

## Oceanic Stochastic Parameterizations in a Seasonal Forecast System

M. ANDREJCZUK,<sup>+</sup> F. C. COOPER, S. JURICKE, T. N. PALMER, A. WEISHEIMER,<sup>#</sup> AND L. ZANNA

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

(Manuscript received 15 July 2015, in final form 12 January 2016)

### ABSTRACT

Stochastic parameterization provides a methodology for representing model uncertainty in ensemble forecasts. Here the impact on forecast reliability over seasonal time scales of three existing stochastic parameterizations in the ocean component of a coupled model is studied. The relative impacts of these schemes upon the ocean mean state and ensemble spread are analyzed. The oceanic variability induced by the atmospheric forcing of the coupled system is, in most regions, the major source of ensemble spread. The largest impact on spread and bias came from the stochastically perturbed parameterization tendency (SPPT) scheme, which has proven particularly effective in the atmosphere. The key regions affected are eddy-active regions, namely, the western boundary currents and the Southern Ocean where ensemble spread is increased. However, unlike its impact in the atmosphere, SPPT in the ocean did not result in a significant decrease in forecast error on seasonal time scales. While there are good grounds for implementing stochastic schemes in ocean models, the results suggest that they will have to be more sophisticated. Some suggestions for next-generation stochastic schemes are made.

### 1. Introduction


Seasonal forecasting with coupled atmosphere–ocean–land surface models has become well established at many numerical weather prediction (NWP) and climate forecast centers during the last two decades (MacLachlan et al. 2015; Molteni et al. 2011; Saha et al. 2014). These coupled systems for seasonal forecasting exploit predictability originating from the ocean and the land surface. More specifically, coupled models allow for predictions of seasonal-to-interannual variations of the climate system such as the El Niño–Southern Oscillation (ENSO) cycle,

which may in turn affect predictions of other long time-scale variations. Coupled models are also increasingly being used for medium-range weather predictions. For example, the European Centre for Medium-Range Weather Forecasts (ECMWF) recently introduced a coupled atmosphere–ocean model into its operational forecasts in November 2013 (Lang et al. 2015; Janssen et al. 2013).

Forecast uncertainties need to be accounted for in these prediction systems. For example, errors in the observations and an incomplete observing system lead to inaccuracies in the initial conditions of forecasts. Because of the chaotic nature of the atmosphere and the earth system, these initial errors tend to grow quickly, reducing the predictive skill of forecast models. Ensembles of model simulations are used to account for initial condition uncertainty. Each ensemble member is initialized with slightly different initial conditions, generated using a perturbation method [e.g., ensembles of data assimilations, singular vectors; see, e.g., Molteni et al. (2011)]. The ensemble as a whole then provides a probabilistic forecast.

In addition to the uncertainties in initial conditions, forecast models themselves are inaccurate due to the

---

 Denotes Open Access content.

---

 Supplemental information related to this paper is available at the Journals Online website: <http://dx.doi.org/10.1175/MWR-D-15-0245.s1>.

<sup>+</sup> Current affiliation: Met Office, Exeter, United Kingdom.

<sup>#</sup> Additional affiliation: European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, United Kingdom.

---

*Corresponding author address:* Laure Zanna, Atmospheric, Oceanic and Planetary Physics, Department of Physics, University of Oxford, Clarendon Lab, Parks Road, Oxford OX1 3PU, United Kingdom.  
E-mail: laure.zanna@physics.ox.ac.uk



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

DOI: 10.1175/MWR-D-15-0245.1

numerical approximations used in the temporal and spatial discretization or due to inaccurate representations of subgrid-scale processes. In the past few decades, different strategies have emerged to account for these model uncertainties. Multimodel ensembles make use of the diverse set of weather and climate models to sample the uncertainty in the model formulation, on time scales of weather (e.g., Mylne et al. 2002; Park et al. 2008), seasonal (Palmer et al. 2004; Weisheimer et al. 2009), decadal (e.g., Smith et al. 2013; Meehl et al. 2014), and climate (e.g., Tebaldi and Knutti 2007; Flato et al. 2013) predictions. Perturbed parameter ensembles, on the other hand, sample the uncertainty in the choice of specific, imperfectly constrained model parameters (e.g., Stainforth et al. 2005). Furthermore, in multi-parameterization ensembles the set of applied parameterization schemes is modified for each ensemble member.

Finally, stochastic parameterizations are becoming an alternative technique for representing model uncertainty, especially for atmospheric NWP (e.g., Shutts 2005; Berner et al. 2009; Palmer et al. 2009, and references therein). By injecting stochastic perturbations into the system at the subgrid scale, uncertainties in closure schemes are incorporated and unresolved subgrid-scale variability may be taken into account. Some of the stochastic schemes such as the stochastically perturbed parameterization tendency (SPPT) scheme have been used successfully in atmospheric forecast models (e.g., Palmer et al. 2009), leading to improved forecast skill up to the seasonal time scales (Weisheimer et al. 2014; Batté and Doblas-Reyes 2015). SPPT is a multiplicative stochastic scheme based on existing deterministic parameterizations. The study of Weisheimer et al. (2011) compared the multimodel, perturbed parameter and atmospheric stochastic physics approach for monthly and seasonal forecasts showing the potential for stochastic parameterizations to outperform the multimodel ensemble.

The motivation for stochastic parameterizations in the atmosphere originates to some degree from the existence of power-law structures and the related rapid up-scale error propagation [see Palmer (2012) for a detailed discussion]. Since similar power-law structures, associated with mesoscale eddies, can also be found in the ocean (LaCasce and Ohlmann 2003; LaCasce 2008), similar arguments for the potential role of stochastic parameterizations may therefore hold.

Stochastic parameterizations have not only been introduced into the atmospheric component of NWP and seasonal forecast models (e.g., Buizza et al. 1999; Palmer et al. 2009), but have more recently also been implemented into the sea ice (Juricke et al. 2013; Juricke and Jung 2014), ocean (e.g., Brankart 2013; Brankart et al. 2015), land surface (MacLeod et al. 2016), and

air–sea coupling (e.g., Williams 2012; Beena and von Storch 2009) components of global general circulation models, to account for the uncertainty in subgrid parameterizations. Although the impact of ocean stochastic parameterizations has been demonstrated in a climatological context (e.g., Brankart 2013), as yet no comparable stochastic parameterization in the ocean has been implemented into operational coupled models. Since some of the proposed stochastic schemes for ocean subgrid-scale parameterizations (e.g., Porta Mana and Zanna 2014) may not be straightforward to implement in complex coupled global models, the purpose of this study is to investigate the impact of simpler stochastic schemes of similar complexity to those used in the atmosphere. These schemes aim at quantifying model uncertainty based on existing parameterizations.

Three methods of stochastic parameterization are considered: a surface flux parameterization similar to Williams (2012), a stochastic perturbation of the equation of state similar to Brankart (2013), and stochastic perturbations of the parameterized tendencies of diffusion—mixing and viscosity—which can be considered to be the application to the ocean of the SPPT scheme used successfully in atmospheric models (e.g., Palmer et al. 2009). The details of the model setup, the stochastic parameterization schemes, and the methods of data analysis are given in section 2. The results of several ensemble integrations over a 10-yr period are presented in section 3 and our conclusions are summarized and discussed in section 4.

## 2. Model configuration and experimental design

Integrations were performed using a variation of the ECMWF seasonal forecast system [system 4; Molteni et al. (2011)]. The ocean component consists of the NEMO v3.0 global ocean primitive equation model (Madec et al. 2008) discretized onto an approximately 1° ORCA tripolar grid (Madec and Imbard 1996) with 42 vertical levels. NEMO is coupled to the ECMWF atmospheric forecast model IFS Cy36r4, integrated at T159 horizontal resolution (reduced from T255 in system 4) with 91 vertical levels. Ocean reanalysis data, used for initial conditions and verification, was provided by the new ECMWF operational ocean reanalysis system (ORAS4; Balmaseda et al. 2013). The ensemble of five reanalyses is driven by sampling uncertainty in winds and in deep ocean initial conditions, and subsampling observation coverage. The ocean analyses are then augmented by applying SST perturbations with an associated subsurface temperature signal (Molteni et al. 2011), as stochastic SST perturbations are linearly decreased to a depth of 80 m. The unperturbed ensemble member of ORAS4 is used for verification. Atmospheric

initial conditions are derived from the ERA-Interim dataset (Dee et al. 2011), without initial or stochastic perturbations applied.

To investigate the potential impact of spatially and temporally correlated noise used for the stochastic schemes, we followed the example of first implementations of SPPT in the atmosphere (see Buizza et al. 1999) and introduced a simple, computationally efficient spatial correlation method. The horizontal domain was divided into a coarse, regular  $L \times L$  latitude longitude grid. In each of the cells of this coarse grid, a single uniformly distributed random number was generated at a given time interval  $I$  and applied uniformly to each of the ocean model grid cells within the respective  $L \times L$  region. Different resolutions for  $L$  and time intervals  $I$  were tested. The diagnostics in this paper focus on the results with  $L = 30^\circ$  and with  $I = 1$  or  $I = 30$  days, as these showed the largest, physically reasonable and computationally stable response in terms of changes to the generated oceanic ensemble spread. Reduced spatial and time scales reduced the impact. Results of a reduced  $20^\circ \times 20^\circ$  spatial correlation grid will be briefly discussed as well.

A 10-member ensemble of experiments, initialized on 1 November 1989 was integrated for 3 months. The same procedure was repeated over 9 more years to generate an ensemble for each of the 10 years 1989–98, for a total of 100 integrations. Our setup is designed to mimic, in a computationally efficient way an operational seasonal forecasting system (e.g., system 4 from ECMWF). We define three regions for the purpose of summarizing the time-dependent nature of the integrations: 1) the North Atlantic subpolar region,  $45^\circ\text{--}65^\circ\text{N}$ ,  $70^\circ\text{W}\text{--}0^\circ$ ; 2) the North Atlantic subtropics region,  $10^\circ\text{--}45^\circ\text{N}$ ,  $70^\circ\text{W}\text{--}0^\circ$ ; and 3) the Southern Ocean,  $35^\circ\text{--}65^\circ\text{S}$ , all longitudes. Although we considered other regions, including the tropical Pacific and whole globe averages, the above regions were chosen in order to focus on the areas where the largest biases in the mean and the largest ensemble variance occur. These are also regions where the schemes examined here have the largest impact. Results of area averages over other regions support the findings described here, or are inconclusive (i.e., cannot be distinguished from the noise). In this context, the criterion to distinguish signal from noise is simply defined as approximately where the mean is greater than two standard deviations in the mean:

$$m > \frac{2\sigma}{\sqrt{n}} \quad (1)$$

[e.g., see section 2.5.9 in von Storch and Zwiers (1999)]. Here  $m$  is the ensemble mean difference between a

stochastically perturbed and a control ensemble integration,  $\sigma$  is the standard deviation of the differences of all start dates, and  $n = 10$  is the number of independent samples, assuming that the model state is independent from one year to the next. Equation (1) is similar to the 95% confidence two tailed  $t$  test. We chose this test statistic because it is simple and to avoid giving the impression of accurate confidence intervals. The  $t$  test is not robust against the sampling assumption that every realization occurs independently of all other realizations [e.g., see section 6.6.1 in von Storch and Zwiers (1999)] and is not appropriate for variances. The choice of a  $2\sigma$  threshold is somewhat arbitrary and provides a rough guide. [See Fig. 3 for an impression of the spread indicated by (1).]

#### a. Stochastic surface flux (SSF)

Stochastic perturbations of the air–sea fluxes were applied to the seasonal forecast model based upon the method described in Williams (2012) who performed experiments with a lower-resolution coupled atmosphere–ocean GCM. Their ocean model (OPA 8.2; Madec et al. 1998) was integrated using the ORCA2 grid, which has approximately a  $2^\circ$  horizontal resolution, with 31 vertical levels. Their atmospheric model (ECHAM 4.6; Roeckner et al. 1996) was integrated using a T30 spectral grid with 19 vertical levels. In Williams (2012), the air–sea freshwater flux  $\Delta S$  and nonsolar heat flux  $\Delta T$ , were perturbed separately in two separate experiments, given the following perturbation scheme:

$$\Delta T \rightarrow (1 + r_T)\Delta T, \quad \text{and} \quad \Delta S \rightarrow (1 + r_S)\Delta S, \quad (2)$$

with  $r_T$  and  $r_S$  being random numbers uniformly distributed between  $\pm 0.5$  and generated at 3-h intervals on the ORCA2 grid scale. For this paper, in addition to moving to a higher resolution, the interval over which random numbers were chosen was increased to  $I = 1$  day and applied using the  $L \times L$  grid defined above with  $L = 30^\circ$ . Also, both fluxes were perturbed simultaneously but with different sequences of random numbers.

#### b. Stochastic equation of state (SES)

The method of stochastic parameterization of the nonlinear equation of state, which relates the density to the temperature, salinity, and pressure, is based upon the method described in Brankart (2013). To simulate the uncertainty related to area-averaged temperature and salinity fields used as input for the equation of state, a first-order autoregressive process perturbs both state variables by an amount proportional to their gradients. The autoregressive process was implemented with a

decay time scale of 12 days, however, integrations did not complete. We found that a decay time of 7.5 days was necessary to keep the model stable. Further investigation, beyond our current scope, is necessary to establish the precise reason for this instability. Ultimately the density at each grid point is perturbed independently of neighboring grid cells as described by Brankart (2013). In Brankart (2013) integrations were performed using NEMO with the same ORCA2 grid as that used by Williams (2012). Their system, however, was forced using climatological atmospheric data without interannual variations (Large and Yeager 2009) rather than the atmosphere from a coupled model.

### c. Stochastically perturbed parameterization tendency

We introduce a SPPT scheme for the ocean similar to that implemented in atmospheric models for ensemble weather forecasts. The subgrid-scale-parameterized tendencies (used to crudely mimic turbulent diffusion, mixing, convection, and viscosity) applied to the zonal velocity  $u$ , meridional velocity  $v$ , salinity  $S$ , and temperature  $T$ , are multiplied by  $(1 + r_X)$  as in (2), with different random sequences  $r_X$  for  $u$ ,  $v$ ,  $S$ , and  $T$ , that is,  $X \in \{u, v, S, T\}$ . For example, the deterministic prognostic equation for  $T$  given by

$$\frac{\partial T}{\partial t} = -\nabla \cdot (T\mathbf{U} + T\mathbf{U}_{\text{GM}}) + D_T + F_T \quad (3)$$

takes the following form when SPPT is implemented

$$\frac{\partial T}{\partial t} = -\nabla \cdot (T\mathbf{U} + T\mathbf{U}_{\text{GM}}) + (1 + r_T)D_T + F_T, \quad (4)$$

where  $\mathbf{U} = (u, v, w)$  is the 3D Eulerian velocity,  $\mathbf{U}_{\text{GM}}$  is the eddy-induced velocity from the Gent–McWilliams parameterization scheme (Gent and McWilliams 1990),  $D_T$  represents the parameterized diffusion and mixing tendencies, and  $F_T$  is the air–sea flux.

Summarized in the respective  $D_X$  and  $X \in \{u, v, S, T\}$ , terms for the momentum and tracer equations are parameterized terms using vertical eddy viscosity and diffusivity coefficients calculated by a turbulent kinetic energy (TKE) closure scheme as well as a double-diffusion mixing scheme. Lateral diffusion and viscosity are also included in  $D_X$  using horizontally varying coefficients for tracers [following Held and Larichev (1996) and Treguier et al. (1997)] as well as a three-dimensional spatially varying viscosity coefficient for the momentum equations. See Madec et al. (2008) for the specification of these schemes. For the results discussed in section 3 the spatial field of the random numbers  $r_X$  was generated at either  $I = 30$  days or alternatively  $I = 1$ -day intervals using the  $L = 30^\circ$  grid

defined above with the same values of  $r_X$  applied on each vertical model level. The four  $r_X$  were applied to the parameterized tendencies of the respective prognostic equations simultaneously for all four fields and were drawn from a uniform distribution between  $\pm 0.8$ . This value for the magnitude of  $r_X$  was found to be the maximum value consistent with model stability. The eddy viscosity term  $D_T$  stabilizes the model, so a value of  $r_u < -0.8$ , held for 30 days, could be an obvious source of model instability.

### 3. Impact of the stochastic parameterizations on model bias and ensemble forecast performance

The control ensemble integrations without any stochastic perturbations exhibit biases relative to the reanalysis. The daily bias is estimated by taking the difference between the reanalysis and the ensemble mean for each start year averaged over all 10 start years. Averaging in time between 60 and 90 days into the integration indicates that this bias, as illustrated for upper 300-m ocean heat content in Fig. 1a, often coincides with the regions of high ensemble variability such as the eastern tropical Pacific, Southern Ocean, Gulf Stream, and Kuroshio regions. The upper 300-m heat content is chosen as it yields a particularly strong signal-to-noise ratio compared to sea surface temperatures in isolation (see the online supplemental material).

The SPPT scheme leads to a change in the bias in the Southern Ocean, particularly in the region of the south coast of Australia, and the North Atlantic (Fig. 1b). Comparing Figs. 1a and 1b indicates that the warm bias along the south coast of Australia is reduced. Taking the area means, specified in section 2, of the daily bias (Fig. 2) highlights a reduction in the mean bias for the upper 300-m ocean heat content and sea surface salinity (SSS) in the North Atlantic subpolar region after the first month. The change in bias of SST found in our study is relatively small (about 10%) in the North Atlantic and the tropical Atlantic over the 3-month integration performed compared to those obtained when increasing the ocean resolution (Scaife et al. 2011) or the atmospheric vertical resolution (Harlasse et al. 2015) for long integrations. The bias is initially not exactly zero due to the random spread in the initial condition perturbations. In the Southern Ocean, while the warm bias has been reduced, no noticeable changes are observed in the SSS bias. In contrast, the bias in the North Atlantic subtropical region has increased due to SPPT. For certain regions the difference in bias remains relatively constant over the length of the integration.



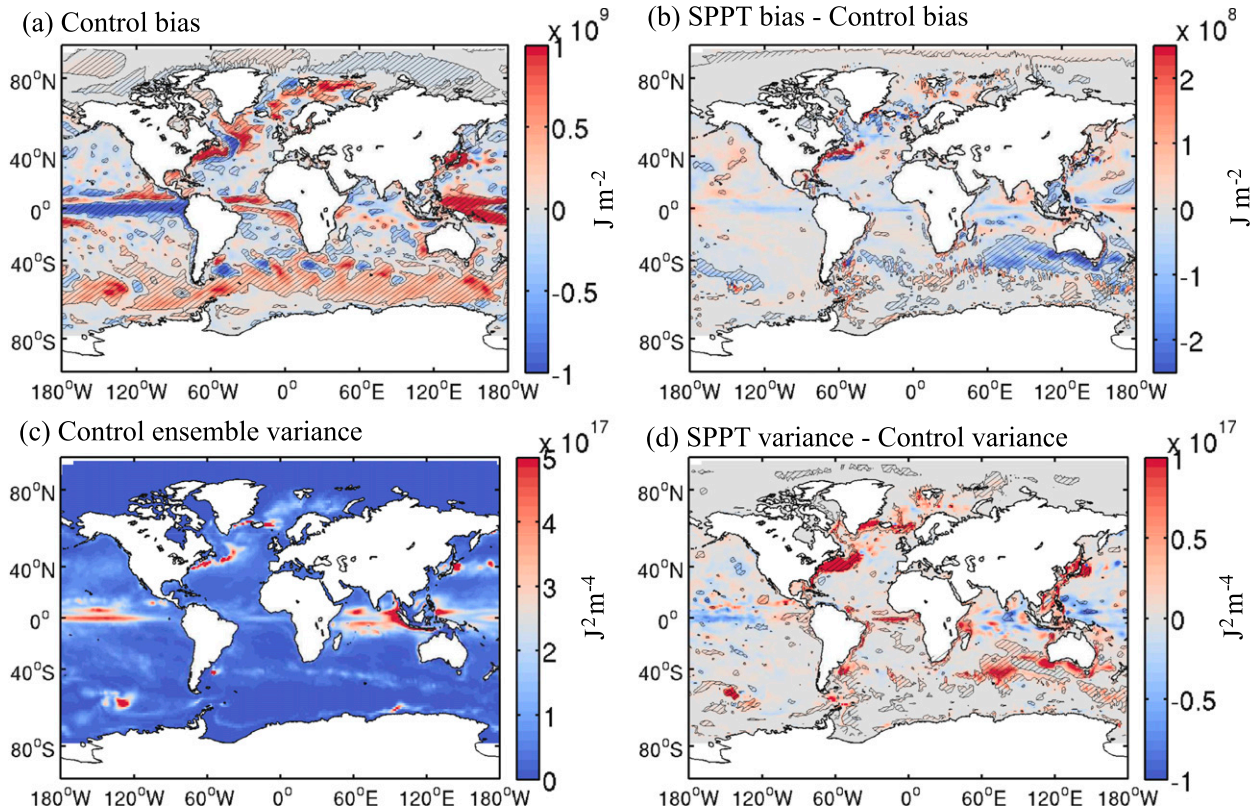


FIG. 1. Statistics of the upper 300-m ocean heat content ( $\text{J m}^{-2}$ ) averaged over days 60–90 of the seasonal forecast integration. (a) Mean bias in the control integration. (b) Difference in the bias between the integration with SPPT and the control. (c) Ensemble spread (variance) in the control integration. (d) Difference in the ensemble spread between the integration with SPPT and the control. Hatched areas indicate regions that are distinguishable from zero using (1). Note the different color scales.

The gray lines in Fig. 2 represent the ensemble mean bias of the control integration for each of the 10 years and can be considered as an indicator of confidence. Changes in the ensemble bias and spread (defined as the ensemble variance) between the control and integrations using the SSF and SES schemes are too small to distinguish from zero. The inherent ensemble spread due to the ocean initial condition perturbations and atmospheric variability is large in comparison. When it comes to the area-averaged quantities, Fig. 2 also confirms that the SSF and SES schemes have little impact upon the bias though there are regions in which the bias appears slightly reduced (e.g., North Atlantic subpolar SSS around 2–3 months for the SES scheme).

The impact of the SPPT scheme upon the ensemble spread is apparent in several key regions (Fig. 1d). The largest and most significant impact is shown in the region of the south coast of Australia with patches of visibly increased spread throughout the Southern Ocean. This is a similar pattern to the changes in vertical mixing induced by a stochastic wind forcing compared to climatological wind forcing [cf. Fig. 8 in Beena and von Storch

(2009)]. In addition to the Southern Ocean, Fig. 1d also reveals increased ensemble spread along western boundary currents, including the Gulf Stream and Kuroshio. Although there are reductions of ensemble spread in parts of the tropics, they are not distinguishable from zero using the criterion in (1). The increase in ensemble spread over the length of the integrations is readily seen in area means of the North Atlantic subtropical and subpolar regions and the Southern Ocean (Figs. 3a–c). For the North Atlantic subtropics (see Fig. 3b), which corresponds to a rather moderate increase in regionally averaged ensemble spread, the relative increase in spread after three months represents about 40% of the mean ensemble spread of the deterministic ensembles. The sharp changes in spread apparent in Figs. 3a–c is due to a new set of random numbers being chosen every 30 days. We expect that substituting the  $r_X$  for autoregressive processes with the equivalent time scales will remove this artifact.

On the other hand, Figs. 3d–f indicates that the SPPT scheme has increased the regional spread at the expense of increased forecast error in the upper ocean heat

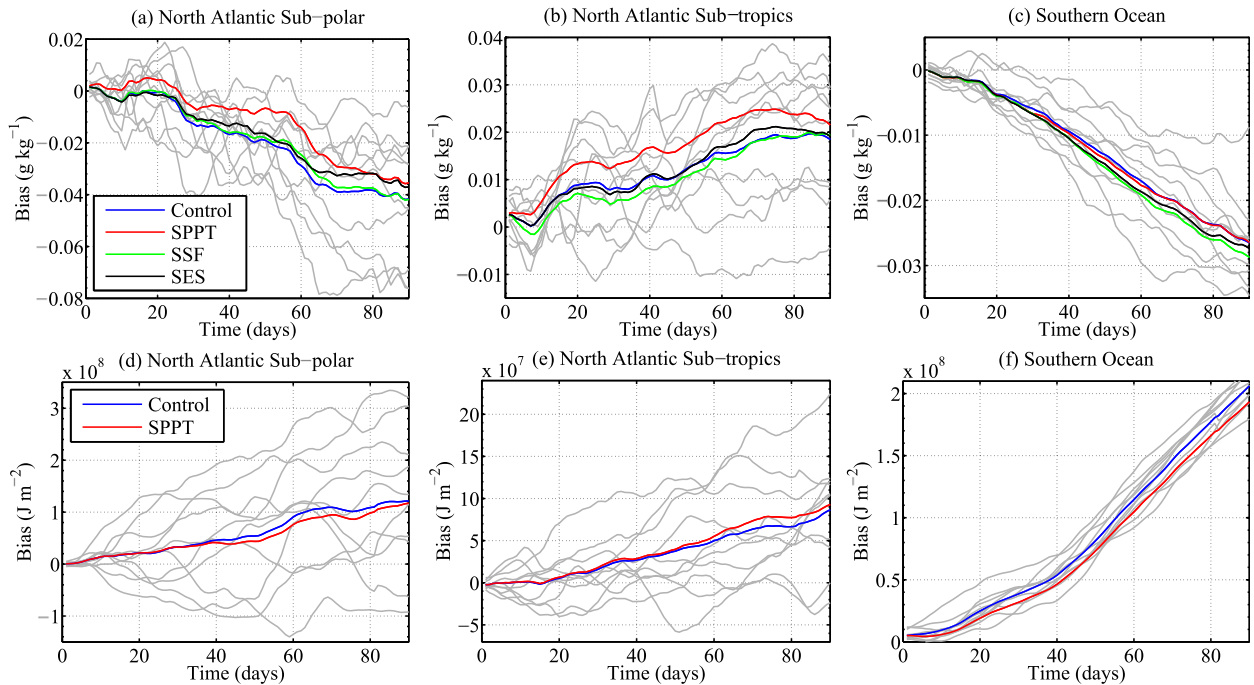


FIG. 2. Bias in the (a)–(c) sea surface salinity and (d)–(f) upper 300-m ocean heat content averaged over (left to right) three regions. See text for details. The colored lines indicate the ensemble mean bias averaged over all 10 start years for the control integrations (blue), the integrations with SPPT (red), SSF (green), and SES (black). Each gray line represents the 10-member ensemble mean bias of the control integration. There is 1 gray line for each of the 10 yr. Note the different y-axis scales.

content in most regions over the length of the integration. The forecast error is defined as the difference between the bias-corrected ensemble mean and the reanalysis. Bias corrected means that the daily bias is subtracted from each integration. The mean squared error is then the squared error, averaged over the region of interest. The forecast error is slightly, but statistically significantly reduced between 2 and 5 weeks in the North Atlantic subtropical region. Comparing Figs. 3a and 3d, Figs. 3b and 3e, and Figs. 3c and 3f, indicates that the additional mean squared error is approximately an order of magnitude lower than the increase in spread. The equivalent results for the SSF and SES schemes show that there is little impact upon the total ensemble spread (see the online supplemental material).

Reducing the time between random numbers from  $I = 30$  days to  $I = 1$  day resulted in a reduced impact for the SPPT scheme (to the point at which changes are almost indistinguishable from zero) when comparing ensemble bias, spread, and mean squared error to those of the control integration. This supports the conventional hypothesis that the impact of the stochastic term is strongly dependent upon its time scale, with longer time scales corresponding to a larger impact. Integrations in which the spatial correlation of the noise is reduced to an  $L = 20^\circ$  grid yield only small changes. From previous

sensitivity studies we would expect that as the correlation length scale is reduced further, there will come a point at which the impact of the random term is strongly reduced (see, e.g., Juricke et al. 2013). (See the online supplemental material for additional figures demonstrating the reduced impact at these time scales and equivalent plots for the SSF and SES schemes and the impact on the sea surface temperature and salinity.)

#### 4. Discussion

In this study, we test three oceanic stochastic parameterization schemes that aim at quantifying model uncertainty: the stochastically perturbed parameterization tendency (SPPT), the stochastic surface flux (SSF) (Williams 2012), and a stochastic equation of state (SES) (Brankart 2013). The relatively simple SPPT scheme, which has proved to be an important element of an ensemble forecast system in numerical weather prediction, injects multiplicative noise into the prognostic equations with an amplitude proportional to the deterministically parameterized tendencies. These three schemes are applied to the ocean component of a state-of-the-art seasonal coupled forecast system to account in part for the uncertainty in subgrid processes. The model considered here exhibits relatively large oceanic variability

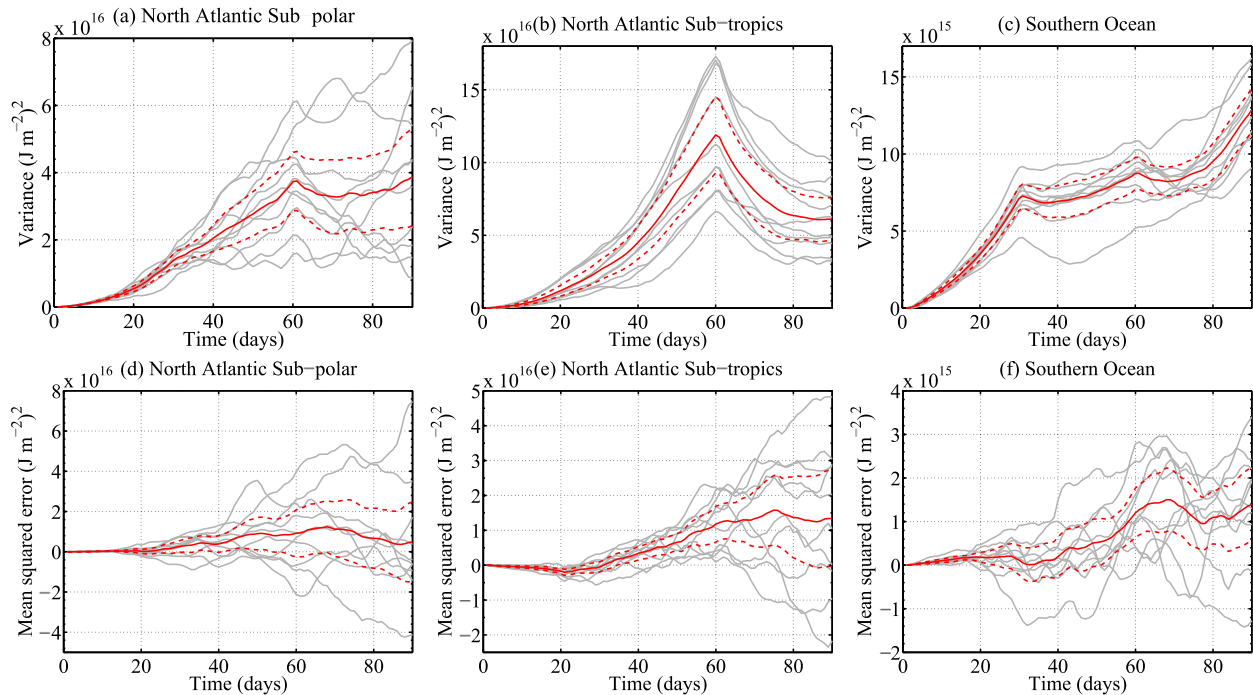


FIG. 3. Area-averaged changes in the (top) ensemble spread and (bottom) mean square error with respect to the control integration due to SPPT of upper 300-m ocean heat content. See text for details. The gray lines indicate the statistics from the ensembles for each of the 10 start years. The solid red line indicates the mean over all years and the dashed lines indicate the uncertainty in the mean (the mean  $\pm 2\sigma/10^{1/2}$ ).

compared to a system run at coarser resolution or without an interactive atmosphere. Using such a model, the impact of the SSF and SES schemes was relatively small at the monthly time scales considered, and may be more visible through longer time-scale integrations. In the case of the SES and SPPT schemes, the amplitude of the stochastic perturbations was limited by model stability.

Results show that compared with the other schemes, SPPT is an effective stochastic parameterization for increasing ensemble spread for variables such as sea surface temperature and salinity, and upper 300-m ocean heat content. The impact of SPPT was found to be particularly marked, and visible above the background variability, in regions of strong eddy activity, such as along western boundary currents in the Gulf Stream and Kuroshio regions, in the North Atlantic subpolar region, and also in parts of the Southern Ocean. In most other regions the atmospherically induced oceanic variability is the major contributor to the ensemble spread. Similar results were found by Juricke et al. (2014) in the context of applying stochastic perturbations to the sea ice strength in seasonal sea ice modeling. They showed that after a few weeks the atmospheric variability is the largest contributor to sea ice ensemble spread.

While SPPT increased ensemble spread for some regions, ensemble-mean forecast skill on the other hand

was not improved by the addition of SPPT, with the exception of the North Atlantic subtropics. For some regions, model bias was made worse. The latter does not necessarily imply that SPPT is degrading model accuracy, since the value of many climate model parameters are found by running the model in deterministic mode and estimating the values that fit the observations best. If a stochastic scheme impacts on the model mean state, then tuning should be performed using the full stochastic model and not a deterministic approximation to it (Palmer 2012). In addition to tuning the deterministic parameters of the model, the impact of a stochastic parameterization may be tuned by adjusting the decorrelation time scale and spatial distribution of the random perturbations as well as their magnitude. These are key parameters of the stochastic parameterization [e.g., see Cooper and Zanna (2015) for a discussion]. For the particular configurations discussed in this study, changes to the time scale appeared more important than the spatial scale. It should also be mentioned that the procedure used to create the random perturbations is crucial if unphysical artifacts such as the sharp changes in spread apparent in Figs. 3a–c are to be avoided. For future implementations of these kinds of stochastic ocean parameterization schemes the pattern generator used to create spatially correlated perturbations should

be able to produce more realistic, smoothly varying random fields, respecting physical constraints such as boundary conditions. Where possible, observational data should be used to get estimates for spatial and temporal decorrelation scales. A more technical constraint, on the other hand, is the computational efficiency of such a random pattern generator, especially when it comes to operational seasonal forecasting.

While the current implementation of SPPT led to an increase in spread, it was not able to demonstrate a reduction in ensemble-mean forecast error, even when the model fields have been bias corrected a posteriori. This result is possibly due to SPPT being too crude a scheme for ocean models. In the ocean, subgrid processes are parameterized largely by diffusion and viscosity. The impact of uncertainty in these terms upon the large-scale oceanic circulation, as simulated by SPPT, may not well reflect the uncertainty in the underlying turbulent processes. What this suggests is that a more positive impact than has been found here requires the development of more sophisticated stochastic schemes for unresolved and missing processes. Nonetheless the simple oceanic SPPT scheme introduced in this study provides a first estimate of the potential impact of stochastic ocean parameterizations in an operational seasonal forecast model.

The development of such stochastic closures to represent unresolved and missing processes should when possible remain consistent with fundamental physical constraints. For example, our initial implementations of a stochastic Gent–McWilliams scheme violated important nondivergent and adiabatic constraints leading to unrealistic upwelling and growing instabilities (see the online supplemental material). Avenues for future investigations might include a stochastic Gent–McWilliams scheme that is stable over long time periods, and remains consistent with the nondivergent and adiabatic constraints of the deterministic scheme. Stochastic schemes should ultimately be guided by observations, however, ocean observations are sparse. As a methodology to develop new parameterizations, high-resolution idealized simulations can be substituted as “truth” (e.g., Berloff 2005, 2015) and optimally coarse grained to derive stochastic terms. This approach can provide a guide for the magnitude, spatial, and temporal patterns of the stochastic terms to be implemented in the numerical model, and can, for example, be decoupled from the background flow (e.g., Cooper and Zanna 2015). A complementary approach is to implement a PDF-based parameterization, such as Porta Mana and Zanna (2014) who have developed a stochastic parameterization of ocean mesoscale eddies that depends on the temporal tendency of potential vorticity.

*Acknowledgments.* This work was funded by the U.K. NERC Grant NE/J00586X/1, along with a generous allocation of supercomputer resources from ECMWF. AW was partly supported by the EU project SPECS funded by the European Commission Seventh Framework Research Programme under the Grant Agreement 308378. TNP was supported under European Research Council Grant PESM. Thanks go to Peter Düben and Aneesh Subramanian for useful discussions and encouragement during the preparation of this paper, to J.M. Brankart for sharing the code associated with his parametrization, and to the reviewers for their useful suggestions. The model data can be obtained by e-mailing Stephan Juricke at Stephan.Juricke@physics.ox.ac.uk.

#### REFERENCES

- 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.
- Batté, L., and F. Doblas-Reyes, 2015: Stochastic atmospheric perturbations in the EC-Earth3 global coupled model: Impact of SPPT on seasonal forecast quality. *Climate Dyn.*, **45**, 3419–3439, doi:10.1007/s00382-015-2548-7.
- Beena, B. S., and J.-S. von Storch, 2009: Effects of fluctuating daily surface fluxes on the time-mean oceanic circulation. *Climate Dyn.*, **33**, 1–18, doi:10.1007/s00382-009-0575-y.
- Berloff, P. S., 2005: Random-forcing model of the mesoscale oceanic eddies. *J. Fluid Mech.*, **529**, 71–95, doi:10.1017/S0022112005003393.
- , 2015: Dynamically consistent parameterization of mesoscale eddies. Part I: Simple model. *Ocean Modell.*, **87**, 1–19, doi:10.1016/j.oceomod.2014.12.008.
- Berner, J., G. J. Shutts, M. Leutbecher, and T. N. Palmer, 2009: A spectral stochastic kinetic energy backscatter scheme and its impact on flow-dependent predictability in the ECMWF ensemble prediction system. *J. Atmos. Sci.*, **66**, 603–626, doi:10.1175/2008JAS2677.1.
- Brankart, J.-M., 2013: Impact of uncertainties in the horizontal density gradient upon low resolution global ocean modeling. *Ocean Modell.*, **66**, 64–76, doi:10.1016/j.oceomod.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/gmd-8-1285-2015.
- Buizza, R., M. Miller, and T. N. Palmer, 1999: Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Quart. J. Roy. Meteor. Soc.*, **125**, 2887–2908, doi:10.1002/qj.49712556006.
- Cooper, F. C., and L. Zanna, 2015: Optimisation of an idealised ocean model, stochastic parameterisation of sub-grid eddies. *Ocean Modell.*, **88**, 38–53, doi:10.1016/j.oceomod.2014.12.014.
- Dee, D. P., and Coauthors, 2011: The ERA-interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, doi:10.1002/qj.828.
- 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. [Available online at <http://www.ipcc.ch/report/ar5/wg1/>.]
- 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.
- Harlasse, J., M. Latif, and W. Park, 2015: Improving climate model simulation of tropical Atlantic sea surface temperature: The importance of enhanced vertical atmosphere model resolution. *Geophys. Res. Lett.*, **42**, 2401–2408, doi:10.1002/2015GL063310.
- Held, I. M., and V. D. Larichev, 1996: A scaling theory for horizontally homogeneous, baroclinically unstable flow on a beta plane. *J. Atmos. Sci.*, **53**, 946–952, doi:10.1175/1520-0469(1996)053<0946:ASTFHH>2.0.CO;2.
- Janssen, P., and Coauthors, 2013: Air-sea interaction and surface waves. ECMWF Tech. Memo. 712, 34 pp.
- Juricke, S., and T. Jung, 2014: Influence of stochastic sea ice parametrization on climate and the role of atmosphere–sea ice–ocean interaction. *Philos. Trans. Roy. Soc. London*, **372A**, 2013 0283, doi: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.
- , H. F. Goessling, and T. Jung, 2014: Potential sea ice predictability and the role of stochastic sea ice strength perturbations. *Geophys. Res. Lett.*, **41**, 8396–8403, doi:10.1002/2014GL062081.
- LaCasce, J. H., 2008: Statistics from Lagrangian observations. *Prog. Oceanogr.*, **77**, 1–29, doi:10.1016/j.pocean.2008.02.002.
- , and C. Ohlmann, 2003: Relative dispersion at the surface of the Gulf of Mexico. *J. Mar. Res.*, **61**, 285–312, doi:10.1357/002224003322201205.
- Lang, S. T. K., M. Bonavita, and M. Leutbecher, 2015: On the impact of re-centring initial conditions for ensemble forecasts. *Quart. J. Roy. Meteor. Soc.*, **141**, 2571–2581, doi:10.1002/qj.2543.
- Large, W. G., and S. G. Yeager, 2009: The global climatology of an interannually varying air-sea flux data set. *Climate Dyn.*, **33**, 341–364, doi:10.1007/s00382-008-0441-3.
- MacLachlan, C., and Coauthors, 2015: Global seasonal forecast system version 5 (GloSea5): A high-resolution seasonal forecast system. *Quart. J. Roy. Meteor. Soc.*, **141**, 1072–1084, doi:10.1002/qj.2396.
- MacLeod, D. A., H. L. Cloke, F. Pappenberger, and A. Weisheimer, 2016: Improved seasonal prediction of the hot summer of 2003 over Europe through better representation of uncertainty in the land surface. *Quart. J. Roy. Meteor. Soc.*, **142**, 79–90, doi:10.1002/qj.2631.
- Madec, G., and M. Imbard, 1996: A global ocean mesh to overcome the North Pole singularity. *Climate Dyn.*, **12**, 381–388, doi:10.1007/BF00211684.
- , P. Delecluse, M. Imbard, and C. Levy, 1998: OPA 8.1 Ocean General Circulation Model reference manual. Rep. 11, France, 91 pp.
- , and Coauthors, 2008: NEMO ocean engine. Note du Pôle de modélisation 1288-1619, Rep. 27, France, 209 pp.
- Meehl, G. A., and Coauthors, 2014: Decadal climate prediction: An update from the trenches. *Bull. Amer. Meteor. Soc.*, **95**, 243–267, doi:10.1175/BAMS-D-12-00241.1.
- Molteni, F., and Coauthors, 2011: The new ECMWF seasonal forecast system (system 4). ECMWF Tech. Memo. 656, 49 pp.
- Mylne, K., R. Evans, and R. Clark, 2002: Multi-model multi-analysis ensembles in quasi-operational medium-range forecasting. *Quart. J. Roy. Meteor. Soc.*, **128**, 361–384, doi:10.1256/00359000260498923.
- 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.
- , and Coauthors, 2004: Development of a European Multi-model Ensemble System for Seasonal-to-Interannual Prediction (DEMETER). *Bull. Amer. Meteor. Soc.*, **85**, 853–872, doi:10.1175/BAMS-85-6-853.
- , R. Buizza, F. Doblas-Reyes, T. Jung, M. Leutbecher, G. Shutts, M. Steinheimer, and A. Weisheimer, 2009: Stochastic parametrization and model uncertainty. ECMWF Tech. Memo. 598, 42 pp.
- Park, Y.-Y., R. Buizza, and M. Leutbecher, 2008: TIGGE: Preliminary results on comparing and combining ensembles. *Quart. J. Roy. Meteor. Soc.*, **134**, 2029–2050, doi:10.1002/qj.334.
- Porta Mana, P. G. L., 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.
- Roeckner, E., and Coauthors, 1996: The atmospheric general circulation model ECHAM-4: Model description and simulation of present day climate. Rep. 218, Hamburg, Germany, 90 pp.
- Saha, S., and Coauthors, 2014: The NCEP Climate Forecast System version 2. *J. Climate*, **27**, 2185–2208, doi:10.1175/JCLI-D-12-00823.1.
- Scaife, A. A., and Coauthors, 2011: Improved Atlantic winter blocking in a climate model. *Geophys. Res. Lett.*, **38**, L23703, doi:10.1029/2011GL049573.
- Shutts, G., 2005: A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Quart. J. Roy. Meteor. Soc.*, **131**, 3079–3102, doi:10.1256/qj.04.106.
- Smith, D. M., and Coauthors, 2013: Real-time multi-model decadal climate predictions. *Climate Dyn.*, **41**, 2875–2888, doi:10.1007/s00382-012-1600-0.
- Stainforth, D. A., and Coauthors, 2005: Uncertainty in predictions of the climate response to rising levels of greenhouse gases. *Nature*, **433**, 403–406, doi:10.1038/nature03301.
- Tebaldi, C., and R. Knutti, 2007: The use of the multi-model ensemble in probabilistic climate projections. *Philos. Trans. Roy. Soc. London*, **365A**, 2053–2075, doi:10.1098/rsta.2007.2076.
- 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.
- von Storch, H., and F. W. Zwiers, 1999: *Statistical Analysis in Climate Research*. Cambridge University Press, 484 pp.
- Weisheimer, A., and Coauthors, 2009: ENSEMBLES: A new multi-model ensemble for seasonal-to-annual predictions—Skill and progress beyond DEMETER in forecasting tropical Pacific SSTs. *Geophys. Res. Lett.*, **36**, L21711, doi:10.1029/2009GL040896.
- , T. N. Palmer, and F. J. Doblas-Reyes, 2011: Assessment of representations of model uncertainty in monthly and seasonal forecast ensembles. *Geophys. Res. Lett.*, **38**, L16703, doi:10.1029/2011GL048123.
- , 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. London*, **372A**, 20130290, doi: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.