Uncertainty and scale interactions in ocean ensembles: from seasonal forecasts to multidecadal climate predictions

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

RESEARCH ARTICLE

Uncertainty and scale interactions in ocean ensembles: From seasonal forecasts to multidecadal climate predictions

L. Zanna[1] | J. M. Brankart[2] | M. Huber[1] | S. Leroux[3] | T. Penduff[2] | P. D. Williams[4]

[1]Atmospheric, Oceanic & Planetary Physics, Department of Physics, University of Oxford, Oxford, UK
[2]MEOM Team, Univ. Grenoble Alpes, CNRS, IRD, Grenoble INP, IGE, Grenoble, France
[3]Ocean Next, Grenoble, France
[4]Department of Meteorology, University of Reading, Reading, UK

Correspondence
L. Zanna, Department of Physics, Atmospheric, Oceanic & Planetary Physics, University of Oxford, Parks Road, Oxford OX1 3PU, UK.
Email: laure.zanna@physics.ox.ac.uk

The ocean plays an important role in the climate system on time-scales of weeks to centuries. Despite improvements in ocean models, dynamical processes involving multiscale interactions remain poorly represented, leading to errors in forecasts. We present recent advances in understanding, quantifying, and representing physical and numerical sources of uncertainty in novel regional and global ocean ensembles at different horizontal resolutions. At coarse resolution, uncertainty in 21st century projections of the upper overturning cell in the Atlantic is mostly a result of buoyancy fluxes, while the uncertainty in projections of the bottom cell is driven equally by both wind and buoyancy flux uncertainty. In addition, freshwater and heat fluxes are the largest contributors to Atlantic Ocean heat content regional projections and their uncertainties, mostly as a result of uncertain ocean circulation projections. At both coarse and eddy-permitting resolutions, unresolved stochastic temperature and salinity fluctuations can lead to significant changes in large-scale density across the Gulf Stream front, therefore leading to major changes in large-scale transport. These perturbations can have an impact on the ensemble spread on monthly time-scales and subsequently interact nonlinearly with the dynamics of the flow, generating chaotic variability on multiannual time-scales. In the Gulf Stream region, the ratio of chaotic variability to atmospheric-forced variability in meridional heat transport is larger than 50% on time-scales shorter than 2 years, while between 40 and 48°S the ratio exceeds 50% on on time-scales up to 28 years. Based on these simulations, we show that air–sea interaction and ocean subgrid eddies remain an important source of error for simulating and predicting ocean circulation, sea level, and heat uptake on a range of spatial and temporal scales. We discuss how further refinement of these ensembles can help us assess the relative importance of oceanic versus atmospheric uncertainty in weather and climate.

KEYWORDS

climate, ensemble simulations, modelling, ocean, stochastic parametrizations, uncertainties

1 | INTRODUCTION

The ocean has a much larger heat capacity than the atmosphere and cryosphere, and plays a leading-order role in weather and climate. The thermal inertia of the ocean provides long-term memory to the climate system. More than 90% of the excess heat in the climate system in the last 50 years has been stored in the ocean (Levitus *et al.*, 2012), with important consequences for sea-level rise. The ocean not only stores heat and other tracers like carbon and oxygen but can also redistribute them, affecting sea level regionally (Gregory *et al.*, 2016) and maintaining the sea-ice edge by transporting heat to high latitudes (for example, Winton, 2003).

Taking the ocean dynamics into account is important for weather and seasonal forecasting, in both the Tropics and extratropics. In the Tropics, for example, the Madden–Julian Oscillation (MJO), which is regarded primarily as an atmospheric phenomenon, is influenced by the ocean and coupled atmosphere–ocean processes, which improve the skill of MJO forecasts (Woolnough *et al.*, 2007; DeMott *et al.*, 2015). The El Niño–Southern Oscillation (ENSO) is another example of a coupled phenomenon in which the ocean dynamics sets the spatial pattern and period of ENSO events (Philander, 1983). The ocean is also important for forecasting the state of the atmosphere in the extratropics. One symptom underlying this importance is the adverse effect of North Atlantic sea-surface temperature (SST) biases on European climate predictions (Scaife *et al.*, 2010).

Scale interactions and two-way feedbacks shape the ocean dynamics. The ocean is driven mainly by atmospheric fluxes of energy, momentum, and matter at the sea surface (Peixoto and Oort, 1984). The ocean dynamics is characterized by a range of phenomena including eddies, zonal jets, internal waves, and mixing. These phenomena are linked, in the sense that eddies may merge and create zonal jets through an anisotropic inverse cascade, eddies may generate internal waves through geostrophic adjustment, and internal waves create mixing when they break (Vallis, 2006). Nonlinear scale interactions and feedbacks between different features in the climate system could involve two processes in the ocean, or a process in the ocean interacting with a process in the atmosphere. For example, in the ocean–ocean category, eddies may influence the large-scale wind- and buoyancy-driven circulation via the spatial inverse cascade associated with two-dimensional geostrophic turbulence (Arbic *et al.*, 2007). This scale interaction is also featured in the temporal domain, with fast nonlinear chaotic fluctuations influencing slow modes of variability (Arbic *et al.*, 2012; Serazin *et al.*, 2018). In the ocean–atmosphere category, turbulent air–sea fluxes lead to coupling at all latitudes and may involve the intertropical convergence zone, Hadley cells, and jet stream (Dong and Sutton, 2002).

A critical limitation when modelling the above phenomena is that—due to the wide range of interactions at different scales and between different components—numerical predictions are inherently uncertain (Palmer *et al.*, 2005). There are different types of uncertainty. Uncertainty in the initial conditions (IC) of oceanic and atmospheric models arises from poor observational coverage or imprecise model initialization. Another important source of uncertainty is the lack of numerical resolution, such that many processes are subgrid-scale and unresolved. The effects of these processes on the resolved scales have to be parametrized. Important subgrid-scale processes include air–sea fluxes (bulk formula), mixed-layer processes (for example, restratification), mesoscale and submesoscale eddies, and diapycnal mixing. The lack of explicit representation of subgrid-scale processes and imperfect parametrizations inhibits many of the scale interactions described above. Therefore, when considering uncertainty related to subgrid or chaotic processes, there will be two components to the error: a random error associated with the impact of unresolved or resolved chaotic processes on the large-scale flow, and systematic errors due to imperfect parametrizations and parameters (Allen *et al.*, 2002).

The discretization in space and time of the equations of motion (Teixeira *et al.*, 2007) yields errors from the truncation and unresolved scales, and from the different numerical schemes and vertical coordinates (for example, Gibson *et al.*, 2017). For example, there have been recent developments in stabilizing the computational mode of the commonly used leapfrog time-stepping scheme using the Robert–Asselin–Williams (RAW) filter (Williams, 2009; 2011; 2013). Specifically, the use of the RAW filter, instead of the Robert–Asselin filter, reduces biases in the location of the Gulf Stream (Young *et al.*, 2014). Therefore changes in the numerics might help correct for the bias, in the same way that stochastic and chaotic variability can alleviate the bias.

None of the errors described above is truly independent. How can we reduce uncertainties in simulations and forecasts associated with random and systematic errors, or at least account for them? First, one could increase the numerical resolution; however, this comes at a computational cost and the models would still be missing processes at finer scales down to the Kolmogorov scale of millimetres to centimetres, below which viscosity ensures that motions are not energized. Second, one could increase model complexity, for example, by adding more components and new parametrizations of missing processes. Complexity again increases computational cost, and the new parametrizations need to be developed such that their implementation does not deteriorate the model's performance. Third, one could include new ways to represent uncertainty, such as perturbed parameters informed by observations or the use of stochastic physics, as is done routinely in several atmospheric models (Leutbecher *et al.*, 2017).

The best way forward is likely a combination of the three solutions mentioned above. In the present article, we will focus on the latter two. We will present newly designed ocean ensembles at coarse and eddy-permitting resolutions for weather and climate. We will concentrate on the representation of uncertainty in the ocean component of the system and discuss the results in terms of scale interactions and feedbacks on the resolved scales.

2 | DIVERSITY OF OCEAN ENSEMBLES AND UNCERTAINTY

Ensembles for weather and seasonal forecasts have a long history, as reported in this special issue. The use of a dynamical ocean model for a range of studies, from weather forecasts to climate predictions, is increasingly common. Below, we briefly review representations of uncertainties and then highlight several applications using ocean ensembles with uncertainty representation.

TABLE 1 Summary of ocean ensemble configurations used in sections 2 and 3

Model name	Configuration	Horizontal resolution (°)	Ensemble set-up	Simulations	Members	Ref. section(s)
MITgcm – CMIP5 forcing	Global	2.8	Deterministic surface forcing	Climatology and climate change	29	2.2.4; 3.1
OPA-ECHAM	Global	2	Stochastic air–sea forcing	Climatology	1	3.1
NEMO-ORCA2	Global	2	Stochastic density for model error	Climatology	1	3.2
NEMO-NATL025	Regional	0.25	Stochastic density for model error	Hindcast	13	2.2.2; 3.2
NEMO-OCCIPUT	Global	0.25	Stochastic density for I.C. generation	Hindcast	50	2.2.3; 3.3

2.1 | Representation and quantification of uncertainty in ensembles

There are standard methods to represent uncertainty in the ocean component of weather and climate models for ensemble simulations and forecasts, namely: perturbed initial state, perturbed physics (parameters and parametrizations), and multi-physics (such as multimodels), and more recently stochastic physics has also gained some traction.

Initial conditions (IC) of oceanic and atmospheric models are inherently uncertain, due to the lack of observations or imperfect model initialization. Oceanic and atmospheric IC uncertainties (together with model uncertainties) are likely to persist on long time-scales. Therefore, these uncertainties need to be simulated and characterized, to assess the robustness of oceanic hindcasts (ocean simulations driven by realistic atmospheric forcings) and the sensitivity of ocean and climate simulations to ocean model uncertainty and initialization. The sensitivity of the ocean state to IC uncertainty may be represented in ensemble simulations by slightly perturbing the ocean model IC, as is done in atmospheric models for weather and seasonal forecasts. Such IC perturbations may be introduced in the ensemble members at different times, to evaluate the fate of IC uncertainties for various background oceanic states (for example, different stratifications, Atlantic Meridional Overturning (AMOC) magnitude). Studies have been carried out to emulate these IC uncertainties and describe their spatio-temporal evolution. Many studies, using laminar ensemble simulations, have identified optimal IC perturbations that maximize the response of climate-relevant oceanic indices (for example, heat content, ENSO, AMOC; see Kleeman and Moore, 1997; Zanna *et al.*, 2011; Sevellec and Fedorov, 2017), their possible impacts in coupled models (see Tziperman *et al.*, 2008; Hawkins and Sutton, 2009; Germe *et al.*, 2017a, 2017b), and their use in targeting new observations for prediction (Zanna *et al.*, 2011; 2012). Other studies (section 3.3) focus on the fate of IC uncertainties and chaotic ocean behaviour in higher-resolution ocean model ensembles, in regimes where nonlinear turbulent scale interactions are at play (for example, Spall, 1996; Dewar, 2003; Berloff *et al.*, 2007).

For each parametrization scheme of a given process, a parameter is introduced. However, such parameters are often poorly known and poorly constrained. Despite this, the weather and climate states are very sensitive to these parameters. In perturbed parameter ensembles, different parameter values are sampled from a distribution representing their uncertainty. The distribution can be ad hoc or created from observations or high-resolution simulations. Each ensemble member is then assigned a different set of parameters. However, while these ensembles target some uncertainty in the parameters, they are not representing the uncertainty in the parametrization itself. An alternative technique for representing uncertainty in parametrization is the use of stochastic schemes. These use random numbers to represent the uncertainty in the parametrization scheme itself, accounting for the unresolved subgrid-scale variability associated with a given process. Many such stochastic physics schemes have been introduced over the years and some will be described in sections 2.2.2, 3.1 and 3.2. Finally, a multimodel approach has become common practice to target structural and parameter uncertainty for climate predictions simultaneously (for example, Kharin and Zwiers, 2002). The mean of the ensemble of Atmosphere–Ocean General Circulation Models (AOGCMs) from different modelling centres is used to provide the "best estimate" forecast, assuming the simulation errors in different AOGCMs are independent. The spread of the ensemble then corresponds to the uncertainty, which encompasses the different numerical schemes, parametrization choice, and parameter values. The ensemble mean and spread are often considered to be an improved basis for probabilistic projections, compared with ensembles based on a single model (Palmer *et al.*, 2005).

While none of the approaches described above represents all uncertainties, they are certainly the most common ones.

2.2 | A range of ocean-based ensembles

Below we provide examples of ensembles used for weather and climate predictions. Some of the ensembles introduced below are then used in section 3 to explore the key sources of uncertainty for ocean states relevant for weather and climate and are summarized in Table 1.

2.2.1 | Initialized ensemble forecasts

Ensembles are in widespread use for operational weather and seasonal forecasts. Most medium-range and seasonal forecast systems involve ensemble prediction to help formulate probabilistic forecasts and to provide users with an indication

of forecast uncertainty. For example, the European Centre for Medium-Range Weather Forecasts (ECMWF) produces weather forecasts for up to 15 days ahead using an ensemble of 52 individual ensemble members twice a day. Optimal perturbations from singular vectors for initial conditions and stochastic schemes have been operational at ECMWF since 1998 (Buizza *et al.*, 1999). The initial states and model physics in the ensemble members are perturbed to explore the currently understood range of uncertainty in the observations and the model. The result is a plume of possible futures, in which the different types of uncertainties are accounted for.

2.2.2 | Links between model and observational uncertainty and the NATL025 ensemble

An essential aspect of a seasonal forecast system is the ocean initialization. Ensemble-based data assimilation methods provide an ensemble of states to initialize the forecasts with state-dependent background-error covariances combining observations and models. Several existing ocean ensemble data assimilation products are widely used (Yin *et al.*, 2011). Different ensembles are designed to explore the role of uncertainty in the initialization or in the products derived for initializations. In addition, each data-assimilation product is influenced by the underlying model used and the errors associated with it. These model errors will then propagate and affect the forecasts on a range of time-scales.

Understanding and representing model uncertainties is necessary to assess how much information from the "true system" is contained in models; this is a key asset for ocean data assimilation and to produce reliable probabilistic hindcasts and forecasts. Ensembles of simulations have been performed to explore the effects of explicit representation of initial conditions and model uncertainties in a range of simulations (Williams, 2012; Brankart, 2013; Andrejczuk *et al.*, 2016; Williams *et al.*, 2016; Juricke *et al.*, 2017). For example, Andrejczuk *et al.* (2016) use a coarse (1°) resolution coupled GCM to investigate the impact of uncertainty in different sub-grid schemes using stochastic physics versus initial-condition perturbations from singular vectors. They showed that the ensemble spread for certain quantities can be increased in regions such as the Gulf Stream on seasonal time-scales. However, the uncertainty from the chaotic atmospheric variability and that of the ocean initial conditions was often dominant over that of stochastically represented ocean model error. This result can be contrasted with other studies using uncoupled models (Brankart, 2013) or even coarser resolutions (Williams, 2012), which found a large impact from stochastically represented ocean model error (nonlinear density equation and air–sea turbulent fluxes, respectively) on the large-scale climatological ocean state.

Recently, novel eddy-permitting (1/4° resolution) ensembles of the North Atlantic have been generated using a configuration of the ocean model Nucleus for European Modeling of the Ocean (NEMO), NEMO-NATL025 (developed by the DRAKKAR consortium: see Barnier *et al.*, 2006), with a focus on the stochastic effect of unresolved scales in the computation of density (as in Brankart *et al.*, 2015; see section 3.2 for results). The original purpose of these ensemble simulations was to investigate to what extent uncertain model operators can be made statistically consistent with observations (for example, satellite altimetry or Array for Real-time Geostrophic Oceanography (ARGO) floats). Inconsistency here would imply that our assumptions about model or observation accuracy are incorrect, and that further fundamental studies are needed to identify the missing sources of random or structural uncertainty associated with the system. Following this approach, ensemble data assimilation experiments have been performed to reduce uncertainties in the NEMO-NATL025 ensemble simulations using altimeter observations (Candille *et al.*, 2015). Candille *et al.* (2015) quantified the reliability and resolution (that is, statistical consistency and information content) of ensemble analyses and short-term forecasts based on probabilistic scores (for example, Andrejczuk *et al.*, 2016). Candille *et al.* (2015) highlight the usefulness of representing model error via stochastic physics to produce a reliable ensemble and the need of initialization for skilful forecasts.

2.2.3 | Dynamical uncertainty in hindcasts and the OCCIPUT ensemble: long-term impacts of ocean chaos on dynamics

The long-term memory of the climate system comes from the ocean. Uncertainties in oceanic and atmospheric IC uncertainties, together with model uncertainties, are likely to persist and influence ocean dynamics on long time-scales. Idealized simulations (see, for example, the review by Dijkstra and Ghil, 2005) have demonstrated that, when the resolution of ocean models is fine enough to represent key nonlinear processes, the ocean itself spontaneously generates a strong chaotic, intrinsic variability with temporal scales reaching from years to decades and spatial scales of thousands of kilometres. Mesoscale turbulence is one of these key processes, and its explicit representation in ocean models can lead to the emergence of low-frequency chaotic variability, even without any low-frequency atmospheric forcing (for example, Penduff *et al.*, 2011; Gregorio *et al.*, 2015). This low-frequency chaotic variability indeed emerges under a repeated seasonal forcing, and locally can reach the amplitude of the variability obtained in oceanic hindcasts, driven by the full range of atmospheric time-scales.

In order to disentangle the chaotic part of the ocean variability and its atmospherically driven (forced) part, the OceaniC Chaos – ImPacts, strUcture, predicTability (OCCIPUT) project has performed a 50 member ensemble of 56 year, global NEMO ocean–sea-ice hindcasts at 1/4° resolution, based on a probabilistic version of the NEMO ocean model

(Penduff *et al.*, 2014; Bessieres *et al.*, 2017). The OCCIPUT experiment is a nonassimilative, global and longer version of the regional ensemble experiment NATL025 (section 2.2.2). The 50 members of OCCIPUT were perturbed stochastically (using the methodology described in section 3.2) for one year only (after a common spinup), and then driven between 1960 and 2015 by the same realistic forcing derived from an atmospheric reanalysis. OCCIPUT was designed to study the long-term fate of small IC uncertainties, or equivalently of the chaotic variability, in climate-relevant oceanic indices over a wide range of scales. These results complement those obtained in lower-resolution (Germe *et al.*, 2017a, 2017b) and idealized (Wilson *et al.*, 2015) oceanic ensembles regarding the fate of IC uncertainties.

2.2.4 | Climatological simulations and forced projections, and the CMIP5-forced MITgcm ensemble

Coupled climate models, in coordinated efforts with a common set of experiments, can produce climatological (control) simulations and projections under a range of scenarios. This multimodel approach, often referred to as an "ensemble of opportunity" (Tebaldi and Knutti, 2007), has helped in sampling initial condition uncertainty, parameter and structural uncertainty in the climate models participating in the Coupled Model Intercomparison Project Phase (CMIP) ensemble. However, this ensemble of opportunity does not help in identifying and understanding sources of uncertainties. Therefore, several new ensembles and protocols have been designed for this purpose.

Small ensembles of realistic models with differing horizontal resolution can target the impact of a specific resolved versus parametrized process on large-scale dynamics. By keeping the same model and changing the resolution and a minimal set of parameters, the effect of resolution can be assessed. For example, increased vertical upward eddy heat and salt fluxes in the ocean interior with increasing ocean horizontal resolution have been demonstrated in both the Geophysical Fluid Dynamics Laboratory (GFDL) and Max Planck Institute (MPI) coupled climate models (Griffies *et al.*, 2015; von Storch *et al.*, 2016). Furthermore, the magnitude of the spatial inverse cascade (that is, kinetic energy flux from small to large scale) increases as resolution increases (Kjellsson and Zanna, 2017), influencing jet dynamics in the Gulf Stream, Kuroshio and Antarctic circumpolar regions. The influence of resolved eddies also increases the surface fluxes out of the ocean in the same regions (Roberts *et al.*, 2016).

Recently, Huber and Zanna (2017) have designed a new low-resolution ocean-only setup to analyse the effects of the uncertainty arising from air-sea fluxes and model parameters on the ocean circulation and heat uptake. They used the surface boundary conditions from CMIP5 models (air–sea fluxes, surface temperature, and salinity) to force the same Massachusetts Institute of Technology General Circulation Model (MITgcm) ocean model. They were able to reproduce most of the CMIP5 ocean interior fields and concluded that surface forcing is the main source of uncertainty, over subgrid-scale parameters, for AMOC under pre-industrial and 1% CO_2 forcing scenarios, and for Atlantic ocean heat content (OHC) change. However parametric uncertainty remains important and dominant in other basins. We will use their design to explore individually the role of momentum and buoyancy forcing on the circulation and heat uptake in section 3.1.

3 | DOMINANT SOURCES OF UNCERTAINTY

Using some of the diverse and novel ensembles described in the previous section, we highlight several key uncertainties in the ocean component of the climate system. We concentrate mainly on the Atlantic ocean basin, with a few exceptions.

3.1 | Air–sea interaction uncertainty

Air–sea fluxes are critical in forcing the ocean circulation and its variability, while feeding information back from the ocean into the atmosphere on a range of scales. Large-scale air–sea fluxes govern the large-scale ocean circulation. Based on the same ensemble design as Huber and Zanna (2017) (described in section 2.2.4), the impact of individual air–sea fluxes on large-scale ocean properties can be assessed, following a protocol inspired by the Flux-Anomaly-Forced Model Intercomparison Project (FAFMIP) experiments (Gregory *et al.*, 2016). By using the surface fluxes and surface properties (SST and sea-surface salinity (SSS)) of different CMIP5 models, we can quantify the uncertainty arising from individual boundary conditions. The ocean-only MITgcm ensemble is forced with the surface boundary conditions from 29 CMIP5 models under 1% CO_2/yr or RCP8.5 forcing scenarios (each ensemble member was first spun up with its associated climatological seasonally dependent surface forcings from CMIP5).

Under climate change scenarios, we find a large uncertainty in AMOC weakening projections (between 0 and 5 Sv) and Antarctic Bottom Water (AABW) change (between −1 Sv and 5 Sv). The AMOC and AABW strength are defined as the maximum of the meridional overturning streamfunction in the Atlantic between 20°N and 60°N and below 500 m, and south of the Equator and below 2000 m, respectively. The spread in AMOC is mostly a result of buoyancy fluxes (90%), rather than arising from the wind forcing (10%), as shown in Figure 1a. The spread in AABW transport is driven equally by both wind and buoyancy uncertainty (although buoyancy might result in more spread over longer integrations). However, note the large decadal variability in both AMOC and AABW. Under RCP8.5, shown in Figure 2, regional changes in OHC are due to a combination of heat, freshwater (FW), and wind. The contribution from the heat flux is fairly uniform, except for the "cold spot" around
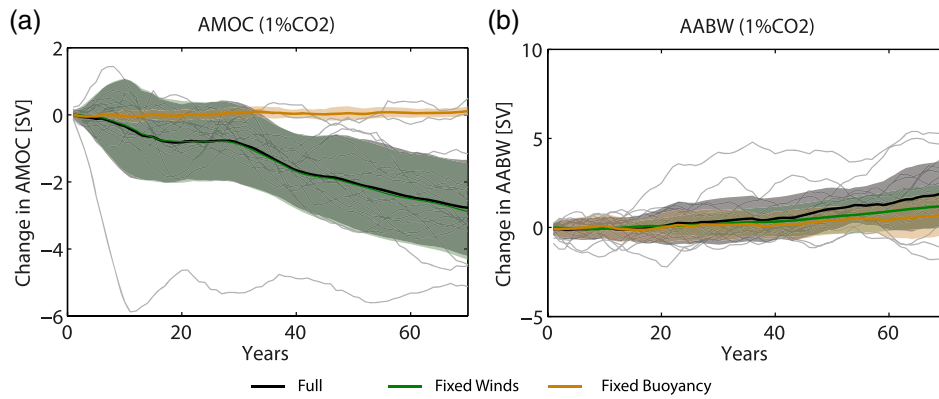
FIGURE 1 MITgcm ocean simulations forced with prescribed surface boundary conditions taken from 29 CMIP5 1% CO_2 experiments for temperature, salinity, and momentum: (a) AMOC and (b) AABW magnitude (see definitions within the main text). The panels show individual model simulations (grey curves) and the ensemble mean (black curves) when the model is forced with all surface fields; ensemble mean and spread (green curve and shading, respectively) when using different wind forcing from the CMIP5 ensemble but the ensemble-mean buoyancy forcing; and ensemble mean and spread (orange curve and shading, respectively) using different buoyancy forcing from the CMIP5 ensemble and the ensemble-mean wind forcing [Colour figure can be viewed at wileyonlinelibrary.com]
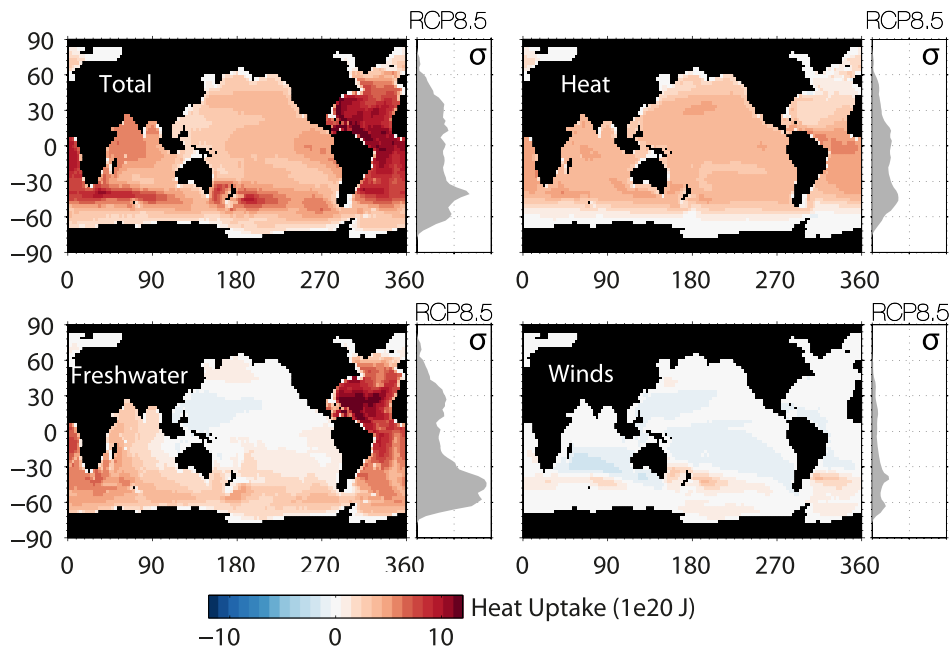

FIGURE 2 Vertically integrated OHU ensemble mean (maps) and standard deviation (panels with grey curves for the zonal integral) in MITgcm simulations under the RCP 8.5 scenario with forcing from all fluxes; heat flux alone; freshwater flux alone; and wind-only from 29 CMIP5 models [Colour figure can be viewed at wileyonlinelibrary.com]

Antarctica and in the North Atlantic. In the Atlantic, FW and heat fluxes are the largest contributors to OHC changes. The strong changes in FW have a strong influence on the ocean circulation in the North Atlantic and the associated heat uptake; however, this signal might be amplified due to the coarse resolution of the MITgcm model setup. In the Southern Ocean, in addition to FW and heat influence, the wind forcing—via Ekman pumping—is significant. The spread (uncertainty) in OHC is due mainly to the difference in FW and heat fluxes—especially in the Southern Ocean.

The patterns resemble those found in Gregory *et al.* (2016) using AOGCMs under the FAFMIP protocol, which uses the same anomalous flux for heat, wind, and FW. Therefore the

ocean-only climate ensemble shows that the uncertainty in air–sea forcing in climate change scenarios at coarse resolution is large and can explain the uncertainty (as measured by the spread between models) of the CMIP5 ensemble. However, the spread measured is not the true uncertainty in the projections, since all simulations are coarse-resolution and lack both turbulent (nonlinear) ocean processes and high-frequency air–sea fluxes.

Different studies have shown the impact of high-frequency air–sea forcing on the upper ocean and large-scale circulation, from idealized box models (Zanna and Tziperman, 2008) to AOGCMs (Williams, 2012). To examine the impacts of unresolved air–sea fluxes on ocean climatology and variability,
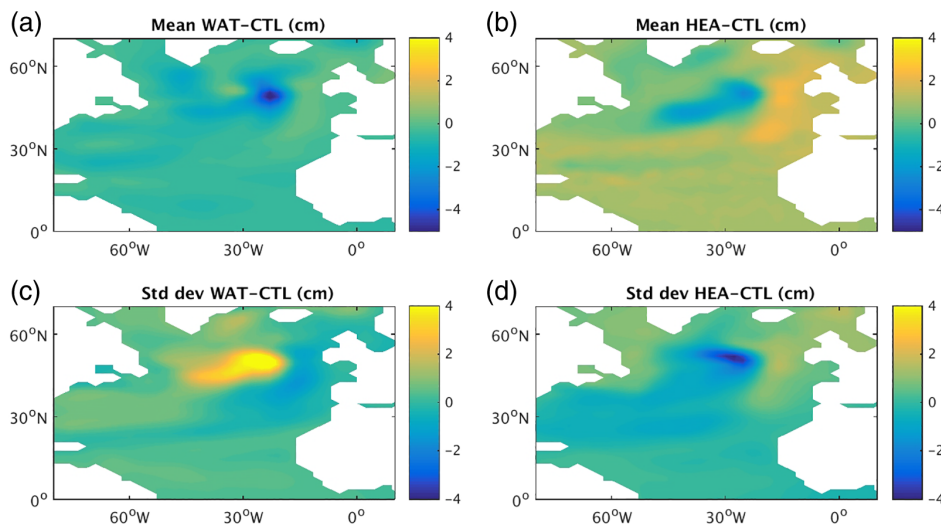
FIGURE 3 Maps of the changes in (a,b) the annual-mean SSH [cm] and (c,d) the standard deviation of annual-mean SSH [cm] in the North Atlantic ocean, calculated from the simulations of Williams (2012). The changes are caused by stochastic perturbations to the air–sea fluxes of (a,c) FW (WAT) and (b,d) heat (HEA). The change refers to the difference in mean and standard deviation between the stochastic simulations and a control simulation [Colour figure can be viewed at wileyonlinelibrary.com]

we revisit the experiments of Williams (2012), in which two crude stochastic parametrizations of air–sea fluxes were implemented. We will examine the impacts on sea-surface height (SSH). The numerical experiments were performed using a coupled AOGCM, consisting of ECHAM4.6 coupled to OPA8.2 (see Table 1). In the first simulation (WAT), the deterministically calculated net FW flux across the air–sea interface is modified stochastically before being passed to the ocean. The second simulation (HEA) is the same, except that the deterministically calculated net heat flux is modified stochastically instead. The two simulations are compared with a control simulation (CTL). The stochastic modifications are achieved through multiplication of the deterministically cal-culated fluxes by spatially uncorrelated white noise drawn from a uniform distribution between 0.5 and 1.5. Each simu-lation is 100 years long and initiated from observations.

The impacts of the stochastic noise on the mean SSH and the SSH inter-annual variability in the North Atlantic are shown in Figure 3. In the central North Atlantic, the stochas-tic noise impacts both the annual-mean SSH and the standard deviation of annual-mean SSH. The changes in each case have magnitudes of up to around 4 cm. These changes demon-strate (a) the sensitivity of mean SSH and SSH variability to stochastic air–sea fluxes, and (b) that the sign and pattern of the response depend on whether the stochastic forcing is applied to the FW flux or heat flux.

3.2 | Uncertainty from subgrid ocean processes

Stochastic forcing as used above is a promising approach to simulate uncertainties resulting from unresolved pro-cesses/scales (Palmer, 2012). One can generate stochastic processes with the same statistics as the unresolved pro-cesses, so that the effect of every particular instance of the

stochastic processes can be computed explicitly and applied as a correction to the nonlinear terms of the model equations. The difficulties with this approach are the following: (a) to obtain a reliable statistical model for unresolved processes, (b) to include adequate dependences for all processes resolved explicitly by the model, and (c) to tune the remaining free parameters.

For instance, Brankart (2013) proposed simulating the effect of unresolved temperature and salinity fluctuations (ΔT and ΔS) in the equation of state $\rho(T, S)$ using a first-order dependence of the fluctuations with respect to the large-scale gradient, such that

$$\overline{\rho} = \frac{1}{p} \sum_{i=1}^{p} \rho \left[\overline{T} + \Delta T_i, \overline{S} + \Delta S_i \right] \qquad (1)$$

with

$$\sum_{i=1}^{p} \Delta T_i = 0, \qquad \sum_{i=1}^{p} \Delta S_i = 0, \qquad (2)$$

where $\overline{\rho}$, $\overline{T}$, and $\overline{S}$ are the large-scale density, temperature, and salinity, respectively, and

$$\Delta T_i = \xi_i(t) \cdot \nabla \overline{T}, \qquad \Delta S_i = \xi_i(t) \cdot \nabla \overline{S}. \qquad (3)$$

The stochastic vectors $\xi_i(t)$ are the same for temperature and salinity, which corresponds roughly to sampling T and S per-turbations along random walks in the neighbourhood of every model grid point. The free parameters, which need to be tuned, are the number p of random walks and the statistics (variance, time correlation, space correlation) of the random walks $\xi_i(t)$.

To evaluate the impact of uncertainties in the horizontal density gradient resulting from unresolved scales, we com-pare results obtained at low resolution with NEMO-ORCA2, a 2°-resolution global ocean model (as in Brankart, 2013), and at eddy-permitting resolution from NEMO-NATL025

TABLE 2 Statistical parameters defining the random walks in Equation 3. Length-scales are in model grid points, ϕ is latitude

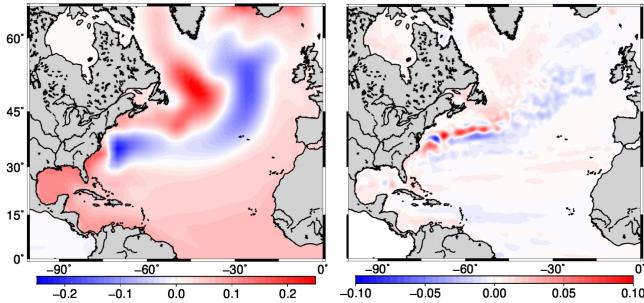|	ORCA2	NATL025
Number of random walks	$p = 6$	$p = 1$
Horizontal standard deviation	$\ell_x = \ell_y = 4.2\|\sin\phi\|$	$\ell_x = \ell_y = \|\sin 2\phi\|$
Vertical standard deviation	$\ell_z = \|\sin\phi\|$	$\ell_z = 0.5\|\sin 2\phi\|$
Correlation time-scale	$\tau = 12$ days	$\tau = 10$ days

FIGURE 4 Impact of uncertainties in the horizontal density gradient on the mean SSH [m] in the North Atlantic. Difference between model simulations with stochastic parametrization of the equation of state and a reference simulation, for (a) ORCA2, at low resolution, and (b) NATL025, at eddy-permitting resolution. The reference simulation for ORCA2 is a climatological run, while the reference run for NATL025 uses stochastic perturbations of lateral diffusion. Note the different colour bars [Colour figure can be viewed at wileyonlinelibrary.com]

(described above in section 2.2.2). Climatological surface forcing is used for ORCA2 and ERA40 is used for NATL025. The statistics of the random walks in Equation 3 are tuned using a reanalysis at 1/4° resolution for implementation into ORCA2 simulations, and model output at 1/20° resolution for NATL025 simulations (see the parameters in Table 2). For ORCA2, the settings are the same as in Brankart (2013), while for NATL025 the amplitude of the fluctuations could not be chosen as large as it should be to fit the diagnosed statistics of the high-resolution model, due to numerical stability during implementation. The variance of the fluctuations has thus been set to the largest possible value that keeps the model stable, with a time step divided by four with respect to the standard NATL025 configuration. It can thus be expected that the impact obtained and illustrated below for these simulations underestimates the real impact of the unresolved scales.

Simulations with ORCA2 are 25 year experiments, with one single member, performed with or without explicit simulation of uncertainties in the computation of density. Simulations with NATL025 are 15 year with 12 member ensemble simulations. The first member of the ensemble is produced using only small perturbations of lateral diffusion, to be used as a reference, and the other members also include the explicit simulation of uncertainties in the computation of density. For all experiments with ORCA2 and NATL025, time averages are performed over the last 10 years of the simulations.

Figure 4 compares the mean effect of a stochastically perturbed density parametrization on the SSH in ORCA2 and NATL025. This result shows that the magnitude of the mean SSH difference is smaller in NATL025 than in ORCA2 by a factor of about 3. The increased resolution of NATL025 compared with ORCA2 has reduced the variance of unresolved temperature and salinity fluctuations, therefore reducing the associated uncertainties in the evaluation of density. However, note that the applied stochastic perturbations are reduced compared with their true estimates due to numerical stability (as described above).

Despite the weaker effect of the unresolved stochastic perturbations in the NATL025 experiments compared with ORCA2, the perturbations still lead to a substantial change in the SSH gradient across the Gulf Stream, therefore accelerating the current and reducing the model bias significantly. Figure 4 shows similar patterns in the correction obtained in ORCA2 and NATL025 (despite the different scales): mainly positive along the northern flank of the Gulf Stream front (except for a negative spot north of Cape Hatteras), and negative along the southern flank of the front. Therefore the representation of unresolved temperature and salinity fluctuations leads to perturbations (for example, smaller eddies) that impact the large-scale density across the front, as expected from an inverse energy cascade argument, and cannot be neglected.

Figure 5 shows the impact of this stochastic parametrization on the interannual variability of the system in ORCA2. The time standard deviation of SSH is displayed with (right panel) and without (left) explicit simulation of uncertainties in the horizontal density gradient. The large-scale SSH variability from ORCA2 is only increased in regions with strong mesoscale activity, such as the Gulf Stream, Kuroshio, Aghulas current, confluence region, and Antarctic circumpolar current with amplitude of about 10 cm, similarly to Andrejczuk *et al.* (2016) and Juricke *et al.* (2017). The perturbations produced by stochastic subgrid-scale parametrizations may lead to an upscale cascade generating large-scale and low-frequency variability through nonlinearities and scale interactions.

Figure 6 illustrates the impact of this parametrization on the SSH ensemble spread at eddy-permitting resolution in NATL025. Without explicit simulation of density uncertainties (top panels) of a few centimetres, small perturbations of lateral diffusion can only generate a very small spread after 3 months (left panel), which is localized around the Gulf Stream separation point from the American coastline. The spread is then slowly amplified and propagated by the chaotic mesoscale flow, as shown for the SSH spread after 1 year and after 15 years (middle and left panels, respectively), reaching 20 cm in large regions.

With explicit representation of density uncertainties (bottom panels), the initial rate of spread increases much faster than for the reference ensemble (compare top and bottom left panels). At the beginning of the experiment when the spread
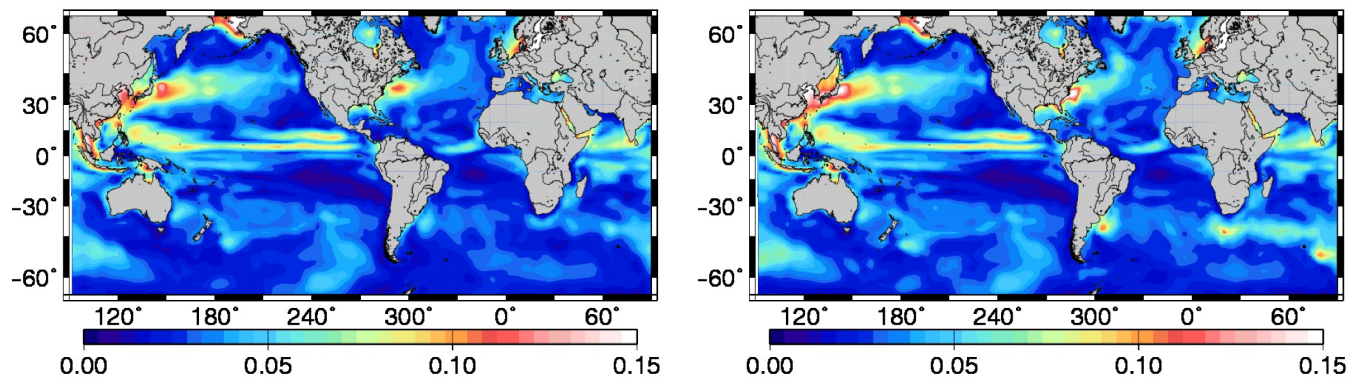
FIGURE 5 Standard deviation of SSH [m] in ORCA2, obtained (a) without and (b) with explicit simulation of uncertainties in the horizontal density gradient [Colour figure can be viewed at wileyonlinelibrary.com]
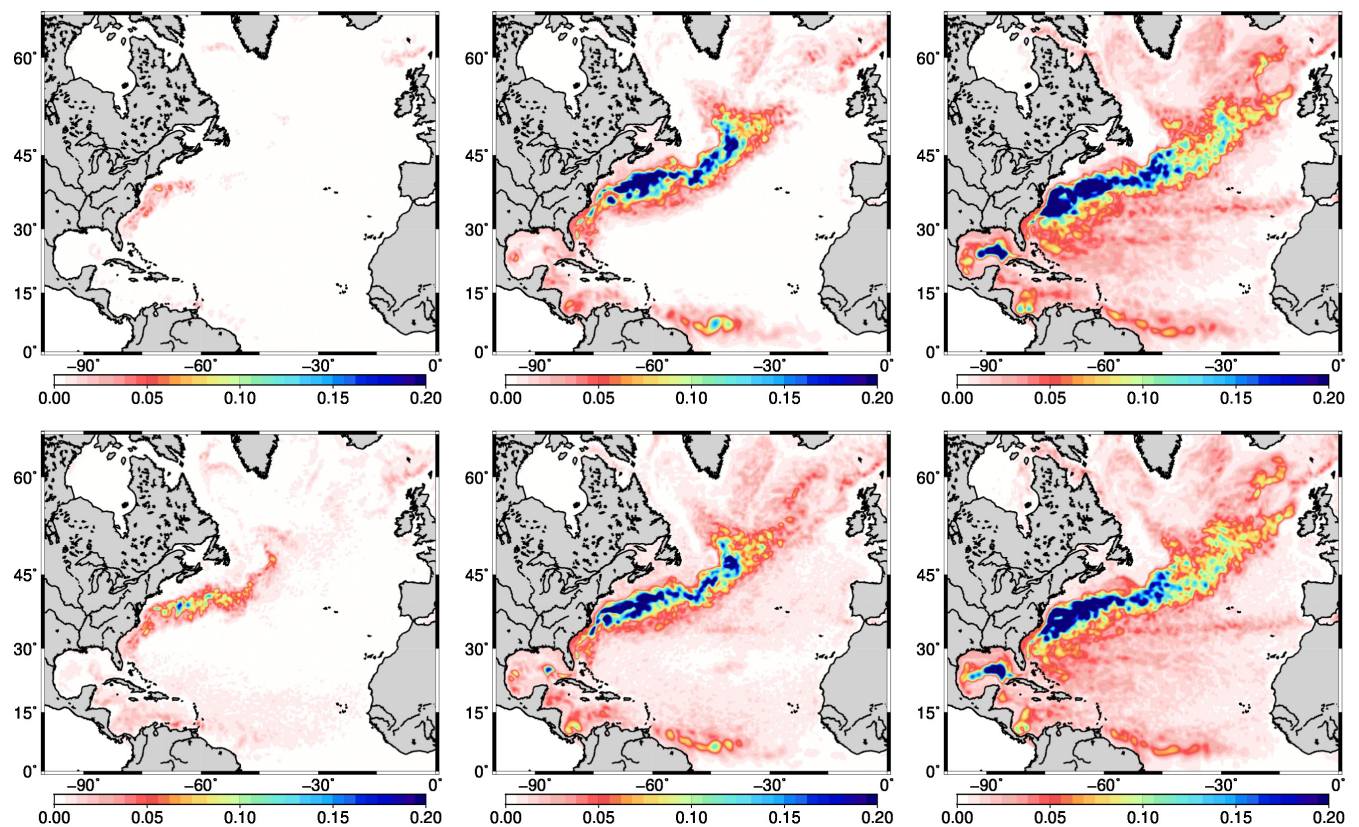

FIGURE 6 Temporal evolution of the standard deviation of SSH [m] in NATL025: after (a,d) 3 months, (b,e) 1 year, and (c,f) 15 years. In the top panels, the model only includes small perturbations of lateral diffusion (our reference state); in the bottom panels, the model also includes both small perturbations of lateral diffusion and explicit representation of uncertainties in the horizontal density gradient [Colour figure can be viewed at wileyonlinelibrary.com]

is still small, the dynamics of the ensemble spread is dominated by the effect of stochastic perturbations in the equation of state (about 10 cm in some regions). After a few months, however, when the spread is larger, the stochastic perturbations lose their dominant role in the dynamics of the spread, to be replaced by the chaotic dynamics of the mesoscale, which is mostly responsible for the subsequent increase of the spread and for its propagation in other regions of the Atlantic. After 1 year (middle panels), the spread is only slightly more developed with the stochastic perturbations; after 15 years (right panels), no significant difference in the spread can be

observed in the figure (even if a difference necessarily exists, since the mean flow is different, as shown in Figure 4).

This example illustrates that a fair approximation of the asymptotic behaviour of the spread, characterizing the attractor of the system, does not guarantee that the same model will correctly represent transient ocean dynamics effects, which are a prerequisite to producing reliable short-term or seasonal ensemble forecasts. Therefore, even at eddy-permitting resolution, the need for representing ocean model uncertainty is crucial for adequate ensemble seasonal forecasts, but also for bias reduction of the climatological state.
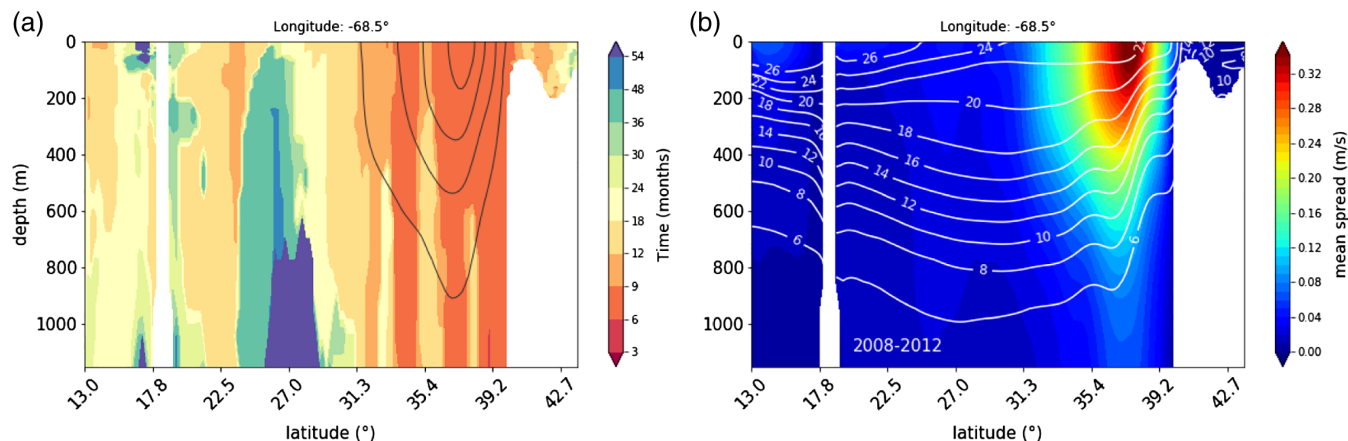
FIGURE 7 (a) Time (in months) over which the ensemble spread grows exponentially in the OCCIPUT ensemble simulation, and (b) time average of the velocity ensemble spread (in m/s) towards the later part of the OCCIPUT ensemble simulation for 2008–2012 (that is, long after the initial growing phase). The sections are taken along 68.5°W longitude as a function of depth and latitude. Contours indicate the ensemble and time average of velocity (left) and temperature (right) [Colour figure can be viewed at wileyonlinelibrary.com]

3.3 | Impacts of initial condition uncertainties on climate-relevant oceanic variables

The impact of initial conditions and transient eddies on large-scale low-frequency signals is now evaluated using the OCCIPUT ensemble, described in section 2.2.3 and based on a global realistic ocean model. To generate different IC among the different members for the OCCIPUT global ensemble, stochastic perturbations in the equation of state (as explained in section 3.2) were activated during the first year (that is, 1960), then switched off until the end of the integration in 2015.

In OCCIPUT, we define two ensemble statistics to provide a simple way to disentangle the chaotic and forced parts of the oceanic variability. The ensemble mean of any simulated quantity at any location captures the part of the oceanic variability that is atmospherically driven and identical within all members. The ensemble spread, measured by the ensemble standard deviation (eSTD), is associated with chaotic fluctuations superimposed on the atmospherically driven ensemble mean, the phases of which differ among the members.

Intermember differences become noticeable after a few days or weeks in the OCCIPUT ensemble at 1/4° resolution. At that time, the ensemble spread enters a phase of exponential growth controlled by barotropic and baroclinic instabilities, which feed the meandering of unstable currents, the growth of mesoscale eddies, and, eventually their shedding from the jets; this makes the ensemble members diverge over the first year (as shown in Figure 6 for the NATL025 ensemble).

Figure 7a shows the duration of this initial phase along a longitudinal section across the Gulf Stream (also illustrated in figure 4 of Bessieres *et al.*, 2017). The eSTD exponential growing phase lasts about 6–9 months in the Gulf Stream down to 1000 m depth (and in other regions with strong currents, such as the Kuroshio and Antarctic Circumpolar Current (ACC)), where strong velocity shears favour the emergence of mesoscale turbulence through instabilities. The eSTD increases much more slowly in quiescent regions, that is, within a few years.

The ensemble spread then enters a second phase, where it saturates and fluctuates around an equilibrium value. Figure 7b shows the eSTD of monthly ocean velocities, along the same section, averaged over several years during this second saturated phase. This represents the root-mean-square of simulated velocities, computed across the ensemble dimension instead of the temporal dimension as is often done. The surface-intensified eddy activity in the Gulf Stream[1] can fluctuate temporally in the saturated phase. These fluctuations are statistically insignificant for the annually averaged AMOC (Leroux *et al.*, 2018), but subannual fluctuations of 5-day-mean SSH eSTD are significant in certain areas, such as upwelling regions, where they reflect the seasonal emergence and subsequent westward propagation of mesoscale activity (not shown).

Figure 8 (left panel) shows that, during the first months or years, the chaotic variability responsible for the ensemble spread in OHC consists mostly of mesoscale structures, with relatively small spatial scales. The scales of the chaotic variability, however, evolve slowly. The right panel in Figure 8 shows that, after a few decades of integration, mesoscale structures still dominate in the eddy-active ACC, but much larger-scale, yet chaotic, fluctuations have populated the midlatitudes. This propensity of mesoscale structures to interact spontaneously and create larger-scale structures is known as the spatial inverse cascade of kinetic energy (for example, Batchelor, 1953). Arbic *et al.* (2012; 2014) showed further that mesoscale variability also feeds lower-frequency chaotic fluctuations through a temporal inverse cascade of kinetic energy. Serazin *et al.* (2018) confirmed that both spatial and

---

[1]Note that eSTDs and other ensemble statistics may be computed at any time from such experiments, providing instantaneous estimates of the turbulent activity.
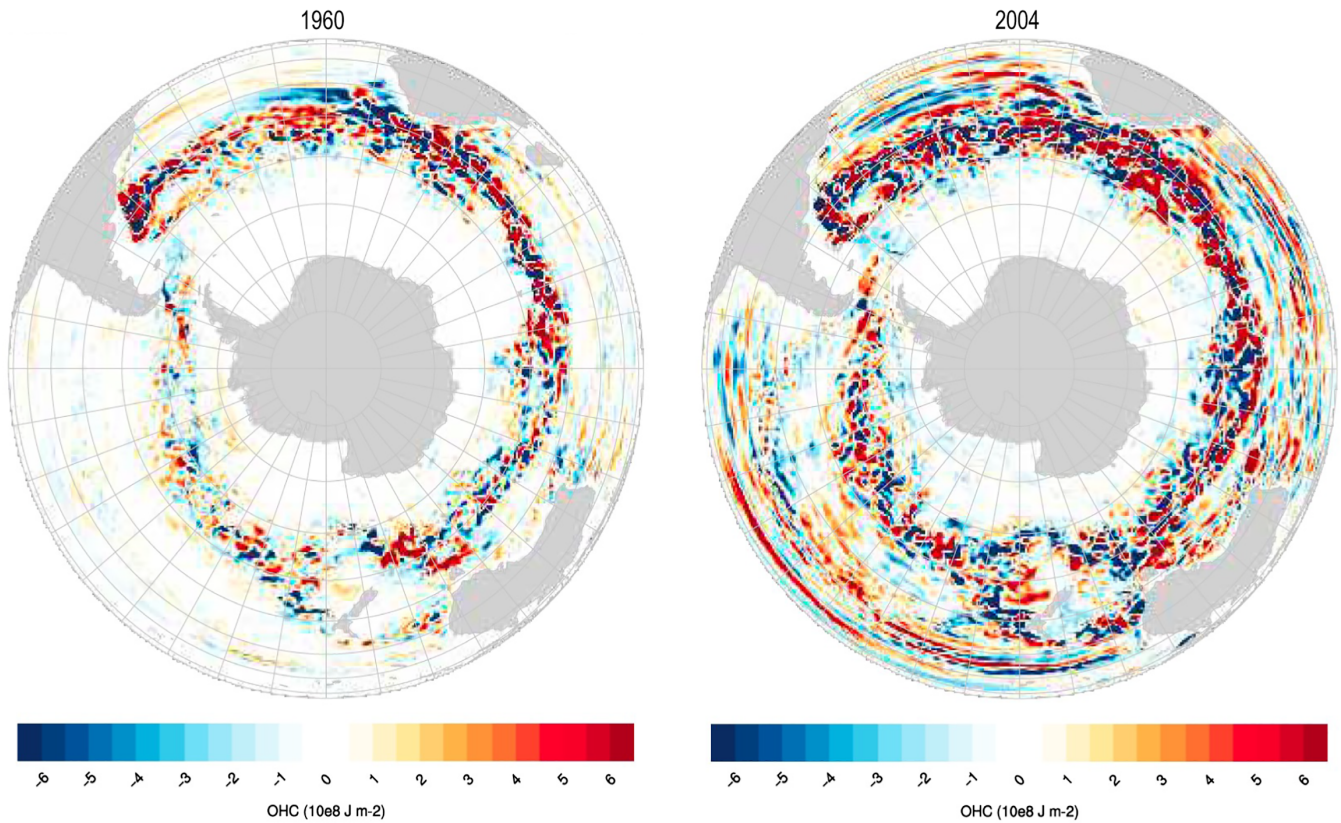
FIGURE 8 Chaotic anomalies (with respect to the ensemble mean) of yearly $OHC_{0-800\,m}$ in member 5 of the ensemble, (a) during the first year of the OCCIPUT simulation (1960), and (b) after 44 years of integration (2004) [Colour figure can be viewed at wileyonlinelibrary.com]
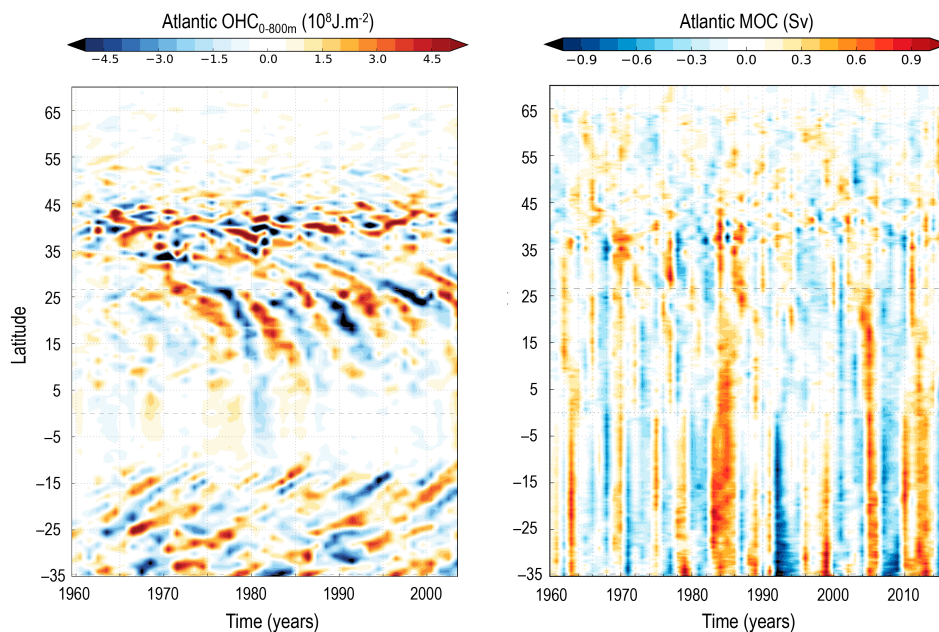

FIGURE 9 Latitude–time Hovmöller diagrams of interannual-to-decadal chaotic anomalies of (a) the Atlantic OHC and (b) Atlantic MOC in member 5 of the OCCIPUT global ensemble simulation [Colour figure can be viewed at wileyonlinelibrary.com]

temporal cascades are indeed at work, and this may potentially explain the progressive emergence of slower, larger-scale chaotic variability. Note that large-scale baroclinic instability processes, distinct from those occurring at the mesoscale, may coexist with mesoscale instabilities (Huck *et al.*, 2015), feed multidecadal ocean chaotic variability (Sevellec and Fedorov,

2013), and potentially contribute to the long-term evolution of the ensemble spread as well.

Figure 9 shows the Hovmöller (latitude–time) diagrams of large-scale, low-frequency chaotic anomalies for OHC (left) and MOC (right) in the Atlantic, built from annual averages. Their phases differ among the OCCIPUT members and are
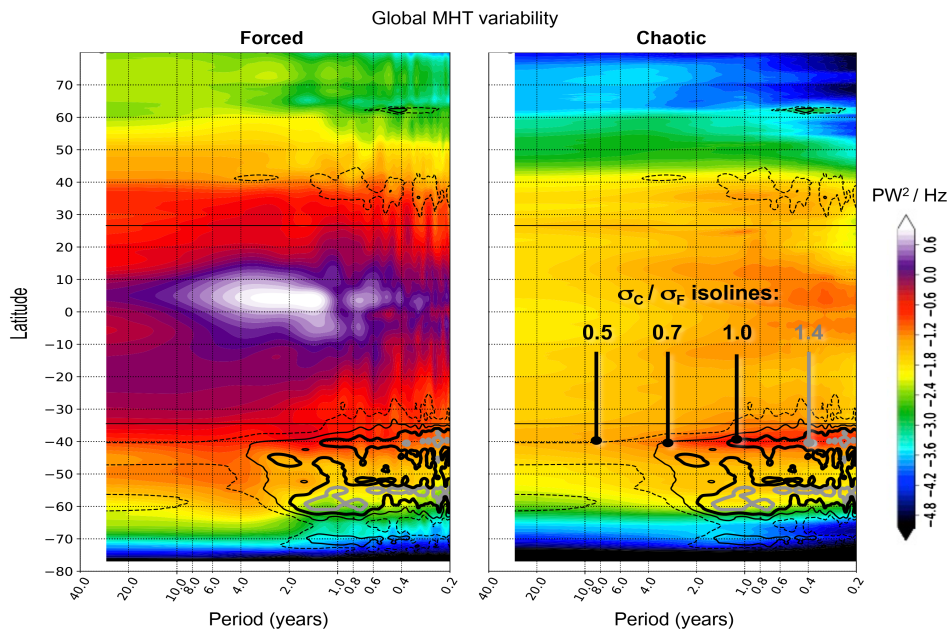
FIGURE 10 Power spectral density of (a) forced (σ_F^2) and (b) chaotic (σ_C^2) variabilities of the global MHT, as a function of period and latitude diagnosed from the OCCIPUT global ensemble simulation. Contours in these two panels indicate selected isolines of the σ_C/σ_F ratio [Colour figure can be viewed at wileyonlinelibrary.com]

independent of the prescribed atmospheric variability. The upper OHC chaotic variability in the Atlantic locally exceeds that of the atmospherically forced variability in the midlatitudes (as in the Southern Ocean, Serazin *et al.*, 2017). The chaotic OHC anomalies (left panel) in the subtropical regions tend to propagate westward and equatorward at phase speeds close to those of Rossby waves. In the midlatitudes, between 33°N and 46°N, there is no indication of propagation, but there is a hint of persistence of anomalies over a few years, despite the noisy fields. AMOC (right panel) exhibits a different structure: large, meridionally coherent, low-frequency chaotic anomalies modulate the AMOC at the basin scale, with intensities reaching their atmospherically forced counterpart around 30°S (Leroux *et al.*, 2018). Leroux *et al.* (2018) and Serazin *et al.* (2017) show that the time-scale of these chaotic signals reaches several decades.

Do the chaotic variabilities of OHC and meridional circulations impact the redistribution of heat at the global scale? Figure 10 compares the frequency spectra of the forced (atmospherically driven) and chaotic variabilities of the meridional heat transport (MHT) integrated globally along latitudes. The absence of contours between 30°S and 30°N indicates that the atmosphere explains most of the global MHT multiscale variability at low latitudes. Around the latitudes of the Gulf Stream and Kuroshio extensions, however, the dashed contours show that the chaotic variability exceeds half of its forced counterpart on time-scales shorter than about 2 years. In the Southern Ocean, this chaotic-to-forced variability fraction exceeds 50% on time-scales up to 2–4 years everywhere between 35 and 72°S, and up to multidecadal time-scales between 40 and 48°S. Around 40°S and 55°S, the global MHT variability is primarily chaotic on time-scales shorter than 1–2 years.

The OCCIPUT ensemble simulation thus shows that, in the presence of mesoscale turbulence, small IC uncertainties evolve within a few decades from the scale of eddies to large spatial and temporal scales. The interannual to multidecadal variability of large-scale climate-relevant oceanic indices in the turbulent regime is not only due to the atmospheric variability or the air–sea coupling: a substantial fraction of it may emerge spontaneously from oceanic nonlinearities with chaotic evolution and a complex spatial structure. Studies are under way to assess the potential impacts of this phenomenon on the other components of the climate system (atmosphere, cryosphere) and on the detection and attribution of climate signals in the real ocean.

4 | CONCLUSIONS

As computational resources increase, our models are increasing in their fidelity with respect to the ocean and climate system. However, models are still only approximate representations of the real world and do not resolve many key processes such as ocean eddies, air–sea fluxes, and upper ocean mixing. Therefore we still rely heavily on parametrizations of these essential, often nonlinear, processes. The chaotic behaviour of these processes when partially resolved, or the error associated with their inaccurate parametrization, leads to uncertainty in the representation of the ocean state and its variability on a wide range of spatio-temporal scales. To provide reliable ocean and climate projections, we must quantify the uncertainty associated with these processes and their representation in models, using ensemble simulations.

In this article, we present a range of new ocean ensembles by targeting time-scales from weeks to decades, using different horizontal resolutions, different surface forcings, and different quantification of model uncertainty. Based on an ocean ensemble at coarse resolution, forced by a range of possible air–sea fields taken from the CMIP5 ensemble, we find a large uncertainty in AMOC weakening projections (between 0 and 5 Sv), due mainly to uncertainties in buoyancy fluxes (90% versus 10% from wind forcing). We also find large uncertainties in AABW projections (between −1 Sv and 5 Sv), driven equally by both wind and buoyancy uncertainty. The uncertainty in projected air–sea forcing translates into an uncertainty in the amplitude and phase of decadal ocean variability and multidecadal projections. The different air–sea fluxes (heat, freshwater, and momentum) have a distinct impact on the global ocean heat uptake and its regional patterns. In the Atlantic, FW and heat fluxes are the largest contributors to regional OHC changes, including via ocean stratification and circulation changes. The spread (uncertainty) in regional OHC in this ensemble is due mainly to the difference in FW and heat fluxes—especially in the Southern Ocean. As shown in Huber and Zanna (2017), the uncertainty in air–sea forcing is of order one in regional predictions of OHU and often more important than parametric uncertainty. The dominance of air–sea flux uncertainty is not surprising in coarse-resolution models, where the stratification and transformation of water masses are strongly influenced by the surface forcing. In higher resolution models, where eddies and their effect on the large-scale flow are partially resolved, the sensitivity of ocean properties might be reduced, as the stratification and its adjustment will depend strongly on eddy and mixing processes (for example, Hallberg and Gnanadesikan, 2006). Therefore, further experiments with differing ocean resolution and spread of surface forcings should be conducted to understand the uncertainty due to eddy mixing and air–sea fluxes on ocean circulation and associated regional heat content change, and ultimately their impact on the atmospheric circulation.

Even as the resolution of the ocean model component increases, subgrid-scale parametrizations are needed. We have shown that representing subgrid-scale uncertainty using stochastic physics gives us a quantification of model error associated with the underlying subgrid processes in ocean ensembles (for example, Berner et al., 2017). Due to the nonlinearity of the model physics, the inclusion of stochastic physics, especially when based on high-resolution data or observations, can lead to significant improvements in the model state, reducing mean biases (for example, Brankart, 2013). For example, we showed that, at both 2° and 1/4° horizontal resolution, the representation of unresolved temperature and salinity fluctuations leads to perturbations that impact the large-scale density field across the Gulf Stream front. Stochastic physics can also improve the representation of ocean variability on seasonal to decadal time-scales and generate spread in ocean ensembles on different scales.

The dynamics of the ensemble spread at eddy-permitting resolution is dominated by the effect of the stochastic perturbations in the equation of state (an impact of about 10 cm on sea level in some regions); this spread can persist for months and can generate interactions with the large-scale flow. The growing phase of the spread in the region of strong current is fed by strong velocity shears, which favour the emergence of mesoscale turbulence through instabilities. Despite this, the impact of stochastic perturbations and their evolution depends strongly on the model horizontal resolution, the time-scale considered, and the atmospheric forcing variability (for example, Williams, 2012; Brankart, 2013; Brankart et al., 2015; Andrejczuk et al., 2016; Williams et al., 2016; Juricke et al., 2017). There are differences in the impact of stochastic perturbations on the flow for different model resolutions, and the impact is altered in the presence of high-frequency atmospheric forcing. From the state-of-the-art OCCIPUT ensemble (Penduff et al., 2014), there is a clear chaotic variability signal, which overwhelms the directly (atmospheric) forced response. For example, around the latitudes of the Gulf Stream and Kuroshio extensions, the chaotic variability in global MHT exceeds half of its forced counterpart on time-scales shorter than about 2 years. In the Southern Ocean, this chaotic-to-forced variability fraction exceeds 50% on time-scales up to 2–4 years everywhere between 35 and 72°S, and up to multidecadal time-scales between 40 and 48°S. This chaotic variability, via nonlinear scale interactions, can therefore give rise to large-scale and low-frequency variability in ocean volume transport, heat content, and heat transport (for example, Leroux et al., 2018; Penduff et al., 2018).

From the range of experiments we have presented, which identify air–sea fluxes, eddy mixing, and jet dynamics as leading sources of uncertainty over several spatio-temporal scales, there is a need to represent these processes and associated uncertainties adequately in models for seasonal forecasts to multidecadal projections. Understanding how eddies impact the large-scale and low-frequency variability is still vastly unexplored. Even at 1/4°, eddies are only partially resolved; therefore the need for parametrization is present. There are several avenues based on the development of scale-aware parametrizations to represent some of the scale interactions (ocean–ocean and ocean–atmosphere and short–long time-scales) described above. Empirical parametrizations have been developed using idealized ocean models (Berloff, 2005; Cooper and Zanna, 2015), while other parametrizations are targeting specific processes misrepresented in coarser resolution models, such as eddy saturation (Mak et al., 2017), eddy variance (Grooms, 2016), energy transfer (Grooms and Majda, 2013; Jansen et al., 2015; Bachman et al., 2017), and upgradient momentum fluxes (Porta Mana and Zanna, 2014; Zanna et al., 2017). New avenues for parametrization could include ensemble superparametrization for vertical mixing, as proposed in atmospheric models (Subramanian and Palmer, 2017). Eddy-permitting or eddy-resolving ensemble ocean simulations are nevertheless mandatory to

assess how these parametrizations mimic eddy interactions properly, and their inverse cascade to climate-relevant variabilities (Penduff *et al.*, 2014; Serazin *et al.*, 2017; 2018).

Representation of missing processes and uncertainty using stochastic physics and parametrizations has proven useful across a range of ensembles. However, it is important to design these parametrizations using statistics that are representative of the missing processes or missing error. There is a need to diagnose the subgrid statistics using observations, higher resolution simulations using coarse-graining, or ensemble data assimilation that combines model and observations, to link the physics of the model to the observed physics (Palmer *et al.*, 2005). With the emergence of machine learning tools, designing parametrizations or building relationships between subgrid and large-scale variables will likely become more efficient.

The different ensembles presented here have highlighted that the nonlinearity and chaotic behaviour of the ocean can be large. The relative importance of the atmospheric versus oceanic chaotic variability remains to be quantified, in order to understand fundamental oceanographic processes and their possible imprints on the climate system. The different ensembles have so far been designed to explore only one aspect: coarse-resolution models with different sources of uncertainty (atmospheric uncertainty, ocean IC, and model uncertainty), or models at an eddy-permitting resolution with only ocean uncertainty. To investigate the relative contribution of atmospheric and oceanic uncertainty, ensemble experiments at eddy-permitting (or resolving) resolution, which allow for rich nonlinear scale interactions, should be employed with different ocean model and initial condition uncertainty representations, together with a representation of atmospheric uncertainty (for example, using singular vectors for initial and boundary condition uncertainty and stochastic physics for model uncertainty). The number of ensemble members remains to be determined and will likely be an important factor in interpreting any result. However, we will be able to move towards a probabilistic view of understanding time-dependent ocean variability and prediction, by estimating the relative uncertainty as a function of resolution and time-scales for a range of processes.

ACKNOWLEDGEMENTS

REFERENCES

Allen, M.R. and Stainforth, D.A. (2002) Towards objective probabilistic climate forecasting. *Nature*, 419, 228.

Allen, M.R., Kettleborough, J.A. and Stainforth, D.A. (2002) Model error in weather and climate forecasting *ECMWF Predictability of Weather and Climate Seminar*. Reading, UK: European Centre for Medium Range Weather Forecasts, 275–294.

Andrejczuk, M., Cooper, F.C., Juricke, S., Palmer, T.N., Weisheimer, A. and Zanna, L. (2016) Oceanic stochastic parametrizations in a seasonal forecast system. *Monthly Weather Review*, 144(5), 1867–1875.

Arbic, B.K., Flierl, G.R. and Scott, R.B. (2007) Cascade inequalities for forced-dissipated geostrophic turbulence. *Journal of Physical Oceanography*, 37(6), 1470–1487.

Arbic, B.K., Scott, R.B., Flierl, G.R., Morten, A.J., Richman, J.G. and Shriver, J.F. (2012) Nonlinear cascades of surface oceanic geostrophic kinetic energy in the frequency domain. *Journal of Physical Oceanography*, 42, 1577–1600.

Arbic, B.K., Müller, M., Richman, J.G., Shriver, J.F., Morten, A.J., Scott, R.B., Serazin, G. and Penduff, T. (2014) Geostrophic turbulence in the frequency-wavenumber domain: Eddy-driven low-frequency variability. *Journal of Physical Oceanography*, 44, 2050–2069.

Bachman, S.D., Fox-Kemper, B. and Pearson, B. (2017) A scale-aware subgrid model for quasi-geostrophic turbulence. *Journal of Geophysical Research: Oceans*, 122(2), 1529–1554.

Barnier, B., Madec, G., Penduff, T., Molines, J.-M., Treguier, A.-M., Le Sommer, J., Beckmann, A., Biastoch, A., Böning, C., Dengg, J., Derval, C., Durand, E., Gulev, S., Remy, E., Talandier, C., Theetten, S., Maltrud, M., McClean, J. and DeCuevas, B. (2006) Impact of partial steps and momentum advection schemes in a global ocean circulation model at eddy permitting resolution. *Ocean Dynamics*, 56, 543–567.

Batchelor, G.K. (1953) *The Theory of Homogeneous Turbulence*, Vol. 79. Cambridge: Cambridge University Press.

Berloff, P.S. (2005) Random-forcing model of the mesoscale oceanic eddies. *Journal of Fluid Mechanics*, 529, 71–95.

Berloff, P.S., Hogg, A.M. and Dewar, W. (2007) The turbulent oscillator: a mechanism of low-frequency variability of the wind-driven ocean gyres. *Journal of Physical Oceanography*, 37, 2363–2386.

Berner, J., Achatz, U., Batt, L., Bengtsson, L., Cmara, A.D., Christensen, H.M., Colangeli, M., Coleman, D.R., Crommelin, D., Dolaptchiev, S.I., Franzke, C.L., Friederichs, P., Imkeller, P., Jrvinen, H., Juricke, S., Kitsios, V., Lott, F., Lucarini, V., Mahajan, S., Palmer, T.N., Penland, C., Sakradzija, M., von Storch, J., Weisheimer, A., Weniger, M., Williams, P.D. and Yano, J. (2017) Stochastic parametrization: toward a new view of weather and climate models. *Bulletin of the American Meteorological Society*, 98, 565–588.

Bessieres, L., Leroux, S., Brankart, J.-M., Molines, J.-M., Moine, M.-P., Bouttier, P.-A., Penduff, T., Terray, L., Barnier, B. and Serazin, G. (2017) Development of a probabilistic ocean modelling system based on NEMO 3.5: application at eddying resolution. *Geoscientific Model Development*, 10, 1091–1106.

Brankart, J.-M. (2013) Impact of uncertainties in the horizontal density gradient upon low resolution global ocean modelling. *Ocean Modeling*, 66, 64–76.

Brankart, J.-M., Candille, G., Garnier, F., Calone, C., Melet, A., Bouttier, P.-A., Brasseur, P. and Verron, J. (2015) A generic approach to explicit simulation of uncertainty in the NEMO ocean model. *Geoscientific Model Development*, 8, 1285–1297.

Buizza, R., Milleer, M. and Palmer, T.N. (1999) Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, 125, 2887–2908.

Candille, G., Brankart, J.-M. and Brasseur, P. (2015) Assessment of an ensemble system that assimilates Jason-1/Envisat altimeter data in a probabilistic model of the North Atlantic ocean circulation. *Ocean Science*, 11, 425–438.

Cooper, F. and Zanna, L. (2015) Optimisation of an idealized ocean model: stochastic parametrisation of sud-grid eddies. *Ocean Modelling*, 88, 38–53.

DeMott, C.A., Klingaman, N.P. and Woolnough, S.J. (2015) Atmosphere-ocean coupled processes in the Madden–Julian oscillation. *Reviews of Geophysics*, 53, 1099–1154.

Dewar, W.K. (2003) Nonlinear midlatitude ocean adjustment. *Journal of Physical Oceanography*, 33, 1057–1082.

Dijkstra, H.A. and Ghil, M. (2005) Low-frequency variability of the large- scale ocean circulation: a dynamical systems approach. *Reviews of Geophysics*, 43.

Dong, B.W. and Sutton, R.T. (2002) Adjustment of the coupled ocean–atmosphere system to a sudden change in the thermohaline circulation. *Geophysical Research Letters*, 29(15).

Fox-Kemper, B., Danabasoglu, G., Ferrari, R. and Hallberg, R.W. (2008) Parameterizing submesoscale physics in global climate models. *CLIVAR Exchanges*, 13(1), 3–5.

Frankignoul, C. and Hasselmann, K. (1977) Stochastic climate models, Part II Application to sea-surface temperature anomalies and thermocline variability. *Tellus*, 29(4), 289–305.

Germe, A., Sevellec, F., Mignot, J., Swingedouw, D. and Nguyen, S. (2017a) On the robustness of near term climate predictability regarding initial state uncertainties. *Climate Dynamics*, 48, 353–366.

Germe, A., Sevellec, F., Mignot, J., Fedorov, A., Nguyen, S. and Swingedouw, D. (2017b) The impacts of oceanic deep temperature perturbations in the North Atlantic on decadal climate variability and predictability. *Climate Dynamics*, 51, 2341–2357.

Gibson, A.H., Hogg, A.M., Kiss, A.E., Shakespeare, C.J. and Adcroft, A. (2017) Attribution of horizontal and vertical contributions to spurious mixing in an Arbitrary Lagrangian–Eulerian ocean model. *Ocean Modelling*, 119, 45–56.

Gregorio, S., Penduff, T., Serazin, G., Molines, J.-M., Barnier, B. and Hirschi, J. (2015) Intrinsic variability of the Atlantic meridional overturning circulation at interannual-to-multidecadal time-scales. *Journal of Physical Oceanography*, 45(7), 1929–1946.

Gregory, J.M., Bouttes, N., Griffies, S.M., Haak, H., Hurlin, W.J., Jungclaus, J., Maxwell, K., Lee, W.G., Marshall, J., Romanou, A. and Saenko, O.A. (2016) The Flux-Anomaly-Forced Model Intercomparison Project (FAFMIP) contribution to CMIP6: investigation of sea-level and ocean climate change in response to CO2 forcing. *Geoscientific Model Development*, 9(11), 3993.

Griffies, S.M., Winton, M., Anderson, W.G., Benson, R., Delworth, T.L., Dufour, C.O., Dunne, J.P., Goddard, P., Morrison, A.K., Rosati, A., Wittenberg, A.T., Yin, J.J. and Zhang, R. (2015) Impacts on ocean heat from transient mesoscale eddies in a hierarchy of climate models. *Journal of Climate*, 28, 952–977.

Grooms, I. (2016) A Gaussian-product stochastic Gent–McWilliams parametrization. *Ocean Modelling*, 106, 27–43.

Grooms, I. and Majda, A.J. (2013) Efficient stochastic superparametrization for geophysical turbulence. *Proceedings Of the National Academy of Sciences of the United States of America*, 110, 4464–4469.

Hallberg, R.W. and Gnanadesikan, A. (2006) On the role of eddies in determining the structure and response of the wind-driven Southern Hemisphere overturning: results from the Modeling Eddies in the Southern Ocean (MESO) project. *Journal of Physical Oceanography*, 36, 2232–2252.

Hawkins, E. and Sutton, R. (2009) Decadal predictability of the Atlantic Ocean in a coupled GCM: forecast skill and optimal perturbations using linear inverse modeling. *Journal of Climate*, 22(14), 3960–3978.

Huber, M. and Zanna, L. (2017) Drivers of uncertainty in simulated ocean circulation and heat uptake. *Geophysical Research Letters*, 44(3), 1402–1413.

Huck, T., Arzel, O. and Sevellec, F. (2015) Multidecadal variability of the overturning circulation in presence of eddy turbulence. *Journal of Physical Oceanography*, 45(1), 157–173.

Jansen, M.F., Held, I.M., Adcroft, A. and Hallberg, R. (2015) Energy budget-based backscatter in an eddy permitting primitive equation model. *Ocean Modelling*, 94, 15–26.

Juricke, S., Palmer, T.N. and Zanna, L. (2017) Stochastic perturbations to subgrid-scale ocean mixing: impacts on low frequency variability. *Journal of Climate*, 30(13), 4997–5019.

Kharin, V.V. and Zwiers, F.W. (2002) Climate predictions with multimodel ensembles. *Journal of Climate*, 15, 793–799.

Kjellsson, J. and Zanna, L. (2017) The impact of horizontal resolution on energy transfers in global ocean models. *Fluids*, 2(3), 45.

Kleeman, R. and Moore, A.M. (1997) A theory for the limitation of ENSO predictability due to stochastic atmospheric transients. *Journal of the Atmospheric Sciences*, 54(6), 753–767.

Leroux, S., Penduff, T., Bessieres, L., Molines, J., Brankart, J., Serazin, G., Barnier, B. and Terray, L. (2018) Intrinsic and atmospherically forced variability of the AMOC: insights from a large-ensemble ocean hindcast. *Journal of Climate*, 31, 1183–1203.

Leutbecher, M., Lock, S., Ollinaho, P., Lang, S.T., Balsamo, G., Bechtold, P., Bonavita, M., Christensen, H.M., Diamantakis, M., Dutra, E., English, S., Fisher, M., Forbes, R.M., Goddard, J., Haiden, T., Hogan, R.J., Juricke, S., Lawrence, H., MacLeod, D., Magnusson, L., Malardel, S., Massart, S., Sandu, I., Smolarkiewicz, P.K., Subramanian, A., Vitart, F., Wedi, N. and Weisheimer, A. (2017) Stochastic representations of model uncertainties at ECMWF: state of the art and future vision. *Quarterly Journal of the Royal Meteorological Society*, 143, 2315–2339. https://doi.org/10.1002/qj.3094.

Levitus, S., Antonov, J.I., Boyer, T.P., Baranova, O.K., Garcia, H.E., Locarnini, R.A., Mishonov, A.V., Reagan, J.R., Seidov, D., Yarosh, E.S. and Zweng, M.M. (2012) World ocean heat content and thermosteric sea level change (0–2000 m), 1955–2010. *Geophysical Research Letters*, 39, L10603.

Mak, J., Marshall, D.P., Maddison, J.R. and Bachman, S.D. (2017) Emergent eddy saturation from an energy constrained eddy parametrization. *Ocean Modelling*, 112, 125–138.

Marzocchi, A., Hirschi, J.J.-M., Holliday, N.P., Cunningham, S.A., Blaker, A.T. and Coward, A.C. (2015) The North Atlantic subpolar circulation in an eddy-resolving global ocean model. *Journal of Marine Systems*, 142, 126–143.

Murphy, J.M., Sexton, D.M.H., Barnett, D.N., Jones, G.S., Webb, M.J., Collins, M. and Stainforth, D.A. (2004) Quantifying uncertainties in climate change using a large ensemble of global climate model predictions. *Nature*, 430, 768–772.

Palmer, T.N. (2012) Towards the probabilistic earth-system simulator: a vision for the future of climate and weather prediction. *Quarterly Journal of the Royal Meteorological Society*, 138(665), 841–861.

Palmer, T.N. and Zanna, L. (2013) Singular vectors, predictability and ensemble prediction for weather and climate. *Journal of Physics A: Mathematical and Theoretical*, 46(2013), 254018.

Palmer, T.N., Shutts, G.J., Hagedorn, R., Doblas-Reyes, F.J., Jung, T. and Leutbecher, M. (2005) Representing model uncertainty in weather and climate prediction. *Annual Review of Earth and Planetary Sciences*, 33, 163–193.

Pearson, B., Fox-Kemper, B., Bachman, S.D. and Bryan, F.O. (2017) Evaluation of scale-aware subgrid mesoscale eddy models in a global eddy-rich model. *Ocean Modelling*, 115, 42–58.

Peixoto, J.P. and Oort, A.H. (1984) Physics of climate. *Reviews of Modern Physics*, 56, 365–429.

Penduff, T., Juza, M., Barnier, B., Zika, J., Dewar, W.K., Treguier, A.-M., Molines, J.-M. and Audiffren, N. (2011) Sea-level expression of intrinsic and forced ocean variabilities at interannual time-scales. *Journal of Climate*, 24, 5652–5670.

Penduff, T., Barnier, B., Terray, L., Bessieres, L., Serazin, G., Gregorio, S., Brankart, J.-M., Moine, M.-P., Molines, J.-M. and Brasseur, P. (2014) Ensembles of eddying ocean simulations for climate. *CLIVAR Exchanges*, 19(2), 65. Special Issue on High Resolution Ocean Climate Modelling.

Penduff, T.G., Sérazin, S., Leroux, S., Close, J.-M., Molines, B., Barnier, L., Bessières, L.T. and Maze, G. (2018) Chaotic variability of ocean heat content: climate-relevant features and observational implications. *Oceanography*, 31(2).

Philander, S.G.H. (1983) El Nino southern oscillation phenomena. *Nature*, 302(5906), 295.

Porta Mana, P. and Zanna, L. (2014) Toward a stochastic parametrization of ocean mesoscale eddies. *Ocean Modelling*, 79(0), 1–20.

Roberts, M.J., Hewitt, H.T., Hyder, P., Ferreira, D., Josey, S.A., Mizielinski, M. and Shelly, A. (2016) Impact of ocean resolution on coupled air–sea fluxes and large-scale climate. *Geophysical Research Letters*, 43, 10,430–10,438.

Scaife, A.A., Woollings, T., Knight, J., Martin, G. and Hinton, T. (2010) Atmospheric blocking and mean biases in climate models. *Journal of Climate*, 23(23), 6143–6152.

Serazin, G., Jaymond, A., Leroux, S., Penduff, T., Bessieres, L., Llovel, W., Barnier, B., Molines, J.-M. and Terray, L. (2017) A global probabilistic study of the ocean heat content low-frequency variability: atmospheric forcing versus oceanic chaos. *Geophysical Research Letters*, 44, 5580–5589.

Serazin, G., Penduff, T., Barnier, B., Molines, J.M., Terray, L. and Arbic, B. (2018) Inverse cascades of kinetic energy as a source of low-frequency intrinsic

variability: a global OGCM study. *Journal of Physical Oceanography*, 48, 1385–1408. https://doi.org/10.1175/JPO-D-17-0136.1.

Sevellec, F. and Fedorov, A.V. (2013) The leading, interdecadal eigenmode of the Atlantic meridional overturning circulation in a realistic ocean model. *Journal of Climate*, 26, 2160–2183.

Sevellec, F. and Fedorov, A.V. (2017) Predictability and decadal variability of the North Atlantic ocean state evaluated from a realistic ocean. *Journal of Climate*, 30(2), 477–498.

Spall, M.A. (1996) Dynamics of the Gulf stream/deep western boundary current crossover. Part II: low-frequency internal oscillations. *Journal of Physical Oceanography*, 26(10), 2169–2182.

Stainforth, D.A., Aina, T., Christensen, C., Collins, M., Faull, N., Frame, D.J., Kettleborough, J.A., Knight, S., Martin, A., Murphy, J.M., Piani, C., Sexton, D., Smith, L.A., Spicer, R.A., Thorpe, A.J. and Allen, M.R. (2005) Uncertainty in predictions of the climate response to rising levels of greenhouse gases. *Nature*, 433, 403–406.

von Storch, J.-S., Haak, H., Hertwig, E. and Fast, I. (2016) Vertical heat and salt fluxes due to resolved and parametrized meso-scale eddies. *Ocean Modelling*, 108, 1–19.

Subramanian, A.C. and Palmer, T.N. (2017) Ensemble super-parametrization versus stochastic parametrization: a comparison of model uncertainty representation in tropical weather prediction. *Journal of Advances in Modeling Earth Systems*, 9, 1231–1250. https://doi.org/10.1002/2016MS000857.

Tebaldi, C. and Knutti, R. (2007) The use of the multimodel ensemble in probabilistic climate projections. *The Royal Society*, 365(1857), 2053–2075.

Teixeira, J., Reynolds, C.A. and Judd, K. (2007) Time step sensitivity of nonlinear atmospheric models: numerical convergence, truncation error growth, and ensemble design. *Journal of the Atmospheric Sciences*, 64, 175–189.

Tziperman, E., Zanna, L. and Penland, C. (2008) Nonnormal thermohaline circulation dynamics in a coupled ocean–atmosphere GCM. *Journal of Physical Oceanography*, 3, 588–604.

Vallis, G.K. (2006) *Atmospheric and Oceanic Fluid Dynamics*. Cambridge: Cambridge University Press.

Venaille, A., Vallis, G.K. and Smith, K.S. (2011) Baroclinic turbulence in the ocean: analysis with primitive equation and quasigeostrophic simulations. *Journal of Physical Oceanography*, 41, 1605–1623.

Williams, P.D. (2009) A proposed modification to the Robert–Asselin time filter. *Monthly Weather Review*, 137, 2538–2546.

Williams, P.D. (2011) The RAW filter: an improvement to the Robert–Asselin filter in semi-implicit integrations. *Monthly Weather Review*, 139, 1996–2007.

Williams, P.D. (2012) Climatic impacts of stochastic fluctuations in air–sea fluxes. *Geophysical Research Letters*, 39, L10705.

Williams, P.D. (2013) Achieving seventh-order amplitude accuracy in leapfrog integrations. *Monthly Weather Review*, 141, 3037–3051.

Williams, P.D., Howe, N.J., Gregory, J.M., Smith, R.S. and Joshi, M.M. (2016) Improved climate simulations through a stochastic parametrization of ocean eddies. *Journal of Climate*, 29(24), 8763–8781.

Wilson, C., Hughes, C.W. and Blundell, J.R. (2015) Forced and intrinsic variability in the response to increased wind stress of an idealized Southern Ocean. *Journal of Geophysical Research Oceans*, 120, 113–130.

Winton, M. (2003) On the climatic impact of ocean circulation. *Journal of Climate*, 16, 2875–2889.

Woolnough, S.J., Vitart, F. and Balmaseda, M.A. (2007) The role of the ocean in the Madden–Julian Oscillation: implications for MJO prediction. *Quarterly Journal of the Royal Meteorological Society*, 133, 117–128.

Yin, Y., Alves, O. and Oke, P.R. (2011) An ensemble ocean data assimilation system for seasonal prediction. *Monthly Weather Review*, 139(3), 786–808.

Young, C., Liang, Y., Tseng, Y. and Chow, C. (2014) Characteristics of the RAW-filtered leapfrog time-stepping scheme in the ocean general circulation model. *Monthly Weather Review*, 142, 434–447.

Zanna, L. and Tziperman, E. (2008) Surface optimal excitation of the thermohaline circulation. *Journal of Physical Oceanography*, 38(8), 1820–1830.

Zanna, L., Heimbach, P., Moore, A.M. and Tziperman, E. (2011) Optimal excitation of interannual Atlantic meridional overturning circulation variability. *Journal of Climate*, 24(2), 413–427.

Zanna, L., Heimbach, P., Moore, A.M. and Tziperman, E. (2012) Upper ocean singular vectors of the North Atlantic climate with implications for linear predictability and variability. *Quarterly Journal of the Royal Meteorological Society*, 138, 500–513.

Zanna, L., Porta Mana, P.G.L., Anstey, J., David, T. and Bolton, T. (2017) Scale-aware deterministic and stochastic parametrizations of eddy-mean flow interaction. *Ocean Modelling*, 111, 66–80.