JAMES | Journal of Advances in Modeling Earth Systems

RESEARCH ARTICLE
10.1029/2022MS003297

Stochastic Data-Driven Parameterization of Unresolved Eddy Effects in a Baroclinic Quasi-Geostrophic Model

Long Li[1], Bruno Deremble[2], Noé Lahaye[1], and Etienne Mémin[1]

[1]Centre Inria de l'Université de Rennes, Rennes, France, [2]Institut des Géosciences de l'Environnement, Grenoble, France

Abstract In this work, a stochastic representation based on a physical transport principle is proposed to account for mesoscale eddy effects on the large-scale oceanic circulation. This stochastic framework arises from a decomposition of the Lagrangian velocity into a smooth-in-time component and a highly oscillating noise term. One important characteristic of this random model, without any external forcing and damping, is that it conserves the total energy of the resolved flow for any realization. The proposed stochastic formulation is successfully implemented in a well established multi-layered quasi-geostrophic dynamical core. The empirical spatial correlation of the unresolved noise is calibrated from the eddy-resolving simulation data. In particular, a stationary correction drift can be introduced in the noise through Girsanov transformation. This non-intuitive term appears to be important in reproducing on a coarse mesh the eastward jet of the wind-driven double-gyre circulation. In addition, a projection method has been proposed to constrain the noise to act along the iso-surfaces of the vertical stratification. The resulting noise enables us to improve the intrinsic low-frequency variability of the large-scale current.

Plain Language Summary Accurate numerical simulation of the ocean and the atmosphere is important for understanding complex patterns in climate science. Unfortunately, the high-dimensional geophysical fluid dynamics contain a wide range of spatio-temporal scales with complex interactions. In order to maintain affordable computational time, only large-scale representations of the geophysical flows of interest are simulated. However, the effect of the unresolved scales on the resolved component must be carefully taken into account to minimize the errors of truncation. In the present work, we stick to a specific stochastic formulation that represents the physical principles of the resolved large scales and parameterizes the statistical properties of the unresolved small scales. Such random formulation is successfully tested in numerical simulations of a simplified ocean mesoscale model, in which the structure of the unresolved flow is calibrated from a finite sequence of reference data. The resulting model enables to represent accurately the statistics of the resolved flow component. In addition, numerical results show that an adequate modeling of the unresolved dynamics leads to a significant improvement of the intrinsic variability for the large-scale ocean models. These findings are important for future developments of ocean and atmosphere climate models.

Key Points:

- A stochastic transport formulation is used to parameterize the effects of the unresolved eddies on the resolved flow
- The structure of unresolved flow is calibrated from data and a projection method is proposed to parameterize its dynamics
- The proposed random models on coarse resolutions enable to reproduce the eastward jet and the intrinsic variability of ocean wind-driven circulation

Correspondence to:
L. Li,
long.li@inria.fr

Citation:
Li, L., Deremble, B., Lahaye, N., & Mémin, E. (2023). Stochastic data-driven parameterization of unresolved eddy effects in a baroclinic quasi-geostrophic model. *Journal of Advances in Modeling Earth Systems*, *15*, e2022MS003297. https://doi.org/10.1029/2022MS003297

Received 14 JUL 2022
Accepted 28 JAN 2023

Author Contributions:
Conceptualization: Etienne Mémin
Data curation: Long Li
Formal analysis: Long Li, Etienne Mémin
Funding acquisition: Etienne Mémin
Methodology: Long Li, Bruno Deremble, Etienne Mémin
Software: Long Li
Supervision: Bruno Deremble, Noé Lahaye, Etienne Mémin
Validation: Bruno Deremble, Noé Lahaye, Etienne Mémin
Visualization: Long Li
Writing – original draft: Long Li
Writing – review & editing: Bruno Deremble, Noé Lahaye, Etienne Mémin

1. Introduction

Ocean mesoscale eddies contain a large proportion of energy and have an important impact on large-scale circulations. They are found everywhere and are particularly intensive in the western boundary currents such as the Gulf Stream and the Kuroshio. However, these eddies are often not resolved in ocean general circulation models, since the deformation radius in the ocean is at most of the order of 100 km. In particular, the effects of the mesoscale eddies need to be properly parameterized in coarse-resolution ocean models.

The most successful parameterization of the mesoscale effect is based on the eddy-induced advection scheme (Gent & McWilliams, 1990; Gent et al., 1995; Griffies, 1998), which mimics the impact of the baroclinic instability by flattening the isopycnal surfaces to transfer the available potential energy (PE) of the resolved flow to the eddy kinetic energy at the sub-grid scales. However, this traditional scheme does not account for the backscattering of kinetic energy (KE) from small to large scales. Some complementary schemes (Bachman, 2019; Bachman et al., 2017; Grooms et al., 2015) have been developed in order to perform more efficient energy transfers for the resolved flow. Classical eddy viscosity models are introduced in coarse models to mimic the mixing action of the unresolved scales. The associated energy dissipation mechanism is often represented by some functional operators (Bachman et al., 2017; Griffies & Hallberg, 2000; Leith, 1971) that depend on the resolved flow. These highly scale-selective, state-dependent friction operators have proven to be very useful for eddy-permitting

models. Nevertheless, they still lead to excessive decreases of the resolved KE in large-scale ocean models. The lack of variability due to energy missing is particularly detrimental in ensemble forecasting and data assimilation applications (Franzke et al., 2015; Karspeck et al., 2013).

To address these shortcomings, stochastic parameterizations are becoming more and more popular and several schemes have been devised and studied for ocean circulation models. These schemes introduce energy backscattering across scales and increase the internal ocean variability. For instance, Berloff (2005) developed a random-forcing model based on the dynamical decomposition of the flow into a large-scale component and an eddy component. Jansen and Held (2014) proposed a stochastic forcing in the form of a negative Laplacian viscosity to inject dissipated energy back to the resolved flow. Grooms and Majda (2014) and Grooms et al. (2015) proposed a stochastic superparameterization scheme that includes stochastic Reynolds stress terms to backscatter KE. Porta Mana and Zanna (2014) and Zanna et al. (2017) proposed an eddy parameterization based on a non-Newtonian stress which depends on the partially resolved scales and their variability.

We propose here a specific stochastic formulation, the so-called location uncertainty (LU) (Mémin, 2014), which introduces a random unresolved flow component and relies on a stochastic transport principle. This random model enables consideration of systems that are more energetic than their deterministic counterparts. Nevertheless, the ability of such models to represent faithfully the uncertainties associated to the actual unresolved small scales highly depends on the definition of the random component and on its evolution in time. Unsurprisingly, stationarity/time-varying and homogeneity/inhomogeneity characteristics of the unresolved flow component have strong influences on the numerical results. For instance, Bauer et al. (2020a) illustrates that the noise inhomogeneity induces a structuration of the large-scale flow reminiscent to the action of the vortex force associated with surface wave-induced Stokes drift. Bauer et al. (2020b) shows that the stochastic barotropic quasi-geostrophic (QG) model under an inhomogeneous and non-stationary noise enables us to reproduce on a coarse mesh the high order statistics of the eddy-resolving data. Brecht et al. (2021) demonstrates that the stochastic shallow water model under the same kind of noise provides a good trade-off between model error representation and ensemble spread.

There exists very similar approaches based on the same type of random flow decomposition such as the stochastic advection by Lie transport (SALT) model proposed by Holm (2015) and Cotter et al. (2019, 2020). Some analytical and numerical comparisons between the LU and SALT models have been already performed in the previous works (Bauer et al., 2020b; Resseguier, Pan, & Fox-Kemper, 2020). In short, the LU models conserve the total energy of the resolved flow, whereas the SALT models preserve the potential enstrophy of the resolved flow.

In this work, we investigate the LU formulation for the baroclinic QG model (Hogg et al., 2004) within an idealized double-gyre circulation configuration. It is the first time that the LU formulation is applied on a vertically dependent model. In particular, we focus on the reproduction of the eastwards jet as well as the prediction of the low-frequency variability for the proposed random model on a coarse resolution grid, for cases in which the baroclinic instability is not resolved. Another important aspect concerns the ability to include in the noise representation a stationary drift component associated to the temporal mean of the high-resolution (HR) fluctuations. As shown in this paper such stationary drift can be elegantly introduced in the noise through Girsanov transformation. In addition, a projection method has been proposed to update in time the noise along the iso-surfaces of the stratification.

The remainder of this paper is structured as follows. Section 2 describes the framework of LU and presents the derived stochastic QG model. Section 3 explains the discretization of the proposed QG model and the parameterization method for the unresolved noise. Section 4 discusses the numerical results with some statistical diagnosis and energetic analysis. In Section 5 we draw some conclusions and provide an outlook for future works. In the Appendices we show briefly the derivation of the proposed QG model, demonstrate the energy conservation of the random system and explore the energy conversion between the ensemble-mean and the ensemble-eddy components.

2. Continuous Models

In this section, we first review the general setting of LU. We then present the resulting stochastic baroclinic QG model for a continuously stratified ocean. Finally we complete the random formulation by including a time-correlated drift into the unresolved flow component through the Girsanov transformation.

2.1. Stochastic Flow

The evolution of Lagrangian particle trajectory ($\boldsymbol{X}_t$) under LU is described by the following stochastic differential equation:

$$\mathrm{d}\boldsymbol{X}_t(\boldsymbol{x}) = \boldsymbol{v}(\boldsymbol{X}_t(\boldsymbol{x}), t)\,\mathrm{d}t + \boldsymbol{\sigma}(\boldsymbol{X}_t(\boldsymbol{x}), t)\,\mathrm{d}\boldsymbol{B}_t, \qquad \boldsymbol{X}_0(\boldsymbol{x}) = \boldsymbol{x} \in \mathcal{D}, \tag{2.1}$$

where $\boldsymbol{v}$ denotes the time-smooth resolved velocity that is both spatially and temporally correlated, $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$ stands for the fast oscillating unresolved flow component (also called *noise* in the following) that is correlated in space yet uncorrelated in time.

We now give the mathematical definition of the noise. The random process $\{\boldsymbol{B}_t\}_{0\le t\le T}$ is an $\boldsymbol{I}_d$-cylindrical Brownian motion (Da Prato & Zabczyk, 2014) on a filtered probability space $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\le t\le T}, \mathbb{P}\right)$ and takes values in the Hilbert space $H := \left(L^2(\mathcal{D})\right)^d$, where $\mathcal{D} \subset \mathbb{R}^d$ ($d = 2$ or 3) is a bounded spatial domain, Ω is a sample space, $\mathcal{F}$ is a collection of subsets of Ω, $\{\mathcal{F}_t\}_{0\le t\le T}$ is a filtration and $\mathbb{P}$ is a probability measure. Informally, one can consider the time-derivative of $\boldsymbol{B}_t$ as a space-time "white noise" (in a distribution sense) with independent components $\left(\boldsymbol{B}_t^i\right)_{i=1,\dots,d}$. We remark that the randomness symbol parameter $\omega \in \Omega$ of the stochastic processes (e.g., $\boldsymbol{X}_t(\boldsymbol{x}) := \boldsymbol{X}(\boldsymbol{x}, t, \omega)$) have been dropped in this work for the sake of simplicity.

The spatial structure of the unresolved flow component is modeled by the correlation operator $\boldsymbol{\sigma}$. In the most general case it is random and time-dependent with some regularity conditions. For each $\omega \in \Omega$ and $t \in [0, T]$, $\boldsymbol{\sigma}(\cdot, t)[\cdot]$ is assumed to be a Hilbert-Schmidt integral operator on H with a bounded matrix kernel $\breve{\boldsymbol{\sigma}} = (\breve{\sigma}_{ij})_{i,j=1,\dots,d}$. The composition of $\boldsymbol{\sigma}(\cdot, t)[\cdot]$ and its adjoint $\boldsymbol{\sigma}^*(\cdot, t)[\cdot]$ is a trace class operator on H and admits eigenfunctions $\boldsymbol{\xi}_n(\cdot, t)$ with eigenvalues $\lambda_n(t)$ satisfying $\sum_{n\in\mathbb{N}} \lambda_n(t) < +\infty$. The spectral decomposition (Da Prato & Zabczyk, 2014) of the noise reads then

$$\boldsymbol{\sigma}(\boldsymbol{x}, t)\,\mathrm{d}\boldsymbol{B}_t = \int_{\mathcal{D}} \breve{\boldsymbol{\sigma}}(\boldsymbol{x}, \boldsymbol{y}, t)\,\mathrm{d}\boldsymbol{B}_t(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} = \sum_{n\in\mathbb{N}} \lambda_n^{1/2}(t)\boldsymbol{\xi}_n(\boldsymbol{x}, t)\,\mathrm{d}\beta_n(t), \tag{2.2}$$

where β_n are independent (one-dimensional) standard Brownian motions.

In addition, we assume that the operator-space-valued process $\{\boldsymbol{\sigma}(\cdot, t)[\cdot]\}_{0\le t\le T}$ is stochastically integrable, that is, $\mathbb{P}\left[\int_0^T \sum_{n\in\mathbb{N}} \lambda_n(t)\,\mathrm{d}t < +\infty\right] = 1$. As such, under the probability measure $\mathbb{P}$, the stochastic integral $\left\{\int_0^t \boldsymbol{\sigma}(\cdot, s)\,\mathrm{d}\boldsymbol{B}_s\right\}_{0\le t\le T}$ is a H-valued Gaussian process of zero mean, that is, $\mathbb{E}_{\mathbb{P}}\left[\int_0^t \boldsymbol{\sigma}(\cdot, s)\,\mathrm{d}\boldsymbol{B}_s\right] = \boldsymbol{0}$, and of bounded (global) variance, that is, $\mathbb{E}_{\mathbb{P}}\left[\left\|\int_0^t \boldsymbol{\sigma}(\cdot, s)\,\mathrm{d}\boldsymbol{B}_s\right\|_H^2\right] < +\infty$.

Moreover, the point-wise ($\boldsymbol{x} \in \mathcal{D}$) and path-wise ($\omega \in \Omega$) strength of the unresolved flow component at each instant ($t \in [0, T]$) is measured by the matrix kernel of the composite operator $\boldsymbol{\sigma}\boldsymbol{\sigma}^*$, and denoted by $\boldsymbol{a}(\boldsymbol{x}, t)$, namely

$$\boldsymbol{a}(\boldsymbol{x}, t) := \int_{\mathcal{D}} \breve{\boldsymbol{\sigma}}(\boldsymbol{x}, \boldsymbol{y}, t)\breve{\boldsymbol{\sigma}}^T(\boldsymbol{y}, \boldsymbol{x}, t)\,\mathrm{d}\boldsymbol{y} = \sum_{n\in\mathbb{N}} \lambda_n(t)\left(\boldsymbol{\xi}_n\boldsymbol{\xi}_n^T\right)(\boldsymbol{x}, t). \tag{2.3}$$

As shown in Bauer et al. (2020a), the process $\int_0^t \boldsymbol{a}(\boldsymbol{x}, s)\,\mathrm{d}s$ actually corresponds to the quadratic variation (Da Prato & Zabczyk, 2014) of $\int_0^t \boldsymbol{\sigma}(\boldsymbol{x}, s)\,\mathrm{d}\boldsymbol{B}_s$, and it is a continuous random process of finite variation (hence time differentiable). In the particular case of a nonrandom $\boldsymbol{\sigma}$, the matrix kernel $\boldsymbol{a}$ reduces then to the local variance of the noise according to the Itô isometry (Da Prato & Zabczyk, 2014). Physically, the symmetric non-negative tensor $\boldsymbol{a}$ represents the (possibly random) friction coefficients of the unresolved fluid motions and the eigenvalues λ_n have the unit of m^2/s.

2.2. Stochastic Transport

The evolution law of a random tracer Θ with extensive property (e.g., temperature, salinity, buoyancy) transported by the stochastic flow, $\Theta(\boldsymbol{X}_{t+\delta t}, t + \delta t) = \Theta(\boldsymbol{X}_t, t)$ with δt an infinitely small time variation, is derived by Mémin (2014) and Bauer et al. (2020a) using the generalized Itô formula (Kunita, 1997). It is described by the following stochastic partial differential equation (SPDE):

$$\mathbb{D}_t\Theta = \mathrm{d}_t\Theta + \left(\boldsymbol{v}^\star\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)\cdot\boldsymbol{\nabla}\Theta - \frac{1}{2}\boldsymbol{\nabla}\cdot(\boldsymbol{a}\boldsymbol{\nabla}\Theta)\,\mathrm{d}t = 0, \tag{2.4}$$

in which $\mathbb{D}_t$ is introduced as a stochastic transport operator and $\mathrm{d}_t\Theta(\boldsymbol{x}) \coloneqq \Theta(\boldsymbol{x}, t+\delta t) - \Theta(\boldsymbol{x}, t)$ stands for the (forward) time-increment of Θ at a fixed point $\boldsymbol{x} \in \mathcal{D}$.

This SPDE encompasses physically meaningful terms. For instance, the third term on the right-hand side (RHS) of Equation 2.4 describes the tracer's advection by the unresolved flow component. As shown in Resseguier et al. (2017a) and Bauer et al. (2020a), the resulting multiplicative noise $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t \cdot \nabla\Theta$ continuously backscatters random energy to the system through the quadratic variation $1/2\ (\nabla\Theta) \cdot (\boldsymbol{a}\nabla\Theta)$ of the random tracer. The last term in Equation 2.4 depicts tracer's diffusion due to the mixing of the unresolved scales. In particular, under specific noise definitions (Mémin, 2014), the resulting diffusion can be connected to the eddy viscosity with flow-dependent coefficients as introduced in many large-scale circulation models (Redi, 1982; Smagorinsky, 1963).

As an additional feature of interest, there exists an effective advection velocity $\boldsymbol{v}^\star$ in Equation 2.4 which is defined as

$$\boldsymbol{v}^\star = \boldsymbol{v} - \frac{1}{2}\boldsymbol{\nabla}\cdot\boldsymbol{a} + \boldsymbol{\sigma}^*(\boldsymbol{\nabla}\cdot\boldsymbol{\sigma}). \tag{2.5}$$

This statistical eddy-induced velocity drift captures the action of inhomogeneity of the random field on the transported tracer and the possible divergence of the unresolved flow component. Bauer et al. (2020a) shows that the *turbophoresis* term $1/2\ \nabla\cdot\boldsymbol{a}$ can be interpreted as a generalization of the Stokes drift, which occurs, for example, in the Langmuir circulation (Craik & Leibovich, 1976; Leibovich, 1980).

Many useful properties of the stochastic transport operator $\mathbb{D}_t$ have been explored by Resseguier et al. (2017a), Resseguier, Li, et al. (2020), and Li (2021). In particular, if a random tracer is transported by the incompressible stochastic flow under suitable boundary conditions, then the path-wise pth moment ($p \geq 1$) of the tracer is materially and integrally invariant, namely

$$\mathbb{D}_t\left(\frac{1}{p}\Theta^p\right) = 0, \quad \mathrm{d}_t\int_D \frac{1}{p}\Theta^p\,\mathrm{d}\boldsymbol{x} = 0. \tag{2.6}$$

2.3. Stochastic Baroclinic QG Model

The LU formulation have been analytically and numerically investigated for different barotropic models in previous works (Bauer et al., 2020a, 2020b; Brecht et al., 2021). A preliminary stochastic baroclinic QG model has been derived by Resseguier et al. (2017b), in which the horizontal component of the noise was assumed to be quasi-constant along z. Moreover, the proposed model has never been numerically assessed. In this work, we follow the same strategies to re-derive the continuous model but the vertically independent assumption is here relaxed by introducing an additive buoyancy noise into the hydrostatic equation. A sketch of the derivation of this stochastic QG model can be found in Appendix A and the final governing equations are summarized as the following system:

Momentum equations

$$\mathbb{D}_t^h\boldsymbol{u} + f_0\boldsymbol{k}\times\boldsymbol{u}_a\,\mathrm{d}t + \beta y\boldsymbol{k}\times(\boldsymbol{u}\,\mathrm{d}t + \boldsymbol{\sigma}_h\mathrm{d}\boldsymbol{B}_t) = -\boldsymbol{\nabla}_h p_a\,\mathrm{d}t, \tag{2.7a}$$

Buoyancy equation

$$\mathbb{D}_t^h b + N^2 w_a\,\mathrm{d}t = 0, \tag{2.7b}$$

Continuity equation

$$\boldsymbol{\nabla}_h\cdot\boldsymbol{u}_a + \partial_z w_a = 0, \tag{2.7c}$$

Geostrophic balances

$$\boldsymbol{u} = \frac{1}{f_0}\boldsymbol{\nabla}_h^\perp p, \quad \boldsymbol{\sigma}_h\mathrm{d}\boldsymbol{B}_t = \frac{1}{f_0}\boldsymbol{\nabla}_h^\perp\mathrm{d}p_t^\sigma, \quad \boldsymbol{\nabla}_h\cdot(\boldsymbol{u}^\star - \boldsymbol{u}) = 0, \tag{2.7d}$$

Hydrostatic balances

$$b = \partial_z p, \quad \mathrm{d}b_t^\sigma = \partial_z\mathrm{d}p_t^\sigma, \tag{2.7e}$$

where $\boldsymbol{u}$ and p denote the geostrophic components of the resolved horizontal velocity and the dynamic pressure, $(\boldsymbol{u}_a, w_a)^T$ and p_a denote the ageostrophic components, $f = f_0 + \beta y$ is the Coriolis parameter using beta-plane approximation, $\boldsymbol{k} = (0, 0, 1)^T$, $\boldsymbol{\nabla}_h = (\partial_x, \partial_y)^T$, $\boldsymbol{\nabla}_h^\perp = (-\partial_y, \partial_x)^T$, $N^2(z) = -g\partial_z\overline{\rho}(z)/\rho_0$ is the Brunt-Väisälä frequency with g the gravity value and $\overline{\rho}$ the stationary density in equilibrium that only depends on height, $b(\boldsymbol{x}, t) = -g\rho'(\boldsymbol{x}, t)/\rho_0$ is the resolved buoyancy associated with the density anomaly ρ' and the background density ρ_0, $\boldsymbol{\sigma}_h \mathrm{d}\boldsymbol{B}_t$ denotes the non-divergent horizontal noise resulting from the derivation, $\boldsymbol{u}^\star := \boldsymbol{u} - 1/2\, \nabla_h \cdot \boldsymbol{a}_h$ stands for the incompressible horizontal effective advection drift with $\boldsymbol{a}_h$ the quadratic variation of $\boldsymbol{\sigma}_h \mathrm{d}\boldsymbol{B}_t$ as defined in Equation 2.3, $\mathrm{d}p_t^\sigma/\mathrm{d}t$ and $\mathrm{d}b_t^\sigma/\mathrm{d}t$ (in a distribution sense) are the fast oscillating component of the pressure fluctuations and the buoyancy fluctuations, and $\mathbb{D}_t^h$ denotes the horizontal component of the stochastic transport operator, which is defined as

$$\mathbb{D}_t^h \Theta = \mathrm{d}_t \Theta + \left(\boldsymbol{u}^\star \,\mathrm{d}t + \boldsymbol{\sigma}_h \mathrm{d}\boldsymbol{B}_t\right) \cdot \boldsymbol{\nabla}_h \Theta - \frac{1}{2} \boldsymbol{\nabla}_h \cdot (\boldsymbol{a}_h \boldsymbol{\nabla}_h \Theta)\, \mathrm{d}t. \tag{2.8}$$

For sake of conciseness of the notations, we drop the horizontal subscript h in the following. We remark that both prognostic variables such as $(\boldsymbol{u}, p, b)$ and unresolved noise components such as $\left(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t, \mathrm{d}p_t^\sigma, \mathrm{d}b_t^\sigma\right)$ are vertically dependent. The unresolved buoyancy noise $\mathrm{d}b_t^\sigma$ is an additional component compared to the random model proposed by Resseguier et al. (2017b). This term can be seen as a thermodynamics noise coming from the nonlinear mapping law of state applied to the noise term of salinity and temperature transports. It is of the same nature as the noise term introduced in Brankart (2013). Combining the above equilibrium (Equations 2.7d and 2.7e), the thermal wind balances, $\partial_z \boldsymbol{u} \perp \nabla b$ and $\partial_z \boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t \perp \boldsymbol{\nabla} \mathrm{d}b_t^\sigma$, can be deduced as well. Note also that the proposed random system (Equations 2.7a–2.7e) is valid only under "moderate" noise scaling (see definition in Appendix A). As shown in Resseguier et al. (2017c), including stronger uncertainty could modify the classical geostrophic balance and a more complex random model would be obtained.

In order to represent the previous random system in terms of potential vorticity (PV) in the same way as in the classical framework (Vallis, 2017), we first take the (horizontal) curl of the first-order momentum (Equation 2.7a), then sum the resulting equation with the vertical derivative of the buoyancy (Equation 2.7b) rescaled by f_0/N^2, together with the equilibrium (Equations 2.7c–2.7e), to obtain finally

$$\begin{aligned} &\underline{\text{Stochastic evolution of PV}} \\ &\mathbb{D}_t q = \sum_{i=1,2} -\mathrm{J}\left((u^\star \,\mathrm{d}t + \sigma \mathrm{d}B_t)^i, u^i\right) + \frac{1}{2}\boldsymbol{\nabla} \cdot \left(\partial_{x_i}^\perp \boldsymbol{a} \boldsymbol{\nabla} u^i\right) \mathrm{d}t - \boldsymbol{\nabla} \cdot (\boldsymbol{a} \boldsymbol{\nabla} f)\, \mathrm{d}t \\ &\quad - \frac{f_0}{N^2} \partial_z (\boldsymbol{u}^\star \,\mathrm{d}t + \boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t) \cdot \boldsymbol{\nabla} b + \frac{f_0}{2N^2} \boldsymbol{\nabla} \cdot (\partial_z \boldsymbol{a} \boldsymbol{\nabla} b)\, \mathrm{d}t, \end{aligned} \tag{2.9a}$$

$$\begin{aligned} &\underline{\text{From PV to streamfunction}} \\ &q = \nabla^2 \psi + \partial_z \left(\frac{f_0}{N^2} b\right) + \beta y, \qquad b = f_0 \partial_z \psi \end{aligned} \tag{2.9b}$$

$$\begin{aligned} &\underline{\text{Incompressible constraints}} \\ &\boldsymbol{u} = \boldsymbol{\nabla}^\perp \psi, \qquad \boldsymbol{\nabla} \cdot \boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t = \boldsymbol{\nabla} \cdot (\boldsymbol{u}^\star - \boldsymbol{u}) = 0, \end{aligned} \tag{2.9c}$$

where q denotes the PV, $\psi := p/f_0$ stands for the streamfunction of the resolved flow, $\mathrm{J}(f, g) = \partial_x f \partial_y g - \partial_x g \partial_y f$ stands for the Jacobian operator, $\nabla^2 = \partial_{xx}^2 + \partial_{yy}^2$ is the two-dimensional (2D) Laplacian operator and the index i denotes the dimension. Note that in the deterministic setting, the sources and sinks of PV in Equation 2.9a cancel due to the anti-symmetry of the Jacobian operator, $\mathrm{J}(u^i, u^i) = 0$, and to the thermal wind balance, $\partial_z \boldsymbol{u} \perp \nabla b$. However, it should be noted that the presence of sub-grid terms in a large-scale deterministic Boussinesq system leads also in general to additional terms in the RHS of the PV equation. The additional terms in the RHS of Equation 2.9a are coming exactly from the same sources of difficulty inherent to large-scale flow dynamics representations (i.e., non-commutation of curl operator with the sub-grid model, breaking of the Jacobian anti-symmetry as well as possibly non-commutation of coarsening processes with derivatives).

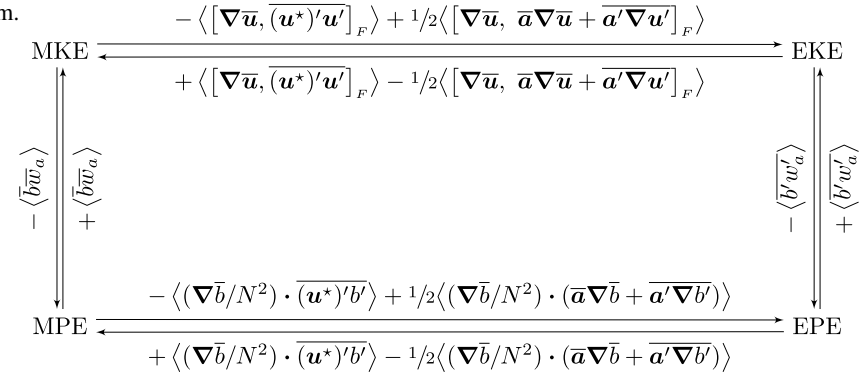
2.4. Energy Conservation and Transfers

One important characteristic of the random system (Equations 2.7a–2.7e) is that it conserves path-wise (i.e., for each realization) the total energy of the resolved geostrophic flow (under natural boundary conditions), namely

$$\mathrm{d}_t \int_D \frac{1}{2}\left(|\boldsymbol{u}|^2 + \frac{b^2}{N^2}\right)\mathrm{d}\boldsymbol{x} = \left(\int_D w_a b\,\mathrm{d}\boldsymbol{x} - \int_D w_a b\,\mathrm{d}\boldsymbol{x}\right)\mathrm{d}t = 0. \tag{2.10}$$

The demonstration can be found in Appendix B. We remark that this conservation property is consistent with the energy conservation of the stochastic barotropic QG system (Bauer et al., 2020a) and of the stochastic rotating shallow-water system (Brecht et al., 2021).

The energy conversions between the ensemble-mean and ensemble-eddy components are in addition explored in Appendix C. Based on the decompositions $\boldsymbol{u} = \mathbb{E}[\boldsymbol{u}] + (\boldsymbol{u} - \mathbb{E}[\boldsymbol{u}]) := \overline{\boldsymbol{u}} + \boldsymbol{u}'$ and $b = \overline{b} + b'$, the mean kinetic energy (MKE), the eddy kinetic energy (EKE), the mean potential energy (MPE) and the eddy potential energy (EPE) are respectively defined by $1/2\langle|\overline{\boldsymbol{u}}|^2\rangle$, $1/2\langle\overline{|\boldsymbol{u}'|^2}\rangle$, $1/2\langle\left(\overline{b}/N\right)^2\rangle$ and $1/2\langle\overline{(b'/N)^2}\rangle$, where $\langle\cdot\rangle$ denotes the integration over domain for simplicity. The energy conversions of the proposed random model (Equations 2.7a–2.7e) are summarized in the following diagram.



Here, $[\boldsymbol{A},\boldsymbol{B}]_F = \mathrm{Tr}(\boldsymbol{A}^T\boldsymbol{B})$ stands for the Frobenius inner product of the tensors $\boldsymbol{A}$ and $\boldsymbol{B}$. The ensemble-mean of the path-wise energy conservation (Equation 2.10) can be recovered by summing all the conversion terms.

Compared to the energy budget of the classical deterministic system, the random system provides additional conversion terms between the mean and eddy components. In the general setting with $\boldsymbol{\sigma}$ random, the classical geostrophic Reynolds stress (Bachman, 2019) is modified to $\overline{(\boldsymbol{u}^\star)'\boldsymbol{u}'}$ due to the possibly inhomogeneous nature of the noise. This term includes an additional term $1/2\nabla\cdot\boldsymbol{a}$ that can be positive or negative depending on its alignment with the large-scale velocity component. The fact that $\overline{\boldsymbol{a}}\,\mathrm{d}t = \mathrm{Var}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)$ is symmetric and non-negative definite, ensures that the integral $\left\langle\left[\nabla\overline{\boldsymbol{u}},\;\overline{\boldsymbol{a}}\nabla\overline{\boldsymbol{u}}\right]_F\right\rangle$ (resp. $\left\langle\left[\nabla\overline{b},\;\overline{\boldsymbol{a}}\nabla\overline{b}\right]_F\right\rangle$) is always positive, hence provides a positive transfer from MKE to EKE (resp. from MPE to EPE). When the correlation tensor $\boldsymbol{\sigma}$ is nonrandom, only the terms depending on $\overline{\boldsymbol{a}}$ remain in addition to the term related to the classical geostrophic Reynolds stress. As the stochastic system (Equations 2.7a–2.7e) does not contain vertical noise components, it does not explicitly modify the conversion mechanisms between KE and PE as observed in the classical system. Nevertheless, the effect of the stochastic terms on this energy transfers is implicitly accounted via the time-integration of the resolved variables b and w_a.

2.5. Girsanov Transformation

The previous formulations (Equations 2.7a–2.7e) and (Equations 2.9a–2.9c) involve a zero-mean and temporally uncorrelated noise. However, this might not be enough in practice and including a mean or time-correlated component of the unresolved flow component could be of crucial importance to obtain a relevant model. This can be found in many works of the literature. For instance, Zanna et al. (2017) propose both deterministic and stochastic parameterizations for their eddy forcing. The mean component of their stochastic eddy forcing keeps the same form as their deterministic parameterization, which is based on a negative Laplacian of the PV transport. For the double-gyre circulation configuration, the considered deterministic parameterization allows to reproduce the eastwards jet for the coarse-resolution model, while the additional zero-mean terms enhance the gyres circulation and improve the flow variability. Similarly, the random-forcing model proposed by Berloff (2005) consists in a space-time correlated stochastic process to enhance the jet extension.

The slow components of the sub-grid scales can be provided by adequate high-pass filtering of HR data on the coarse grid. We aim in this work at investigating the incorporation of such slow components within the LU frame-

work. However, the derivation of LU models (Bauer et al., 2020a; Mémin, 2014; Resseguier et al., 2017a) relies on the martingale properties of the centered noise and we need hence to properly handle non-centered Brownian terms. The Girsanov transformation (Da Prato & Zabczyk, 2014) provides a theoretical tool that fully warrants such a superposition: by a change of the probability measure, the composed noise can be centered with respect to a new probability measure while an additional drift term appears, which pulls back time-correlated sub-grid-scale components into the dynamical system. The associated mathematical description is given as follows.

Let $\boldsymbol{\Gamma}_t$ be an H-valued $\mathcal{F}_t$-predictable process satisfying $\mathbb{E}_{\mathbb{P}}\left[\exp\left(\frac{1}{2}\int_0^T \|\boldsymbol{\Gamma}_t\|_H^2 \,\mathrm{d}t\right)\right] < +\infty$, then the process $\left\{\tilde{\boldsymbol{B}}_t := \boldsymbol{B}_t + \int_0^t \boldsymbol{\Gamma}_s \,\mathrm{d}s\right\}_{0\le t\le T}$ is an H-valued cylindrical Wiener process on $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\le t\le T}, \tilde{\mathbb{P}}\right)$ by a change of the probability measure, $\mathrm{d}\tilde{\mathbb{P}}/\mathrm{d}\mathbb{P} = \exp\left(-\int_0^T \langle \boldsymbol{\Gamma}_t, \mathrm{d}\boldsymbol{B}_t\rangle_H - \frac{1}{2}\int_0^T \|\boldsymbol{\Gamma}_t\|_H^2 \,\mathrm{d}t\right)$, where $\langle \boldsymbol{f}, \boldsymbol{g}\rangle_H = \int_D \left(\boldsymbol{f}^T \boldsymbol{g}\right)(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$ denotes the inner product on $H = \left(L^2(D)\right)^d$. Under this new probability measure $\tilde{\mathbb{P}}$, the previous system (Equations 2.9a–2.9c) must be adapted to the modified noise $\boldsymbol{\sigma}\mathrm{d}\tilde{\boldsymbol{B}}_t$ with $\mathbb{E}_{\tilde{\mathbb{P}}}\left[\boldsymbol{\sigma}\mathrm{d}\tilde{\boldsymbol{B}}_t\right] = \boldsymbol{0}$ by a correction of the effective drift

$$\widetilde{\boldsymbol{u}^\star} := \boldsymbol{u}^\star - \boldsymbol{\sigma}\boldsymbol{\Gamma}_t. \tag{2.11}$$

Hereafter, the resulting stochastic system will be adopted with $\boldsymbol{\sigma}\boldsymbol{\Gamma}_t$ referred to as the *Girsanov drift*. We remark that the proposed change of probability measure will not alter any other terms in Equations 2.9a and 2.9c, as all the noise terms are driven by the same Brownian motions and both $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$, $\boldsymbol{\sigma}\mathrm{d}\tilde{\boldsymbol{B}}_t$, $\boldsymbol{\sigma}\boldsymbol{\Gamma}_t$ share the same spatial structure imposed through the correlation tensor $\boldsymbol{\sigma}$.

3. Numerical Models

In this section, we first describe the discretization of the continuous stochastic QG system (under a simplification) that will be used later for the numerical simulations. Then, we present numerical methods to estimate the spatial correlation functions of the unresolved flow component based on the spectral decomposition (Equation 2.2).

3.1. Discrete QG Model

We consider here a vertically discretized QG model. This formulation is quite common in geophysical fluid dynamics and the derivation follows the standard methods (McWilliams, 2006; Vallis, 2017). As illustrated in Figure 1, such model consists of n ocean layers with constant thickness H_k and density ρ_k in each layer k. The prognostic variables (q_k, ψ_k, $\boldsymbol{u}_k$) and the unresolved flow structures ($\boldsymbol{\sigma}_k$, $\boldsymbol{a}_k$) are assumed to be layer-averaged quantities.

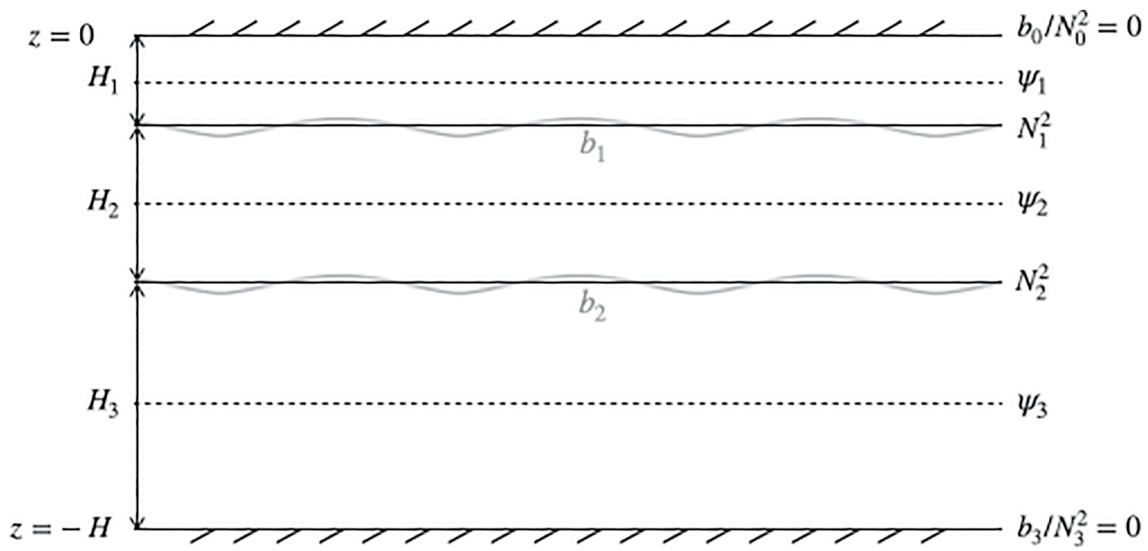


Figure 1. Illustration for the vertical discretization of a three-layered quasi-geostrophic system.

We first apply finite differences for the buoyancy frequency N^2 and anomaly b on the interface between layers k and $k+1$ (for $k = 1, \ldots, n-1$):

$$b_k = \frac{f_0(\psi_k - \psi_{k+1})}{(H_k + H_{k+1})/2}, \qquad N_k^2 = \frac{g'_k}{(H_k + H_{k+1})/2}, \tag{3.1}$$

where $g'_k := -g(\rho_k - \rho_{k+1})/\rho_0$ is introduced as the reduced gravity across the interface k. Applying subsequently finite differences for the stratification term $\partial_z(f_0 b/N^2)$ with the zero boundary conditions at ocean surface and bottom, $b_0/N_0^2 = b_n/N_n^2 = 0$, Equation 2.9b can be vertically discretized by

$$\nabla^2\psi_k + \frac{f_0^2}{H_k}\left(\frac{\psi_{k-1} - \psi_k}{g'_{k-1}} - \frac{\psi_k - \psi_{k+1}}{g'_k}\right) = q_k - \beta y. \tag{3.2a}$$

In this work, we adopt the dynamical core of the QG coupled model (Q-GCM) developed by Hogg et al. (2003), which integrates numerically the state variables q and ψ. The RHS terms of Equation 2.9a have much higher complexity than the left-hand side (LHS) terms. In practice, they require more interpolations in both horizontal and vertical directions and it is difficult to find an equivalent conservative scheme (will be discussed later) as used for the LHS terms. In order to avoid the additional numerical difficulties, we do not include these RHS terms of Equation 2.9a in this work. These PV source terms will be naturally considered when working in $\boldsymbol{u}$–b formulations such as in the primitive equations dynamical core (Bachman, 2019) or for multi-layered shallow water systems (Adcroft et al., 2019). In our case, for the type of noises considered (with moderated amplitude so as to respect the QG scaling performed in this work) these additional noise terms are small. For stronger noise, with de facto other scalings with potential modification of the hydrostatic balances these additional noise terms have to be considered in the PV equation.

Taking into account the numerical dissipation and the vertical entrainment e_k (across interfaces), the simplified stochastic evolution of PV reads

$$\tilde{\mathbb{D}}_t^k q_k = \left(-A_4\nabla^4\left(\nabla^2\psi\right) + \frac{f_0}{H_k}(e_{k-1} - e_k)\right)\mathrm{d}t, \tag{3.2b}$$

where $\tilde{\mathbb{D}}_t^k$ denotes the kth layer stochastic transport operator under the probability measure $\tilde{\mathbb{P}}$ (hence includes the Girsanov drift described in Section 2.5), $\nabla^4 = \nabla^2(\nabla^2)$ is the bi-Laplacian operator associated with the dissipation coefficient A_4.

As for the spatial discretization, we propose to use a conservative flux form $\nabla\cdot(\boldsymbol{u}q)$ that corresponds exactly to the 9-points Arakawa Jacobian scheme (Arakawa & Lamb, 1981) of J(ψ, q), in which the zonal and meridional components of the advection flux are defined as $2/3\left(\overline{\overline{u}^x}^y\overline{q}^x\right) + 1/3\left(\overline{\overline{u}^x}^d\right)$ and $2/3\left(\overline{\overline{v}^y}^x\overline{q}^y\right) + 1/3\left(\overline{\overline{v}^x}^d\right)$. Here, $\bar{\cdot}^x$, $\bar{\cdot}^y$ and $\bar{\cdot}^d$ denote the central averaging operations in the zonal, meridional and diagonal directions respectively. We use this conservative form to discretize the advection of PV by the noise and the Girsanov drift. On the other hand, we employ simply a central winding scheme to discretize the inhomogeneous diffusion term $\nabla\cdot(\boldsymbol{a}\nabla q)$. We adopt a stochastic Leapfrog scheme (Ewald & Témam, 2005) for the time-stepping of Equation 3.2b. The inversion of Equation 2.9b is carried out with a discrete Fourier transform method (Hogg et al., 2003).

The vertical entrainment e in the classical QG models (Berloff et al., 2007; Grooms et al., 2015; Marshall et al., 2012; Zanna et al., 2017) accounts only for the Ekman pumping e_0 and e_n for the upper and bottom boundary layers, that is, $e_k = 0$ for $k = 1, \ldots, n-1$. The Ekman pumpings are defined by $e_0 = \nabla^\perp\cdot\boldsymbol{\tau}/f_0$ and $e_n = 1/2\,\delta_{ek}\nabla^2\psi_n$, where $\boldsymbol{\tau}$ denotes a given surface wind stress and δ_{ek} stands for a prescribed thickness of the bottom Ekman layer. In the Q-GCM formulation (Hogg et al., 2004), an additional entrainment e_1 is considered due to the Ekman pumping of mixed layer temperature. This will be further detailed in Section 4.3. All the parameter values of the model configuration will be provided later in Section 4.1.

3.2. Parameterizations of Unresolved Flow

This section describes the numerical methods of the parameterization for the unresolved flow structure. We first review the empirical orthogonal functions (EOF) method which relies on an assumption of quasi-stationarity of

the noise covariance. To relax this stationarity constraint, we then propose a projection method enabling to update on-line the noise spatial correlations.

Let $\{\boldsymbol{u}_{\mathrm{HR}}(\boldsymbol{x},t_i)\}_{i=1,\dots,N}$ be the finite set of velocity snapshots provided by a HR simulation. We first build the spatial local fluctuations $\boldsymbol{u}_f(\boldsymbol{x},t_i)$ of each snapshot at the coarse-grid points. In particular, for the QG system, one can first perform a high-pass filtering with a 2D Gaussian convolution kernel G on each HR streamfunction ψ_{HR}, to obtain the streamfunction fluctuations, $\psi_f(\boldsymbol{x},t_i)=((\mathrm{I}-\mathrm{G})\star\psi_{\mathrm{HR}})(\boldsymbol{x},t_i)$ (only for the coarse-grid points $\boldsymbol{x}$). Then, the geostrophic velocity fluctuations can be derived by $\boldsymbol{u}_f=\boldsymbol{\nabla}^{\perp}_{\mathrm{LR}}\psi_f$. We next center the data set by $\boldsymbol{u}'_f=\boldsymbol{u}_f-\overline{\boldsymbol{u}_f}^t$ (with $\overline{\cdot}^t$ the temporal mean) and perform the EOF procedure (Sirovich, 1987) to get a set of orthogonal temporal modes $\{\alpha_n\}_{n=1,\dots,N}$ and orthonormal spatial modes $\{\boldsymbol{\phi}_n\}_{n=1,\dots,N}$ satisfying $\boldsymbol{u}'_f(\boldsymbol{x},t_i)=\sum_{n=1}^{N}\alpha_n(t_i)\boldsymbol{\phi}_n(\boldsymbol{x})$, where $\langle\boldsymbol{\phi}_m,\boldsymbol{\phi}_n\rangle_{\mathcal{H}}=\delta_{m,n}$, $\overline{\alpha_m\alpha_n}^t=\lambda_n\delta_{m,n}$, $\mathcal{H}:=L^2(\mathcal{D})$ and $\mathcal{D}=\mathcal{A}\times[-H,0]$ denotes the 3D domain with the horizontal area $\mathcal{A}$ and the total depth H.

Truncating subsequently the modes (with $M\ll N$) and rescaling them by a small-scale decorrelation time τ, the kth layer noise and its variance are built by

$$\boldsymbol{\sigma}_k(\boldsymbol{x})\,\mathrm{d}\tilde{\boldsymbol{B}}_t=\tau^{1/2}\sum_{n=1}^{M}\lambda_n^{1/2}\boldsymbol{\phi}_{n,k}(\boldsymbol{x})\,\mathrm{d}\beta_n(t), \tag{3.3a}$$

$$\boldsymbol{a}_k(\boldsymbol{x})=\tau\sum_{n=1}^{M}\lambda_n\boldsymbol{\phi}_{n,k}(\boldsymbol{x})\boldsymbol{\phi}_{n,k}^{T}(\boldsymbol{x}). \tag{3.3b}$$

Note that this time scale τ is used to match the fact that the noise in Equation 2.1 admits units of meters. In practice, we often consider the coarse-grid simulation time-step Δt_{LR}.

In this work, we simply propose to represent the Girsanov drift included in Equation 2.11 by the temporal mean of the unresolved spatial fluctuations:

$$\boldsymbol{\sigma}_k(\boldsymbol{x})\boldsymbol{\Gamma}_t=\sum_{n=1}^{N}\langle\overline{\boldsymbol{u}_f}^t,\boldsymbol{\phi}_n\rangle_{\mathcal{H}}\boldsymbol{\phi}_{n,k}(\boldsymbol{x}). \tag{3.3c}$$

We remark that in future works this stationary drift term could be generalized to slowly varying fluctuation components or to drift correction with respect to observations (Dufée et al., 2022).

The previous EOF procedure is an efficient off-line learning method, yet it relies on a strong stationary assumption, and hence leads to a sequence of random flow fields with no connection with the resolved dynamics. In the following, we propose a novel approach that projects the EOF-based noise on a resolved state variable. In order to enforce the noise to act only on the resolved momentum (Equation 2.7a) and provide efficient backscattering of KE, we consider constraining the noise along the iso-surfaces of the stratification, $\tilde{\boldsymbol{\sigma}}\mathrm{d}\tilde{\boldsymbol{B}}_t\perp\boldsymbol{\nabla}\partial_z\tilde{b}$ with $\tilde{b}:=f_0b/N^2$. This projection procedure reads

$$\tilde{\boldsymbol{\sigma}}_k(\boldsymbol{x},t)\,\mathrm{d}\tilde{\boldsymbol{B}}_t=A^{1/2}\,\mathbf{P}_k(\boldsymbol{x},t)\,\boldsymbol{\sigma}_k(\boldsymbol{x})\,\mathrm{d}\tilde{\boldsymbol{B}}_t,\qquad \mathbf{P}_k:=\mathbf{I}_2-\frac{\boldsymbol{\nabla}\left(\partial_z\tilde{b}\right)_k\left(\boldsymbol{\nabla}\left(\partial_z\tilde{b}\right)_k\right)^T}{|\boldsymbol{\nabla}\left(\partial_z\tilde{b}\right)_k|^2}, \tag{3.4}$$

where $\boldsymbol{\sigma}_k\mathrm{d}\tilde{\boldsymbol{B}}_t$ is generated as in Equation 3.3a and $A:=\|\boldsymbol{\sigma}\mathrm{d}\tilde{\boldsymbol{B}}_t\|^2_{\mathcal{H}}\,/\|\boldsymbol{P}\,\boldsymbol{\sigma}\mathrm{d}\tilde{\boldsymbol{B}}_t\|^2_{\mathcal{H}}$ is a path-wise and time-dependent scaling factor to ensure that the noise amplitude is projection invariant. This rescaling will allow us in particular to evaluate the net performance of the projection method. From the definition of the noise quadratic variation (Equation 2.3), tensor $\boldsymbol{a}$ defined through Equation 3.3b is updated as $\tilde{\boldsymbol{a}}_k=A\,\mathbf{P}_k\,\boldsymbol{a}_k\,\mathbf{P}_k^T$. In the next section, we show that the resulting non-stationary noise leads to significant improvement of variability for coarse-resolution simulations.

4. Numerical Results

In this section, we discuss and compare numerical results of several models run within different configurations. The objective consists in improving the variability of low-order models at a low Reynolds numbers flow, and in the particular case of unresolved baroclinic instabilities.

Table 1
Common Parameters for All the Models

Parameters	Value	Description
$X \times Y$	(3,840 × 4,800) km	Domain size
H_k	(350, 750, 2,900) m	Mean layer thickness
ρ	1,000 kg m^3	Density
g'_k	(0.025, 0.0125) m s^{-2}	Reduced gravity
δ_{ek}	2 m	Bottom Ekman layer thickness
τ_0	2×10^{-5} m^2 s^{-2}	Wind stress magnitude
α_{bc}	0.2	Mixed boundary condition coefficient
f_0	9.375×10^{-5} s^{-1}	Mean Coriolis parameter
β	1.754×10^{-11} (m s)$^{-1}$	Coriolis parameter gradient
L_d	(39, 22) km	Baroclinic Rossby radii

4.1. Model Configurations

We consider here a finite box ocean at mid-latitude driven by an idealized (stationary and symmetric) wind stress $\boldsymbol{\tau} = (-\tau_0 \cos(2\pi y)/L_y, 0)^T$. A mixed horizontal boundary condition is used for the kth layer streamfunction: $\psi_k|_{\partial\mathcal{A}} = f_k(t)$ and $\partial^2_{\boldsymbol{nn}}\psi_k|_{\partial\mathcal{A}} = -(\alpha_{bc}/\Delta x)\partial_{\boldsymbol{n}}\psi_k|_{\partial\mathcal{A}}$ (same for the 4th order normal derivative). Here, $\partial_{\boldsymbol{n}} := \boldsymbol{n} \cdot \nabla$ denotes the derivative in the unit normal vector direction $\boldsymbol{n}$, the value of the time-dependent function $f_k(t)$ is constrained by mass conservation following McWilliams (1977) and is the same for all boundaries, Δx stands for the horizontal resolution and α_{bc} is a nondimensional coefficient associated to the slip boundary conditions (Haidvogel et al., 1992).

A deterministic eddy-resolving (Δx = 5 km) model is first simulated and referred to as the references (REF). It is then compared to several coarse-resolution (Δx = 40 km, 80 km, 120 km) models: the benchmark deterministic model (DET), two stochastic models one with a stationary EOF-based noise (STO-EOF) and one with a non-stationary projection noise (STO-EOF-P). The eddy-resolving model starts from a quiescent initial condition, whereas spin-up conditions downsampled from REF (after 90-years integration) are adopted for all the coarse-resolution models. The common parameters for all the simulations are listed in Table 1, whereas resolution-dependent parameters are presented separately in Table 2.

Snapshots of the relative vorticity on the two top layers are shown in Figure 2. The dynamics of the REF model (Figures 2a and 2e) is mainly characterized by the energetic eastward jet with adjacent recirculations and by the fast westward Rossby waves. The effect of mesoscale scale eddies via nonlinear barotropic and baroclinic instabilities are mostly captured in the highly turbulent jet region. However, the baroclinic instability cannot be resolved once the horizontal resolution is similar or greater than the baroclinic deformation radius maximum (39 km here). For instance, without any eddy parameterization, the DET (40 km) simulation (Figures 2b and 2f) only generates a smooth field due to an over-dissipation effect.

To calibrate the structure of the unresolved flow (noise and Girsanov drift), the REF data are sampled at each day during the first 10 years after the spin-up, this is followed by the procedure presented in Section 3.2 with a fixed 200 km wide Gaussian filter. Considering the computational cost, we employ only 50 EOFs to build the noises (Equation 3.3a) for the stochastic simulations at different resolutions. However, the eigenvalues of these truncated EOFs are amplified in order to capture 95% energy of the full set of modes. An example of the first EOF $\sqrt{\lambda_1}\boldsymbol{\phi}_1$ is illustrated by Figures 3a and 3b and the resulting variance $\boldsymbol{a}$ of the noise is shown in Figures 3c and 3d. The noise structure is mainly characterized by the most energetic fluctuations in the jet region.

Including the additional advection of PV by the calibrated unresolved flow component, both STO-EOF and STO-EOF-P models are able to reproduce the eastward jet on the coarse mesh, and the latter seems to enhance fluctuations everywhere. For instance, such result is illustrated at a resolution of 40 km by Figures 2c, 2d, 2g, and 2h. The ability of these coarse models to reproduce the statistical properties of the REF model will be diagnosed and analyzed more precisely in the next section.

Table 2
Values of Grid Varying Parameters

Δx (km)	Δt (s)	A_4 (m^4 s^{-1})
5	600	2×10^9
40	1,200	5×10^{11}
80	1,440	5×10^{12}
120	1,800	1×10^{13}

4.2. Statistical Diagnostics

To investigate the prediction of the long-term statistics, we sample the data (streamfunction) of each run at 15-day intervals over the 120 years (after the spin-up) and compute the statistics over the last 100 years.

We first compare the temporal mean of the modal streamfunctions for both coarse models to that subsampled (in space) of the eddy-resolving model. We observe that both barotropic (Figure 4a) and baroclinic (Figure 4e) modal mean of the REF characterize the eastwards jet, and the two stochastic models (Figures 4c, 4d, 4g, and 4h) at 40 km enable to reproduce qualitatively the

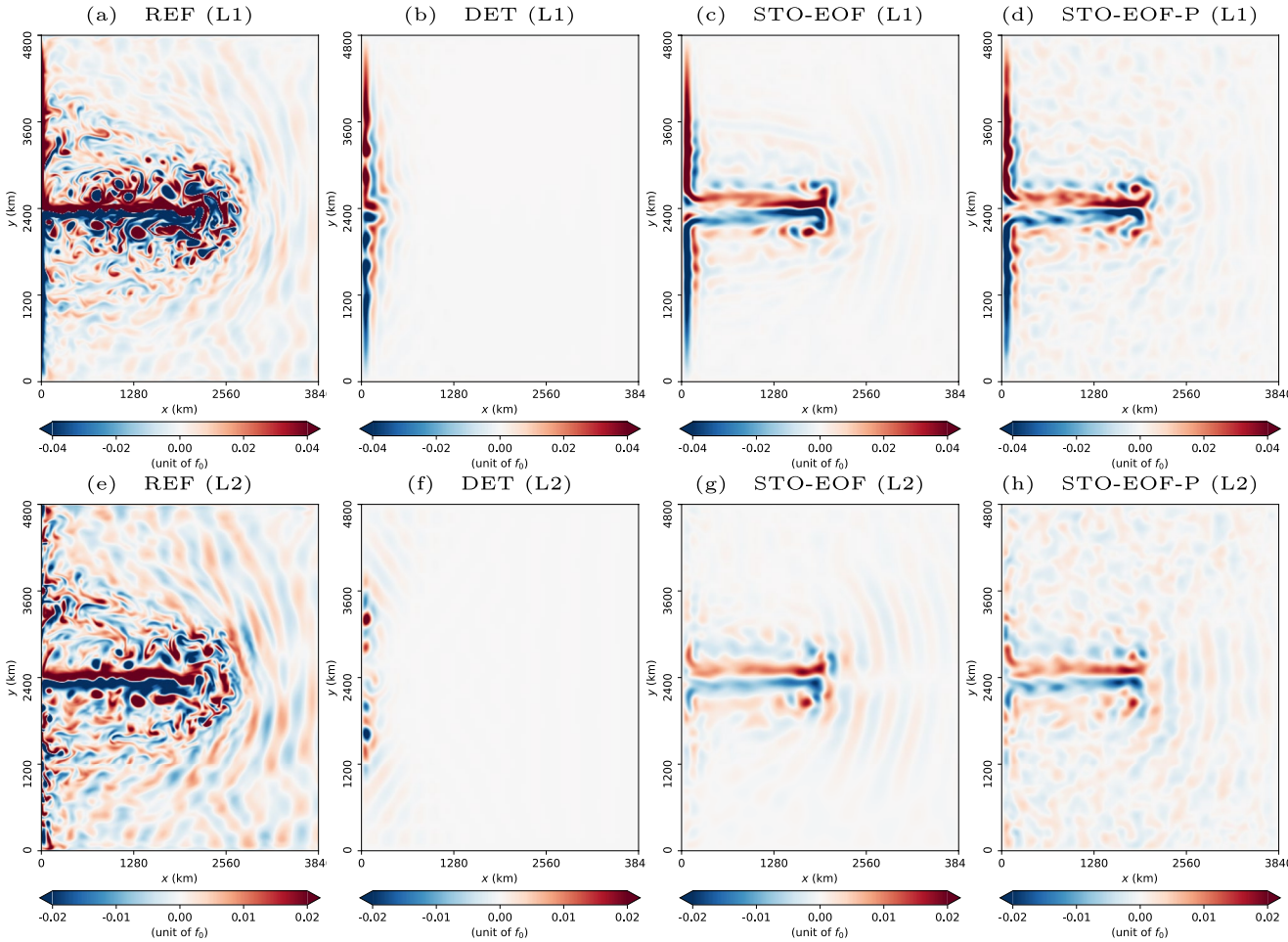


Figure 2. Snapshots of relative vorticity (divided by f_0) for the two upper layers (L1 and L2 by rows) provided by different simulations (by columns) after 60-years integration. All the coarse models plotted (DET, STO-EOF, and STO-EOF-P) are at a resolution of 40 km.

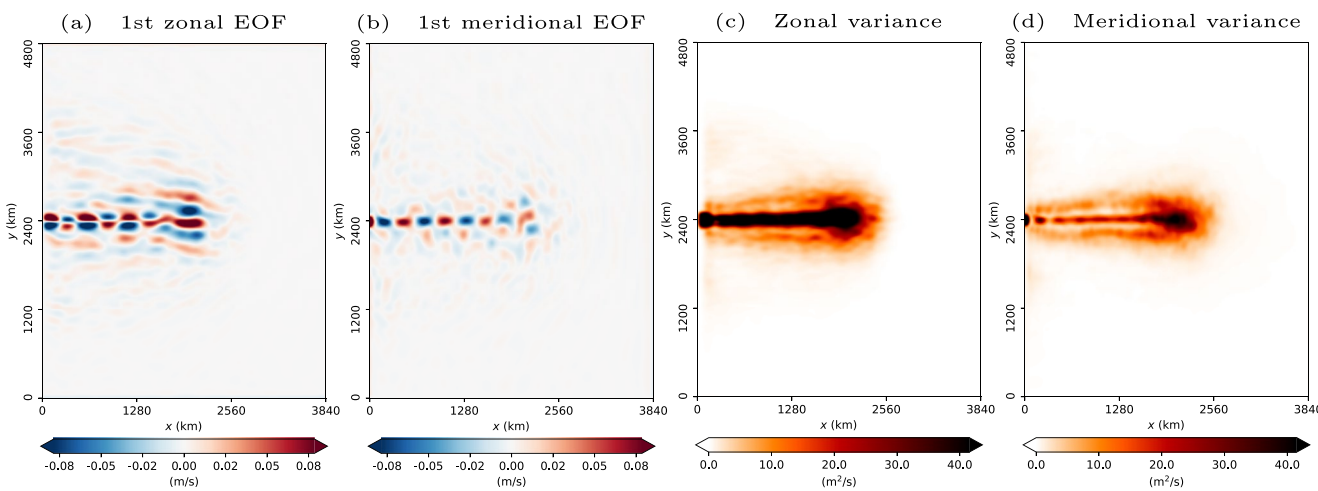


Figure 3. Illustrations for (a) first empirical orthogonal functions (EOF) of zonal noise component $\sqrt{\lambda_1}\phi_1^x$, (b) first EOF of meridional noise component $\sqrt{\lambda_1}\phi_1^y$, (c) variance of zonal noise component a_{xx}, (d) variance of meridional noise component a_{yy}. We recall that the variance tensor $\boldsymbol{a}$ in this case is defined in Equation 3.3b.

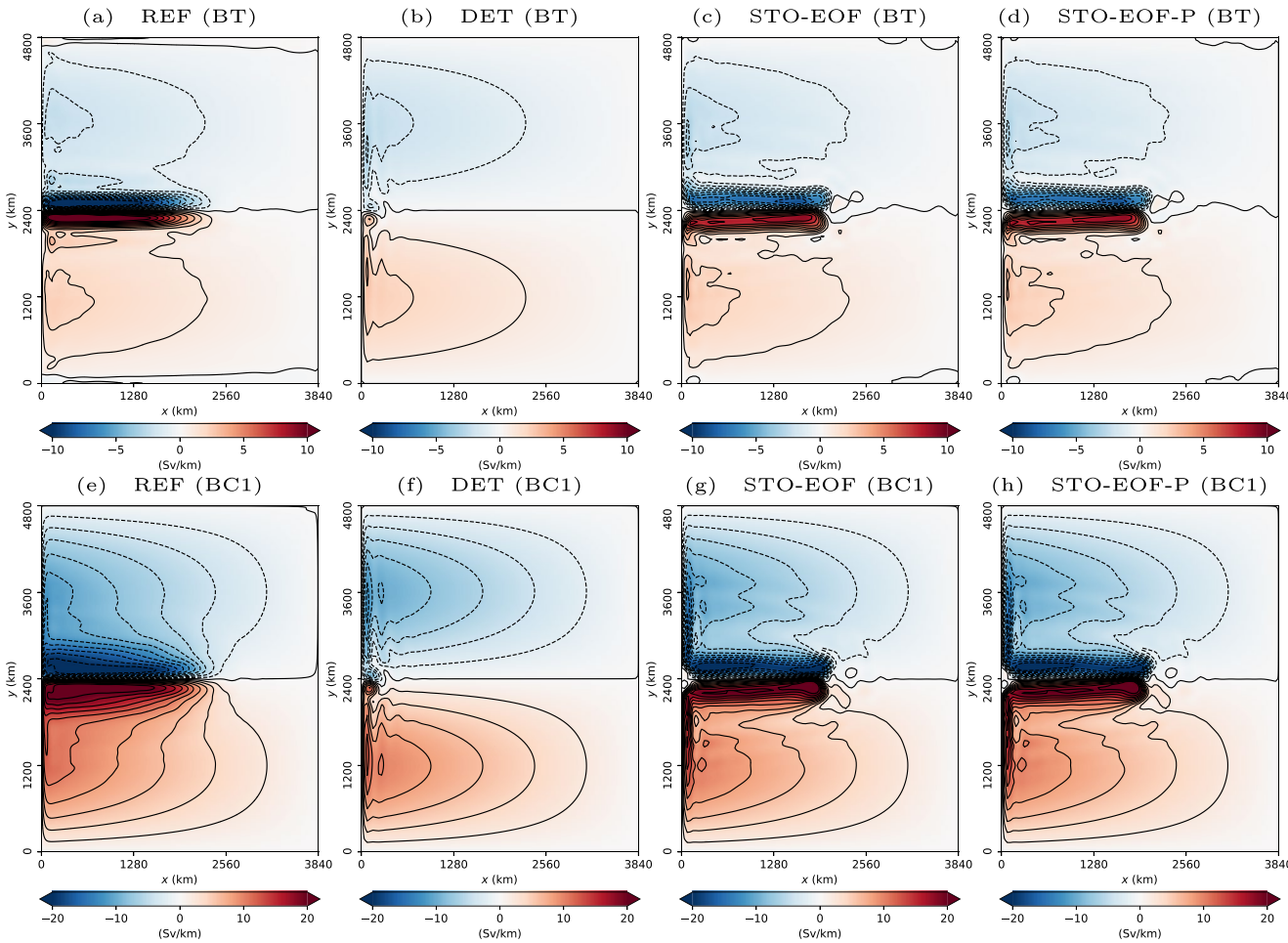


Figure 4. Comparison of time-mean contour of barotropic (BT, top row) and first baroclinic (BC1, bottom row) streamfunctions (in Sv/km with 1 Sv = 10^6 m^3/s) for different models (by columns). All the coarse models (DET, STO-EOF, and STO-EOF-P) plotted are at a resolution of 40 km.

local structures of both vertical modes predicted by the REF. Conversely, the DET model (Figures 4b and 4f) can only capture the symmetric double-gyre structure. We note that both STO-EOF (Figure 4g) and STO-EOF-P (Figure 4h) models produce stronger circulation along the western boundary than the DET model (Figure 4f). This may be related to boundary condition effects spreading too far into the domain due to the coarse resolution.

To quantify the models' statistics at different resolutions, we then employ the spatial root-mean-square error (RMSE) of the temporal mean and the standard deviation (std) between each coarse model and the subsampled REF one. To assess in a single measure the mean and the variance reconstruction, we adopt the Gaussian relative entropy (GRE) proposed by Grooms et al. (2015) that is defined as

$$\mathrm{GRE} = \frac{1}{|\mathcal{D}|}\int_{\mathcal{D}} \frac{1}{2}\left(\frac{\left(\overline{\psi^{\mathrm{M}}}^{t} - \overline{\psi^{\mathrm{R}}}^{t}\right)^2}{\sigma^2\left[\psi^{\mathrm{M}}\right]} + \frac{\sigma^2\left[\psi^{\mathrm{R}}\right]}{\sigma^2\left[\psi^{\mathrm{M}}\right]} - 1 - \log\left(\frac{\sigma^2\left[\psi^{\mathrm{R}}\right]}{\sigma^2\left[\psi^{\mathrm{M}}\right]}\right)\right) \mathrm{d}\boldsymbol{x}, \tag{4.1}$$

where σ denotes the std, ψ^{M} (resp. ψ^{R}) stands for the streamfunction of the coarse model (resp. for that subsampled from the REF). The lower is the value of this criterion, the better are the reconstructions of both the mean and variance.

These three global criterions are computed for all the coarse models at different resolutions and the results are summarized in Figure 5. For all the resolutions tested the two random models have lower errors than the

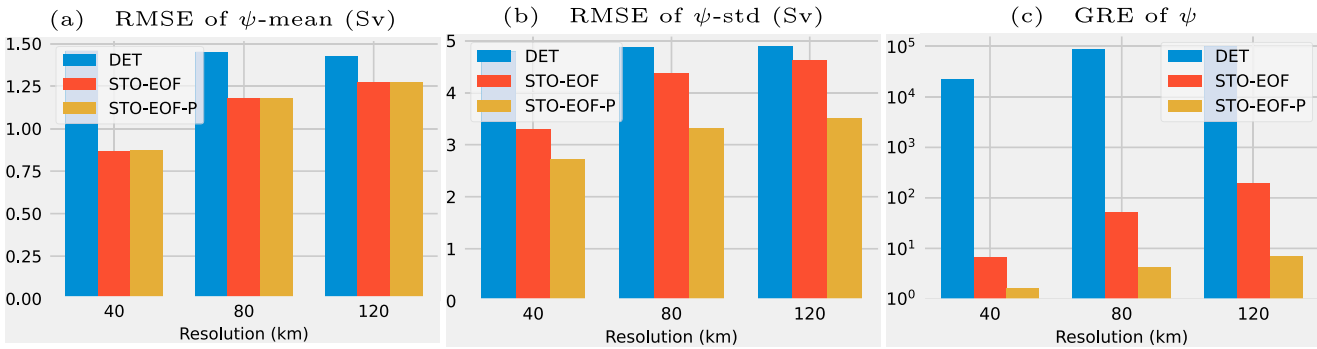


Figure 5. Comparison of variability measures (by sub-figures) for different coarse models (by colors and groups in each sub-figure). Note that the *y*-axis of the last sub-figure is in log-scale.

DET model both in terms of mean (Figure 5a) and std (Figure 5b). In addition, the STO-EOF-P models enable to reduce the RMSE of the std and the GRE (Figure 5c), which means an improvement on the prediction of variability.

We investigate subsequently the models' variability based on the energy analysis. The time-averaged KE spectra provided by different models are compared in Figure 6. Unsurprisingly, compared to the REF, the DET coarse models produce a severe lack of resolved KE (Arbic et al., 2013; Kjellsson & Zanna, 2017) due to the excessive dissipation without any eddy parameterization. Both stochastic models bring more energy over all the wavenumbers, particularly at large scales. Moreover, when the resolution increases (from Figures 6a–6c), their spectral slopes in the inertial range become gradually closer to that of the REF model.

We compare finally the energy decomposition as described in Section 2 for all the models. However, we focus only on a single realization here and the eddy-mean decomposition of $\boldsymbol{u}$ and b is simply build from a global time-average. We observe that both random models have higher MKE (Figure 7a) and MPE (Figure 7c) than the DET model. In particular, at resolution 40 km, they are almost at the same order as that of the REF model. Compared to the STO-EOF models, STO-EOF-P models produce globally higher EKE (Figure 7b) and EPE (Figure 7d) at all the different resolutions. However, they remain much lower than that of the REF ones.

Furthermore, the EKE and EPE density provided by different models are compared in Figure 8. It shows that the STO-EOF-P model (Figures 8d and 8h) improves locally the eddy energy almost everywhere except in the jet region. A more precise parameterization of the unresolved flow is required to explore in future works how to improve locally the energy transfers.

In summary, the diagnostic results of this section demonstrate that including the unresolved flow structure inferred from the data into the proposed random models enable to reproduce on the coarse mesh the mean flow of the eddy-resolving simulation. In addition, as performed here through a projection onto the iso-surfaces of vertical stratification, providing an adequate dynamics for the unresolved flow structure is a good strategy to increase the variability of the coarse models.

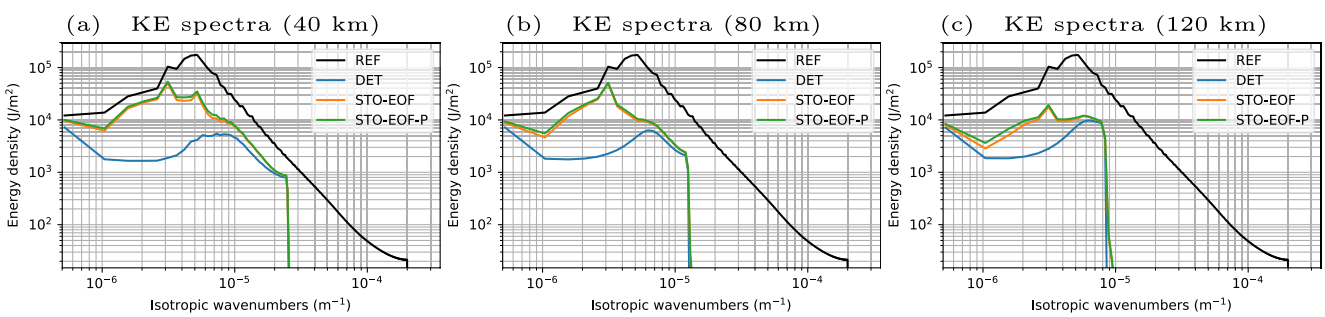


Figure 6. Temporal mean of vertically integrated kinetic energy (KE) spectrum for different models (by colors) at different resolutions: (a) 40 km, (b) 80 km, and (c) 120 km. The same KE spectrum (black line) of the REF (5 km) is added in each panel.

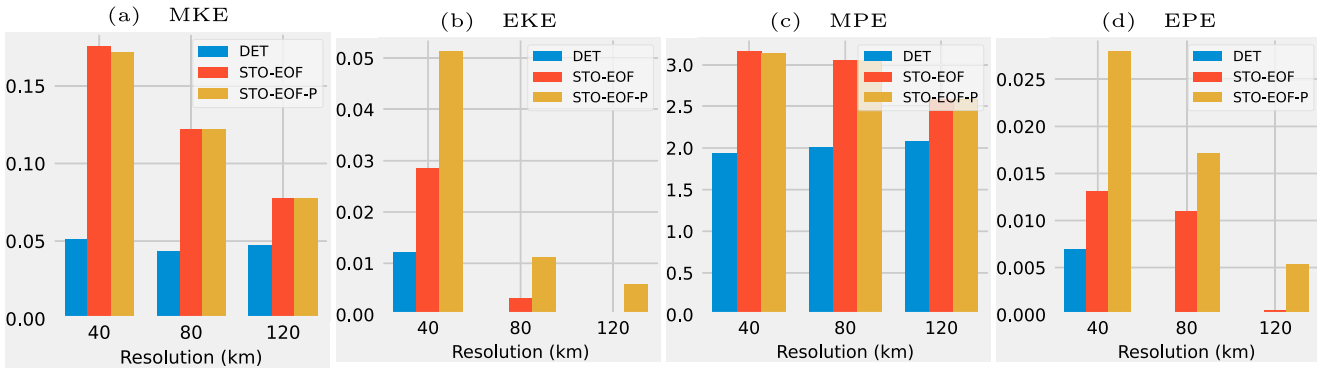


Figure 7. Comparison of (a) mean kinetic energy (MKE), (b) eddy kinetic energy (EKE), (c) mean potential energy (MPE) and (d) eddy potential energy (EPE) (in J/m^3) for different coarse models (by colors and groups in each panel). The corresponding values of the REF model: MKE = 0.23 J/m^3, EKE = 0.62 J/m^3, MPE = 4.01 J/m^3, and EPE = 0.34 J/m^3.

4.3. Including Mixed Layer Dynamics

In this section, we follow further the Q-GCM formulation (Hogg et al., 2004) to embed a mixed layer in the upper QG layer. This brings an additional diabatic forcing into the PV (Equation 3.2b) due to the Ekman pumping of sea surface temperature (SST). The objective is to enhance the small-scale eddies effect and the intrinsic variability

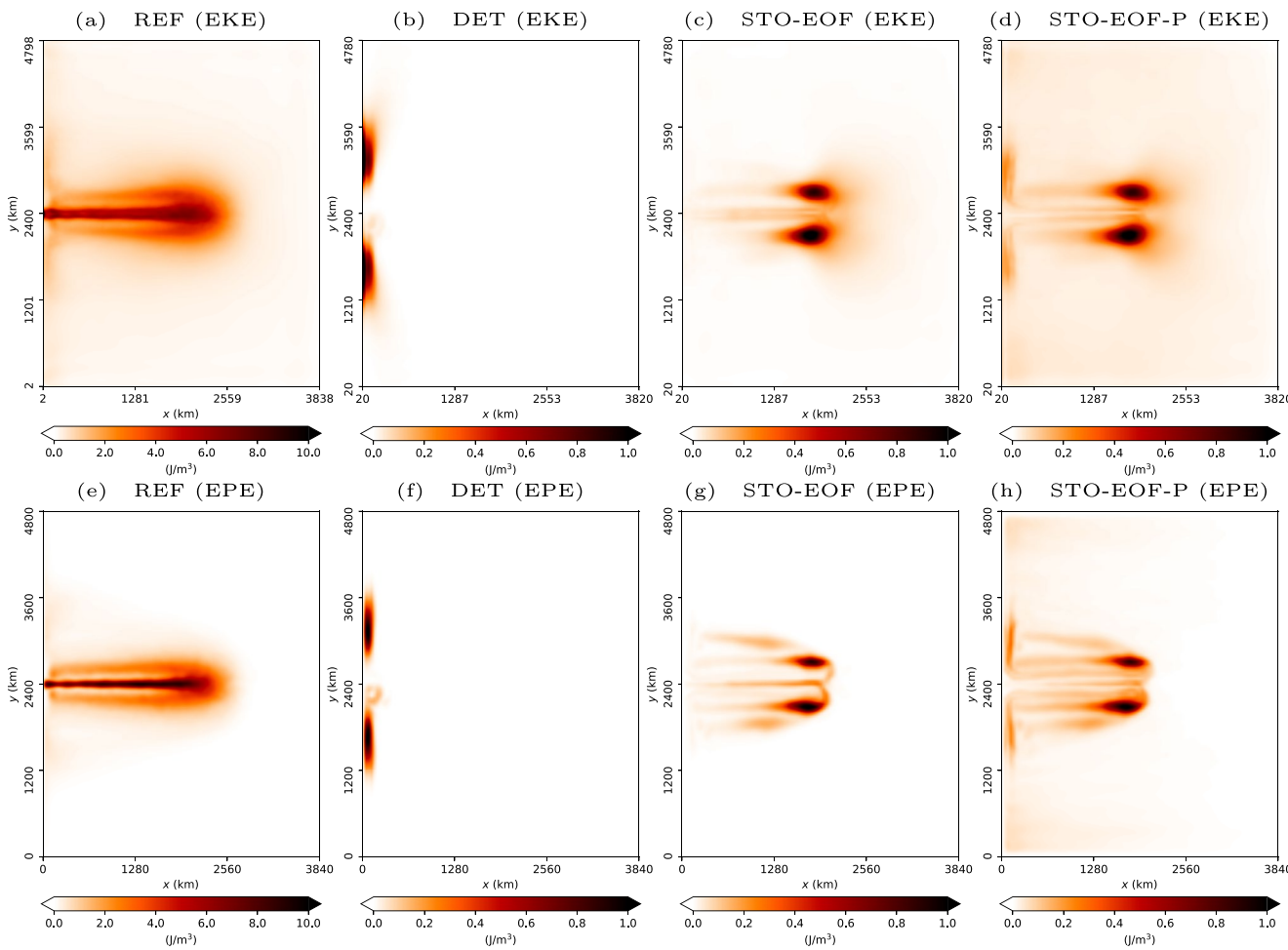


Figure 8. Comparison of the vertically integrated eddy kinetic energy density (top row) and eddy potential energy density (bottom row) for different models (by columns). Both coarse models (DET, STO-EOF, and STO-EOF-P) are at a resolution of 40 km. Note that the colorbar of coarse models is different than that of the REF model.

of the reference model. The resulting EOFs of the noise will contain more realistic data information. In view of the application of data assimilation in future works, it also allows us to incorporate realistic SST observations. Moreover, this constitutes a good basis for the development of a fully ocean-atmosphere coupled stochastic model in subsequent works.

The thickness (H_m) of the mixed layer is fixed here to be 100 m. The vertical velocity across the first ocean interface is $e_1 = -1/2\ (\Delta_m T/\Delta_1 T)\ e_0$, where $\Delta_m T = T_m - T_1$ (resp. $\Delta_1 T = T_1 - T_2$) denotes the temperature difference across the bottom boundary of the mixed layer (resp. of the upper QG layer). Here, the QG potential temperature are fixed to be (287, 282, 276) K.

The additional forcing e_1 evolves in time as the SST does. In the present work, we only considered the geostrophic component of the unresolved noise in the mixed layer. In this case, the stochastic evolution of SST reads

$$\tilde{\mathbb{D}}_t^m T_m + T_m \boldsymbol{\nabla} \cdot \boldsymbol{u}_m \, \mathrm{d}t = \left(-K_4 \nabla^4 T_m + \frac{T_m + T_1}{2H_m} e_0 \right) \mathrm{d}t, \tag{4.2}$$

where $\boldsymbol{u}_m = \nabla^\perp \psi_1 - \boldsymbol{\tau}^\perp/(f_0 H_m)$ is the divergent mixed layer velocity, $\tilde{\mathbb{D}}_t^m$ is the mixed layer stochastic transport operator under probability measure $\tilde{\mathbb{P}}$ (associated to $\boldsymbol{u}_m$) and $K_4 = A_4$ (see values in Table 2). Note that this shallow-water type equation is derived by vertical integration of a 3D heat equation. Equation 4.2 is numerically discretized by the central winding scheme (in space) and the stochastic Leapfroq scheme (in time).

As expected, the inclusion of the additional forcing coming from the SST evolution generates more small-scale eddies almost everywhere in REF (Figures 9a and 9e). The improvements of the two stochastic models observed in the previous test case still hold here (Figures 9c, 9d, 9g, and 9h), namely the zonal jet is reproduced and perturbed. In addition, the variability of the gyres seems also to be enhanced in this case.

Subsequently, an analog procedure to that presented in Section 4.1 is performed for all the coarse-resolution models in this configuration. For instance, the results of the time-mean streamfunctions at 40 km are provided in Figure 10 and a conclusion consistent with the previous test case is recovered.

As illustrated in Figure 11, the STO-EOF-P models, with a noise along the iso-surfaces of vertical stratification, involve an additional SST random forcing that brings a higher KE backscattering than the stationary noise model at both resolutions. This seems to highlight the importance of the non-stationary characteristic of the noise brought by the projection.

As shown in Figure 12, the STO-EOF-P models provide the best results for all the metrics of variability at different resolutions. The local structures of the eddy energy are also better than those recovered in the other coarse models, which is demonstrated by Figure 13. These results confirm the benefits of slaving the noise to the large-scale stratification dynamics.

5. Conclusions and Discussion

A stochastic parameterization of the unresolved eddy field is successfully implemented in a well-known QG model. The empirical spatial correlation of the small-scale noise has been first estimated from eddy-resolving simulation data. An additional correction drift that is fully justified in the stochastic setting considered and ensuing from a change of probability measure has been introduced. The resulting random model enables to reproduce the eastward jet for coarse-grid simulations within the wind-driven double-gyre circulation. In addition, a projection method has been proposed to constrain the noise living along the iso-surfaces of the vertical stratification. The resulting non-stationary noise model enables us to improve the intrinsic variability of the large-scale resolved flow. This improvement has been demonstrated through some statistical diagnostics. We have highlighted that both stationary and non-stationary noise models provide improvements at ocean climatic scale in terms of variability metrics, compared to a corresponding deterministic coarse model. At that resolution, the effects of the mesoscale eddies within the ocean basin at midlatitudes (of deformation radius around 30–40 km) are better represented by the random models, even though the baroclinic instability can not be resolved.

The numerical results presented in this work encourage us to implement the proposed random model on more complex and realistic flow configurations. A following study for a fully coupled model of the ocean and atmosphere (Hogg et al., 2003) is already in progress. Including random forcing into the bottom layer of the atmosphere

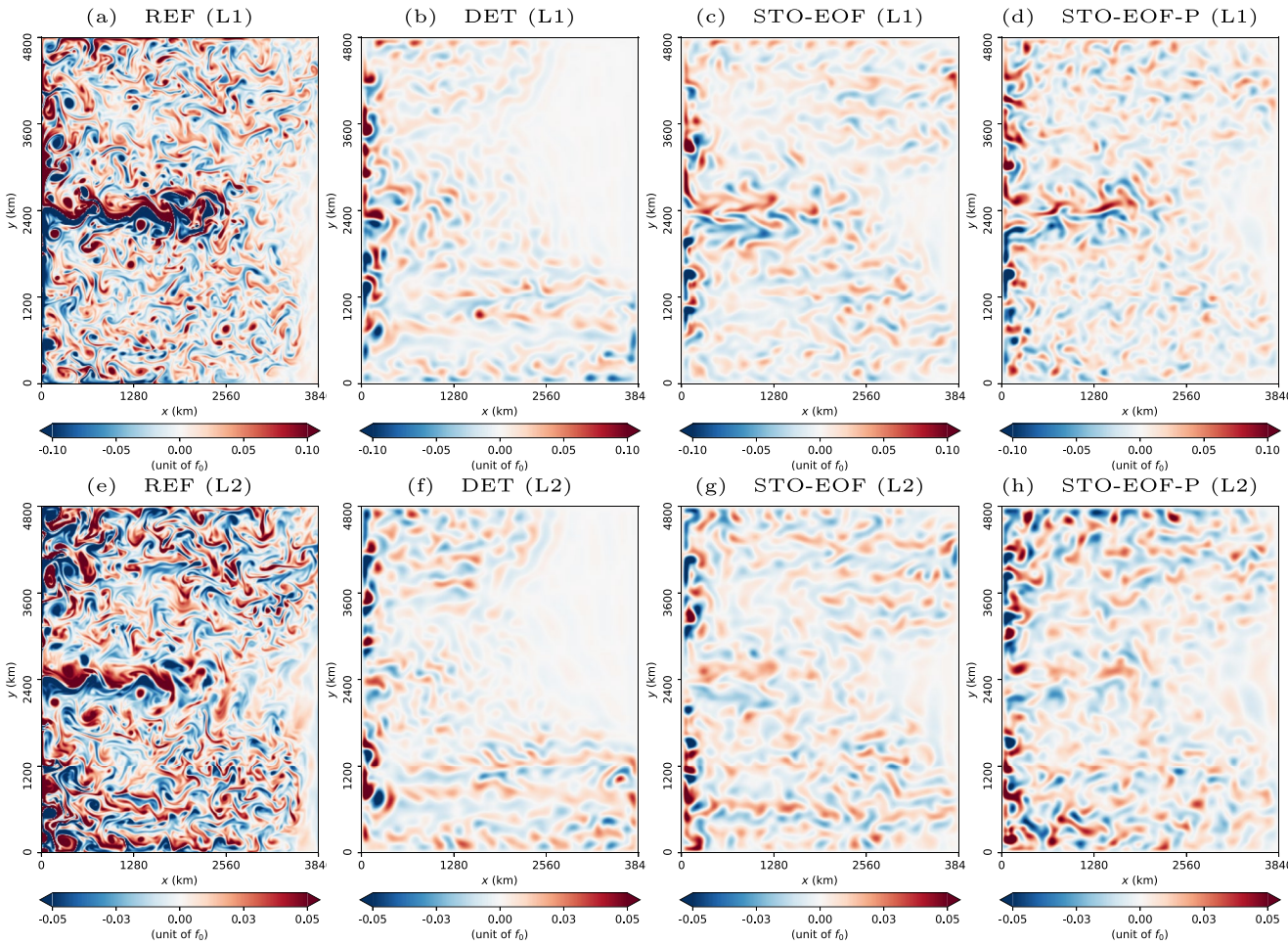


Figure 9. Snapshots of relative vorticity (divided by f_0) for the two upper layers (L1 and L2 by rows) provided by different simulations (by columns) with an embedded mixed layer after 60-years integration. All the coarse models plotted (DET, STO-EOF, and STO-EOF-P) are at a resolution of 40 km.

leads to an unsteady wind stress and to a random diabatic forcing due to Ekman pumping of the temperature in both atmospheric and oceanic mixed layers. Thus, it will be interesting to verify the response of the ocean dynamics in terms of variability and energy to such highly variable forcing, and particularly to compare the results to a stationary wind forced ocean dynamics as well as to an unforced dynamics. The LU formulation for the ocean primitive model has been already derived in Resseguier et al. (2017a) and Li (2021). The numerical assessment of this stochastic formulation in some well established ocean circulation models such as Nucleus for European Modelling of the Ocean (NEMO) is ongoing (Tucciarone et al., 2023). In that case, some additional techniques have to be taken into account. In the stratified interior of the ocean, the noise should be projected on the evolving isopycnal surfaces in order to minimize the possible diapycnal diffusion from $\nabla \cdot (\boldsymbol{a}\nabla\Theta)$. Due to the horizontal and vertical interpolations within the staggered grid, a specific rescaling parameter could be introduced in order to mimic the balance between the quadratic variation of the advection noise and the energy lost by the corresponding diffusion. Moreover, the random pressure $\mathrm{d}p_t^\sigma$ has to be adequately modeled. It could be either constrained by the conservation of the total energy of the resolved flow, or computed by the modified hydrostatic balances based on a proper scaling of the unresolved component.

Another possible research axis focuses on new parameterization methods for the unresolved flow component. Techniques relying on the dynamic mode decomposition (DMD) (Gugole & Franzke, 2020; Schmid, 2010), or on the spectral analysis of the Koopman operator (Mezić, 2013) could be explored in practice to estimate the dynamics of the noise and the Girsanov drift from HR data. The objective will consist in evaluating their ability to characterize the long-terms impacts of some small-scale events. In that spirit, the DMD procedure is adopted in

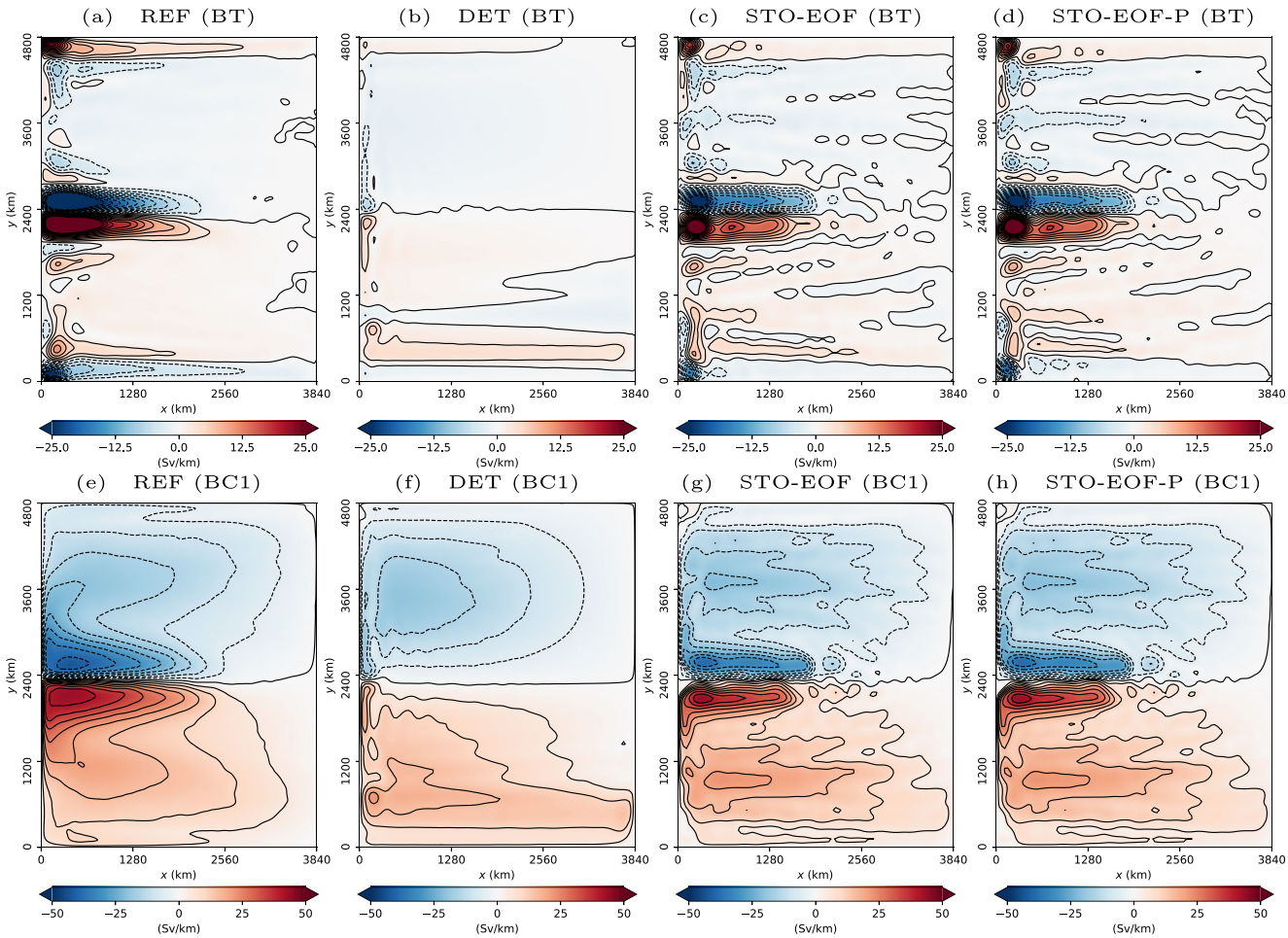


Figure 10. Comparison of time-mean contour of barotropic (BT, top row) and first baroclinic (BC1, bottom row) streamfunctions for different models (by columns) with an embedded mixed layer. All the coarse models plotted (DET, STO-EOF, and STO-EOF-P) are at a resolution of 40 km.

Li et al. (2023) to calibrate the noise structure with time-harmonic property and very promising results have been obtained. More general parameterization methods based on the resolved variables will be investigated as well. A first idea is to estimate the noise covariance from a moving temporal/ensemble fluctuations of the resolved flow in the simulation. In that case, stochastic eddy-resolving simulations could be tested and some physical analysis (Hogg & Blundell, 2006; Hogg et al., 2004) should be performed by comparing the results to that of the deterministic model at the same resolution.

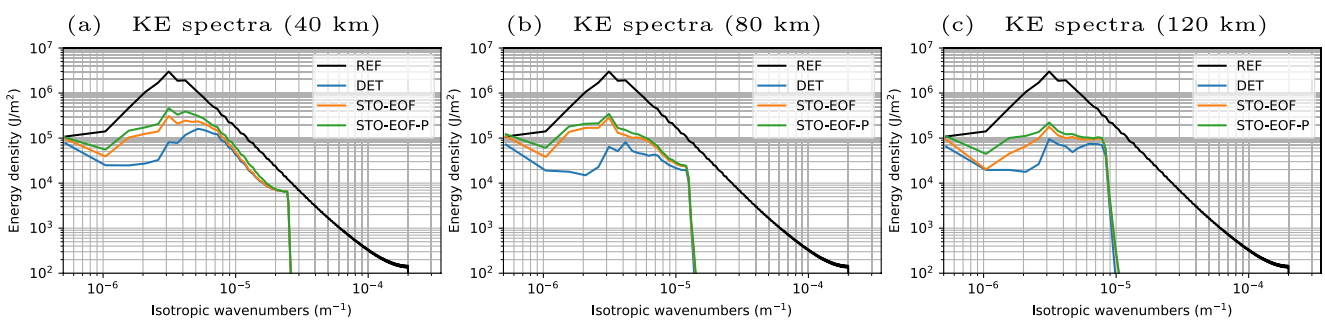


Figure 11. Temporal mean of vertically integrated KE spectra for different models (by colors) with an embedded mixed layer at different resolutions: (a) 40 km, (b) 80 km, and (c) 120 km. The same KE spectra (black line) of the REF (5 km) is added in each panel.

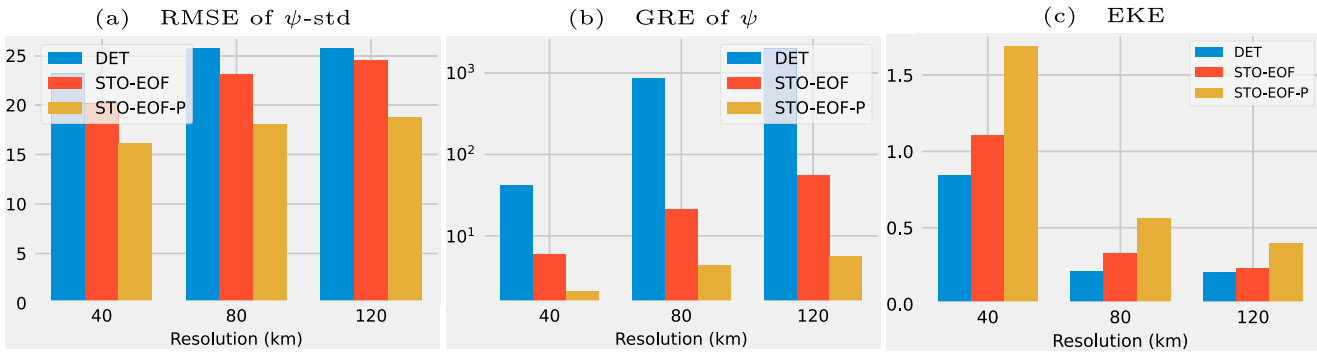


Figure 12. Comparison of variability measures (by sub-figures) for different coarse models (by colors and groups in each sub-figure) with an embedded mixed layer. The eddy kinetic energy of the REF model is 8.26 J/m^3.

We plan also to explore the connection between the LU formulation to other physical parameterizations. For instance, by projecting the noise on the isopycnals of the primitive model, the resulting tracer diffusion $\nabla \cdot (\boldsymbol{a}\nabla\Theta)$ plays exactly the same role as the isoneutral diffusion (Redi, 1982). Moreover, one could reconstruct the Girsanov drift by projecting a meaningful bolus velocity (Gent & McWilliams, 1990) on the noise eigenfunctions (Equation 3.3c). We remark that the proposed advection and diffusion scheme acts on the momentum equation

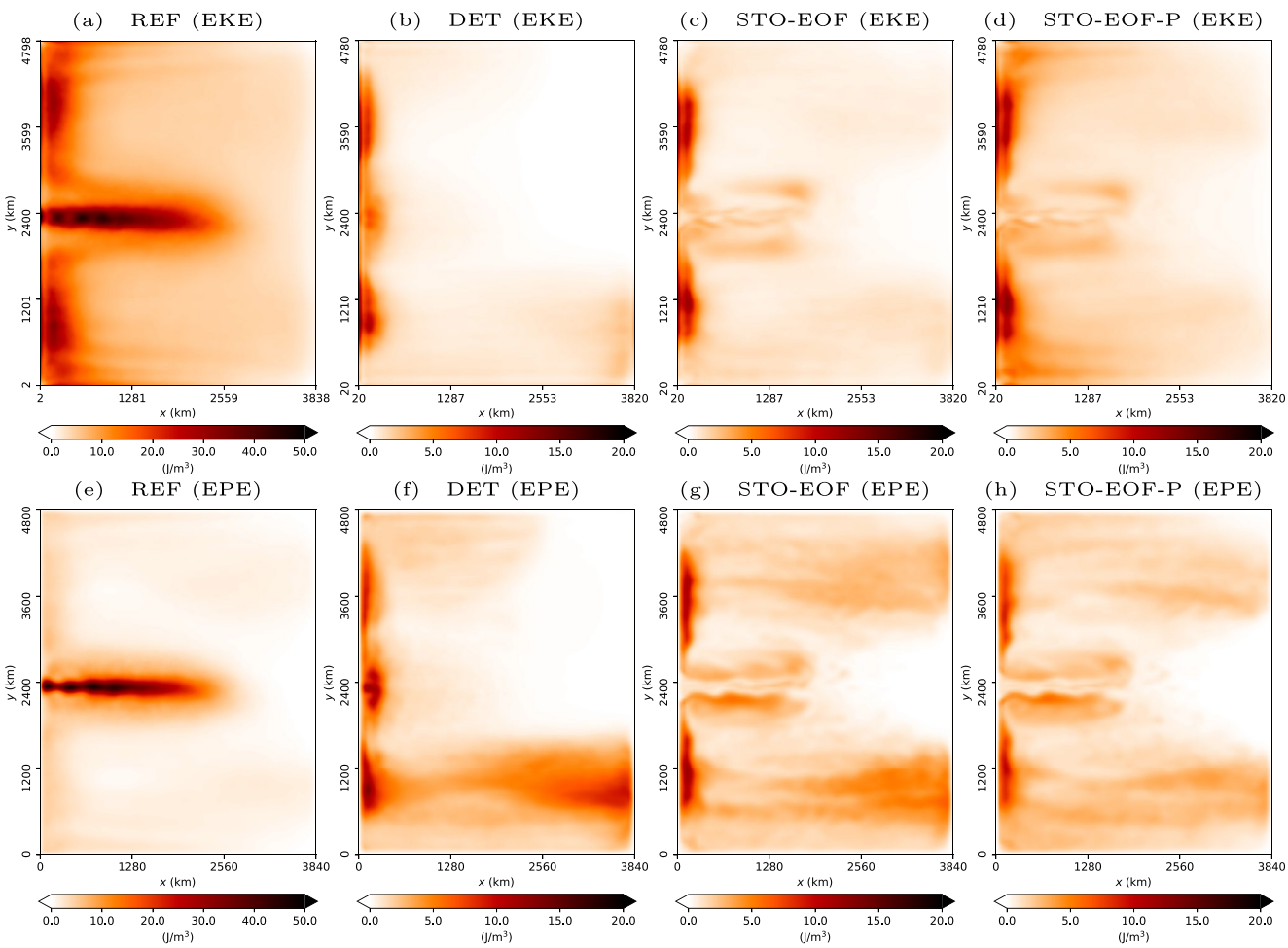


Figure 13. Comparison of the vertically integrated eddy kinetic energy density (top row) and eddy potential energy density (bottom row) for different models (by columns) with an embedded mixed layer. All the coarse models plotted (DET, STO-EOF, and STO-EOF-P) are at a resolution of 40 km. Note that the colorbar of coarse models is different than the one of the REF model.

in addition to the tracer equation, which is hence more general than the traditional methods. We could also investigate a dynamical closure for the noise amplitude $\mathrm{Tr}(\boldsymbol{a})$ by including a prognostic equation of the EKE (Eden & Greatbatch, 2008; Jansen et al., 2019; Mak et al., 2018). It will be also interesting to design specific noises for more efficient energy conversion (Bachman, 2019) based on the analysis presented in this work. The objective is to release the available PE through the noise quadratic variation (Equation 2.7b) and to backscatter the unresolved KE to the resolved KE through the advection of momentum by the noise (Equation 2.7a).

Appendix A: Derivation of Stochastic Baroclinic QG Model

In this section, we describe briefly the derivation for the stochastic baroclinic QG system (Equations 2.7a–2.7e). The derivation of the stochastic geophysical models under the LU framework follows almost exactly the same path as the deterministic derivation (Vallis, 2017). We start from a slightly different stochastic Euler-Boussinesq model than the one proposed by Resseguier et al. (2017a):

$$\underline{\text{Momentum equations}}$$

$$\mathbb{D}_t \boldsymbol{v} + f\,\boldsymbol{k}\times(\boldsymbol{u}\,\mathrm{d}t + \boldsymbol{\sigma}_h \mathrm{d}\boldsymbol{B}_t) = \left(b\,\mathrm{d}t + \mathrm{d}b_t^{\sigma}\right)\boldsymbol{k} - \boldsymbol{\nabla}\left(p\,\mathrm{d}t + \mathrm{d}p_t^{\sigma}\right), \tag{A1a}$$

$$\underline{\text{Continuity equations}}$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{v}^{\star} = 0, \qquad \boldsymbol{\nabla}\cdot\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t = 0, \tag{A1b}$$

$$\underline{\text{Thermodynamic equation}}$$

$$\mathbb{D}_t b + N^2(w^{\star}\,\mathrm{d}t + \sigma_z \mathrm{d}B_t) = \frac{1}{2}\boldsymbol{\nabla}\cdot\left(\boldsymbol{a}_{\cdot z}N^2\right)\mathrm{d}t. \tag{A1c}$$

Here, $\boldsymbol{v} = (\boldsymbol{u}, w)^T$, $\boldsymbol{v}^{\star} = (\boldsymbol{u}^{\star}, w^{\star})^T$, $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t = (\boldsymbol{\sigma}_h\mathrm{d}\boldsymbol{B}_t, \sigma_z\mathrm{d}B_t)^T$, $\boldsymbol{a} = \begin{pmatrix} \boldsymbol{a}_h & \boldsymbol{a}_{hz} \\ \boldsymbol{a}_{hz}^T & a_{zz} \end{pmatrix}$ with the 2D vector fields $\boldsymbol{u}$, $\boldsymbol{u}^{\star}$, $\boldsymbol{\sigma}_h\mathrm{d}\boldsymbol{B}_t$, $\boldsymbol{a}_{hz}$, the scalar fields w, $w^{\star}$, $\sigma_z\mathrm{d}B_t$, a_{zz} and the 2D tensor $\boldsymbol{a}_h$; $f = f_0 + \beta y$ is the Coriolis parameter on beta-plane and $\boldsymbol{k} = (0,0,1)^T$; p and $\mathrm{d}p_t^{\sigma}/\mathrm{d}t$ are the time-smooth component and the fast oscillating noise of the pressure fluctuations rescaled by the background density ρ_0; $N^2(z) = -g\partial_z\overline{\rho}(z)/\rho_0$ is the Brunt-Väisälä frequency with g the gravity value and $\overline{\rho}$ the stationary density that only depends on height; $b(\boldsymbol{x}, t) = -g\rho'(\boldsymbol{x}, t)/\rho_0$ is the resolved buoyancy associated with the density anomaly ρ' and $\mathrm{d}b_t^{\sigma}/\mathrm{d}t$ is a zero-mean noise that models the unresolved buoyancy fluctuations. This thermodynamics noise comes from the assumption of a highly fluctuating uncertainty on the equation of state. It does not directly ensue from the turbulent small-scale velocity fluctuations. The introduction of this uncertainty term has strong consequences on the vertical structure of the noise and constitutes the main difference with the Euler-Boussinesq system in Resseguier et al. (2017a).

We next adimensionalize the previous system. Following Vallis (2017) the basic variables are scaled as:

$$\boldsymbol{x} = L\,\hat{\boldsymbol{x}}, \quad \boldsymbol{u} = U\,\hat{\boldsymbol{u}}, \quad t = \frac{L}{U}\hat{t}, \quad z = H\,\hat{z}, \quad f = f_0\,\hat{f}, \quad N = N_0\,\hat{N}, \tag{A2a}$$

where the capital letters are used for the scales of variables and $\hat{\cdot}$ stand for nondimensional variables. The scales of vertical velocity, dynamical pressure and buoyancy variable are suggested respectively by the classical continuity equation, geostrophic and hydrostatic balances:

$$w = \frac{H}{L}U\,\hat{w}, \quad p = f_0 U L\,\hat{p}, \quad b = \frac{f_0 U L}{H}\hat{b}, \tag{A2b}$$

which implies $\partial_z b/N^2 \sim \mathrm{R_o}/\mathrm{B_u}$ with $R_o = U/(f_0 L)$ the Rossby number, $\mathrm{B_u} = (L_d/L)^2$ the Burger number and $L_d = N_0 H/f_0$ the deformation radius.

Under the LU framework, the noise introduces an additional degree of freedom that must be appropriately accounted for. Following Brecht et al. (2021) and Resseguier et al. (2017b), the horizontal components of the quadratic variation are first scaled as

$$\boldsymbol{a}_h = \epsilon U L\,\hat{\boldsymbol{a}}_h, \quad \epsilon = \frac{T^{\sigma}}{T}\frac{\mathrm{EKE}}{\mathrm{MKE}}, \tag{A2c}$$

where MKE $= U^2$, ϵ is a scaling factor and T^σ denotes the small-scale correlation time. From the definitions (Equations 2.2 and 2.3), the horizontal components of the noise can be scaled as

$$\boldsymbol{\sigma}_h \mathrm{d}\boldsymbol{B}_t = \epsilon^{1/2} L\, \hat{\boldsymbol{\sigma}}_h \mathrm{d}\boldsymbol{B}_{\hat{t}}. \tag{A2d}$$

As a result, the greater this scaling number ϵ, the stronger the uncertainty. As shown in Resseguier et al. (2017c), using different levels of ϵ in the stochastic system enables to model different physical regimes of the large-scale flow. In order to keep the assumption of flat isopycnal (small variation of stratification) in the QG theory, Resseguier et al. (2017b) proposed to scale the ratio between the vertical scale and the horizontal scale of the noise as follows:

$$\frac{\hat{\sigma}_z \mathrm{d}B_{\hat{t}}}{\hat{\boldsymbol{\sigma}}_h \mathrm{d}\boldsymbol{B}_{\hat{t}}} \sim \delta \frac{\mathrm{R_o}}{\mathrm{B_u}}, \tag{A2e}$$

where $\delta = H/L$ denotes the small aspect ratio. From the definition (Equation 2.3), one can deduce the scales for the cross and vertical components of the quadratic variation:

$$\boldsymbol{a}_{hz} = \epsilon \frac{\mathrm{R_o}}{\mathrm{B_u}} U H\, \hat{\boldsymbol{a}}_{hz}, \qquad a_{zz} = \epsilon\delta \left(\frac{\mathrm{R_o}}{\mathrm{B_u}}\right)^2 U H\, \hat{a}_{zz}. \tag{A2f}$$

Moreover, the unresolved pressure and buoyancy noises are scaled as

$$\mathrm{d}p_t^\sigma = \epsilon^{1/2} f_0 L^2\, \mathrm{d}\hat{p}_{\hat{t}}^\sigma, \qquad \mathrm{d}b_t^\sigma = \epsilon^{1/2} \frac{f_0 L^2}{H}\, \mathrm{d}\hat{b}_{\hat{t}}^\sigma. \tag{A2g}$$

Substituting the above scalings in Equations A1a–A1c, we deduce the following nondimensional stochastic Boussinesq system

Momentum equations

$$\mathrm{R_o}\, \hat{\mathbb{D}}_{\hat{t}}^\epsilon \hat{\boldsymbol{u}} + \hat{f}\hat{\boldsymbol{k}} \times \left(\hat{\boldsymbol{u}}\, \mathrm{d}\hat{t} + \epsilon^{1/2}\, \hat{\boldsymbol{\sigma}}_h \mathrm{d}\boldsymbol{B}_{\hat{t}}\right) = -\hat{\boldsymbol{\nabla}}_h \left(\hat{p}\, \mathrm{d}\hat{t} + \epsilon^{1/2}\, \mathrm{d}\hat{p}_{\hat{t}}^\sigma\right), \tag{A3a}$$

Hydrostatic equation

$$\mathrm{R_o}\delta^2\, \hat{D}_{\hat{t}}^\epsilon \hat{w} = \left(\hat{b}\, \mathrm{d}\hat{t} + \epsilon^{1/2}\, \mathrm{d}\hat{b}_{\hat{t}}^\sigma\right) - \partial_{\hat{z}} \left(\hat{p}\, \mathrm{d}\hat{t} + \epsilon^{1/2}\, \mathrm{d}\hat{p}_{\hat{t}}^\sigma\right), \tag{A3b}$$

Continuity equations

$$\begin{aligned} &\hat{\boldsymbol{\nabla}}_h \cdot \left(\hat{\boldsymbol{u}} - \frac{\epsilon}{2}\left(\hat{\boldsymbol{\nabla}}_h \cdot \hat{\boldsymbol{a}}_h + \frac{\mathrm{R_o}}{\mathrm{B_u}}\, \partial_{\hat{z}} \hat{\boldsymbol{a}}_{hz}\right)\right) + \partial_{\hat{z}} \left(\hat{w} - \frac{\epsilon}{2} \frac{\mathrm{R_o}}{\mathrm{B_u}} \left(\hat{\boldsymbol{\nabla}}_h \cdot \hat{\boldsymbol{a}}_{hz} + \frac{\mathrm{R_o}}{\mathrm{B_u}} \partial_{\hat{z}} \hat{a}_{zz}\right)\right) = 0, \\ &\hat{\boldsymbol{\nabla}}_h \cdot \hat{\boldsymbol{\sigma}}_h \mathrm{d}\boldsymbol{B}_{\hat{t}} + \frac{\mathrm{R_o}}{\mathrm{B_u}}\, \partial_{\hat{z}} \hat{\sigma}_z \mathrm{d}B_{\hat{t}} = 0, \end{aligned} \tag{A3c}$$

Thermodynamic equation

$$\mathrm{R_o}\, \hat{\mathbb{D}}_{\hat{t}}^\epsilon \hat{b} + \hat{N}^2 \left(\mathrm{B_u}\, \hat{w}\, \mathrm{d}\hat{t} - \mathrm{R_o}\epsilon \left(\hat{\boldsymbol{\nabla}}_h \cdot \hat{\boldsymbol{a}}_{hz} + \partial_{\hat{z}} \hat{a}_{zz}\right) \mathrm{d}\hat{t} + \mathrm{R_o}\epsilon^{1/2}\, \hat{\sigma}_z \mathrm{d}B_{\hat{t}}\right) = 0, \tag{A3d}$$

where the nondimensional stochastic transport operator $\hat{\mathrm{D}}_{\hat{t}}^\epsilon$ is defined as

$$\begin{aligned} \hat{D}_{\hat{t}}^\epsilon \hat{b} := \mathrm{d}_{\hat{t}} \hat{b} \; &+ \left(\hat{\boldsymbol{u}}\, \mathrm{d}\hat{t} - \frac{\epsilon}{2}\left(\hat{\boldsymbol{\nabla}}_h \cdot \hat{\boldsymbol{a}}_h + \frac{\mathrm{R_o}}{\mathrm{B_u}}\, \partial_{\hat{z}} \hat{\boldsymbol{a}}_{hz}\right) \mathrm{d}\hat{t} + \epsilon^{1/2}\, \hat{\boldsymbol{\sigma}}_h \mathrm{d}\boldsymbol{B}_{\hat{t}}\right) \cdot \hat{\boldsymbol{\nabla}}_h \hat{b} \\ &+ \left(\hat{w}\, \mathrm{d}\hat{t} - \frac{\epsilon}{2} \frac{\mathrm{R_o}}{\mathrm{B_u}} \left(\hat{\boldsymbol{\nabla}}_h \cdot \hat{\boldsymbol{a}}_{hz} + \frac{\mathrm{R_o}}{\mathrm{B_u}}\, \partial_{\hat{z}} \hat{a}_{zz}\right) \mathrm{d}\hat{t} + \epsilon^{1/2} \frac{\mathrm{R_o}}{\mathrm{B_u}}\, \hat{\sigma}_z \mathrm{d}B_{\hat{t}}\right) \partial_{\hat{z}} \hat{b} \\ &- \frac{\epsilon}{2} \left(\hat{\boldsymbol{\nabla}}_h \cdot \left(\hat{\boldsymbol{a}}_h \hat{\boldsymbol{\nabla}}_h \hat{b}\right) + \frac{\mathrm{R_o}}{\mathrm{B_u}} \left(\hat{\boldsymbol{\nabla}}_h \cdot \left(\hat{\boldsymbol{a}}_{hz} \partial_{\hat{z}} \hat{b}\right) + \partial_{\hat{z}} \left(\hat{\boldsymbol{a}}_{hz} \hat{\boldsymbol{\nabla}}_h \hat{b}\right) + \frac{\mathrm{R_o}}{\mathrm{B_u}}\, \partial_{\hat{z}} \left(\hat{a}_{zz} \partial_{\hat{z}} \hat{b}\right)\right)\right) \mathrm{d}\hat{t}. \end{aligned}$$

We now make the following assumptions for the QG approximation: The Rossby number is small, $\mathrm{R_o} \ll 1$; The scale of motion is similar to that of the deformation scale, $\mathrm{B_u} = \mathcal{O}(1)$ or $\mathrm{R_o}/\mathrm{B_u} = \mathcal{O}(\mathrm{R_o})$; The variation of Coriolis

parameter stays small, $\beta L < f_0$; The uncertainty is "moderate," $\epsilon = \mathcal{O}(1)$ or $\epsilon \mathrm{R_o} = \mathcal{O}(\mathrm{R_o})$. We then expand the nondimensional resolved variables in power series of the Rossby number (Vallis, 2017):

$$\left(\widehat{\boldsymbol{u}}, \widehat{w}, \widehat{p}, \widehat{b}\right)^T = \sum_{k\geq 0} \mathrm{R_o}^k \left(\widehat{\boldsymbol{u}}^{(k)}, \widehat{w}^{(k)}, \widehat{p}^{(k)}, \widehat{b}^{(k)}\right)^T, \quad \widehat{f} = \widehat{f_0} + \mathrm{R_o}\,\widehat{\beta}\widehat{y}, \tag{A4}$$

where $\cdot^{(k)}$ denotes the kth order quantity in the Rossby number, $\widehat{f_0} = f_0/f_0 = 1$ and $\widehat{\beta} = \beta L/(\mathrm{R_o} f_0)$. Substituting subsequently these expressions into the nondimensional system (Equations A3a–A3d) and using the above QG assumptions to truncate the equations at zeroth order, we obtain

$$\underline{\text{Geostrophic balances}}$$
$$\widehat{\boldsymbol{u}}^{(0)} = \frac{1}{\widehat{f_0}}\widehat{\boldsymbol{\nabla}}_h^{\perp}\widehat{p}^{(0)}, \quad \widehat{\boldsymbol{\sigma}}_h \mathrm{d}\boldsymbol{B}_{\widehat{t}} = \frac{1}{\widehat{f_0}}\widehat{\boldsymbol{\nabla}}_h^{\perp}\mathrm{d}\widehat{p}_{\widehat{t}}^{\sigma}, \quad \widehat{\boldsymbol{\nabla}}_h \cdot \left(\widehat{\boldsymbol{\nabla}}_h \cdot \widehat{\boldsymbol{a}}_h\right) = 0, \tag{A5a}$$

$$\underline{\text{Hydrostatic balances}}$$
$$\widehat{b}^{(0)} = \partial_{\hat{z}}\widehat{p}^{(0)}, \quad \mathrm{d}\widehat{b}_{\widehat{t}}^{\sigma} = \partial_{\hat{z}}\mathrm{d}\widehat{p}_{\widehat{t}}^{\sigma}, \tag{A5b}$$

where $\widehat{\boldsymbol{\nabla}}_h^{\perp} := (-\partial_{\hat{y}}, \partial_{\hat{x}})^T$. Equations 2.7d and 2.7e are recovered by restoring the dimensions of variables. Note that the geostrophic balances (Equation A5a) together with a zero boundary condition for the vertical noise component at the ocean surface or bottom leads to $\widehat{\sigma}_z \mathrm{d}\boldsymbol{B}_{\widehat{t}} = 0$ and $\widehat{\boldsymbol{a}}_{hz} = \widehat{a}_{zz} = 0$. Subtracting Equations A5a and A5b from Equations A3a–A3d, then truncating the resulting equations at first order using the above assumptions, we have

$$\underline{\text{Momentum equations}}$$
$$\widehat{D}_{\widehat{t}}^{\epsilon(0)}\widehat{\boldsymbol{u}}^{(0)} + \widehat{f}_0\boldsymbol{k}\times\widehat{\boldsymbol{u}}^{(1)}\,\mathrm{d}\widehat{t} + \widehat{\beta}\widehat{y}\boldsymbol{k}\times\left(\widehat{\boldsymbol{u}}^{(0)}\,\mathrm{d}\widehat{t} + \widehat{\boldsymbol{\sigma}}_h \mathrm{d}\boldsymbol{B}_{\widehat{t}}\right) = -\widehat{\boldsymbol{\nabla}}_h\widehat{p}^{(1)}\,\mathrm{d}\widehat{t}, \tag{A5c}$$

$$\underline{\text{Buoyancy equation}}$$
$$\widehat{D}_{\widehat{t}}^{\epsilon(0)}\widehat{b}^{(0)} + \widehat{N}^2\widehat{w}^{(1)}\,\mathrm{d}\widehat{t} = 0, \tag{A5d}$$

$$\underline{\text{Continuity equation}}$$
$$\widehat{\boldsymbol{\nabla}}_h \cdot \widehat{\boldsymbol{u}}^{(1)} + \partial_{\hat{z}}\widehat{w}^{(1)} = 0, \tag{A5e}$$

where $\widehat{\mathbb{D}}_{\widehat{t}}^{\epsilon(0)}$ denotes the nondimensional horizontal stochastic transport operator which is defined as

$$\widehat{\mathbb{D}}_{\widehat{t}}^{\epsilon(0)}\widehat{b} := \mathrm{d}_{\widehat{t}}\widehat{b} + \left(\left(\widehat{\boldsymbol{u}}^{(0)} - \frac{\epsilon}{2}\widehat{\boldsymbol{\nabla}}_h \cdot \widehat{\boldsymbol{a}}_h\right)\mathrm{d}\widehat{t} + \epsilon^{1/2}\,\widehat{\boldsymbol{\sigma}}_h \mathrm{d}\boldsymbol{B}_{\widehat{t}}\right)\cdot\widehat{\boldsymbol{\nabla}}_h\widehat{b} - \frac{\epsilon}{2}\widehat{\boldsymbol{\nabla}}_h \cdot \left(\widehat{\boldsymbol{a}}_h\widehat{\boldsymbol{\nabla}}_h\widehat{b}\right)\mathrm{d}\widehat{t}. \tag{A5f}$$

Restoring the dimensions of variables, the dynamical QG model (Equations 2.7a and 2.7b) can be found. In this work, we only focus on noise satisfying the classical geostrophic equilibrium (Equation A5a). It means that we model the mesoscale flow as a stochastic component given a large-scale flow. However, such random equilibrium is not necessary in a general setting and could be relaxed either by changing the scaling parameter ϵ or by expanding the noise components in power series of Rossby number as well.

Appendix B: Conservation of Path-Wise Energy

In this section, we show that the stochastic Boussinesq QG system (Equations 2.7a–2.7e), without considering the beta effect on the noise, preserves along time the global energy of the resolved geostrophic flow component for each realization. To this end, we first recall the Itô's integration-by-part formula (Kunita, 1997): let $\mathrm{d}_t X = F\,\mathrm{d}t + G\,\mathrm{d}B_t$ and $\mathrm{d}_t Y = F'\,\mathrm{d}t + G'\,\mathrm{d}B_t$ be two S(P)DEs driven by a Brownian motion B satisfying $\mathbb{E}[G\mathrm{d}B_t] = \mathbb{E}\left[G'\mathrm{d}B_t\right] = 0$, then

$$\begin{aligned}\mathrm{d}_t(XY) &= X\mathrm{d}_t Y + Y\mathrm{d}_t X + \mathrm{d}\langle X, Y\rangle_t \\ &= X\mathrm{d}_t Y + Y\mathrm{d}_t X + GG'\,\mathrm{d}t,\end{aligned} \tag{B1}$$

where $\langle X,Y\rangle_t \overset{\mathbb{P}}{=} \lim_{n\to+\infty}\sum_{i=1}^{p_n}\left(X^n_{t_i}-X^n_{t_{i-1}}\right)\left(Y^n_{t_i}-Y^n_{t_{i-1}}\right)$ stands for the joint quadratic variation between the processes X and Y with $0=t^n_0<t^n_1<\cdots<t^n_{p_n}=t$ a partition of the interval $[0,\,t]$, and the limit, if it exists, is defined in the sense of convergence in probability. In particular, we have $\langle B,B\rangle_t = t$.

We now expand Equation 2.7a for the resolved geostrophic velocity $\boldsymbol{u}=(u,v)^T$ and Equation 2.7b for the resolved buoyancy b,

$$\mathrm{d}_t u = -\left(\boldsymbol{u}^\star\,\mathrm{d}t+\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)\cdot\boldsymbol{\nabla}u+\left(\frac{1}{2}\boldsymbol{\nabla}\cdot(\boldsymbol{a}\boldsymbol{\nabla}u)+f_0v_a+\beta yv-\partial_x p_a\right)\mathrm{d}t \tag{B2a}$$

$$\mathrm{d}_t v = -\left(\boldsymbol{u}^\star\,\mathrm{d}t+\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)\cdot\boldsymbol{\nabla}v+\left(\frac{1}{2}\boldsymbol{\nabla}\cdot(\boldsymbol{a}\boldsymbol{\nabla}v)-f_0u_a-\beta yu-\partial_y p_a\right)\mathrm{d}t \tag{B2b}$$

$$\mathrm{d}_t b = -\left(\boldsymbol{u}^\star\,\mathrm{d}t+\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)\cdot\boldsymbol{\nabla}b+\frac{1}{2}\boldsymbol{\nabla}\cdot(\boldsymbol{a}\boldsymbol{\nabla}b)\,\mathrm{d}t-N^2w_a\,\mathrm{d}t, \tag{B2c}$$

together with the geostrophic equilibrium (Equation 2.7d), the hydrostatic balances (Equation 2.7e) and the continuity (Equation 2.7c). We recall that $\boldsymbol{a}\,\mathrm{d}t$ defined in Equation 2.3 is the quadratic variation of the noise $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$. Applying Equation B1 for the density of available PE, $b^2/(2N^2)$, and integrating subsequently over the ocean domain $\mathcal{D}=\mathcal{A}\times[0,-H]$, we have

$$\begin{aligned}\int_{\mathrm{D}}\mathrm{d}_t\left(\frac{b^2}{2N^2}\right)\mathrm{d}\boldsymbol{x} &= \int_{\mathcal{D}}\frac{1}{N^2}\left(b\mathrm{d}_t b+\frac{1}{2}\mathrm{d}\langle b,b\rangle_t\right)\mathrm{d}\boldsymbol{x}\\ &= \int_{\mathcal{D}}\frac{1}{N^2}\left(-b\left(\boldsymbol{u}^\star\,\mathrm{d}t+\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)\cdot\boldsymbol{\nabla}b+\frac{1}{2}b\boldsymbol{\nabla}\cdot(\boldsymbol{a}\boldsymbol{\nabla}b)\,\mathrm{d}t\right.\\ &\quad\left.-N^2w_ab\,\mathrm{d}t+\frac{1}{2}(\boldsymbol{\nabla}b)\cdot\boldsymbol{a}\boldsymbol{\nabla}b\,\mathrm{d}t\right)\mathrm{d}\boldsymbol{x}\\ &= -\int_{\mathcal{D}}\boldsymbol{\nabla}\cdot\left(\left(\boldsymbol{u}^\star\,\mathrm{d}t+\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)\frac{b^2}{2N^2}\right)\mathrm{d}\boldsymbol{x}-\int_{\mathcal{D}}w_ab\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t\\ &= -\int_{\mathcal{D}}w_ab\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t,\end{aligned} \tag{B3}$$

where the third equality results from the geostrophic balances (Equation 2.7d) and the last equality comes from the divergence theorem with ideal boundary conditions, $\boldsymbol{u}\cdot\boldsymbol{n}\,|_{\partial\mathcal{A}}=\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{n}\,|_{\partial\mathcal{A}}=0$ (which leads to $\boldsymbol{u}^\star\cdot\boldsymbol{n}\,|_{\partial\mathcal{A}}=0$ from Equations 2.5 and 2.3). Similarly to the previous calculation, one can show that

$$\int_{\mathcal{D}}\mathrm{d}_t\left(\frac{u^2}{2}\right)\mathrm{d}\boldsymbol{x}=\int_{\mathcal{D}}u(f_0v_a+\beta yv-\partial_x p_a)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t, \tag{B4a}$$

$$\int_{\mathcal{D}}\mathrm{d}_t\left(\frac{v^2}{2}\right)\mathrm{d}\boldsymbol{x}=-\int_{\mathcal{D}}v(f_0u_a+\beta yu+\partial_y p_a)\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t. \tag{B4b}$$

Summing these two equations, we deduce the evolution of KE:

$$\begin{aligned}\int_{\mathcal{D}}\mathrm{d}_t\left(\frac{|\boldsymbol{u}|^2}{2}\right)\mathrm{d}\boldsymbol{x} &= -\int_{\mathcal{D}}\left(\left(\boldsymbol{\nabla}^\perp p\right)\cdot\boldsymbol{u}_a^\perp+\boldsymbol{\nabla}\cdot(\boldsymbol{u}p_a)\right)\mathrm{d}\boldsymbol{x}\,\mathrm{d}t\\ &= -\int_{\mathcal{D}}\left(\boldsymbol{\nabla}^\perp\cdot\left(p\boldsymbol{u}_a^\perp\right)-p\boldsymbol{\nabla}\cdot\boldsymbol{u}_a\right)\mathrm{d}\boldsymbol{x}\,\mathrm{d}t\\ &= -\int_{\mathcal{A}}\int_{-H}^{0}p\partial_z w_a\,\mathrm{d}z\,\mathrm{d}\mathcal{A}\,\mathrm{d}t\\ &= \int_{\mathcal{A}}\left([pw_a]^0_{-H}-\int_{-H}^{0}w_a\partial_z p\,\mathrm{d}z\right)\mathrm{d}\mathcal{A}\,\mathrm{d}t\\ &= \int_{\mathcal{D}}w_ab\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t,\end{aligned} \tag{B5}$$

where the first equality comes from the geostrophic balances (Equation 2.7d), the third equality results from both the Stokes theorem with ideal boundary condition of $\boldsymbol{u}_a$ and the continuity (Equation 2.7c), the fourth equality

derives from the vertical integration-by-parts formula, and the last equality is based on both the vertical boundary conditions, $w_a(z=0) = w_a(z=-H) = 0$ and the hydrostatic balances (Equation 2.7e). Finally, by summing Equations B5 and B3, one can show that the stochastic system (Equations B2a–B2c) conserves the path-wise total energy of the resolved geostrophic flow component, namely

$$\mathrm{d}_t \int_D \frac{1}{2}\left(|\boldsymbol{u}|^2 + \frac{b^2}{N^2}\right)\mathrm{d}\boldsymbol{x} = \int_D \mathrm{d}_t\left(\frac{|\boldsymbol{u}|^2}{2}\right)\mathrm{d}\boldsymbol{x} + \int_D \mathrm{d}_t\left(\frac{b^2}{2N^2}\right)\mathrm{d}\boldsymbol{x} = 0. \tag{B6}$$

Since this conservation property is path-wise, it holds also for their ensemble-mean.

Appendix C: Conversion of Ensemble Energies

In this section, we show briefly the conversions between the ensemble-variance and the energy of the ensemble-mean. To that end, we first decompose each prognostic variable Θ (could be u, v, b or $\boldsymbol{a}$) of Equations B2a–B2c into a mean component $\overline{\Theta} := \mathbb{E}[\Theta]$ and an eddy component $\Theta' := \Theta - \mathbb{E}[\Theta]$. The other time-smooth components $F\,\mathrm{d}t$ apart from the geostrophic transport operator are decomposed in the same way. For instance, $F = -N^2\left(\overline{w_a} + w_a'\right)$ in the buoyancy equation. This way, Equations B2a–B2c have exactly the same form. However, the noise is only an eddy component since $\mathbb{E}[\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t] = 0$. Substituting such decomposition into Equations B2a–B2c and taking subsequently the ensemble-mean of the expanded equation, we obtain

$$\mathrm{d}_t\overline{\Theta} = \left(\overline{F} - \overline{\boldsymbol{u}^\star}\cdot\boldsymbol{\nabla}\overline{\Theta} + \frac{1}{2}\boldsymbol{\nabla}\cdot\left(\overline{\boldsymbol{a}}\boldsymbol{\nabla}\overline{\Theta}\right) - \overline{(\boldsymbol{u}^\star)'\cdot\boldsymbol{\nabla}\Theta'} + \frac{1}{2}\boldsymbol{\nabla}\cdot\left(\overline{\boldsymbol{a}'\boldsymbol{\nabla}\Theta'}\right)\right)\mathrm{d}t, \tag{C1a}$$

where the second and the third terms on the RHS describe the mean-mean interactions, whereas the last two terms represent the effect of eddy-eddy interactions on the ensemble-mean component. Under natural boundary conditions, on can easily show the time evolution of the global energy for $\overline{\Theta}$:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int_D \frac{\overline{\Theta}^2}{2}\mathrm{d}\boldsymbol{x} = &\int_D \overline{F}\,\overline{\Theta}\,\mathrm{d}\boldsymbol{x} + \int_D \overline{(\boldsymbol{u}^\star)'\Theta'}\cdot\boldsymbol{\nabla}\overline{\Theta}\,\mathrm{d}\boldsymbol{x}\\ &- \frac{1}{2}\int_D\left(\boldsymbol{\nabla}\overline{\Theta}\right)\cdot\left(\overline{\boldsymbol{a}}\boldsymbol{\nabla}\overline{\Theta} + \overline{\boldsymbol{a}'\boldsymbol{\nabla}\Theta'}\right)\mathrm{d}\boldsymbol{x}.\end{aligned} \tag{C1b}$$

Applying this result for $\overline{u}$ and $\overline{v}$ (resp. for $\overline{b}$) together with Equation B5 (resp. with Equation B3), one can deduce the evolution of the MKE (resp. MPE), and the conversion terms are summarized in the diagram in Section 2.4.

Subtracting next Equation C1a from Equations B2a–B2c (where Θ could be u, v, b), we deduce the evolution of Θ':

$$\begin{aligned}\mathrm{d}_t\Theta' \;=\; &-\left(\overline{\boldsymbol{u}^\star}\cdot\boldsymbol{\nabla}\Theta' - \frac{1}{2}\boldsymbol{\nabla}\cdot\left(\overline{\boldsymbol{a}}\boldsymbol{\nabla}\Theta'\right)\right)\mathrm{d}t\\ &-\left(\left(\boldsymbol{u}^\star\right)'\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)\cdot\boldsymbol{\nabla}\overline{\Theta} - \frac{1}{2}\boldsymbol{\nabla}\cdot\left(\boldsymbol{a}'\boldsymbol{\nabla}\overline{\Theta}\right)\mathrm{d}t + F'\,\mathrm{d}t\\ &-\left(\left(\boldsymbol{u}^\star\right)'\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)\cdot\boldsymbol{\nabla}\Theta' - \frac{1}{2}\boldsymbol{\nabla}\cdot\left(\boldsymbol{a}'\boldsymbol{\nabla}\Theta'\right)\mathrm{d}t\\ &+\left(\overline{(\boldsymbol{u}^\star)'\cdot\boldsymbol{\nabla}\Theta'} - \frac{1}{2}\boldsymbol{\nabla}\cdot\overline{(\boldsymbol{a}'\boldsymbol{\nabla}\Theta')}\right)\mathrm{d}t,\end{aligned} \tag{C2a}$$

where the first two lines describe the eddy-mean interactions, whereas the last two lines represent the effect of eddy-eddy interactions on the ensemble-eddy component. To evaluate the ensemble-mean of the energy for Θ', the Itô's integration-by-part formula (Equation B1) is required again. In this case, one can first show that the mean of the joint quadratic variation process for Θ' reads

$$\begin{aligned}\overline{\mathrm{d}\langle\Theta',\Theta'\rangle_t} \;=\; &\overline{(\boldsymbol{\nabla}\Theta')\cdot\left(\overline{\boldsymbol{a}}\boldsymbol{\nabla}\Theta'\right)} + \overline{(\boldsymbol{\nabla}\Theta')\cdot\left(\boldsymbol{a}'\boldsymbol{\nabla}\Theta'\right)} + \overline{(\boldsymbol{\nabla}\Theta')\cdot\left(\boldsymbol{a}'\boldsymbol{\nabla}\overline{\Theta}\right)}\\ &+ \left(\boldsymbol{\nabla}\overline{\Theta}\right)\cdot\left(\overline{\boldsymbol{a}'\boldsymbol{\nabla}\Theta'}\right) + \left(\boldsymbol{\nabla}\overline{\Theta}\right)\cdot\left(\overline{\boldsymbol{a}}\boldsymbol{\nabla}\overline{\Theta}\right).\end{aligned} \tag{C2b}$$

Subsequently, we can deduce the evolution of the global energy for Θ', namely

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_D \frac{\overline{(\Theta')^2}}{2}\,\mathrm{d}\boldsymbol{x} = \int_D \overline{F'\Theta'}\,\mathrm{d}\boldsymbol{x} - \int_D \overline{(\boldsymbol{u}^\star)'\Theta'}\cdot\boldsymbol{\nabla}\overline{\Theta}\,\mathrm{d}\boldsymbol{x} + \frac{1}{2}\int_D \boldsymbol{\nabla}\left(\overline{\Theta}\right)\cdot\left(\overline{\boldsymbol{a}}\boldsymbol{\nabla}\overline{\Theta} + \overline{\boldsymbol{a}'\boldsymbol{\nabla}\Theta'}\right)\mathrm{d}\boldsymbol{x}. \tag{C2c}$$

Applying similarly this result for $\overline{u}$ and $\overline{v}$ (resp. for $\overline{b}$), we can deduce the evolution of the EKE (resp. EPE), and the results are shown in the diagram in Section 2.4.

Data Availability Statement

The research data are available at https://doi.org/10.5281/zenodo.6818751 (Li et al., 2022) and the source code can be found in https://github.com/matlong/qgcm_lu.

Acknowledgments
The authors gratefully acknowledge the three anonymous reviewers for their insightful comments and suggestions, which helped to significantly improve the manuscript. We also wish to acknowledge the financial support provided by the ERC EU project 856408-STUOD. Long would like to thank Gilles Tissot for helpful discussions and Charles Deltel for technical support.

References

Adcroft, A., Anderson, W., Balaji, V., Blanton, C., Bushuk, M., Dufour, C. O., et al. (2019). The GFDL global ocean and sea ice model OM4.0: Model description and simulation features. *Journal of Advances in Modeling Earth Systems*, *11*(10), 3167–3211. https://doi.org/10.1029/2019ms001726

Arakawa, A., & Lamb, V. R. (1981). A potential enstrophy and energy conserving scheme for the shallow water equations. *Monthly Weather Review*, *109*(1), 18–36. https://doi.org/10.1175/1520-0493(1981)109<0018:apeaec>2.0.co;2

Arbic, B. K., Polzin, K. L., Scott, R. B., Richman, J. G., & Shriver, J. F. (2013). On eddy viscosity, energy cascades, and the horizonal resolution of gridded stallite altimeter products. *Journal of Physical Oceanography*, *43*(2), 283–300. https://doi.org/10.1175/jpo-d-11-0240.1

Bachman, S. D. (2019). The GM+E closure: A framework for coupling backscatter with the gent and McWilliams parameterization. *Ocean Modelling*, *136*, 85–106. https://doi.org/10.1016/j.ocemod.2019.02.006

Bachman, S. D., Fox-Kemper, B., & Pearson, B. (2017). A scale-aware subgrid model for quasi-geostrophic turbulence. *Journal of Geophysical Research: Oceans*, *122*(2), 1529–1554. https://doi.org/10.1002/2016jc012265

Bauer, W., Chandramouli, P., Chapron, B., Li, L., & Mémin, E. (2020a). Deciphering the role of small-scale inhomogeneity on geophysical flow structuration: A stochastic approach. *Journal of Physical Oceanography*, *50*(4), 983–1003. https://doi.org/10.1175/jpo-d-19-0164.1

Bauer, W., Chandramouli, P., Chapron, B., Li, L., & Mémin, E. (2020b). Stochastic representation of mesoscale eddy effects in coarse-resolution barotropic models. *Ocean Modelling*, *151*(4), 101646. https://doi.org/10.1016/j.ocemod.2020.101646

Berloff, P. (2005). Random-forcing model of the mesoscale oceanic eddies. *Journal of Fluid Mechanics*, *529*, 71–95. https://doi.org/10.1017/s0022112005003393

Berloff, P., Hogg, A. M., & Dewar, W. K. (2007). The turbulent oscillator: A mechanism of low-frequency variability of the wind-driven ocean gyres. *Journal of Physical Oceanography*, *37*(9), 2363–2386. https://doi.org/10.1175/jpo3118.1

Brankart, J. (2013). Impact of uncertainties in the horizontal density gradient upon low resolution global ocean modeling. *Ocean Modelling*, *66*, 64–76. https://doi.org/10.1016/j.ocemod.2013.02.004

Brecht, R., Li, L., Bauer, W., & Mémin, E. (2021). Rotating shallow water flow under location uncertainty with a structure-preserving discretization. *Journal of Advances in Modeling Earth Systems*, *13*(12), e2021MS002492. https://doi.org/10.1029/2021ms002492

Cotter, C., Crisan, D., Holm, D. D., Pan, W., & Shevchenko, I. (2019). Numerically modeling stochastic Lie transport in fluid dynamics. *Multiscale Modeling and Simulation*, *17*(1), 192–232. https://doi.org/10.1137/18m1167929

Cotter, C., Crisan, D., Holm, D. D., Pan, W., & Shevchenko, I. (2020). Modelling uncertainty using stochastic transport noise in a 2-layer quasi-geostrophic model. *Foundations of Data Science*, *2*, 173–205. https://doi.org/10.3934/fods.2020010

Craik, A., & Leibovich, S. (1976). A rational model for Langmuir circulations. *Journal of Fluid Mechanics*, *73*(3), 401–426. https://doi.org/10.1017/s0022112076001420

Da Prato, G., & Zabczyk, J. (2014). *Stochastic equations in infinite dimensions* (2nd ed.). Cambridge University Press.

Dufée, B., Mémin, E., & Crisan, D. (2022). Stochastic parametrization: An alternative to inflation in ensemble Kalman filters. *Quarterly Journal of the Royal Meteorological Society*, *148*(744), 1075–1091. https://doi.org/10.1002/qj.4247

Eden, C., & Greatbatch, R. J. (2008). Towards a mesoscale eddy closure. *Ocean Modelling*, *20*(3), 223–239. https://doi.org/10.1016/j.ocemod.2007.09.002

Ewald, B., & Témam, R. (2005). Numerical analysis of stochastic schemes in geophysics. *SIAM Journal on Numerical Analysis*, *42*(6), 2257–2276. https://doi.org/10.1137/s0036142902418333

Franzke, C. E., O'Kane, T. J., Berner, J., Williams, P. D., & Lucarini, V. (2015). Stochastic climate theory and modeling. *Wiley Interdisciplinary Reviews: Climate Change*, *6*(1), 63–78. https://doi.org/10.1002/wcc.318

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

Gent, P. R., Willebrand, J., McDougall, T. J., & McWilliams, J. C. (1995). Parameterising eddy-induced tracer transports in ocean circulation models. *Journal of Physical Oceanography*, *25*(4), 463–474. https://doi.org/10.1175/1520-0485(1995)025<0463:peitti>2.0.co;2

Griffies, S. M. (1998). The Gent-McWilliams skew flux. *Journal of Physical Oceanography*, *28*(5), 831–841. https://doi.org/10.1175/1520-0485(1998)028<0831:tgmsf>2.0.co;2

Griffies, S. M., & Hallberg, R. W. (2000). Biharmonic friction with a Smagorinsky-like viscosity for use in large-scale eddy-permitting ocean models. *Monthly Weather Review*, *128*(8), 2935–2946. https://doi.org/10.1175/1520-0493(2000)128<2935:bfwasl>2.0.co;2

Grooms, I., & Majda, A. J. (2014). Stochastic superparameterization in quasigeostrophic turbulence. *Journal of Computational Physics*, *271*, 78–98. https://doi.org/10.1016/j.jcp.2013.09.020

Grooms, I., Majda, A. J., & Smith, K. S. (2015). Stochastic superparameterization in a quasigeostrophic model of the Antarctic Circumpolar Current. *Ocean Modelling*, *85*, 1–15. https://doi.org/10.1016/j.ocemod.2014.10.001

Gugole, F., & Franzke, C. E. (2020). Spatial covariance modeling for stochastic subgrid-scale parameterizations using dynamic mode decomposition. *Journal of Advances in Modeling Earth Systems*, *12*(8), e2020MS002115. https://doi.org/10.1029/2020ms002115

Haidvogel, D. B., McWilliams, J. C., & Gent, P. R. (1992). Boundary current separation in a quasigeostrophic, eddy-resolving ocean circulation model. *Journal of Physical Oceanography*, *22*(8), 882–902. https://doi.org/10.1175/1520-0485(1992)022<0882:bcsiaq>2.0.co;2

Hogg, A. M., & Blundell, J. R. (2006). Interdecadal variability of the southern ocean. *Journal of Physical Oceanography*, *36*(8), 1626–1645. https://doi.org/10.1175/jpo2934.1

Hogg, A. M., Dewar, W. K., Killworth, P. D., & Blundell, J. R. (2003). A quasi-geostrophic coupled model (Q-GCM). *Monthly Weather Review*, *131*(10), 2261–2278. https://doi.org/10.1175/1520-0493(2003)131<2261:aqcmq>2.0.co;2

Hogg, A. M., Killworth, P. D., Blundell, J. R., & Dewar, W. K. (2004). Mechanisms of decadal variability of the wind-driven ocean circulation. *Journal of Physical Oceanography*, *35*(4), 512–530. https://doi.org/10.1175/jpo2687.1

Holm, D. D. (2015). Variational principles for stochastic fluid dynamics. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, *471*(2176), 20140963. https://doi.org/10.1098/rspa.2014.0963

Jansen, M. F., Adcroft, A., Khani, S., & Kong, H. (2019). Toward an energetically consistent, resolution aware parameterization of ocean mesoscale eddies. *Journal of Advances in Modeling Earth Systems*, *11*(8), 2844–2860. https://doi.org/10.1029/2019ms001750

Jansen, M. F., & Held, I. M. (2014). Parameterizing subgrid-scale eddy effects using energetically consistent backscatter. *Ocean Modelling*, *80*, 36–48. https://doi.org/10.1016/j.ocemod.2014.06.002

Karspeck, A. R., Yeager, S., Danabasoglu, G., Hoar, T., Collins, N., Raeder, K., et al. (2013). An ensemble adjustment Kalman filter for the CCSM4 ocean component. *Journal of Climate*, *26*(19), 7392–7413. https://doi.org/10.1175/jcli-d-12-00402.1

Kjellsson, J., & Zanna, L. (2017). The impact of horizontal resolution on energy transfers in global ocean models. *Fluids*, *2*(3), 45. https://doi.org/10.3390/fluids2030045

Kunita, H. (1997). *Stochastic flows and stochastic differential equations* (Vol. 24). Cambridge University Press.

Leibovich, S. (1980). On wave-current interaction theories of Langmuir circulations. *Journal of Fluid Mechanics*, *99*(04), 715–724. https://doi.org/10.1017/s0022112080000857

Leith, C. E. (1971). Atmospheric predictability and two-dimensional turbulence. *Journal of the Atmospheric Sciences*, *28*(2), 145–161. https://doi.org/10.1175/1520-0469(1971)028<0145:apatdt>2.0.co;2

Li, L. (2021). Stochastic modeling and numerical simulation of ocean dynamics. (Theses, Université Rennes 1). Retrieved from https://hal.archives-ouvertes.fr/tel-03207741/file/PhD-thesis-Li2021.pdf

Li, L., Deremble, B., Lahaye, N., & Mémin, E. (2022). Stochastic data-driven parameterization of unresolved eddy effects in a baroclinic quasi-geostrophic model [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.6818751

Li, L., Mémin, E., & Tissot, G. (2023). Stochastic parameterization with dynamic mode decomposition. In B. Chapron, D. Crisan, D. Holm, E. Mémin, & A. Radomska (Eds.), *Stochastic transport in upper ocean dynamics* (pp. 179–193). Springer International Publishing.

Mak, J., Maddison, J., Marshall, D., & Munday, D. R. (2018). Implementation of a geometrically informed and energetically constrained mesoscale eddy parameterization in an ocean circulation model. *Journal of Physical Oceanography*, *48*(10), 2363–2382. https://doi.org/10.1175/jpo-d-18-0017.1

Marshall, D. P., Maddison, J. R., & Berloff, P. (2012). A framework for parameterizing eddy potential vorticity fluxes. *Journal of Physical Oceanography*, *42*(4), 539–548. https://doi.org/10.1175/jpo-d-11-048.1

McWilliams, J. C. (1977). A note on a consistent quasigeostrophic model in a multiply connected domain. *Dynamics of Atmospheres and Oceans*, *1*(5), 427–441. https://doi.org/10.1016/0377-0265(77)90002-1

McWilliams, J. C. (2006). *Fundamentals of geophysical fluid dynamics*. Cambridge University Press.

Mémin, E. (2014). Fluid flow dynamics under location uncertainty. *Geophysical & Astrophysical Fluid Dynamics*, *108*(2), 119–146. https://doi.org/10.1080/03091929.2013.836190

Mezić, I. (2013). Analysis of fluid flows via spectral properties of the Koopman operator. *Annual Review of Fluid Mechanics*, *45*(1), 357–378. https://doi.org/10.1146/annurev-fluid-011212-140652

Porta Mana, P., & Zanna, L. (2014). Toward a stochastic parameterization of ocean mesoscale eddies. *Ocean Modelling*, *79*, 1–20. https://doi.org/10.1016/j.ocemod.2014.04.002

Redi, M. H. (1982). Oceanic isopycnal mixing by coordinate rotation. *Journal of Physical Oceanography*, *12*(10), 1154–1158. https://doi.org/10.1175/1520-0485(1982)012<1154:oimbcr>2.0.co;2

Resseguier, V., Li, L., Jouan, G., Derian, P., Mémin, E., & Chapron, B. (2020). New trends in ensemble forecast strategy: Uncertainty quantification for coarse-grid computational fluid dynamics. *Archives of Computational Methods in Engineering*, *28*(1), 1886–1784. https://doi.org/10.1007/s11831-020-09437-x

Resseguier, V., Mémin, E., & Chapron, B. (2017a). Geophysical flows under location uncertainty, part I: Random transport and general models. *Geophysical & Astrophysical Fluid Dynamics*, *111*(3), 149–176. https://doi.org/10.1080/03091929.2017.1310210

Resseguier, V., Mémin, E., & Chapron, B. (2017b). Geophysical flows under location uncertainty, part II: Quasi-geostrophic models and efficient ensemble spreading. *Geophysical & Astrophysical Fluid Dynamics*, *111*(3), 177–208. https://doi.org/10.1080/03091929.2017.1312101

Resseguier, V., Mémin, E., & Chapron, B. (2017c). Geophysical flows under location uncertainty, part III: SQG and frontal dynamics under strong turbulence. *Geophysical & Astrophysical Fluid Dynamics*, *111*(3), 209–227. https://doi.org/10.1080/03091929.2017.1312102

Resseguier, V., Pan, W., & Fox-Kemper, B. (2020). Data-driven versus self-similar parameterizations for stochastic Lie transport and location uncertainty. *Nonlinear Processes in Geophysics*, *27*(2), 209–234. https://doi.org/10.5194/npg-27-209-2020

Schmid, P. (2010). Dynamic mode decomposition of numerical and experimental data. *Journal of Fluid Mechanics*, *656*, 5–28. https://doi.org/10.1017/s0022112010001217

Sirovich, L. (1987). Turbulence and the dynamics of coherent structures, part I: Coherent structures. *Quarterly of Applied Mathematics*, *45*(3), 561–571. https://doi.org/10.1090/qam/910462

Smagorinsky, J. (1963). General circulation experiments with the primitive equations. *Monthly Weather Review*, *91*(3), 99–164. https://doi.org/10.1175/1520-0493(1963)091<0099:gcewtp>2.3.co;2

Tucciarone, F., Li, L., & Mémin, E. (2023). Primitive equations under location uncertainty: Analytical description and model development. In B. Chapron, D. Crisan, D. Holm, E. Mémin, & A. Radomska (Eds.), *Stochastic transport in upper ocean dynamics* (pp. 287–300). Springer International Publishing.

Vallis, G. K. (2017). *Atmospheric and oceanic fluid dynamics: Fundamentals and large-scale circulation* (2nd ed.). Cambridge University Press.

Zanna, L., Porta Mana, P., Anstey, J., David, T., & Bolton, T. (2017). Scale-aware deterministic and stochastic parametrizations of eddy-mean flow interaction. *Ocean Modelling*, *111*, 66–80. https://doi.org/10.1016/j.ocemod.2017.01.004