Tracer and observationally-derived constraints on horizontal and diapycnal diffusivities in ocean models

David Samuel Trossman[1], Thomas W N Haine[2], Amy Frances Waterhouse[3], Arash Bigdeli[1], Matthew R. Mazloff[4], Caitlin Whalen[5], An T Nguyen[1], Patrick Heimbach[6], and Robin M Kovach[7]

[1]University of Texas-Austin
[2]Johns Hopkins University
[3]University of California, San Diego
[4]UCSD
[5]University of Washington
[6]University of Texas at Austin
[7]Science Systems and Applications

November 23, 2022

Abstract

Mixing parameters can be inaccurate in ocean data assimilation systems, even if there is close agreement between observations and mixing parameters in the same modeling system when data are not assimilated. To address this, we investigate whether there are additional observations that can be assimilated by ocean modeling systems to improve their representation of mixing parameters and thereby gain knowledge of the global ocean's mixing parameters. Observationally-derived diapycnal diffusivities–using a strain-based parameterization of finescale hydrographic structure–are included in the Estimating the Circulation & Climate of the Ocean (ECCO) framework and the GEOS-5 coupled Earth system model to test if adding observational diffusivities can reduce model biases. We find that adjusting ECCO-estimated and GEOS-5-calculated diapycnal diffusivity profiles toward profiles derived from Argo floats using the finescale parameterization improves agreement with independent diapycnal diffusivity profiles inferred from microstructure data. Additionally, for the GEOS-5 hindcast, agreement with observed mixed layer depths and temperature/salinity/stratification (i.e., hydrographic) fields improves. Dynamic adjustments arise when we make this substitution in GEOS-5, causing the model's hydrographic changes. Adjoint model-based sensitivity analyses suggest that the assimilation of dissolved oxygen concentrations in future ECCO assimilation efforts would improve estimates of the diapycnal diffusivity field. Observationally-derived products for horizontal mixing need to be validated before conclusions can be drawn about them through similar analyses.

Tracer and observationally-derived constraints on horizontal and diapycnal diffusivities in ocean models

D. S. Trossman,[1,2] C. B. Whalen,[3] T. W. N. Haine,[4], A. F. Waterhouse,[5] A. Bigdeli,[1] A. T. Nguyen,[1] M. Mazloff,[5] P. Heimbach,[1,6] R. M. Kovach[7]

[1]Oden Institute for Computational Engineering and Sciences, University of Texas, Austin, USA
[2]Global Science & Technology, NOAA STAR/NESDIS, College Park, USA
[3]Applied Physics Laboratory, University of Washington, Seattle, USA
[4]Department of Earth and Planetary Sciences, Johns Hopkins University, Baltimore, USA
[5]Scripps Institution of Oceanography, University of California, San Diego, USA
[6]Jackson School of Geosciences & Institute for Geophysics, University of Texas, Austin, USA
[7]Science Systems and Applications, Inc., NASA Goddard Space Flight Center, Greenbelt, USA

Key Points:

- Model-calculated diapycnal diffusivities disagree with microstructure observations, but this can be improved with multiple data sources
- Adjusting model-calculated diapycnal diffusivities primarily affects resolved advection of heat and salt via dynamic adjustment
- Adjoint-based data assimilation of biogeochemical tracers could potentially help estimate more accurate ocean mixing parameters

Corresponding author: D. S. Trossman, `david.s.trossman@noaa.gov`

Abstract

Mixing parameters can be inaccurate in ocean data assimilation systems, even if there is close agreement between observations and mixing parameters in the same modeling system when data are not assimilated. To address this, we investigate whether there are additional observations that can be assimilated by ocean modeling systems to improve their representation of mixing parameters and thereby gain knowledge of the global ocean's mixing parameters. Observationally-derived diapycnal diffusivities–using a strain-based parameterization of finescale hydrographic structure–are included in the Estimating the Circulation & Climate of the Ocean (ECCO) framework and the GEOS-5 coupled Earth system model to test if adding observational diffusivities can reduce model biases. We find that adjusting ECCO-estimated and GEOS-5-calculated diapycnal diffusivity profiles toward profiles derived from Argo floats using the finescale parameterization improves agreement with independent diapycnal diffusivity profiles inferred from microstructure data. Additionally, for the GEOS-5 hindcast, agreement with observed mixed layer depths and temperature/salinity/stratification (i.e., hydrographic) fields improves. Dynamic adjustments arise when we make this substitution in GEOS-5, causing the model's hydrographic changes. Adjoint model-based sensitivity analyses suggest that the assimilation of dissolved oxygen concentrations in future ECCO assimilation efforts would improve estimates of the diapycnal diffusivity field. Observationally-derived products for horizontal mixing need to be validated before conclusions can be drawn about them through similar analyses.

Plain Language Summary

How the ocean mixes across space and time is not yet adequately simulated by models. One way to estimate this mixing is to use a framework that minimizes a function of the disagreements between observations and the model. However, there are many other variables the model needs to estimate and there are observations of relatively few variables. Currently, this model only constrains ocean mixing with observations of the warmth and saltiness of the ocean. To help the model estimate more realistic ocean mixing, some theories can be used to quantify ocean mixing from observations. Here, we show evidence that at least one of these theories is realistic, but because there are large uncertainties with the estimates from these theories, here we test whether there are measured variables with relatively small uncertainties that can be used to constrain ocean mixing in the model. We find some evidence that aiming to achieve better agreement between a model's oxygen concentrations and those from observations could help reduce the errors in ocean mixing in the model.

1 Introduction

In this paper, we consider whether additional observations may aid in representing mixing in ocean data assimilation systems. Previous studies have documented the importance of ocean mixing in setting the general circulation of the ocean and its role in global climate variability. Ocean mixing is typically conceptualized in terms of diffusion along and across isopycnal surfaces, as well as associated with the transport of isopycnal thickness (or bolus). Ocean models often represent mixing with three parameters: the across-isopycnal mixing parameter (diapycnal diffusivity; *Munk and Wunsch*, 1998), the along-isopycnal mixing parameter (Redi coefficient; *Redi*, 1982), and the eddy isopycnal thickness transport parameter (Gent-McWilliams coefficient; *Gent and McWilliams*, 1990). Mixing across isopycnal surfaces is an essential ingredient to explain the observed oceanic stratification (Munk & Wunsch, 1998; Gnanadesikan, 1999; J. R. Scott & Marotzke, 2002). Changes in background mixing across isopycnals (Dalan et al., 2005; Krasting et al., 2018; Sinha et al., 2020), mixing along isopycnals (Gnanadesikan et al., 2015; Ehlert et al., 2017), and eddy isopycnal thickness transport (Danabasoglu & McWilliams, 1995) each have a profound influence on climate simulations through alterations in the response to surface flux perturbations and changes in ventilation rates.

Ocean models must parameterize the unresolved turbulent diffusion of oceanic tracers since they are unable to resolve the scales of the processes responsible for mixing. However, it has been a challenge to observe, calculate, and assess the three ocean mixing parameters mentioned above. The Redi coefficient is still set to be globally constant in many ocean models, even though several observational studies have found evidence of substantial spatial and/or temporal variability in mixing along isopycnals (R. P. Abernathey & Marshall, 2013; Forget et al., 2011; Cole et al., 2015; Busecke & Abernathey, 2019). Despite theoretical progress (Bates et al., 2014; Groeskamp et al., 2020), the vertical structure of the Redi coefficients remains unknown. The Gent-McWilliams coefficient is known to be very similar to the Redi coefficient (Bachman et al., 2020), except in the vicinity of intensified jets, where multiple models set the Gent-McWilliams coefficient to be unequal to the Redi coefficient. While complex parameterizations for the diapycnal diffusivity field (Gaspar et al., 1990; Large et al., 1994) have allowed models to use spatiotemporally-varying diapycnal diffusivities for decades now, studies have only begun to use observations to improve the diapycnal diffusivities in ocean models. For instance, Zhu and Zhang (2020) and Zhu et al. (2020) have shown that diapycnal diffusivities derived from Argo floats can be used to improve some variables in ocean models. Also, Pollmann et al. (2017) and de Lavergne et al. (2020) have evaluated global internal wave mixing schemes using observationally-derived diapycnal diffusivities.

We use an ocean parameter and state estimation framework to evaluate how near-global, observationally-derived estimates of mixing can improve ocean models. The aim of this framework is to reconstruct the recent history of the ocean (the "state estimate") by filling in the gaps between incomplete observations–often sparse and aliased ones–through data assimilation techniques. The state estimate is much like a reanalysis product, but the state estimation framework overcomes some shortcomings by requiring dynamical and kinematical consistency (Stammer et al., 2016). The state estimate is achieved by fitting a general circulation model to available observations in a weighted least-squares sense (Wunsch, 2006). The model-data misfit (objective or "cost function") is minimized by varying (i.e., inverting for) a set of uncertain control variables, all of which are independent inputs to the model equations being solved. Importantly for our goal of parameter estimation, the set of control variables may consist not only of initial and boundary conditions, but also of (spatially-varying) model parameters, such as the ones used to represent ocean mixing. To provide accurate ocean mixing parameter estimates, the framework should minimize numerical diffusion.

Previously, the only available observational information about ocean mixing came from tracer release experiments (Ledwell & Watson, 1991; Polzin et al., 1997; Messias et al., 2008) and microstructure (i.e., the scales over which molecular viscosity and diffusion are important) measurements of velocity shear (e.g., *Waterhouse et al.*, 2014). These data are infrequently sampled and cover a much smaller portion of the ocean than the more recent global mixing data products that have made combined use of finestructure data and parameterizations mentioned above. None of these observations have been assimilated in existing ocean state estimation frameworks to constrain the diapycnal diffusivity field. Each of the three ocean mixing parameters have been included as control parameters, but the only constraints provided to any of them come from hydrographic (i.e., temperature, salinity, and pressure) observations. C. Liu et al. (2012) found that including the three ocean mixing parameters as control parameters in the optimization of an ocean state estimate can reduce the total cost function, a measure of model performance relative to observations, over the entire ocean from 1992 to 2001 by 10% compared with only including surface fluxes as control parameters. However, with a similar ocean state estimation framework but different model configuration, Forget et al. (2015)

suggests more than twice as large of an effect over the entire ocean from 1992 to 2011 (see their Table 8).

An open question is what observations (other than temperature, salinity, and pressure) can provide useful constraints on ocean mixing parameters. To do this, we must first perform comparisons of the ocean mixing parameters from an ocean state estimate with observations, which have not previously been performed. We examine whether the diapycnal diffusivities from an observationally-derived data product have smaller biases relative to microstructure observations than the diapycnal diffusivities from an ocean state estimate (Sections 3.1 and 3.2). We argue that because large biases remain in the diapycnal diffusivities from a recent ocean parameter and state estimate compared with microstructure observations, assimilation of hydrographic observations is insufficient to estimate ocean mixing parameters using the ocean parameter and state estimation framework. We also assess whether a coupled earth system model's hydrography is improved relative to observational climatologies when its diapycnal diffusivities are substituted with ones derived from Argo floats (Section 3.3) and what the implications for steric sea level are (Section 3.4). We analyze the steric sea level budget for each coupled earth system model simulation because this framework provides us with an understanding of how the model's dynamics change upon variation of the diapycnal diffusivity field. Finally, we perform model experiments in forward plus backward ("adjoint") mode to determine whether biogeochemical tracer data and observationally-derived diapycnal diffusivities would provide similar constraints on ocean mixing when assimilated (Section 3.5). The latter exercise is repeated for an Argo-derived Redi coefficient field, but not for the third mixing parameter, the Gent-McWilliams coefficient, because this parameter cannot be directly compared with our model's Gent-McWilliams coefficient and is known to be very similar to the Redi coefficient (Bachman et al., 2020). These experiments allow us to conclude whether biogeochemical tracer data could be assimilated in a future optimization of an ocean state estimation framework to better estimate either of the two ocean mixing parameters considered here.

2 Methods

2.1 Observationally-derived data products and measured data

2.1.1 Diapycnal Diffusivities

The diapycnal diffusivities in our model simulations use data sets of diapycnal diffusivities derived from observations. (We distinguish between "observations" that are measured quantities using in situ instruments and observationally-derived values, which use measured quantities and a theory to derive values. The former data have only measurement uncertainties, while the latter data have both measurement and structural uncertainties.) These data sets contain values equatorwards of 75^{o}S and 75^{o}N–no shallower than about about 250 meters because the method does not yield accurate results in the presence of strong upper-ocean density variability (e.g., *D'Asaro*, 2014). The diapycnal diffusivities are derived from finestructure observations using a strain-based finescale parameterization, which has been developed and implemented in different ways (Henyey et al., 1996; Gregg, 1989; Polzin et al., 1995, 2014) but typically assumes a mixing efficiency of 0.2 (St. Laurent & Schmitt, 1999; Gregg et al., 2018). The finescale parameterization assumes that 1) the production of turbulent energy at small scales is due to an energy transfer driven by wave-wave interactions down to a wave breaking scale; 2) nonlinearities in the equation of state, double diffusion, downscale energy transports, and mixing associated with boundary layer physics and hydraulic jumps are neglected; and 3) stationary turbulent energy balance exists where production is matched by dissipation and a buoyancy flux in fixed proportions (Polzin et al., 2014). The implementation by *Whalen et al.* (2015) assumes a shear-to-strain variance ratio of 3 and a flux Richardson number, $R_f = 0.17$, is used to determine the fraction of turbulent production that goes into the buoyancy flux and the rest for dissipation. The finestructure method is not expected to be valid in equatorial regions of the ocean, but nevertheless, the diapycnal diffusivity product compares well with microstructure near the equator (Whalen et al., 2015). We use the 2006-2014 climatology of Whalen et al. (2015)–referred to as $\kappa_{\rho,W15}$ hereafter–which is a gridded product on an approximately $1^o \times 1^o$ horizontal grid and has three vertical levels: 250-500 meters, 500-1000 meters, and 1000-2000 meters depth. Whalen et al. (2015) found that 81% (96%) of their Argo-derived diapycnal diffusivities from the finescale parameterization are within a factor of two (three) of the microstructure measurements. We use this as the basis for the factor of 2-3 uncertainty we cite hereafter. We also use the implementation of Whalen et al. (2015) to construct a time-varying Argo-derived diapycnal diffusivity data set from 2001 to 2016–referred to as $\kappa_{\rho,t}$ hereafter. (In 2001, the Profiling Autonomous Lagrangian Circulation Explorer (PALACE) floats (Davis, 1991; Davis et al., 1992) are used, and they supplement the Argo data through 2006.) In addition to the Argo-derived diapycnal diffusivities, there are ship-based Conductivity, Temperature, and Depth (CTD) hydrography-derived diapycnal diffusivity field (Kunze, 2017)–referred to as $\kappa_{\rho,K17}$ hereafter–that uses the same finestructure parameterization as the $\kappa_{\rho,W15}$ product is included (see Section 2.3). The vertical resolution of the $\kappa_{\rho,K17}$ product is 256 meters and horizontal resolution is the spacing between each CTD profile.

Microstructure-inferred diapycnal diffusivities (Osborn, 1980; Lueck et al., 1997; Gregg, 1989; Moum et al., 2002; Waterhouse et al., 2014) are used to evaluate each model's diapycnal diffusivities. (We further distinguish "observationally-inferred" values, which are from the currently accepted method of observing a quantity such as a diapycnal diffusivity but are not measured, and "observationally-derived" values because the latter data depend on a method that requires additional assumptions.) The microstructure-inferred diapycnal diffusivities are based on an expression for the isotropic turbulence field, which is proportional to the viscosity of water and the velocity shear resolved to dissipative scales (*Thorpe*, 2007; and references therein). The depth ranges of the data collected by Waterhouse et al. (2014)–referred to as $\kappa_{\rho,micro}$ hereafter–go from the upper several hundred meters to the full water column. The profiles are seasonally aliased at higher latitudes and span decades. There are thousands of vertical profiles that comprise this data set, samples being taken in North Pacific Ocean, North Atlantic Ocean, tropical Pacific, near Drake Passage, near the Kerguelen Plateau, and in the South Atlantic Ocean. Many of the profiles were taken in regions with both smooth and rough bottom topography. To compare the microstructure profiles with model output, the nearest neighbors to each model's grid are selected, which reduces the data set to 42 profiles.

We use a consistent comparison method for both ECCO and GEOS-5 output by accounting for the fact that the GEOS-5-calculated diapycnal diffusivities are time-varying and the ECCO-estimated diapycnal diffusivities are not. The comparison method described below nudges each model's diapycnal diffusivity field closer to the $\kappa_{\rho,W15}$ product at microstructure profile observation locations. Each model's initial diapycnal diffusivity profiles and their nudged diapycnal diffusivity profiles are then compared to microstructure profile observations at the same locations. This comparison allows us to assess whether the bias in each model's diapycnal diffusivity profiles is reduced when nudged closer to the $\kappa_{\rho,W15}$ product. We use the below nudging method because there are only three points in the vertical in the $\kappa_{\rho,W15}$ product and the nudging effectively simulates how a model's diapycnal diffusivity profile would respond to the assimilation of the $\kappa_{\rho,W15}$ product. We nudge the model-calculated diapycnal diffusivity field's temporal mean by applying an adjustment derived from the $\kappa_{\rho,W15}$ product to get a new diapycnal diffusivity field, $\hat{\kappa}_\rho$, according to Equation (A1) from Zhang et al. (2001),

$$\hat{\kappa}_\rho = \begin{cases} \kappa_\rho + \frac{\overline{\kappa}_{\rho,Argo} - \overline{\kappa}_\rho}{\kappa_0 - \overline{\kappa}_\rho}(\kappa_0 - \kappa_\rho), & \text{if } \overline{\kappa}_{\rho,Argo} > \overline{\kappa}_\rho \\ \kappa_\rho + \frac{\overline{\kappa}_{\rho,Argo} - \overline{\kappa}_\rho}{\overline{\kappa}_\rho}\kappa_\rho, & \text{if } \overline{\kappa}_{\rho,Argo} \le \overline{\kappa}_\rho \end{cases} \tag{1}$$

Here, κ_0 is set to be the maximum possible diapycnal diffusivity found in the model, κ_ρ is the monthly averaged diapycnal diffusivity for temperature/salinity calculated from model output, $\kappa_{\rho,Argo}$ has a yearly mean equal to $\kappa_{\rho,W15}$ and seasonal cycle set by the model, $\overline{\kappa}_\rho$ is the 24-year averaged model output of the diapycnal diffusivity, $\overline{\kappa}_{\rho,Argo} = \kappa_{\rho,W15}$, and $\hat{\kappa}_\rho$ is the diapycnal diffusivity field used in the simulations that utilize the diapycnal diffusivity increment. Equation (1) nudges the model-based diapycnal diffusivity field (κ_ρ) so that its long-term mean is closer to the Argo-derived diapycnal diffusivity field from the finescale parameterization ($\kappa_{\rho,W15}$). Because Equation (1) ensures that the extreme values of $\hat{\kappa}_\rho$ are non-negative and never exceed κ_0, the distribution of $\hat{\kappa}_\rho$ in time at each grid point may be skewed relative to its initial distribution. We leave the model's diapycnal diffusivities unchanged in the mixed layer since if we override the diapycnal diffusivities in the mixed layer, the model will cease to convect, even under convection-favorable conditions.

2.1.2 Along-Isopycnal Diffusivities: Redi Coefficients

An Argo- and ocean state estimate-derived Redi coefficient field from mixing length theory Cole et al. (2015), both above and below the mixed layer depth, can also be used in our model simulations. The Redi coefficients are computed as the product of a mixing length scale, characteristic velocity scale, and a mixing efficiency–assumed to be one. Cole et al. (2015) used Argo observations to compute the mixing length scale (see *Cole et al.*, 2015 - see their Eq. 1a). The mixing length scale is computed as the ratio of the temporal standard deviation of the salinity field over the horizontal gradient of the mean salinity field from the Argo data. Cole et al. (2015) used output from a nominally $1/4^o$ ocean state estimate (ECCO2) to calculate the characteristic velocity scale (equal to the ECCO2's root-mean-square velocity field). The primary differences between ECCO2 and the ocean state estimate configuration we run for the purposes of this manuscript are that ECCO2 is eddy-permitting, on a cube-sphere grid, runs over 2005-2012, uses a Green's function approach to adjust a small number of control parameters (Menemenlis et al., 2005). The final Cole et al. (2015) product–referred to as $\kappa_{Redi,C15}$ hereafter–is a climatology with 1 meter vertical resolution between 2000 meters depth and close to the surface. This product is on an approximately $1^o\times1^o$ horizontal grid, matching the model resolution of the model we compare it to. There are very few independent observationally-inferred data sets (e.g., NATRE and DIMES) with which to pursue an assessment of the Redi coefficient field (Groeskamp et al., 2020), like we have with microstructure for assessment of the diapycnal diffusivity field, so we only compare the model output with the $\kappa_{Redi,C15}$ product. Also, while there are Gent-McWilliams coefficients derived from Argo observations (Katsumata, 2016), the treatment of the rotational component of their estimated eddy transport has a different treatment from that in C. Liu et al. (2012), which uses the same treatment as the model we use here. Thus, we exclude consideration of the Gent-McWilliams coefficients altogether from this study.

2.1.3 Biogeochemical constraints

In addition to the mixing products, we perform similar analyses using oxygen, phosphate, and alkalinity as other potential constraints on ocean mixing. Oxygen has vertical gradients that can be resolved by ocean models, has future changes projected to be dependent upon mixing across and along isopycnals (Palter & Trossman, 2018; Couespel et al., 2019), and has been shown to depend upon the Redi coefficients (Gnanadesikan et al., 2015; Bahl et al., 2019; Rudnickas et al., 2019; Bahl et al., 2020) due to their ability to modulate deep wintertime convection. Further, alkalinity is known to be sensitive to fresh/salty water perturbations due to the contributing dilution/concentration of charge (Jiang et al., 2014; Kakehi et al., 2017), and phosphate is a function of its supply regions, which could provide an imprint of how water mixes (Paytan & McLaughlin, 2007). Thus, we choose to include the oxygen, alkalinity, and phosphate climatologies from the World Ocean Atlas (2013) in our simulations.

2.2 Modeling systems

Details about the model simulations we perform are first summarized in Section 2.3 and in Table 1. The first modeling system uses time-varying diapycnal diffusivities calculated from a suite of parameterizations, where the diapycnal diffusivities associated with temperature and salinity differ due to double-diffusive processes, but the Redi coefficient field is constant everywhere for all times. The second modeling system uses time-invariant but spatially varying diapycnal diffusivity and Redi coefficient fields, each estimated with an optimization procedure, where the diapycnal diffusivities associated with temperature and salinity are assumed to be identical. We describe how each of these modeling systems are used in combination with several observationally-derived products, listed in Table 2, and in situ measurements in Section 2.1.

2.3 Model experiments

2.3.1 GEOS-5

The GEOS-5 modeling system is comprised of multiple components. GEOS-5 includes a global, finite volume atmospheric general circulation model that is used for numerical weather prediction, seasonal-to-decadal forecasts, and as the background field for atmospheric reanalyses (Molod et al., 2015). The ocean is represented by the GFDL Modular Ocean Model (Griffies et al., 2015), version 5 (MOM5) and the Los Alamos Community Ice CodE sea ice model (Hunke et al., 2013), version 4.1 (CICE4.1). We use a configuration of the GEOS-5 modeling system with a 1^o (0.5^o at equator) resolution on a tripolar (*Murray*, 1996) staggered Arakawa B-grid (*Mesinger and Arakawa*, 1976) and 50 geopotential levels for MOM5, 2^o resolution and 24 pressure levels for the atmospheric model, and 1^o resolution and 3 layers for CICE4.1. Historical aerosols (sulfate, dust, and sea salt) and biomass burning emissions (black and organic carbon) updated from the Goddard Chemistry Aerosol Radiation and Transport (GOCART) model (Chin et al., 2002) are used over the time period 1992 through 2016. Initial conditions are based on a long spin-up that used MOM4 coupled to one version of the GEOS-5 atmosphere model (Molod et al., 2012) and hundreds of additional years of spin-up that used MOM4 coupled to a slightly different version of the GEOS-5 atmosphere model. The differences between the two versions of the GEOS-5 atmospheric model used in the two phases of spin-up include developments in cloud microphysics and atmospheric chemistry.

The diapycnal diffusivities, Redi coefficients, and Gent-McWilliams coefficients are determined in MOM5 as follows. Diapycnal diffusivities in MOM5 are represented by the K-Profile Parameterization (KPP; *Large et al.*, 1994) and a parameterization for mixing due to internal tides (Simmons et al., 2004). Shear-driven mixing, gravitational instabilities that can cause vertical convection, and double-diffusive processes, which can cause the temperature diffusivity to be different from the salinity diffusivity, are accounted for in the interior (*Large et al.*, 1994). The resulting diapycnal diffusivities spatio-temporally vary. However, this combination of parameterizations does not make use of an explicit energy budget that accounts for conversion between kinetic and potential energy when determining the diapycnal diffusivities. The Redi coefficients (Redi, 1982) and Gent-McWilliams coefficients of the Gent and McWilliams (1990) parameterization for mesoscale eddies are, by default, prescribed to be 600 $m^2\ s^{-1}$ everywhere, except for some variation in western boundary current regions for the Gent-McWilliams coefficients. The Redi coefficients and Gent-McWilliams coefficients are, thus, constant in time and in most locations. A mixed layer instability scheme for the submesoscale transport by *Fox-Kemper et al.* (2011) is used.

Multiple coupled simulations are run using the GEOS-5 modeling system. We use GEOS-5 because it accounts for coupled feedbacks, such as the sea ice-albedo and cloud feedbacks, that, in addition to ocean dynamics, contribute to internal variability of the Earth system model. We inquire whether the error associated with the diapycnal diffusivity parameter is an important source of model error, relevant on the timescales of our simulation. We take the approach of substituting the diapycnal diffusivities with time-varying ones where and whenever they are available for the following reason. Substitution of the model-calculated diapycnal diffusivities with κ_{ρ_t} allows for spatial as well as temporal variations found in the Argo-derived data. The $\kappa_{\rho,W15}$ product does not capture temporal variations in the diapycnal diffusivity field, which are likely important in locations such as the tropical Pacific Ocean (Warner & Moum, 2019).

We perform and analyze the following GEOS-5 experiments:

- **G-CTRL** - a 25 year in length (1992-2016) hindcast run that substitutes the diapycnal diffusivities computed online with $\kappa_{\rho,MOM5}$
- **DIFF** - a 25 year in length (1992-2016) hindcast run that substitutes the diapycnal diffusivities computed online with $\kappa_{\rho,t}$ except where they are not available, in which case $\kappa_{\rho,MOM5}$ are used; "DIFF" here stands for diffusivity, not difference
- **BKG** - twenty-one background free-running simulations that are each 1 year in length and identical except in their initialization; each starts from a different time (each month of 1992)
- **GMAO S2S Ocean Analysis** - a reanalysis product using the GEOS-5 modeling system, but with data assimilation from May of 2012 to March of 2019 (see Section 2.3.1.2)

G-CTRL is compared with DIFF instead of the 25 year in length free-running simulation because the diapycnal diffusivities are substituted in the same way, but with different values. Here, G-CTRL is not necessarily the same as the free-running simulation because the frequency of variability in the diapycnal diffusivity field differs. Sub-monthly variability in the diapycnal diffusivities is suppressed in G-CTRL because $\kappa_{\rho,MOM5}$ is an averaged monthly output field from the model that is interpolated in time as the model runs. The wind forcing, for instance, could cause sub-monthly variability in the model-calculated diapycnal diffusivities. Sub-monthly frequencies in the diapycnal diffusivities may be important in the real ocean due to internal tidal breaking during the spring tide, but these higher frequency effects are not sufficiently represented by the model. In G-CTRL and DIFF, the diapycnal diffusivities associated with temperature and salinity are different and have a time-varying components, as calculated using the KPP (Large et al., 1994) scheme. In DIFF, the diapycnal diffusivities associated with temperature calculated with the GEOS-5 model are substituted with $\kappa_{\rho,t}$, except where there are insufficient observations, in which case $\kappa_{\rho,MOM5}$ is used. The differences between the diapycnal diffusivities associated with temperature and those associated with salinity in G-CTRL are preserved in the DIFF experiment wherever substitutions are made. The finescale turbulence parameterization does not distinguish between the diapycnal diffusivities associated with temperature and those associated with salinity so their difference as calculated in G-CTRL is assumed to be the same in DIFF.

2.3.1.1 Diagnostic for understanding dynamical impacts of diapycnal diffusivity changes: steric sea level budget

In order to better understand how diapycnal diffusivity changes dynamically impact the ocean in the present study, we analyze a model's buoyancy budget, which is broken down into heat and salt budgets and used to calculate the steric sea level budget. The tracer tendency equation terms required for the heat and salt budgets were computed as the reanalysis was produced. The tracer equations can be broken down into individual contributions (Palter et al., 2014),

$$\rho\frac{d\Theta}{dt} + \rho A^{\Theta} = -\nabla\cdot\mathbf{J}^{\Theta} + \rho Q^{\Theta} \qquad (2)$$
$$\rho\frac{dS}{dt} + \rho A^{S} = -\nabla\cdot\mathbf{J}^{S} + \rho Q^{S},$$

where $d/dt = \partial/\partial t + (\mathbf{v}+\mathbf{v}^*)\cdot\nabla$ is the material derivative, $\mathbf{v}$ is the resolved velocity field, $\mathbf{v}^*$ is the eddy-induced or quasi-Stokes velocity field that represents parameterized motions, Θ is the potential temperature, S is the salinity, ρ is the locally referenced potential density, $\mathbf{J}^{\Theta}$ and $\mathbf{J}^{S}$ are the parameterized along-isopycnal and diapycnal mixing fluxes associated with potential temperature and salinity, Q^{Θ} and Q^{S} are the sums of sources and sinks of potential temperature and salinity, and A^{Θ} and A^{S} are the analysis increments for potential temperature and salinity due to the assimilation of data by a sequential filter-based data assimilation ocean modeling system. The analysis increments in the sequential filter-based data assimilation system–such as the one described below, in Section 2.3.1.2–obscure the physics so we do not analyze its output.

The heat and salt budget terms summarized by Equation (2) are computed as follows. The resolved, mesoscale, and submesoscale transports are accounted for in the material derivatives Θ and S, the neutral and diapycnal diffusion of Θ and S are accounted for by $\mathbf{J}^{\Theta}$ and $\mathbf{J}^{S}$, and the analysis increments of Θ and S are accounted for by A^{Θ} and A^{S}. The neutral diffusion term includes cabbeling, thermobaricity, and a dianeutral contribution that mixes properties by providing for the exponential transition to horizontal diffusion in regions of steep isoneutral slopes according to Treguier (1992) and Ferrari et al. (2008, 2010) where the surface boundary layer is encountered and following Gerdes et al. (1990) next to solid walls. The diapycnal diffusion term is not added to the vertical component of the along-isopycnal diffusion term, but because of convention (e.g., *Palter et al.*, 2014) is nevertheless referred to as the vertical diffusion term hereafter. The vertical diffusion term also includes penetrating shortwave radiation flux. The sources and sinks of Θ and S accounted for by Q^{Θ} and Q^{S} include nonlocal convection (the transport where turbulent fluxes don't depend upon local gradients in Θ or S because buoyant water gets entrained into the mixed layer when the surface buoyancy forcing drives convection above a stratified water column); surface buoyancy fluxes (latent, sensible, shortwave, longwave, and frazil heat fluxes); precipitation minus evaporation; runoff mixing (mixes properties associated with river outflows); downslope mixing (mixes properties downslope to represent the overflow dense waters from marginal seas); sigma-diffusion (mixing properties along terrain-following coordinates in regions with partial bottom cells); numerical smoothing of the free surface (intended to reduce B-grid checkerboard noise); numerical sponge (intended to absorb the Kelvin waves set off by the assimilation of some data); calving of land ice; and frazil ice formation. The runoff mixing, downslope mixing, and sigma-diffusion terms are considered sources or sinks here because they are associated with numerical schemes that aim to resolve problems created by coarse model resolution, the vertical coordinate system used near boundary layers, and imperfect bathymetry. There is no geothermal heating included in the GMAO S2S Ocean Analysis. The vertical diffusion term includes a subsurface shortwave heating contribution to a function of the diapycnal diffusivity field, the mesoscale transport term assumes constant Gent-McWilliams coefficients, and the neutral diffusion term assumes constant Redi coefficients, explaining why each of these three terms are non-zero globally (Table 3).

At each time step, the model evaluates a tendency term for every process that contributes to (2) from their parameterized or dynamically calculated values, their units are converted to W m^{-2} and kg m^{-2} s^{-1} for Θ and S, and their monthly averages are saved to the output files used in this analysis. Implicit in these output tendency terms is that each term is weighted by the thicknesses of each layer as the model runs and writes the output to file. The heat and salt budget terms saved to file are used to calculate the steric sea level budget as follows. The steric sea level budget terms are computed by scaling the heat tendency terms by α/C_p and the salt tendency terms by -1000β, where C_p (units in J kg^{-1} K^{-1}) is the specific heat of seawater, $\alpha = -[1/\rho](\partial\rho/\partial T)$ (units in K^{-1}) is the thermal expansion coefficient, and $\beta = [1/\rho](\partial\rho/\partial S)$ (units in kg g^{-1}) is the haline contraction coefficient. In order to get a longitude-latitude map of the terms that depend upon depth shown here, we integrate over depth by summing over the depth dimension. We only analyze the steric sea level budgets of G-CTRL and DIFF here in order to interpret the dynamical changes upon adjusting the diapycnal diffusivities.

2.3.1.2 Comparison with a reanalysis product

Before we present some dynamical impacts of perturbed diapycnal diffusivities and ultimately examine how to better constrain ocean mixing parameters in ocean data assimilation systems, we present an example of why there could be a need for better constraints on ocean mixing parameters. To do this, we compare the diapycnal diffusivities from multiple GEOS-5 simulations performed without data assimilation (G-CTRL and DIFF) with those from a reanalysis product that uses the same underlying modeling system called the Global Modeling and Assimilation Office sub-seasonal to seasonal (GMAO S2S) Ocean Analysis. This comparison highlights how the diapycnal diffusivities can behave due to the disruption of dynamical balance that can be the result of the use of a sequential data assimilation system (Stammer et al., 2016; Pilo et al., 2018). The GMAO S2S Ocean Analysis is used to demonstrate what can happen to the diapycnal diffusivities when only hydrographic information is assimilated using a sequential data assimilation framework.

The NASA GMAO has recently updated their GEOS-5 sub-seasonal to seasonal forecast system (S2S-v2.1; https://gmao.gsfc.nasa.gov/cgi-bin/products/climateforecasts/geos5/S2S_2/index.cgi). This new system is the current contribution of the GMAO to the North American Multi-Model project (http://www.cpc.ncep.noaa.gov/products/NMME/about.html) and NOAA's experimental sub-seasonal ensemble project (http://cola.gmu.edu/kpegion/subx/index.html). A configuration of the modeling system is used that is nominally 0.5^o resolution on a tripolar (Murray, 1996) staggered Arakawa B-grid (Mesinger & Arakawa, n.d.) and 40 geopotential levels for MOM5, and 0.5^o resolution and 5 layers for CICE4.1 with atmospheric forcing from MERRA-2 (Modern-Era Retrospective analysis for Research and Applications, Version 2) reanalysis (Gelaro et al., 2017). The GMAO S2S Ocean Analysis (Molod et al., 2020) is a reanalysis product that uses a system similar to the Local Ensemble Transform Kalman Filter (LETKF) data assimilation procedure described by Penny et al. (2013), but where the background error is calculated offline using ensemble members of freely coupled simulations. The background error does not explicitly account for uncertainties in the ocean mixing parameters, as it is only a function of the observed and background temperatures and salinities. The temperature and salinity would change and so would the calculated covariances if the mixing parameterizations were changed, but each of the 21 background free-running simulations (BKG) have the same mixing parameterization, as they only differ in their initialization.

The following datasets were used by the GMAO S2S data assimilation modeling system. A relaxation procedure, or update, is applied towards the MERRA-2 sea surface temperatures and sea ice fraction from the NASA TEAM-2 product (Markus & Cavalieri, 2009) at a 5-day assimilation cycle. No ocean mixing parameter data are assimilated. Assimilated in situ observational data that provide temperatures and salinities come from TAO, PIRATA, RAMA, XBT, CTD, and Argo instruments. Satellite altimetry data that provide sea level anomalies come from TOPEX, ERS-1+2, Geosat FO, Jason-1, Jason-2, Jason-3, Envisat, Cryosat-2, Saral, HY-2A, and Sentinel 3A. The absolute dynamic topography is calculated as the sum of the sea level anomaly and the mean dynamic topography, which is estimated using GOCE and GRACE data, all available altimetry, and in situ data. Absolute dynamic topography data are assimilated into the model system using the same method as for the in situ data, except these data are thinned along-track and a Gaussian weighted mean using a decorrelation scale of 1000 km is calculated prior to assimilation. In addition, the global trend was removed from the absolute dynamic topography before assimilation and zero net input of water was applied. Precipitation is corrected using the Global Precipitation Climatology Project version 2.1 (GPCPv2.1, provided by the NASA/Goddard Space Flight Center's Laboratory for Atmospheres, which calculates the dataset as a contribution to the GEWEX Global Precipitation Climatology Project) and Climate Prediction Center (CPC) Merged Analysis of Precipitation (CMAP, provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their website at http://www.esrl.noaa.gov/psd/), as described by Reichle et al. (2011) except for MERRA-2 instead of MERRA data. All other atmospheric forcing fields used in the construction of the reanalysis came from MERRA-2. The GMAO S2S modeling system is an update to the one described in Borovikov et al. (2017). As such, the model only ran for the period: May of 2012 to March of 2019.

2.3.2 ECCO

The other model system used here is the ECCO-Production, version 4 in revision 3 (ECCOv4r3; *Fukumori et al.*, 2017). The underlying ocean-sea ice model is based on the Massachusetts Institute of Technology general circulation model (MITgcm), which is a global finite volume model. The ECCOv4r3 global configuration uses curvilinear Cartesian coordinates (*Forget et al.*, 2015a - see their Figs. 1-3) at a nominal 1^o (0.4^o at equator) resolution and rescaled height coordinates (Adcroft & Campin, 2004) with 50 vertical levels and a partial cell representation of bottom topography (Adcroft et al., 1997). The MITgcm uses a dynamic/thermodynamic sea ice component (*Menemenlis et al.*, 2005; *Losch et al.*, 2010; *Heimbach et al.*, 2010) and a nonlinear free surface with freshwater flux boundary conditions (Campin et al., 2004). The wind speed and wind stress are specified as 6-hourly varying input fields over 24 years (1992-2015). There are 14-day adjustments to the wind stress, wind speed, specific humidity, shortwave downwelling radiation, and surface air temperature. These adjustments are based on estimated prior uncertainties for the chosen atmospheric reanalysis (Chaudhuri et al., 2013), which is ERA-Interim (Dee et al., 2011). The net heat flux is then computed via a bulk formula (Large & Yeager, 2009). A parameterization of the effects of geostrophic eddies (Gent & McWilliams, 1990) is used. Mixing along isopycnals is according to the framework provided by Redi (1982). Vertical mixing–diapycnal plus the vertical component of the along-isopycnal tensor–is determined according to the Gaspar et al. (1990) mixed layer turbulence closure and simple convective adjustment.

Initial conditions and model parameters for the runs performed here are from ECCOv4r3. The least squares problem solved by the ECCO model uses the method of Lagrange multipliers through iterative improvement, which relies upon a quasi-Newton gradient search (Nocedal, 1980; Gilbert & Lemarechal, 1989). Algorithmic (or automatic) differentiation tools (Griewank, 1992; Giering & Kaminski, 1998) have allowed for the practical use of Lagrange multipliers in a time-varying non-linear inverse problem such as the one for the ocean because the discretized adjoint equations no longer need to be explicitly hand-coded. Contributions of observations to the model-data misfit function are weighted by best-available estimated data and model representation error variance (Wunsch & Heimbach, 2007). The observational data assimilated into the ECCO state estimate are discussed in Forget et al. (2015) and Fukumori et al. (2017). These data include satellite-derived ocean bottom pressures, sea ice concentrations, sea surface temperatures, sea surface salinities, sea surface height anomalies, and mean dynamic topography, as well as profiler- and mooring-derived temperatures and salinities (Fukumori et al., 2017). No ocean mixing parameter or biogeochemical tracer data are used in the ECCO assimilation. The control variables that are inverted and optimized for by ECCO include the initial condition of the sea surface heights, ocean velocities, temperatures, and salinities; time-mean three-dimensional distribution of along-isopycnal diffusion (Redi coefficients–*Redi*, 1982), Gent-McWilliams (Gent & McWilliams, 1990) coefficients, and diapycnal diffusivities (Gaspar et al., 1990); and time-varying two-dimensional surface forcing fields. Fifty-nine iterations in the optimization run of ECCO were performed to arrive at the ECCOv4r3 solution we start from for our adjoint sensitivity experiments. ECCO avoids some pitfalls of sequential data assimilation systems because adjustments are only applied to the input parameters and ocean-sea ice state evolves through the entire model trajectory (1992–2015) without added artificial sources/sinks.

There are three ways to run ECCO: 1) an optimization run of the model in forward plus adjoint modes, where data are assimilated and new values of the models input "control parameters" (ocean mixing parameters, initial conditions, and forcing fields) are estimated; 2) an adjoint sensitivity run of the optimized state estimate in forward plus adjoint modes, where data are included in the cost function but not technically "assimilated" because the model input parameters do not change; and 3) a re-run of the optimized state estimate in forward mode, like most ocean models except all control parameters are set to be their estimated values from the optimization run. We perform (2) and (3) in this study.

In order to assess whether the assimilation of a particular data set would lead to a more accurate estimate of ocean mixing parameter K (either κ_ρ or κ_{Redi}), two conditions must be satisfied. The first condition is that the observationally-derived ocean mixing parameter K has a smaller bias with respect to independent observations than the model's estimate of K. We devote the first portion of our study to determining whether this is true for the diapycnal diffusivities. We use microstructure to assess whether the model-calculated diapycnal diffusivities (unconstrained) have smaller biases when nudged to be closer to $\kappa_{\rho,W15}$ (constrained) than they would without the nudging; i.e.,

$|\kappa_{\rho,unconstrained} - \kappa_{\rho,micro}|/\sigma_\kappa \leq |\kappa_{\rho,constrained} - \kappa_{\rho,micro}|/\sigma_\kappa$, for uncertainty in the observationally-derived values σ_κ. We do not assess this first condition for observationally-derived Redi coefficients due to the dearth of independent observations and the magnitudes of their uncertainties.

The second condition is that the "adjoint sensitivities" from two different experiments have the same sign in the majority of locations. An adjoint sensitivity is essentially the sensitivity of one variable to another, computed by making use of the model's adjoint. Formally, an adjoint sensitivity is $\partial J/\partial X$, where the cost function J is a sum of weighted misfits to observations and a control variable X is a variable that the model estimates by making use of its adjoint and observations–see Section 2.3.2.1. The adjoint sensitivities provide information about which directions the model should change X in order to minimize J (see below). The experiments performed in this study always use $X = K$, one of the ocean mixing parameters, but X could be a different variable. To gauge the adjoint sensitivities, we perform new experiments that include observationally-derived ocean mixing parameters–from either a finescale parameterization or mixing length theory–in ECCO's cost function. One of these experiments compares observationally-derived ocean mixing parameters with the ECCOv4r3 solution's ocean mixing parameters. The other experiment compares observed with simulated biogeochemical oceanic tracer distributions. This is repeated for three different biogeochemical tracers to see whether any of these tracers provide information about ocean mixing–along or across isopycnals. Several tracers are simulated using Biogeochemistry with Light, Iron, Nutrients and Gases (BLING) model (Galbraith et al., 2015). BLING is an intermediate complexity biogeochemistry model that uses several prognostic tracers and parameterized, implicit representations of iron, macronutrients, and light limitation and photoadaptation. BLING has been shown to compare well with the Geophysical Fluid Dynamics Laboratory's full-complexity biogeochemical model, TOPAZ (Galbraith et al., 2015), and has been adapted for use in the MITgcm with its adjoint (Verdy & Mazloff, 2017).

The following is a summary of the ECCO experiments we run:

- **E-CTRL** - a forward ECCOv4 simulation that uses the parameters from ECCOv4r3; this simulation can be referred to as a "re-run"
- **Dmisfit** - an adjoint sensitivity (with respect to $X = log_{10}(\kappa_\rho)$) experiment in which only the base-10 logarithm of the $\kappa_{\rho,W15}$ and $\kappa_{\rho,K17}$ products are included as observations and compared to the ECCOv4r3 solution's diapycnal diffusivities in J
- **Rmisfit** - an adjoint sensitivity (with respect to $X = \kappa_{Redi}$) experiment in which only the $\kappa_{Redi,C15}$ product is included as observations and compared to the ECCOv4r3 solution's Redi coefficients in J
- **Omisfit** - an adjoint sensitivity (with respect to $X = log_{10}(\kappa_\rho)$ and $X = \kappa_{Redi}$) experiment in which only oxygen concentrations from the World Ocean Atlas (2013) climatology are included as observations and compared to those simulated using BLING with the ECCOv4r3 solution in J
- **Amisfit** - an adjoint sensitivity (with respect to $X = log_{10}(\kappa_\rho)$ and $X = \kappa_{Redi}$) experiment in which only alkalinities from the World Ocean Atlas (2013) climatology are included as observations and compared to those simulated using BLING with the ECCOv4r3 solution in J
- **Pmisfit** - an adjoint sensitivity (with respect to $X = log_{10}(\kappa_\rho)$ and $X = \kappa_{Redi}$) experiment in which only phosphate concentrations from the World Ocean Atlas (2013) climatology are included as observations and compared to those simulated using BLING with the ECCOv4r3 solution in J

We take the ECCOv4r3 solution as initial conditions and perform an adjoint calculation for each of the five experiments. Only one year was run for each of these simulations because we are using time-invariant climatologies, and one year suffices to demonstrate the point that the assimilation of a biogeochemical tracer may reduce the bias in the ocean mixing parameter estimates. The adjoint sensitivities from Dmisfit and Rmisfit are not sensitive to their initial conditions or run length due to the lack of time-dependence of the ocean mixing parameters. While the adjoint sensitivities from Omisfit, Amisft, and Pmisfit are sensitive to initial conditions, we begin from a previously-derived product that has been spun-up from an initial GLobal Ocean Data Analysis Project version 2 (GLODAPv2) climatology (Dutkiewicz et al., 2005) and our results are not qualitatively sensitive to the run length. It is important to note that a base-10 logarithm of the diapycnal diffusivities–which are positive definite–is taken in each simulation, which stabilizes the numerics of the model and reduces the adjoint sensitivities relative to using the untransformed diapycnal diffusivities.

2.3.2.1 Adjoint sensitivity analyses

In order to further understand whether ocean mixing parameters could be estimated more accurately through data assimilation of biogeochemical tracers, we perform multiple adjoint sensitivity experiments with ECCO. We define the objective (or cost) function here to more formally explain what the adjoint sensitivity is. ECCO calculates the cost function to be minimized, J, as (Stammer et al., 2002):

$$J = \sum_{t=1}^{t_f} [\mathbf{y}(t) - \mathbf{E}\tilde{\mathbf{x}}(t)]^T \mathbf{W}(t) [\mathbf{y}(t) - \mathbf{E}\tilde{\mathbf{x}}(t)] \tag{3}$$

where t_f is the final time step, $\tilde{\mathbf{x}}$ is the model-based estimate of the state vector ($\mathbf{x}$), $\mathbf{E}$ is the observation matrix that relates the model state vector to observed variables (such that $\mathbf{E}\tilde{\mathbf{x}}$ is the model-based estimate of the observables $\mathbf{y}$), and $\mathbf{W}$ is the weight (inverse square of the uncertainty) of the observations. In each of our adjoint sensitivity experiments, the misfit to a single data set is included in the cost function; all other terms in the cost function are zero.

While the adjoint sensitivities from Omisfit, Amisfit, and Pmisfit must be computed online using ECCO, the adjoint sensitivities from Dmisfit and Rmisfit can either be computed online using ECCO or come from using an analytical equation offline. The adjoint sensitivities computed in this study are the derivatives of J in Eq. 3 with respect to one of the ocean mixing parameters: the base-10 logarithm of the diapycnal diffusivity ($\log_{10}(\kappa_\rho)$) or the Redi coefficient (κ_{Redi}). The adjoint sensitivity runs with the ocean mixing parameters included in the misfit calculation (Dmisfit and Rmisfit) have adjoint sensitivities that can be computed offline (i.e., using the output of E-CTRL instead of running Dmisfit or Rmisfit), using:

$$\frac{\partial J}{\partial K} = -2\frac{(K_{obs} - K_{model})}{\sigma_K^2}. \quad (4)$$

Here, K is either κ_{Redi} or $\log_{10}(\kappa_\rho)$, $\mathbf{y} = K_{obs}$ is the observationally-derived value of K described in the previous section, $\mathbf{E}\tilde{\mathbf{x}} = K_{model}$ is the value that ECCO estimates for K, and σ_K^2 (entries of $\mathbf{W}$) is taken to be $3{*}K_{obs}$ (or the base-10 logarithm of this in the case of the diapycnal diffusivities) due to the factor of 2-3 uncertainty. The offline adjoint sensitivities using Eq. 4 and the adjoint sensitivities using ECCO have been verified to be in agreement. Values of $\kappa_{\rho,ECCO}$ and $\kappa_{Redi,ECCO}$ from ECCOV4r3 are used for K_{model} in these adjoint sensitivity simulations and offline calculations (Eq. 4).

The main finding of this study comes from our test to see whether $\partial J/\partial K$ in Dmisfit and in Rmisfit is of the same sign as $\partial J/\partial K$ in Omisfit, Amisfit, and/or Pmisfit. For example, say that the $\kappa_{\rho,W15}$ and $\kappa_{\rho,K17}$ products are in close agreement with microstructure-inferred diapycnal diffusivities. Then if Dmisfit and Omisfit each show that $\partial J/\partial \log_{10}(\kappa_\rho) < 0$ (i.e., the diapycnal diffusivities need to be increased to lower the cost) in the same locations, then it is preferable to assimilate the more accurately known oxygen concentrations instead of the diapycnal diffusivities in a new optimization. Pmisfit and Omisfit are expected to provide similar information because of the phosphate to oxygen Redfield ratio, but we test this expectation by including Pmisfit here. Note that $\partial J/\partial \kappa_{Redi} = 0$ in Dmisfit and $\partial J/\partial \log_{10}(\kappa_\rho) = 0$ in Rmisfit because J is a function of only one of the ocean mixing parameters in each experiment (i.e., no other observations are included in the cost function) and each ocean mixing parameter is simply read in, as opposed to dynamically calculated.

In order to compare the adjoint sensitivities across different experiments, a normalization factor must be computed. After weighting by the grid cell volume to make each grid cell comparable to another, the adjoint sensitivities can be normalized in two ways. One way is to non-dimensionalize the sensitivities by multiplying them by a representative value for the variable the sensitivity is taken with respect to and then weighting by the inverse square of an estimate of the temporal variability in the field computed in the misfit calculation. The second way to normalize the adjoint sensitivities is to simply divide the adjoint sensitivities by the cost function of each respective experiment. We choose to use this second method (results presented in Section 3.5), but the first method produces qualitatively similar results. Table 4 tabulates the data sources, described in Section 2.1, and cost functions used to normalize the adjoint sensitivities for each ECCO experiment, summarized in Table 1 and Section 2.3.

3 Results

Our first goal is to determine if using an observationally-derived diapycnal diffusivity from the finescale parameterization reduces biases in the diapycnal diffusivity field with respect to independent observational data. To address this goal, we take one direct approach–through comparison with microstructure observations–and the other indirect approach–involving the adjustment of the diapycnal diffusivities in the GEOS-5 simulations. Next, we use the observationally-derived diapycnal diffusivities from the finescale parameterization and Redi coefficients from mixing length theory to investigate whether biogeochemical tracers could be assimilated to better estimate ECCO's ocean mixing parameters in a future optimization at global 1^o resolution. Specifically, we run several adjoint sensitivity experiments in which either an ocean mixing parameter or a biogeochemical tracer is included in the misfit calculation to guide constraints on ocean mixing parameters.

3.1 Assessments of diapycnal diffusivities from models and finescale parameterization

Previous studies have shown that $\kappa_{\rho,micro}$ and $\kappa_{\rho,W15}$ agree within a factor of 2-3 and exhibit no systematic high or low bias in open ocean conditions from below the mixed layer to a depth of 2 km (Whalen et al., 2015). Here, we compare the average model-calculated profiles–with and without nudging to the $\kappa_{\rho,W15}$ product–and the average $\kappa_{\rho,micro}$ profile that is comprised of 24 campaigns worth of data *Waterhouse et al.* (2014 - see their Fig. 6; black curve in Fig. 1). A geometric average is taken for each profile because a geometric average is more representative than an arithmetic average for a small sample size and when the data are not normally distributed (Manikandan, 2011), like the log-normal distribution of diapycnal diffusivities.

In general, a dearth of mixing at intermediate depths ($250-1500$ meters depth) and at abyssal depths (> 3500 meters depth) is found in the ECCO solutions. The initial guess (pre-optimized) values (grey curve in Fig. 1) are even smaller than the ECCOv4r3 solution: E-CTRL (red curve in Fig. 1). When $\kappa_{\rho,ECCO}$ is adjusted towards $\kappa_{\rho,W15}$, the average profiles from the model and $\kappa_{\rho,micro}$ in the upper 2000 meters agree more closely (blue curve in Fig. 1). The blue curve sits on top of the red curve in Fig. 1 below 2000 meters, by construction.

$\kappa_{\rho,micro}$ is also compared with an averaged diapycnal diffusivity profile from G-CTRL, DIFF, and a reanalysis product (the GMAO S2S Ocean Analysis). The diapycnal diffusivity profiles from the G-CTRL and DIFF simulations (red and blue curves in Fig. 2) and from the GMAO S2S Ocean Analysis (green curve in Fig. 2) are sampled at the same locations as the microstructure observations. Nudging towards $\kappa_{\rho,W15}$ tends to increase $\kappa_{\rho,MOM5}$ between $750-1750$ meters depth. As a result, when nudged towards $\kappa_{\rho,W15}$, the model's average profile does not agree better with $\kappa_{\rho,micro}$ between $750-1750$ meters depth (red and blue curves in Fig. 2), likely due to differences in spatial coverage between the Argo and microstructure observations. However, on average, the disagreement with $\kappa_{\rho,micro}$ is no worse in the full adjusted model-calculated diapycnal diffusivity profile than in the full unadjusted model-calculated average diapycnal diffusivity profile. The differences between the full adjusted model-calculated and $\kappa_{\rho,micro}$ profiles are well within the uncertainty of the $\kappa_{\rho,W15}$ product.

While the average diapycnal diffusivity profile in the model is fairly accurate, particularly below 500 meters depth, in each of the GEOS-5 simulations we ran without the GMAO's data assimilation system (red and blue curves in Fig. 2), the GMAO S2S Ocean Analysis product has diapycnal diffusivities that are too small below (large above) about 500 meters depth (green curve in Fig. 2). Potential reasons for this discrepancy include dynamical adjustments due to the analysis increments, or inconsistencies between the model's atmosphere and ocean due to the strong relaxation to sea surface temperatures, and fixed zero net water input for global sea level. We only include the GMAO S2S Ocean Analysis result here to suggest that data assimilation systems, particularly ones that are based on filter-based sequential data assimilation methods, may require stronger constraints on their diapycnal diffusivities to prevent them from becoming too unrealistic. One way to do this is to assimilate ocean mixing parameters. Another possible method is to assimilate a biogeochemical tracer, which is proposed later in this study.

3.2 Model- vs finescale parameterization-derived ocean mixing parameter comparisons

We next present the $\kappa_{\rho,W15}$ product (Figs. 3a,d,g) and the percent differences between their product and ECCO-estimated diapycnal diffusivities ($\kappa_{\rho,ECCO}$; Figs. 3b,e,h). Blue regions in Figs. 3b,e,h indicate where $\kappa_{\rho,ECCO}$ are too small and red regions indicate where $\kappa_{\rho,ECCO}$ are too large. The regions with the largest disagreement below the mixed layer are in the Atlantic and Indian sectors of the Southern Ocean, the tropical Pacific Ocean, the Atlantic Ocean, and the Kuroshio Extension between 500-1000 meters depth (Figs. 3b,e,h). The values of $\kappa_{\rho,ECCO}$ are smaller than those in the observational product in the Kursoshio Extension (500-1000 meters depth), subpolar North Atlantic (500-1000 meters depth), Southern Ocean, and equatorial regions in the Atlantic and shallow (250-500 meters depth) Indian and eastern Pacific Oceans (Figs. 3b,e,h). The errors in $\kappa_{\rho,ECCO}$ could be partially compensating for errors in the vertical component of the along-isopycnal diffusivity tensor and/or numerical diffusion (see later).

The base-10 logarithm of the diapycnal diffusivity field at different depth levels from the $\kappa_{\rho,W15}$ product (Figs. 3a,d,g) is also compared with the time-averaged GEOS-5-calculated diapycnal diffusivity field ($\kappa_{\rho,MOM5}$; Figs. 3c,f,i). The sign of the discrepancy between the values of $\kappa_{\rho,MOM5}$ and the observations varies regionally, but the disagreements tend to be smaller than those for $\kappa_{\rho,ECCO}$. The regions with the largest disagreement are along the equator, in the Southern Ocean, in the Labrador and Irminger Seas, and in the Gulf Stream and Kuroshio Extensions (Figs. 3c,f,i). Along the equator the values of $\kappa_{\rho,MOM5}$ tend to be larger than the observational product, but the discrepancy changes sign slightly poleward in the near-equator tropics. The values of $\kappa_{\rho,MOM5}$ are smaller than the observations both in regions where deep convection is prevalent and in the vicinity of the Antarctic Circumpolar Current (ACC). In the Gulf Stream Extension region, the Malvinas Current region, part of the Kuroshio Extension region, and the Indian Ocean sector of the ACC above 500 meters depth, the values of $\kappa_{\rho,MOM5}$ are too large because the mixed layer depth can be deeper than 250 meters. In these regions, the model-calculated diapycnal diffusivities can be much increased due to vertical convection. One likely source of these errors in the abyssal diapycnal diffusivities is the improper treatment of remote internal tide-induced mixing, discussed in Melet et al. (2016), but several other processes, such as the wind-driven near-inertial waves (Alford et al., 2016), can impact the diapycnal diffusivities in the upper water column. MacKinnon et al. (2017) discusses other candidates for more accurate representation of ocean mixing. The values of $\kappa_{\rho,ECCO}$ could be worse than those of $\kappa_{\rho,MOM5}$ in comparison to $\kappa_{\rho,micro}$ and $\kappa_{\rho,W15}$ because $\kappa_{\rho,ECCO}$ is primarily constrained by assimilated hydrographic observations, which are sparse below 2000 meters depth and likely insufficient in near-coastal areas, where internal wave-induced mixing can be important.

We next compare the Redi coefficient field from the $\kappa_{Redi,C15}$ product (Figs. 4a,d,g) and the percent differences between their product and the ECCO-estimated Redi coefficients ($\kappa_{Redi,ECCO}$; Figs. 4b,e,h). As in Figs. 3b,e,h the regions that are red in Figs. 4b,e,h are locations where $\kappa_{Redi,ECCO}$ are too small and blue regions are where $\kappa_{Redi,ECCO}$ are too large. The regions with the largest disagreement are along the equator, in intensified jet regions, and in the Labrador and Irminger Seas (Figs. 4b,e,h). The values of $\kappa_{Redi,ECCO}$ are too large in the Kuroshio Extension and subpolar North Atlantic Ocean (Figs. 4b,e,h). In most other locations, $\kappa_{Redi,ECCO}$ are too small. The exaggeration of Redi coefficients in western boundary current regions and underestimates of Redi coefficients elsewhere in ECCOv4r3 are likely compensating for errors in other model parameters, such as the Gent-McWilliams coefficients, and further arises due to errors in horizontal gradients of dynamical fields such as salinity. The Gent-McWilliams coefficient can impact horizontal gradients due to its impact on the slope of isopycnals, and the horizontal gradients determine the slope of the tensor that sets the direction in which the Redi coefficients diffuse tracers. This makes Redi coefficients susceptible to errors in Gent-

McWilliams coefficients and vice-versa; an error in one parameter may be the result of a trade-off in errors in another parameter.

We also present the base-10 logarithm of the Redi coefficient field from the $\kappa_{Redi,C15}$ product via mixing length theory (Figs. 4a,d,g) and the base-10 logarithm of the ratios of the assumed Redi coefficient field (600 m^2 s^{-1}) of the GEOS-5 model to the $\kappa_{Redi,C15}$ product (Figs. 4c,f,i). Assuming that the $\kappa_{Redi,C15}$ product is accurate, the regions that are red in Figs. 4c,f,i are locations where the model's ocean mixing parameters are too large and blue regions are where the model's ocean mixing parameters are too small. The values $\kappa_{Redi,MOM5} = 600$ m^2 s^{-1} of the GEOS-5 model are too small in the upper 2000 meters of every region except for the North Pacific Ocean and Weddell Sea. The largest disagreements occur in the jets (Figs. 4c,f,i). While it is well-known that the Redi coefficients should not be constant (R. P. Abernathey & Marshall, 2013; Forget et al., 2011), the observational bias and uncertainty in the Redi coefficient field is not very well-known. For example, Roach et al. (2018) found values for the Redi coefficients that are a factor of 2-3 less than those of $\kappa_{Redi,C15}$ at 1000 meters depth when drifter observations were used instead of high resolution ECCO2 output. Because the order of magnitude disagreement shown in many regions of Figs. 4c,f,i is larger than this approximate factor of 2-3 bias and uncertainty in the $\kappa_{Redi,C15}$ product, the Redi coefficient estimates may improve data assimilation if their uncertainties are accounted for.

3.3 Dynamical impacts on GEOS-5

We compare the model output from our simulations with and without $\kappa_{\rho,t}$ substituting the model's diapycnal diffusivity field to assess whether the diapycnal diffusivities derived using the finescale parameterization has smaller biases than the model-calculated diapycnal diffusivities. Hereafter, we refer to the difference between the model-calculated diapycnal diffusivity field, $\kappa_{\rho,MOM5}$ and $\kappa_{\rho,t}$ to be the "adjustment" $\Delta\kappa_{\rho,MOM5}$. We next show that GEOS-5 is improved by using the adjustment $\Delta\kappa_{\rho,MOM5}$, which suggests that diapycnal diffusivity products can be derived using the finescale parameterization and used to constrain diapycnal diffusivities in models.

This internal variability of the GEOS-5 modeling system is first shown here and compared with the changes from applying $\Delta\kappa_{\rho,MOM5}$. The spread in the vertically and zonally averaged anomalies in temperature (Fig. 5a) and salinity (Fig. 5b) relative to the ensemble mean from the 21 free-running simulations that only differ in their initial conditions (BKG) is first compared with the difference in temperature and salinity from use of $\Delta\kappa_{\rho,MOM5}$. Each of the GEOS-5 simulation results were averaged over their final five years. Finding one time period where the changes in the GEOS-5 simulations are larger than the spread in the BKG anomalies is sufficient to show that the internal variability associated with initial conditions is smaller than that associated with adjusting the model's diapycnal diffusivity field. Changing the diapycnal diffusivities can lead to vertically and zonally averaged temperature (salinity) differences in DIFF relative to G-CTRL. These differences can be greater than 0.1^{o}C (0.05 PSS-1978) in some high latitude regions, which is greater than any of the anomalies in BKG (Fig. 5). In the subpolar North Atlantic Ocean, use of $\Delta\kappa_{\rho,MOM5}$ induces the largest temperature and salinity changes, each well beyond the level of internal variability. Use of $\Delta\kappa_{\rho,MOM5}$ also alters the temperature and salinity by more than the level of internal variability in other high latitude regions and change the salinity beyond the level of internal variability in the tropics. These findings demonstrate that the adjustments in ocean mixing can induce changes in temperature and salinity that are larger than the internal variability of the model.

Next, we assess whether using $\Delta\kappa_{\rho,MOM5}$ causes changes to the temperature and salinity that improves the free-running modeling system relative to a number of climatologies. Local changes in vertical heat and salt transport lead to convergences and divergences of heat and salt, which influences the temperature (Fig. 5a), salinity (Fig. 5b), and stratification (Hieronymus et al., 2019) but this can lead to greater agreement or disagreement with observational climatologies. At least three factors explain why the temperature, salinity, and stratification fields can sometimes disagree more with observations at some locations: the spatiotemporal gaps in the diapycnal diffusivity substitutions, feedbacks as a result of air-sea flux changes in regions with deep convection (Wang et al., 2018; W. Liu et al., 2019; Putrasahan et al., 2019; Kostov et al., 2019), and not accounting for differences in diapycnal diffusivities of temperature and salinity in the finescale parameterization. On average, the temperature, salinity, and stratification fields each agree more with observational climatologies in DIFF than in G-CTRL over the last 16 years of each simulation (2001-2016). The mean-square error relative to Levitus and et al. (2012) observations in temperature (salinity) over the upper 1500 meters is 0.81^oC^2 (0.010 g^2 kg^{-2}) in G-CTRL and 0.27% smaller (1.0% smaller) in DIFF. Also, the mean-square error relative to Levitus and et al. (2012) observations in stratification over the upper 1500 meters is $1.14 \times 10^{-10}s^{-2}$ in G-CTRL and 0.36% smaller in DIFF. These mean-square errors are dependent on the time period used in our simulations, but over most continuous subsets of the final 16 years of our simulations, the mean-square errors are smaller in DIFF than in G-CTRL.

We additionally use diagnostics from our simulations that account for atmosphere-ocean feedbacks (G-CTRL shown in Figs. 6a and 7a) and compare them with their equivalent observational climatologies: the mixed layer depths from Holte et al. (2017) and the sea surface temperatures from Reynolds et al. (2007). The sea surface temperature changes and coinciding sea level pressure changes due to substituting $\kappa_{\rho,MOM5}$ with $\kappa_{\rho,t}$ are shown in Fig. 7b. The locations with blue coloring shown in Figs. 6b and 7c are improved relative to a given observational product when $\kappa_{\rho,MOM5}$ is substituted with $\kappa_{\rho,t}$. The maximum yearly mixed layer depths and sea surface temperatures are mostly improved upon adjustment of the diapycnal diffusivity field (Figs. 6b and 7c). The largest maximum yearly mixed layer depths changes occur in the Norwegian Sea–more than 50 meters deeper in DIFF than in G-CTRL–because deep convection is altered there (not shown). These changes and other smaller ones–such as improvements in most equatorial regions, in subtropical gyres, and in the vicinity of intensified jets–are improvements almost everywhere in the maximum yearly mixed layer depths (Fig. 6b). Diapycnal diffusivity changes at depth also have consequences at the surface, even though the diapycnal diffusivity field is never altered above the mixed layer depth (Fig. 7b). The effects on sea surface temperature are particularly pronounced in the Southern Ocean where upwelling occurs and the diapycnal diffusivity changes tend to be deeper due to deeper mixed layer depths. There is a hemispheric dipole pattern in the sea surface temperature changes, which aligns well regionally with sea level pressure changes (Fig. 7b). This suggests that some of the surface flux changes (Fig. 8e) due to adjusting the diapycnal diffusivities are caused by both sea surface temperature alterations and atmospheric circulation differences. The sea surface temperature changes are mainly improvements (Fig. 7c), which tend to lie within distinct regions where the diapycnal diffusivities were changed at depth. The margins of these regions of improvement see degraded sea surface temperatures relative to Reynolds et al. (2007). The gold contours in Fig. 7c indicate the depth- and time-averaged $\Delta\kappa_{\rho,MOM5}$ field, which tend to line up with the improved/degraded agreement patterns more closely in the Northern Hemisphere because these are primarily regions where deep water formation occurs. Thus, use of $\Delta\kappa_{\rho,MOM5}$ improves the mixed layer depth and sea surface temperature fields by changing diapycnal diffusion at the base of the mixed layer, which alters the sea surface temperatures and can then cause atmosphere-ocean feedbacks.

3.4 Steric sea level impacts in GEOS-5

Next, we analyze the steric sea level budget, as described earlier, in order to better understand how the diapycnal diffusivity adjustments change the dynamics. Since the thermal expansion coefficient and haline contraction coefficient vary with depth, any changes in the diapycnal diffusivities will alter the vertical transport of heat and therefore the steric sea level.

The simulation that uses $\kappa_{\rho,MOM5}$ (G-CTRL) is first discussed. Similarities between the steric sea level budget's vertical diffusion term and the steric sea level budgets have been described by previous studies. Using the same ocean model, but different atmospheric and sea ice models, Palter et al. (2014) found that vertical diffusion and surface flux terms dominate the steric sea level budget. Consistent with the findings of Palter et al. (2014), the resolved advection, neutral and vertical diffusion, and surface flux terms are among the most locally important physical terms in the steric sea level budget (Figs. 8a-c; vertical diffusion not shown; Table 3). The resolved advection term globally volume-averages to nearly zero, but not exactly zero partially due to the downward resolved heat advection below 2000 meters depth. These findings are also in agreement with *Hieronymus and Nycander* (2013). The resolved advection term also has a large amount of spatiotemporal variability (Table 3), consistent with the findings of Piecuch and Ponte (2011, 2014). The largest regional tendency terms in the GEOS-5 simulation's steric sea level budget are the resolved advection, surface heat flux, vertical diffusion, and neutral diffusion terms (Figs. 8a-c). Of lesser importance to the regional steric sea level budget are the remaining numerical and parameterized terms, precipitation minus evaporation, and contributions from (land and sea) ice.

When $\kappa_{\rho,t}$ is used in another simulation (DIFF), the resolved advection term locally changes by nearly 10% (Fig. 8d), with the other terms changing by less than 1% (Figs. 8e-f). The resolved advection and neutral diffusion term changes look similar, as they are largest in the vicinity of subtropical gyres (Figs. 8d,f). However, the globally averaged resolved advection term changes by about 10% and the globally averaged neutral diffusion term changes by less than 1%. This is because the resolved advection and neutral diffusion terms depend upon the geostrophic velocities, which are altered due to the differences in the vertical transport of heat and changes in isopycnal slopes, but the neutral diffusion term also depends upon the Redi coefficients, which do not change. While locally the steric sea level tendencies can change by $> 100\%$ due to use of $\Delta\kappa_{\rho,MOM5}$ (not shown), the globally averaged steric sea level tendency is increased by 5.35% in DIFF relative to G-CTRL. This global change is dominated by the changes in the resolved advection term. The largest surface flux term changes are in tropical regions and in jet regions (Fig. 8e) with a globally averaged change of less than 1%. The changes in some of these terms due to substituting the diapycnal diffusivity field in a time-varying manner can be larger than equivalent terms in the GMAO S2S Ocean Analysis (e.g., some coastal current regions in Figs. 8d,f; not shown for the reanalysis because the analysis increment is comparably large as the resolved advection term, which confounds physical interpretation). Thus, the diapycnal diffusivity adjustments primarily affect the resolved advection in the model as a result of mostly redistributing heat and salt, which leads to dynamic adjustment, and secondarily impact heat uptake/loss at the sea surface.

3.5 Adjoint sensitivities in ECCO

3.5.1 Diapycnal diffusivity: κ_ρ

We describe results from the adjoint sensitivity calculation using Eq. 4 for the base-10 logarithm of diapycnal diffusivity ($\log_{10}(\kappa_\rho)$) misfits (Dmisfit). Because $\kappa_{\rho,W15}$ and $\kappa_{\rho,K17}$ are not normally distributed, we focus on $\log_{10}(\kappa_\rho)$ misfits. A region with black dots with a red plus sign surrounded by a grey contour in Figs. 9-12 implies that the misfit can be decreased by decreasing the ocean mixing parameter because the adjoint sensitivity, $\partial J/\partial \log_{10}(\kappa_\rho)$, is positive. The radii of the circles in Fig. 9a indicate the horizontal extents over which changing $\kappa_{\rho,ECCO}$ can influence the model's misfits, which are primarily determined by the model's resolution; the model's dynamics are less important in determining these radii over short time intervals such as 1992-2015. When physically interpreting the adjoint sensitivities, it should be noted that $\kappa_{\rho,W15}$ and $\kappa_{\rho,K17}$ can obtain values different from 10^{-5} m^2 s^{-1}, but the default value for $\kappa_{\rho,ECCO}$ where there are few observational constraints is 10^{-5} m^2 s^{-1}. Simply because of the chosen default value for $\kappa_{\rho,ECCO}$, some regions with few observations–such as the Arctic Ocean (Chanona et al., 2018)–can have positive adjoint sensitivities and other regions–for example, near the seafloor (Polzin et al., 1997; Waterhouse et al., 2014)–can have negative adjoint sensitivities with respect to the base-10 logarithm of the diapycnal diffusivities ($\partial J/\partial \log_{10}(\kappa_\rho)$) in Dmisfit.

κ_ρ needs to be decreased in many regions at depths shallower than 500 meters to agree better with $\kappa_{\rho,W15}$ and $\kappa_{\rho,K17}$, but the regions where κ_ρ should be increased (dots with red plus signs surrounded by grey contours in Fig. 9a) tend to be in locations where microstructure measurements (used for Fig. 1) were taken. These are regions where coastal wind-driven mixing occurs and the centers of subtropical gyres. Inadequate resolution and parameterization of mixing across isopycnals can cause too little mixing to occur in these regions as well as in the Southern Ocean and along mid-ocean ridges (MacKinnon et al., 2017). $\partial J/\partial \log_{10}(\kappa_\rho)$ tend to be larger at higher latitudes (Figs. 9a,c). $\partial J/\partial \log_{10}(\kappa_\rho)$ are relatively small wherever they are positive, except for regions where the mixed layer can get relatively deep ($\sim$ 1000 meters; Fig. 9a). The signs of $\partial J/\partial \log_{10}(\kappa_\rho)$ shown in Fig. 9a are consistent with the signs of disagreement shown in Fig. 3b, by construction, and those shown in Fig. 9c generally agree with the disagreements with microstructure shown in Fig. 1.

We now compare $\partial J/\partial \log_{10}(\kappa_\rho)$ from Dmisfit with those of the experiments that include biogeochemical tracers in the misfit calculation (Omisfit, Amisfit, Pmisfit). The locations of the positive/negative signs of $\partial J/\partial \log_{10}(\kappa_\rho)$ are not the same everywhere between the Dmisfit and the Omisfit, Amisfit, and Pmisfit experiments, but they generally agree in many regions. The percent of ocean volume where sufficient observations exist to derive an ocean mixing parameter in which the signs of the adjoint sensitivities agree across experiments are tabulated in Table 5. The signs of $\partial J/\partial \log_{10}(\kappa_\rho)$ from Dmisfit and the signs of $\partial J/\partial \log_{10}(\kappa_\rho)$ from Omisfit agree over more than two-thirds of the ocean's volume. This is greater than the percent volume over which there is agreement between $\partial J/\partial \log_{10}(\kappa_\rho)$ from Omisfit and $\partial J/\partial \log_{10}(\kappa_\rho)$ from Pmisfit. The percent volume over which there is agreement between $\partial J/\partial \log_{10}(\kappa_\rho)$ from Dmisfit and $\partial J/\partial \log_{10}(\kappa_\rho)$ from Amisfit and Pmisfit is also smaller. The vast majority of the locations where disagreements occur in the signs of $\partial J/\partial \log_{10}(\kappa_\rho)$ from Dmisfit and the signs of $\partial J/\partial \log_{10}(\kappa_\rho)$ from Omisfit are in places with small differences between $\kappa_{\rho,ECCO}$ and $\kappa_{\rho,W15}$ supplemented with $\kappa_{\rho,K17}$. In Omisfit, $\partial J/\partial \log_{10}(\kappa_\rho)$ are negative almost exclusively in coastal oxygen minimum zones (Wyrtki, 1962; Schmidtko et al., 2017) (Figs. 9b,d). In Amisfit, $\partial J/\partial \log_{10}(\kappa_\rho)$ are negative in most open ocean (non-coastal) regions in the Northern Hemisphere and in about half of the Southern Hemisphere (Figs. 10a,c). In Pmisfit, $\partial J/\partial \log_{10}(\kappa_\rho)$ are negative in most open ocean and many coastal upwelling locations (Figs. 10b,d).

The zonally averaged $\partial J/\partial \log_{10}(\kappa_\rho)$ patterns alternate between positive and negative across latitudes for each experiment, but each experiment tends to agree that diapycnal diffusivities are too small near the seafloor at low Northern Hemisphere latitudes, where internal tide breaking is important (Arbic et al., 2004; Nycander, 2005; Melet et al., 2013; MacKinnon et al., 2017) and beneath the Antarctic Circumpolar Current (ACC), where lee wave breaking is important (Nikurashin & Ferrari, 2011; R. B. Scott et al., 2011; Naveira Garabato et al., 2013; Melet et al., 2014; Wright et al., 2014; Trossman et al., 2013, 2016; Yang et al., 2018). With the exception of Amisfit in the tropical Pacific Ocean, each experiment generally agrees that diapycnal diffusivities are too large in the model's equatorial regions, where the intermittency of tropical instability wave-induced mixing is likely not accounted for with a time-invariant diapycnal diffusivity field. This was unexpected because the fidelity of $\kappa_{\rho,W15}$ supplemented with $\kappa_{\rho,K17}$ is unknown near the equator. All four experiments also generally agree that diapycnal diffusivities are too large in the model's upper several hundred meters and very high latitude regions, where the most pronounced errors in the stratification occur (not shown).

The magnitudes of the normalized $\partial J/\partial \log_{10}(\kappa_\rho)$ attain their maxima in different locations across experiments. The magnitudes are largest close to the Antarctic coast in Amisfit, in many open ocean regions in Dmisfit, and in tropical regions in both Omisfit and Pmisfit. A future optimization of ECCO would not need to change the diapycnal diffusivity field very much in these locations to achieve better agreement with observations, so if both alkalinity and oxygen, for example, were assimilated, then alkalinity (oxygen) would limit the extent to which the diapycnal diffusivity gets changed near the Antarctic coast (coastal oxygen minimum zones). Further, despite expectations, these results suggest that oxygen and phosphate concentrations would not provide similar information about the diapycnal diffusivities in a future optimization of ECCO, except possibly in the tropical ocean's upper 2000 meters. It is suggested here that dissolved oxygen concentrations and possibly other biogeochemical tracers could be used to more accurately estimate ocean mixing parameters in a newly optimized ECCO solution.

3.5.2 Redi coefficient: κ_{Redi}

While we currently do not know how realistic the $\kappa_{Redi,C15}$ product is in the same way as the $\kappa_{\rho,W15}$ product, we repeat the above exercise for the Redi coefficients by first inspecting the adjoint sensitivities with respect to the Redi coefficients ($\partial J/\partial\kappa_{Redi}$) in Rmisfit and later comparing $\partial J/\partial\kappa_{Redi}$ across experiments. Consistent with Figs. 4b,e,h, the ECCO-estimated Redi coefficients ($\kappa_{Redi,ECCO}$) are too small almost everywhere (Figs. 11a,c), which is a function of the resolution of ECCOv4r3 but also due to factors in the mixing length theory (see below). According to Rmisfit, $\partial J/\partial\kappa_{Redi}$ are only positive in the region where deep convection occurs and in the deep Southern Ocean, but we mask out depths deeper than 2000 meters and the highest latitudes because we lack $\kappa_{Redi,C15}$ in these regions (Fig. 11c). The values of $\kappa_{Redi,ECCO}$ are too large (i.e., there are positive adjoint sensitivities) in the Southern Ocean, which could be due to eddy-diffusive transport significantly contributing to southward eddy heat transport (Dufour et al., 2015), the enhancement of mesoscale eddy stirring (R. Abernathey et al., 2013), and/or the larger effects from the nonlinearities in the equation of state (Palter et al., 2014) near the fronts of the ACC. The values of $\kappa_{Redi,ECCO}$ are also too large in the subpolar North Atlantic Ocean, which could be related to how the horizontal eddy diffusivity field influences the overturning circulation (Marshall et al., 2017) and the fraction of North Atlantic Deep Water that gets to the deep Pacific (Jones & Abernathey, 2019). The adjoint sensitivities are larger in the tropics near the surface and in the vicinity of intensified jets (Figs. 11a). The fact that these patterns emerge in locations where horizontal eddy transport is known to be important suggests that at least the spatial patterns of $\kappa_{Redi,C15}$ are insightful.

Lastly, in order to show whether a biogeochemical tracer can be used to help with more accurate estimation of the Redi coefficients in a future optimization of ECCO, we compare the normalized $\partial J/\partial\kappa_{Redi}$ from Rmisfit with those from Omisfit, Amisfit, and Pmisfit. The signs of $\partial J/\partial\kappa_{Redi}$ from Rmisfit and the signs of $\partial J/\partial\kappa_{Redi}$ from Omisfit, Amisfit, and Pmisfit agree over about half of the ocean's volume (Table 5). This is greater than the volume over which there is agreement between $\partial J/\partial\kappa_{Redi}$ from Omisfit and $\partial J/\partial\kappa_{Redi}$ from Pmisfit. The spatial patterns of the signs of $\partial J/\partial\kappa_{Redi}$ are approximately consistent across the experiments with biogeochemical tracer concentrations compared in the misfit calculation (Omisfit, Amisfit, and Pmisfit). However, there are many locations where $\partial J/\partial\kappa_{Redi}$ are positive in Omisfit, Amisfit, and Pmisfit are places where $\partial J/\partial\kappa_{Redi}$ are negative in Rmisfit.

If the magnitudes of the $\kappa_{Redi,C15}$ product are reduced by a factor of 2-3, as Roach et al. (2018) suggest, then more agreement in sign between $\partial J/\partial \kappa_{Redi}$ in Rmisfit and those in Omisfit, Amisfit, and Pmisfit is found (not shown). This suggests that the magnitudes of the $\kappa_{Redi,C15}$ product may be inaccurate. The results of Roach et al. (2018) suggest that this may be due to the dearth of kinetic energy that Cole et al. (2015) used from the ECCO2 product. Canuto et al. (2019) further suggest that the mixing efficiency that Cole et al. (2015) may be too large and recommend deriving the mixing use efficiency using sea surface kinetic energy spectra. However, it is also possible that $\kappa_{Redi,C15}$ are only appropriate for models with horizontal resolutions that are different from 1^o. $\partial J/\partial \kappa_{Redi}$ in Omisfit are largest in coastal areas, intensified jet regions, and a few other open ocean regions (e.g., the Norwegian Sea, the subpolar North Atlantic Ocean, and North Pacific Ocean; Figs. 11b,d). $\partial J/\partial \kappa_{Redi}$ in Pmisfit are largest in similar regions near the surface to $\partial J/\partial \kappa_{Redi}$ in Omisfit, but $\partial J/\partial \kappa_{Redi}$ in Pmisfit are also largest south of where North Atlantic Deep Water is formed in the deep ocean. In contrast, $\partial J/\partial \kappa_{Redi}$ are largest where Antarctic Bottom Water is formed and at high latitudes in Amisfit. It is less clear (than for the diapycnal diffusivities) whether it would be beneficial for placing constraints on the Redi coefficients if dissolved oxygen concentrations and possibly other biogeochemical tracers were assimilated in a future optimized state estimate.

4 Conclusions

This study evaluated the potential to affect the diapycnal diffusivities in multiple ocean modeling frameworks using tracer and observationally-derived information; the Redi coefficients were also considered using an ocean state estimation framework. The fidelity of the diapycnal diffusivities derived from finestructure observations was assessed in a couple of ways, building upon the results of Whalen et al. (2015). Comparisons were performed between the average observed microstructure-inferred diapycnal diffusivity profile and the average diapycnal diffusivity profiles from two different models. This comparison was repeated using the average profiles from the models after adjusting them based on the observationally-derived values from a parameterization. The profiles that included substitutions of the observationally-derived values from the parameterization were in better, or at least no worse, agreement with the microstructure observations than the model-calculated values with no substitutions. A coupled earth system model's diapycnal diffusivities were overridden by the observationally-derived diapycnal diffusivities from the finescale parameterization as the model ran. On global average, the model showed improvement in some metrics and no worse disagreement in other metrics. The diapycnal diffusivity substitutions in the coupled model redistribute heat and salt, altering the resolved advection term in the steric sea level budget and leading to dynamic adjustment. The temperature and salinity changes are significant because they exceed the range observed in the model under different initial conditions.

Adjoint sensitivity experiments were used to determine if the misfits of either of two ocean mixing parameters could be improved by assimilating biogeochemical tracers in a future optimization of an ocean state estimate. While we further established that the diapycnal diffusivities derived from finestructure observations are more realistic than diapycnal diffusivities from each model considered here, the uncertainties in the observationally-inferred ocean mixing parameters are fairly large (here, approximately a factor of 2-3). Therefore, three biogeochemical tracers were proposed as potential constraints on ocean mixing in the ocean state estimate. Three adjoint sensitivity experiments were performed using one biogeochemical tracer at a time in the misfit calculation of the model: oxygen concentrations, alkalinities, or phosphate concentrations. These adjoint sensitivity experiments were compared with ones that used one ocean mixing parameter at a time in the misfit calculation: diapycnal diffusivities or Redi coefficients. The spatial distributions of the signs of the adjoint sensitivities with respect to the two different ocean mixing parameters from each of the simulations were compared. The signs of the adjoint sensitivities with respect to the diapycnal diffusivities generally agreed well across one pair of these experiments in the upper ocean and at many deeper depths, but the signs of the adjoint sensitivities with respect to the Redi coefficients did not. These results suggest that the assimilation of dissolved oxygen concentrations could improve estimates of the diapycnal diffusivity field in an ocean state estimate optimization, which is the main result of this study. It is less clear whether the Redi coefficient would also be more accurate upon optimization.

4.1 Caveats and future directions

Many factors–including a dearth of independent observations for assessment, a combination of measurement and structural errors, numerical diffusion in our simulations, and unconstrained parameters in the biogeochemical modules–have stymied progress in state estimation of ocean mixing parameters. First, only one ocean mixing parameter–namely, the diapycnal diffusivity–has been compared with independent observational data–specifically, microstructure. The Redi coefficients derived from Argo observations and ECCO2 output have not been independently validated. It is not clear whether the Osborn-Cox diffusivities from R. P. Abernathey and Marshall (2013) and Busecke and Abernathey (2019) could be used to assess the accuracy of the Redi coefficient product, nor is it obvious whether the NATRE and DIMES observations Groeskamp et al. (2020) used are sufficient for validation. Second, the ECCO-estimated diapycnal diffusivities account for other (e.g., structural) model error, which explains some of their biases relative to microstructure observations. For instance, the ocean mixing parameters in ECCO should be time-dependent as well as spatially-varying, but they are only spatially-varying. Currently, only numerical diffusion varies in time, which could confound some physical inferences about the model (e.g., regarding how sensitive the model's state is to diapycnal diffusion relative to along-isopycnal diffusion). Lastly, there are several unconstrained parameters in biogeochemical modules used to calculate biogeochemical tracers (Verdy & Mazloff, 2017), so some of the disagreements in signs of the adjoint sensitivities found here could be associated with other inaccurate parameters.

These challenges can continue to be overcome by allowing models and observations to inform each other. First, the observationally-derived diapycnal diffusivity from the finescale parameterization could be further scrutinized using ship-based CTD data taken concurrently with microstructure velocity shear data. A preliminary analysis suggests that the percent difference between the full depth-averaged microstructure CTD-derived diapycnal diffusivities from the finescale parameterization and the microstructure-inferred diapycnal diffusivities is indistinguishable from zero (1.68%), but the quality of the the microstructure CTD data has not been fully assessed. Second, we will need to account for the time-dependence of each ocean mixing parameter in a future ocean state estimate. The underdetermined nature of the parameter estimation procedure makes this difficult. These efforts would also benefit from minimizing numerical diffusion, but with added computational expense. It is possible that we can achieve a more accurate ocean state estimate if we calculate a time-dependent, dynamically active diapycnal diffusivity field using a suite of parameterizations instead of allowing the diapycnal diffusivity to be treated as a control parameter. However, we showed that it may not be advisable to rely solely on the parameterizations for diapycnal diffusivities in ocean data assimilation systems, as in the case of the GMAO S2S Ocean Analysis. This was either because of the model's analysis increments or its use of atmospheric forcing fields that were inconsistent with the model's sea surface conditions. Third, unconstrained parameters in the biogeochemical modules could potentially be circumvented. One potential way to do this is by assimilating preformed oxygen (i.e., oxygen without any biological influence, making it a passive tracer) instead of oxygen concentrations. Observationally-derived transit-time distributions with a maximum entropy-based method from previous studies (e.g., *Khatiwala et al.*, 2009; *Zanna et al.*, 2019) can help derive preformed oxygen from oxygen concentration observations. Lastly, the (imperfectly-known) initial conditions of each bio-

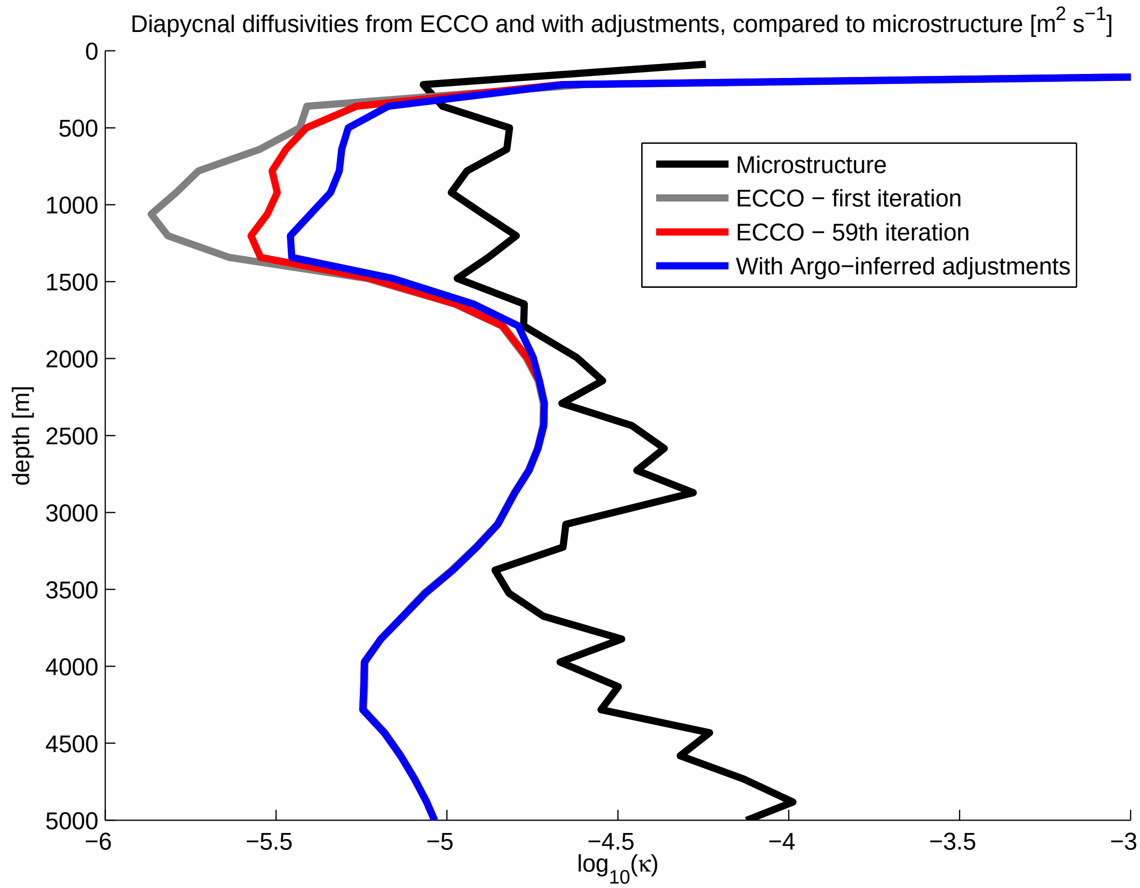


Figure 1. The diapycnal diffusivity profiles averaged over all microstructure observation locations and over the length of the ECCO simulations from the first iteration of the optimization (E-CTRL$_0$ - grey curve), from the (final) fifty-ninth iteration of the optimization (E-CTRL - red curve), and from an ECCO re-run with Argo-derived nudges using Eq. B.1 (blue curve). Also shown is the average of the diapycnal diffusivity profiles from the 24 full-depth microstructure observations (black curve) presented in *Waterhouse et al.* (2014 - see their Fig. 6). At each location, the simulated profiles are extracted and the base-10 logarithms of the geometric averages of the observed and ECCO-estimated diapycnal diffusivities (units in m^2 s^{-1}) are shown.

geochemical tracer will also need to be included in the input control vector during optimization of the ocean state estimate. Our results suggest that the assimilation of biogeochemical tracers will help build a more complete representation and understanding of ocean mixing, and the next step is to perform another optimization of the ocean state estimate including these tracers observations.

Acknowledgments

The authors thank Yury V. Vikhliaev for his help in setting up the GEOS-5 configuration used in this study, Sylvia Cole for providing the Argo- and ECCO2-derived Redi coefficients using mixing length theory, and the reviewers of this manuscript for their suggestions. David Trossman was supported by the Goddard Earth Sciences Technol-

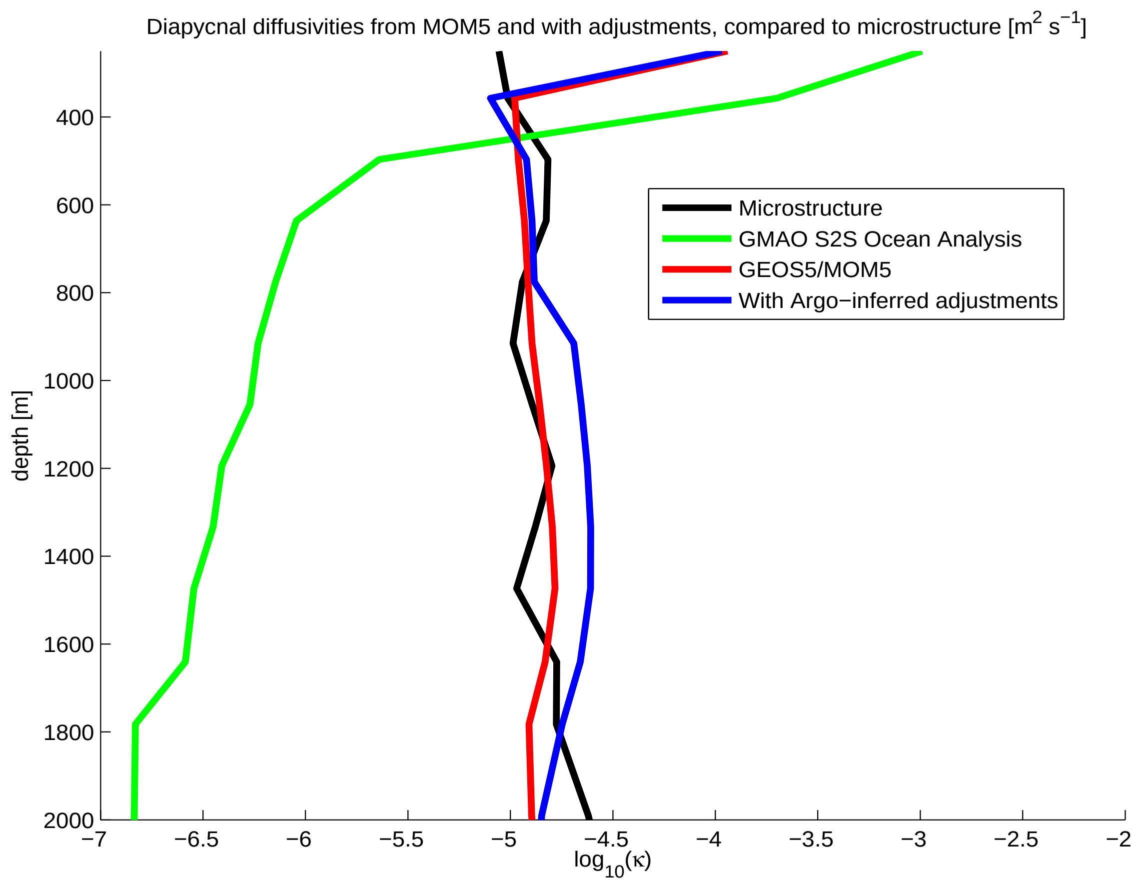


Figure 2. Same as Fig. 1, except for the depth range and shown are the average profiles of diapycnal diffusivity from the GMAO S2S Ocean Analysis (green curve), from the G-CTRL simulation (red curve), from the GEOS-5 simulation with Argo-derived nudges using Eq. B.1 (blue curve), and from the microstructure (black curve), geometrically averaged over the 24 full-depth microstructure observation locations. The Argo-derived nudges used here are from the Whalen et al. (2015) climatology, not the time-varying ones used in DIFF.

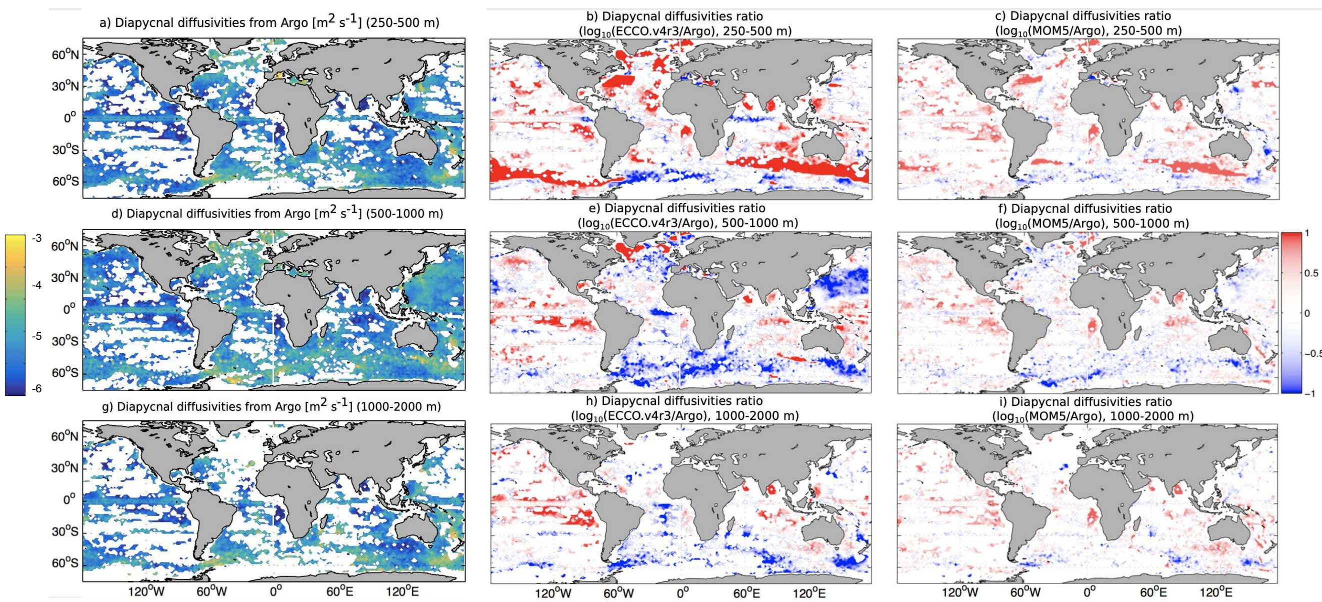


Figure 3. Shown are (a,d,g) the base-10 logarithms of the diapycnal diffusivities (units in $m^2\ s^{-1}$) from the Argo observations (Whalen et al., 2015), (b,e,h) the base-10 logarithms of the ratios of the time-averaged diapycnal diffusivities associated from E-CTRL to those from the Argo-derived product using the finescale parameterization, and (c,f,i) the base-10 logarithms of the ratios of the time-averaged diapycnal diffusivities associated from G-CTRL to those from the Argo-derived product using the finescale parameterization. Panels a-c show an average over 250-500 meters depth. Panels d-f show an average over 500-1000 meters depth. Panels g-i show an average over 1000-2000 meters depth. White areas in the ocean indicate insufficient Argo data to derive a diapycnal diffusivity.

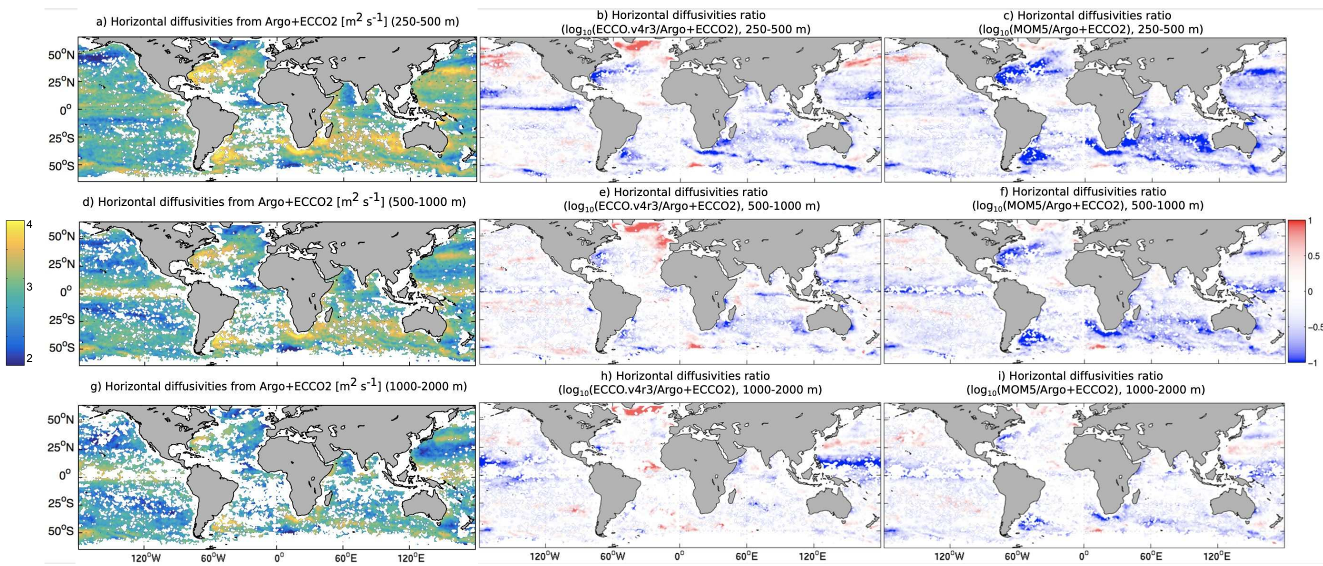


Figure 4. Shown are (a,d,g) the base-10 logarithms of the Redi coefficients (units in m^2 s^{-1}) from the Argo observations and ECCO2 (Cole et al., 2015), (b,e,h) the base-10 logarithms of the ratios of the Redi coefficients from E-CTRL to the Redi coefficients from the Argo- and ECCO2-derived product using mixing length theory, and (c,f,i) the base-10 logarithms of the ratios of the Redi coefficients from G-CTRL to the Redi coefficients from the Argo- and ECCO2-derived product using mixing length theory. Panels a-c show an average over 250-500 meters depth. Panels d-f show an average over 500-1000 meters depth. Panels g-i show an average over 1000-2000 meters depth. White areas in the ocean indicate insufficient Argo data to calculate a Redi coefficient.

Table 1. Listed are the model simulations performed and analyzed in the present study as well as the observationally-derived data or measured data included in each simulation. Only observationally-derived data are included in the GEOS-5 simulations through substitution, only measured data are included through assimilation in the case of the GMAO S2S Ocean Analysis, and either observationally-derived data or measured data are included in the ECCO simulations through its misfit calculation (Eq. 3). Here, κ_ρ denotes an observationally-derived diapycnal diffusivity product from the finescale parameteration, κ_{Redi} indicates the observationally-derived Redi coefficient product from mixing length theory, O_2 is the climatology of measured oxygen concentrations, Alk. is the climatology of measured alkalinities, and PO_4 is the climatology of measured phosphate concentrations.

modeling system	experiment	observationally-derived data	measured data
GEOS-5	G-CTRL	N/A	N/A
GEOS-5	DIFF	κ_ρ (Whalen et al., 2015)	N/A
GEOS-5	BKG	N/A	N/A
GEOS-5	GMAO S2S Ocean Analysis	N/A	see Section 2.3.1
ECCO	E-CTRL	N/A	see Section 2.3.2
ECCO	Dmisfit	κ_ρ (Whalen et al., 2015; Kunze, 2017)	N/A
ECCO	Rmisfit	κ_{Redi} (Cole et al., 2015)	N/A
ECCO	Omisfit	N/A	O_2 [*WOA*, 2013]
ECCO	Amisfit	N/A	Alk. [*WOA*, 2013]
ECCO	Pmisfit	N/A	PO_4 [*WOA*, 2013]

Table 2. The latitude and depth ranges of each observationally-derived product from a parameterization used in this study. The longitude range for each dataset spans $(180^oE, 180^oW)$. Also listed is the time period of the observations each product is based on and the range of values in each product (to the nearest order of magnitude in units of m^2s^{-1}).

data source	range [m^2s^{-1}]	latitude range	depth range	time period
Argo ($\kappa_{\rho,W15}$)	$(10^{-7}, 10^{-2})$	$(75^oS, 75^oN)$	(250,2000)	2006-2014
(P)ALACE and Argo ($\kappa_{\rho,t}$)	$(10^{-7}, 10^{-2})$	$(75^oS, 75^oN)$	(125,2000)	2001-2016
Ship-based CTD hydrography ($\kappa_{\rho,K17}$)	$(10^{-8}, 10^{-3})$	$(77.35^oS, 78.70^oN)$	(173,6044.5)	1981-2010
Argo and ECCO2 ($\kappa_{Redi,C15}$)	$(10^1, 10^5)$	$(61^oS, 62^oN)$	(20, 1920)	2005-2012

Table 3. The globally volume-averaged steric sea level (units in m yr^{-1}) budget terms (and their temporal standard deviations in parentheses) over the length of the G-CTRL and DIFF minus G-CTRL simulations. Here, "surface fluxes" includes shortwave (accounting for the penetrating contribution), longwave, latent, sensible, and frazil heat flux contributions. Contributions not listed here include calving of land ice and frazil ice formation, which approximately equal the differences between the diagnosed total and total tendencies. The terms in **bold** are numerical terms.

term	G-CTRL [mm yr^{-1}]	DIFF minus G-CTRL [mm yr^{-1}]
resolved advection	−0.67 (1.68)	−0.13 (0.57)
neutral diffusion	−3.30 (0.36)	−0.017 (0.052)
vertical diffusion	−15.9 (2.29)	0.060 (0.14)
mesoscale transport	−1.26 (0.11)	−0.0090 (0.010)
submesoscale transport	−0.031 (0.14)	0.010 (0.029)
nonlocal convection (KPP)	−0.23 (0.15)	−0.0077 (0.0079)
sigma-diffusion	0.0015 (0.011)	−0.0049 (0.0049)
downslope mixing	0.06 (0.052)	0.0017 (0.0015)
precipitation minus evaporation	−2.66 (0.56)	−0.16 (0.14)
surface flux	28.1 (52.6)	0.27 (3.38)
runoff mixing	5.04 (1.34)	0.081 (0.19)
smoother	−0.015 (0.0086)	0.00087 (0.011)
diagnosed total	9.11 (52.2)	0.095 (3.08)
total	9.11 (52.2)	0.095 (3.08)

Table 4. The cost functions of the five adjoint sensitivity ECCO runs for each data sources. Listed are the globally computed values, which are used to normalize the adjoint sensitivities shown in Figs. 9-12, and the number of data points used.

experiment	data source	cost function	number of data points
Dmisfit	Argo	1.91×10^{17}	5.933×10^4
Dmisfit	Ship-based CTD hydrography	2.89×10^{18}	7.3806×10^4
Rmisfit	Argo and ECCO2	5.32×10^5	1.5045×10^4
Omisfit	O_2 WOA (2013)	7.71×10^4	7.9752×10^4
Amisfit	Alkalinity WOA (2013)	9.56×10^{14}	6.7104×10^4
Pmisfit	PO_4 WOA (2013)	6.37×10^{11}	3.0382×10^4

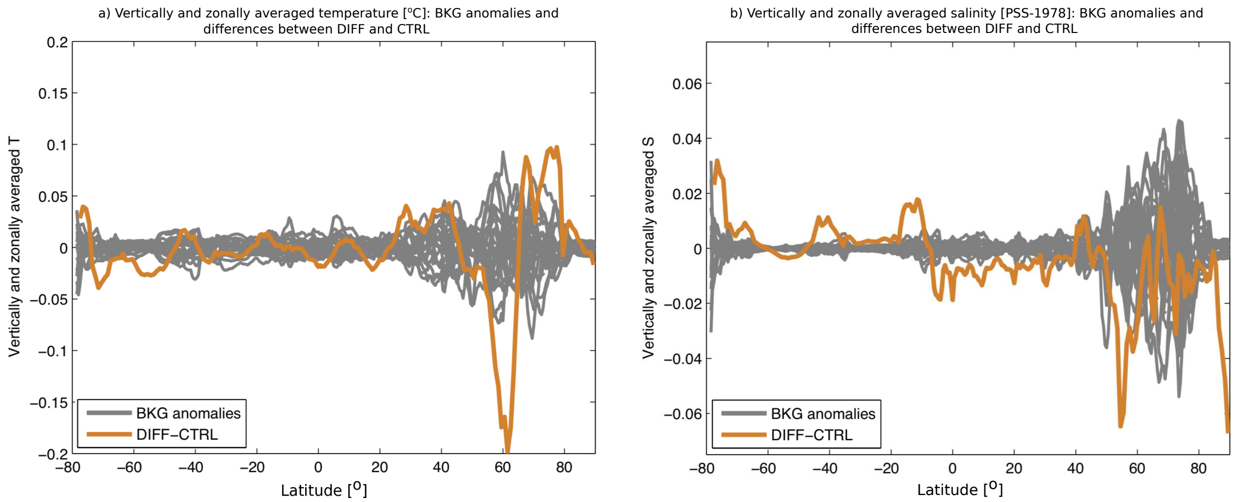


Figure 5. Shown are the vertically and zonally averaged temperature (units in ^{o}C - panel a) and salinity (in PSS-1978 - panel b) anomalies from the average of the 21 free-running simulations (BKG) used to compute the background error covariances, each starting from different initial conditions (grey curves). Also shown are the differences in vertical and zonally averaged temperature (panel a) and salinities (panel b) between the simulations with the time-varying diapycnal diffusivity overrides (DIFF) and without (G-CTRL) (brown/tan-ish curves). An average from the May of the sixth-to-final year to the September of the final year of the GEOS-5 simulations have been taken.

ogy And Research (GESTAR) cooperative agreement between the GMAO of the NASA Goddard Space Flight Center base and Johns Hopkins University as well as NASA SLCT grant 80NSSC17K0675 at the University of Texas-Austin. Thomas W. N. Haine was supported by NOAA award NA15OAR4310172 and NSF award OCE-1338814. Amy Waterhouse was supported by NSF award OCE-0968721. Caitlin Whalen was supported by the Applied Physics Laboratory SEED fellowship. Patrick Heimbach was supported by the ECCO project through a JPL/Caltech subcontract.

The authors acknowledge the Texas Advanced Computing Center (TACC) at The University of Texas at Austin for providing HPC resources for the ECCO simulations (URL: http://www.tacc.utexas.edu) and the NASA Center for Climate Simulation (NCCS) for the computer time spent on the GEOS-5 simulations that have contributed to the research results reported within this paper. The Argo-derived diapycnal diffusivities from the finescale parameterization and GMAO S2S Ocean Analysis data used in this study are available at:

`ftp://gmaoftp.gsfc.nasa.gov/pub/data/kovach/S2S_OceanAnalysis/`

. The hydrography-derived diapycnal diffusivities from the finescale parameterization used in this study, courtesy of Eric Kunze, are available by logging in as a guest at:

`ftp://ftp.nwra.com/outgoing/kunze/iwturb/`

. The microstructure data used in this study are available at:

`https://microstructure.ucsd.edu/`

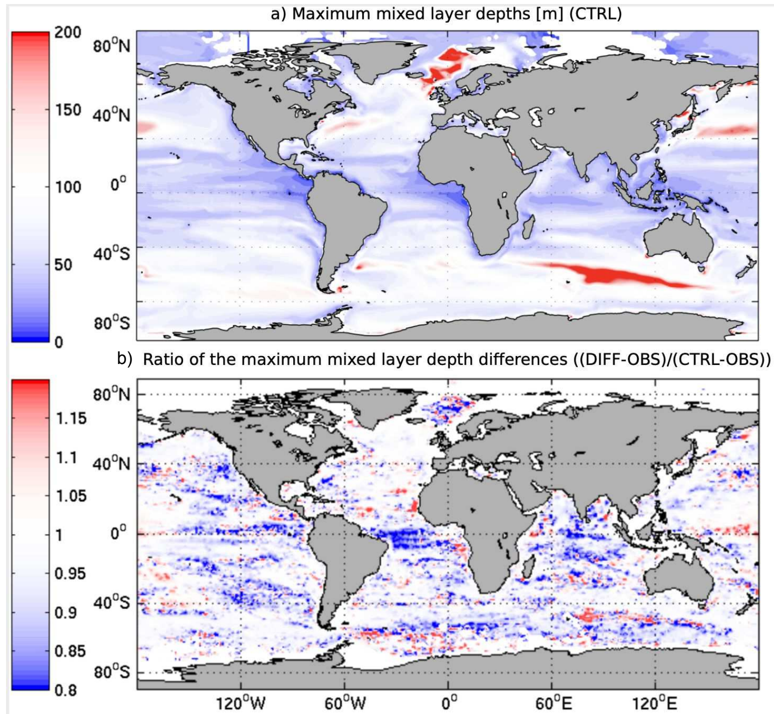


Figure 6. Shown are the maximum yearly mixed layer depths (units in meters - panel a) in G-CTRL, averaged over the final 16 years. Also shown are the ratios of the differences between the maximum mixed layer depths from the density-based algorithm of *Holte et al.* (2017) using Argo observations (panel b) (OBS) and DIFF to the differences between those from OBS and G-CTRL, averaged over the final 16 years of the simulations. Blue colors in panel b imply that ocean mixing parameter adjustment results in better agreement with observations.

.

References

Abernathey, R., Marshall, J., Mazloff, M., & Shuckburgh, E. (2013). Enhancement of mesoscale eddy stirring at steering levels in the southern ocean. *Journal of Physical Oceanography*, *40*, 170–184.

Abernathey, R. P., & Marshall, J. (2013). Global surface eddy diffusivities derived from satellite altimetry. *Journal of Geophysical Research: Oceans*, *118*, 901–916. doi: https://doi.org/10.1002/jgrc.20066

Adcroft, A., & Campin, J.-M. (2004). Rescaled height coordinates for accurate representation of free-surface flows in ocean circulation models. *Ocean Modelling*, *7*, 269–284.

Adcroft, A., Hill, C., & Marshall, J. (1997). The representation of topography by shaved cells in a height coordinate model. *Mon. Wea. Rev.*, *125*, 2293–2315.

Alford, M. H., MacKinnon, J. A., Simmons, H. L., & Nash, J. D. (2016). Near-inertial internal gravity waves in the ocean. *Annual Review of Marine Science*, *8*, 95–123.

Arbic, B. K., Garner, S. T., Hallberg, R. W., & Simmons, H. L. (2004). The accuracy of surface elevations in forward global barotropic and baroclinic tide models. *Deep Sea Research, Part II*, *51*, 3069–3101. doi: http://dx.doi.org/10.1016/j.dsr2.2004.09.014

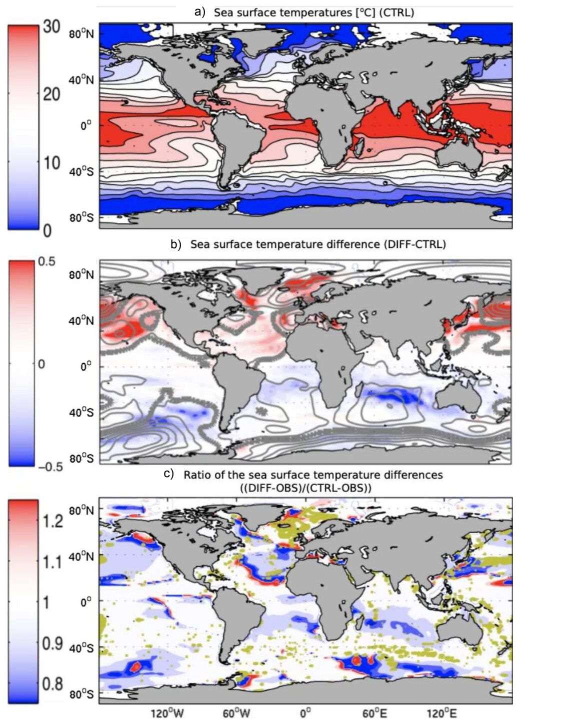


Figure 7. Shown are the sea surface temperatures (units in ^{o}C - panel a) in G-CTRL, averaged over the final 16 years. Also shown are the changes in the sea surface temperatures (units in ^{o}C - panel b) in DIFF relative to G-CTRL, averaged over the final 16 years The grey contours in panel b indicate the magnitude of sea level pressure changes (in 25 Pa contour levels); the thickest grey contours indicate the zero change contour for sea level pressure. Lastly, shown are the ratios of the differences between the sea surface temperatures from the Reynolds product of *Reynolds et al.* (2007) (panel c) (OBS) and DIFF to the differences between those from OBS and G-CTRL, averaged over the final 16 years of the simulations. Blue colors in panel c imply that ocean mixing parameter adjustment results in better agreement with observations. The gold contours in panel c indicate depth-integrated and temporally-averaged diapycnal diffusivity changes at contour intervals of 5×10^{-5} m^2 s^{-1}.

Bachman, S. D., Fox-Kemper, B., & Bryan, F. O. (2020). A diagnosis of anisotropic eddy diffusion from a high-resolution global ocean model. *JAMES*, *12*, e2019MS001904. doi: https://doi.org/10.1029/2019MS001904

Bahl, A., Gnanadesikan, A., & Pradal, M.-A. (2019). Variations in ocean deoxygenation across earth system models: Isolating the role of parameterized lateral mixing. *Global Biogeochemical Cycles*, *33*, 703–724. doi: https://doi.org/10.1029/2018GB006121

Bahl, A., Gnanadesikan, A., & Pradal, M.-A. (2020). Scaling global warming impacts on ocean ecosystems: Lessons from a suite of earth system models. *Frontiers in Marine Science*, *7*(698). doi: https://doi.org/10.3389/fmars.2020.00698

Bates, M., Tulloch, R., Marshall, J., & Ferrari, R. (2014). Rationalizing the spatial distribution of mesoscale eddy diffusivity in terms of mixing length theory. *Journal of Physical Oceanography*, *44*(6), 1523–1540. doi: https://doi.org/10.1175/jpo-d-13-0130.1

Borovikov, A., Cullather, R., Kovach, R., Marshak, J., Vernieres, G., Vikhliaev, Y.,

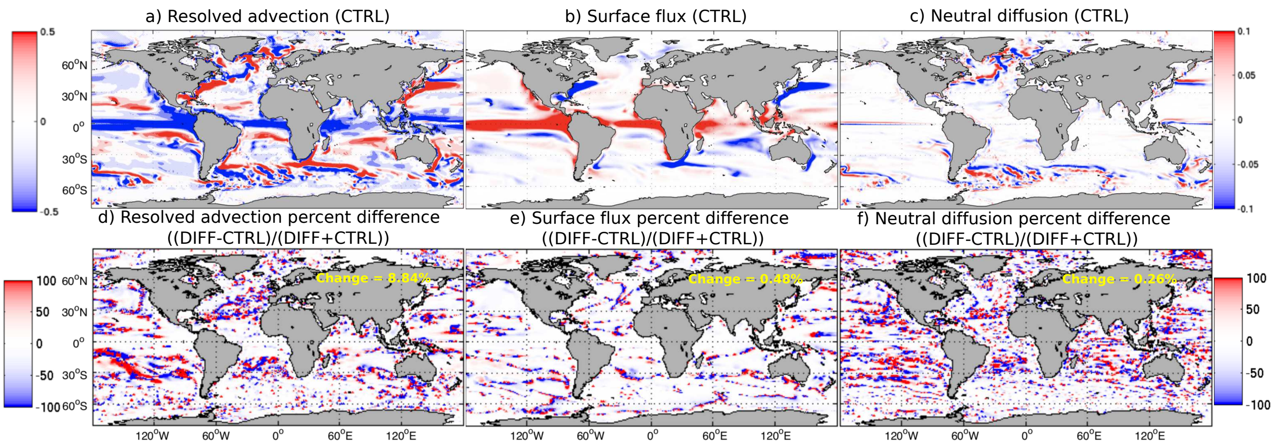


Figure 8. Shown are the cumulative steric sea level changes [units in meters] over the final five years and five months of the G-CTRL simulation for (a) the resolved advection term, (b) the surface flux term, and (c) the neutral diffusion term, described in Section 2.3.1. Also shown are the percent differences (ratio of their differences to their sum) between the cumulative steric sea level changes [units in meters] over the final five years and five months of the DIFF and G-CTRL simulations for the same terms (panels d-f). Globally area-weighted averages of percent changes between DIFF and G-CTRL (ratio of their differences to their sums) are listed in yellow.

… Li, Z. (2017). Geos-5 seasonal forecast system. *Climate Dynamics*, 1–27. doi: https://doi.org/10.1007/s00382-017-3835-2

Busecke, J. J. M., & Abernathey, R. P. (2019). Ocean mesoscale mixing linked to climate variability. *Science Advances*, *5*. doi: 10.1126/sciadv.aav5014

Campin, J.-M., Adcroft, A., Hill, C., & Marshall, J. (2004). Conservation of properties in a free surface model. *Ocean Modelling*, *6*, 221–244.

Canuto, V. M., Cheng, Y., Howard, A. M., & Dubovikov, M. S. (2019). Three-dimensional, space-dependent mesoscale diffusivity: derivation and implications. *Journal of Physical Oceanography*, *49*(4), 1055–1074. doi: https://doi.org/10.1175/JPO-D-18-0123.1

Chanona, M., Waterman, S., & Gratton, Y. (2018). Variability of internal wave-driven mixing and stratification in the canadian arctic ocean. *Journal of Geophysical Research-Oceans*, *123*, 9178–9195. doi: https://doi.org/10.1029/2018JC014342

Chaudhuri, A. H., Ponte, R. M., Forget, G., & Heimbach, P. (2013). A comparison of atmospheric reanalysis surface products over the ocean and implications for uncertainties in air-sea boundary forcing. *Journal of Climate*, *26*, 153–170.

Chin, M., Ginoux, P., Kinne, S., Holben, B. N., B. N. Duncan, Martin, R. V., Logan, J. A., … Nakajima, T. (2002). Tropospheric aerosol optical thickness from the gocart model and comparisons with satellite and sunphotometer measurements. *Journal of Atmospheric Sciences*, *59*, 461–483.

Cole, S. T., Wortham, C., Kunze, E., & Owens, W. B. (2015). Eddy stirring and horizontal diffusivity from argo float observations: Geographic and depth variability. *Geophysical Research Letters*, *42*, 3989–3997. doi: https://doi.org/10.1002/2015GL063827

Couespel, D., Lévy, M., & Bopp, L. (2019). Major contribution of reduced upper ocean oxygen mixing to global ocean deoxygenation in an earth system model. *Geophysical Research Letters*, *46*, 12239–12249. doi: https://doi.org/10.1029/2019GL084162

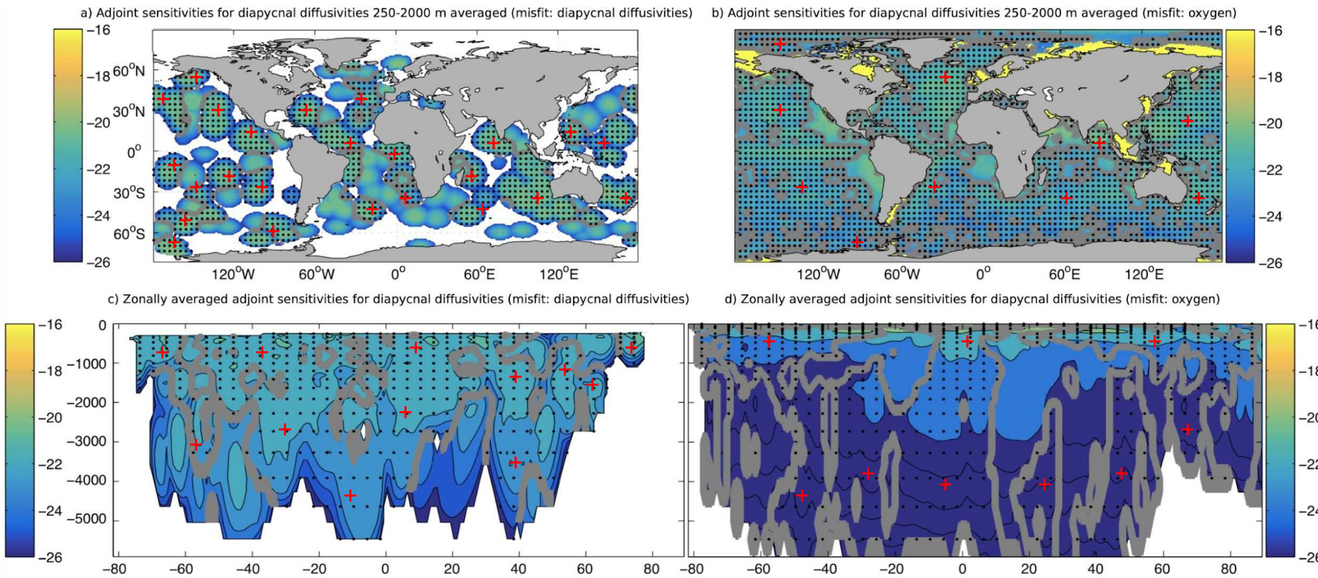


Figure 9. Results from Dmisfit (panels a and c) and Omisfit (panels b and d) are shown. The base-10 logarithms of the absolute values of the adjoint sensitivities (units in s m^{-2}) with respect to the diapycnal diffusivities are shown: averaged over 250-2000 meters depth (panels a-b) and zonally averaged (panels c-d) in the misfit calculation. The black dots with a red plus sign surrounded by grey contours mean that the adjoint sensitivities are positive ($\partial J/\partial K > 0$); elsewhere show negative adjoint sensitivities. $\kappa_{\rho,W15}$ and $\kappa_{\rho,K17}$ are the only quantities used in the misfit calculation of an adjoint run shown in panels a and c. The climatological oxygen concentrations from the World Ocean Atlantic (2013) are the only observations used in the misfit calculation of a separate adjoint run shown in panels b and d. The adjoint sensitivities in panels a and c are computed offline (i.e., not using ECCO, but by plugging in the value the model reads in for the base-10 logarithm of the diapycnal diffusivities and comparing that with the above observationally-derived base-10 logarithm of the diapycnal diffusivity products using the finescale parameterization via Eq. 4). The white regions in panels a and c are locations where there is insufficient data or where there is bathymetry. The adjoint sensitivities in panels b and d are computed online (i.e., using ECCO, which computes the misfits between the base-10 logarithm of the diapycnal diffusivities it reads in and the observationally-derived base-10 logarithm of the diapycnal diffusivities using the finescale parameterization). The white regions in panels b and d are locations with bathymetry or insufficient observations. The adjoint sensitivities at each grid point are divided by the volume of each grid cell and then scaled by the cost function (Table 4) for each respective experiment in order to make each point more comparable with another. The adjoint sensitivities are calculated over just one year (1992)

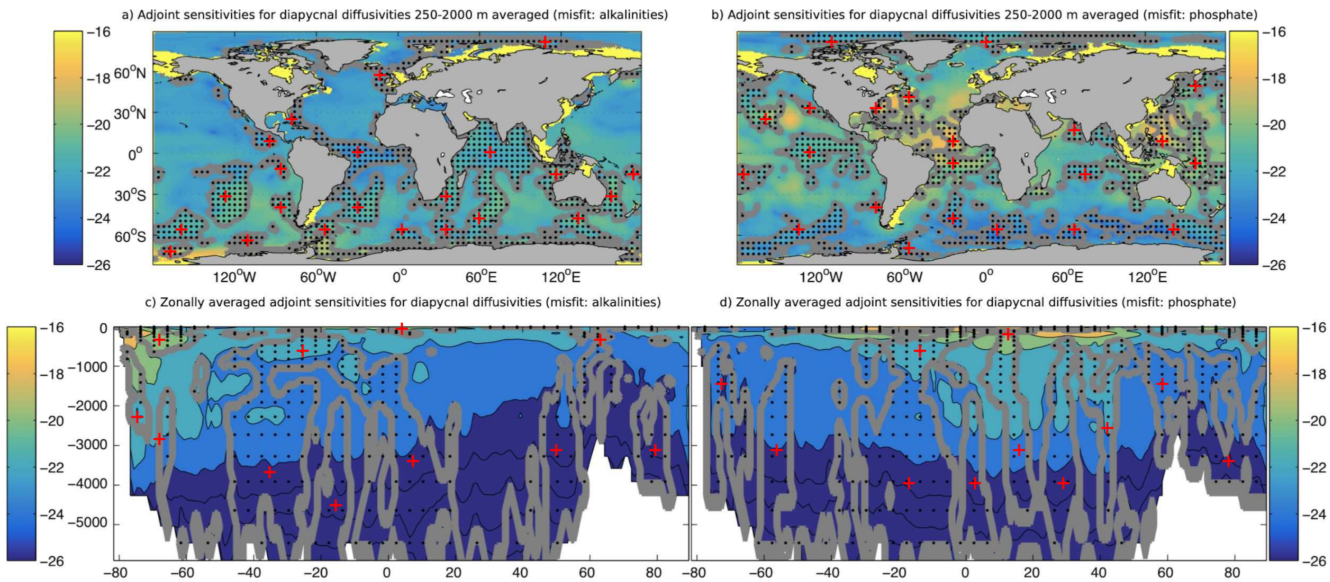


Figure 10. Same as Figs. 9, except the only observations used in the misfit calculation of the adjoint runs are the climatological alkalinities (panels a and c) or phosphate concentrations (panels b and d) from the World Ocean Atlas (2013). Each of these runs compute the adjoint sensitivities online.

Table 5. Listed are the percent volumes where the signs of the adjoint sensitivities across pair-wise model simulations agree. The percentages are only calculated where sufficient observations are available to derive an ocean mixing parameter using a parameterization and where the difference between the model-calculated and observationally-derived ocean mixing parameter using a parameterization is greater than the uncertainty (i.e., three times the observationally-derived ocean mixing parameter using a parameterization). The percentages are smaller (by up to 20%) if all locations where sufficient observations are available to derive an ocean mixing parameter using a parameterization are included, suggesting that the disagreements tend to be in locations where the model's diapycnal diffusivity bias relative to the observationally-derived value from a parameterization is insignificant from zero. The adjoint sensitivities with respect to the diapycnal diffusivity (κ_ρ) or Redi coefficient (κ_{Redi}) are specified.

experiments	$\partial J/\partial \log_{10}(\kappa_\rho)$ or $\partial J/\partial \kappa_{Redi}$	percent of ocean volume with agreement
Dmisfit, Omisfit	$\partial J/\partial \log_{10}(\kappa_\rho)$	70.8%
Dmisfit, Amisfit	$\partial J/\partial \log_{10}(\kappa_\rho)$	41.8%
Dmisfit, Pmisfit	$\partial J/\partial \log_{10}(\kappa_\rho)$	33.2%
Omisfit, Pmisfit	$\partial J/\partial \log_{10}(\kappa_\rho)$	42.3%
Rmisfit, Omisfit	$\partial J/\partial \kappa_{Redi}$	47.8%
Rmisfit, Amisfit	$\partial J/\partial \kappa_{Redi}$	49.6%
Rmisfit, Pmisfit	$\partial J/\partial \kappa_{Redi}$	51.2%
Omisfit, Pmisfit	$\partial J/\partial \kappa_{Redi}$	44.8%

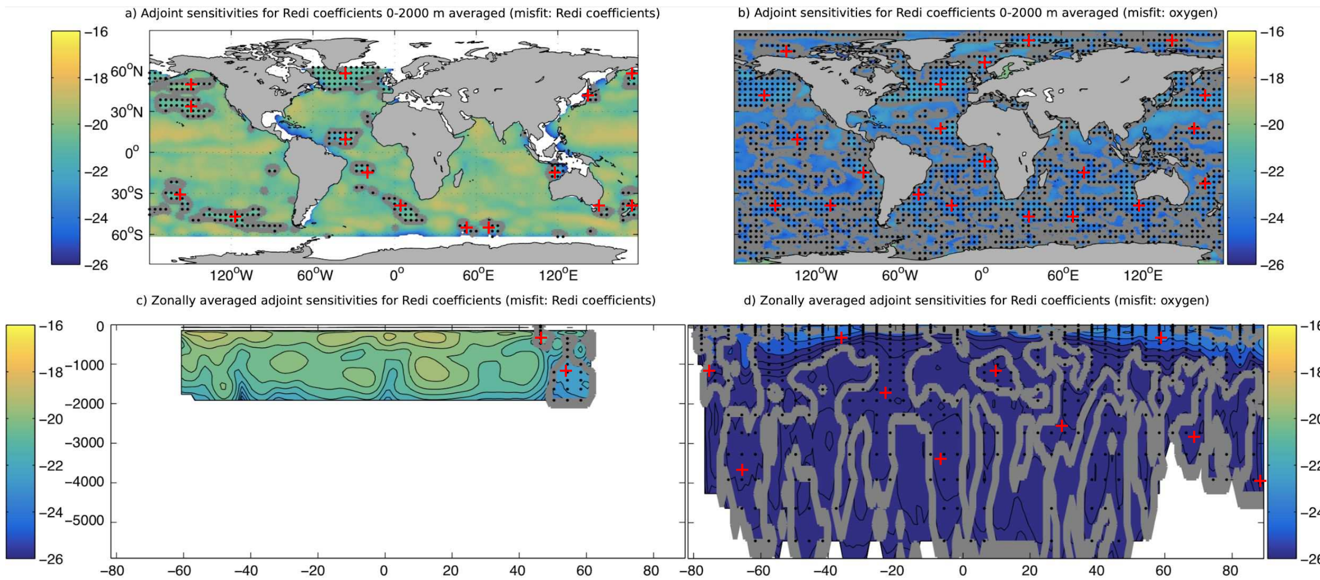


Figure 11. Results from Rmisfit (panels a and c) and Omisfit (panels b and d) are shown. The base-10 logarithms of the absolute values of the adjoint sensitivities (units in s m^{-2}) with respect to the Redi coefficients are shown: averaged over 0-2000 meters depth (panels a-b) and zonally averaged (panels c-d) in the misfit calculation. The black dots with a red plus sign surrounded by grey contours mean that the adjoint sensitivities are positive ($\partial J/\partial K > 0$); elsewhere show negative adjoint sensitivities. $\kappa_{Redi,C15}$ is the only quantity used in the misfit calculation of an adjoint run shown in panels a and c. The climatological oxygen concentrations from the World Ocean Atlantic (2013) are the only observations used in the misfit calculation of a separate adjoint run shown in panels b and d. The adjoint sensitivities in panels a and c are computed offline (i.e., not using ECCO, but by plugging in the value the model reads in for the Redi coefficient and comparing that with $\kappa_{Redi,C15}$ via Eq. 4). The adjoint sensitivities in panels b and d are computed online (i.e., using ECCO, which computes the misfits between the Redi coefficients it reads in and the observationally-derived Redi coefficients using mixing length theory). The adjoint sensitivities at each grid point are divided by the volume of each grid cell and then scaled by the cost function (Table 4) for each respective experiment in order to make each point more comparable with another. The adjoint sensitivities are calculated over just one year (1992).

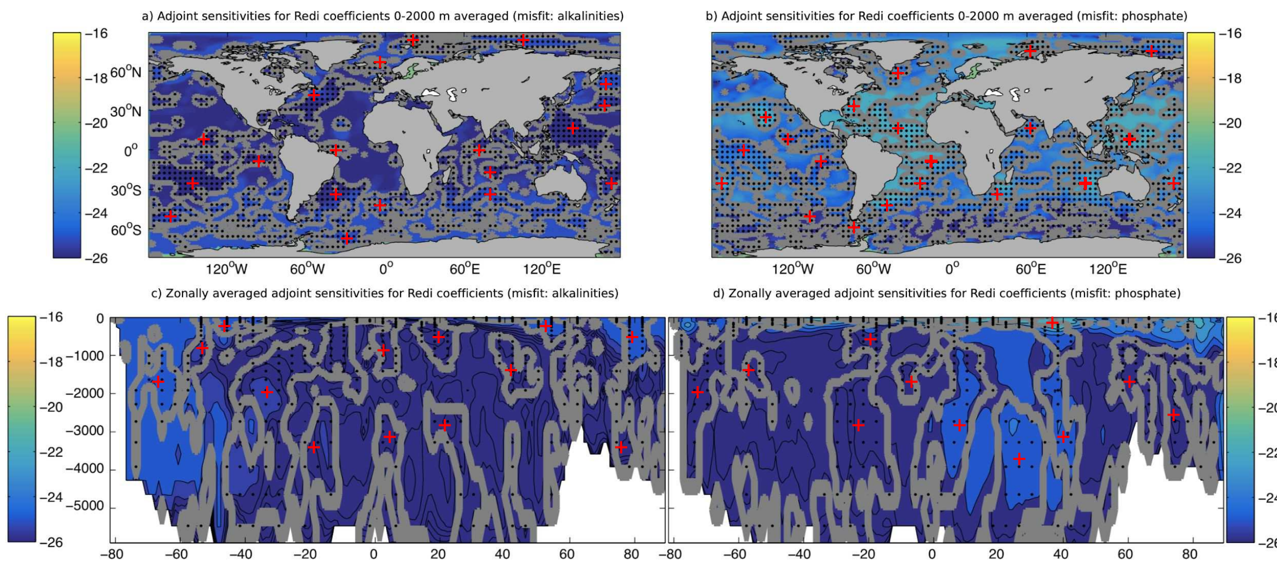


Figure 12. Same as Figs. 11, except the only observations used in the misfit calculation of the adjoint runs are the climatological alkalinities (panels a and c) or phosphate concentrations (panels b and d) from the World Ocean Atlas (2013). Each of these runs compute the adjoint sensitivities online.

Dalan, F., Stone, P. H., & Sokolov, A. P. (2005). Sensitivity of the ocean's climate to diapycnal diffusivity in an emic. part ii: Global warming scenario. *Journal of Climate*, *18*, 2482–2496.

Danabasoglu, G., & McWilliams, J. C. (1995). Sensitivity of the global ocean circulation to parameterizations of mesoscale tracer transports. *J. Clim.*, *8*, 2967–2987.

Davis, R. E. (1991). Observing the general circulation with floats. *Deep Sea Research Part A. Oceanographic Research Papers*, *38(S1)*, S531–S571. doi: https://doi.org/10.1016/S0198-0149(12)80023-9

Davis, R. E., Regier, L. A., Dufour, J., & Webb, D. C. (1992). The autonomous lagrangian circulation explorer (alace). *Journal of Atmospheric and Oceanic Technology*, *9*, 264–285. doi: https://doi.org/10.1175/1520-0426(1992)009⟨0264:TALCE⟩2.0.CO;2

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., ... Vitart, F. (2011). The era-interim reanalysis: configuration and performance of the data assimilation system. *Q. J. Royal Met. Soc.*, *137*, 553–597.

de Lavergne, C., Vic, C., Madec, G., Roquet, F., Waterhouse, A. F., Whalen, C. B., ... Hibiya, T. (2020). Sensitivity of the global ocean circulation to parameterizations of mesoscale tracer transports. *JAMES*, *12*, e2020MS002065. doi: https://doi.org/10.1029/2020MS002065

Dufour, C. O., Griffies, S. M., de Souza, G. F., Frenger, I., Morrison, A. K., Palter, J. B., ... Slater, R. D. (2015). Role of mesoscale eddies in cross-frontal transport of heat and biogeochemical tracers in the southern ocean. *Journal of Physical Oceanography*, *45*, 3057–3081.

Dutkiewicz, S., Sokolov, A., Scott, J., & Stone, P. (2005). A three-dimensional ocean-sea ice-carbon cycle model and its coupling to a two-dimensional atmospheric model: Uses in climate change studies. *Tech. rep., MIT Joint Program of the Science and Policy of Global Change*, *122*. doi: http://web.mit.edu/globalchange/www/MITJPSPGC_{}Rpt122.pdf

Ehlert, D., Zickfeld, K., Eby, M., & Gillett, N. (2017). The sensitivity of the proportionality between temperature change and cumulative co_2 emissions to ocean mixing. *Journal of Climate*, *30*, 2921–2935.

Ferrari, R., Griffies, S. M., Nurser, A. J. G., & Vallis, G. K. (2010). A boundary-value problem for the parameterized mesoscale eddy transport. *Ocean Modelling*, *32*, 143–156.

Ferrari, R., McWilliams, J. C., Canuto, V. M., & Dubovikov, M. (2008). Parameterization of eddy fluxes near oceanic boundaries. *Journal of Climate*, *21*, 2770–2789.

Forget, G., Campin, J. M., Heimbach, P., Hill, C. N., Ponte, R. M., & Wunsch, C. (2015). ECCO version 4: an integrated framework for non-linear inverse modeling and global ocean state estimation. *Geosci. Model Dev.*, *8*, 3071-3104. doi: https://doi.org/10.5194/gmd-8-3071-2015

Forget, G., Ferreira, D., & Liang, X. (2011). On the observability of turbulent transport rates by argo: supporting evidence from an inversion experiment. *Ocean Science*, *11*, 839–853. doi: http://doi.org/10.5194/os-11-839-2015

Fukumori, I., Wang, O., Fenty, I., Forget, G., Heimbach, P., & Ponte, R. M. (2017). Ecco version 4 release 3. *DSpace MIT*. doi: http://hdl.handle.net/1721.1/110380

Galbraith, E. D., Dunne, J. P., Gnanadesikan, A., Slater, R. D., Sarmiento, J. L., Dufour, C. O., ... Marvasti, S. S. (2015). Complex functionality with minimal computation: Promise and pitfalls of reduced-tracer ocean biogeochemistry models. *Journal of Advances in Modeling Earth Systems*, *7*, 20122028. doi: https://doi.org/10.1002/2015MS000463

Gaspar, P., Grégoris, Y., & LeFevre, J.-M. (1990). A simple eddy kinetic energy model for simulations of the oceanic vertical mixing: tests at station papa and long-term upper ocean study site. *Journal of Geophysical Research*, *95*, 16,179–16,193.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., ... Zhao, B. (2017). The modern-era retrospective analysis for research and applications, version 2 (merra-2). *Journal of Climate*, *30*, 5419–5454. doi: https://dx.doi.org10.1175/JCLI-D-16-0758.1

Gent, P. R., & McWilliams, J. C. (1990). Isopycnal mixing in ocean circulation models. *Journal of Physical Oceanography*, *20*(1), 150–155. doi: https://doi.org/10.1175/1520-0485(1990)020⟨0150:IMIOCM⟩2.0.CO;2

Gerdes, R., Köberle, C., & Willebrand, J. (1990). The influence of numerical advection schemes on the results of ocean general circulation models. *Climate Dynamics*, *5*, 211–226.

Giering, R., & Kaminski, T. (1998). Recipes for adjoint code construction. *ACM Transactions on Mathematical Software*, *24*, 437–474.

Gilbert, J. C., & Lemarechal, C. (1989). Some numerical experiments with variable-storage quasi-newton algorithms. *Math. Program.*, *45*, 407–435.

Gnanadesikan, A. (1999). A simple predictive model for the structure of the oceanic pycnocline. *Science*, *283*, 2077.

Gnanadesikan, A., Pradal, M.-A., & Abernathey, R. (2015). Isopycnal mixing by mesoscale eddies significantly impacts oceanic anthropogenic carbon uptake. *Geophysical Research Letters*, *42*, 4249–4255. doi: https://doi.org/10.1002/2015GL064100

Gregg, M. C. (1989). Scaling turbulent dissipation in the thermocline. *Journal of Geophysical Research*, *94*, 9686–9698.

Gregg, M. C., D'Asaro, E. A., Riley, J. J., & Kunze, E. (2018). Mixing efficiency in the ocean. *Annual Review of Marine Science*, *10*, 443–473.

Griewank, A. (1992). Achieving logarithmic growth of temporal and spatial complexity in reverse automatic differentiation. *Optimization Methods and Software*, *1*, 35–54. doi: https://doi.org/10.1080/10556789208805505

Griffies, S. M., Winton, M., Anderson, W. G., Benson, R., Delworth, T. L., Dufour, C. O., ... Zhang, R. (2015). Impacts on ocean heat from transient mesoscale eddies in a hierarchy of climate models. *Journal of Climate*, *28*, 952–977.

Groeskamp, S., LaCasce, J. H., McDougall, T. J., & Rogé, M. (2020). Fulldepth global estimates of ocean mesoscale eddy mixing from observations and theory. *Geophysical Research Letters*, *47*, e2020GL089425. doi: https://doi.org/10.1029/2020GL089425

Henyey, F. S., Wright, J., & Flatté, S. M. (1996). Energy and action flow through the internal wave field: an eikonal approach. *Journal of Geophysical Research*, *91*, 8487–8495.

Hieronymus, M., Nycander, J., Nilsson, J., Döös, K., & Hallberg, R. (2019). Oceanic overturning and heat transport: the role of background diffusivity. *Journal of Climate*, *32*, 701–716.

Holte, J., Talley, L. D., Gilson, J., & Roemmich, D. (2017). An argo mixed layer climatology and database. *Geophysical Research Letters*, *44*, 5618 5626. doi: https://doi.org/10.1002/2017GL073426

Hunke, E. C., Lipscomb, W. H., Turner, A. K., Jeffery, N., & Elliott, S. (2013). Cice: the los alamos sea ice model documentation and software user's manual version 5.0. *Los Alamos National Laboratory*, *LA-CC-06-012*.

Jiang, Z.-P., Tyrrell, T., Hydes, D. J., Dai, M., & Hartman, S. E. (2014). Variability of alkalinity and the alkalinity-salinity relationship in the tropical and subtropical surface ocean. *Global Biogeochemical Cycles*, *28*, 729–742. doi: https://doi.org/10.1002/2013GB004678

Jones, C. S., & Abernathey, R. P. (2019). Isopycnal mixing controls deep ocean ventilation. *Geophysical Research Letters*, *46*. doi: https://doi.org/10.1029/2019GL085208

Kakehi, S., Ito, S.-I., & Wagawa, T. (2017). Estimating surface water mixing ratios using salinity and potential alkalinity in the kuroshio-oyashio mixed water regions. *Journal of Geophysical Research-Oceans*, *122*, 1927–1942. doi: https://doi.org/10.1002/2016JC012268

Katsumata, K. (2016). Eddies observed by argo floats. part i: Eddy transport in the upper 1000 dbar. *Journal of Physical Oceanography*, *46*, 3471–3486. doi: https://doi.org/10.1175/JPO-D-16-0150.1

Kostov, Y., Johnson, H. L., & Marshall, D. P. (2019). Amoc sensitivity to surface buoyancy fluxes: the role of air-sea feedback mechanisms. *Climate Dynamics*, *53*, 4521–4537. doi: https://doi.org/10.1007/s00382-019-04802-4

Krasting, J., Stouffer, R., Griffies, S., Hallberg, R., Malyshev, S., Samuels, B., & Sentman, L. (2018). Role of ocean model formulation in climate response uncertainty. *Journal of Climate*, *31*, 9313–9332.

Kunze, E. (2017). Internal-wave-driven mixing: Global geography and budgets. *Journal of Physical Oceanography*, *47*, 1325–1345.

Large, W. G., McWilliams, J. C., & Doney, S. C. (1994). Oceanic vertical mixing: a review and a model with a nonlocal boundary layer parameterization. *Reviews of Geophysics*, *32*, 363–403.

Large, W. G., & Yeager, S. G. (2009). The global climatology of an interannually varying air-sea flux data set. *Climate Dynamics*, *33*, 341–364.

Ledwell, J. R., & Watson, A. J. (1991). The santa monica basin tracer experiment: A study of diapycnal and isopycnal mixing. *Journal of Geophysical Research*, *96*, 8695–8718. doi: https://doi.org/10.1029/91JC00102

Levitus, S., & et al. (2012). World ocean heat content and thermosteric sea level change (0-2000 m), 1955-2010. *Geophysical Research Letters*, *39*. doi: https://doi.org/10.1029/2012GL051106

Liu, C., Kohl, A., & Stammer, D. (2012). Adjoint-based estimation of eddy-induced tracer mixing parameters in the global ocean. *J. Phys. Oceanogr.*, *42*, 1186–1206.

Liu, W., Fedorov, A., & Sévellec, F. (2019). The mechanisms of the atlantic meridional overturning circulation slowdown induced by arctic sea ice decline. *Journal of Climate*, *32*, 977–996.

Lueck, R. G., Huang, D., Newman, D., & Box, J. (1997). Turbulence measurement with a moored instrument. *Journal of Atmospheric and Oceanic Technology*, *14*, 143–161. doi: https://doi.org/10.1175/1520-0426(1997)014⟨0143:TMWAMI⟩2.0.CO;2

MacKinnon, J., Zhao, Z., Whalen, C. B., Waterhouse, A. F., Trossman, D. S., Sun, O. M., ... Alford, M. H. (2017). Climate process team on internal-wave driven ocean mixing. *Bulletin of the American Meteorological Society*, *98*, 2429–2454. doi: http://dx.doi.org/10.1175/BAMS-D-16-0030.1

Manikandan, S. (2011). Measures of central tendency: The mean. *Journal of Phamacol Pharmacother*, *2*, 140–142.

Markus, T., & Cavalieri, D. J. (2009). The amsr-e nt2 sea ice concentration algorithm: its basis and implementation. *Journal of The Remote Sensing Society of Japan*, *29*, 216–225. doi: https://doi.org/10.11440/rssj.29.216

Marshall, J., Scott, J. R., Romanou, A., Kelley, M., & Leboissetier, A. (2017). The dependence of the ocean's moc on mesoscale eddy diffusivities: a model study. *Ocean Modelling*, *111*, 1–8. doi: https://doi.org/10.1016/j.ocemod.2017.01.001

Melet, A., Hallberg, R., Legg, S., & Nikurashin, M. (2014). Sensitivity of the ocean state to lee wave-driven mixing. *Journal of Physical Oceanography*, *44*, 900–921.

Melet, A., Legg, S., & Hallberg, R. (2016). Climatic impacts of parameterized local and remote tidal mixing. *Journal of Climate*, *29*, 3473–3500. doi: http://dx.doi.org/10.1175/JCLI-D-15-0153.1

Melet, A., Nikurashin, M., Muller, C., Falahat, S., Nycander, J., Timko, P. G., ... Goff, J. A. (2013). Internal tide generation by abyssal hills using analytical theory. *Journal of Geophysical Research-Oceans*, *118*, 6303–6318.

Menemenlis, D., Fukumori, I., & Lee, T. (2005). Using green's functions to calibrate and ocean general circulation model. *Monthly Weather Review*, *133*(5), 1224-1240. doi: https://doi.org/10.1175/MWR2912.1

Mesinger, F., & Arakawa, A. (n.d.). Numerical methods used in atmospheric models. In *Wmo/icsu joint organizing committee* (p. 64). GARP Publ. Series.

Messias, M.-J., Watson, A. J., Johannessen, T., Oliver, K. I. C., Olsson, K. A., Fogelqvist, E., ... Ledwell, J. R. (2008). The greenland sea tracer experiment 1996-2002: Horizontal mixing and transport of greenland sea intermediate water. *Progress in Oceanography*, *78*, 85–105. doi: https://doi.org/10.1016/j.pocean.2007.06.005

Molod, A., Hackert, E., Vikhliaev, Y., Zhao, B., Barahona, D., Vernieres, G., & et al. (2020). Geos-s2s version 2: The gmao high-resolution coupled model and assimilation system for seasonal prediction. *Journal of Geophysical Research: Atmospheres*, *125*, e2019JD031767. doi: https://doi.org/10.1029/2019JD031767

Molod, A., Takacs, L., Suarez, M., & Bacmeister, J. (2015). Development of the geos-5 atmospheric general circulation model: evolution from merra to merra-2. *Geoscientific Model Development*, *8*, 1339–1356. doi: https://doi.org/10.5194/gmd-8-1339-2015

Molod, A., Takacs, L., Suarez, M., Bacmeister, J., Song, I.-S., & Eichmann, A. (2012). The geos-5 atmospheric general circulation model: Mean climate and development from merra to fortuna. *Technical Report Series on Global Modeling and Data Assimilation*, *28*.

Moum, J. N., Caldwell, D. R., Nash, J. D., & Gunderson, G. D. (2002). Observations of boundary mixing over the continental slope. *Journal of Physical Oceanography*, *32*, 2113–2130. doi: https://doi.org/10.1175/1520-0485(2002)032⟨2113:OOBMOT⟩2.0.CO;2

Munk, W., & Wunsch, C. (1998). Abyssal recipes ii: Energetics of tidal and wind mixing. *Deep Sea Research, Part I*, *45*, 1977–2010.

Murray, R. J. (1996). Explicit generation of orthogonal grids for ocean models. *Journal of Computational Physics*, *126*, 251–273.

Naveira Garabato, A. C., Nurser, A. G., Scott, R. B., & Goff, J. A. (2013). The impact of small-scale topography on the dynamical balance of the ocean. *Journal of Physical Oceanography*, *43*, 647–668.

Nikurashin, M., & Ferrari, R. (2011). Global energy conversion rate from geostrophic flows into internal lee waves in the deep ocean. *Geophysical Research Letters*, *38*, L08610. doi: http://doi.org/10.1029/2011GL046576

Nocedal, J. (1980). Updating quasi-newton matrices with limited storage. *Mathematics of Computation*, *35*, 773–782.

Nycander, J. (2005). Generation of internal waves in the deep ocean by tides. *Journal of Geophysical Research*, *110*, C10028. doi: http://doi.org/10.1029/2004JC002487

Osborn, T. R. (1980). Estimates of the local rate of vertical diffusion from dissipation measurements. *Journal of Physical Oceanography*, *10*, 83–89.

Palter, J. B., Griffies, S. M., Galbraith, E. D., Gnanadesikan, A., Samuels, B., & Klocker, A. (2014). The driving processes of the deep ocean buoyancy budget and their temporal variability. *Journal of Climate*, *27*, 551–573.

Palter, J. B., & Trossman, D. S. (2018). The sensitivity of future ocean oxygen to changes in ocean circulation. *Global Biogeochemical Cycles*, *32*, 738–751. doi: https://doi.org/10.1002/2017GB005777

Paytan, A., & McLaughlin, K. (2007). The oceanic phosphorus cycle. *Chem. Rev.*, *107*, 563–576.

Penny, S. G., Kalnay, E., Carton, J. A., Hunt, B. R., Ide, K., Miyoshi, T., & Chepurin, G. A. (2013). The local ensemble transform kalman filter and the running-in-place algorithm applied to a global ocean general circulation model. *Nonlinear Processes in Geophysics*, *20*, 1031–1046. doi: http://doi.org/10.5194/npg-20-1031-2013

Piecuch, C. G., & Ponte, R. M. (2011). Mechanisms of interannual steric sea level variability. *Geophysical Research Letters*, *38*, L15605. doi: http://doi.org/10.1029/2011GL048440

Piecuch, C. G., & Ponte, R. M. (2014). Mechanisms of global-mean steric sea level change. *Journal of Climate*, *27*, 824–834. doi: http://doi.org/10.1175/JCLI-D-13-00373.1

Pilo, G. S., Oke, P. R., Coleman, R., Rykova, T., & Ridgway, K. (2018). Impact of data assimilation on vertical velocities in an eddy resolving ocean model. *Ocean Modelling*, *131*, 71–85.

Pollmann, F., Eden, C., & Olbers, D. (2017). Evaluating the global internal wave model idemix using finestructure methods. *Journal of Physical Oceanography*, *47*, 2267–2289.

Polzin, K. L., Naveira Garabato, A. C., Huussen, T. N., Sloyan, B. N., & Waterman, S. (2014). Finescale parameterizations of turbulent dissipation. *Journal of Geophysical Research-Oceans*, *119*, 1383–1419. doi: https://doi.org/10.1002/2013JC008979

Polzin, K. L., Toole, J. M., Ledwell, J. R., & Schmitt, R. W. (1997). Spatial variability of turbulent mixing in the abyssal ocean. *Science*, *276*, 93–96. doi: https://doi.org/10.1126/science.276.5309.93

Polzin, K. L., Toole, J. M., & Schmitt, R. W. (1995). Finescale parameterizations of turbulent dissipation. *Journal of Physical Oceanography*, *25*, 306–328.

Putrasahan, D. A., Lohmann, K., von Storch, J.-S., Jungclaus, J. H., Gutjahr, O., & Haak, H. (2019). Surface flux drivers for the slowdown of the atlantic meridional overturning circulation in a high-resolution global coupled climate model. *JAMES*, *11*, 1349–1363. doi: https://doi.org/10.1029/2018MS001447

Redi, M. H. (1982). Oceanic isopycnal mixing by coordinate rotation. *J. Phys. Oceanogr.*, *12*, 1154–1158.

Reichle, R., Koster, R., De Lannoy, G., Forman, B., Liu, Q., Mahanama, S., & Touré, A. (2011). Assessment and enhancement of merra land surface hydrology estimates. *Journal of Climate*, *24*, 6322–6338. doi: https://doi.org/10.1175/JCLI-D-10-05033.1

Reynolds, R. W., Smith, T. M., Liu, C., Chelton, D. B., Casey, K. S., & Schlax, M. G. (2007). Daily high-resolution blended analyses for sea surface temperature. *Journal of Climate*, *20*, 54735496. doi: https://doi.org/10.1175/2007JCLI1824.1

Roach, C. J., Balwada, D., & Speer, K. (2018). Global observations of horizontal mixing from argo float and surface drifter trajectories. *Journal of Geophysical Research: Oceans*, *123*. doi: https://doi.org/10.1029/2018JC013750

Rudnickas, D. J., Palter, J., Hebert, D., & Rossby, H. T. (2019). Isopycnal mixing in the north atlantic oxygen minimum zone revealed by rafos floats. *Journal of Geophysical Research: Oceans*, *124*. doi: https://doi.org/10.1029/2019JC015148

Schmidtko, S., Stramma, L., & Visbeck, M. (2017). Decline in global oceanic oxygen content during the past five decades. *Nature Letters*, *542*, 335–339. doi: https://doi.org/10.1038/nature21399

Scott, J. R., & Marotzke, J. (2002). The location of diapycnal mixing and the meridional overturning circulation. *Journal of Physical Oceanography*, *32*, 3578–3595.

Scott, R. B., Goff, J. A., Naveira-Garabato, A. C., & Nurser, A. J. G. (2011). Global rate and spectral characteristics of internal gravity wave generation by geostrophic flow over topography. *Journal of Geophysical Research-Oceans*, *116*, C09029. doi: https://doi.org/10.1029/2011JC007005

Simmons, H. L., Jayne, S. R., St. Laurent, L. C., & Weaver, A. J. (2004). Tidally driven mixing in a numerical model of the general circulation. *Ocean Modelling*, *6*, 245–263.

Sinha, B., Sévellec, F., Robson, J., & Nurser, G. (2020). Surging of global surface temperature due to decadal legacy of ocean heat uptake. *Journal of Climate*. doi: https://doi.org/10.1175/JCLI-D-19-0874.1

Stammer, D., Balmaseda, M., Heimbach, P., Köhl, A., & Weaver, A. (2016). Ocean data assimilation in support of climate applications: status and perspectives. *Ann. Rev. Mar. Sci.*, *8*, 491–518. doi: https://doi.org/10.1146/annurev-marine-122414-034113

Stammer, D., Wunsch, C., Giering, R., Eckert, C., Heimbach, P., Marotzke, J., … Marshall, J. (2002). Global ocean circulation during 1992-1997, estimated from ocean observations and a general circulation model. *Journal of Geophysical Research*, *107*, 3118. doi: https://doi.org/10.1029/2001/JC000888

St. Laurent, L., & Schmitt, R. (1999). The contribution of salt fingers to vertical mixing in the north atlantic tracer release experiment. *Journal of Physical Oceanography*, *29*, 1404–1424.

Treguier, A. M. (1992). Kinetic energy analysis of an eddy resolving, primitive equation model of the north atlantic. *Journal of Geophysical Research*, *97*, 687–701.

Trossman, D. S., Arbic, B. K., Garner, S. T., Goff, J. A., Jayne, S. R., Metzger, E. J., & Wallcraft, A. J. (2013). Impact of parameterized lee wave drag on the energy budget of an eddying global ocean model. *Ocean Modelling*, *72*, 119–142.

Trossman, D. S., Arbic, B. K., Richman, J. G., Garner, S. T., Jayne, S. R., & Wallcraft, A. J. (2016). Impact of topographic internal lee wave drag on an eddying global ocean model. *Ocean Modelling*, *97*, 109–128.

Verdy, A., & Mazloff, M. R. (2017). A data assimilating model for estimating southern ocean biogeochemistry. *Journal of Geophysical Research-Oceans*, *122*, 6968–6988. doi: https://doi.org/10.1002/2016JC012650

Wang, H., Legg, S., & Hallberg, R. (2018). The effect of arctic freshwater pathways on north atlantic convection and the atlantic meridional overturning circulation. *Journal of Climate*, *31*, 5165–5188.

Warner, S. J., & Moum, J. N. (2019). Feedback of mixing to enso phase change. *Geophysical Research Letters*, *46*, 13920–13927. doi: https://doi.org/10.1029/2019GL085415

Waterhouse, A. F., MacKinnon, J. A., Nash, J. D., Alford, M. H., Kunze, E., Simmons, H. L., … Lee, C. M. (2014). Global patterns of diapycnal mixing from measurements of the turbulent dissipation rate. *Journal of Physical Oceanography*, *44*, 1854–1872.

Whalen, C. B., MacKinnon, J. A., Talley, L. D., & Waterhouse, A. F. (2015). Estimating the mean diapycnal mixing using a finescale strain parameterization. *Journal of Physical Oceanography*, *45*, 1174–1188.

Wright, C. J., Scott, R. B., Ailliot, P., & Furnival, D. (2014). Lee wave generation rates in the deep ocean. *Geophysical Research Letters*, *41*, 2434-2440. doi: 10.1002/2013GL059087

Wunsch, C. (2006). In *The ocean circulation inverse problem.* Cambridge University Press. doi: 10.1017/CBO9780511629570

Wunsch, C., & Heimbach, P. (2007, June). Practical global oceanic state estimation. *Physica D: Nonlinear Phenomena*, *230*(1-2), 197–208. doi: 10.1016/j.physd.2006.09.040

Wyrtki, K. (1962). The oxygen minima in relation to ocean circulation. *Deep Sea Research*, *9*, 11–23.

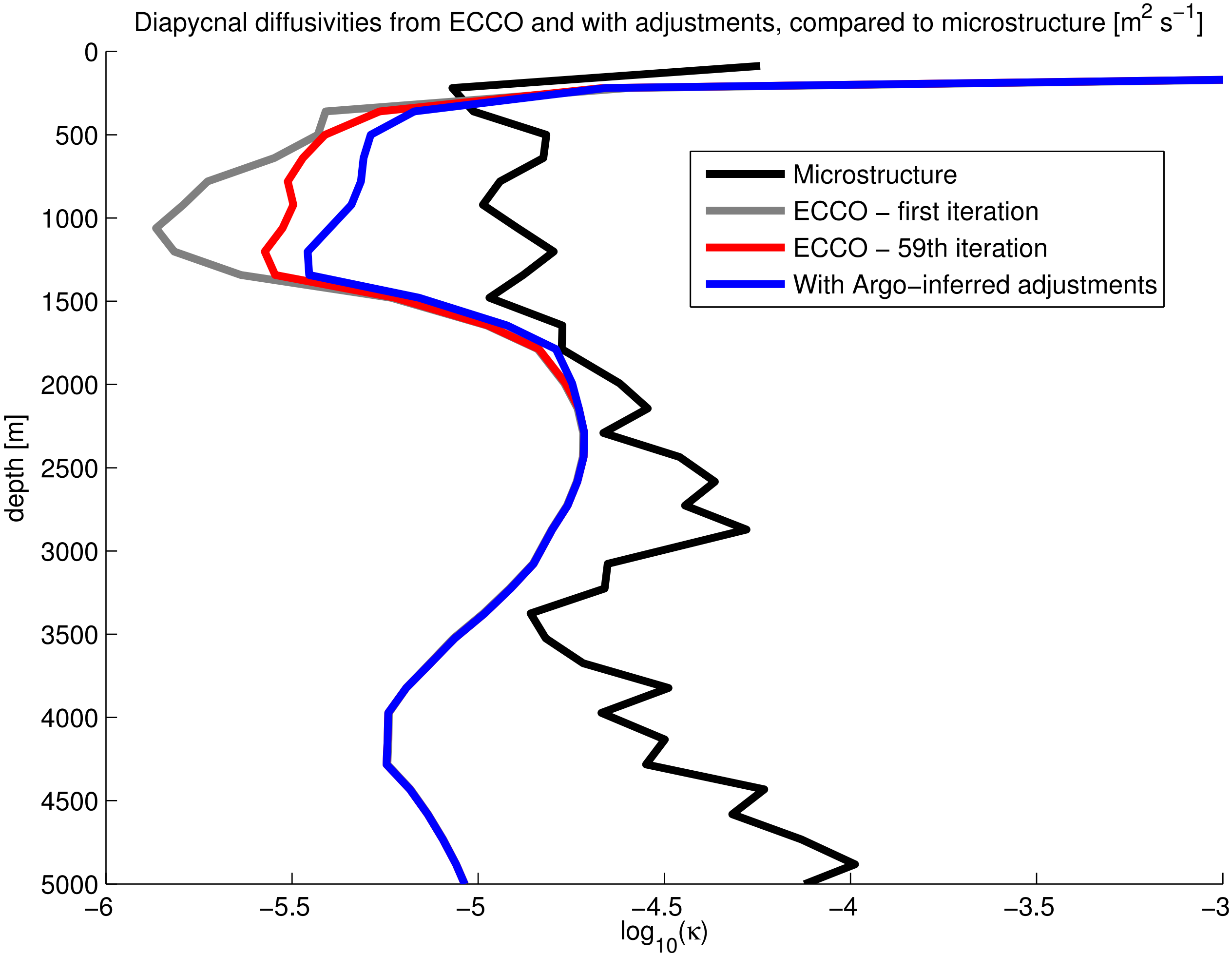
Yang, L., Nikurashin, M., Hogg, A. M., & Sloyan, B. M. (2018). Energy loss from transient eddies due to lee wave generation in the southern ocean. *Journal of Physical Oceanography*, *48*, 2867–2885.

Zhang, R., Kang, S. M., & Held, I. M. (2001). Sensitivity of climate change induced by the weakening of the atlantic meridional overturning circulation to cloud feedback. *Journal of Climate*, 378–389.

Zhu, Y., & Zhang, R.-H. (2020). A modified vertical mixing parameterization for its improved ocean and coupled simulations in the tropical pacific. *Journal of Physical Oceanography*, *49*, 21–37. doi: https://doi.org/10.1175/JPO-D-18-0100.1

Zhu, Y., Zhang, R.-H., & Sun, J. (2020). North pacific upper-ocean cold temperature biases in cmip6 simulations and the role of regional vertical mixing. *Journal of Climate*, *33*, 7523–7538. doi: https://doi.org/10.1175/JCLI-D-19-0654.1

Figure 1.



Diapycnal diffusivities from ECCO and with adjustments, compared to microstructure [m^2 s^{-1}]
Microstructure
ECCO – first iteration
ECCO – 59th iteration
With Argo–inferred adjustments
depth [m]
0
500
1000
1500
2000
2500
3000
3500
4000
4500
5000
−6
−5.5
−5
−4.5
−4
−3.5
−3
$\log_{10}(\kappa)$

Figure 2.

Diapycnal diffusivities from MOM5 and with adjustments, compared to microstructure [$m^2\ s^{-1}$]

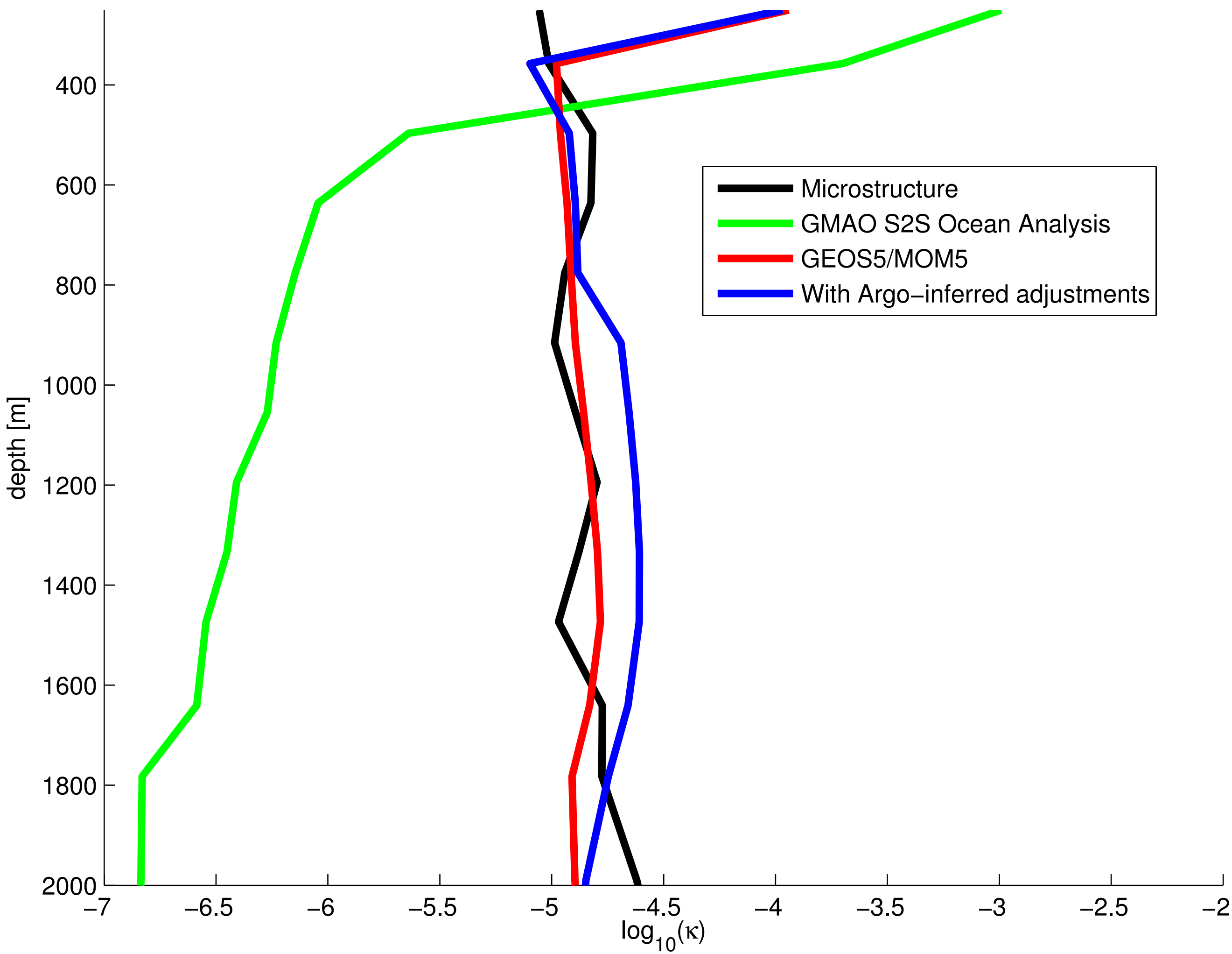
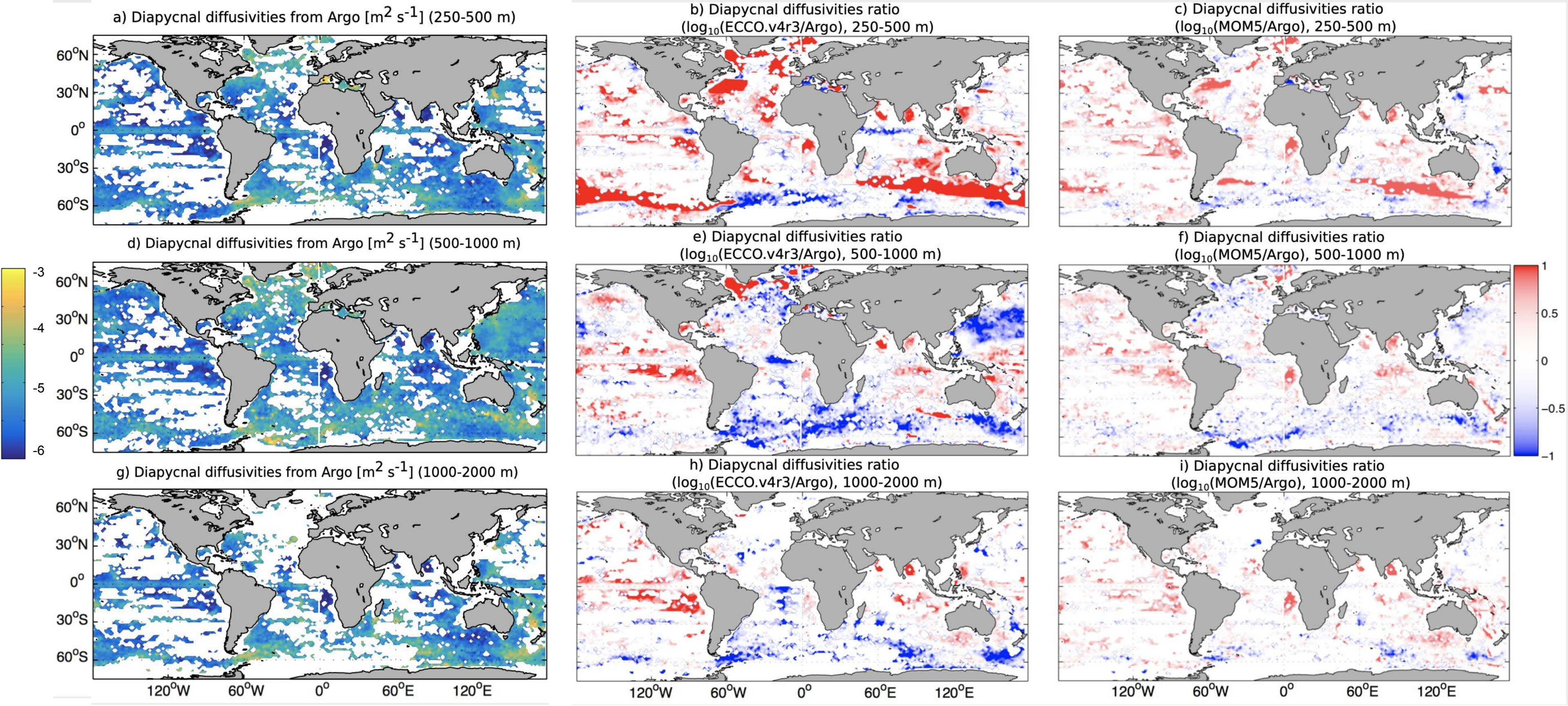
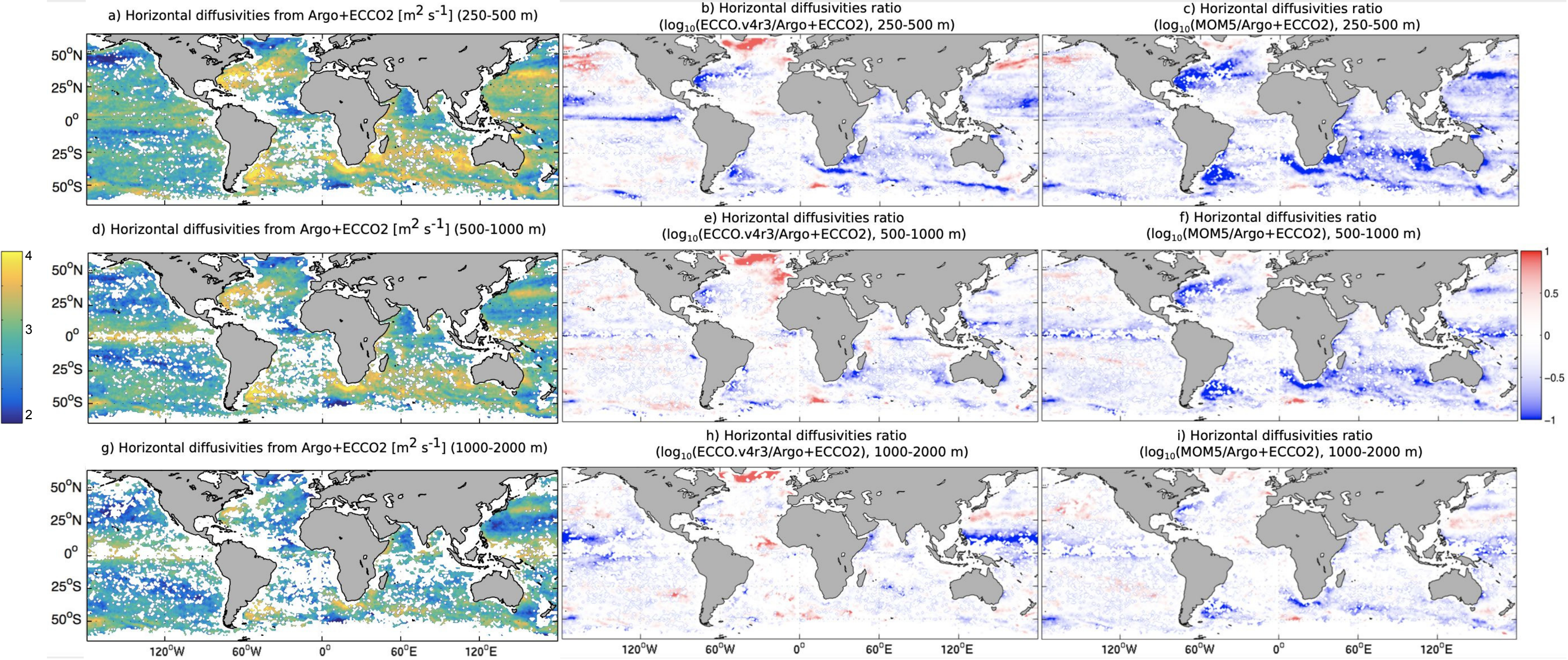


Figure 3.



a) Diapycnal diffusivities from Argo [m2 s-1] (250-500 m)
b) Diapycnal diffusivities ratio
(log10(ECCO.v4r3/Argo), 250-500 m)
c) Diapycnal diffusivities ratio
(log10(MOM5/Argo), 250-500 m)
d) Diapycnal diffusivities from Argo [m2 s-1] (500-1000 m)
e) Diapycnal diffusivities ratio
(log10(ECCO.v4r3/Argo), 500-1000 m)
f) Diapycnal diffusivities ratio
(log10(MOM5/Argo), 500-1000 m)
g) Diapycnal diffusivities from Argo [m2 s-1] (1000-2000 m)
h) Diapycnal diffusivities ratio
(log10(ECCO.v4r3/Argo), 1000-2000 m)
i) Diapycnal diffusivities ratio
(log10(MOM5/Argo), 1000-2000 m)
60°N
30°N
0°
30°S
60°S
120°W
60°W
0°
60°E
120°E
-3
-4
-5
-6
1
0.5
0
−0.5
−1

Figure 4.



a) Horizontal diffusivities from Argo+ECCO2 [m2 s-1] (250-500 m)
b) Horizontal diffusivities ratio (log10(ECCO.v4r3/Argo+ECCO2), 250-500 m)
c) Horizontal diffusivities ratio (log10(MOM5/Argo+ECCO2), 250-500 m)
d) Horizontal diffusivities from Argo+ECCO2 [m2 s-1] (500-1000 m)
e) Horizontal diffusivities ratio (log10(ECCO.v4r3/Argo+ECCO2), 500-1000 m)
f) Horizontal diffusivities ratio (log10(MOM5/Argo+ECCO2), 500-1000 m)
g) Horizontal diffusivities from Argo+ECCO2 [m2 s-1] (1000-2000 m)
h) Horizontal diffusivities ratio (log10(ECCO.v4r3/Argo+ECCO2), 1000-2000 m)
i) Horizontal diffusivities ratio (log10(MOM5/Argo+ECCO2), 1000-2000 m)
50°N
25°N
0°
25°S
50°S
120°W
60°W
0°
60°E
120°E
4
3
2
1
0.5
0
−0.5
−1

Figure 5.


a) Vertically and zonally averaged temperature [ºC]: BKG anomalies and differences between DIFF and CTRL
Vertically and zonally averaged T
Latitude [º]
BKG anomalies
DIFF−CTRL
b) Vertically and zonally averaged salinity [PSS-1978]: BKG anomalies and differences between DIFF and CTRL
Vertically and zonally averaged S
Latitude [º]
BKG anomalies
DIFF−CTRL

Figure 6.

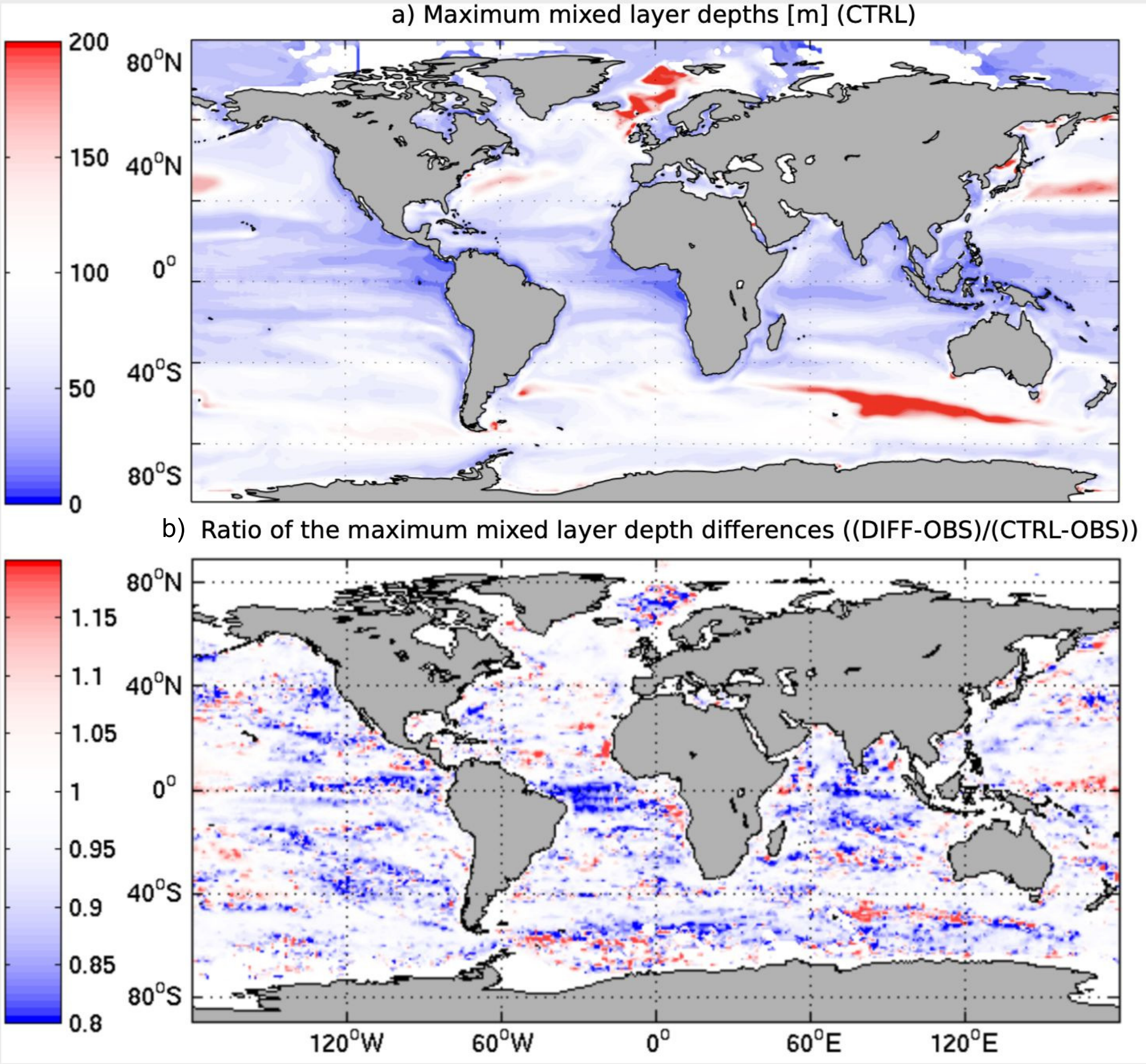
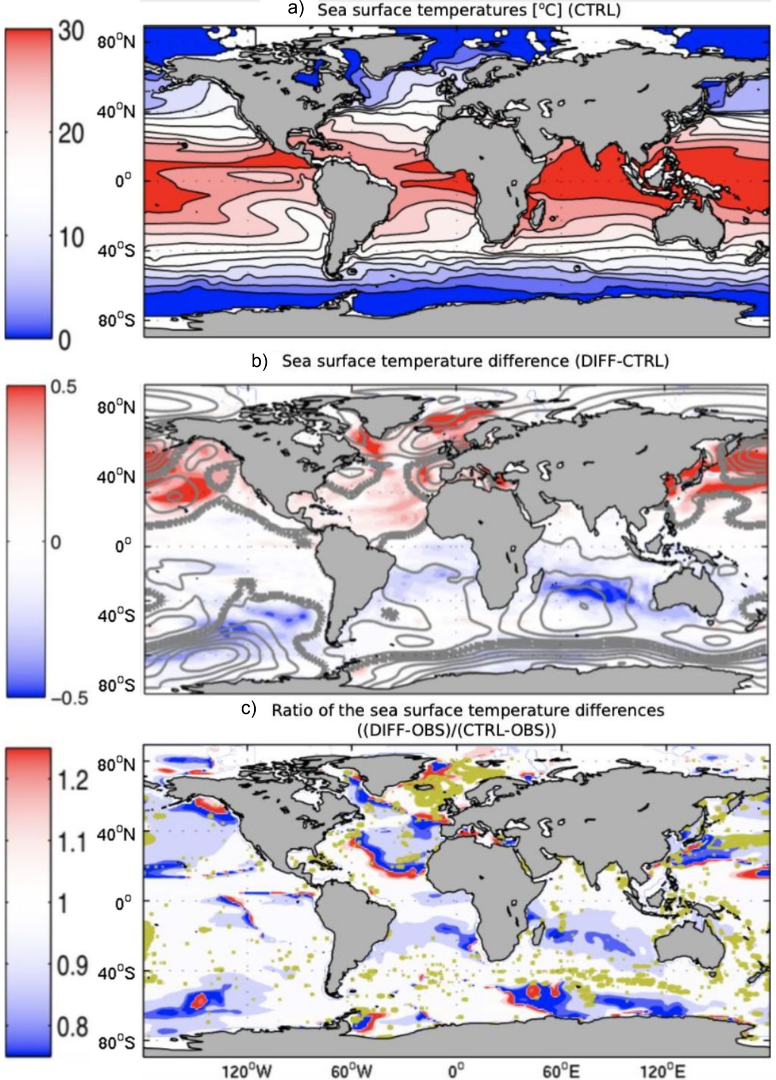

a) Maximum mixed layer depths [m] (CTRL)
200
150
100
50
0
80°N
40°N
0°
40°S
80°S
b) Ratio of the maximum mixed layer depth differences ((DIFF-OBS)/(CTRL-OBS))
1.15
1.1
1.05
1
0.95
0.9
0.85
0.8
80°N
40°N
0°
40°S
80°S
120°W
60°W
0°
60°E
120°E

Figure 7.

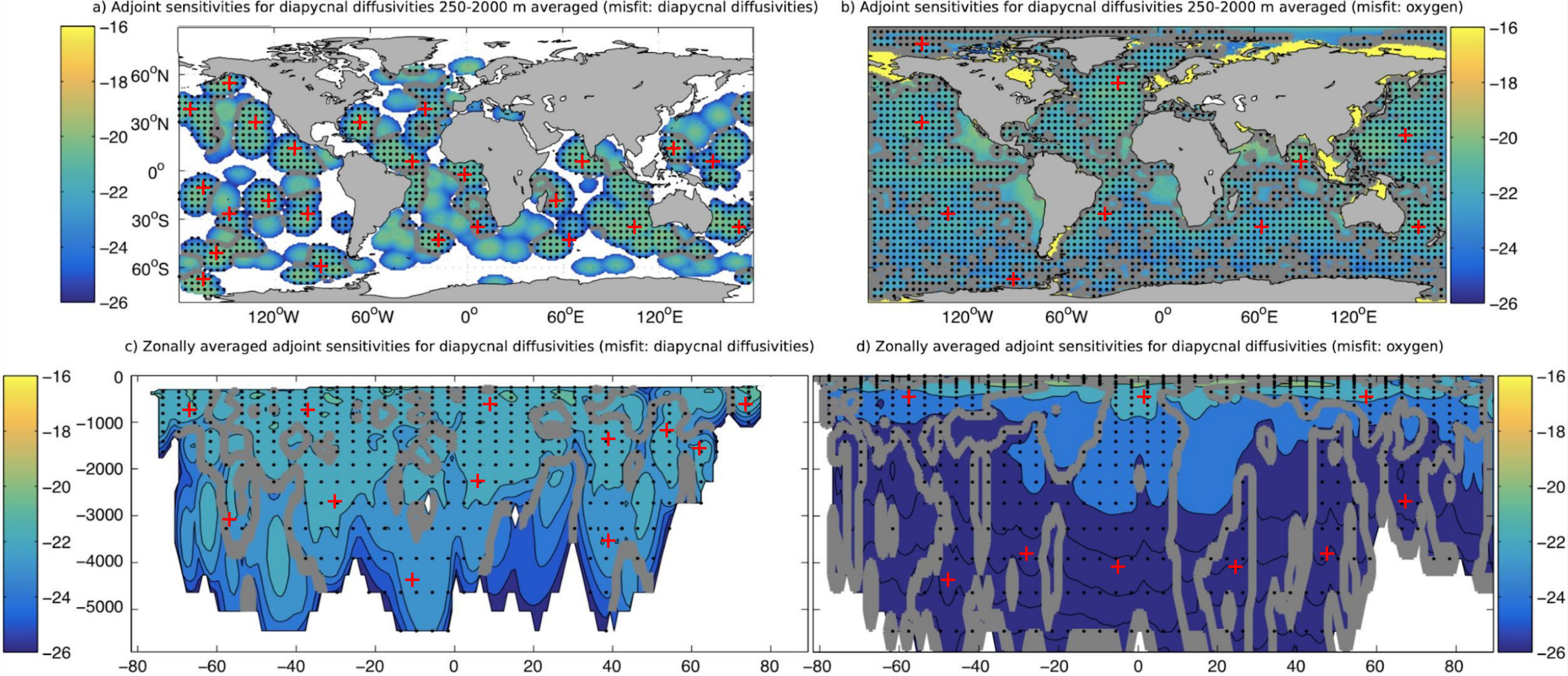


a) Sea surface temperatures [°C] (CTRL)
30
20
10
0
80°N
40°N
0°
40°S
80°S
b) Sea surface temperature difference (DIFF-CTRL)
0.5
0
−0.5
80°N
40°N
0°
40°S
80°S
c) Ratio of the sea surface temperature differences
((DIFF-OBS)/(CTRL-OBS))
1.2
1.1
1
0.9
0.8
80°N
40°N
0°
40°S
80°S
120°W
60°W
0°
60°E
120°E

Figure 8.

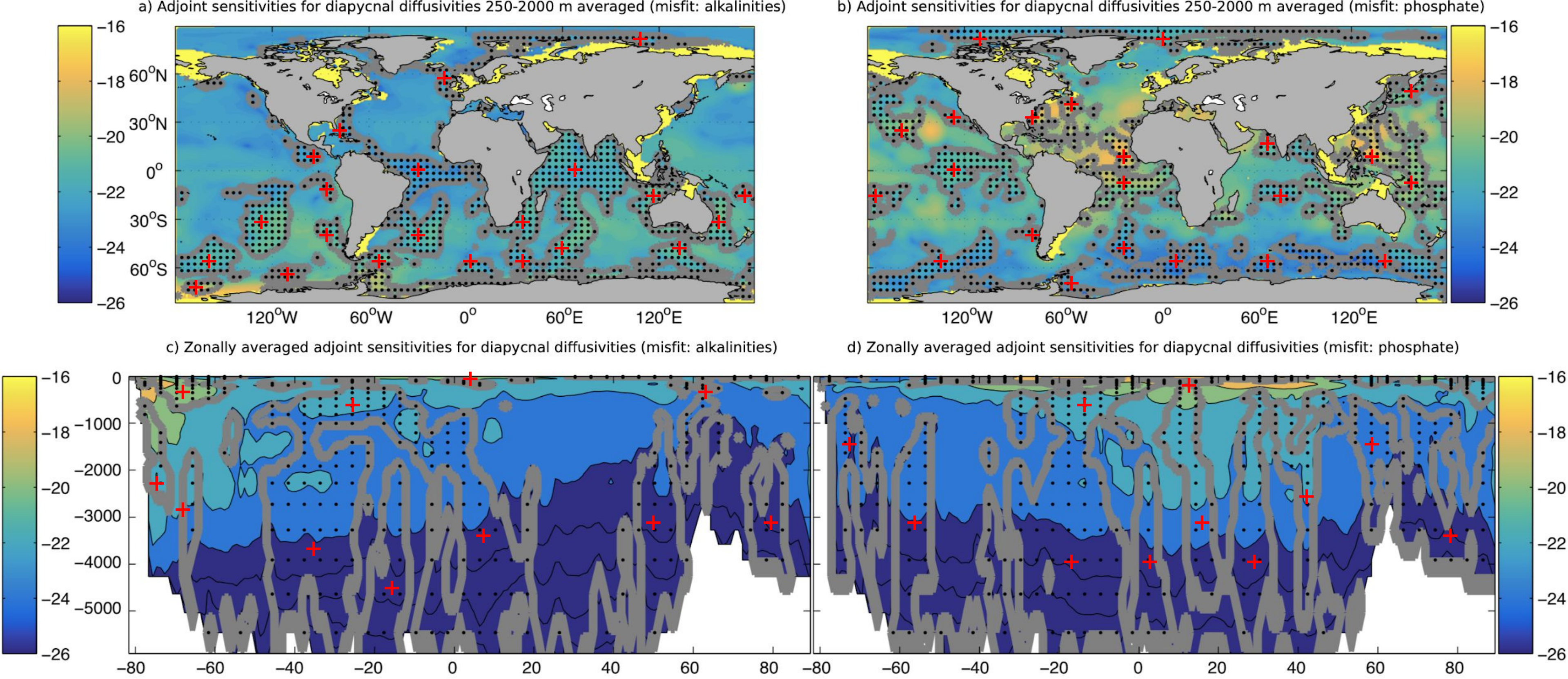

a) Resolved advection (CTRL)
b) Surface flux (CTRL)
c) Neutral diffusion (CTRL)
d) Resolved advection percent difference
((DIFF-CTRL)/(DIFF+CTRL))
e) Surface flux percent difference
((DIFF-CTRL)/(DIFF+CTRL))
f) Neutral diffusion percent difference
((DIFF-CTRL)/(DIFF+CTRL))
Change = 8.84%
Change = 0.48%
Change = 0.26%
0.5
0
-0.5
0.1
0.05
0
-0.05
-0.1
100
50
0
-50
-100
60°N
30°N
0°
30°S
60°S
120°W
60°W
0°
60°E
120°E

Figure 9.



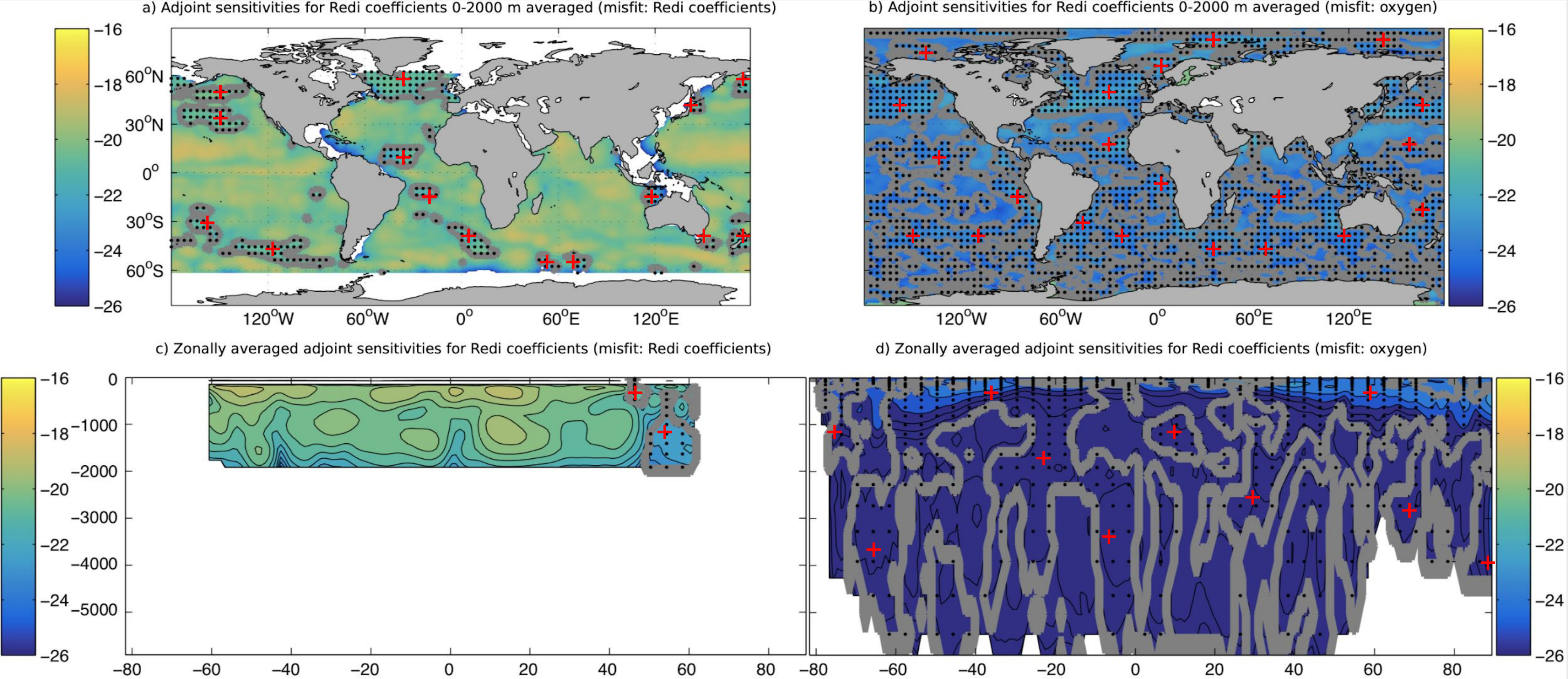
a) Adjoint sensitivities for diapycnal diffusivities 250-2000 m averaged (misfit: diapycnal diffusivities)
b) Adjoint sensitivities for diapycnal diffusivities 250-2000 m averaged (misfit: oxygen)
-16
-18
-20
-22
-24
-26
60°N
30°N
0°
30°S
60°S
120°W
60°W
0°
60°E
120°E
c) Zonally averaged adjoint sensitivities for diapycnal diffusivities (misfit: diapycnal diffusivities)
d) Zonally averaged adjoint sensitivities for diapycnal diffusivities (misfit: oxygen)
0
-1000
-2000
-3000
-4000
-5000
-80
-60
-40
-20
0
20
40
60
80

Figure 10.



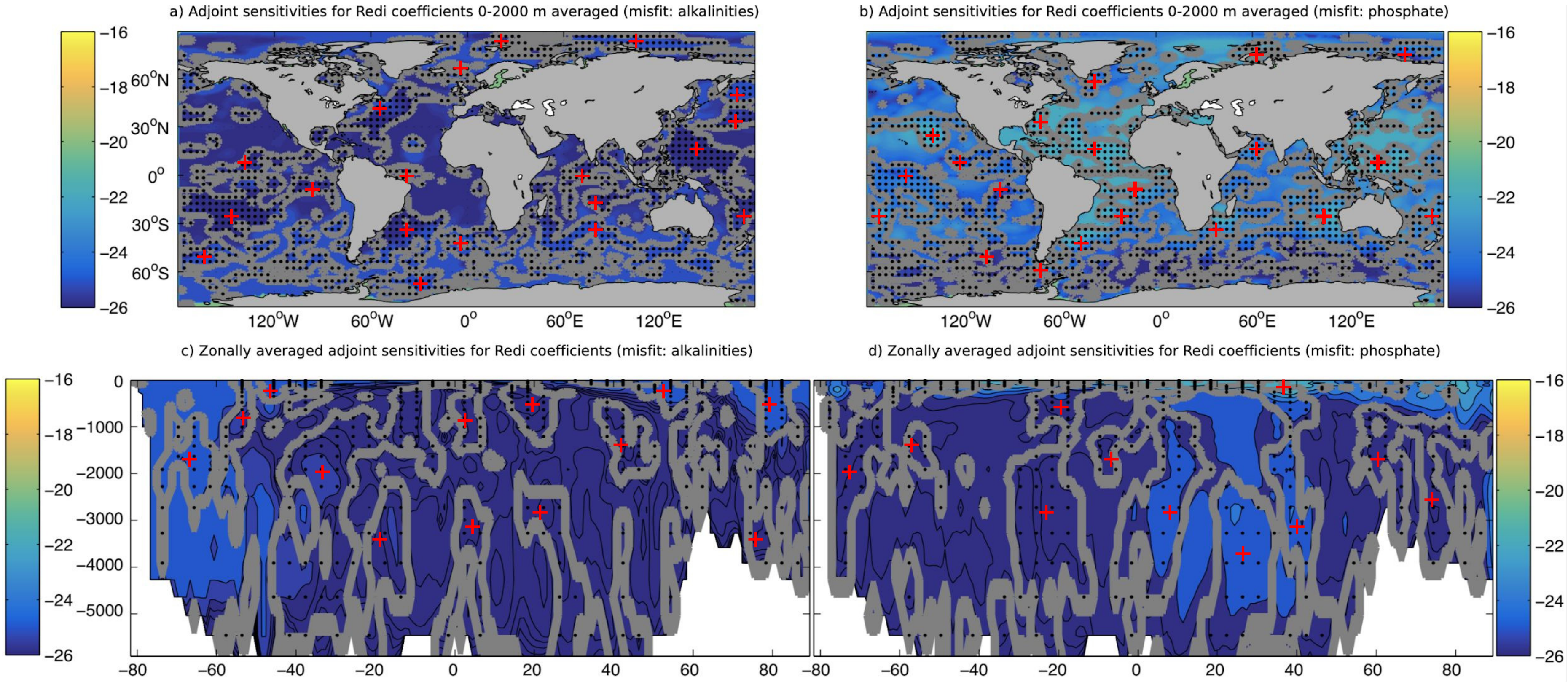
a) Adjoint sensitivities for diapycnal diffusivities 250-2000 m averaged (misfit: alkalinities)
b) Adjoint sensitivities for diapycnal diffusivities 250-2000 m averaged (misfit: phosphate)
60^{0}N
30^{0}N
0^0
30^{0}S
60^{0}S
120^{0}W
60^{0}W
0^0
60^{0}E
120^{0}E
−16
−18
−20
−22
−24
−26
c) Zonally averaged adjoint sensitivities for diapycnal diffusivities (misfit: alkalinities)
d) Zonally averaged adjoint sensitivities for diapycnal diffusivities (misfit: phosphate)
0
−1000
−2000
−3000
−4000
−5000
−80
−60
−40
−20
0
20
40
60
80

Figure 11.



a) Adjoint sensitivities for Redi coefficients 0-2000 m averaged (misfit: Redi coefficients)
b) Adjoint sensitivities for Redi coefficients 0-2000 m averaged (misfit: oxygen)
c) Zonally averaged adjoint sensitivities for Redi coefficients (misfit: Redi coefficients)
d) Zonally averaged adjoint sensitivities for Redi coefficients (misfit: oxygen)
60°N
30°N
0°
30°S
60°S
120°W
60°W
0°
60°E
120°E
−16
−18
−20
−22
−24
−26
0
-1000
-2000
-3000
-4000
-5000
-80
-60
-40
-20
0
20
40
60
80

Figure 12.


a) Adjoint sensitivities for Redi coefficients 0-2000 m averaged (misfit: alkalinities)
b) Adjoint sensitivities for Redi coefficients 0-2000 m averaged (misfit: phosphate)
c) Zonally averaged adjoint sensitivities for Redi coefficients (misfit: alkalinities)
d) Zonally averaged adjoint sensitivities for Redi coefficients (misfit: phosphate)
−16
−18
−20
−22
−24
−26
60°N
30°N
0°
30°S
60°S
120°W
60°W
0°
60°E
120°E
0
-1000
-2000
-3000
-4000
-5000
-80
-60
-40
-20
0
20
40
60
80