Science,* 363,516-521. DOI:10.1126/science.aau6592.

Roemmich, D., et al. (2019), On the Future of Argo: A Global, Full-Depth, Multi-Disciplinary Array. *Frontiers in Marine Science*, 6, https://doi.org/10.3389/fmars.2019.00439.

Schmitt, R.W. (2008), Salinity and the global water cycle. *Oceanography*, 21, 12–19, https://doi.org/10.5670/oceanog.2008.63.

Schweiger, A., R. Lindsay, J. Zhang, M. Steele, H. Stern, & R. Kwok (2011), Uncertainty in modeled Arctic sea ice volume, *J. Geophys. Res.*, 116, C00D06, doi:10.1029/2011JC007084.

Shiklomanov, A.I., R.M. Holmes, J.W. McClelland, S.E. Tank, and R.G.M. Spencer (2020), Arctic Great Rivers Observatory. Discharge Dataset, Version 20200625. https://www.arcticrivers.org/data. [Accessed 07 July 2020].

Stammer, D., C. Wunsch, R. Giering, C. Eckert, P. Heimbach, J. Marotzke, A. Adcroft, C. N. Hill, & J. Marshall (2002), Global ocean circulation during 1992–1997, estimated from ocean observations and a general circulation model, *J. Geophys. Res.*, 107, 3118, doi:10.1029/2001JC000888.

Storto, A., et al. (2019), Ocean Reanalyses: Recent Advances and Unsolved Challenges, *Front. Mar. Sci.*, 6, doi:10.3389/fmars.2019.00418.

Susanto, R.D., Z. Wei, T.R. Adi, Q. Zheng, G. Fang, F. Bin, A. Supangat, T. Agustiadi, S. Li, M. Trenggono, & A. Setiawan (2016), Oceanography surrounding Krakatau Volcano in the Sunda Strait, Indonesia. *Oceanography*, 29, 264–172, https://doi.org/10.5670/oceanog.2016.31.

Taburet, G., A. Sanchez-Roman, M. Ballarotta, M.-I. Pujol, J.-F. Legeais, F. Fournier, Y. Faugere, & G. Dibarboure (2019), DUACS DT2018: 25 years of reprocessed sea level altimetry products, *Ocean Sci.*, 15, 1207–1224, https://doi.org/10.5194/os-15-1207-2019, 2019.

Trenberth, K. E., & Y. Zhang (2019), Observed Interhemispheric Meridional Heat Transports and the Role of the Indonesian Throughflow in the Pacific Ocean. *J. Climate*, 32, 8523–8536, https://doi.org/10.1175/JCLI-D-19-0465.1.

Volkov, D. L., T. Lee, & L.-L. Fu (2008), Eddy-induced meridional heat transport in the ocean, *Geophys. Res. Lett.*, 35, L20601, doi:10.1029/2008GL035490.

Winton, M. (2000), A reformulated three-layer sea ice model, J. Atmos. Ocean Technol., 17, 525–531. DOI: 10.1175/1520-0426(2000),017<0525:ARTLSI>2.0.CO;2

Winton, M., W. G. Anderson, T. L Delworth, S. M. Griffies, W. J Hurlin, & A. Rosati (2014), Has coarse ocean resolution biased simulations of transient climate sensitivity? *Geophys. Res. Letts.*, 41, 8522–8529. https://doi.org/10.1002/2014GL061523

Zhang, J., & D.A. Rothrock (2003), Modeling Global Sea Ice with a Thickness and Enthalpy Distribution Model in Generalized Curvilinear Coordinates. *Mon. Wea. Rev.*, 131, 845–861, https://doi.org/10.1175/1520-0493(2003),131<0845:MGSIWA>2.0.CO;2.