

Stochastic Subgrid-Scale Ocean Mixing: Impacts on Low-Frequency Variability

STEPHAN JURICKE, TIM N. PALMER, AND LAURE ZANNA

Atmospheric, Oceanic and Planetary Physics, Department of Physics, University of Oxford, Oxford, United Kingdom

(Manuscript received 22 July 2016, in final form 1 March 2017)

ABSTRACT


In global ocean models, the representation of small-scale, high-frequency processes considerably influences the large-scale oceanic circulation and its low-frequency variability. This study investigates the impact of stochastic perturbation schemes based on three different subgrid-scale parameterizations in multidecadal ocean-only simulations with the ocean model NEMO at 1° resolution. The three parameterizations are an enhanced vertical diffusion scheme for unstable stratification, the Gent–McWilliams (GM) scheme, and a turbulent kinetic energy mixing scheme, all commonly used in state-of-the-art ocean models. The focus here is on changes in interannual variability caused by the comparatively high-frequency stochastic perturbations with subseasonal decorrelation time scales. These perturbations lead to significant improvements in the representation of low-frequency variability in the ocean, with the stochastic GM scheme showing the strongest impact. Interannual variability of the Southern Ocean eddy and Eulerian streamfunctions is increased by an order of magnitude and by 20%, respectively. Interannual sea surface height variability is increased by about 20%–25% as well, especially in the Southern Ocean and in the Kuroshio region, consistent with a strong underestimation of interannual variability in the model when compared to reanalysis and altimetry observations. These results suggest that enhancing subgrid-scale variability in ocean models can improve model variability and potentially its response to forcing on much longer time scales, while also providing an estimate of model uncertainty.


1. Introduction

One of the big challenges in ocean modeling is the development of parameterizations that can accurately represent the impact of the unresolved subgrid-scale processes on resolved scales. An inadequate representation of such processes can have an impact on the simulated climatic mean state and variability (see, e.g., Kirtman et al. 2012) as well as the models' climatic response to forcing (e.g., Griffies et al. 2015). The ocean varies on long time scales and exhibits mesoscale eddies that are much smaller than those found in the atmosphere. Consequently it is very difficult and computationally expensive to integrate

the governing equations with an ocean model resolution that adequately resolves mesoscale or submesoscale processes on time scales from years to decades. This is reflected in the rather coarse ocean resolutions in the last coupled model intercomparison project (CMIP5; see Flato et al. 2013).

However, high-resolution model simulations have been carried out, analyzed (e.g., Bishop et al. 2016), and compared with lower-resolution simulations, revealing significant differences between simulations with the same model but differing grid resolution (e.g., Seo et al. 2006; Kirtman et al. 2012; Griffies et al. 2015). Additionally, idealized high-resolution ocean models have been used in conjunction with theoretical considerations to gain further insight into mesoscale and submesoscale processes and to investigate new parameterization frameworks (e.g., Marshall et al. 2012). Such simulations have guided developments in so-called scale-aware parameterizations (e.g., Fox-Kemper et al. 2014; Porta

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

 Supplemental information related to this paper is available at the Journals Online website: <http://dx.doi.org/10.1175/JCLI-D-16-0539.s1>.

Corresponding author: Stephan Juricke, stephan.juricke@physics.ox.ac.uk

DOI: 10.1175/JCLI-D-16-0539.1

© 2017 American Meteorological Society



This article is licensed under a [Creative Commons Attribution 4.0 license](http://creativecommons.org/licenses/by/4.0/) (<http://creativecommons.org/licenses/by/4.0/>).

Mana and Zanna 2014), which scale with the model resolution as more and more of the parameterized processes become explicitly resolved with finer grids. In addition, global satellite and in situ observational data can be used to calibrate and develop parameterizations, as well as to validate full climate models. Ocean observations of the subsurface are still rather sparse, even if the advent of satellite measurements allows for large-scale coverage of high-resolution surface properties [e.g., the Archiving, Validation and Interpretation of Satellite Oceanographic (AVISO) data; <http://www.aviso.altimetry.fr>]. However, even when constrained by high-resolution model simulations and observational estimates, uncertainties in ocean models remain large. This is reflected in the inaccuracies of seasonal forecasts (Kim et al. 2012) as well as climate model biases (Richter 2015) and the large spread in ocean reanalysis products (Balmaseda et al. 2015; Karspeck et al. 2017).

In the atmospheric modeling community, especially in weather forecasting, it has become common standard to account for uncertainties in forecasts by running ensemble simulations (e.g., Molteni et al. 1996). At first, such ensembles only perturbed initial conditions. But in the past two decades there has also been an increasing effort to develop and implement schemes and methods that account for the uncertainties of the models themselves. This may be tackled by running multimodel ensembles on climate time scales, as is done in CMIP5 and has been used for climate projection purposes by the Intergovernmental Panel on Climate Change (IPCC; Flato et al. 2013), which also account for uncertainties in the boundary conditions such as future anthropogenic greenhouse gas and aerosol forcings. To account for uncertainties in the parameterized processes, or in the choice of specific model parameters, multiparameterization (also called multiphysics; e.g., Berner et al. 2011) or perturbed parameter ensembles (e.g., Murphy et al. 2004) have been carried out. Finally, another approach to account for model uncertainty originating from the imperfect representation of the subgrid scales is stochastic parameterizations. These have been successfully used in atmospheric models to increase ensemble spread and therefore model reliability (e.g., Palmer et al. 2005) by providing a better estimate of the forecast error and uncertainty. It has been shown that stochastic parameterizations can reduce forecast errors as well as climate biases (e.g., Weisheimer et al. 2014) by better representing the variability of the unresolved scales.

For ocean models, similar developments have been made in recent years, ranging from stochastic parameterizations in simplified setups (e.g., Pasquero and Tziperman 2007; Kitsios et al. 2013; Grooms and Majda 2013; Grooms 2016) and theoretical frameworks

(e.g., Berloff and McWilliams 2002) to implementations in global uncoupled (Brankart 2013; Brankart et al. 2015) and coupled general circulation models (GCMs) (Williams 2012; Andrejczuk et al. 2016; Williams et al. 2016). These studies assume that the unresolved variability in the subgrid scales and the related error may have an impact on the resolved flow variability and its mean state. This becomes especially important when a general assumption of deterministic parameterizations—each model grid box contains a sufficiently large number of events of a given subgrid-scale process—breaks down, so that averaging may not be an accurate representation of the cumulative effect of subgrid-scale processes anymore [for an expanded discussion, see Palmer (2012)].

In this study we investigate whether and how adding stochastic subgrid-scale variability in mixing and eddy parameterizations of a global non-eddy-resolving ocean model affects the model integrations, especially with regards to low-frequency variability (i.e., interannual and longer time scale variability). We follow an approach that largely keeps the formulation of the models parameterizations untouched, but identifies important and imperfectly constrained parameters or tendencies in these parameterizations and adds stochastic variability. Therefore, the deterministic averaging used to parameterize effects of the unresolved process is not altered, but unresolved variability is reintroduced to the parameterization. Instead of developing and implementing a new parameterization, we will specifically focus on adding unresolved variability of eddy and convective time scales in the already existing parameterization formulations. This approach also provides a quantification of model uncertainty arising from the imperfect representation of the respective parameterized subgrid-scale process.

The paper is organized as follows. In section 2 we briefly describe the general model setup of the ocean model known as NEMO. Section 3 introduces the stochastic perturbation approach and its application to three different parameterization schemes used in NEMO. The results are presented in section 4 and summarized in section 5, in which we discuss the implications of the results and future perspectives.

2. Model setup

The ocean model used for this study is the Nucleus for European Modelling of the Ocean (NEMO), version 3.3.1 (Madec 2008). This specific version is used in the coupled climate model EC-EARTH, version 3.1 (see www.ec-earth.org; Hazeleger et al. 2012). NEMO is a primitive equation GCM of the ocean. The model grid is the ORCA1L46 with the horizontal discretization

carried out on the approximately 1° tripolar grid and with 46 levels in the vertical, using partial step z coordinates. The horizontal grid has a meridional refinement from 1° to $\frac{1}{3}^\circ$ near the equator (Madec and Imbard 1996). The model time step is 1 h. Different parameterizations can be chosen to account for unresolved processes, some of which will be described in more detail in the next section. For a detailed discussion of the model discretization, boundary conditions, and the parameterization schemes available for NEMO, we refer the reader to Madec (2008). The ocean model is coupled to the Louvain-la-Neuve Sea Ice Model, version 3 (LIM3; Vancoppenolle et al. 2009). The atmospheric forcing applied to drive the ocean model is the Drakkar Forcing Set, version 4.3 (DFS4.3; Brodeau et al. 2010), making use of the Co-ordinated Ocean–Ice Reference Experiments (CORE) bulk formulation (Griffies et al. 2009). The model uses surface salinity restoring and also sets the global mean freshwater budget (evaporation minus precipitation minus runoff) to zero at each time step to ensure that there is no continuous drift in sea surface height (SSH).

The default model setup is integrated for 30 yr starting from 1 January 1960 where the temperature and salinity are taken from the January climatology of the *World Ocean Atlas 2009* (Locarnini et al. 2010; Antonov et al. 2010). To test the stochastic perturbation schemes, multiple simulations were carried out after this initial spinup. They are described in the respective sections below. Most diagnostics presented here—as also stated in the respective figure captions—are based on the 5-day averaged output from these model simulations. A mean state comparison of the model simulations (described in section 4b) with two reanalysis datasets is presented in appendix A.

3. Stochastic perturbation schemes

a. General perturbation approach

Here we discuss the inclusion of subgrid-scale variability and uncertainty estimation in some of the deterministic parameterizations of the ocean model. The general approach we follow in this study is to take a deterministic parameter or a tendency and perturb it by multiplying it with a random number, that is,

$$P^{\text{sto}}(i, j) = [1 + \xi(i, j)]P^{\text{ref}}(i, j), \quad (1)$$

where i is the time step, j is the grid node index, P^{ref} is the deterministic parameter or tendency that may depend on time and location, ξ is a random number sampled from a symmetric distribution with zero mean, and P^{sto} is the resulting stochastically perturbed parameter or

tendency. The multiplication with the random number ξ takes place at every time step and at every grid node. This general approach has been used in previous studies by, for example, Juricke et al. (2013) for sea ice modeling and Ollinaho et al. (2016) for atmospheric modeling.

There are a few properties of the random number that need to be defined, namely temporal and spatial correlations, and the shape of the distribution. The temporal correlation is given by a first-order autoregressive process:

$$\tilde{\xi}(i, j) = (1 - \Delta t/\tau)\tilde{\xi}(i - 1, j) + \eta = \alpha\tilde{\xi}(i - 1, j) + \eta, \quad (2)$$

where Δt is the model time step and η is a random number with a Gaussian distribution of zero mean and standard deviation $\sigma\Delta t$, with σ being a tunable noise amplitude parameter. The temporal decorrelation time is given by τ , with $\alpha = (1 - \Delta t/\tau)$. Finally, $\tilde{\xi}$ is a Gaussian random number with variance

$$\sigma_{\text{max}}^2 = (\sigma\Delta t)^2/(1 - \alpha^2), \quad (3)$$

for $i \rightarrow \infty$. The spatial correlation is introduced to the system by η . There are no variations in the random numbers in the vertical (i.e., at a specific horizontal location the same random number is used for the full vertical column). The horizontal spatial correlation of η is then defined by a correlation matrix \mathbf{C} with entries

$$c_{nm} = \exp(-r_{nm}/\Delta r), \quad (4)$$

where n and m are horizontal indices of a coarse-resolution grid, r_{nm} is the distance between the two points n and m , and Δr is a spatial decorrelation length scale. The coarse-resolution grid in this study is a regular $6^\circ \times 6^\circ$ grid, to reduce the computational burden of calculating the spatial correlations. After the spatial correlation matrix is generated at the beginning of the simulation, a Cholesky decomposition of the matrix is carried out. After that a matrix–vector multiplication can generate a vector of spatially correlated random numbers from a vector of statistically independent Gaussian random numbers at every time step [see also Juricke and Jung (2014)]. The spatially correlated random numbers of the $6^\circ \times 6^\circ$ grid are then interpolated to the $1^\circ \times 1^\circ$ NEMO grid by calculating a weighted average of the three nearest grid nodes of the coarse-resolution grid. Figure S3 in the supplementary material illustrates such a spatially correlated pattern.

The spatial correlation generated by Eq. (4) is simulating nonlocal properties of the unresolved subgrid turbulence. Zidikheri and Frederiksen (2009) derived and analyzed the vertical and horizontal structure of

their stochastic subgrid parameterization in a simplified spectral model with two vertical levels. Aside from nonlocal effects in the horizontal related to the wave-number dependence of the subgrid-scale interactions, vertical variations should be considered as well. However, as mentioned earlier, in the present study we neglect a more complex vertical structure of the perturbations and assume perfect correlation (i.e., the same random number throughout the vertical water column).

Using Eq. (2) spatially and temporally correlated random numbers ξ can be generated. At the beginning of each simulation Eq. (2) is run a few hundred times to converge the variance of the process to the one described by Eq. (3).

After the ξ values have been generated, they need to be transformed into a distribution with a limited range $(-a, a)$, with $0 < a < 1$, to avoid $1 + \xi(i, j)$ becoming negative. For this purpose we apply the transformation described by Juricke et al. (2013):

$$\xi(i, j) = -a + (2a)/\{1 + \exp[-\beta\tilde{\xi}(i, j)]\}, \quad (5)$$

where β is a regularization factor to ensure a Gaussian-like shape of the distribution of $\xi(i, j)$, which is set to be $1.2/\sigma_{\max}$ [see appendix A of Juricke et al. (2013) for more details].

Therefore the crucial parameters of the perturbation scheme that need to be defined are temporal decorrelation time scale τ , spatial decorrelation length scale Δr , and perturbation amplitude a . Note that the noise amplitude σ does not have an impact on the actual size of the perturbations ξ since the regularization factor β takes the value of σ into account through σ_{\max} , thereby making the amplitude of the perturbations solely dependent on a .

Following this explanation of the general approach, the ways in which we perturbed the specific parameterizations are discussed in the following subsections. The dependence of the parameters on (i, j) are no longer included even though the equations should always be viewed in the discretized framework.

b. Stochastic perturbations to Gent–McWilliams parameterization amplitude

The Gent–McWilliams parameterized bolus or eddy-induced velocity is implemented in NEMO as an additional advective term in the prognostic tracer equations for salinity and temperature. The implementation in NEMO is based on Gent and McWilliams (1990, hereafter GM) with a spatially and temporally varying coefficient (Treguier et al. 1997). The bolus velocity represents a tracer advection contribution originating from unresolved mesoscale eddies. The tracer

equations for temperature T and salinity S read, in simplified form,

$$\frac{\partial X}{\partial t} = -\nabla \cdot (X\mathbf{U} + X\mathbf{U}_{\text{GM}}) + M_X + F_X, \quad (6)$$

where X is either salinity or temperature, t is time, $\mathbf{U} = (u, v, w)$ is the three-dimensional ocean velocity vector, $\mathbf{U}_{\text{GM}} = (u_{\text{GM}}, v_{\text{GM}}, w_{\text{GM}})$ is the three-dimensional GM bolus velocity, M_X is the sum of all contributions from subgrid-scale parameterizations of three-dimensional mixing and diffusion, and F_X are the remaining forcings. In these equations the GM parameterization adds another advective term with an advective bolus velocity \mathbf{U}_{GM} . The bolus or eddy-induced velocity is described by

$$u_{\text{GM}} = -\frac{\partial}{\partial z}(A_{\text{GM}}S_x), \quad (7)$$

$$v_{\text{GM}} = -\frac{\partial}{\partial z}(A_{\text{GM}}S_y), \quad \text{and} \quad (8)$$

$$w_{\text{GM}} = +\frac{\partial}{\partial x}(A_{\text{GM}}S_x) + \frac{\partial}{\partial y}(A_{\text{GM}}S_y), \quad (9)$$

where (x, y, z) is the three-dimensional Cartesian coordinate vector, and S_x and S_y are the horizontal isoneutral slopes in the x and y direction, respectively. The GM eddy-induced velocity is nondivergent and its amplitude is defined by the parameterization coefficient A_{GM} , which in turn is (for the specific configuration of NEMO used in this study) two-dimensional and temporally varying [adapted from Treguier et al. (1997)], while constant throughout the same vertical column, and approximated by

$$A_{\text{GM}} = \lambda^2 \Lambda^{-1}, \quad (10)$$

with λ being the local Rossby radius of deformation and Λ the time scale of baroclinic eddies. The coefficient is tapered near the boundaries and additional adjustments are applied in the tropics, where the coefficient is reduced. Additionally, the steepness of the slopes is bounded everywhere in the ocean. We refer the reader to Madec (2008) and the NEMO documentation within the NEMO source code for a more detailed description of the actual implementation.

In summary, the parameterization represents the impact of the subgrid mesoscale eddies on the mixing of salinity and temperature in a non-eddy-resolving model. Given a certain mean background state in temperature, velocity, and salinity, the parameterization provides a deterministic mean eddy-induced transport, designed to reduce mean potential energy. As a representation of the mean effect of baroclinic instability, it converts available potential energy into kinetic energy. While

Eq. (10) provides an estimate of the order of magnitude of the GM coefficient, it is largely based on an empirical choice with many attempts to find best fits (Ferreira and Marshall 2006; Bachman et al. 2017, and references therein). Bachman et al. (2017) as well as others show that there is a large scatter in the GM coefficient over an order of magnitude and more. Furthermore, the GM parameterization is only representing the aforementioned mean effect of eddy induced transport.

Consequently, the stochastic perturbation schemes described in the previous subsection can be used to incorporate two effects into the ocean simulations, namely, the subgrid-scale variability and uncertainty of the GM coefficient as well as a dynamically evolving state-dependent uncertainty estimation of the parameterization as a whole. In the case of the GM parameterization, we choose the deterministic coefficient $A_{\text{GM}}^{\text{ref}}$ [see Eq. (10) and remarks above] to be perturbed as

$$A_{\text{GM}}^{\text{sto}} = (1 + \xi_{\text{GM}})A_{\text{GM}}^{\text{ref}}, \quad (11)$$

following Eq. (1). The generation of the random number ξ_{GM} is described by Eqs. (2) and (5).

Effectively the perturbation to the GM coefficient is designed to simulate variations and uncertainties in the strength of the GM parameterization. More specifically, in the parameterization of the GM coefficient used in this study [i.e., Eq. (10)] the perturbations enhance variations in the time scale of baroclinic eddies Λ , which is calculated as

$$\Lambda = (\sqrt{N^2 S})^{-1}, \quad (12)$$

with N as the local Brunt–Väisälä frequency. Here the overbar denotes the vertical average for N^2 and $S = S_x^2 + S_y^2$ over all vertical levels. Because of this averaging process and the related uncertainties in the calculation of the mean slope and the mean buoyancy frequency over each grid box, the resulting time scale is smoothed in its temporal and spatial variation, especially in regions of strong local variability in these properties. Previous studies of eddy statistics (e.g., Cheng et al. 2014; Samelson et al. 2016) have shown that statistics of eddy lifetime and amplitude are highly non-Gaussian and skewed with long tails for high values. Applying a Gaussian-like perturbation to $A_{\text{GM}}^{\text{ref}}$ corresponds to a skewed perturbation to Λ , as $A_{\text{GM}}^{\text{ref}}$ is inversely proportional to Λ (see Fig. B1). Following this approach, we keep the perturbations to $A_{\text{GM}}^{\text{ref}}$ symmetric, reducing potential mean drifts in the stochastic setup compared to the deterministic setup, while still simulating a skewed perturbation to the eddy time scales. The assumption here is that especially the long

tail in the eddy statistics is crucially underestimated in the 1° ocean model. (See also appendix B for further discussion.) The effective symmetric perturbation amplitude also agrees with the estimates of GM coefficient uncertainty from aforementioned studies (e.g., Ferreira and Marshall 2006; Bachman et al. 2017). Applying perturbations to the eddy time scale in the calculation of the GM coefficient might be tied to this specific choice of parameterizing the GM coefficient. The general resulting amplitude of the coefficient perturbation, however, can be easily extended to other setups where the coefficient is estimated differently.

A similar perturbation strategy as well as an alternative approach based on products of Gaussian variables to perturb GM has concurrently been developed and tested by Grooms (2016) in an idealized box ocean model configuration. Furthermore, Jansen (2017) suggested as a reply and extension to the study by Grooms (2016) that a stochastic mixing length model should be used to compute eddy fluxes. For the mean buoyancy fluxes such a model reduces to the GM scheme.

Perturbations of the GM coefficient are only simulating some of the variability and uncertainty in the very specific eddy process that is parameterized by GM. We are not simulating other eddy effects on low-frequency ocean variability such as those generated by eddy interaction with baroclinic Rossby waves. The aim of the present approach is to use the current parameterization setup and enhance subgrid-scale parameter variability rather than to develop a new parameterization. However, these aspects are certainly worth further investigation.

c. Stochastic perturbations to the turbulent kinetic energy vertical mixing scheme

The NEMO configuration of this study uses a prognostic equation of turbulent kinetic energy (TKE) to estimate the amplitude of the vertical diffusion and viscosity coefficients for the diffusion of tracers and momentum, respectively (Bougeault and Lacarrere 1989; Gaspar et al. 1990; Blanke and Delecluse 1993; Madec et al. 1998). The TKE is evolved through

$$\begin{aligned} \frac{\partial e}{\partial t} = & A_{\text{vm}} \left[\left(\frac{\partial u}{\partial z} \right)^2 + \left(\frac{\partial v}{\partial z} \right)^2 \right] - A_{\text{vt}} N^2 \\ & + \frac{\partial}{\partial z} \left(A_{\text{vm}} \frac{\partial e}{\partial z} \right) - c_\epsilon \frac{e^{3/2}}{l_\epsilon}, \end{aligned} \quad (13)$$

where e is the mean turbulent kinetic energy, A_{vm} and A_{vt} are the vertical eddy viscosity and diffusivity coefficients for the momentum and tracer equations,

respectively, l_ε is the turbulent dissipation length scale, and $c_\varepsilon = 0.7$ is a constant coefficient. The first term on the right-hand side corresponds to the production of TKE by vertical shear, the second term is a TKE destruction term through stratification, the third term is viscosity, and the last term corresponds to Kolmogorov dissipation (Kolmogorov 1942).

The coefficients A_{vm} and A_{vT} are then calculated by

$$A_{vm} = c_k l_k \sqrt{e} \quad \text{and} \quad A_{vT} = A_{vm} / P_{rt}, \quad (14)$$

with $c_k = 0.1$ a constant, l_k the turbulent mixing length scale, and P_{rt} the Prandtl number (see also Madec et al. 1998).

There are several other parts to this turbulent closure, such as minimum background vertical diffusion and viscosity coefficients, different options to parameterize the vertical turbulent length scales, and additional parameterizations for surface wave breaking at the upper boundary and the effects of Langmuir cells (see Madec 2008). However, since these additions to the TKE scheme do not directly affect the stochastic perturbation scheme introduced in the next paragraph, we refer the reader to Madec (2008).

The stochastic perturbations to the TKE scheme are applied to the tendencies of the vertical shear and stratification separately, in a similar fashion as the perturbations to the GM coefficient in the last subsection, that is,

$$\begin{aligned} P_{tend}^{sto} &= (1 + \xi_{ProdTKE}) P_{tend}^{ref} \\ &= (1 + \xi_{ProdTKE}) \left\{ A_{vm} \left[\left(\frac{\partial u}{\partial z} \right)^2 + \left(\frac{\partial v}{\partial z} \right)^2 \right] \right\} \quad \text{and} \end{aligned} \quad (15)$$

$$D_{tend}^{sto} = (1 + \xi_{DestTKE}) D_{tend}^{ref} = (1 + \xi_{DestTKE}) A_{vT} N^2, \quad (16)$$

with P_{tend}^{sto} and P_{tend}^{ref} being the perturbed and unperturbed TKE production tendencies through shear, D_{tend}^{sto} and D_{tend}^{ref} the perturbed and unperturbed TKE destruction tendencies through stratification, and $\xi_{ProdTKE}$ and $\xi_{DestTKE}$ the respective locally and temporally varying random numbers for both terms, whose evolution is described by Eqs. (2) and (5). The perturbations simulate unresolved subgrid-scale variability and uncertainty in the production and destruction terms of TKE. These contributions to the TKE budget have been studied based on local measurements by, for example, Goodman et al. (2006) for two locations along the northeast coast of the United States. Observed along-track estimates of the two tendencies show large variations on scales of hundreds of meters [see Figs. 9 and 15 in Goodman et al. (2006)]. This study motivates our choice of the

perturbations and associated amplitudes of up to 80% to both tendencies. Also, since these terms are both solved with an explicit forward scheme, implementing the perturbations is straightforward.

d. Stochastic perturbations to the enhanced vertical diffusion scheme

The third parameterization scheme that was stochastically perturbed was the enhanced vertical diffusion (EVD) scheme in situations of unstable stratification. If the local Brunt–Väisälä frequency is negative (i.e., in the case of instability), the vertical tracer diffusivity coefficient A_{vT} is increased strongly, but not the vertical momentum diffusivity coefficient. Here the value for tracer diffusivity is set to $A_{vT}^{refMAX} = 50 \text{ m}^2 \text{ s}^{-1}$ in those situations.

The perturbation scheme is then applied to A_{vT}^{refMAX} :

$$A_{vT}^{stoMAX} = (1 + \xi_{EVD}) A_{vT}^{refMAX}, \quad (17)$$

where ξ_{EVD} is the locally and temporally varying random number for the EVD scheme, and A_{vT}^{stoMAX} is the perturbed diffusivity coefficient. The perturbations can be seen as a representation of subgrid-scale variability in the time scale of vertical adjustment to unstable stratification. Note that A_{vT}^{refMAX} is proportional to $\Delta z^2 / \Lambda_{vT}$ with Δz as the vertical scale of the instability and Λ_{vT} as the time scale of vertical mixing. Therefore, perturbing A_{vT}^{refMAX} by a bounded, positive random number $1 + \xi_{EVD}$ that is symmetric and has a Gaussian-like distribution function is equivalent to perturbing the time scale Λ_{vT} using a skewed random number with a long tail for large values (similar to the explanation in section 3b). The motivation for doing so is based on 1) the assumption that the convective events should be skewed toward long time scales and 2) observations discussed by, among others, Lavender et al. (2002). Lavender et al. (2002) analyzed deep convection in the Labrador Sea and looked at the vertical water parcel displacement over time at 400-m depth (see their Fig. 7). These displacements show a wide variety of time scales for a similar displacement length scale (e.g., 1000 m), with some very fast events, a majority of medium time scales, and a few very long time scales. Perturbing Λ_{vT} by a skewed random number will simulate the skewed variations of the deep convection time scales that are not resolved by a constant A_{vT}^{refMAX} . However, it will still result in a symmetric perturbation to A_{vT}^{refMAX} itself. Furthermore, by choosing the amplitude of the perturbations as discussed in the next section and in appendix C, we specifically target convective events that have a vertical extent of 250 m and beyond (i.e., deep convection of a specific amplitude). For such events the

TABLE 1. The names and configurations of the different sensitivity test cases for the different schemes; see text for details on the schemes. Yes means the stochastic perturbations are applied; No means they are not applied. Integration length for all simulations was 15 yr, starting from 1 Jan 1990.

Name	A_{GM}^{sto}	P_{tend}^{sto}	D_{tend}^{sto}	A_{VT}^{stoMAX}	τ (days)	Δr (km)	a
STO_TKE_D	No	No	Yes	No	60	3000	0.8
STO_TKE_P	No	Yes	No	No	60	3000	0.8
STO_TKE	No	Yes	Yes	No	60	3000	0.8
STO_EVD	No	No	No	Yes	60	3000	0.8
STO_GM	Yes	No	No	No	60	3000	0.8
STO	Yes	Yes	Yes	Yes	60	3000	0.8

perturbations lead to actual differences in the vertical mixing profile between stochastic and deterministic simulations when a model time step of 1 h is used. Events with a smaller vertical extent will mix within one time step regardless of the random perturbation (see Fig. C1).

In the following section we will discuss the impacts of the three schemes introduced here, in a generalized way for short sensitivity studies and in a more comprehensive discussion of long multidecadal simulations.

4. Results

a. Sensitivity studies

For this section a set of short sensitivity simulations has been carried out to investigate the impact of each of the perturbation schemes as well as the impact of the parameter choices. Readers solely interested in the final setup of the perturbation schemes and its impact on low-frequency variability can skip this part and move straight to section 4b.

There are a number of parameters that need to be set for the stochastic perturbation schemes. Those are the bounds a of the bounded distribution (i.e., the amplitude of the perturbations), the temporal decorrelation τ , and the spatial decorrelation Δr , as well as the shape of the distribution function controlled by β in Eq. (5). The latter value is fixed to $1.4/\sigma_{\max}$ where σ_{\max} is calculated by Eq. (3).

In the first set of sensitivity simulations each scheme was applied separately, with the other schemes switched off, and simulations were carried out for 15 yr after the spinup. We also carried out an integration where both perturbations to the TKE scheme were used, and one where all perturbation schemes were applied (see Table 1). All schemes used the same set of parameter values for τ , Δr , and a (Table 1) to reduce the parameter space. These integrations were then analyzed with regards to the change in annual mean SSH standard deviation in the stochastic simulations compared to the deterministic reference simulation (REF), shown in Fig. 1. This provides a computationally cheap first estimate of where

the schemes act and the potential amplitude of the induced changes, specifically focusing on interannual variability (IV). Following these sensitivity simulations, the multidecadal simulations discussed in section 4b will provide more robust diagnostics.

The impact of the perturbations to EVD (section 3d) in Fig. 1d is rather small (on the order of a few percent change in IV) and localized mostly in the North Atlantic region. The reason for this is that even though EVD perturbations have the potential to alter overturning through vertical mixing time scales, bounding them to convective events of an extent of 250 m and beyond (i.e., vertical diffusion within $[10, 90] \text{ m}^2 \text{ s}^{-1}$) does not have a very strong effect in the 1° model. Only in cases of very strong deep convection the actual stochastic choice of the vertical mixing coefficient has an impact on the depth of the mixed layer (not shown). But since these events are usually very temporally and spatially localized, the system is pulled back to its mean state within a few days after the event. In the current uncoupled setup the perturbations also do not seem to have a big impact on the preconditioning of deep convective events, potentially influencing the stratification of the ocean and as a consequence the strength, location, and duration of deep convection. Supported by observational estimates of, for example, Lavender et al. (2002), it is possible to also focus on uncertainty in shallower convection by increasing the amplitude of the perturbations or by reducing the default value for EVD. This would most likely lead to a more pronounced response in the 1° simulations.

However, in a coupled system the current setup of the stochastic perturbations might also show an enhanced impact, since the perturbations might affect the preconditioning of convective events more strongly and cause changes in the time scales of deep convection. Both these aspects influence the uptake of heat and carbon at the interface between the atmosphere and ocean and are crucially dependent on the coupling with the atmosphere.

Figures 1a–c show the impact of the perturbations to the TKE mixing scheme. As for the EVD perturbations,

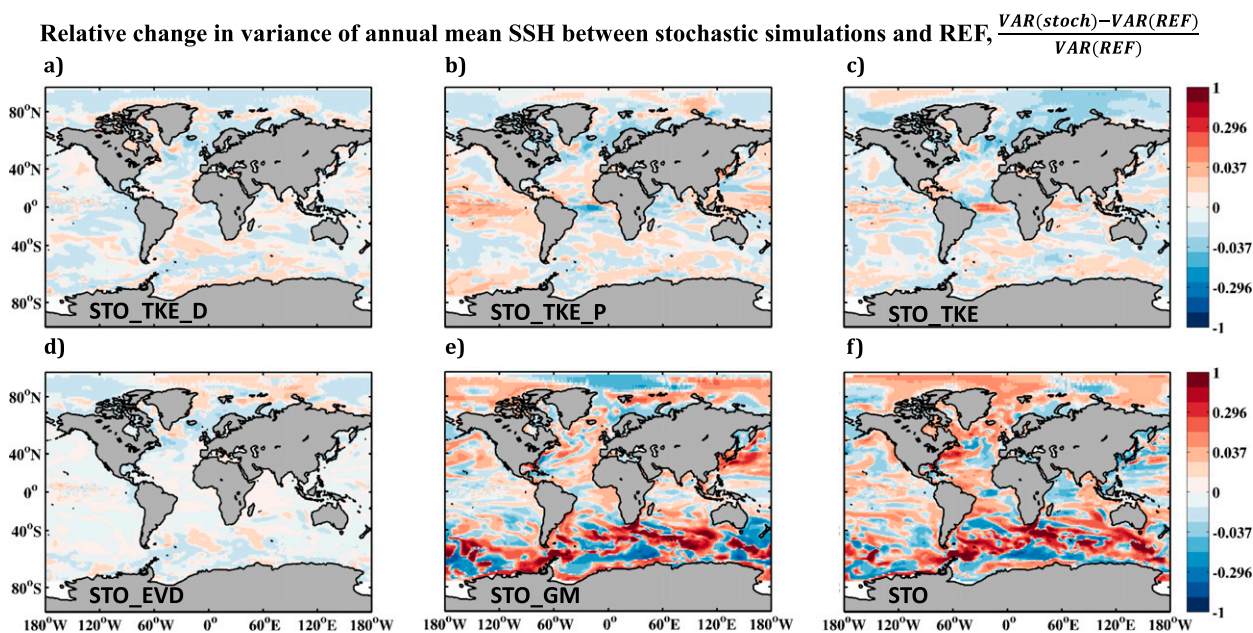


FIG. 1. Relative change in variance of linearly detrended annual mean SSH between the different stochastic schemes and REF (normalized by the variance of REF). (a) Stochastic perturbations to the kinetic energy destruction tendency due to stratification. (b) Stochastic perturbations to the kinetic energy production tendency due to shear. (c) Stochastic perturbations to both kinetic energy destruction and production tendencies. (d) Stochastic perturbations to the EVD. (e) Stochastic perturbations to the Gent–McWilliams coefficient. (f) All stochastic perturbations mentioned before. Note the nonlinear color bars.

the perturbations for TKE were bounded by $a = 0.8$ based on spatial subgrid variations of the TKE tendencies estimated by Goodman et al. (2006). The perturbations mostly act in regions of strong surface mixing induced by buoyancy or shear forcing. This is the case in the tropics, the North Atlantic, and some parts of the Southern Ocean. The effect is much stronger for the perturbations to the shear tendency (about 5%–10% change in IV; Fig. 1b) then it is for the buoyancy perturbations (a few percent change in IV; Fig. 1a), especially in the tropics. Switching both perturbations on (Fig. 1c) slightly reduces the impact in the tropical Pacific when compared to the case where perturbations are only applied to the shear tendency but increases the impact in most other regions. It is likely that the induced variabilities due to each scheme may counteract each other in some locations, while amplifying each other in other regions. The counteracting effect could also be a sampling issue, since sensitivity studies were carried out for only 15 yr.

Finally, the perturbations to the GM parameterization in Fig. 1e show by far the largest impact (about 50%–70% change in IV), but predominantly in the Southern Ocean, and, to a lesser degree, in the western boundary currents (about 30% change in IV). These are the regions where the GM parameterization is active. However, although the other schemes by themselves have a rather small impact, especially in the Southern

Ocean, combined with the perturbations to the GM parameterization (STO; Fig. 1f) the effects seem to be strongly enhanced (more than 100% change in IV). Therefore we concentrate our analysis with all parameterizations turned on simultaneously to investigate their combined impact on low-frequency variability.

In a separate set of sensitivity studies we performed simulations with all schemes used at the same time but with different parameter choices for τ , Δr , and a . A subset of these simulations is presented in Fig. 2, with the values for τ and Δr shown and $a = 0.8$. The simulations were carried out for 5 yr after the spinup to allow us to investigate a larger parameter space without using too large amounts of computing time. Figure 2 shows the annual mean difference or anomaly of the last year of integration for SSH between the stochastic simulations and the deterministic reference simulation REF. These diagnostics provide only an indication of the efficiency with which the stochastic schemes generate low-frequency variability. They do not provide any information on the robustness of these anomalies. However, it allows for a rather straightforward comparison of the effect of the different parameter choices in terms of spatial distribution and strength of anomalies. A similar strategy was used by Juricke et al. (2013) to investigate the impact of spatial and temporal correlations of perturbations to the sea ice strength parameterization in a sea ice model.

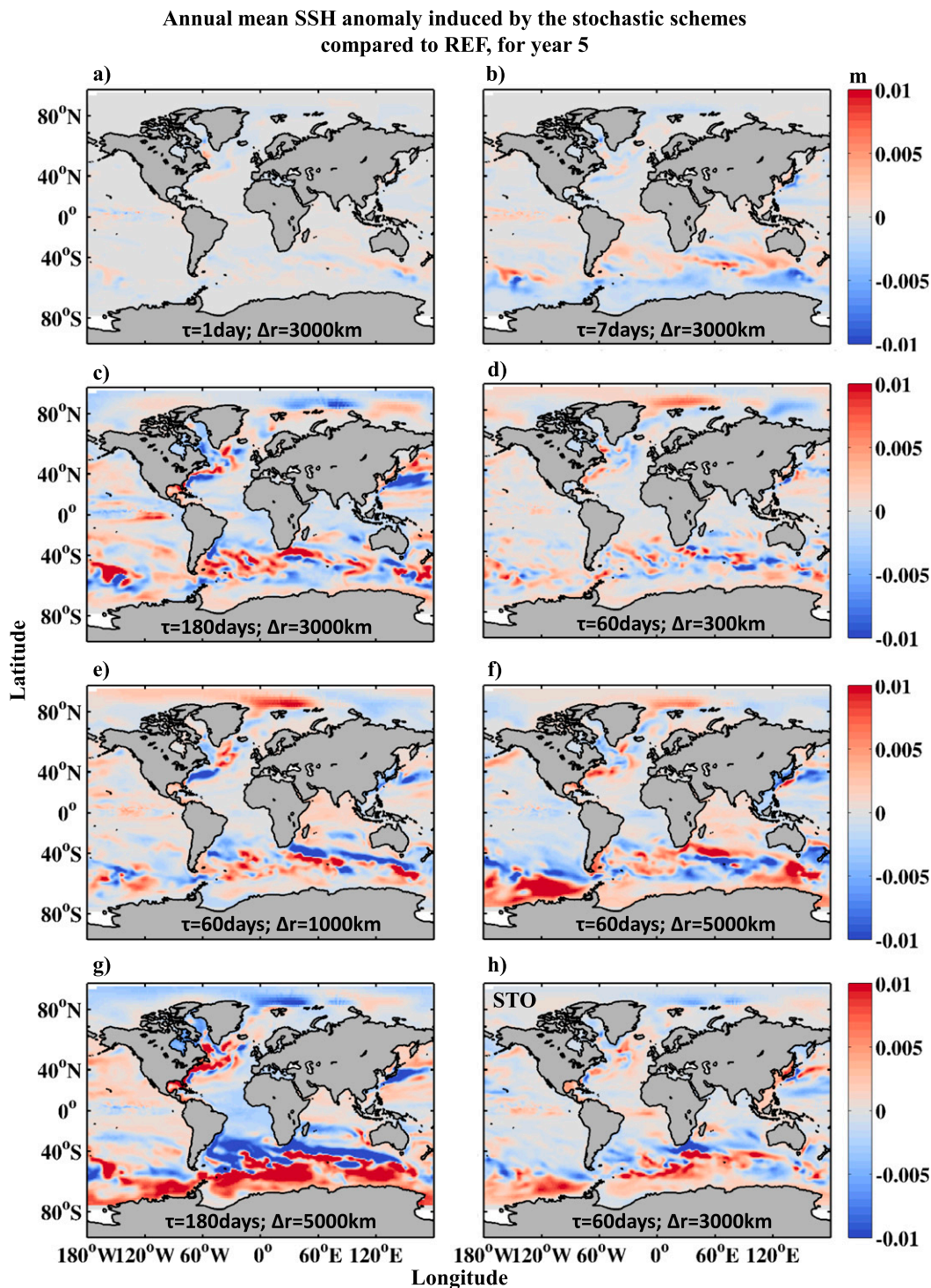


FIG. 2. Annual mean difference in SSH (m) between REF and the stochastic sensitivity simulations for the fifth year of each integration. Parameter values for the temporal and spatial correlations τ and Δr are shown. (h) Differences between REF and the default setup for STO.

The impact of the stochastic reference simulation (STO) used for the multidecadal simulations in section 4b is shown in Fig. 2h. Reducing the temporal correlation strongly, as is done in Figs. 2a,b, reduces the amplitude of the stochastic schemes. However, comparing Figs. 2c and 2h shows that increasing temporal correlations beyond a certain threshold does not change the effect of the perturbations much. The same is true for the spatial correlation when comparing Figs. 2d–f. The size of the perturbations has a direct impact on the size of the anomalies created. However, the amplitude of the perturbations is not affected much. Comparing Figs. 2e and 2h shows that the spatial correlation might not even matter much at all if chosen within a specific range. Finally Fig. 2g shows that the impact of the perturbations can be further enhanced if both spatial and temporal correlations are increased. We have also run additional simulations where we investigated the impact of the perturbation amplitude a (not shown). Values tested were in the range of [0.6, 1]. The amplitude has a direct impact on the strength of the generated anomalies.

The final values chosen for the reference stochastic parameterization of the next section are to some extent an ad hoc decision based on a reasonable impact of the scheme, but also on the physically plausible choice of each parameter as discussed in the previous section and in appendixes B and C. All schemes were given the same parameters. The parameter choices are similar to what was used by Andrejczuk et al. (2016). Length scales of the schemes are quite large, but it should be kept in mind that the effective resolution of the model is much coarser than 1° . The perturbation statistics are not supposed to reflect eddy time and length scales, since eddies are not resolved outside the tropics. Instead, the perturbation statistics are connected with unresolved large-scale anomalies in the water mass properties and the forcing. An analysis of the Estimating the Circulation and Climate of the Ocean, phase 2 (ECCO2), ocean state estimate data (Menemenlis et al. 2008) has shown that length and time scales of the leading EOFs of the North Atlantic are of the same order of magnitude as the values chosen for the perturbations in this study. However, we acknowledge that in the current setup the parameter choices are also made as part of a tuning decision and can be optimized in the future, making use of better observational estimates and high-resolution model studies. We will briefly discuss steps needed to better constrain these parameters in the last section.

b. Multidecadal integrations

As mentioned in the last subsection, after the sensitivity studies had been carried out a certain set of parameters for the stochastic schemes was decided upon to

run a multidecadal simulation. These parameter values are $\tau = 60$ days, $\Delta r = 3000$ km, and $a = 0.8$ for an experiment henceforth named STO. All perturbation schemes were used simultaneously, with the same parameters but different sequences of random numbers. The simulation was carried out for 105 yr (15 + 45 + 45 yr), forced by 1990–2004 atmospheric forcing, and two repeated cycles of 1959–2004 atmospheric forcing, where the year 1959 was neglected to reduce the initial shock created by the change in the atmospheric forcing. A deterministic reference simulation REF of the same length was also carried out.

1) GENERAL COMPARISON BETWEEN REF AND STO

In this section we will focus on the differences in interannual variability between the deterministic reference simulation REF and the stochastic simulation STO. Figures 3a,c,e show the zonally averaged streamfunction for the eddy (i.e., GM) Ψ_{GM} , the Eulerian Ψ_{Euler} , and the residual (i.e., Eulerian + eddy) Ψ_{res} component of the flow, for the reference simulation REF and focused on the Southern Ocean part of the Atlantic basin. Since the impact of the perturbations to the GM parameterization proved to be the strongest in the sensitivity studies, we will mostly focus on the changes in the Southern Ocean.

There are two maxima in interannual streamfunction variability in the Atlantic, one between 60° and 40° S as shown in Fig. 3, at the location of the polar fronts, and another one between 30° and 60° N in the North Atlantic, signifying the location of maximum overturning (not shown). While the Eulerian and residual streamfunctions have the strongest variability in the Southern Ocean near the surface, mostly due to the interannually varying wind forcing, the eddy streamfunction varies mostly in the subsurface as a response to the tilting of the isopycnals caused by the Eulerian flow field and due to the fact that GM is tapered near the boundaries. The additional strong variability of Ψ_{GM} at the bottom of the basin is most likely a topographic effect and its realism is difficult to quantify.

Figures 3b,d,f show the relative changes caused by the stochastic perturbations in the IV of the respective streamfunctions. IV for Ψ_{GM} is increased by a factor of 10 and more in some regions, especially in regions where the variability was the largest in REF. Therefore IV is hugely increased by more than an order of magnitude. We remind the reader that the time scale of the perturbations applied is defined by τ . Therefore the perturbations are applied on a seasonal time scale with a decorrelation of 60 days and the effect seen in Ψ_{GM} on interannual time scales is the consequence of the slow

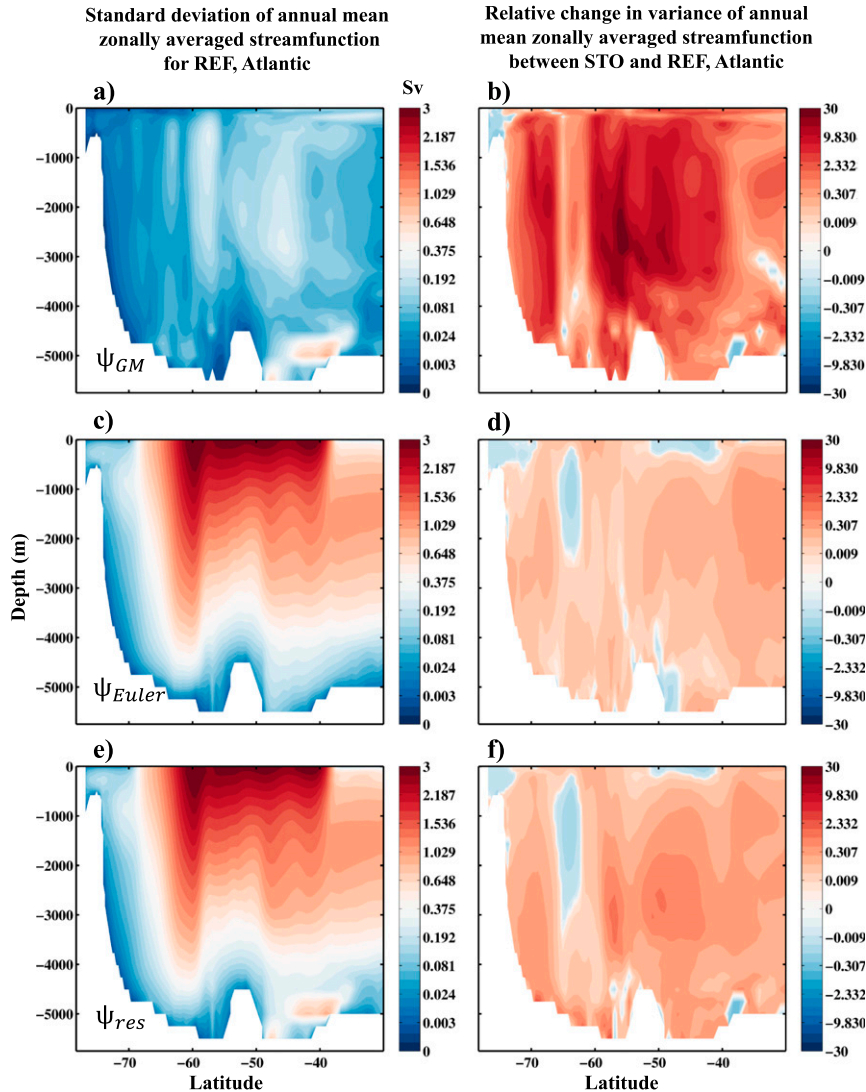


FIG. 3. Standard deviation of linearly detrended annual mean zonally averaged streamfunction (Sv ; $1 \text{ Sv} \equiv 10^6 \text{ m}^3 \text{ s}^{-1}$) and relative changes in variance due to the stochastic perturbations in the Atlantic basin. (a),(c),(e) Standard deviation of the reference simulation REF, and (b),(d),(f) the relative change in variance between STO and REF (normalized by the variance of REF), for the (top) eddy-induced streamfunction Ψ_{GM} , (middle) Eulerian streamfunction Ψ_{Euler} , and (bottom) residual streamfunction Ψ_{res} . Data are calculated from the 5-daily averaged values of the velocity fields. Values above around 0.2 in (b),(d),(f) are significant at the 0.18 level, and values above around 0.3 at the 0.09 level according to the one-tailed F test. Note the nonlinear color bars.

dynamic response of the system to the comparatively high-frequency perturbations.

The strong increase in parameterized eddy variability impacts on IV of both Ψ_{Euler} and Ψ_{res} . While the relative impact for Ψ_{Euler} is strongest just to the north of the previous maximum in IV of REF (between 30° and 40°S), the relative changes for Ψ_{res} are dominated by the direct impact of the increased IV of Ψ_{GM} (between 60° and 40°S). The increase in IV for Ψ_{GM} is around 20%–30%,

which corresponds to a relatively low significance level of 0.1–0.2 given the one-tailed F test. The northern part of the changes, around 30° – 40°S , is strongest mostly because the effect of the normalization by the reference IV is smaller here. In general the IV for Ψ_{Euler} is increased throughout most parts of the Southern Ocean, also near the surface where the reference IV is strongest at around 60°S . This is a consistent response to the increased IV in Ψ_{GM} . For Ψ_{res} the relative IV increase is about 60% and

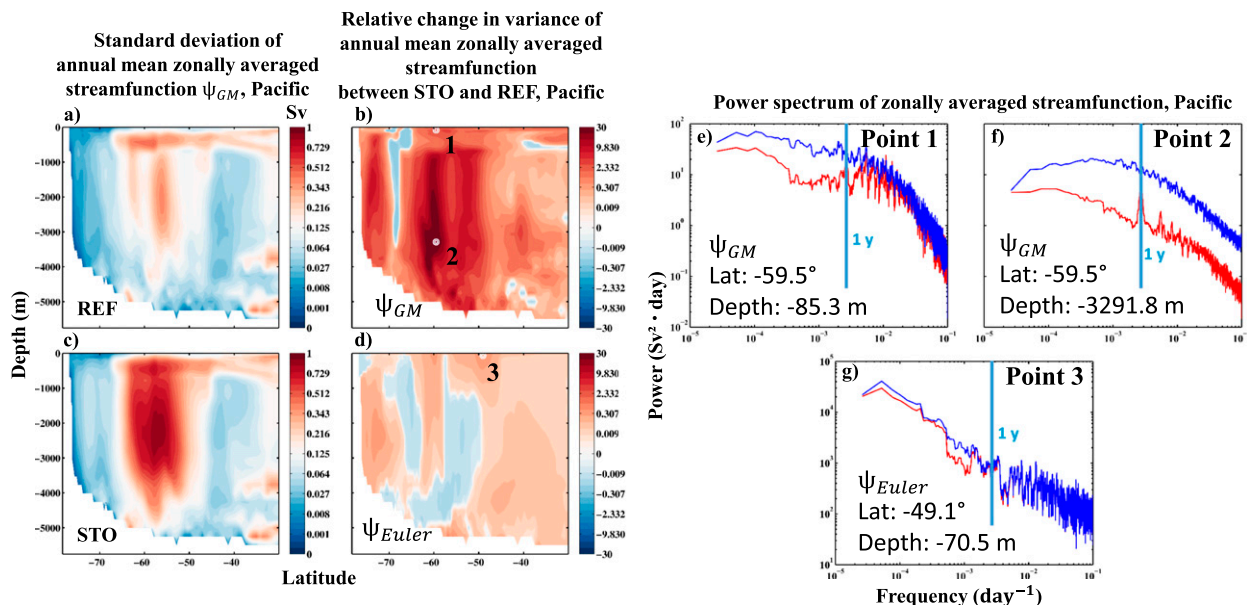


FIG. 4. Changes in standard deviation of linearly detrended annual mean zonally averaged streamfunction (Sv) for the Pacific basin and changes in the power spectrum. Interannual variability of Ψ_{GM} in (a) REF and (c) STO, and the relative change in variance between STO and REF (normalized by the variance of REF) for (b) Ψ_{GM} and (d) Ψ_{Euler} , calculated from the 5-daily averaged values of the velocity field. Also shown are power spectra ($Sv^2 \text{ day}$) at three different locations in the Pacific basin for (red) REF and (blue) STO and for (e), (f) Ψ_{GM} and (g) Ψ_{Euler} . White dots in (b) and (d) mark the locations to which the spectra correspond. Spectra are plotted in a log-log scale, calculated from the 5-daily averaged values of the velocity fields. The climatological seasonal cycle is removed. The spectra are smoothed by a running mean with window size 13 (65 days). Note the nonlinear color bars for (a)–(d).

more, which is highly significant using the one-tailed F test.

Not shown here is the impact of the perturbations to EVD and TKE on the near-surface variability (upper one hundred meters) in the tropics and in the regions of deep convection in the North Atlantic. Variability is only slightly increased by a few percent in the deep North Atlantic, but more strongly by around 20% near the surface at the equator. There are also (spurious) increases of variability along the bottom topography throughout the Atlantic basin, in areas where there is already comparatively high variability in the deterministic setup—a potentially spurious effect of parameterization schemes such as GM (Marshall et al. 2012).

Figures 4a–d show similar results for Ψ_{GM} and Ψ_{Euler} in the Pacific. The major difference here is that there is a much stronger relative signal for Ψ_{Euler} near the surface. To analyze whether these changes are only on the interannual time scale we have chosen two points in the Pacific for Ψ_{GM} and another for Ψ_{Euler} and Ψ_{res} , marked in Figs. 4b,d, to look at the power spectrum, shown in Figs. 4e–g. Figures 4e,f show that Ψ_{GM} has higher energy for STO than REF throughout the entire frequency range, on weekly up to decadal time scales, by about an order of magnitude. This is more pronounced for the point in the deep ocean than the one near the surface

where the relative impact of the perturbations is weaker, but still significant in the low-frequency range (i.e., starting with the interannual time scale). We again note that the decorrelation time scale of the perturbations is only 60 days. The strong increase in IV and low-frequency variability is caused by the dynamic response of the system to the increased variability on the seasonal time scale.

For Ψ_{Euler} and Ψ_{res} the relative impact on IV in the Pacific is specifically strong in the upper ocean near 50°S. The power spectrum at a location there shows that low-frequency variability is enhanced for Ψ_{Euler} (Fig. 4f), with a basically indistinguishable spectrum for Ψ_{res} . The impact is not as large as it is for Ψ_{GM} , but it consistently shows an increase in variability on the interannual time scale and beyond, reflecting the dynamic response to the increased high-frequency variability of the eddy-induced advection.

Similar responses to the stochastic perturbations as the one described so far for the Atlantic and the Pacific can be observed in the Indian Ocean as well, and also appear in the globally averaged streamfunction, even though the amplitudes for Ψ_{Euler} and Ψ_{res} are somewhat reduced by the global averaging.

We also analyzed the impact of the stochastic perturbation schemes on zonally averaged temperature and salinity (not shown). The relative changes in IV are of

Relative change in variance of annual mean temperature between STO and REF, for the Kuroshio region and different depths

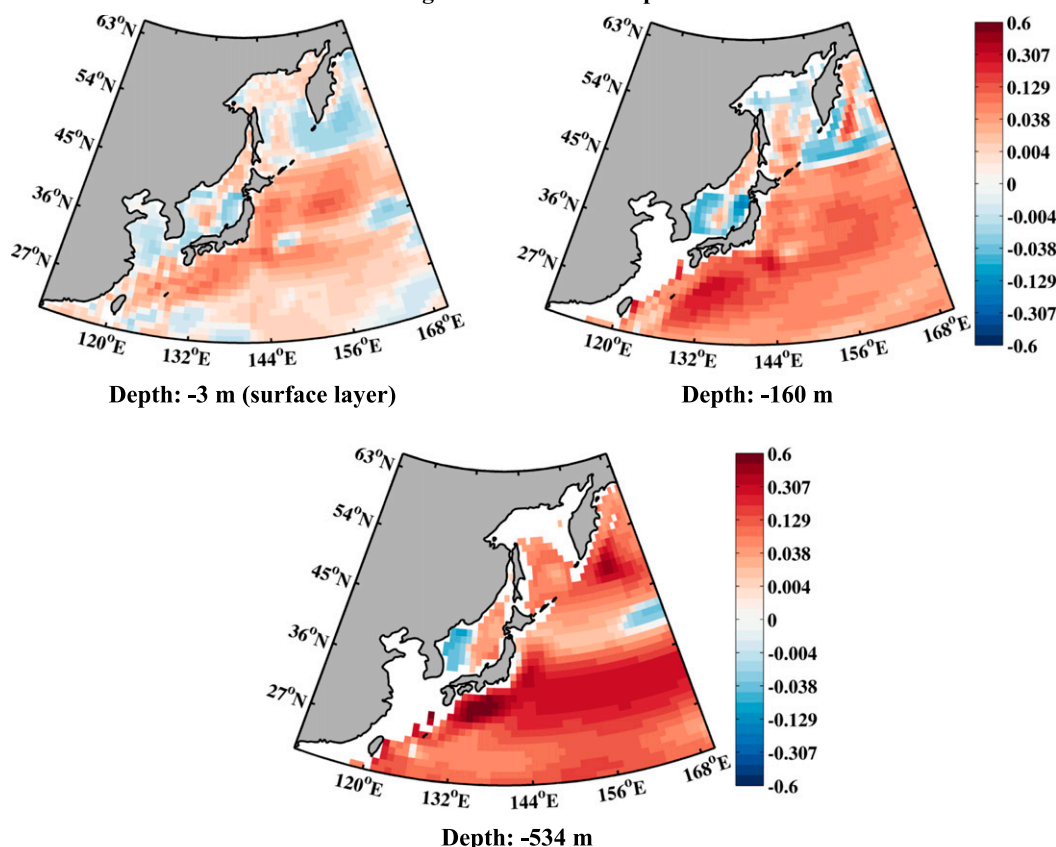


FIG. 5. Relative change in variance of linearly detrended annual mean temperature between STO and REF (normalized by the variance of REF), at different depths: (a) -3 , (b) -160 , and (c) -534 m. Interpolated to a $1^\circ \times 1^\circ$ grid. Note the nonlinear color bars.

similar amplitude than is the case for Ψ_{Euler} and occur over large areas, but they are less coherent. We see an increase in IV in the deep Southern Ocean of most basins for both temperature and salinity up to about 30%. Some additional increase in IV can be observed in the upper ocean along the equator, as well as in a subsurface region at the Kuroshio and the Gulf Stream, especially strong for the former. Figure 5 shows the relative difference in ocean temperature IV between STO and REF at different depth levels for the Kuroshio region. While the relative difference is around 20% near the surface, it increases to more than 60% at a depth of 534 m. Similar results can be observed in the Southern Ocean.

For the zonally averaged temperature and salinity some slightly increased variability can also be observed in the North Atlantic, as well as some more spurious variability increase near topography (not shown). In some areas, especially close to regions where variability is increased, IV is reduced for both salinity and

temperature. This reduction is not as strong as the increase, but on the order of 10%. It could potentially be due to shifting variability patterns, or suppressed variability induced by an increased variability in a bordering region.

2) COMPARISON TO REANALYSIS AND OBSERVATIONAL DATA

After analyzing the effects of the stochastic perturbations on the low-frequency variability in the streamfunctions, it is of interest to see whether these effects improve the model when compared to available observations and reanalysis products.

To this end we compare the simulated Pacific's Ψ_{Euler} and global SSH IV with two different reanalysis products, the Ocean Reanalysis Pilot 5 (ORAP5; see Zuo et al. 2017) and Ocean Reanalysis System 4 (ORAS4; see Mogenssen et al. 2012; Balmaseda et al. 2013), in Figs. 6 and 7. Both reanalysis products are issued by the European Centre for Medium-Range Weather

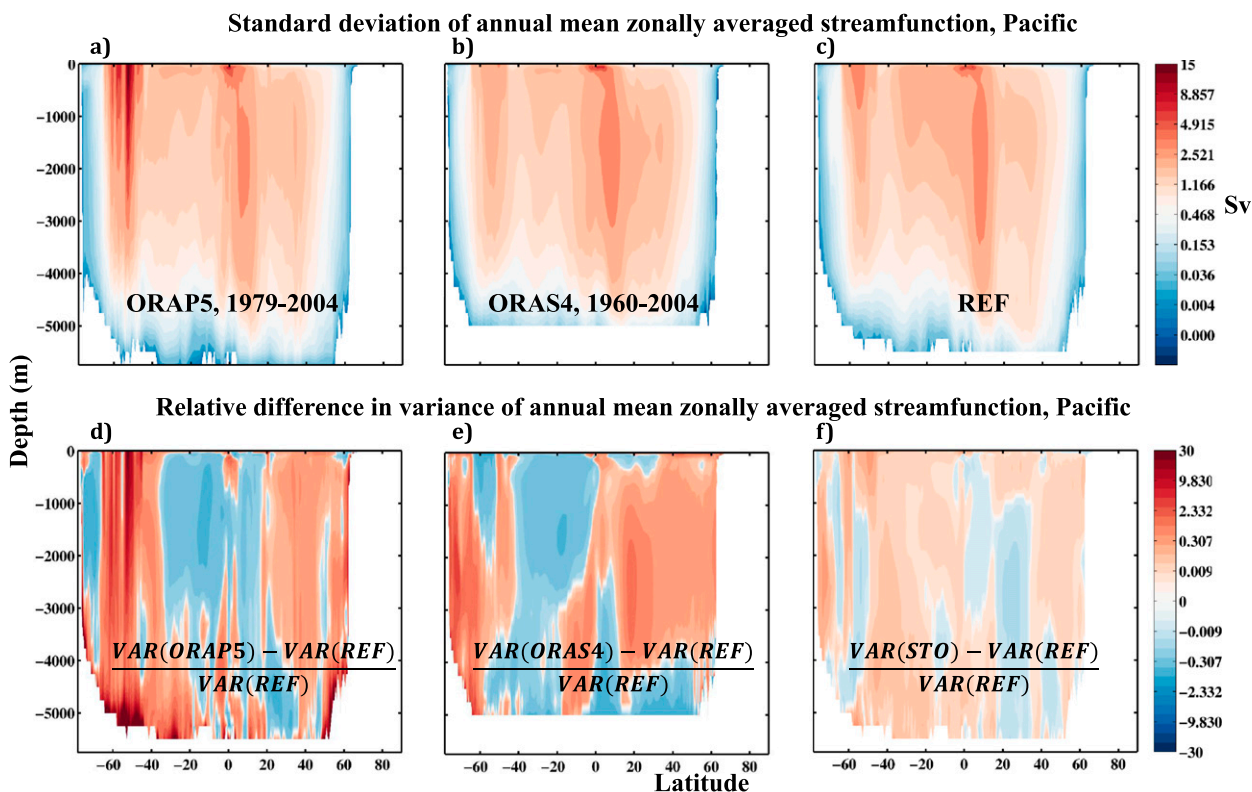


FIG. 6. Standard deviation of linearly detrended annual mean zonally averaged streamfunction (Sv) for Ψ_{Euler} in the Pacific basin for (a) ORAP5 (1979–2004), (b) ORAS4 (1960–2004), and (c) REF (105 yr). Relative change in variance of annual mean zonally averaged streamfunction between REF and (d) ORAP5, (e) ORAS4, and (f) STO (normalized by the variance of REF). Data are calculated from monthly mean values of the velocity fields and interpolated to the grid of REF. Note the nonlinear color bars.

Forecasts (ECMWF) and assimilate SST and SSH [see Zuo et al. (2017) for more details on the differences between the two products]. A major difference between the two products is the resolution of the ocean model, which is also NEMO. For ORAP5 the ocean grid used is ORCA025, with a nominal resolution of about $1/4^\circ$ and 73 vertical layers, while ORAS4 uses the same ORCA1 grid as the simulations carried out in this study, but with only 42 vertical layers. ORAS4 and REF compare generally quite well in terms of IV of Ψ_{Euler} (Fig. 6e). In ORAP5, however, the Ψ_{Euler} and SSH IV in the Southern Ocean is for most parts much stronger than in the other two simulations (cf. Figs. 6d and 6e and Figs. 7d and 7e). REF underestimates the IV of ORAP5 in some regions of the Southern Ocean by more than an order of magnitude (Figs. 6d and 7d). Similarly, ORAS4 has a strongly reduced IV in the same regions when compared with ORAP5. A recent reanalysis intercomparison study by Karspeck et al. (2017) for the North Atlantic has shown that the representation of IV between different reanalysis products varies strongly. One reason for this is that the subsurface state of the ocean is badly constrained in ocean reanalysis because

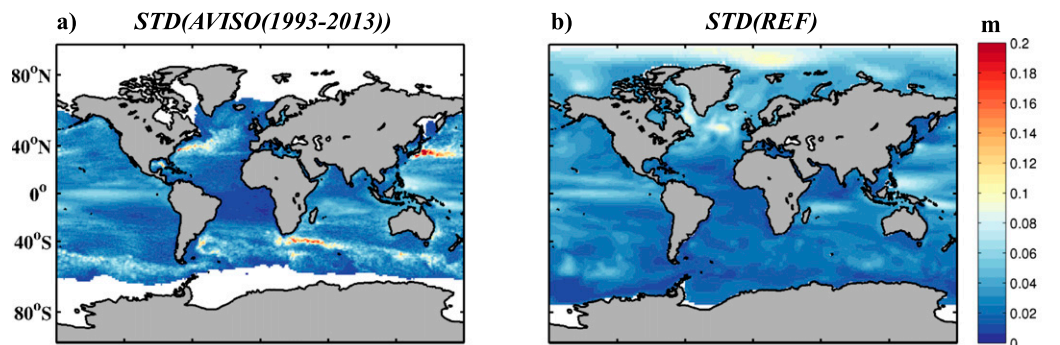
of missing observational data. Nevertheless, we generally assume that an increased resolution reanalysis product should have an increased ability to represent even the subsurface state of the ocean.

Comparing STO with REF in Fig. 6f—which is similar to Fig. 4d—and Fig. 7f reveals an increase in IV for STO where REF is missing IV when compared with ORAP5. Therefore STO increases IV in the right locations if we assume that ORAP5 has a better representation of the overturning and SSH in the Southern Ocean compared to ORAS4. However, the increase in IV in STO is only about 20%–30%, while REF is missing IV on the order of more than 1000% when compared with ORAP5.

To compare the changes caused by the stochastic perturbations to better constrained observational estimates, Fig. 7a shows the IV in SSH for the AVISO SSH altimetry observational dataset.¹ SSH includes information about integrated vertical salinity and

¹ See Ssalto/Duacs altimeter products at <http://www.aviso.altimetry.fr>.

Standard deviation of annual mean SSH



Relative difference in variance of annual mean SSH

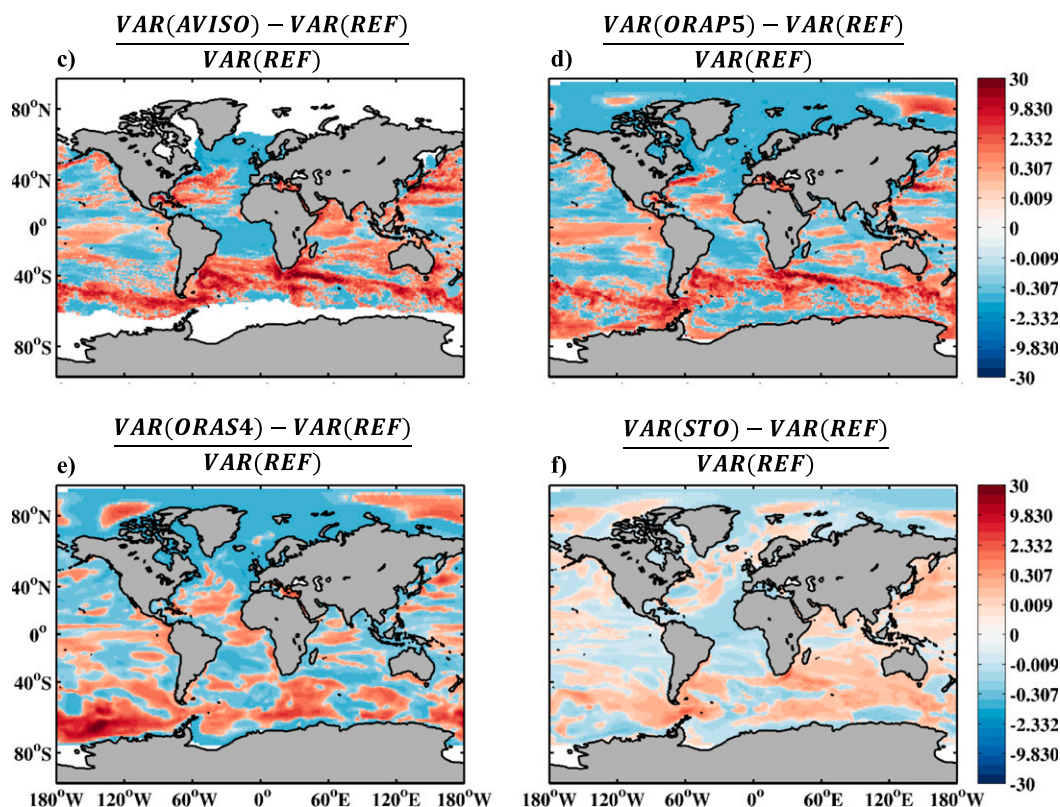


FIG. 7. Standard deviation of linearly detrended annual mean SSH (m) for (a) the AVISO observations (1993–2013) and (b) REF (105 yr), and relative change in variance of annual mean SSH between REF and (c) AVISO, (d) ORAP5, (e) ORAS4, and (f) STO (normalized by the variance of REF). Interpolated to a $1^\circ \times 1^\circ$ grid. Note the nonlinear color bars for (c)–(f).

temperature profiles through its steric component as well as circulation properties but can still be measured at the surface. It therefore provides more integrated information about the ocean state than, for example, SST.

In addition to the Southern Ocean, SSH IV in REF is strongly underestimated by an order of magnitude in other eddy-active regions such as the Gulf Stream and

Kuroshio (see Figs. 7a,b,e), which is at least partly due to the fact that REF at a 1° resolution does not resolve eddies in those latitudes. Note that ORAP5 is much better in representing IV than ORAS4 when compared to AVISO, with ORAS4 generally showing much less IV. The improved representation of IV of ORAP5 compared to ORAS4 is probably at least in parts due to the increased resolution in ORAP5. This might be seen

as a reason to trust the ORAP5 representation of $\Psi_{\text{Euler}}^{\text{IV}}$ in Fig. 6 more than the IV produced by ORAS4. However, it should be noted that even ORAP5 is still underestimating SSH IV in most regions even though the underestimation is considerably reduced especially in regions of strong eddy activity.

Similar to the improvement in the Southern Ocean, SSH IV in the Kuroshio region is better represented in STO, with a comparable increase in variability of about 20%–30%. But again this increase is much less than the actually missing IV compared to AVISO.

Extending the low-frequency analysis to variability of 5- and 10-yr averages of SSH showed that variability on those time scales is increased by the stochastic schemes in very similar regions, but with a reduced amplitude (not shown). This is consistent with the spectral analysis in Fig. 4.

Finally we investigated the impact of the stochastic schemes on IV in heat content (300 and 700 m) and SST (not shown). In both cases variability is increased in similar regions as for SSH and similar relative amplitude, somewhat less for SST. Since SST is strongly coupled to the atmospheric forcing, SST variability is largely driven by the atmospheric coupling, a conclusion also drawn by Andrejczuk et al. (2016). However, comparing SST IV between REF and the National Oceanic and Atmospheric Administration (NOAA) high-resolution SST data (Reynolds et al. 2007) showed that the underestimation of IV in REF is not nearly as strong as it is for SSH (not shown). While there is still an underestimation of variability in the Southern Ocean and parts of the western boundary currents, REF actually simulates far too much variability in the North Atlantic. STO generally improves SST IV in the regions mentioned previously when discussing SSH IV. In a smaller fraction of areas the increase of SST IV in STO, however, actually leads to a small overestimation of IV. But this might also be due to the fact that the climatological mean SST field shows biases between REF and the NOAA data, which might lead to an increase in variability in STO in the wrong regions. Additionally, as mentioned above, the atmospheric forcing largely controls SST variability. While the surface variability is strongly controlled by the atmospheric forcing, internal dynamics become more important when looking at depth-integrated quantities or variables at layers farther removed from the surface (see Fig. 5).

5. Discussion and conclusions

In this study we have introduced stochastic perturbation schemes to three commonly used subgrid-scale

mixing parameterizations in the global 1° ocean model, NEMO, and investigated their impact on low-frequency (i.e., interannual and longer time scale) ocean variability. We showed that the stochastic schemes lead to a considerable improvement of interannual variability in the model.

The perturbations are applied to the turbulent kinetic energy scheme used to calculate vertical mixing coefficients, the enhanced vertical diffusion scheme used to calculate the vertical diffusion coefficient in cases of unstable stratification, and the Gent–McWilliams parameterization (GM) used to calculate an eddy advection term in the tracer equations. The stochastic perturbations were symmetric and correlated in time—with a subseasonal decorrelation time scale—and in space. They were applied to investigate the impact of unresolved (missing) subgrid-scale variability and mixing time scale variations on low-frequency variability in NEMO. Stochastic symmetric perturbations to the vertical diffusion and GM coefficients correspond to skewed perturbations to the related vertical mixing and eddy time scales, respectively.

Applying the comparatively high-frequency perturbations especially to the GM scheme increased zonally averaged eddy streamfunction variability in the Southern Ocean considerably by about a factor of 10, throughout all temporal frequencies. The perturbations also increased low-frequency variability for the Eulerian and residual streamfunction in the Southern Ocean, by up to 30% and 70% on the interannual time scale, respectively. Moreover, interannual sea surface height variability was increased in the Southern Ocean and the Kuroshio region by 20%–25%, improving the representation of low-frequency variability in those regions when compared to a high-resolution, 1/4° reanalysis product, ORAP5, and the AVISO altimetry data. Even though this increase is not sufficient to compensate for the missing interannual variability in these regions, it shows that stochastic, high-frequency perturbations applied in a physically consistent way can help to represent and improve missing low-frequency variability modes in the ocean. This is especially noteworthy since the 1° ocean model does not resolve crucial eddy processes. In this context it is worth referring to the results of Kitsios et al. (2013), who developed both deterministic and stochastic oceanic subgrid parameterizations. They investigated the impact of the schemes on the kinetic energy spectra in a simplified model of the Antarctic Circumpolar Current [see Fig. 12 in Kitsios et al. (2013)]. Kitsios et al. (2013) showed that when baroclinic instability is not

explicitly resolved, a stochastic backscatter scheme allowing negative viscosity will improve the results. However, their results suggest that an additional deterministic negative viscosity term is required to reproduce the missing unresolved variability. But such a term can lead to numerical model instabilities unless the viscosity coefficient is constrained by the available energy (Jansen and Held 2014). Furthermore, there is certainly a need to include other unresolved or strongly dampened eddy effects in the 1° model when it comes to ocean time scales (such as generation of and interaction with baroclinic Rossby waves). GM is designed to only capture the mean effect of baroclinic instability on the large scale. Consequently, perturbations to GM cannot entirely compensate for the generally missing unresolved eddy variability and the extensive damping at 1° when compared with observations and the $1/4^\circ$ reanalysis. However, in our scheme, the perturbed eddy induced velocity field can excite buoyancy anomalies, especially in regions characterized with steep isopycnal slopes. These anomalies can propagate as baroclinic Rossby waves and further excite anomalies propagating with a range of time scales, including interannual time scales.

Nevertheless, the approach followed in this study aims at applying stochastic perturbations to enhance variability and provide uncertainty estimates in existing, already implemented and widely used schemes, which increases the applicability. Constraints to the empirical perturbations are provided by previous observational and modeling studies and could be recalibrated in light of new theoretical insights (e.g., Jansen 2017) and observations.

We speculate here that a higher-resolution model with an eddy-permitting grid of $1/4^\circ$ will profit even more from an inclusion of unresolved subgrid-scale variability. This is because viscosity and diffusion are much reduced in higher-resolution simulations and eddy effects are partially resolved. This in turn allows for an increased nonlinear behavior of the resulting flow field, which interacts in a more complex way with symmetric stochastic perturbations. While we saw only minor, if mostly robust changes to the climatic mean ocean state caused by the stochastic perturbation schemes in the 1° model, the increased nonlinearities in a higher-resolution model might lead to a stronger nonlinear rectification effect which has been observed in atmospheric models (e.g., Weisheimer et al. 2014). Figure 8 shows the mean change in SSH between the stochastic simulation and the deterministic reference, normalized by the interannual standard deviation of the reference. While the absolute value of the mean change is comparatively small

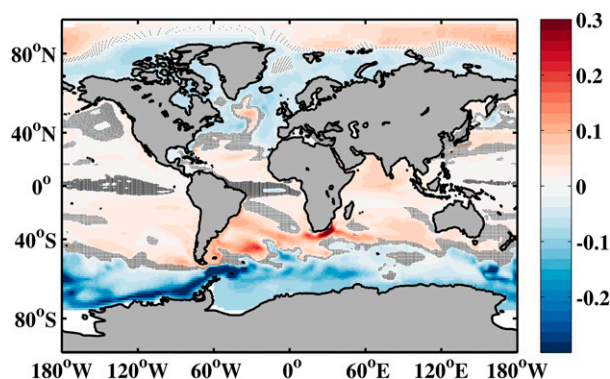


FIG. 8. Relative change in mean SSH between STO and REF (normalized by the linearly detrended interannual standard deviation of REF). Interpolated to a $1^\circ \times 1^\circ$ grid. Gray areas are not significant at the 0.05 level according to the two-sided Student's t test.

(not shown), it occurs largely in regions where the interannual standard deviation is underestimated by more than one order of magnitude. Normalization by the interannual standard deviation in Fig. 8 shows that the difference between the two simulations can reach up to 20%–30% of the interannual variability. In a more variable and nonlinear system these nonlinear rectification effects are likely to increase. The changes in the mean are also largely improving the mean state of the system when compared to observations (not shown), with increased SSH in the subtropics and decreased SSH in the high latitudes. The consequences are increased SSH gradients, but the improvement is far too small to compensate for the large systematic errors of the model. These changes, while small, are encouraging and we expect that for higher-resolution simulations it will also be of interest to investigate the mean response, in addition to the changes in low-frequency variability. This might also be true for coupled climate simulations, even with a 1° ocean model, since nonlinear feedbacks with the atmosphere might respond to the changes in low-frequency ocean variability.

In addition, while different magnitudes for temporal and spatial correlations as well as perturbation amplitude were tested in sensitivity experiments in this study, using high-resolution simulations or observations in combination with, for example, downscaling methods (e.g., Porta Mana and Zanna 2014) could in the future lead to better constraints of the parameters for the stochastic schemes. This might also lead to an improved impact of the schemes. Extensive observational estimates of convective and eddy time scales would be very useful, as well as an improved estimation of turbulent kinetic energy tendencies on larger scales, especially in the tropics.

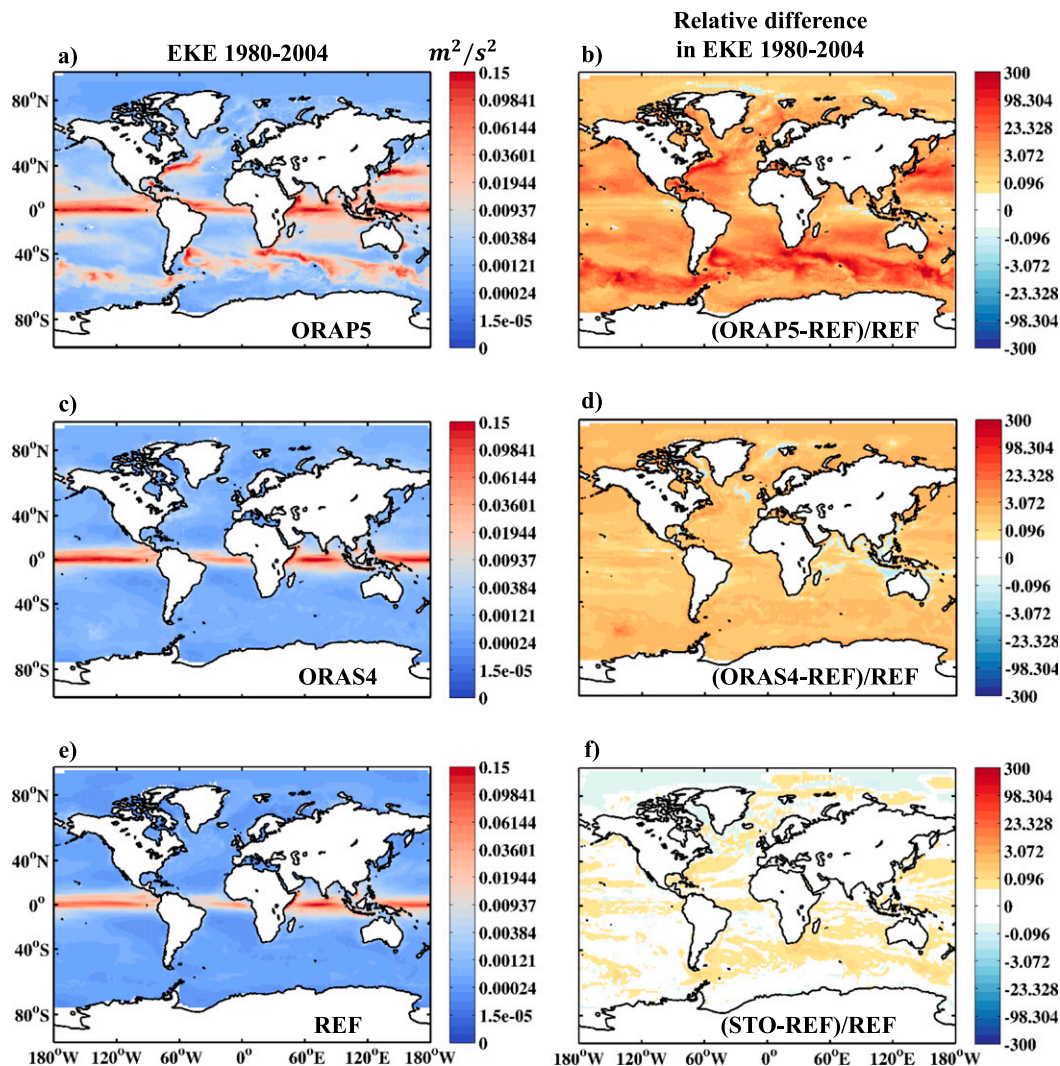


FIG. A1. EKE ($\text{m}^2 \text{s}^{-2}$) for (a) ORAP5, (c) ORAS4, and (e) REF, and the relative difference in EKE between REF and (b) ORAP5, (d) ORAS4, and (f) STO, based on the time period 1980–2004. For REF and STO, one simulation cycle from 1980 to 2004 is used. Data are calculated from the 5-daily averaged values of the velocity fields. For ORAS4, REF, and STO, the EKE is calculated from the residual velocities. Note the nonlinear color bars.

The question remains open whether we can accurately model low-frequency variability in a low-resolution 1° ocean model that is generally lacking variability and uses excessive dissipation for reasons of numerical stability. However, somewhat independent of this question, this study shows that there is a potential benefit in introducing subgrid-scale variability to the ocean system, which can improve the variability of the model climate when compared with observations. This could even lead so far that the effect of increased resolution can also be simulated by stochastic parameterizations as discussed, for example, by Berner et al. (2011) and Palmer (2012). This in

turn would reduce computational costs, because even though stochastic parameterizations and their stochastic pattern generators may add additional operations to the model, increasing resolution is far more expensive.

When it comes to ensemble simulations and forecasts, the stochastic schemes not only improve the low-frequency variability in the ocean model, but also provide an estimate of ocean model uncertainty. This is particularly important for seasonal to decadal forecasts, where it is essential to consider all sources of forecast uncertainty to provide accurate and reliable probabilistic forecasts of the coupled climate system. A similar

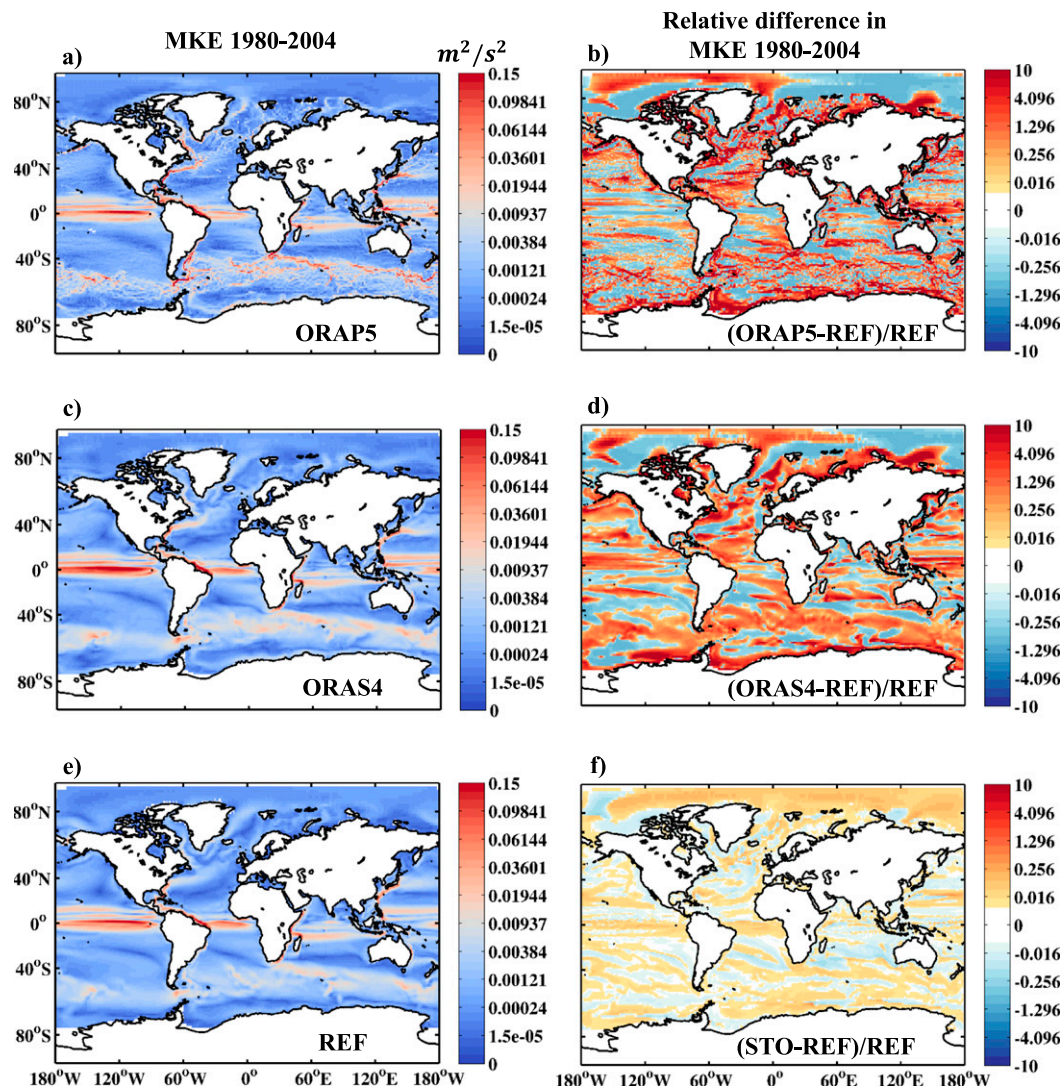


FIG. A2. As in Fig. A1, but for MKE.

argument holds for the usefulness of the stochastic schemes in ensemble data assimilation.

Finally, we note that accurately representing variability and time scales in a climate model is of great importance when it comes to analyzing climate sensitivity (see, e.g., Munday et al. 2013), predictability, and generally signal-to-noise ratios. An insufficient representation of the probability density function and the time scales of the climate system in models will not allow for meaningful predictions of climate change impacts. Therefore, both of these aspects need to be improved in addition to the mean state of the climate.

Acknowledgments. This work was supported by the UK NERC Grants NE/K013548/1 and NE/J00586X/1. The computations were carried out with ECMWF supercomputing resources. SJ and TNP were supported

under the European Research Council Grant PESM 291406. The Ssalto/Duacs altimeter products were produced and distributed by the Copernicus Marine and Environment Monitoring Service (CMEMS; <http://marine.copernicus.eu/>). We would like to thank the two anonymous reviewers for their helpful and constructive comments.

APPENDIX A

Model Mean State Performance

To illustrate that both the deterministic and the stochastic model simulations produce a reasonable climatology, Figs. A1 and A2 compare the model simulations to the two reanalysis products ORAS4 (1° resolution;

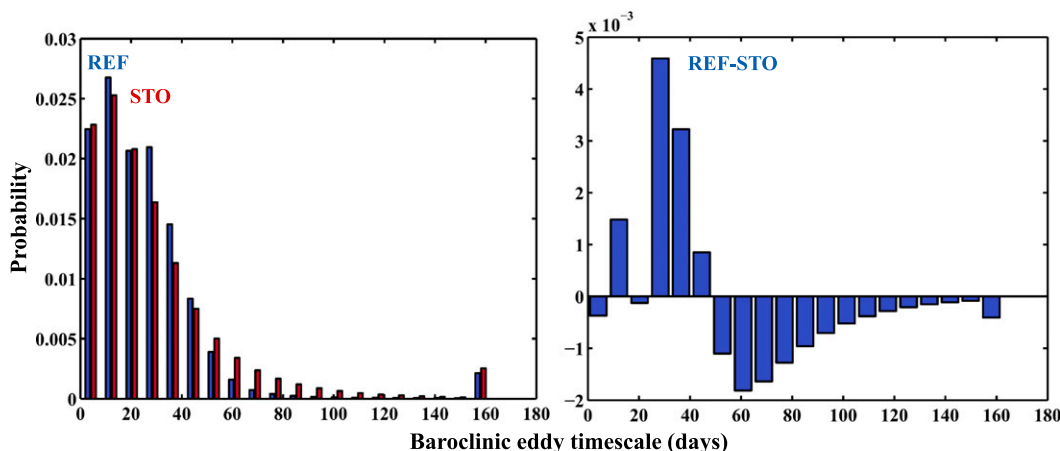


FIG. B1. Baroclinic eddy time scales (left) for a deterministic reference simulation (blue) and the stochastically perturbed simulation (red), and (right) the difference in time scales between deterministic and stochastic. The distributions are based on the global model output of daily snapshots of eddy time scales generated by sensitivity simulations for a period of 5 yr.

Mogensen et al. 2012; Balmaseda et al. 2013) and ORAP5 ($1/4^\circ$ resolution; Zuo et al. 2017) for the time period 1980–2004. The variables under consideration are eddy kinetic energy (EKE; Fig. A1) and mean kinetic energy (MKE; Fig. A2). The supplementary material also provides comparisons for SST and SSH (Figs. S1 and S2, respectively).

The mean SST and SSH climatologies of both the stochastic and deterministic model simulations agree reasonably well with the reanalysis products (Figs. S1 and S2). In line with well-known model biases, errors are relatively large in the Southern Ocean and the North Atlantic as well as along the western boundary currents. These are also regions where horizontal gradients are too small. For EKE and MKE (Figs. A1 and A2) there are resolution-related biases especially in the high latitudes due to the missing eddy activity in the 1° model, but also in the 1° reanalysis when compared with the $1/4^\circ$ reanalysis (cf. Figs. A1a and A1c and Figs. A1b and A1d). The model generally strongly underestimates both MKE and EKE. The stochastic simulation exhibits improved EKE with an increase of 10%–20% especially in the Southern Ocean, the Gulf Stream, and the Kuroshio when compared to the deterministic reference simulation. There is also an increase in EKE in the tropics. These increases are of similar amplitude as those in interannual SSH variability (section 4b), leading to a reduction of the EKE model bias. The EKE in the reanalysis is one to two orders of magnitude larger (see Figs. A1b and A1d) than the EKE simulated by REF.

The impacts of the perturbations on MKE are less uniform despite the general underestimation of MKE compared to the reanalysis (Fig. A2e). While there is

an increase in MKE in the tropics (Fig. A2f), the Southern Ocean and Kuroshio exhibit increases in MKE accompanied by a decrease in a neighboring region, suggesting a slight shift in the position of currents.

APPENDIX B

Eddy Time Scale Perturbations

The stochastic perturbations to the $A_{\text{GM}}^{\text{ref}}$ correspond to a skewed perturbation of the baroclinic eddy time scale Λ . Choosing a perturbation amplitude of $a = 0.8$ corresponds to a perturbation range for the eddy time scale from $1/(1 + 0.8) = 0.556$ to $1/(1 - 0.8) = 5$. This is a moderate perturbation given that eddy statistics are skewed with a long tail of long eddy lifetimes and high eddy amplitudes (Cheng et al. 2014; Samelson et al. 2016). However, choosing this range corresponds to an already relatively large symmetric perturbation of $A_{\text{GM}}^{\text{ref}}$ itself. An illustration of the effect of the perturbations on baroclinic eddy time scales is given in Fig. B1.

We would like to point out that the perturbations of $A_{\text{GM}}^{\text{ref}}$ through Λ can be viewed as an extension of the work done by Brankart (2013). Brankart (2013) investigated the nonlinear response of the equation of state to subgrid variations in temperature and salinity. However, in the implementation of these perturbations in NEMO he neglected the effect of density variations on the buoyancy frequency N , which affects the calculation both of the baroclinic eddy time scale Λ as well as the buoyancy term in the TKE Eq. (13). An extension to the work of Brankart (2013) could therefore be an

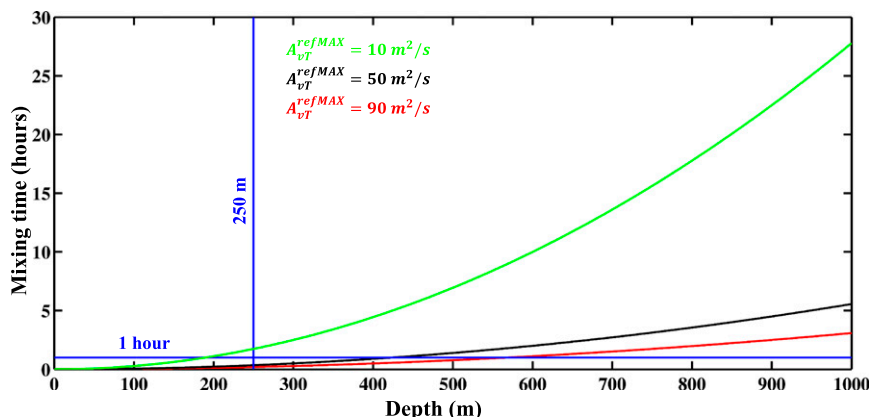


FIG. C1. Convection time scales (h) for different depths (m) and a range of vertical diffusion coefficients: $A_{vT}^{\text{refMAX}} = 50 \text{ m}^2 \text{ s}^{-1}$ (default deterministic value; black), $A_{vT}^{\text{refMAX}} = 10 \text{ m}^2 \text{ s}^{-1}$ (green), and $A_{vT}^{\text{refMAX}} = 90 \text{ m}^2 \text{ s}^{-1}$ (red) (minimum and maximum value for stochastic perturbations, respectively). The horizontal and vertical blue lines represent a time scale of 1 h and a depth of 250 m, respectively.

inclusion of the density perturbations in the local calculation of N . This might effectively also perturb the deep convection parameterization through the stability criterion evaluated for the water column at each grid point.

APPENDIX C

Convection Time Scale Perturbations

In cases of unstable stratification vertical diffusion is increased to a value of $A_{vT}^{\text{refMAX}} = 50 \text{ m}^2 \text{ s}^{-1}$ in the deterministic model. For a deep convection event that acts over a depth of 400 m this coefficient corresponds to a time scale of about 1 h for the water column to mix and stabilize. Variations to A_{vT}^{refMAX} affect this time scale as illustrated by Fig. C1. Shown are the maximum and minimum time scales Λ_{vT} generated by the symmetric stochastic perturbations to the vertical diffusion coefficient. Especially the low values of A_{vT}^{refMAX} will lead to strongly increased mixing time scales for deep convection events. The perturbations to the vertical diffusion coefficient impact deep convection events that affect a water column of at least around 250-m height, given a model time step of 1 h (highlighted by the blue lines in Fig. C1).

REFERENCES

- Andrejczuk, M., F. C. Cooper, S. Juricke, T. N. Palmer, A. Weisheimer, and L. Zanna, 2016: Oceanic stochastic parameterizations in a seasonal forecast system. *Mon. Wea. Rev.*, **144**, 1867–1875, doi:10.1175/MWR-D-15-0245.1.
- Antonov, J. I., and Coauthors, 2010: *Salinity*. Vol. 2, *World Ocean Atlas 2009*, S. Levitus, Ed., NOAA Atlas NESDIS 68, 184 pp.
- Bachman, S., D. Marshall, J. Maddison, and J. Mak, 2017: Evaluation of a scalar eddy transport coefficient based on geometric constraints. *Ocean Modell.*, **109**, 44–54, doi:10.1016/j.ocemod.2016.12.004.
- Balmaseda, M. A., K. Mogensen, and A. T. Weaver, 2013: Evaluation of the ECMWF ocean reanalysis system ORAS4. *Quart. J. Roy. Meteor. Soc.*, **139**, 1132–1161, doi:10.1002/qj.2063.
- , and Coauthors, 2015: The Ocean Reanalyses Intercomparison Project (ORA-IP). *J. Oper. Oceanogr.*, **8**, s80–s97, doi:10.1080/1755876X.2015.1022329.
- Berloff, P. S., and J. C. McWilliams, 2002: Material transport in oceanic gyres. Part II: Hierarchy of stochastic models. *J. Phys. Oceanogr.*, **32**, 797–830, doi:10.1175/1520-0485(2002)032<0797:MTIOGP>2.0.CO;2.
- Berner, J., S.-Y. Ha, J. P. Hacker, A. Fournier, and C. Snyder, 2011: Model uncertainty in a mesoscale ensemble prediction system: Stochastic versus multiphysics representations. *Mon. Wea. Rev.*, **139**, 1972–1995, doi:10.1175/2010MWR3595.1.
- Bishop, S. P., P. R. Gent, F. O. Bryan, A. F. Thompson, M. C. Long, and R. Abernathey, 2016: Southern Ocean overturning compensation in an eddy-resolving climate simulation. *J. Phys. Oceanogr.*, **46**, 1575–1592, doi:10.1175/JPO-D-15-0177.1.
- Blanke, B., and P. Delecluse, 1993: Variability of the tropical Atlantic Ocean simulated by a general circulation model with two different mixed-layer physics. *J. Phys. Oceanogr.*, **23**, 1363–1388, doi:10.1175/1520-0485(1993)023<1363:VOTTAO>2.0.CO;2.
- Bougeault, P., and P. Lacarrere, 1989: Parameterization of orography-induced turbulence in a mesobeta-scale model. *Mon. Wea. Rev.*, **117**, 1872–1890, doi:10.1175/1520-0493(1989)117<1872:POOITI>2.0.CO;2.
- Brankart, J.-M., 2013: Impact of uncertainties in the horizontal density gradient upon low resolution global ocean modelling. *Ocean Modell.*, **66**, 64–76, doi:10.1016/j.ocemod.2013.02.004.
- , G. Candille, F. Garnier, C. Calone, A. Melet, P.-A. Bouttier, P. Brasseur, and J. Verron, 2015: A generic approach to explicit simulation of uncertainty in the NEMO ocean

- model. *Geosci. Model Dev.*, **8**, 1285–1297, doi:[10.5194/gmdd-8-615-2015](https://doi.org/10.5194/gmdd-8-615-2015).
- Brodeau, L., B. Barnier, A.-M. Treguier, T. Penduff, and S. Gulev, 2010: An ERA40-based atmospheric forcing for global ocean circulation models. *Ocean Modell.*, **31**, 88–104, doi:[10.1016/j.ocemod.2009.10.005](https://doi.org/10.1016/j.ocemod.2009.10.005).
- Cheng, Y.-H., C.-R. Ho, Q. Zheng, and N.-J. Kuo, 2014: Statistical characteristics of mesoscale eddies in the North Pacific derived from satellite altimetry. *Remote Sens.*, **6**, 5164–5183, doi:[10.3390/rs6065164](https://doi.org/10.3390/rs6065164).
- Ferreira, D., and J. Marshall, 2006: Formulation and implementation of a “residual-mean” ocean circulation model. *Ocean Modell.*, **13**, 86–107, doi:[10.1016/j.ocemod.2005.12.001](https://doi.org/10.1016/j.ocemod.2005.12.001).
- Flato, G., and Coauthors, 2013: Evaluation of climate models. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 741–866.
- Fox-Kemper, B., S. Bachman, B. Pearson, and S. Reckinger, 2014: Principles and advances in subgrid modelling for eddy rich simulations. *CLIVAR Exchanges*, No. 65, International CLIVAR Project Office, Southampton, United Kingdom, 42–46.
- Gaspar, P., Y. Grégoris, and J.-M. Lefevre, 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. *J. Geophys. Res.*, **95**, 16 179–16 193, doi:[10.1029/JC095iC09p16179](https://doi.org/10.1029/JC095iC09p16179).
- Gent, P. R., and J. C. McWilliams, 1990: Isopycnal mixing in ocean circulation models. *J. Phys. Oceanogr.*, **20**, 150–155, doi:[10.1175/1520-0485\(1990\)020<0150:IMIOCM>2.0.CO;2](https://doi.org/10.1175/1520-0485(1990)020<0150:IMIOCM>2.0.CO;2).
- Goodman, L., E. R. Levine, and R. G. Lueck, 2006: On measuring the terms of the turbulent kinetic energy budget from an AUV. *J. Atmos. Oceanic Technol.*, **23**, 977–990, doi:[10.1175/JTECH1889.1](https://doi.org/10.1175/JTECH1889.1).
- Griffies, S. M., and Coauthors, 2009: Coordinated Ocean-ice Reference Experiments (CORES). *Ocean Modell.*, **26**, 1–46, doi:[10.1016/j.ocemod.2008.08.007](https://doi.org/10.1016/j.ocemod.2008.08.007).
- , and Coauthors, 2015: Impacts on ocean heat from transient mesoscale eddies in a hierarchy of climate models. *J. Climate*, **28**, 952–977, doi:[10.1175/JCLI-D-14-00353.1](https://doi.org/10.1175/JCLI-D-14-00353.1).
- Grooms, I., 2016: A Gaussian-product stochastic Gent–McWilliams parameterization. *Ocean Modell.*, **106**, 27–43, doi:[10.1016/j.ocemod.2016.09.005](https://doi.org/10.1016/j.ocemod.2016.09.005).
- , and A. J. Majda, 2013: Efficient stochastic superparameterization for geophysical turbulence. *Proc. Natl. Acad. Sci. USA*, **110**, 4464–4469, doi:[10.1073/pnas.1302548110](https://doi.org/10.1073/pnas.1302548110).
- Hazeleger, W., and Coauthors, 2012: EC-Earth v2.2: Description and validation of a new seamless Earth system prediction model. *Climate Dyn.*, **39**, 2611–2629, doi:[10.1007/s00382-011-1228-5](https://doi.org/10.1007/s00382-011-1228-5).
- Jansen, M. F., 2017: A note on “A Gaussian-product stochastic Gent–McWilliams parameterization.” *Ocean Modell.*, **110**, 49–51, doi:[10.1016/j.ocemod.2016.12.005](https://doi.org/10.1016/j.ocemod.2016.12.005).
- , and I. M. Held, 2014: Parameterizing subgrid-scale eddy effects using energetically consistent backscatter. *Ocean Modell.*, **80**, 36–48, doi:[10.1016/j.ocemod.2014.06.002](https://doi.org/10.1016/j.ocemod.2014.06.002).
- Juricke, S., and T. Jung, 2014: Influence of stochastic sea ice parameterization on climate and the role of atmosphere–sea–ocean interaction. *Philos. Trans. Roy. Soc.*, **372A**, 20130283, doi:[10.1098/rsta.2013.0283](https://doi.org/10.1098/rsta.2013.0283).
- , P. Lemke, R. Timmermann, and T. Rackow, 2013: Effects of stochastic ice strength perturbation on Arctic finite element sea ice modeling. *J. Climate*, **26**, 3785–3802, doi:[10.1175/JCLI-D-12-00388.1](https://doi.org/10.1175/JCLI-D-12-00388.1).
- Karspeck, A. R., and Coauthors, 2017: Comparison of the Atlantic meridional overturning circulation between 1960 and 2007 in six ocean reanalysis products. *Climate Dyn.*, doi:[10.1007/s00382-015-2787-7](https://doi.org/10.1007/s00382-015-2787-7), in press.
- Kim, H.-M., P. J. Webster, and J. A. Curry, 2012: Seasonal prediction skill of ECMWF system 4 and NCEP CFSv2 retrospective forecast for the Northern Hemisphere winter. *Climate Dyn.*, **39**, 2957–2973, doi:[10.1007/s00382-012-1364-6](https://doi.org/10.1007/s00382-012-1364-6).
- Kirtman, B. P., and Coauthors, 2012: Impact of ocean model resolution on CCSM climate simulations. *Climate Dyn.*, **39**, 1303–1328, doi:[10.1007/s00382-012-1500-3](https://doi.org/10.1007/s00382-012-1500-3).
- Kitsios, V., J. Frederiksen, and M. Zidikheri, 2013: Scaling laws for parameterisations of subgrid eddy–eddy interactions in simulations of oceanic circulations. *Ocean Modell.*, **68**, 88–105, doi:[10.1016/j.ocemod.2013.05.001](https://doi.org/10.1016/j.ocemod.2013.05.001).
- Kolmogorov, A. N., 1942: Equations of turbulent motion of an incompressible turbulent fluid. *Izv. Akad. Nauk SSSR*, **6**, 1–2.
- Lavender, K. L., R. E. Davis, and W. B. Owens, 2002: Observations of open-ocean deep convection in the Labrador Sea from subsurface floats. *J. Phys. Oceanogr.*, **32**, 511–526, doi:[10.1175/1520-0485\(2002\)032<0511:OOODC>2.0.CO;2](https://doi.org/10.1175/1520-0485(2002)032<0511:OOODC>2.0.CO;2).
- Locarnini, R. A., and Coauthors, 2010: *Temperature*. Vol. 1, *World Ocean Atlas 2009*, S. Levitus, Ed., NOAA Atlas NESDIS 68, 184 pp.
- Madec, G., 2008: NEMO ocean engine. Note du Pôle de Modélisation de l’Institut Pierre-Simon Laplace 27, 209 pp.
- , and M. Imbard, 1996: A global ocean mesh to overcome the North Pole singularity. *Climate Dyn.*, **12**, 381–388, doi:[10.1007/BF00211684](https://doi.org/10.1007/BF00211684).
- , P. Delecluse, M. Imbard, and C. Levy, 1998: OPA 8.1 Ocean General Circulation Model reference manual. Note du Pole de Modélisation 11, 91 pp.
- Marshall, D. P., J. R. Maddison, and P. S. Berloff, 2012: A framework for parameterizing eddy potential vorticity fluxes. *J. Phys. Oceanogr.*, **42**, 539–557, doi:[10.1175/JPO-D-11-048.1](https://doi.org/10.1175/JPO-D-11-048.1).
- Menemenlis, D., J. Campin, P. Heimbach, C. Hill, T. Lee, A. Nguyen, M. Schodlock, and H. Zhang, 2008: ECCO2: High resolution global ocean and sea ice data synthesis. Mercator Ocean Quarterly Newsletter, Mercator Ocean, Ramonville Saint-Agne, France, 13–21.
- Mogensen, K., M. A. Balmaseda, and A. Weaver, 2012: The NEMOVAR ocean data assimilation system as implemented in the ECMWF ocean analysis for System 4. ECMWF Tech. Memo. 668, 59 pp.
- Molteni, F., R. Buizza, T. N. Palmer, and T. Petroliagis, 1996: The ECMWF ensemble prediction system: Methodology and validation. *Quart. J. Roy. Meteor. Soc.*, **122**, 73–119, doi:[10.1002/qj.49712252905](https://doi.org/10.1002/qj.49712252905).
- Munday, D. R., H. L. Johnson, and D. P. Marshall, 2013: Eddy saturation of equilibrated circumpolar currents. *J. Phys. Oceanogr.*, **43**, 507–532, doi:[10.1175/JPO-D-12-095.1](https://doi.org/10.1175/JPO-D-12-095.1).
- Murphy, J. M., D. M. H. Sexton, D. N. Barnett, G. S. Jones, M. J. Webb, M. Collins, and D. A. Stainforth, 2004: Quantification of modelling uncertainties in a large ensemble of climate change simulations. *Nature*, **430**, 768–772, doi:[10.1038/nature02771](https://doi.org/10.1038/nature02771).
- Ollinaho, P., and Coauthors, 2016: Towards process-level representation of model uncertainties: Stochastically perturbed parametrisations in the ECMWF ensemble. *Quart. J. Roy. Meteor. Soc.*, **143**, 408–422, doi:[10.1002/qj.2931](https://doi.org/10.1002/qj.2931).
- Palmer, T. N., 2012: Towards the probabilistic Earth-system simulator: A vision for the future of climate and weather prediction. *Quart. J. Roy. Meteor. Soc.*, **138**, 841–861, doi:[10.1002/qj.1923](https://doi.org/10.1002/qj.1923).

- , G. Shutts, R. Hagedorn, F. Doblas-Reyes, T. Jung, and M. Leutbecher, 2005: Representing model uncertainty in weather and climate prediction. *Annu. Rev. Earth Planet. Sci.*, **33**, 163–193, doi:[10.1146/annurev.earth.33.092203.122552](https://doi.org/10.1146/annurev.earth.33.092203.122552).
- Pasquero, C., and E. Tziperman, 2007: Statistical parameterization of heterogeneous oceanic convection. *J. Phys. Oceanogr.*, **37**, 214–229, doi:[10.1175/JPO3008.1](https://doi.org/10.1175/JPO3008.1).
- Porta Mana, P., and L. Zanna, 2014: Toward a stochastic parameterization of ocean mesoscale eddies. *Ocean Modell.*, **79**, 1–20, doi:[10.1016/j.ocemod.2014.04.002](https://doi.org/10.1016/j.ocemod.2014.04.002).
- Reynolds, R. W., T. M. Smith, C. Liu, D. B. Chelton, K. S. Casey, and M. G. Schlax, 2007: Daily high-resolution-blended analyses for sea surface temperature. *J. Climate*, **20**, 5473–5496, doi:[10.1175/2007JCLI1824.1](https://doi.org/10.1175/2007JCLI1824.1).
- Richter, I., 2015: Climate model biases in the eastern tropical oceans: Causes, impacts and ways forward. *Wiley Interdiscip. Rev.: Climate Change*, **6**, 345–358, doi:[10.1002/wcc.338](https://doi.org/10.1002/wcc.338).
- Samelson, R. M., M. G. Schlax, and D. B. Chelton, 2016: A linear stochastic field model of midlatitude mesoscale variability. *J. Phys. Oceanogr.*, **46**, 3103–3120, doi:[10.1175/JPO-D-16-0060.1](https://doi.org/10.1175/JPO-D-16-0060.1).
- Seo, H., M. Jochum, R. Murtugudde, and A. J. Miller, 2006: Effect of ocean mesoscale variability on the mean state of tropical Atlantic climate. *Geophys. Res. Lett.*, **33**, L09606, doi:[10.1029/2005GL025651](https://doi.org/10.1029/2005GL025651).
- Treguier, A. M., I. M. Held, and V. D. Larichev, 1997: Parameterization of quasigeostrophic eddies in primitive equation ocean models. *J. Phys. Oceanogr.*, **27**, 567–580, doi:[10.1175/1520-0485\(1997\)027<0567:POQEIP>2.0.CO;2](https://doi.org/10.1175/1520-0485(1997)027<0567:POQEIP>2.0.CO;2).
- Vancoppenolle, M., T. Fichefet, H. Goosse, S. Bouillon, G. Madec, and M. A. M. Maqueda, 2009: Simulating the mass balance and salinity of Arctic and Antarctic sea ice. 1. Model description and validation. *Ocean Modell.*, **27**, 33–53, doi:[10.1016/j.ocemod.2008.10.005](https://doi.org/10.1016/j.ocemod.2008.10.005).
- Weisheimer, A., S. Corti, T. Palmer, and F. Vitart, 2014: Addressing model error through atmospheric stochastic physical parametrizations: Impact on the coupled ECMWF seasonal forecasting system. *Philos. Trans. Roy. Soc.*, **372A**, 20130290, doi:[10.1098/rsta.2013.0290](https://doi.org/10.1098/rsta.2013.0290).
- Williams, P. D., 2012: Climatic impacts of stochastic fluctuations in air–sea fluxes. *Geophys. Res. Lett.*, **39**, L10705, doi:[10.1029/2012GL051813](https://doi.org/10.1029/2012GL051813).
- , N. J. Howe, J. M. Gregory, R. S. Smith, and M. M. Joshi, 2016: Improved climate simulations through a stochastic parameterization of ocean eddies. *J. Climate*, **29**, 8763–8781, doi:[10.1175/JCLI-D-15-0746.1](https://doi.org/10.1175/JCLI-D-15-0746.1).
- Zidikheri, M. J., and J. S. Frederiksen, 2009: Stochastic subgrid parameterizations for simulations of atmospheric baroclinic flows. *J. Atmos. Sci.*, **66**, 2844–2858, doi:[10.1175/2009JAS3036.1](https://doi.org/10.1175/2009JAS3036.1).
- Zuo, H., M. A. Balmaseda, and K. Mogensen, 2017: The new eddy-permitting ORAP5 ocean reanalysis: Description, evaluation and uncertainties in climate signals. *Climate Dynamics*, doi:[10.1007/s00382-015-2675-1](https://doi.org/10.1007/s00382-015-2675-1), in press.