Geophysical and Astrophysical Fluid Dynamics
Vol. 00, No. 00, 00 Month 2018, 1–20

An eddifying Stommel model: Fast eddy effects in a two-box ocean

William Barham and Ian Grooms*

Department of Applied Mathematics, University of Colorado, Boulder, CO 80309

(*March 27, 2018*)

A system of stochastic differential equations is formulated describing the heat and salt content of a two-box ocean. Variability in the heat and salt content and in the thermohaline circulation between the boxes is driven by fast Gaussian atmospheric forcing and by ocean-intrinsic, eddy-driven variability. The eddy forcing of the slow dynamics takes the form of a colored, non-Gaussian noise. The qualitative effects of this non-Gaussianity are investigated by comparing to two approximate models: one that includes only the mean eddy effects (the 'averaged model'), and one that includes an additional Gaussian white-noise approximation of the eddy effects (the 'Gaussian model'). Both of these approximate models are derived using the methods of fast averaging and homogenization.

In the parameter regime where the dynamics has a single stable equilibrium the averaged model has too little variability. The Gaussian model has accurate second-order statistics, but incorrect skew and rare-event probabilities. In the parameter regime where the dynamics has two stable equilibria the eddy noise is much smaller than the atmospheric noise. The averaged, Gaussian, and non-Gaussian models all have similar stationary distributions, but the jump rates between equilibria are too small for the averaged and Gaussian models.

Keywords: Slow-fast systems; averaging; homogenization; stochastic differential equations; ocean modeling

1. Introduction

H. Stommel (1961) developed a conceptual model of the global ocean thermohaline circulation that consists of a system of ordinary differential equations modeling the heat and salt content of two containers ('boxes'). One box models the equatorial ocean, and the other models the extra-tropical ocean. The boxes exchange heat and freshwater with each other and with the atmosphere. The rate of flow between the boxes is proportional to the density difference between the boxes, and a major result of Stommel's investigation was that in some parameter regimes the system exhibits two equilibria: one analogous to the current climate, with dense cold water sinking at high latitudes, and one corresponding to a very different regime with dense salty water sinking in the equatorial ocean. In general, the goal of studies using extremely simplified models like Stommel's is to observe and understand qualitative features that might inform and guide subsequent studies using more complete and more complex models. The qualitative predictions of Stommel's model have since been verified using more complete ocean models, e.g. Rahmstorf (1995) and Deshayes *et al.* (2013).

The present investigation develops a model closely related to Stommel's where the slow, density-driven exchange of heat and salt between the boxes is augmented by fast, non-Gaussian stochastic processes representing eddy-driven heat and salt transport. Eddies smaller than the grid scale of comprehensive numerical ocean (and atmosphere) models can have significant impacts on the global circulation, and modeling the impacts of these unresolved eddies is a topic of continuing research; Berner *et al.* (2017) and Leutbecher *et al.* (2017) contain reviews

*Corresponding author. Email: ian.grooms@colorado.edu

of stochastic models of eddy effects from an operational modeling perspective.

The second author recently proposed a non-Gaussian model of the heat and salt transport associated with unresolved ocean eddies (Grooms 2016). In this model, the eddy velocity and density fields (the latter linearly related to temperature and salinity) are modeled as centered Gaussian random fields, and the transports are modeled as the product of eddy velocity and density. The product of centered, jointly Gaussian random variables has a distinctive, non-Gaussian probability density, with a logarithmic singularity at the origin and skewed, algebraically-modulated exponential decay in the tails. This non-Gaussian model is significantly different from recent Gaussian stochastic models of eddy transport, e.g. Andrejczuk *et al.* (2016), Williams *et al.* (2016) and Juricke *et al.* (2017). The present investigation is motivated by the desire to observe the qualitative effects of the kind of non-Gaussian transport from Grooms (2016) in an extremely simple model, in particular by comparison to Gaussian stochastic models, with the expectation of informing future investigations using more complex models.

A very wide range of stochastic parameterizations for ocean models of various resolutions with various kinds of Gaussian and non-Gaussian noise are currently under development, e.g. Porta Mana and Zanna (2014), Zanna *et al.* (2017), Mémin (2014), Resseguier *et al.* (2017), Grooms *et al.* (2015), Holm (2015), Cotter *et al.* (2017), Cooper (2017), and Brankart *et al.* (2015), in addition to those cited previously and many more too numerous to cite. The present study is intended to investigate the qualitative differences between a stochastic parameterization with a specific kind of non-Gaussian noise from Grooms (2016), a deterministic parameterization, and a Gaussian stochastic parameterization in a highly idealized model. As noted by Held (2005), the relationship of highly idealized models like the Stommel model to more complex and comprehensive climate models is analogous to the relationship between the fruit fly *Drosphila melanogaster* and *Homo sapiens*. Very few specific conclusions about the latter can be drawn from the former, but the study of the former is nevertheless invaluable in developing a broader understanding of generic features of biology.

Several authors have developed stochastic versions of Stommel's model to investigate the slow response of the ocean thermohaline circulation to fast atmospheric forcing, e.g. Cessi (1994), Vélez-Belchí *et al.* (2001), Monahan (2002), Monahan *et al.* (2002) and Monahan and Culina (2011). In these stochastic Stommel models the atmospheric heat and freshwater fluxes in Stommel's model are replaced by Gaussian stochastic noise terms, resulting in a system of stochastic differential equations (SDEs). The model developed here attempts to understand a qualitatively different physical process: fast eddy transport. Since the eddies are typically faster than the global thermohaline circulation, the new model has the form of a slow-fast system, where eddy variables evolve on a fast time scale and converge towards a jointly Gaussian distribution conditioned on the slow variables. The slow variables (the heat and salt difference between the boxes) are impacted by quadratic products of fast variables modeling the fast eddy transport. The formal theory of fast averaging (Papanicolaou and Kohler 1974, Pavliotis and Stuart 2008, Freidlin and Wentzell 2012), is used to generate approximate slow systems for comparison: one with a drift correction and one with both drift and diffusion corrections derived from the eddy dynamics. These approximate systems qualitatively represent more complete ocean models with, respectively, deterministic and Gaussian stochastic models of the eddy transport.

A new stochastic Stommel model including fast eddy transport is developed in §2. The two approximate models of the slow system are derived in §3. The numerical methods and experimental configuration are described in §4 and the results of these simulations are described in §5. A slightly different model with two stable equilibria is formulated and simulated in §6. The results and their implications are discussed in §7.

2. Formulating a Slow-Fast Two-Box Stochastic Ocean Model

This section recalls the derivation of the original Stommel model by considering the conservation of heat and salt in an ocean basin divided into two subdomains that exchange heat and salt with each other, and are forced by heat and freshwater fluxes from the atmosphere. The novel component of the derivation is to add stochastic eddy-driven fluxes between the subdomains. Consistent with the goal of this investigation the eddy-driven exchange is constructed as the product of centered Gaussian eddy velocity, heat, and salt anomalies; the flux distribution is thus qualitatively similar to the flux distributions recently observed by Grooms (2016). Naturally, other eddy flux models are possible; a Gaussian white noise model is, for example, derived in §3.

Consider a domain $[0, L_x] \times [0, L_y] \times [0, H]$ representing an ocean basin, and let this domain be partitioned into two subdomains $[0, L_x] \times [0, \ell] \times [0, H]$ and $[0, L_x] \times [\ell, L_y] \times [0, H]$ with volumes $V_1 = L_x \ell H$ and $V_2 = L_x(L_y - \ell)H$. The first box (index 1) will represent the equatorial side of the ocean basin, and the second (index 2) will represent the poleward side. The domain is filled with a fluid whose density is related to its temperature and salinity via

$$\rho = \rho_0[1 + \alpha_S(S - S_0) - \alpha_T(T - T_0)]$$

where $\rho_0 = 1029$ kg/m^3 is a constant reference density, $T_0 = 5$ C and $S_0 = 35$ psu are a constant reference temperature and salinity (psu are practical salinity units; for the present purposes it is reasonable to use the simplification 1 psu = 1 g/kg), and $\alpha_S = 7.5 \times 10^{-4}$ psu^{-1} and $\alpha_T = 1.7 \times 10^{-4}$ C^{-1} are coefficients of haline and thermal expansion. The conservation equations for heat are in the form of a system of two differential equations

$$\frac{\mathrm{d}T_1}{\mathrm{d}t} = -\frac{1}{\tau_T}(T_1 - T_1^*) - \frac{F_T}{\rho_0 c_p V_1}, \quad \frac{\mathrm{d}T_2}{\mathrm{d}t} = -\frac{1}{\tau_T}(T_2 - T_2^*) + \frac{F_T}{\rho_0 c_p V_2}$$

where T_1 and T_2 are the mean temperature in each box, τ_T is the timescale of relaxation towards an externally-specified atmospheric temperature T_i^*, $c_p = 4000$ J/kg is the heat capacity of seawater (e.g. $\rho_0 c_p V_1 T_1$ is the heat content of the equatorial box), and F_T is the heat flux from the equatorial box to the poleward box. The total heat content $\rho_0 c_p(V_1 T_1 + V_2 T_2)$ thus depends only on the external forcing.

Similarly, the conservation equations for salt are

$$\frac{\mathrm{d}S_1}{\mathrm{d}t} = \frac{1}{2}F(t) - F_S, \quad \frac{\mathrm{d}S_2}{\mathrm{d}t} = -\frac{1}{2}F(t) + F_S$$

where $F(t)/2$ is the external freshwater forcing in the equatorial box (e.g. rain, runoff, evaporation) and F_S is the salt flux from the equatorial box to the poleward box. The external freshwater forcing is assumed not to change the net salt content, so that $S_1 + S_2$ remains constant in time.

Following Stommel (1961), the heat and salt fluxes between the boxes are assumed to depend only on the temperature and salinity differences between the boxes. As a result, the temperature and salinity differences between the boxes decouple from the net heat and salt content. Defining $\Delta T = T_1 - T_2$ and $\Delta S = S_1 - S_2$,

$$\frac{\mathrm{d}\Delta T}{\mathrm{d}t} = -\frac{1}{\tau_T}(\Delta T - \Delta T^*) - \left(\frac{1}{\rho_0 c_p V_1} + \frac{1}{\rho_0 c_p V_2}\right) F_T$$

$$\frac{\mathrm{d}\Delta S}{\mathrm{d}t} = F(t) - 2F_S.$$

Similar to Cessi (1994) and Vélez-Belchí *et al.* (2001), the atmospheric temperature difference ΔT^* and external freshwater forcing $F(t)$ are here modeled as constant mean terms plus

Gaussian white noise, leading to

$$
\mathrm{d}\Delta T = \left[-\frac{1}{\tau_T}(\Delta T - \overline{\Delta T^*}) - \left(\frac{1}{\rho_0 c_p V_1} + \frac{1}{\rho_0 c_p V_2}\right) F_T\right] \mathrm{d}t + \frac{\sigma_{\Delta T}}{\sqrt{\tau_T}}\mathrm{d}W_{\Delta T}
$$

$$
\mathrm{d}\Delta S = \left[\overline{F} - 2F_S\right] \mathrm{d}t + \frac{\sigma_{\Delta S}}{\sqrt{\tau_d}}\mathrm{d}W_{\Delta S}\,.
$$

The amplitude of the atmospheric heat flux noise forcing is here scaled by $\sqrt{\tau_T}$ so that it generates temperature perturbations of amplitude $\sigma_{\Delta T}$ over a time period of length τ_T; the atmospheric freshwater flux noise is similarly scaled to generate perturbations of amplitude $\sigma_{\Delta S}$ over a diffusive time τ_d, defined below.

In Stommel's original model the fluxes between the boxes consist of diffusive fluxes proportional to the temperature and salinity differences, and advective fluxes associated with the large-scale ocean circulation whose rate is proportional to the magnitude of the density difference between the boxes

$$
\left(\frac{1}{\rho_0 c_p V_1} + \frac{1}{\rho_0 c_p V_2}\right) F_T = \left(\frac{1}{\tau_d} + \frac{1}{\tau_a \rho_0 \alpha_T \overline{\Delta T^*}}|\Delta\rho|\right) \Delta T
$$

$$
2F_S = \left(\frac{1}{\tau_d} + \frac{1}{\tau_a \rho_0 \alpha_T \overline{\Delta T^*}}|\Delta\rho|\right) \Delta S
$$

where τ_d is the time scale of diffusive transport, τ_a is the time scale of advective transport, and

$$
\Delta\rho = \rho_0[\alpha_S \Delta S - \alpha_T \Delta T]
$$

is the density difference between the boxes. Cessi (1994) used a smoother formulation, which does not qualitatively change the results

$$
\left(\frac{1}{\rho_0 c_p V_1} + \frac{1}{\rho_0 c_p V_2}\right) F_T = \left(\frac{1}{\tau_d} + \frac{1}{\tau_a (\rho_0 \alpha_T \overline{\Delta T^*})^2}\Delta\rho^2\right) \Delta T
$$

$$
2F_S = \left(\frac{1}{\tau_d} + \frac{1}{\tau_a (\rho_0 \alpha_T \overline{\Delta T^*})^2}\Delta\rho^2\right) \Delta S.
$$

The novel contribution to the model made here consists of the addition of fast variables crudely representing eddy velocity v_e, temperature T_e and salinity S_e anomalies at the interface between the boxes. The eddy-induced fluxes between the boxes will be modeled as an addition to the slow diffusive and advective fluxes

$$
\left(\frac{1}{\rho_0 c_p V_1} + \frac{1}{\rho_0 c_p V_2}\right) F_T = \left(\frac{1}{\tau_d} + \frac{1}{\tau_a (\rho_0 \alpha_T \overline{\Delta T^*})^2}\Delta\rho^2\right) \Delta T + \left(\frac{1}{\ell} + \frac{1}{L_y - \ell}\right) v_e T_e
$$

$$
2F_S = \left(\frac{1}{\tau_d} + \frac{1}{\tau_a (\rho_0 \alpha_T \overline{\Delta T^*})^2}\Delta\rho^2\right) \Delta S + \left(\frac{1}{\ell} + \frac{1}{L_y - \ell}\right) v_e S_e.
$$

The prefactors $\ell^{-1} + (L_y - \ell)^{-1}$ account for the fact that the boxes need not have equal volume, and that total heat and salt need to be conserved. For simplicity, only $\ell = L_y/2$ is considered from here on.

In general the flux between the boxes should be described by $\int_0^{R_L}\int_0^{R_H} vT\,\mathrm{d}z\,\mathrm{d}x$ where v and T are evaluated at $y = \ell = L_y/2$. Our formulation amounts to a severe simplification that ignores the spatial structure of the eddy velocity and temperature perturbations between the boxes, and considers them only as zero-mean jointly-Gaussian variables. This level of simplification is consistent with the simplification of the ocean to two well-mixed boxes in the original Stommel model, and is guided by the desire to investigate the qualitative effects of Gaussian-product noise, since eddy noise with this structure was recently observed by Grooms (2016).

The fast eddy velocity will be modeled as an Ornstein-Uhlenbeck process

$$\mathrm{d}v_e = -\frac{1}{\tau_e}v_e\mathrm{d}t + \sqrt{\frac{2}{\tau_e}}\sigma_v\mathrm{d}W_v$$

where τ_e is the eddy time scale and σ_v is the eddy velocity scale, chosen to be 15 days and 10 cm/s, respectively (Stammer 1997). The eddy velocity can here be thought of as being set by wind-driven processes independent of the density difference between the boxes. This is a simplification of the more complex reality where eddy kinetic energy and time scale depend also on the large-scale density gradient. The following model of the eddy dynamics is perhaps more qualitatively appropriate

$$\mathrm{d}v_e = -\frac{v_e}{\tau_e}\mathrm{d}t + \sqrt{\frac{2(1+\mu\Delta\rho^2)}{\tau_e}}\mathrm{d}W_v$$

where $\mu > 0$ is a parameter representing the sensitivity of the eddy variance to the large-scale density gradient. This model is not pursued further here, in part because of the difficulties in guaranteeing its ergodicity and in finding a robust numerical method for its solution.

The eddy temperature and salinity anomalies will be modeled as resulting from eddy transport across the large-scale gradients

$$\frac{\mathrm{d}T_e}{\mathrm{d}t} = -\frac{T_e}{\tau_e} - v_e\frac{2\Delta T}{L_y},$$

$$\frac{\mathrm{d}S_e}{\mathrm{d}t} = -\frac{S_e}{\tau_e} - v_e\frac{2\Delta S}{L_y}.$$

The relaxation towards zero on a time scale τ_e qualitatively represents the full range of dissipative processes acting on temperature and salinity anomalies: cascade towards small scales and eventual diffusion, and atmospheric damping of thermal anomalies, etc. The time scale τ_e should not be associated with any particular physical process, but instead guarantees decorrelation of eddy anomalies on the time scale τ_e. Note that the lack of white noise forcing in the equations for T_e and S_e implies that the amplitude of the eddy terms is governed by σ_v; if $\sigma_v = 0$ then the eddy terms disappear, leaving the usual Stommel model.

The governing equations are nondimensionalized using the diffusive time scale for t, the external constant atmospheric temperature difference $\overline{\Delta T^*} \approx 20^\circ$C for large-scale temperature, and the convenient salinity scale $\alpha_T\overline{\Delta T^*}/\alpha_S \approx 4.5$ psu for large-scale salinity. The mean atmospheric forcing $\bar{F}$ is assumed to be 4.5 psu per diffusion time so that its nondimensional value is 1, following Cessi (1994) and Vélez-Belchí *et al.* (2001). The eddy velocity v_e is nondimensionalized using the eddy velocity scale σ_v. It will be convenient to scale the eddy temperature and salinity variables differently; specifically, T_e will have dimensions $\overline{\Delta T^*}L_y/(\sigma_v\tau_d)$ and S_e will have dimensions $\alpha_T\overline{\Delta T^*}L_y/(\alpha_S\sigma_v\tau_d)$. The reason for this unexpected scaling will be commented on shortly.

Following traditional notation, the nondimensional temperature difference will be denoted x and the nondimensional salt difference will be denoted y. The nondimensional eddy variables will drop their subscripts e so that, e.g., the nondimensional eddy velocity is simply v. Risking confusion, the nondimensional time will still be denoted t. The complete nondimensional system is therefore

$$\mathrm{d}x = \left(-\frac{1}{\eta_T}(x-1) - [1 + P_a^2(x-y)^2]x + 4vT\right)\mathrm{d}t + \sqrt{\frac{1}{\eta_T}}\sigma_x \mathrm{d}W_x \tag{1a}$$

$$\mathrm{d}y = \left(1 - [1 + P_a^2(x-y)^2]y + 4vS\right)\mathrm{d}t + \sigma_y \mathrm{d}W_y \tag{1b}$$

$$\mathrm{d}v = -\frac{v}{\xi}\mathrm{d}t + \sqrt{\frac{2}{\xi}}\mathrm{d}W_v \tag{1c}$$

$$\mathrm{d}T = -\frac{1}{\xi}\left(T + 2P_e^2 vx\right)\mathrm{d}t \tag{1d}$$

$$\mathrm{d}S = -\frac{1}{\xi}\left(S + 2P_e^2 vy\right)\mathrm{d}t \tag{1e}$$

where

$$\eta_T = \frac{\tau_T}{\tau_d},\quad \xi = \frac{\tau_e}{\tau_d},\quad P_a = \frac{\tau_d}{\tau_a},\quad P_e = \frac{\sigma_v \tau_d}{L_y},\quad P = \sqrt{P_e}.$$

P_a and P_e are Péclet numbers comparing the time scales of large-scale advective transport and fast eddy transport to the time scale of diffusion, respectively. The nondimensional noise amplitudes are $\sigma_x = \sigma_{\Delta T}/\overline{\Delta T^*}$ and $\sigma_y = \alpha_S \sigma_{\Delta S}/(\alpha_T \overline{\Delta T^*})$.

The following parameter estimates are drawn from Cessi (1994) and Vélez-Belchí *et al.* (2001), and are consistent with the more recent observational analysis of Schmitt (2008). The diffusive time scale τ_d is approximately 220 years, and the time scale of large scale advection τ_a is approximately 35 years. Cessi (1994) estimates τ_T to be 25 days, but Vélez-Belchí *et al.* (2001) argue convincingly that large-scale temperature anomalies are damped on a slower time scale of approximately 220 days. Vélez-Belchí *et al.* (2001) used salinity noise whose nondimensional amplitude is here $\sigma_y = 0.15$, and assuming that fast atmospheric temperature fluctuations lead to perturbations on the order of 0.07 C implies nondimensional thermal noise has amplitude $\sigma_x = 0.005$. Finally, using a length scale appropriate to the global oceans, $L_y \approx 8,250$ km leads to the following set of parameters which are adopted for the remainder of the investigation

$$\eta_T = \frac{1}{400},\quad \xi = \frac{1}{5000},\quad P_a = 6,\ P_e = 80,\ \sigma_x = 0.005,\ \sigma_y = 0.15. \tag{2}$$

The reason for scaling S and T differently from ΔS and ΔT should now be clear: $2P_e^2$ is the same order of magnitude as ξ^{-1}, implying that both terms in the evolution equations for S and T are of comparable magnitude.

For the parameters (2) the system (1) has three equilibria, two of which are stable. The equilibria all have $v, T, S = 0$, and the stable equilibria occur at $(x, y) \approx (0.989, 0.22)$ and $(x, y) \approx (0.998, 1.00)$. In the absence of eddy dynamics, one would expect small atmospheric noise to lead to jumping between the two stable equilibria of the system; this was the focus of Cessi (1994), Monahan (2002), Monahan *et al.* (2002) and Monahan and Culina (2011).

The existence of multiple equilibria is intrinsically tied to the nonlinear terms that model slow advective exchange between the boxes. As the exchange between the boxes becomes dominated by diffusion instead of advection ($P_a \to 0$) one of the stable equilibria disappears in a reverse saddle-node bifurcation leaving a single stable equilibrium.

Equations (1d) and (1e) lack noise terms, implying that the classical conditions for ergodicity (Khasminskii 2012) do not apply. Conditions for ergodicity of this type of system of SDEs can be found in Mattingly *et al.* (2002). The first condition is that there is an inner-product norm $\|\cdot\|$ such that $\langle \boldsymbol{u}, \boldsymbol{F}(\boldsymbol{u})\rangle \leq \alpha - \beta\|\boldsymbol{u}\|^2$ for some $\alpha, \beta > 0$ where $\boldsymbol{u}$ is a vector containing the dependent variables and $\boldsymbol{F}(\boldsymbol{u})$ is the drift. It is straightforward to verify that $\|\boldsymbol{u}\|^2 = x^2+y^2+v^2+(2/P^2)(T^2+S^2)$ satisfies this condition. The second condition is that the vectors $\{\boldsymbol{\rho}_i, [[\boldsymbol{F}, \boldsymbol{\rho}_j], \boldsymbol{\rho}_k]\}$ span $\mathbb{R}^5$ where $\boldsymbol{\rho}_i$ $i = 1, 2, 3$ are the columns of the diffusion matrix, which are here proportional to the first three standard basis vectors, and $[\cdot, \cdot]$ is a Lie bracket. Since $[[\boldsymbol{F}, \boldsymbol{\rho}_1], \boldsymbol{\rho}_3]$ and $[[\boldsymbol{F}, \boldsymbol{\rho}_2], \boldsymbol{\rho}_3]$ are proportional to the fourth and fifth standard basis vectors, respectively, the system satisfies the conditions of Mattingly *et al.* (2002) for ergodicity.

3. Two Approximate Slow Models

In this section two systems of SDEs are derived approximating the evolution of the slow variables x and y in (1). The system of SDEs (1) with parameters (2) has three time scales since $\epsilon < \epsilon_T \ll 1$: x evolves significantly more quickly than y, yet slower than the eddy variables v, T, and S. Many previous investigations (which lacked the eddy variables) accounted for the scale separation somewhat crudely by setting $x = 1$, and focused on the dynamics of the slowest variable y, e.g. Cessi (1994), Monahan (2002), Monahan *et al.* (2002), and Monahan *et al.* (2008). The analysis of Monahan and Culina (2011) is more careful, employing the same methods used here but for the system without eddy variables and in the limit $\epsilon_T \to 0$. This section considers the limit $\epsilon \to 0$ while holding ϵ_T fixed.

The two approximate models are derived using standard approximations for slow-fast systems (Papanicolaou and Kohler 1974, Pavliotis and Stuart 2008, Freidlin and Wentzell 2012). The presentation here follows the convenient review found in Bouchet *et al.* (2016); the formulas are derived in a straightforward manner using formal asymptotic methods applied to the backwards Kolmogorov equation for the system (for details, see the appendices of Bouchet *et al.* (2016)).

The first approximation is derived via simple averaging. In the limit $\epsilon \to 0$ the eddy variables are well approximated as solutions to (1c)–(1e) with x and y considered constant. Curiously, although the full system (1) has a smooth invariant measure the system (1c)–(1e) does not: the long-time limiting distribution of v, T, and S is jointly Gaussian with a singular covariance matrix. In light of this, the following noise-augmented system is considered instead

$$\mathrm{d}v = -\frac{v}{\epsilon}\mathrm{d}t + \sqrt{\frac{2}{\epsilon}}\mathrm{d}W_v \tag{3a}$$

$$\mathrm{d}T = -\frac{1}{\epsilon}\left(T + 2P^2vx\right)\mathrm{d}t + \sqrt{\frac{2}{\epsilon}}\sigma\,\mathrm{d}W_T \tag{3b}$$

$$\mathrm{d}S = -\frac{1}{\epsilon}\left(S + 2P^2vy\right)\mathrm{d}t + \sqrt{\frac{2}{\epsilon}}\sigma\,\mathrm{d}W_S \tag{3c}$$

and the limit $\sigma \to 0$ is taken after the fact.

The invariant measure of (3) is Gaussian with zero mean and covariance

$$\begin{bmatrix} 1 & -P^2x & -P^2y \\ -P^2x & 2P^4x^2+\sigma^2 & 2P^4xy \\ -P^2y & 2P^4xy & 2P^4y^2+\sigma^2 \end{bmatrix}. \tag{4}$$

The averages of the terms vT and vS in the slow equations with respect to the invariant measure of the fast system are simply $-P^2x$ and $-P^2y$, respectively. It is worth noting that these values are independent of the auxiliary noise amplitude σ. Inserting these into the slow equations leads to the following approximate model

Deterministic Approximation

$$\mathrm{d}x = \left(-\frac{1}{T}(x-1) - [1 + P_a(x-y)^2]x - 4P^2x\right)\mathrm{d}t + \sqrt{\frac{1}{\epsilon T}}\sigma_x \mathrm{d}W_x \tag{5a}$$

$$\mathrm{d}y = \left(1 - [1 + P_a(x-y)^2]y - 4P^2y\right)\mathrm{d}t + \sigma_y \mathrm{d}W_y. \tag{5b}$$

The model (5) is referred to as the 'deterministic' or 'averaged' approximation since it models the eddy terms vT and vS as deterministic functions of x and y. It is straightforward to verify that this model is ergodic under the classical conditions of Khasminskii (2012).

As described in Bouchet *et al.* (2016), one can derive equations that approximate the variations of the true solution to (1) around the solution of the approximate model (5). Combining the equations for the variations with the deterministic approximation leads to further corrections in both the drift and diffusion, of order ϵ and $\sqrt{\epsilon}$, respectively. The drift correction is significantly smaller than the leading-order drift. But the leading-order diffusion terms in the x and y equations are of order ≈ 0.1, and corrections of order $\sqrt{\epsilon}$ may be of comparable magnitude.

In order to compute the diffusion corrections, it is convenient to define some notation. Let $\boldsymbol{Y} = (v, T, S)^T$ denote the solution to the noise-augmented system (3). Define constant matrices

$$\mathbf{M} = -\frac{1}{\epsilon}\begin{bmatrix} 1 & 0 & 0 \\ 2P^2x & 1 & 0 \\ 2P^2y & 0 & 1 \end{bmatrix}, \quad \mathbf{G} = \sqrt{\frac{2}{\epsilon}}\begin{bmatrix} 1 & 0 & 0 \\ 0 & \sigma & 0 \\ 0 & 0 & \sigma \end{bmatrix}$$

such that the fast system (3) may be written $\mathrm{d}\boldsymbol{Y} = \mathbf{M}\boldsymbol{Y} + \mathbf{G}\mathrm{d}\boldsymbol{W}$, where $\mathrm{d}\boldsymbol{W}$ is a vector of independent Gaussian white noises. The solution is thus

$$\boldsymbol{Y}(\tau) = e^{\mathbf{M}\tau}\boldsymbol{Y}_0 + \int_0^\tau e^{\mathbf{M}(\tau-s)}\mathbf{G}\mathrm{d}\boldsymbol{W}. \tag{6}$$

The deviations of the eddy terms vT and vS from their conditional means are denoted

$$\boldsymbol{f}(x, y, \boldsymbol{Y}) = \begin{pmatrix} vT + 4P^2x \\ vS + 4P^2y \end{pmatrix}.$$

According to Bouchet *et al.* (2016), the diffusion-corrected model for the slow variables has the form

$$\begin{aligned}\mathrm{d}x = {} & \left(-\frac{1}{T}(x-1) - [1 + P_a(x-y)^2]x - 4P^2x\right)\mathrm{d}t \\ & + \sqrt{\bar{a}_{xx}(x,y)}\mathrm{d}\hat{W}_x + \sqrt{\bar{a}_{xy}(x,y)}\mathrm{d}\hat{W}_y + \sqrt{\frac{1}{\epsilon T}}\sigma_x\mathrm{d}W_x \\ \mathrm{d}y = {} & \left(1 - [1 + P_a(x-y)^2]y - 4P^2y\right)\mathrm{d}t \\ & + \sqrt{\bar{a}_{yx}(x,y)}\mathrm{d}\hat{W}_x + \sqrt{\bar{a}_{yy}(x,y)}\mathrm{d}\hat{W}_y + \sigma_y\mathrm{d}W_y.\end{aligned}$$

where the matrix

$$\mathbf{A} = \begin{pmatrix} a_{xx} & a_{xy} \\ a_{yx} & a_{yy} \end{pmatrix}$$

is any square root of the following symmetric positive definite matrix

$$\mathbf{C} = \int_0^\infty \mathrm{E}^{\boldsymbol{Y}_0}\, \mathrm{E}^{\boldsymbol{Y}(\tau)} \left[\boldsymbol{f}(x, y, \boldsymbol{Y}(\tau))\boldsymbol{f}^T(x, y, \boldsymbol{Y}_0) + \boldsymbol{f}(x, y, \boldsymbol{Y}_0)\boldsymbol{f}^T(x, y, \boldsymbol{Y}(\tau)) \right] \mathrm{d}\tau.$$

The matrix **C** is the integral of the time-lagged auto-covariance of $\boldsymbol{f}$ with x and y considered constant. In the above expression, $\mathrm{E}^{\boldsymbol{Y}(\tau)}$ denotes the expectation on $\boldsymbol{Y}(\tau)$ conditioned on the initial condition $\boldsymbol{Y}_0$; the distribution is Gaussian with mean and covariance implied by (6). $\mathrm{E}^{\boldsymbol{Y}_0}$ denotes expectation on $\boldsymbol{Y}_0$ whose distribution is the stationary distribution of the fast process, in this case a zero-mean Gaussian with covariance (4). The calculation for the system under consideration here is particularly straightforward since it requires only higher moments of jointly-Gaussian variables. The matrix **C** is found to have the form

$$\mathbf{C} = \begin{pmatrix} 16(5P^4x^2 + \sigma^2) & 80P^4xy \\ 80P^4xy & 16(5P^4y^2 + \sigma^2) \end{pmatrix}.$$

In this case (unlike the leading-order drift term) the limit $\sigma \to 0$ is singular in the sense that the matrix **C** becomes positive semi-definite. Nevertheless, a square root matrix **A** exists; in the limit $\sigma \to 0$ it has the form

$$\mathbf{A} = 4\sqrt{5}P^2 \begin{pmatrix} x & 0 \\ y & 0 \end{pmatrix}.$$

The model for the slow variables with leading-order drift and diffusion corrections (but ignoring the order- drift correction) is thus

Gaussian Stochastic Approximation

$$\mathrm{d}x = \left[\frac{1}{\tau}(1 - x) - [1 + a(x - y)^2]x - 4P^2x\right]\mathrm{d}t + 4\sqrt{5}P^2x\,\mathrm{d}\hat{W} + \sqrt{\frac{1}{\tau}}\sigma_x\,\mathrm{d}W_x \qquad (7a)$$

$$\mathrm{d}y = \left[1 - [1 + a(x - y)^2]y - 4P^2y\right]\mathrm{d}t + 4\sqrt{5}P^2y\,\mathrm{d}\hat{W} + \sigma_y\,\mathrm{d}W_y. \qquad (7b)$$

For $x \approx 1$ the noise amplitude associated with the eddies is ≈ 0.16, which is slightly larger than the 'atmospheric' noise $\sigma/\sqrt{\tau} = 0.1$. The order- drift corrections have also been calculated, but they are small in comparison with the leading-order terms, and have been left out of the model for simplicity. This system of SDEs is interpreted in the Ito sense; while the drift corrections in slow-fast systems with one slow degree of freedom can be interpreted as a correction from Stratonovich to Ito, this is no longer generally true in systems with multiple slow degrees of freedom (Pavliotis and Stuart 2008, Freidlin and Wentzell 2012). It is straightforward to verify that this model is ergodic under the classical conditions of Khasminskii (2012).

It is interesting to note that the Gaussian stochastic model replaces the eddy terms $4vT$ and $4vS$ by $-4P^2x(\mathrm{d}t + \sqrt{5}\,\mathrm{d}\hat{W})$ and $-4P^2y(\mathrm{d}t + \sqrt{5}\,\mathrm{d}\hat{W})$. This form of subgrid-scale parameterization is qualitatively the same as that proposed in Buizza *et al.* (1999), where it was proposed to multiply a deterministic parameterization (here $-4P^2x$) by a stochastic process (here $1 + \sqrt{5}\,\dot{\hat{W}}$). This style of stochastic parameterization has been widely used in atmospheric models (Berner *et al.* (2017) provides a review), and much has been made of the role of multiplicative noise by, e.g. Sura *et al.* (2005). The above derivation gives an example where this style of *ad hoc* parameterization is rigorously justified, though multiplicative noise with a linear coefficient is certainly not the universal form of eddy-induced noise (see e.g. Monahan and Culina 2011, for a counterexample).

Recall that for the parameters (2) the system (1) has only three equilibria, two of which are stable. The equilibria all have $v, T, S = 0$, and the stable equilibria occur at $(x, y) \approx (0.989, 0.22)$ and $(x, y) \approx (0.998, 1.00)$. The deterministic and Gaussian stochastic models have the same drift, which has only one equilibrium at $(x, y) \approx (0.974, 0.093)$. As will be verified by the results in §5, the inclusion of nonlinear eddy effects completely changes the regime of the ocean model from a regime of multiple equilibria to a regime with a single stable equilibrium.

The averaged drift has a single stable equilibrium for all P greater than approximately 0.117; below this value the drift undergoes a saddle-node bifurcation that creates a pair of equilibria near $x = 1$ and $y = 1$. To achieve such small values of P would require reducing the eddy velocity scale from 10 cm/s to 1 cm/s, which is unrealistically small. The approximate models derived in this section show that the mean effect of eddies is linear and diffusive. Since a linear diffusive effect is already present in the equations (the terms $-x$ and $-y$ in (1a) and (1b)), the mean eddy effect could be viewed as a double-counting of eddy-induced diffusive exchange between the boxes. This can be rectified by eliminating the mean diffusion terms, and such a model is formulated and studied in §6. By avoiding a double-counting of diffusive exchange, the model in §6 allows multiple equilibria with small, yet realistic eddy amplitudes.

4. Numerical Methods

Numerical methods are needed to compare the qualitative behavior of the three models (1), (5), and (7) . Many methods are derived based on the assumption that the drift is globally-Lipschitz (Kloeden and Platen 1992), which is not the case here. Several more recent investigations have analyzed numerical methods for SDEs whose drift satisfies a one-sided Lipschitz condition (e.g. Higham *et al.* (2002) and Mao and Szpruch (2013)), but none of the models in consideration here satisfy such a condition. A method appropriate to polynomial drifts is derived by Lamba *et al.* (2007), but their analysis requires an invertible diffusion matrix, which the model (1) does not have. The Euler-Maruyama method may be appropriate, but is known to behave poorly in problems with polynomial drift (Mattingly *et al.* 2002, Hutzenthaler *et al.* 2011). In light of this, the 'backward Euler' (BE) method is used here for all three models. For a general system of SDEs of the form

$$d\boldsymbol{X} = \boldsymbol{b}(\boldsymbol{X})\mathrm{d}t + \boldsymbol{\Sigma}(\boldsymbol{X})\mathrm{d}\boldsymbol{W}$$

the BE method takes the following form

$$\boldsymbol{X}_{n+1} - \Delta t \boldsymbol{b}(\boldsymbol{X}_{n+1}) = \boldsymbol{X}_n + \boldsymbol{\Sigma}(\boldsymbol{X}_n)\Delta \boldsymbol{W}_n \tag{8}$$

where Δt is the time step. In every simulation presented here $\Delta t = 2 \times 10^{-6}$, which is significantly smaller than the smallest time scale of the system $\approx 2 \times 10^{-4}$. Mattingly *et al.* (2002) prove that the method is ergodic (for sufficiently small Δt) and that the invariant measure of the numerical method converges to that of the SDE as $\Delta t \to 0$. Though the analysis of Mattingly *et al.* (2002) focuses on models with additive noise, the BE method is nevertheless applied here to the model (7) with multiplicative noise.

For the model (1), a two-step process is used to generate solutions of the nonlinear system of equations (8). First, an asymptotic approximation in the limit $\Delta t \to 0$ is computed that has the form $\boldsymbol{X}_* = \boldsymbol{X}_n + \boldsymbol{\Sigma}(\boldsymbol{X}_n)\Delta \boldsymbol{W}_n + O(\Delta t)$; this approximation is followed by a single Newton step. For the systems (5) and (7), approximate solutions to the nonlinear systems were generated using 10 fixed-point iterations started at $\boldsymbol{X}_n$. Given the small step size, the resulting approximations solve their respective nonlinear systems with high accuracy;

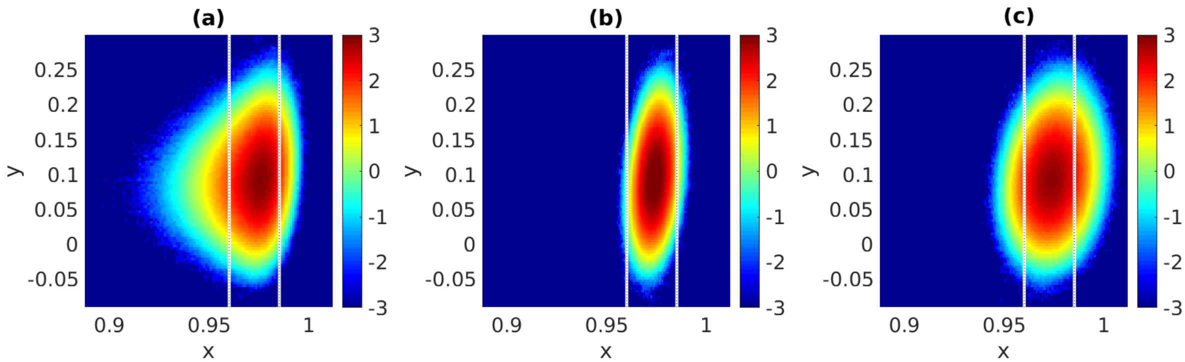

Figure 1. Base-10 logarithm of the climatological joint probability density functions of x and y for (a) Full model (1), (b) Deterministic approximation (5), and (c) Gaussian-stochastic approximation (7). The vertical lines are placed at $x = 0.96$ and $x = 0.985$.

residuals are typically on the order of 10^{-11}.

5. Results

5.1. *Climatology*

A suite of 10,000 independent simulations was run starting from $x, y, v, T, S = 0$ at $t = 0$. Data were saved for the time interval $t \in [4, 10]$, saving every 100th time step for a spacing of 2×10^{-4}. The mean and covariance appeared to have stabilized by $t = 4$, suggesting that the data in $t \in [4, 10]$ represents the stationary climatological distribution of the system. Recalling that the dimensional time unit is 220 years, this amounts to 1,320 years of data saved approximately twice per month. The three models all have the same mean of $(x, y) \approx (0.974, 0.094)$, which is very close to the equilibrium of the deterministic and Gaussian stochastic models at (0.974, 0.093). All three models have the same marginal standard deviation of y approximately equal to 0.034. This can be explained by the fact that the amplitude of the eddy noise in the y equation is estimated in the Gaussian stochastic model to be $\sqrt{5\sigma}P^2 y/4 \approx 0.015$ for $y = 0.093$, which is much less than the atmospheric noise with amplitude $\sigma_y = 0.15$. The parameter values (2) derived from the literature are necessarily imprecise, but the order of magnitude difference between the eddy noise and the atmospheric noise in the y equation suggests that the effects of eddy noise (Gaussian or otherwise) on the salinity dynamics of the real ocean may be small in comparison with atmospheric forcing.

The climatological distributions of the models differ in other respects. For example, the marginal standard deviation of x is 0.0063, 0.0035, and 0.0065 in the full, deterministic, and Gaussian stochastic models, respectively. The eddy noise in the x equation is of comparable size to the atmospheric noise, and has a significant impact on the variability; the deterministic model lacks this eddy noise, and has too little variability. The lack of eddy noise in the x equation of the deterministic model also leads to an overestimate of the correlation between x and y: the full and Gaussian stochastic models have correlations 0.15 and 0.14, respectively, while the deterministic model has correlation 0.23. The most-probable values of the distributions are $(x, y) \approx (0.976, 0.092)$ for the full model, (0.974, 0.091) for the deterministic model, and (0.973, 0.093) for the Gaussian stochastic model; the differences in the y value are negligible, but the differences in the x value are up to half of a standard deviation.

Time-lagged correlation functions were computed, for example $\text{Corr}[x(t), x(t + \tau)] = C(\tau)$ (stationarity is assumed). The correlation functions are all very similar across the models (not shown). The correlation functions all decay monotonically to zero, so it is natural to define a decorrelation time by $\int_0^\infty C(\tau)\,d\tau$. The correlation functions for y in all three models are very similar, with decorrelation time approximately 22 years. The correlation function for x exhibits similar rapid initial decay in all three models. The correlation function for x in the deterministic model has a long tail, with larger long-lag correlations than the other two models, leading to a decorrelation time of 1.6 years, which is longer than the decorrelation times of the full model and Gaussian stochastic model, both of which are approximately 1 year.

A simple binning procedure was used to generate approximations to the climatological probability density function (pdf) for each model; results are shown in Fig. 1, with panels (a)–(c) presenting the full model, deterministic model, and Gaussian stochastic model, respectively. It has already been noted that the three models have the same marginal variance for y, and indeed the range of y in the three models is quite similar. The deterministic model is clearly under-dispersed with respect to x. The climatological distribution of the Gaussian stochastic model has a more-accurate core, but is not skewed in the same way as the full model.

It is possible that minor deficiencies near the core of the distribution could be corrected by adding order- corrections to the drift of the Gaussian stochastic model, but the results of Bouchet *et al.* (2016) indicate that such corrections will not generate correct rare-event probabilities even in the limit $\to 0$. To emphasize differences in the rare event probabilities, the probabilities of $x \leq 0.96$ and $x \geq 0.985$ were calculated for the three models (these x values are indicated by vertical lines in Fig. 1). The small-event probabilities are 0.039 for the full model, less than 10^{-4} for the deterministic model, and 0.022 for the Gaussian stochastic model. The large-event probabilities are 0.016 for the full model, less than 10^{-3} for the deterministic model, and 0.048 for the Gaussian stochastic model. Not surprisingly, the deterministic approximation has too-small rare event probabilities. The Gaussian-stochastic model is more accurate, but is still incorrect by nearly a factor of 2 for small-event probabilities, and a factor of 3 for large-event probabilities.

The system (1) has two stable equilibria, near $(x, y) \approx (1, 1)$ and $(1, 0.22)$. The simulations described above had no trajectories near the stable equilibrium at (1,1); to verify that the system does not remain near the stable equilibrium of (1) at $(x, y) \approx (1, 1)$, a set of 1,000 simulations of (1) was run with initial condition $(x, y, v, T, S) = (1, 1, 0, 0, 0)$. These simulations were again run for the interval $t \in [0, 10]$, saving the output from $t \in [4, 10]$. The stationary distribution did not display a secondary peak near (1, 1), indicating that the two stable equilibria of the full model are largely irrelevant to the dynamics of the system.

5.2. *Rare event forecasting*

The previous section examined only the stationary climatological distributions of the three models. Within a climate prediction scenario, short-term behavior is also important. Given that the climatological distributions differ mainly in their rare event probabilities, a separate set of experiments was used to investigate the ability of the models to predict rare events over a shorter time interval. The goal was to test how accurately the approximate models forecast the probability of the unusually large and small x values over a range of forecast lead times. Two trajectories of the system (1) were selected out of the 10,000 discussed above: one reaching a value of $x \leq 0.96$ and one reaching a value of $x \geq 0.985$. These trajectories are shown in Fig. 2 panels (a) and (b). Note that the large-x trajectory passes the threshold of 0.985 approximately half a year before the final time, whereas the small-x trajectory crosses the 0.96 threshold only at the last time step. Ensembles of 10,000 independent forecasts for all three models were initialized from the true trajectory for a range of lead times out to 2 years. Thus, for each of the three models a 10,000 member ensemble forecast was initialized at $t = -2$ years and run until $t = 0$, and another 10,000 member ensemble forecast was initialized at $t = -1$ year and run until $t = 0$, etc. These ensembles were used to estimate

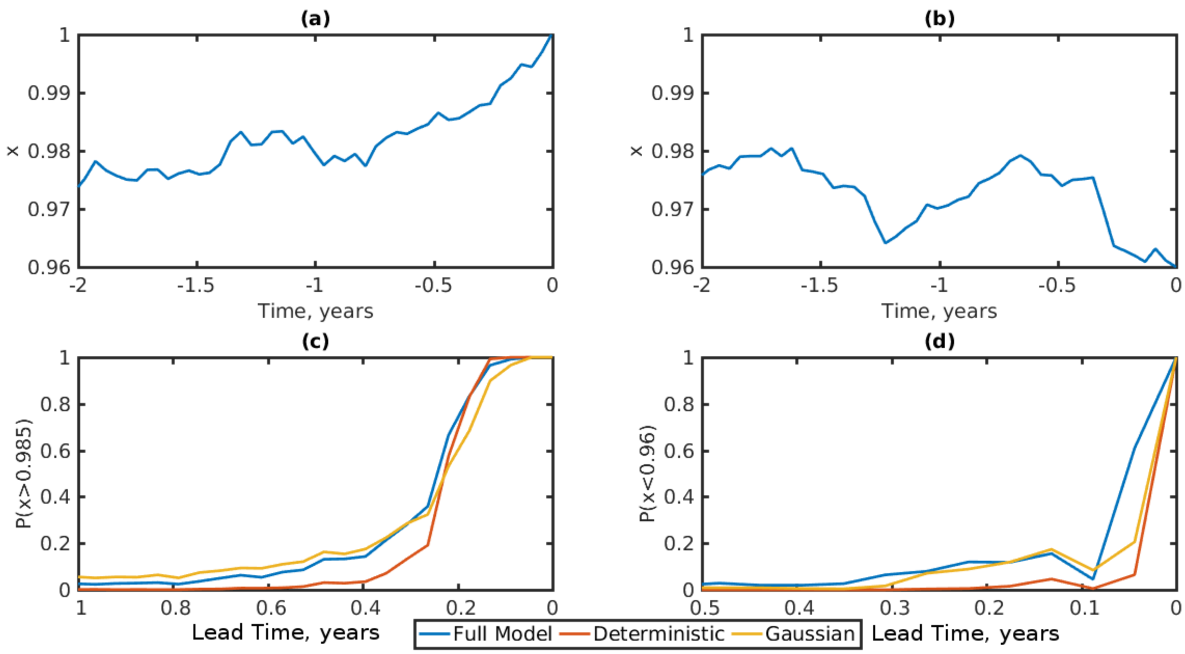

Figure 2. (a) and (b): x trajectories of the full model (1). (c) probability that $x > 0.985$ at $t = 0$, and (d) probability that $x < 0.96$ at $t = 0$ for forecasts initialized from the trajectories in (a) and (b), respectively. Note that the time axes in (c) and (b) are different from each other and from those in (a) and (b).

the probabilities $P\,(x(t = 0) \leq 0.96)$ for the small-event case and $P\,(x(t = 0) \geq 0.985)$ for the large-event case. The probabilities shown in Fig. 2c correspond to the large-x trajectory, and those in Fig. 2d correspond to the small-x trajectory. Since the large-x trajectory crosses the threshold nearly half a year before the final time, all 10,000 of the forecasts initialized at any lead time less than half a year in advance are already above threshold; nevertheless, the probability at the final time is less than one because many of the trajectories cross the threshold back towards smaller values of x.

In both cases the forecast by the deterministic model is significantly worse that the other two models at all but the shortest lead times. The rare-event probability forecast by the Gaussian stochastic model, in contrast, begins to increase from its climatological value at approximately the same time that the true forecast probability begins to increase, between 0.8 and 0.6 years in advance for the large-x event and around 0.3 years in advance for the small-x event. Although the actual probability assigned by the Gaussian stochastic model at relatively long lead times is incorrect, the fact that it begins to increase at the right time could still be used qualitatively to predict whether the model is getting close to a rare event. Once the probability of a rare event increases past about 20%, the Gaussian stochastic model uniformly under-predicts the correct probability, despite having over-predicted the climatological probability for $x > 0.985$. For example, with a lead time of about 2.5 months the Gaussian stochastic model predicts the large-x event with probability only 53% while the true probability is in fact 67%; with a lead time of half a month the Gaussian stochastic model predicts the small-x event with probability only 21% while the true probability is 61%. Differences in the small-event and large-event predictability for these two cases are probably less related to intrinsic predictability than to the fact that the true trajectory remains above threshold for half a year before the forecast verification time $t = 0$ in the large-event case, while in the small-event case the true trajectory reaches threshold only at $t = 0$.

In summary, the deterministic model is essentially useless for rare-event forecasting, while the Gaussian stochastic model is only qualitatively useful, predicting whether a rare event is more likely but not with a robust uncertainty estimate.

6. A model without mean diffusion

As noted at the end of §3, the averaged effect of the eddies is linear and diffusive. Linear diffusive terms are already included in the budgets of heat and salt, with the result that the averaged models have only one stable equilibrium unless the eddies are assumed to be extremely weak, κ, with velocities on the order of 1 cm/s. If one assumes that linear diffusive exchange between the boxes is entirely eddy-driven then one can drop the mean diffusion terms from the governing equations of the full model, i.e. equations (1a) and (1b) are changed to

$$\mathrm{d}x = \left[-\frac{1}{T}(x-1) - P_a(x-y)^2 x + 4vT\right]\mathrm{d}t + \sqrt{\frac{1}{T}}\,\sigma_x \mathrm{d}W_x \tag{9}$$

and

$$\mathrm{d}y = \left[1 - P_a(x-y)^2 y + 4vS\right]\mathrm{d}t + \sigma_y \mathrm{d}W_y \tag{10}$$

respectively. The eddy reductions proceed as before, so that the $-x$ and $-y$ terms are similarly dropped from the deterministic (5) and Gaussian (7) models. The resulting model is much more amenable to multiple equilibria. For P greater than about 0.514 there is a single stable equilibrium with $x \approx 1$ and $y \approx 0.25$. Below this value of P the system undergoes a saddle-node bifurcation that creates a pair of equilibria near $(x, y) = (1, 1)$; the saddle then moves down towards the original equilibrium, which it joins in a reverse saddle-node bifurcation at P approximately 0.301, below which there remains only a single equilibrium. We investigate the system at a value of $P_e = 32$, i.e. $P \approx 0.45$, where there are three equilibria: a stable one at (.99, .24), a saddle at (1.00, .65), and another stable one at (1.00, 1.11).

6.1. *Ergodicity*

Recall that there are two conditions for ergodicity of hypoelliptic SDEs in Mattingly *et al.* (2002). The first condition is that there is an inner-product norm $\|\cdot\|$ such that $\langle \boldsymbol{u}, \boldsymbol{F}(\boldsymbol{u})\rangle \le \alpha - \beta\|\boldsymbol{u}\|^2$ for some $\alpha, \beta > 0$ where $\boldsymbol{u}$ is a vector containing the dependent variables and $\boldsymbol{F}(\boldsymbol{u})$ is the drift. The second condition is that the vectors $\{\boldsymbol{\rho}_i, [\boldsymbol{F}, \boldsymbol{\rho}_j], [\boldsymbol{\rho}_j, \boldsymbol{\rho}_k]\}$ span $\mathbb{R}^5$ where $\boldsymbol{\rho}_i$, $i = 1, 2, 3$ are the columns of the diffusion matrix, and $[\cdot,\cdot]$ is a Lie bracket. It is straightforward to verify that the second condition is met in this model in the same way that it is met in the original model (1).

The first condition is more difficult. We will use the inner product $\langle \boldsymbol{u}, \boldsymbol{v}\rangle = u_1v_1 + u_2v_2 + u_3v_3 + (2/P^2)(u_4v_4 + u_5v_5)$, so we must show that there are $\alpha, \beta > 0$ such that

$$\langle \boldsymbol{u}, \boldsymbol{F}(\boldsymbol{u})\rangle - \alpha + \beta\|\boldsymbol{u}\|^2 \le 0$$

i.e.

$$\begin{aligned}-\alpha + y - x(x-1)/T - P_a(x-y)^2(x^2+y^2) - v^2 - (2/P^2)(T^2+S^2)\\ + \beta(x^2+y^2+v^2+(2/P^2)(T^2+S^2)) \le 0.\end{aligned}$$

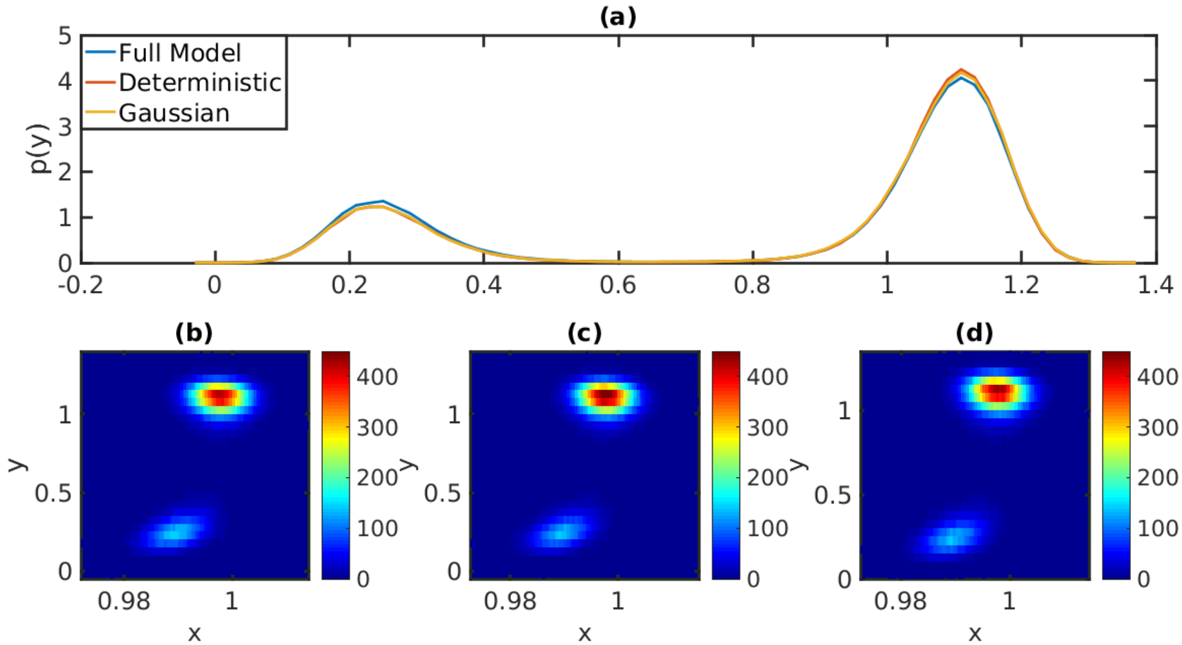

Figure 3. (a) Climatological marginal probability density functions $p(y)$ for the three models without mean diffusion. Climatological joint probability density functions $p(x, y)$ for (b) Full model, (c) Deterministic approximation, and (d) Gaussian-stochastic approximation.

The terms involving the eddy variables (T, v, and S) will clearly pose no problem provided that $\beta < {}^{-1}$. It therefore remains to see whether one can choose α, β such that

$$-\alpha + y - x(x-1)/\gamma - P_a(x-y)^2(x^2+y^2) + \beta(x^2+y^2) \leq 0.$$

Consider the behavior along a line through the origin in the (x, y) plane: along any line except $y = x$ the function is a quartic polynomial that can be made negative by choosing α sufficiently large. Along the line $y = x$ the condition reduces to

$$-\alpha + x - x(x-1)/\gamma + 2\beta x^2 \leq 0.$$

As long as $\beta < 1/(2\gamma)$ it will be possible to choose α sufficiently large that this condition is met. The model without mean diffusion terms is therefore still ergodic. Ergodicity is important because it implies that there is a single climatological distribution independent of the initial condition; the conditions of Mattingly *et al.* (2002) further guarantee that the distribution collapses exponentially quickly towards the climatological distribution.

6.2. *Numerical experiments*

Ensemble simulations for the three models without mean diffusion were run with 1000 ensemble members each; all parameters are the same as in §5 except $P_e = 32$. The deterministic and Gaussian approximate models were initialized with $x = 1$, $y = 0.6$, while the full model was initialized with $x = 1$, $y = 0.65$, and $v, T, S = 0$. After a burn-in of 4 nondimensional time units, the simulations were run for 500 more time units, i.e. about 110,000 years. Although the models are geometrically ergodic, with distributions collapsing exponentially quickly towards the invariant distribution, this was not enough time for the approximate models to reach the invariant distribution. These models were then extended for a further 500 time units, during which time their distributions converged. The full model was initialized closer to the saddle point, so its distribution converged within the first 504 time units.

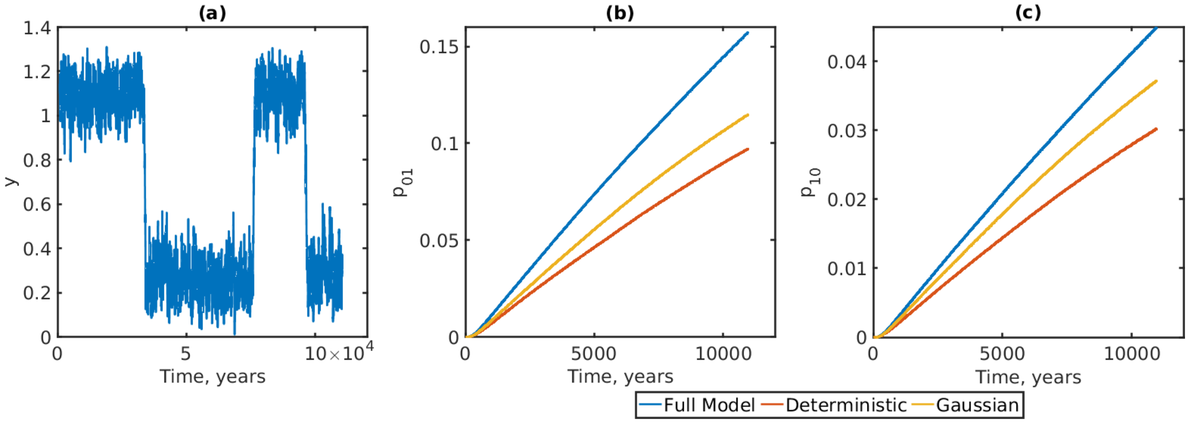

Figure 4. Regime transitions for the three models without mean diffusion. (a) A single $y(t)$ trajectory from the full system showing jumps between regimes. (b) The probability $p_{01}(\tau)$ of a transition from $y(t) < 0.5$ to $y(t+\tau) > 0.8$. (c) The probability $p_{10}(\tau)$ of a transition from $y(t) > 0.8$ to $y(t+\tau) < 0.5$.

The climatological distributions of the slow variables are shown in Fig. 3. Panel (a) shows the marginal y distributions of the three models, while panels (b)–(d) show the joint (x, y) distributions. The three models are remarkably similar. Though is the same as in the previous case, P is smaller. The diffusion correction in the x equation of the Gaussian-stochastic approximation has amplitude $\sqrt{5/4}\,P^2 x \approx 0.026$ which is smaller than the atmospheric noise amplitude $\sigma_x/\sqrt{\tau} = 0.1$; the atmospheric noise similarly dominates the y equation. As a result, the effects of eddy noise are not seen in the equilibrium distributions of the three models.

The noise levels are low enough that the system trajectories make rare transitions between the neighborhoods of the two stable equilibria. Fig. 4 panel (a) shows a system trajectory y from the full model that jumps between regimes. The rates and paths of these transitions are the subject of large deviation theory (Freidlin and Wentzell 2012). The methods of Bouchet *et al.* (2016) to analyze the transitions do not seem to apply directly here because of the inclusion of noise forcing in the slow dynamics. In any case, it is not difficult to estimate the transition probabilities from simulations. For practical purposes it was convenient to estimate the following probabilities $p_{01}(\tau) = P(y(t+\tau) > 0.8 \mid y(t) < 0.5)$ and $p_{10}(\tau) = P(y(t+\tau) < 0.5 \mid y(t) > 0.8)$. These transition probabilities are plotted for the three models in Fig. 4 panels (b) and (c), respectively. The effects of differences in the eddy noise are clear: the deterministic model has the lowest transition probabilities; the Gaussian stochastic model has higher transition probabilities; the full model has the highest transition probabilities.

7. Conclusions

This paper formulates a stochastic two-box ocean model modeled after Stommel's (1961); the model consists of a system of 5 SDEs (1). Previous stochastic Stommel models (e.g. Cessi 1994, Vélez-Belchí *et al.* 2001, Monahan *et al.* 2002, Monahan 2002, Monahan and Culina 2011), modeled the atmospheric heat and freshwater forcing as Gaussian stochastic processes, and the exchange of heat and salt between the boxes as a nonlinear drift term corresponding to the large-scale overturning thermohaline circulation. The novelty of the formulation here is that a fast, eddy-driven component is added to the the exchange between the boxes. The terms modeling the eddy-driven exchange are quadratic products of approximately Gaussian random variables; products of jointly-Gaussian random fields were recently found to be an accurate model of eddy-driven exchanges in Grooms (2016).

In more complete and complex ocean models, fast eddy effects are frequently modeled deterministically. Stochastic parameterizations have recently been developed that multiply these deterministic eddy parameterizations by Gaussian random fields (Andrejczuk *et al.* 2016, Juricke *et al.* 2017), which is a popular approach for atmospheric models based on the work of Buizza *et al.* (1999) and Sura *et al.* (2005). A key benefit of stochastic parameterizations in comparison to deterministic ones is the former's ability to induce realistic variability in the resolved scales. Models with realistic variability are needed for making forecasts with robust uncertainty estimates, which explains the wide adoption of stochastic parameterizations in weather forecasting (Buizza *et al.* 1999, Orrell *et al.* 2001, Palmer *et al.* 2005, Berner *et al.* 2017, Leutbecher *et al.* 2017).

Using methods of averaging and homogenization for slow-fast systems (Pavliotis and Stuart 2008, Freidlin and Wentzell 2012, Bouchet *et al.* 2016), two models were derived approximating the evolution of the slow components (the difference in heat and salt content of the two boxes). The first model (5) replaces the fast eddy-driven exchange terms by a fixed 'deterministic' drift term, analogous to the standard approach of deterministic parameterization in more complex ocean models. The second model (7) adds an additional multiplicative noise term accounting for fast variations in the eddy-driven flux. A suite of simulations of each of the three models was used to compare their qualitative behavior in a parameter regime with a single stable equilibrium. All three models were then altered by removing an explicit representation of diffusion and allowing all diffusive effects to be achieved completely by the eddies. Numerical simulations of these models were used to compare their qualitative behavior in a parameter regime with two stable equilibria.

The main results are as follows. There is little qualitative difference in the core of the stationary distributions of the full, non-Gaussian model and the Gaussian multiplicative approximation. In the regime with a single equilibrium the deterministic model has too little variability, but the Gaussian model gives an accurate climatological mean and covariance. In the regime with two stable equilibria the climatological distribution of the three models is nearly the same. In the regime with two stable equilibria the amplitude of the eddies is smaller than in the regime with a single equilibrium, which could perhaps account for the fact that the deterministic model is more accurate in the former regime. Observational estimates suggest that up to 30% of the variability of the Atlantic Meridional Overturning Circulation (AMOC) is driven by ocean eddies, with the rest driven by atmospheric noise (Hirschi *et al.* 2013, Sonnewald *et al.* 2013).

Though the Gaussian stochastic model gives a good approximation of the core of the climatological distribution, the rare event probabilities are inaccurate. In the single-equilibrium regime there is no clear trend in the behavior: the Gaussian model overestimates rare event probabilities on one side of the mean, and underestimates on the other side. This inaccuracy manifests for short time, transient behavior too: even with a short lead time, the Gaussian model gives inaccurate predictions of the probability of a rare event. Surprisingly, despite overestimating the climatological rare event probability for one kind of event, in a rare event forecasting configuration the Gaussian model systematically underestimates the rare event probability for both kinds of events (i.e. events above and below the climatological mean).

In the regime with two stable equilibria the rare events of interest are the transitions between the two. Despite the fact that the amplitude of the eddy noise in this regime is smaller than the amplitude of the atmospheric noise, clear differences were observed in the rates of transition from the neighborhood of one equilibrium to another: the deterministic model had the rarest transitions, and the Gaussian model still made transitions less frequently than the full model.

In the single-equilibrium regime significant differences in the rare-event dynamics of the three models were only found in the x variable, which describes the temperature difference between the poleward and equatorial boxes. The amplitude of the eddy noise in the salinity equation was an order of magnitude smaller than the amplitude of the atmospheric noise, and the latter dominated despite the long-tailed non-Gaussian statistics of the noise in the full model. In contrast, the amplitude of the eddy noise in the temperature equation was closer to the amplitude of the atmospheric noise, and the effects of non-Gaussianity in the noise were evident in the rare-event statistics. In the regime with two stable equilibria the rare events are transitions between neighborhoods of the two equilibria, and they are most prominent in the salinity rather than the temperature. The amplitude of the eddy noise in this regime is smaller than the amplitude of the atmospheric noise by a factor of about 6, but the long-tailed non-Gaussianity of the eddy noise is still able to have an impact on the rare event probability.

The goal of the investigation was to investigate the qualitative impacts of non-Gaussian eddy noise of the type observed by Grooms (2016) in a simple model, and to compare to models with Gaussian noise and without eddy noise. The extreme simplicity of the model precludes confident extrapolation to more complex and comprehensive ocean models. Nevertheless, the results suggest that Gaussian stochastic parameterizations in ocean general circulation models may be able to successfully produce the day-to-day variability associated with the core of the climatological distribution, but that more accurate non-Gaussian models may be needed to correctly model rare events. Such rare events include extreme behavior like droughts and heat waves, as well as abrupt transitions between climate regimes. The impact of stochastic parameterizations on rare event distributions in climate models has only recently begun to be investigated (Tagle *et al.* 2016).

The qualitative impact of non-Gaussian eddy noise seems to depend on the relative amplitude of that noise in comparison with atmospheric noise forcing. If the eddy noise is significantly smaller than the atmospheric noise, then it will presumably have little impact on the variability of the system. The parameters used here (2) to describe the amplitude of atmospheric and eddy noise are drawn from the literature, but are necessarily imprecise. Hirschi *et al.* (2013) and Sonnewald *et al.* (2013) argue on the basis of observations that up to 30% of the variability of the Atlantic Meridional Overturning Circulation (AMOC) is driven by ocean eddies, with the rest driven by atmospheric noise. Our results suggest that there should be qualitative differences between the rare event probabilities of systems with Gaussian and non-Gaussian eddy models even for noise as small as 30%.

Acknowledgments

IG is grateful to F. Bouchet for a discussion of the methods from Bouchet *et al.* (2016) that are applied here. IG is supported by NSF OCE grant 1736708. WB is supported by NSF EXTREEMS grant 1407340.

References

Andrejczuk, M., Cooper, F., Juricke, S., Palmer, T., Weisheimer, A. and Zanna, L., Oceanic stochastic parameterizations in a seasonal forecast system. *Mon. Weather Rev.*, 2016, **144**, 1867–1875.

Berner, J., Achatz, U., Batt, L., Bengtsson, L., de la Cmara, A., Christensen, H.M., Colangeli, M., Coleman, D.R.B., Crommelin, D., Dolaptchiev, S.I., Franzke, C.L.E., 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.S., Weisheimer, A., Weniger, M., Williams, P.D. and Yano, J.I., Stochastic Parameterization: Toward a New View of Weather and Climate Models. *Bull. Amer. Meteorol. Soc.*, 2017, **98**, 565–588.

Bouchet, F., Grafke, T., Tangarife, T. and Vanden-Eijnden, E., Large deviations in fast–slow systems. *J. Stat. Phys.*, 2016, **162**, 793–812.

Brankart, J.M., Candille, G., Garnier, F., Calone, C., Melet, A., Bouttier, P.A., Brasseur, P. and Verron, J., A generic approach to explicit simulation of uncertainty in the NEMO ocean model. *Geosci. Model Dev.*, 2015, **8**, 1285.

Buizza, R., Miller, M. and Palmer, T., Stochastic Simlulation of Model Uncertainties in the ECWMF Ensemble Prediction System. *QJR Meteorol. Soc*, 1999, **125**, 28887–2908.

Cessi, P., A simple box model of stochastically forced thermohaline flow. *J. Phys. Ocean.*, 1994, **24**, 1911–1920.

Cooper, F.C., Optimisation of an idealised primitive equation ocean model using stochastic parameterization. *Ocean Model.*, 2017, **113**, 187–200.

Cotter, C.J., Gottwald, G.A. and Holm, D.D., Stochastic partial differential fluid equations as a diffusive limit of deterministic Lagrangian multi-time dynamics. *P. R. Soc. A*, 2017, **473**.

Deshayes, J., Tr´eguier, A.M., Barnier, B., Lecointre, A., Sommer, J.L., Molines, J.M., Penduff, T., Bourdall´e-Badie, R., Drillet, Y., Garric, G. *et al.*, Oceanic hindcast simulations at high resolution suggest that the Atlantic MOC is bistable. *Geo. Res. Lett.*, 2013, **40**, 3069–3073.

Freidlin, M. and Wentzell, A., *Random perturbations of dynamical systems*, Third revised and enlarged edition edn, 2012 (Springer).

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

Grooms, I., Lee, Y. and Majda, A.J., Numerical schemes for stochastic backscatter in the inverse cascade of quasigeostrophic turbulence. *Multiscale Model. Sim.*, 2015, **13**, 1001–1021.

Held, I.M., The gap between simulation and understanding in climate modeling. *B. Am. Meteorol. Soc.*, 2005, **86**, 1609–1614.

Higham, D.J., Mao, X. and Stuart, A.M., Strong convergence of Euler-type methods for nonlinear stochastic differential equations. *SIAM J. Numerical Analysis*, 2002, **40**, 1041–1063.

Hirschi, J., Blaker, A., Sinha, B., Coward, A., de Cuevas, B., Alderson, S. and Madec, G., Chaotic variability of the meridional overturning circulation on subannual to interannual timescales. *Ocean Science*, 2013, **9**, 805–823.

Holm, D.D., Variational principles for stochastic fluid dynamics. *P. R. Soc. A*, 2015, **471**, 20140963.

Hutzenthaler, M., Jentzen, A. and Kloeden, P.E., Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally Lipschitz continuous coefficients. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 2011, **467**, 1563–1576.

Juricke, S., Palmer, T.N. and Zanna, L., Stochastic Subgrid-Scale Ocean Mixing: Impacts on Low-Frequency Variability. *J. Climate*, 2017, **30**, 4997–5019.

Khasminskii, R., *Stochastic stability of differential equations*, Second edition edn, 2012 (Springer).

Kloeden, P.E. and Platen, E., *Numerical Solution of Stochastic Differential Equations*, 1992 (Springer).

Lamba, H., Mattingly, J.C. and Stuart, A.M., An adaptive Euler–Maruyama scheme for SDEs: convergence and stability. *IMA journal of numerical analysis*, 2007, **27**, 479–506.

Leutbecher, M., Lock, S.J., 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., Stochastic representations of model uncertainties at ECMWF: State of the art and future vision. *QJR Meteorol. Soc*, 2017.

Mao, X. and Szpruch, L., Strong convergence and stability of implicit numerical methods for stochastic differential equations with non-globally Lipschitz continuous coefficients. *Journal of Computational and Applied Mathematics*, 2013, **238**, 14–28.

Mattingly, J.C., Stuart, A.M. and Higham, D.J., Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise. *Stochastic processes and their applications*, 2002, **101**, 185–232.

M´emin, E., Fluid flow dynamics under location uncertainty. *Geophysical & Astrophysical Fluid Dynamics*, 2014, **108**, 119–146.

Monahan, A.H., Alexander, J. and Weaver, A.J., Stochastic models of the meridional overturning circulation: time scales and patterns of variability. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 2008, **366**, 2525–2542.

Monahan, A.H. and Culina, J., Stochastic averaging of idealized climate models. *J. Climate*, 2011, **24**, 3068–3088.

Monahan, A.H., Correlation effects in a simple stochastic model of the thermohaline circulation. *Stochastics and Dynamics*, 2002, **2**, 437–462.

Monahan, A.H., Timmermann, A. and Lohmann, G., Comments on Noise-induced transitions in a simplified model of the thermohaline circulation. *J. Phys. Ocean.*, 2002, **32**, 1112–1116.

Orrell, D., Smith, L., Barkmeijer, J. and Palmer, T., Model error in weather forecasting. *Nonlinear Proc. Geoph.*, 2001, **8**, 357–371.

Palmer, T., Shutts, G., Hagedorn, R., Doblas-Reyes, F., Jung, T. and Leutbecher, M., Representing model uncertainty in weather and climate prediction. *Annu. Rev. Earth Planet. Sci.*, 2005, **33**, 163–193.

Papanicolaou, G. and Kohler, W., Asymptotic theory of mixing stochastic ordinary differential equations. *Communications on Pure and Applied Mathematics*, 1974, **27**, 641–668.

Pavliotis, G. and Stuart, A., *Multiscale Methods Averaging and Homogenization*, 2008 (Springer).

Porta Mana, P. and Zanna, L., Toward a stochastic parameterization of ocean mesoscale eddies. *Ocean Model.*, 2014, **79**, 1–20.

Rahmstorf, S., Bifurcations of the Atlantic thermohaline circulation in response to changes in the hydrological cycle. *Nature*, 1995, **378**, 145.

Resseguier, V., M´emin, E. and Chapron, B., Geophysical flows under location uncertainty, Part I Random transport and general models. *Geophys. Astro. Fluid*, 2017, **111**, 149–176.

Schmitt, R.W., Salinity and the Global Water Cycle. *Oceanography*, 2008, **21**, 12.

Sonnewald, M., Hirschi, J.J.M., Marsh, R., McDonagh, E.L. and King, B.A., Atlantic meridional ocean heat transport at 26N: impact on subtropical ocean heat content variability. *Ocean Science*, 2013, **9**, 1057–1069.

Stammer, D., Global characteristics of ocean variability estimated from regional TOPEX/POSEIDON altimeter measurements. *J. Phys. Ocean.*, 1997, **27**, 1743–1769.

Stommel, H., Thermohaline convection with two stable regimes of flow. *Tellus*, 1961, **13**, 224–230.

Sura, P., Newman, M., Penland, C. and Sardeshmukh, P., Multiplicative noise and non-Gaussianity: A paradigm for atmospheric regimes?. *J. Atmos. Sci.*, 2005, **62**, 1391–1409.

Tagle, F., Berner, J., Grigoriu, M.D., Mahowald, N.M. and Samorodnitsky, G., Temperature Extremes in the Community Atmosphere Model with Stochastic Parameterizations. *J. Climate*, 2016, **29**, 241–258.

Vélez-Belchí, P., Alvarez, A., Colet, P., Tintoré, J. and Haney, R., Stochastic resonance in the thermohaline circulation. *Geo. Res. Lett.*, 2001, **28**, 2053–2056.

Williams, P.D., Howe, N.J., Gregory, J.M., Smith, R.S. and Joshi, M.M., Improved Climate Simulations through a Stochastic Parameterization of Ocean Eddies. *J. Climate*, 2016, **29**, 8763–8781.

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