Check for updates

JAMES | Journal of Advances in Modeling Earth Systems

RESEARCH ARTICLE
10.1029/2024MS004783

Derivation and Numerical Assessment of a Stochastic Large–Scale Hydrostatic Primitive Equations Model

Francesco L. Tucciarone[1], **Long Li**[1], **Etienne Mémin**[1], **and Pranav Chandramouli**[2]

[1]University Rennes, Inria/Irmar, Campus Universitaire de Beaulieu, Rennes, France, [2]Dexter Energy, Amsterdam, The Netherlands

Key Points:

- Introduces a stochastic parametrization framework for the Hydrostatic Primitive Equations within the ocean general circulation model NEMO
- Allows for multiple noise definitions, both data—driven and data—agnostic, which can be combined to target diverse oceanic processes
- Successfully tested on an idealized double—gyre configuration, a well—established setup in oceanographic research

Correspondence to:

F. L. Tucciarone,
francesco.tucciarone@inria.fr

Citation:

Tucciarone, F. L., Li, L., Mémin, E., & Chandramouli, P. (2025). Derivation and numerical assessment of a stochastic large–scale hydrostatic primitive equations model. *Journal of Advances in Modeling Earth Systems*, *17*, e2024MS004783. https://doi.org/10.1029/2024MS004783

Received 21 OCT 2024
Accepted 20 APR 2025

Author Contributions:

Conceptualization: Francesco L. Tucciarone, Long Li, Etienne Mémin
Formal analysis: Francesco L. Tucciarone, Long Li, Etienne Mémin
Funding acquisition: Etienne Mémin
Investigation: Francesco L. Tucciarone
Methodology: Francesco L. Tucciarone, Long Li, Etienne Mémin
Software: Francesco L. Tucciarone, Pranav Chandramouli
Supervision: Etienne Mémin
Validation: Francesco L. Tucciarone, Long Li, Etienne Mémin
Visualization: Francesco L. Tucciarone
Writing – original draft: Francesco L. Tucciarone
Writing – review & editing: Francesco L. Tucciarone, Long Li, Etienne Mémin

© 2025 The Author(s). Journal of Advances in Modeling Earth Systems published by Wiley Periodicals LLC on behalf of American Geophysical Union. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Abstract Planetary flows are shaped by interactions at scales much smaller than the flows themselves, with mesoscale and sub–mesoscale eddies playing key roles in mixing, particle transport and tracer dispersion. To capture these effects, we introduce a stochastic formulation of the primitive equations within the Location Uncertainty (LU) framework. Derived from conservation principles via a stochastic Reynolds transport theorem, this approach decomposes velocity into a smooth–in–time large–scale component and a random–in–time field representing unresolved scales effects. To model the velocity noise term, we develop two data–driven methods based on Proper Orthogonal Decomposition (POD) and Dynamic Mode Decomposition (DMD) and extend this to hybrid approaches combining model– and data–driven constraints. Simulations show that the LU framework enhances gyre flow predictions, improving mixing, jet structure, and tracer transport while revealing the interplay between small– and large–scale dynamics.

Plain Language Summary Ocean currents are influenced by small, swirling water movements known as eddies, which play a crucial role in mixing large–scale flows and transporting particles and substances. Achieving accurate ocean simulations within feasible computational limits requires careful consideration of how these unresolved eddies impact the overall system. In our study, we present an innovative approach grounded in stochastic transport principles to embed uncertainties, addressing the challenge of unresolved dynamics within an operational Ocean Circulation Model. This method provides a physically consistent representation of small–scale eddies that are typically omitted in conventional simulations. We tested this stochastic formulation within a well–established Ocean General Circulation Model (OGCM), employing efficient data–driven techniques and model–based noise representations to account for the unresolved eddy component across resolutions ranging from coarse (35.3 km) to fine (11.9 km). Our results demonstrate that this approach significantly improves the prediction of ocean currents compared to traditional methods, marking a substantial advancement in accurately simulating the complex dynamics of ocean circulation.

1. Introduction

Despite the long history and continuous development of ocean and climate numerical models since the pioneering work of Charney and von Neumann (Charney, 1951), along with incredible advancements in computational facilities, fully capturing the ocean dynamics and its intricate space–time scale interactions under complex forcing remains unattainable and requires relying on approximations or simplified models at low resolution.

Increasing the resolution of flow dynamics comes with a proportional rise in the Reynolds number, R_e, and an exponential increase in the degrees of freedom, scaling as R_e^3 when considering both spatial and temporal modes. For fully developed turbulent fluids such as the ocean—characterized by extremely high Reynolds numbers and a vast range of scales, approximately proportional to $R_e^{9/4}$– the complexity is immense. These scales span phenomena ranging from decadal dynamic recurring events to minute–scale wind gust effect and forcing mechanisms ranging from global radiative processes (10,000 km) down to dissipation scales (1 mm). This vast range illustrates the extraordinary breadth of interactions at play. Given this staggering complexity, achieving a comprehensive representation of the ocean system remains unfeasible in the foreseeable future.

Truncation of ocean models at scales above the Rossby radius of deformation (mesoscale) has been proposed using efficient and innovative parametrization techniques. These approaches aim to organize energy transfers from available potential energy to kinetic energy (Gent & Mcwilliams, 1990). Such parametrizations incorporate redistribution processes, including the introduction of a "bolus" velocity and isoneutral diffusion (Redi, 1982). However, the appropriate scaling between these two processes and the extent to which both mechanisms should

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

apply to all transport equations remain unclear from a theoretical perspective. They often also lack of mechanism triggering energy backscatter. To that end deterministic model identifying missing energy transfer (Bachman, 2019; Bachman et al., 2017) or stochastic techniques triggering mechanism that can be interpreted as inverse random diffusion have since long been proposed in the literature (Leith, 1990; Mason & Thomson, 1992).

Recent studies have highlighted the crucial role of sub–mesoscale dynamics in oceanic flows (Capet et al., 2008; Callies et al., 2015; Lévy et al., 2010; J. C. McWilliams, 2016). These sub–mesoscale eddies, with characteristic scales ranging from $\mathcal{O}(1)$ to $\mathcal{O}(10)$ kilometers, have been observed to sustain the energetic mesoscale eddies, which are approximately an order of magnitude larger ($\mathcal{O}(100)$ km). Moreover, sub–mesoscale eddies play significant roles in tracer mixing in the upper ocean and contribute to the energy cascade (Capet et al., 2008). Resolving sub–mesoscale eddies explicitly in numerical simulations is computationally feasible only under limited scenarios (Hurlburt & Hogan, 2000; Lévy et al., 2010; Siegel et al., 2001), as it requires grid sizes smaller than 1 km (Fox-Kemper et al., 2019). Truncating portions of the sub–mesoscale range introduces new challenges, both numerically and physically (Griffies & Treguier, 2013), and necessitates a precise representation of a broader energy spectrum (Lemarié et al., 2015). As numerical resolution increases, the dynamics of energy transfer between resolved and unresolved scales must be carefully evaluated (Sanderson, 1998; Soufflet et al., 2016; Thuburn et al., 2014), while addressing the issue of spurious dianeutral mixing remains particularly challenging in the presence of mesoscale eddies (Ilıcak et al., 2012).

As discussed earlier, an appealing alternative for modeling eddy contributions is the introduction of stochastic models. Such frameworks have been explored in particular for developing large–scale reduced–order climate models or for weather forecast (Berner et al., 2017; Franzke et al., 2015; Gottwald et al., 2017; Leutbecher et al., 2017), employing techniques such as Mori–Zwanzig projection methods (Lucarini & Chekroun, 2023), homogenization (Majda et al., 1999), quasi–Gaussian closures in the spectral domain (Kitsios et al., 2023; O'Kane et al., 2023), randomization of (deterministic) subgrid parametrization (Buizza et al., 1999; Shutts, 2005) and stochastic forcing (Hasselmann, 1976). These stochastic closure methods have also been extended to ocean models (Bagaeva et al., 2024; Berloff, 2005; Grooms & Majda, 2013; O'Kane et al., 2023; Porta Mana & Zanna, 2014) to represent unresolved processes at large–scale, provide initial perturbation conditions, or introduce random forcing for higher–resolution simulations (Storto & Andriopoulos, 2021).

However, most of these frameworks lack generality, as they do not enable a formal stochastic derivation of geophysical dynamical models directly from physical conservation principles. Instead, randomness is typically added to the model post hoc, rather than being introduced at the foundational level. This ad hoc approach may result in physically inconsistent models, potentially causing uncontrolled variance growth, which can compromise stability and the well–posedness of the underlying deterministic dynamics. To mitigate such issues, it is crucial to adhere closely to deterministic principles while allowing a smooth transition from deterministic to stochastic representations.

In this context, two complementary methodologies have been proposed by Mémin (2014) and Holm (2015), providing rigorous frameworks for defining stochastic large–scale representations of geophysical fluid dynamics, grounded in Newtonian and Hamiltonian frameworks, respectively. Both approaches rely on a stochastic decomposition of the Lagrangian trajectory into a smooth–in–time, large–scale component, and a fast, uncorrelated random displacement noise. For the Navier–Stokes equations and primitive equations, the Location Uncertainty (LU) model has been shown to preserve (in a probabilistic sense) the convergence properties of their deterministic counterparts. Furthermore, both models converge to the deterministic systems as noise vanishes (Debussche et al., 2023, 2025), ensuring that the stochastic system remains consistent with deterministic physics for small noise levels. This property is not ensured in ad hoc stochastic approaches, which can exhibit discrepancies, as demonstrated in the case of the Lorenz system (Chapron et al., 2018).

The versatility and solid theoretical foundation of the LU framework enables the development of large–scale representations with a stochastic component that accounts for subgrid contributions. This approach introduces additional degrees of freedom, which can be effectively leveraged to model specific phenomena, such as large–scale components (Tucciarone et al., 2023, 2024), small–scale turbulence (Chandramouli et al., 2020; Harouna & Mémin, 2017), boundary layer effects (Pinier et al., 2019), convection processes (Jamet et al., 2024), or to construct intermediate models (Bauer, Chandramouli, Li, Mémin, et al., 2020; Chapron et al., 2018; Cintolesi & Mémin, 2020; Harouna & Mémin, 2017; Pinier et al., 2019).

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The LU framework has been successfully applied to various geophysical models, including barotropic and baroclinic quasi–geostrophic model (Bauer et al., 2020a, 2020b; Li, Deremble, et al., 2023), the rotating shallow water model (Brecht et al., 2021), the surface quasi–geostrophic model (Resseguier et al., 2021), the hydrostatic primitive equations (Tucciarone et al., 2023, 2024), and, more recently, the non–hydrostatic Boussinesq equations (Tissot et al., 2024). These applications demonstrate the LU framework's efficacy in structuring large–scale flows (Bauer, Chandramouli, Li, Mémin, et al., 2020), reproducing long–term flow statistics (Bauer, Chandramouli, Li, & Mémin, 2020; Li, Deremble, et al., 2023), and achieving an effective balance between model error representation and ensemble spread (Brecht et al., 2021; Resseguier et al., 2021).

In this study, the LU framework is applied to the primitive equations in an idealized 3D double–gyre configuration. This stochastic ocean model has been integrated into the operational hydrostatic ocean code NEMO (Madec et al., 2019). Various types of noise, ranging from data–driven specifications to noise derived from large–scale resolved variables, are evaluated for both eddy–permitting and eddy–resolving resolutions. Notably, the inclusion of simplified data–driven dynamics for the noise is investigated within this double–gyre configuration, a benchmark model known for its strong sensitivity to grid resolution (Lévy et al., 2010).

This paper is organized as follows: Section 2 provides an overview of the LU framework and introduces the derived stochastic primitive equations. Section 3 details the numerical modeling of uncertainty, including both data–driven and model–driven approaches. Section 4 presents the numerical results for the previously introduced noise models applied to the double–gyre configuration in NEMO, with analyses across eddy–permitting and eddy–resolving regimes. Finally, Section 5 summarizes our conclusions. The Appendices include a concise derivation of the stochastic Boussinesq equations, an outline of the filtering procedures for high–resolution data, and a description of the time–stepping method implemented for the stochastic formulation.

2. Modeling Under Location Uncertainty

The stochastic modeling framework used in this study is briefly summarized below. For a detailed description, we refer to Bauer, Chandramouli, Li, Mémin, et al. (2020), Resseguier et al. (2017b), Tissot et al. (2024), with additional details about the noise properties provided in Appendix A.

In contrast to classical fluid mechanics, where displacement is fully defined by a predictable velocity function, the LU paradigm is built on the idea that displacement is composed of a smooth–in–time predictable term and an highly oscillating (Brownian) uncertainty term. In this case, the fluid flow map $\boldsymbol{X}_t : \mathcal{S} \times \mathbb{R}^+ \to \mathcal{S}$, that is the trajectory followed by fluid particles starting at initial map $\boldsymbol{X}|_{t=0} = \boldsymbol{x}_0$ of the bounded domain $\mathcal{S} \subset \mathbb{R}^3$, is described by

$$\mathrm{d}\boldsymbol{X}_t = \boldsymbol{v}(\boldsymbol{X}_t, t)\,\mathrm{d}t + \boldsymbol{\sigma}_t(\boldsymbol{X}_t)\,\mathrm{d}\boldsymbol{B}_t. \tag{1}$$

The first component, $\boldsymbol{v}\,\mathrm{d}t$ is the smooth–in–time response of the equations of motion and it is supposed to be both spatially and temporally correlated. The second term, denoted as $\boldsymbol{\sigma}_t\mathrm{d}\boldsymbol{B}_t$ is a stochastic process describing the unresolved components, turbulent effects, and uncertainties in the flow. In this noise definition, $\{\boldsymbol{B}_t\}_{0 \le t \le T}$ is a functional (cylindrical) Wiener process. In physical terms, the grounding definition of trajectory is now such that with each displacement the arrival position is no more $\boldsymbol{X}_t = \boldsymbol{X}_{t-\mathrm{d}t} + \boldsymbol{v}\mathrm{d}t$ but rather in its (Gaussian) neighborhood.

Despite being temporally uncorrelated, the fast scales introduced by $\boldsymbol{\sigma}_t\mathrm{d}\boldsymbol{B}_t$ are spatially correlated. The spatial correlation of noise component is formulated through the integral operator $\boldsymbol{\sigma}_t$, which incorporates spatial and temporal dependencies. This operator is defined by a symmetric positive definite kernel $\check{\boldsymbol{\sigma}} = \left(\check{\boldsymbol{\sigma}}_{ij}\right)_{1 \le i,j \le 3}$, which is assumed bounded in both space and time:

$$\left(\boldsymbol{\sigma}_t \boldsymbol{f}\right)(\boldsymbol{x}) = \int_{\mathcal{S}} \check{\boldsymbol{\sigma}}(\boldsymbol{x}, \boldsymbol{y}, t)\, \boldsymbol{f}(\boldsymbol{y})\mathrm{d}\boldsymbol{y}, \quad \boldsymbol{f} \in H, \tag{2}$$

where H denotes a relevant Hilbert space of functions. With these assumptions the correlation operator is a positive definite Hilbert–Schmidt operator. It admits a spectral decomposition with a set of eigenfunctions

$\{\boldsymbol{\varphi}_n\}_{n\in\mathbb{N}}$, rescaled by their corresponding eigenvalues, forming an orthonormal basis, such that the noise can be written as

$$\boldsymbol{\sigma}_t(\boldsymbol{x})\,\mathrm{d}\boldsymbol{B}_t = \sum_{n\in\mathbb{N}} \boldsymbol{\varphi}_n(\boldsymbol{x},t)\,\mathrm{d}\beta_n, \tag{3}$$

where $(\beta_n)_{n\in\mathbb{N}}$ denote independent standard Brownian motions. Associated with the correlation operator $\boldsymbol{\sigma}_t$, we define a tensor $\boldsymbol{a} = (a_{ij})_{1\le i,j\le 3}$:

$$a_{ij}(\boldsymbol{x},t) = \sum_{n\in\mathbb{N}} \varphi_n^i(\boldsymbol{x},t)\,\varphi_n^j(\boldsymbol{x},t), \tag{4}$$

which, in the general case, corresponds to the quadratic variation of the noise. For a deterministic correlation operator, it can be understood as the one–point covariance tensor and for that reason we refer to $\boldsymbol{a}$ as the variance tensor by abuse of language. Notably, this matrix has the units of viscosity in $\mathrm{m}^2\mathrm{s}^{-1}$ and it is a symmetric, positive semi–definite matrix at all spatial points, corresponding to a generalized matrix–valued diffusion tensor. This property has been notably used to design efficient subgrid terms built from velocity fluctuations (Chandramouli et al., 2018; Harouna & Mémin, 2017; Resseguier et al., 2017). The trace of this tensor represents the magnitude of the noise.

As introduced above, the noise term $\boldsymbol{\sigma}_t\mathrm{d}\boldsymbol{B}_t$ is a zero–mean Wiener process. However, this definition can be generalized to non–centered noises (Li, Deremble, et al., 2023; Tucciarone et al., 2024), a concept extensively used in Section 3. The LU framework is here tested in the context of the primitive equations. These equations are generally used for large–scale ocean modeling, and their implementation is usually performed on a fixed grid of relatively coarse resolution (regarding the effective Reynolds number of the flow). Depending on the chosen resolution, sub–grid processes can be significantly correlated and cannot be efficiently represented as a Brownian motion. For this reason, it is appropriate to design a noise that contains a (random) correlated component through non–centered noise. A non–centered noise term shifted by a random process $(\tilde{\boldsymbol{Y}}_t)_{t\in[0,T]}$ can be defined through Girsanov theorem (Li, Deremble, et al., 2023) as:

$$\boldsymbol{\sigma}_t\,\mathrm{d}\boldsymbol{B}_t = \boldsymbol{\sigma}_t\,\mathrm{d}\tilde{\boldsymbol{B}}_t - \boldsymbol{\sigma}_t\,\tilde{\boldsymbol{Y}}_t\,\mathrm{d}t, \tag{5}$$

so that the Lagrangian displacement under LU can be written as

$$\mathrm{d}\boldsymbol{X}_t = \left[\boldsymbol{v}(\boldsymbol{X}_t,t) - \boldsymbol{\sigma}_t\,\tilde{\boldsymbol{Y}}_t(\boldsymbol{X}_t)\right]\mathrm{d}t + \boldsymbol{\sigma}_t\,\mathrm{d}\tilde{\boldsymbol{B}}_t(\boldsymbol{X}_t). \tag{6}$$

The term $\boldsymbol{\sigma}_t\,\tilde{\boldsymbol{Y}}_t(\boldsymbol{X}_t)$ is termed Girsanov drift, and in the following it will be split into a time–independent part $\boldsymbol{u}^\flat$ named *bias* and a time–dependent part, using thus the following notation:

$$\boldsymbol{\sigma}_t\,\tilde{\boldsymbol{Y}}_t = \boldsymbol{u}^\flat + \boldsymbol{\sigma}_t\,\boldsymbol{Y}_t. \tag{7}$$

2.1. Stochastic Reynolds Transport Theorem and Stochastic Transport Operator

The inclusion of a non–differentiable term in the displacement under LU necessitates a modification of the transport theorem, which forms the foundation of fluid mechanics. The Stochastic Reynolds Transport Theorem (SRTT) provides the Eulerian representation of the stochastic Lagrangian flow described in Mémin (2014), Tissot et al. (2024). Considering a physical quantity $\theta(\boldsymbol{x},t)$ within a material volume $V_t \subset \mathbb{R}^3$, transported by a stochastic flow as described by Equation (1) the stochastic transport theorem provides a means to interpret the stochastic balance

$$\mathrm{d}\int_{V_t} \theta\,\mathrm{d}\boldsymbol{x} = \int_{V_t} (\Theta_t\,\mathrm{d}t + \boldsymbol{\Theta}_\sigma\cdot\mathrm{d}\boldsymbol{B}_t)\,\mathrm{d}\boldsymbol{x}, \tag{8}$$

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

where the left–hand side is the temporal evolution of volume–integrated scalar θ, and the right–hand side captures the stochastic forcing, split into its finite variation and martingale components. The (extended) stochastic Reynolds' transport theorem in advective form reads:

$$\int_{V_t} \left\{ \mathbb{D}_t\theta + \theta\nabla\cdot\left[\boldsymbol{v}^\star\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right] + \mathrm{Tr}\left(\boldsymbol{\sigma}_t^{\mathrm{T}}\nabla\boldsymbol{\Theta}_\sigma\right)\mathrm{d}t \right\}\mathrm{d}\boldsymbol{x} = \int_{V_t} \Theta_t\,\mathrm{d}t + \boldsymbol{\Theta}_\sigma\cdot\mathrm{d}\boldsymbol{B}_t\,\mathrm{d}\boldsymbol{x}. \tag{9}$$

Here, the transport operator $\mathbb{D}_t\theta$, playing the role of the material derivative, is given by the expression

$$\mathbb{D}_t\theta = \mathrm{d}_t\theta + \left[\boldsymbol{v}^\star\,\mathrm{d}t + \boldsymbol{\sigma}\,\mathrm{d}\boldsymbol{B}_t\right]\cdot\nabla\theta - \frac{1}{2}\nabla\cdot(\boldsymbol{a}\nabla\theta)\,\mathrm{d}t. \tag{10}$$

The term enclosed in the square brackets represents a stochastic advection displacement, which involves the fast evolving noise term and a modified advection. This modified advection, expressed as

$$\boldsymbol{v}^\star = \boldsymbol{v} - \frac{1}{2}\nabla\cdot\boldsymbol{a} + \boldsymbol{\sigma}^{\mathrm{T}}(\nabla\cdot\boldsymbol{\sigma}), \tag{11}$$

combines a time correlated field, $\boldsymbol{v}^s = \frac{1}{2}\nabla\cdot\boldsymbol{a}$, referred to as the Itô–Stokes drift (Bauer, Chandramouli, Li, Mémin, et al., 2020), and a compressibility term related to the correlation tensor. We note that the stochastic Reynolds transport theorem is analogous to the deterministic Reynolds Transport Theorem (see Section 5.3 of Borisenko et al., 1979), and has an equivalent flux form given by:

$$\int_{V_t} \mathrm{d}_t\theta + \nabla\cdot[\theta(\boldsymbol{v}-\boldsymbol{v}^s)\,\mathrm{d}t + \theta\boldsymbol{\sigma}_t\,\mathrm{d}\boldsymbol{B}_t] - \frac{1}{2}\nabla\cdot(\boldsymbol{a}\nabla\theta)\,\mathrm{d}t + \nabla\cdot(\boldsymbol{\sigma}_t\boldsymbol{\Theta}_\sigma)\,\mathrm{d}t\,\mathrm{d}\boldsymbol{x} = \int_{V_t} \Theta_t\,\mathrm{d}t + \boldsymbol{\Theta}_\sigma\cdot\mathrm{d}\boldsymbol{B}_t\,\mathrm{d}\boldsymbol{x}.$$

All the terms of the SRTT can be physically interpreted. The first term on the right–hand side of (Equation 10) represents the *increment in time* of the process θ at a fixed location. This is expressed as $\mathrm{d}_t\theta = \theta(\boldsymbol{x}, t+\mathrm{d}t) - \theta(\boldsymbol{x}, t)$, and it plays the role of the partial time derivative for θ. However, since θ, is a random process that is not differentiable in time, the notion of a standard derivative is replaced by this stochastic increment.

The second term, made explicit in (Equation 11), consists of an effective advection displacement. It involves the usual large–scale velocity $\boldsymbol{v}$, and modified by a statistical drift related to the noise correlation compressibility and Itô–Stokes drift. As shown in Bauer, Chandramouli, Li, Mémin, et al. (2020), this latter term—which vanishes for homogeneous noise since $\boldsymbol{aa}$ becomes constant—triggers secondary circulations in a manner similar to the Stokes drift, with its associated vortex force and Bernoulli pressure head, as introduced in the Craik–Leibovich theory (Craik & Leibovich, 1976; J. McWilliams et al., 1997). Its importance in a fluid–mechanical context is highlighted in Resseguier et al. (2017), Chandramouli et al. (2018), Pinier et al. (2019).

The third term, $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\nabla\theta$, represents scalar advection by the small–scales. It is noteworthy that this term, involving the product of the solution gradient with a Gaussian random field, is itself not Gaussian. Consequently, while the random fluctuations are Gaussian fields, the advected quantity θ exhibits a non–Gaussian distribution characterized by heavy tails.

The last term corresponds to a diffusion that mixes θ (as the variance tensor is positive definite). It accounts for the mixing of the scalar θ associated with the small–scale random component and serves a role similar to the eddy diffusivity models introduced in classical large–scale representations (Bardina et al., 1980; Gent & Mcwilliams, 1990; Lilly, 1992; Smagorinsky, 1963). However, our approach is more general as it does not rely on the Boussinesq eddy viscosity analogy. Instead, the subgrid term is a general diffusion that arises from the small–scale decorrelation assumption, with directions aligned with the principal axes of the uncertainty variance

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

(Resseguier et al., 2017). From this perspective, it is interesting to note that the Redi isoneutral diffusion tensor (Redi, 1982) can be derived by constraining the noise to remain confined to the isopycnal surfaces (Mémin, 2014).

Finally, the two terms $\mathrm{Tr}\left(\boldsymbol{\sigma}_t^{\mathrm{T}} \nabla \boldsymbol{\Theta}_\sigma\right)$ and $\nabla \cdot (\boldsymbol{\sigma}_t \boldsymbol{\Theta}_\sigma)$ are two equivalent representations of the interaction between the small scales and the forcing.

It is important to emphasize that systems governed solely by stochastic transport operators with non–divergent noise, such as the stochastic primitive quasi–geostrophic model (Li, Deremble, et al., 2023) and the stochastic shallow water model (Brecht et al., 2021), have been shown to conserve total energy for each realization. This conservation arises from the balance between the Itô diffusion term and the noise–advection term, ensuring a direct consistency with the fluctuation–dissipation theorem. This balance is carefully maintained in the proposed numerical scheme (Appendix D). While total energy is globally conserved, local redistribution occurs across scales and energy components (Li, Deremble, et al., 2023). In more complex stochastic models with additional random forcing, energy is not strictly conserved but remains bounded, ensuring mathematical well–posedness (Debussche et al., 2025). In contrast, kinetic energy (KE) backscatter parametrizations (Bachman, 2019; Jansen & Held, 2014), commonly used in coarse–resolution ocean models, mitigate excessive KE dissipation by reinjecting energy into the resolved scales. While serving a different purpose, our stochastic framework provides an alternative by representing subgrid–scale effects through a physically consistent formulation.

Another important remark can be made to conclude this presentation of LU models. The unresolved noise term relies on the assumption of time decorrelation. This corresponds obviously to an idealized simplification of the interactions between large– and small–scale structures in real turbulent flows. However, as described above, this assumption, which enables differential formulations through stochastic calculus, rigorously introduces additional advection and diffusion terms that are correlated in time. These terms includes correlated bias terms for non–centered noise as well as terms generalizing the wave–induced Stokes drift and the Boussinesq eddy viscosity assumption (which with additional assumptions on the noise comes back to Redi diffusion). Notably, the introduction and derivation of a generalized Boussinesq eddy viscosity assumption are rigorously justified through this decorrelation assumption. As for the Itô–Stokes, it aligns with the well established Generalized Lagrangian Mean (GLM) theory (Andrews & Mcintyre, 1978). However, unlike GLM, it is not restricted to wave dynamics but arises more generally when the small–scale component is (statistically) inhomogeneous in space (Bauer, Chandramouli, Li, Mémin, et al., 2020). Additionally, this decorrelation assumption incorporates directly the fluctuation dissipation relationship of statistical physics within this large scale representation. The relaxation of this decorrelation assumption is an ongoing research topic, and first positive results have been recently obtained to include correlated fluctuations, while preserving the advantages of the LU frameworks (Debussche & Mémin, 2024).

2.2. Boussinesq Equations Under Location Uncertainty

The SRTT and the transport operator enable us to derive stochastic representation of geophysical fluid flow dynamics (Bauer, Chandramouli, Li, Mémin, et al., 2020; Mémin, 2014; Resseguier et al., 2017a, 2017b). Their construction closely follows the same approach as in the deterministic case, where the variation in volume concentration is equated to both internal and external processes. Within the LU framework, the stochastic Boussinesq equations can be written as

$$\underline{\text{Horizontal momentum:}}\quad \mathbb{D}_t \boldsymbol{u} + f \boldsymbol{e}_3 \times \left(\boldsymbol{u}\ \mathrm{d}t + \boldsymbol{\sigma}_t \mathrm{d}\boldsymbol{B}_t^{\mathrm{H}}\right) = -\nabla_{\mathrm{H}}\left(p'\ \mathrm{d}t + \mathrm{d}p_t^{\sigma}\right), \tag{12}$$

$$\underline{\text{Vertical momentum:}}\quad \mathbb{D}_t w = -\frac{\partial}{\partial z}\left(p'\ \mathrm{d}t + \mathrm{d}p_t^{\sigma}\right) + b\ \mathrm{d}t, \tag{13}$$

$$\underline{\text{Temperature and salinity:}}\quad \mathbb{D}_t T = \kappa_T \Delta T\ \mathrm{d}t, \tag{14}$$

$$\mathbb{D}_t S = \kappa_S \Delta S\ \mathrm{d}t, \tag{15}$$

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

$$\underline{\text{Incompressibility:}} \qquad \nabla \cdot [\boldsymbol{v} - \boldsymbol{v}^s] = 0, \qquad \nabla \cdot \boldsymbol{\sigma}_t \mathrm{d}\boldsymbol{B}_t = 0, \tag{16}$$

$$\underline{\text{Equation of state:}} \qquad b = b(T, S, z), \tag{17}$$

with the convention $\boldsymbol{v} = (\boldsymbol{u}, w)$, $\nabla_{\mathrm{H}} = (\partial_x, \partial_y)$, and the buoyancy defined as $b = -g\frac{\rho - \rho_0}{\rho_0}$. Metric terms related to Earth's rotation should also conform to the stochastic Frenet–Serret formula $\mathrm{d}\boldsymbol{C} = \boldsymbol{\Omega}\,\mathrm{d}t \times \boldsymbol{C}$, when applied to planetary scale simulations. These equations where derived in Tucciarone et al. (2023) using asymptotic analysis starting from the stochastic Navier–Stokes presented in Mémin (2014). A novel derivation, based on the compressible Navier–Stokes equations (Tissot et al., 2024), is provided in Appendix B. Temperature T and salinity S are active tracers transported by the stochastic flow, influencing the momentum equation through the (deterministic) equation of state. The zero–mean martingale pressure term, $\mathrm{d}p_t^{\sigma}$, termed *stochastic pressure*, is associated to the divergence–free condition of the noise component.

The stochastic primitive equations are derived from the stochastic Boussinesq equations by neglecting the temporal variation and advection of w by the resolved flow, retaining all noise–dependent terms, including noise–induced advection and diffusion. This approach yields a generalized hydrostatic balance:

$$[\boldsymbol{\sigma}_t \mathrm{d}\boldsymbol{B}_t - \boldsymbol{u}^s\,\mathrm{d}t] \cdot \nabla w - \frac{1}{2}\nabla \cdot (\boldsymbol{a}\nabla w)\,\mathrm{d}t = -\frac{\partial p}{\partial z}\,\mathrm{d}t - \frac{\partial \mathrm{d}p_t^{\sigma}}{\partial z} + b, \tag{18}$$

where the pressure and stochastic pressure are related to the vertical component of the diagnosed large–scale velocity as:

$$p'(\boldsymbol{x}) = \int_{\eta_b}^{z} b + \boldsymbol{u}^s \cdot \nabla w + \frac{1}{2}\nabla \cdot (\boldsymbol{a}\nabla w)\,\mathrm{d}\zeta, \tag{19}$$

$$\mathrm{d}p_t^{\sigma}(\boldsymbol{x}) = \int_{\eta_b}^{z} \boldsymbol{\sigma}_t \mathrm{d}\boldsymbol{B}_t \cdot \nabla w\,\mathrm{d}\zeta, \tag{20}$$

with w obtained by integrating the continuity equation:

$$w(x, y, z) = w(\eta_b) - \int_{\eta_b}^{z} \nabla_{\mathrm{H}} \cdot \boldsymbol{u} + \nabla \cdot \boldsymbol{v}^s\,\mathrm{d}\zeta, \tag{21}$$

where η_b denotes the ocean bottom.

In this extended balance, the classical hydrostatic equilibrium is corrected to account for the advection of vertical velocity by the Itô–Stokes drift, the diffusion of vertical velocity, and stochastic compressibility. The stochastic compressibility is reflected in the divergence of the Itô–Stokes drift, which modifies the vertical velocity. This extended balance has been explored in the context of convection modeling (Jamet et al., 2024) and analyzed mathematically in Debussche et al. (2025). It may hold particular relevance for modeling physical processes near the limit of validity of the traditional hydrostatic assumption (Auclair et al., 2011; Klingbeil & Burchard, 2013; Marsaleix et al., 2019). In that context, the stochastic pressure term can be interpreted as a correction to the classical hydrostatic assumption.

The incompressibility condition of the noise (Equation 13) is prescribed via vertical integration of the horizontal divergence of the noise as:

$$\sigma_t \mathrm{dB}_t^z(\boldsymbol{x}, z) = -\int_{\eta_b}^{z} \nabla_{\mathrm{H}} \cdot \boldsymbol{\sigma}_t \mathrm{d}\boldsymbol{B}_t^{\mathrm{H}}(\boldsymbol{x}, \zeta)\,\mathrm{d}\zeta. \tag{22}$$

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

In the following sections, we describe several practical approaches to define the noise term. These methods can be categorized as either data–driven or model–driven. The data–driven approach leverages high–resolution numerical data to inform the noise structure, whereas the model–driven approach assumes a specific form of noise, typically dependent on the resolved variables. Additionally, a hybrid approach combining elements of both types of noise will also be discussed.

3. Numerical Modeling of Uncertainty

As previously seen the random field and the corresponding variance tensor can be written in spectral form as:

$$\boldsymbol{\sigma}_t \, \mathrm{d}\boldsymbol{B}_t(\boldsymbol{x}) = \sum_{n\in\mathbb{N}} \boldsymbol{\varphi}_n(\boldsymbol{x},t) \, \mathrm{d}\beta_t^n, \tag{23}$$

$$\boldsymbol{a}(\boldsymbol{x},t) = \sum_{n\in\mathbb{N}} \boldsymbol{\varphi}_n \boldsymbol{\varphi}_n^{\dagger}(\boldsymbol{x},t), \tag{24}$$

where $\boldsymbol{\varphi}_n$ are the orthogonal eigenfunctions of the correlation tensor, scaled by their respective eigenvalues, β_t^n are independent standard (scalar) Brownian motions and with the upperscript † denoting transposed complex conjugation. Note that if the empirical covariance tensor (instead of its symmetric square root) is employed, as it is usual, the scaling should be in square root of the eigenvalues. In the following, we specify both data–driven and model–driven approaches to calibrate or parameterize the basis functions $\boldsymbol{\varphi}_n$.

3.1. Data–Driven Noise Models

Operationally, a finite set of basis functions $\{\boldsymbol{\varphi}_n, n \in N\}$ is specified with different data–driven approaches, defined from reference off–line data set. These methods are described in details in the following subsections. In this work, the off–line data set comes from a higher resolution simulation, but has to be intended as a proxy for data available from observations. Regardless of the approach chosen to construct the eigenpairs, the off–line data is filtered and re–gridded from the original higher resolution (Δ_{HR}) to the target resolution (Δ_{LR}). The subsampling operator is denoted with $\overline{\boldsymbol{u}}^{\downarrow_{\mathrm{LR}}}$, where LR is the target resolution. This operator always acts from the higher resolution to a single, fixed resolution grid. For this reason, we will omit the notation of the target resolution LR as it will be clear from the context of Sections 4.2 and 4.3. A spatial low–pass filter, denoted $\overline{\boldsymbol{u}}^{\,c_x}$, is defined, with its corresponding high–pass filter constructed as a residual: $\tilde{\boldsymbol{u}}^{c_x}(\boldsymbol{x},t) = [\boldsymbol{u} - \overline{\boldsymbol{u}}^{\,c_x}](\boldsymbol{x},t)$. These spatial filter are built using a Gaussian filter with integer width $c_x = \Delta_{\mathrm{HR}}/2\Delta_{\mathrm{LR}}$. Additionally, a temporal filter, $\overline{\boldsymbol{u}}^{\,b,\,\epsilon}$, is defined to represent the data $\boldsymbol{u}(\cdot,t)$ filtered in time with a 5th order (low–pass) Butterworth windowing function with a cutoff frequency ϵ. The associated temporal high–pass filter is similarly defined as: $\tilde{\boldsymbol{u}}^{\,b,\epsilon}(\boldsymbol{x},t) = \left[\boldsymbol{u} - \overline{\boldsymbol{u}}^{\,b,\epsilon}\right](\boldsymbol{x},t)$. These spatial and temporal low–pass and high–pass filters can be combined with a downsampling operator. When this is the case, they are denoted as:

$$\overline{\boldsymbol{u}}^{\,c_x\downarrow_{\mathrm{LR}}}, \quad \overline{\tilde{\boldsymbol{u}}^{c_x}}^{\,c_x\downarrow_{\mathrm{LR}}}, \quad \overline{\overline{\boldsymbol{u}}^{\,b,\,\epsilon}}^{\,c_x\downarrow_{\mathrm{LR}}}, \quad \overline{\tilde{\boldsymbol{u}}^{\,b,\,\epsilon}}^{\,c_x\downarrow_{\mathrm{LR}}}. \tag{25}$$

Here, the sequence of filtering operations is interpreted as being applied in order from the innermost operator to the outermost. A detailed explanation of these procedures is provided in Appendix C. Finally, the time–averaged data set $\boldsymbol{u}(\boldsymbol{x},t)$ over a period T is denoted $\overline{\boldsymbol{u}}^{\,T}$.

3.1.1. Modeling the *Bias*

In Section 2 the Girsanov drift was split into a time–dependent and a time–independent component. This latter was denoted as $\boldsymbol{u}^{\flat}$ and called *bias*. This term is operationally defined with the difference of the two time averages, at the high resolution and at the target resolution, such as

$$\boldsymbol{u}^{\flat} = \alpha\left[\overline{\overline{\boldsymbol{u}}_{\mathrm{HR}}^{\,T}}^{\,c_x\,\downarrow_{\mathrm{LR}}} - \overline{\boldsymbol{u}}_{\mathrm{LR}}^{\,T}\right]. \tag{26}$$

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The purpose of this term is to steer the large–scale features toward a reference average dynamics, with the parameter α serving to adjust the energy input of the term. By defining the bias in this way, the simulation can be directed where it is most needed—specifically, in regions where the two averages diverge significantly—while disregarding areas of the domain where the difference between the two time averages is minimal.

3.1.2. Off–Line Noise Modeling Through POD

We begin by describing a representation of stationary basis functions using the Karhunen–Loève decomposition, also referred to as proper orthogonal decomposition (POD) in fluid mechanics (Holmes et al., 1996) or empirical orthogonal function (EOF) in climate sciences (Hannachi et al., 2007). In this first case, the finite sets of eigenfunctions $\{\boldsymbol{\phi}_n(\boldsymbol{x}), n \in [1,N]\}$ and of eigenvalues $\{\lambda_n, n \in [1,N]\}$ are computed through proper orthogonal decomposition (POD) (Holmes et al., 1996). This approach provides the following decomposition for the filtered velocity:

$$\overline{\boldsymbol{u}}^{\,c_x\downarrow}(\boldsymbol{x},t) \approx \boldsymbol{\phi}_0(\boldsymbol{x}) + \sum_{n=1}^{N} \boldsymbol{\phi}_n(\boldsymbol{x})\alpha_n(t), \tag{27}$$

where $\boldsymbol{\phi}_0(\boldsymbol{x})$ represents the temporal average of $\overline{\boldsymbol{u}}^{\,c_x\downarrow}$. The eigenfunctions and their corresponding temporal modes $\{\alpha_n, n \in [1,N]\}$ satisfy the following orthogonality and variance properties:

$$\langle \boldsymbol{\phi}_m, \boldsymbol{\phi}_n \rangle = \int_{\Omega} \boldsymbol{\phi}_m^{\mathrm{T}} \boldsymbol{\phi}_n(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = \delta_{mn}, \quad \overline{\alpha_m \alpha_n}^{\,T} = \lambda_m \delta_{mn}. \tag{28}$$

The noise field is defined by rescaling the spatial modes as $\boldsymbol{\varphi}_n = \sqrt{\lambda_n}\boldsymbol{\phi}_n$, and replacing the temporal modes with random modes as

$$\boldsymbol{\sigma}_t \, \mathrm{d}\tilde{\boldsymbol{B}}_t(\boldsymbol{x}) = \sqrt{\tau} \sum_{n=1}^{N} \boldsymbol{\varphi}_n(\boldsymbol{x}) \mathrm{d}\beta_t^n, \tag{29}$$

where τ is a correlation time introduced to ensure consistency with a length scale. It is important to emphasize that in this model, the spatial modes $\boldsymbol{\phi}_n(\boldsymbol{x})$ remain fixed over time, as they are derived under the assumption of statistically stationary velocity fields. Consequently, POD is less effective at capturing recurrent slow oscillations or non–stationary temporal variations present in data sets. Nevertheless, it proves well–suited for describing stationary phenomena, with temporal occurrences governed by their associated energy levels.

A slightly different approach involves using POD to reconstruct a fluctuation component derived from high–pass temporal filtering of the original data (using the aforementioned Butterworth filter). This is expressed as

$$\overline{\tilde{\boldsymbol{u}}^{t,\epsilon}}^{\,c_x\downarrow}(\boldsymbol{x},t) = \overline{\boldsymbol{u}(\boldsymbol{x},t) - \overline{\boldsymbol{u}}^{\,b,\epsilon}(\boldsymbol{x},t)}^{\,c_x\downarrow} \approx \boldsymbol{\phi}_0(\boldsymbol{x}) + \sum_{n=1}^{N} \boldsymbol{\phi}_n(\boldsymbol{x})\alpha_n(t), \tag{30}$$

as opposed to the spatially filtered velocity in (Equation 22). This approach retains the assumption of stationarity.

Another related approach to construct noise involves applying POD across multiple realizations of the target simulation and averaging the resulting modes over the ensemble of realizations. This can be achieved by first high–pass filtering the original data set to isolate the fast dynamics. The data set is then divided into slices of a fixed temporal length (e.g., 1 year), forming an ensemble denoted as $\mathcal{I}$. Subsequently, POD is applied independently to each slice $i \in \mathcal{I}$, as

$$\overline{\tilde{\boldsymbol{u}}_i^{t,\epsilon}}^{\,c_x\downarrow}(\boldsymbol{x},t) \approx \boldsymbol{\phi}_0^i(\boldsymbol{x}) + \sum_{n=1}^{N} \boldsymbol{\phi}_n^i(\boldsymbol{x})\alpha_n^i(t), \quad t \in [(i-1)\cdot 1\mathrm{Y},\ i\cdot 1\mathrm{Y}],\ i \in \mathcal{I}, \tag{31}$$

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

to construct an ensemble of modes $\{\boldsymbol{\phi}_n^i, \lambda_n^i,\ n = 1 \ldots N\}_{i \in \mathcal{I}}$, which are then averaged over the ensemble. This approach aims to describe the fast scales with more regular structures due to the averaging process. Regardless of the method used to construct the POD modes, the associated variance tensor remains a truncated, real version of (Equation 19). The Girsanov drift is chosen to consist solely of the stationary field, as in Equation 21, with the base–mode $\boldsymbol{\phi}_0(\boldsymbol{x})$ always discarded to avoid interference with the bias.

3.1.3. Off–Line Noise Modeling Through DMD

As a second approach, Dynamical Mode Decomposition is considered (Kutz et al., 2016). DMD is a methodology used to construct a proxy linear dynamical system that approximates an unknown non–linear dynamics. In this study, DMD is applied to the time evolution of the filtered velocity, which is approximated as:

$$\boldsymbol{u}(\boldsymbol{x}, t_{i+1}) \approx A\boldsymbol{u}(\boldsymbol{x}, t_i). \tag{32}$$

Assuming that A is diagonalizable, this finite dimensional linear dynamical system has the general solution:

$$\boldsymbol{u}(\boldsymbol{x}, t) \approx \sum_{m=1}^{N} b_m \exp(\mu_m t)\boldsymbol{\psi}_m(\boldsymbol{x}), \tag{33}$$

where $\boldsymbol{\psi}_m(\boldsymbol{x}) \in \mathbb{C}^3$ are the eigenvectors of A corresponding to the eigenvalues $\mu_m \in \mathbb{C}$ and $b_m \in \mathbb{C}$ are amplitudes. Specifically, as $\mu_m = \sigma_m + \mathrm{i}\omega_m$, the real part σ_m is the growth rate and ω_m is the frequency of the mode m z fields, the eigenvectors, eigenvalues and amplitudes occur in complex conjugate pairs: $\boldsymbol{\psi}_{2p} = \boldsymbol{\psi}^*_{2p+1}$. Following the approach proposed by Li, Mémin, and Tissot (2023), we split the DMD modes into slowly evolving and rapidly evolving modes to define the Girsanov drift using the slow component of the dynamics, and the random noise using the fast component. Specifically, by introducing a temporal separation scale, τ_c, two sets of modes can be defined: $\mathcal{M}^n$ for the noise and $\mathcal{M}^d$ for the drift part, where.

$$\mathcal{M}^n = \left\{\boldsymbol{\psi}_m^u,\ m \in [1, N^u]\ :\ |\mu_m^u| \approx 1,\ |\omega_m^u| > \frac{\pi}{\tau_c},\ |b_m^u| \geq C^u\right\}, \tag{34}$$

$$\mathcal{M}^d = \left\{\boldsymbol{\psi}_m^c,\ m \in [1, N^c]\ :\ |\mu_m^c| \approx 1,\ |\omega_m^c| \leq \frac{\pi}{\tau_c},\ |b_m^c| \geq C^c\right\}, \tag{35}$$

with C^u, C^c denoting two empirical cut–off of amplitudes. A visual representation of the aforementioned procedure is given in Figure 1. The procedure is applied to two complementary data sets, each corresponding to a different set of modes. Specifically, DMD is performed on the following two sets of data:

$$\overline{\boldsymbol{u}}^{\,c_x\downarrow}(\boldsymbol{x}, t) \approx \sum_{m=1}^{N^c} b_m^c \exp(\mu_m^c t)\boldsymbol{\psi}_m^c(\boldsymbol{x}), \tag{36}$$

$$\overline{\widetilde{\boldsymbol{u}}^{c_x}}^{\,c_x\downarrow}(\boldsymbol{x}, t) \approx \sum_{m=1}^{N^u} b_m^u \exp(\mu_m^u t)\boldsymbol{\psi}_m^u(\boldsymbol{x}), \tag{37}$$

where $\overline{\widetilde{\boldsymbol{u}}^{c_x}}^{\,c_x\downarrow}$ stands for the high–passed velocity field based on the Gaussian field of width c_x. This procedure aims at splitting the large scale structures from the small scale structures in the data set. Temporal separation is then achieved through DMD, as explained earlier. This results in two sets of basis functions: One for the large–in–space and slow–in–time dynamics, and the other for the small–in–space and fast–in–time scales. As the DMD modes are not orthogonal, an $H-$optimal projection is employed, as described in Li, Mémin, and Tissot (2023). This procedure, which minimizes the reconstruction error of the truncated basis, is applied separately for $\mathcal{M} = \mathcal{M}^n$ and $\mathcal{M} = \mathcal{M}^d$. Next, by scaling the DMD modes with their respective amplitudes, $\boldsymbol{\varphi}_m = b_m\boldsymbol{\psi}_m$, the noise and Girsanov drift are defined as:

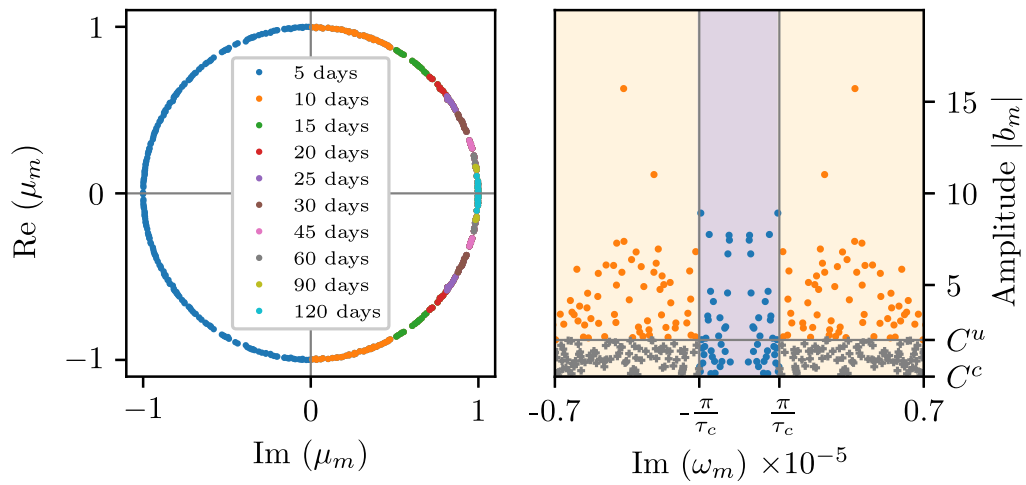


Figure 1. Illustration of the selection of the Dynamic Mode Decomposition modes. On the left, frequencies of the modes are plotted on the unitary circle; they are colored differently to represent their characteristic physical time scales. At this point, a threshold $\tau_c = 45$ days is chosen to differentiate the slowly evolving from the rapidly evolving modes. On the right, over the violet background the slowly evolving modes are plotted, while the rapidly evolving modes are plotted over the orange background. The amplitude threshold for the slowly evolving mode C^c is set to zero, while for rapidly evolving modes C^u is set to 2. The gray dots represent the set of rapidly evolving modes below this threshold, that are thus discarded.

$$\boldsymbol{\sigma}_t \, \mathrm{d}\tilde{\boldsymbol{B}}_t \, (\boldsymbol{x}) = \sqrt{\tau} \sum_{m \in \mathcal{M}^n} \exp\left(\mathrm{i}\omega_m t\right) \boldsymbol{\varphi}_m(\boldsymbol{x}) \mathrm{d}\beta_m, \tag{38}$$

$$\boldsymbol{\sigma}_t \, \boldsymbol{Y}_t \, (\boldsymbol{x}) = \boldsymbol{u}^\flat + \sum_{m \in \mathcal{M}^d} \exp\left(\mathrm{i}\omega_m t\right) \boldsymbol{\varphi}_m(\boldsymbol{x}). \tag{39}$$

Note that the Girsanov drift is defined with the bias term introduced in Equation 21. For this reason, the DMD procedure is carried out on each field after removing the temporal mean. Each pair of complex Brownian motions are conjugates, with their real and imaginary parts being independent. Therefore, both the noise and Girsanov drift are real–valued fields. The quadratic variation of this noise remains stationary:

$$\boldsymbol{a}(\boldsymbol{x}) = \tau \sum_{m \in \mathcal{M}^n} \boldsymbol{\varphi}_m(\boldsymbol{x}) \boldsymbol{\varphi}_m^\dagger(\boldsymbol{x}). \tag{40}$$

However, it is important to highlight that the eigenvalues correspond to oscillating modes over time. In contrast to POD procedures, this type of noise introduce slow oscillating energy modes along the time axis.

3.1.4. Isopycnal Projection

After constructing the noise through Equations 24 and 33 with the off–line data, the small–scale process $\boldsymbol{\sigma}_t \mathrm{d}\boldsymbol{B}_t$ is constrained to live on the tangent space of the isopycnal surfaces. This is achieved by applying the isopycnal projection operator P^ρ, given by

$$\mathrm{P}^\rho = I - \frac{\nabla\rho(\nabla\rho)^{\mathrm{T}}}{|\nabla\rho|^2}, \tag{41}$$

to the noise. As the density function depends on the temperature and salinity, $\rho = \rho(T, S, z)$, the isopycnal projection operator carries information about the current state of the simulation. The projected noise $\boldsymbol{\sigma}_t \mathrm{d}\boldsymbol{B}_t^\rho(\boldsymbol{x}) = \mathrm{P}^\rho \, \boldsymbol{\sigma}_t \mathrm{d}\boldsymbol{B}_t(\boldsymbol{x})$ is strongly tied to the evolution of the flow density, making it intrinsically nonstationary.

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Furthermore, since T and S are advected by the stochastic flow (Equation 1), they themselves become stochastic processes, and consequently, the projector P^{ρ} is also stochastic. A deterministic field projected along the isopycnal surfaces thus becomes a stochastic process. This point is crucial when discussing the Girsanov drift, as it becomes random once projected through P^{ρ}.

3.2. Model–Driven Noise Models

Data–driven noise methodologies are primarily designed to introduce unresolved processes into the simulation. These methods are particularly useful in low–resolution, non–eddy resolving simulations. However, when operating at eddy–permitting resolutions, data–driven approaches may become less effective, as the simulation is able to partially resolve the dynamics on its own. In such cases, a different class of noise models is necessary. These models should not rely on offline data but instead utilize the model–resolved fields.

3.2.1. Vertical Profile Prescription Noise

One simple approach to model–based noise design involves defining the horizontal noise components in (Equation 2) using simple Gaussian radial basis functions with a given standard deviation d:

$$\boldsymbol{\sigma}_t \, \mathrm{d}\boldsymbol{B}_t^{\mathrm{H}} \, (\boldsymbol{x}, z) = c(\boldsymbol{x}, z) \, (\mathcal{G} * \mathrm{d}\boldsymbol{B}_t)(\boldsymbol{x}), \tag{42}$$

$$(\mathcal{G} * \mathrm{d}B_t)^i(\boldsymbol{x}) := \frac{1}{\sqrt{2\pi} d} \sum_k \exp\left[-\frac{|\boldsymbol{x} - \boldsymbol{y}_k|^2}{2d^2}\right] \mathrm{d}\beta_t^{k,i}. \tag{43}$$

In the above equations, $\boldsymbol{x}$ and $\boldsymbol{y}_k$ represent horizontal coordinates, with $\boldsymbol{y}_k$ denoting the collocation points of the Gaussian functions within their support. The horizontal noise component can thus be interpreted as the discrete convolution of a set of Brownian motions with Gaussian blur functions. The function $c(\boldsymbol{x}, z)$ applies a vertical profile to the noise, linking it to the current state of the flow at each simulation timestep. Specifically, the vertical profile is derived by computing the normalized vertical kinetic energy at time n:

$$c(\boldsymbol{x}, z) = \frac{1}{\|c(\boldsymbol{x}, z)\|} \frac{1}{\eta - \eta_b} \int_{\eta_b}^{z} \left[\frac{1}{|\nu(\boldsymbol{x})|} \int_{\nu(\boldsymbol{x})} \mathbb{1}_{\nu} w^{(n)}(\boldsymbol{\xi}, t) \, \mathrm{d}\boldsymbol{\xi} \right]^2 \mathrm{d}\zeta, \tag{44}$$

where the local spatial averaging within a neighborhood $\nu(\boldsymbol{x})$ around $\boldsymbol{x}$ helps stabilize the method by smoothing small fluctuations. We will refer to this method as Vertical Profile Prescription (VPP).

3.2.2. Combining Different Noise Models

In this study, we focus on idealized configurations that serve as proxies for realistic systems, such as the Gulf Stream. However, our future work will address more complex configurations involving real–world basins and forcing conditions. In such cases, a variety of physical processes—such as seasonal convection events (Jamet et al., 2024), deep–water currents, internal waves interactions and others—will require parametrization. These processes may involve distinct noise characteristics, and consequently, there will be a need to combine different noise models to represent their effects.

As an example, we consider a specific noise specified from the geostrophic velocity $\boldsymbol{u}_{\eta}$, which is reconstructed from the sea surface elevation η as

$$u_{\eta} = -\frac{g}{f} \frac{\partial \eta}{\partial y}, \qquad v_{\eta} = \frac{g}{f} \frac{\partial \eta}{\partial x}. \tag{45}$$

DMD is then applied to this temporal field to capture the large–scale, slow, and correlated dynamics of η, with a Girsanov drift defined accordingly. The small–scale, fast, and uncorrelated dynamics are modeled through a vertical profile prescription (VPP) noise. This leads to a noise ansatz of the following form:

$$\boldsymbol{\sigma}_t \mathrm{d}\boldsymbol{B}_t = -\left[\boldsymbol{u}_\eta^\flat + \sum_{m\in\mathcal{M}^d} \exp\left(\mathrm{i}\omega_m^\eta t\right)\boldsymbol{\varphi}_m^\eta(\boldsymbol{x})\right]\mathrm{d}t + c(\boldsymbol{x},z)\,\mathcal{G} * \mathrm{d}\boldsymbol{B}_t(\boldsymbol{x}), \tag{46}$$

where $\boldsymbol{u}_\eta^\flat$ corresponds to the bias as defined by Equation 21 but computed with geostrophic velocities as in Equation 40. The data η considered for this procedure is taken from high–resolution reference data. To minimize errors due to interpolation of η onto the velocity grid, the geostrophic velocity is computed on the high–resolution grid, then filtered and downsampled for DMD application. In real–world applications, sea surface height (SSH) data is commonly used in data assimilation procedures, owing to its availability at high resolution from modern satellite missions. Hence, this noise model represents an optimal candidate for future implementations, especially in configurations with access to real observations, such as those provided by the SWOT satellite.

4. Numerical Experiments in NEMO

The experiments discussed are performed using a state–of–the–art modeling framework for oceanography and climatology, NEMO, that stands for "Nucleus for European Modeling of the Ocean" (Madec et al., 2019). The "blue ocean" engine of NEMO solves the primitive equations of ocean thermodynamics on a curvilinear orthogonal Arakawa C–grid with full or partial step s–coordinate, allowing the usage of z– or σ– coordinates. The prognostic variables are the horizontal velocity field, a linear or non–linear sea surface height, the conservative temperature and the absolute salinity. The non–linear sea surface height is discretized in time with a three–step Generalized Forward–Backward algorithm based on the third order Adams–Bashford and fourth order Adams–Moulton schemes, as proposed by Shchepetkin and McWilliams (2009). In the following, we first provide a brief description of the double–gyre configuration in NEMO, followed by a discussion of the numerical results obtained in both eddy–permitting and eddy–resolving regimes. The focus is on evaluating the performance of the proposed stochastic formulation under various noise models.

4.1. Double–Gyre Configuration

The domain configuration is a rectangular basin on the β–plane centered at $\sim$30°N, rotated by 45° degrees, 3180 km long, 2120 km wide and 4 km deep, bounded by vertical walls and a flat bottom. Seasonal winds (surface wind stress τ) and buoyancy changes (freshwater flux F, restoring toward an apparent air temperature $T^\star$ and penetrative heat Q) are imposed as external forcings to induce the creation of a strong jet that separates a cold sub–polar gyre from a warm sub–tropical gyre. This jet starts at the westernmost point of the domain and moves toward the interior of the domain. A recirculation gyre is also visible in the southern corner. The complete details of this configuration are given in Lévy et al. (2010, 2012), while Figure 2 shows the geographical location of the basin, the applied forcings and provides a first glimpse of the eastwards jet in terms of sea surface height.

The vertical domain, for all experiments, is discretized into 31 levels of increasing thickness with depth. Near the ocean surface, the thickness is $\sim$10 m and it increases to $\sim$500 m at the ocean floor. Experiments at varying horizontal resolutions (R3, R9, R27) are performed, while the vertical discretization is kept constant, as increase was observed to have minimal effect by Lévy et al. (2010). The R3 is a coarse simulation with a horizontal resolution of 35.3 km (1/3°), which corresponds roughly to a central value of mesoscale eddy–permitting resolution. The finer resolution at R9 of 11.8 km (1/9°) is approaching realistic eddy–resolving ocean models (Maltrud & McClean, 2005; Sasaki et al., 2008). The finest resolution at R27 with 3.9 km (1/27°) is at a resolution where a significant portion of sub–mesoscale eddies would be resolved. The choice of resolutions corresponds to configurations that are well documented in the literature (Lévy et al., 2010, 2012) and for which values of the dissipation coefficient have been established. Additionally, the choice of resolutions was dictated by major differences observed in the classical deterministic simulation when the resolution is coarsened from R9 to R3 (Lévy et al., 2012).

The first run is initialized with a 1/3° resolution simulation, spun up for 1,000 years. The resulting state is then interpolated on the 1/9° and 1/27° grids to initialize the runs. The R27 simulation has been spun–up for 100 years before collecting data for the LU framework. Similarly, an additional 1/9° deterministic simulation has been spun up for 100 years under similar conditions to construct an initial state for the deterministic and stochastic 1/3° simulations. In order to assess the benefits brought by this stochastic approach, each stochastic simulation is compared to its deterministic counterpart at the same resolution and at the highest resolution (1/27°).

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

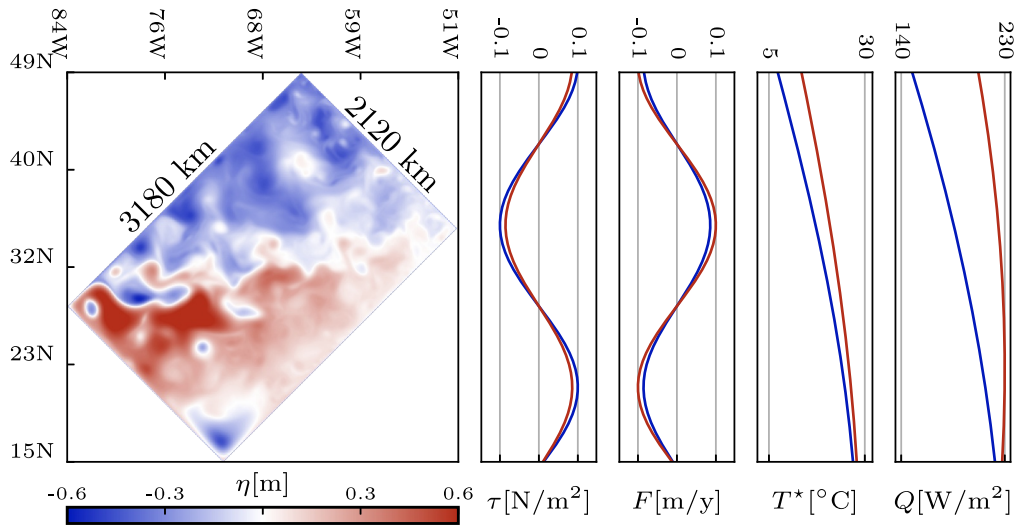


Figure 2. Geographical collocation of the domain with its external forcings: Surface wind stress τ imposed in the momentum equations, freshwater flux F imposed in the sea surface equation, restoring toward an apparent air temperature $T^{\star}$ and penetrative heat Q in the temperature equation. The red curve represents the summer forcing, while the blue curve represents the winter forcing. During the simulation year, the forcing oscillates regularly between these two extremes. Inside the domain, sea surface height η is plotted, showing the anomaly caused by the eastwards jet and thus providing its localization.

The model parameters for the three resolutions are given in Table 1. A bi–harmonic horizontal momentum diffusion (characterized by κ_M) is implemented for all the resolutions. For the tracers, that is, temperature and salinity, diffusion is implemented along isopycnal surfaces without horizontal background for R3 while for the finer resolution, a bi–harmonic horizontal diffusion (characterized by κ_T) is implemented. The values for the diffusion/friction coefficients are obtained from Lévy et al. (2010, 2012). Vertical diffusion for all resolutions are implemented via the turbulent closure model developed by Bougeault and Lacarrère (1989). The model is implemented with a background value of $10^{-5}\mathrm{m}^2\mathrm{s}^{-1}$. Partial slip boundary conditions are implemented in all the experiments. In the deterministic case, the Leapfrog method is employed for time–stepping, while a Milstein–type method is used to integrate the noise advection and Itô diffusion terms in the stochastic formulation. Details on these numerical methods are provided in Appendix D. Let us stress that the discrete spatial scheme remain unchanged.

4.2. Eddy–Permitting Regime

The primary objective of stochastic modeling is to reproduce large scale features in coarse–resolution models. As shown in Figure 3, the deterministic simulation (second column) lacks essential large–scale features such as the wind–driven eastwards jet. Therefore, the aim of the following experiments is to enrich the behavior of the flow with large scale characteristics, together with an enhancement of the fine scales variability. The choice of the noise model must be calibrated for these low–resolution simulations. On–line data models are unsuitable for this

Table 1
Model Parameters for Different Horizontal Resolutions

	R3	R9	R27
Horizontal resolution (in°)	1/3°	1/9°	1/27°
Horizontal resolution (in km)	35.3 km	11.8 km	3.9 km
Mesh grid (Nx × Ny)	90 × 60	270 × 180	820 × 540
Vertical levels (Nz)	31	31	31
Time step (in s)	3,600	1,200	300
Eddy Viscosity (κ_M)	$-10^{12}\mathrm{m}^4s^{-1}$	$-5 \times 10^{10}\mathrm{m}^4\mathrm{s}^{-1}$	$-5 \times 10^{9}\mathrm{m}^4\mathrm{s}^{-1}$
Eddy diffusivity (κ_T)	$300\ \mathrm{m}^2\mathrm{s}^{-1}$	$-5 \times 10^{10}\mathrm{m}^4\mathrm{s}^{-1}$	$-10^{9}\mathrm{m}^4\mathrm{s}^{-1}$

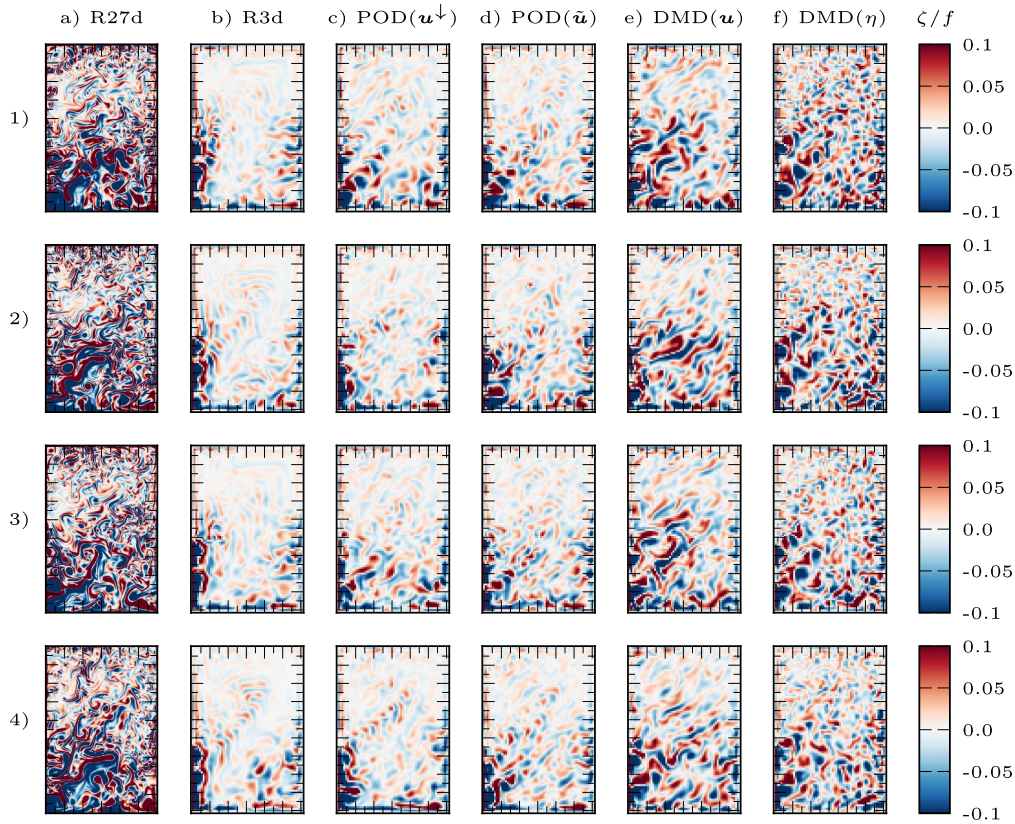


Figure 3. A sequence of snapshots of relative vorticity (ζ/f) maps for the four stochastic cases and the deterministic control case. Each row corresponds to a different time instant, chosen to be well representative of the instantaneous behavior: The first row shows one of the best snapshots for POD($\boldsymbol{u}^{\downarrow}$), the second shows one of the best snapshots for POD($\tilde{\boldsymbol{u}}$), the third shows one of the best snapshots for DMD($\boldsymbol{u}$) and finally the fourth row shows one of the best snapshots for DMD(η). We have chosen to show all models at each instant to show that the stochastic models are consistent in their temporal evolution.

purpose, as the simulation is not rich enough to sustain itself at such resolutions. Consequently, off–line data models are preferred.

Four experiments with different noise models are conducted in this regime, as summarized in Table 2. Two of these experiments are performed with noise based on POD, structured as described in Section 3: POD($\boldsymbol{u}^{\downarrow}$) is defined upon Equation 22 and thus contains all energy scales except the temporal average energetic scale; POD($\tilde{\boldsymbol{u}}$) is defined through Equation 25 and it is based on a high–pass temporal filter with 45 days frequency cut–off. Both POD–based models employ a Girsanov drift consisting only of a stationary bias. The third experiment relies on Dynamical Mode Decomposition noise, DMD($\boldsymbol{u}$), as defined in Equations 33 and 34, with temporal separation at 45 days. Finally, the last experiment, DMD(η), is performed with the noise described in Equation 41

Table 2
Summary of Noise Models Employed in Eddy–Permitting Regime

		Girsanov drift			Noise		
Name	Bias	Technique	Field	Filtering	Technique	Field	Filtering
POD($\boldsymbol{u}^{\downarrow}$)	$\boldsymbol{u}^{\flat}$ (26)	–	–	–	POD (27)	$\boldsymbol{u}$	Low–pass (space)
POD($\tilde{\boldsymbol{u}}$)	$\boldsymbol{u}^{\flat}$ (26)	–	–	–	POD (30)	$\boldsymbol{u}$	High–pass (time)
DMD($\boldsymbol{u}$)	$\boldsymbol{u}^{\flat}$ (26)	DMD (39)	$\boldsymbol{u}$	Low–pass (space)	DMD (38)	$\boldsymbol{u}$	High–pass (space)
DMD(η)	$\boldsymbol{u}^{\flat}$ (46)	DMD (46)	η	Low–pass (space)	VPP (42)	w	Box–Filter

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

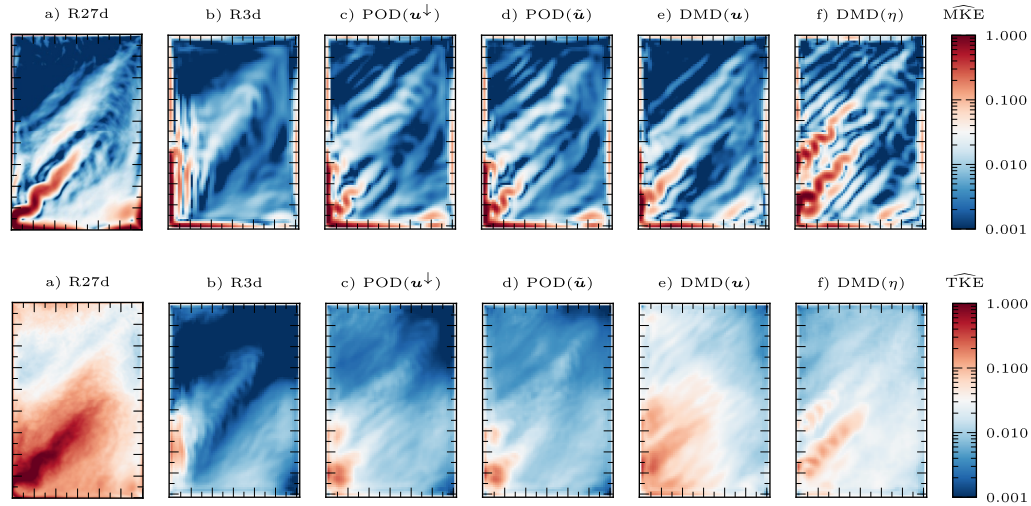


Figure 4. Comparison of the 10–years averaged mean kinetic energy ($\overline{\mathrm{MKE}}^{10Y}$, top) and turbulent kinetic energy ($\overline{\mathrm{TKE}}^{10Y}$, bottom) for the different models: (a) The deterministic high–resolution reference R27d; (b) the deterministic low–resolution benchmark R3d; (c) POD($\boldsymbol{u}^{\downarrow}$), constructed using a spatially low–pass filtered velocity field; (d) POD($\tilde{\boldsymbol{u}}$), constructed using a temporally high–pass filtered velocity field; (e) DMD($\boldsymbol{u}$), constructed using two velocity fields, one spatially low–pass filtered and the other spatially high–pass filtered; and (f) DMD(η), constructed using sea surface height and Vertical Profile Prescription. The kinetic energy fields are normalized with respect to the maximum energy value of the high–resolution deterministic simulation (R27d), such that $\hat{f} = f/\|f_{\mathrm{R27d}}\|_{\infty}$.

For all these noises, with the exception of the VPP component in DMD(η), both Girsanov and noise components are projected onto isopycnal surfaces.

In the following, we compare the performance of these stochastic models to the benchmark deterministic R3 run, using the reference R27 data for validation. The comparison encompasses both qualitative and quantitative analyses of spatial structures, temporal behavior, spectral density, and statistical metrics.

4.2.1. Qualitative Analyses

A sequence of vorticity snapshots from different simulations is presented in Figure 3. All stochastic models successfully reproduce the large–scale meandering jet on the coarse mesh at different time instances. The DMD–based noise models generate more small–scale structures compared to the POD–based methods.

This result is further confirmed and analyzed using the 10–years averaged fields.

Figure 4 shows the 10–year average of mean kinetic energy $\overline{\mathrm{MKE}}^{10Y} = \left[\left(\overline{u}^{10Y}\right)^2 + \left(\overline{v}^{10Y}\right)^2\right]/2$ and the turbulent kinetic energy $\overline{\mathrm{TKE}}^{10Y} = \left[\overline{(u')^2 + (v')^2}\right]^{10Y}/2$ (where u' stands for the fluctuations with respect to $\overline{u}^{10Y}$). The difference between the stochastic models and the deterministic counterpart is striking.

From the MKE maps, it can be readily observed that all noise models have a remarkable imprinting of the data in the long term statistics of the simulation, even though a systematic bias in the position of the jet is present for the DMD–based models. At this resolution, they all serve well the task of enhancing the representation of large scale features.

The TKE map shows also differences in variability: POD noise is more tied to large scale structures and to the most energetic areas by construction, thus the noise effect concentrated close to the boundary and the jet, while DMD noise is much more spread over the domain, with also strong impact along the jet. Concerning SSH noise, the Girsanov drift enhances the variability in the jet region, while the Gaussian filtered noise raises the level of energy uniformly in the domain, which was expected as there is no particular spatial structure for this latter model. The SSH noise, built from the Sea Surface Height, has a stronger signature on the secondary jet than the other noises defined from velocity data.

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

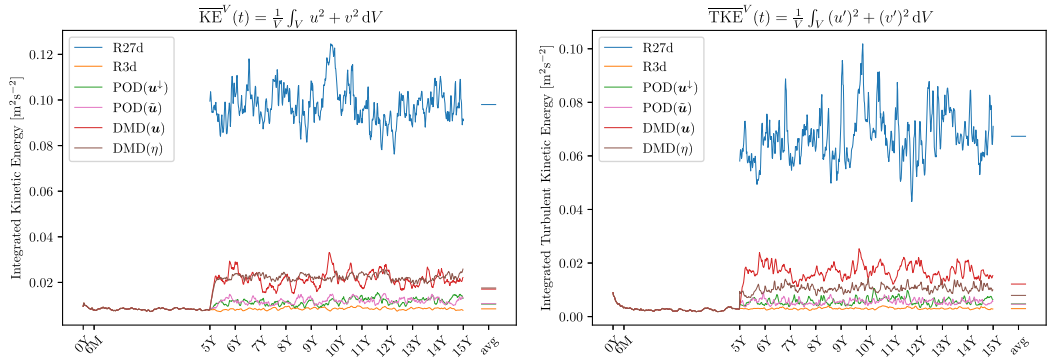


Figure 5. Comparison of the temporal evolution of the partial volume–integrated total kinetic energy (left) and turbulent kinetic energy (right) for different eddy–permitting models and the high–resolution reference.

4.2.2. Temporal Analysis

In this section, we consider the temporal behavior of kinetic energy integrated in the first 6 layers of the model, spanning approximately the upper 60 m. The integrated kinetic energy in time is defined as

$$\mathrm{KE}(t) = \frac{1}{2V}\int_V (u^2 + v^2)(\boldsymbol{x}, z, t)\,\mathrm{d}V, \tag{47}$$

where $V = \Omega \times [\eta_0, \eta_{-5}]$ and $\mathrm{d}V = \mathrm{d}\boldsymbol{x}\mathrm{d}z$. Consequently, the integrated turbulent kinetic energy in time is

$$\mathrm{TKE}(t) = \frac{1}{2V}\int_V \left[(u')^2 + (v')^2\right](\boldsymbol{x}, z, t)\,\mathrm{d}V. \tag{48}$$

Figure 5 shows the temporal behavior of the integrated kinetic energy in the top plot and that of turbulent kinetic energy in the bottom plot. All the simulations are run for a total of 15 years, where the stochastic components are switched on after 5 years of adjustment from the initial condition. For consistency, the fluctuations of the first 5 years are computed with respect of the average of the first 5 years only, while after the introduction of the stochastic parametrization the fluctuations are computed with respect to the average of the stochastic period alone, hence from year 5 to year 15. The energy of the high resolution R27 velocity data is shown in blue and it corresponds to the first 10 years but it has been shifted for visualization purposes.

A first trivial observation is that during the adjustment, all the low–resolution models provide the same integrated KE and TKE, as expected. During this period, KE remains of the same magnitude, while TKE suffers a steep decay in the initial 6 months, due to the dissipation of the high energy of the initial condition. After this highly dissipative period, TKE stabilizes around a much lower energy level than the initial condition.

Considering the deterministic case R3det, both the levels of KE and TKE persist until the end of the simulation almost unvaried. This is not the case of the stochastic simulations, where an increase of total and turbulent kinetic energy is visible at the beginning of year 5. Beyond energy exchanges, the additive Coriolis term in Equation 12, associated with the noise and the Girsanov drift, can also contribute to such increase in KE compared to the deterministic case.

All the stochastic simulations show a rapid increase of KE after the stochastic parametrization is started, with the DMD–based methods performing sensibly better than the POD methods for both total kinetic energy and turbulent kinetic energy. The POD based stochastic models share a similar behavior, with increased MKE and TKE with respect to the control simulation, while it is not recognizable if one method outperforms the others.

Finally, on the right of the plots the average level of energy is shown, with time average computed on the period from 5 to 15 years only, to leave aside the common deterministic initial period. DMD($\boldsymbol{u}$) introduces the higher turbulent kinetic energy, so the model outperforms all the others in terms of energy fluctuations. However, in

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

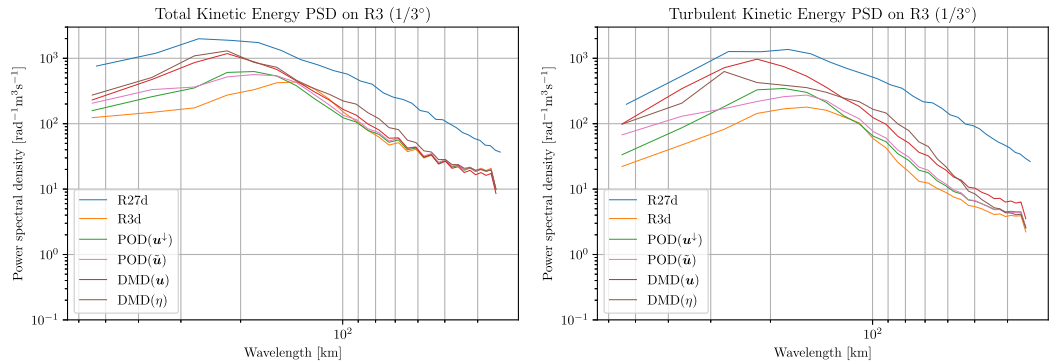


Figure 6. Comparison of the 10–years averaged total kinetic energy ($\overline{\text{KE}}^{10Y}$, left) and turbulent kinetic energy ($\overline{\text{TKE}}^{10Y}$, right) power spectral density for different eddy–permitting models with the high–resolution reference.

terms of total kinetic energy, the wide oscillations of KE render a lower average kinetic energy than DMD(η), that shows a more consistent total kinetic energy during the simulation, but lacks the energy provided by the velocity data for the noise construction. Nonetheless, the energy of the noise is much higher than that of the deterministic simulation. The two POD–based methods have a higher energy than the deterministic counterpart, even though they do not perform as well as the two DMD–based models.

4.2.3. Spectral Analyses

Next, we analyze the spectral distribution of energy and enstrophy in terms of wavenumbers and frequencies for the different models. These spectra provide insight into how different noise formulations impact the energy distribution across spatial and temporal scales.

Figure 6 shows the power spectral density for different fields: On the left, the total Kinetic Energy $\overline{\text{KE}}^{10Y} = \overline{\left[(u)^2 + (v)^2\right]}^{10Y}/2$, and on the right, the turbulent kinetic energy $\overline{\text{TKE}}^{10Y}$. Compared to the baseline deterministic model, all stochastic models significantly improve the representation of total KE at large scales and TKE across all scales. This improvement is particularly important for large–scale ocean models, which often face challenges in accurately representing mesoscale dynamical variability.

Among the stochastic models, the DMD–based models outperform the POD–based methods. This result is expected, as the DMD approach incorporates a richer, dynamically evolving Girsanov drift by construction. However, the two DMD–based models exhibit substantial differences: DMD($\boldsymbol{u}$) primarily enhances the TKE, while DMD(η) increases the total KE overall, with a partial contribution to the TKE through its VPP component. This behavior is consistent with their respective constructions. Specifically, DMD($\boldsymbol{u}$) is derived from the three–dimensional velocity field of high–resolution data, capturing both the Girsanov drift and noise components, making it a robust representation of the linear dynamics of the original baroclinic process. In contrast, DMD(η) relies on lower–dimensional data for its Girsanov drift and incorporates a separate Gaussian process for its noise component, leading to its distinct spectral characteristics.

In Figure 7 the temporal TKE and enstrophy spectra of the coarse–resolution simulations are shown, in comparison to the high–resolution simulation. The first observable feature is that the low–resolution deterministic simulation exhibits a spectrum shape that departs significantly from the high–resolution spectrum. At temporal scales smaller than 100 days the spectrum shows a steep decrease in kinetic energy intensity and a similar decay is shown for scales smaller than 50 days for the enstrophy, while the high resolution exhibits strong energetic and enstrophy patterns at all scales above 30 days. This highlights the low capability of the coarse–resolution simulation at resolving the ocean mesoscales.

Conversely, the stochastic parametrizations here proposed show a greater effect in the mesoscale range, from 20 to 100 days DMD(η) shows a remarkable resemblance with DMD($\boldsymbol{u}$), highlighting that sea surface elevation data can provide huge benefits at a reduced cost, with the VPP component sustaining better the energetic content at fast scales. Let us remark that the energy and enstrophy peaks visible around 25 days in the DMD(η) spectra are not

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

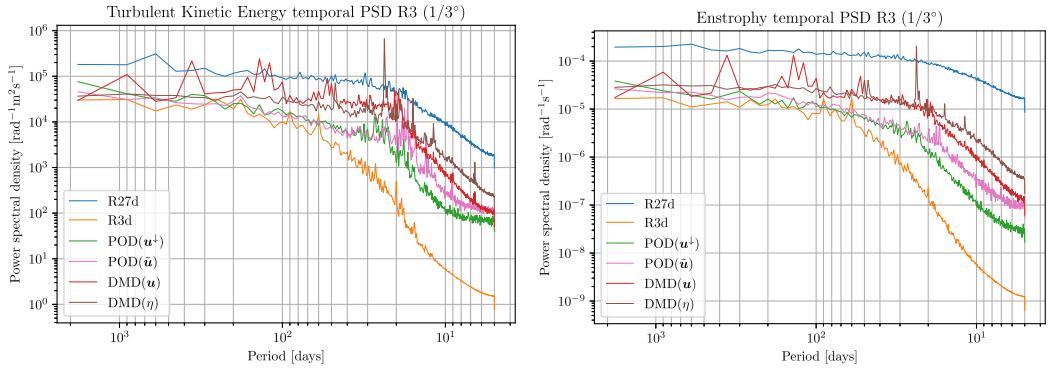


Figure 7. Comparison of turbulent kinetic energy (left) and enstrophy (right) temporal power spectral density (PSD) for different eddy–permitting models with the high–resolution reference. The temporal PSD is computed at each point in the domain and subsequently averaged in space.

present in the SSH data of the R27 and while we do not have an explanation for such behavior, it will be examined in further research.

The energy based noises, that is, the POD–based noises, show a general improvements compared to R3det, especially at fast scales. Once again, the selection of fast scales in POD($\tilde{\boldsymbol{u}}$) is visible in the enhancement of all scales faster than the 45 days threshold, consistently with the chosen windowing. In conclusion, the off–line data based stochastic parametrization does not serve only as a simple energetic offset toward a state of higher energy, but rather enhances the energetic spectra coherently with the scales selected in the preprocessing. Obviously, these improvements depend on the quality of the data used.

4.2.4. Statistical Analysis

To complete our analysis, we quantify the predictive accuracy of the model statistics relative to the high–resolution reference. To assess both the reconstruction of the temporal mean and variance within a single metric, we employ the Gaussian Relative Entropy (GRE), as proposed by Grooms et al. (2014). The GRE is defined as:

$$\mathrm{GRE} = \frac{1}{2}\left[\frac{\left(\overline{f}^{\,T}_{\mathrm{HR}} - \overline{f}^{\,T}_{\mathrm{LR}}\right)^2}{\sigma^2_{f,\mathrm{LR}}} + \frac{\sigma^2_{f,\mathrm{HR}}}{\sigma^2_{f,\mathrm{LR}}} - 1 - \ln\left(\frac{\sigma^2_{f,\mathrm{HR}}}{\sigma^2_{f,\mathrm{LR}}}\right)\right], \tag{49}$$

where $\overline{f}^{\,T} = T^{-1}\int_T f\,\mathrm{d}t$ denotes the time average and $\sigma^2_f = T^{-1}\int_T \left(f - \overline{f}^{\,T}\right)^2 \mathrm{d}t$ represents the temporal variance.

The first term on the right–hand side of GRE quantifies the error in the mean, weighted by the variance of the model. The remaining terms measure the error in model variability, commonly referred to as "dispersion". This non–dimensional criterion is minimized when the mean is perfectly reconstructed and the variance of the reference model matches that of the tested coarse–resolution model. A lower GRE value thus indicates a more accurate reconstruction.

Figure 8 presents the corresponding GRE maps for kinetic energy and enstrophy across different eddy–permitting models. The value displayed above each plot represents the domain–averaged GRE score. This statistical criterion confirms the previous findings. The two POD–based noise models yield better GRE scores than the benchmark deterministic model, particularly in the jet region. The DMD–based methods further enhance the results, showing improvements throughout the domain. Notably, the DMD(η) model performs almost as well as the DMD($\boldsymbol{u}$) model, which is highly encouraging in the context of using real–world altimetric observational data.

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

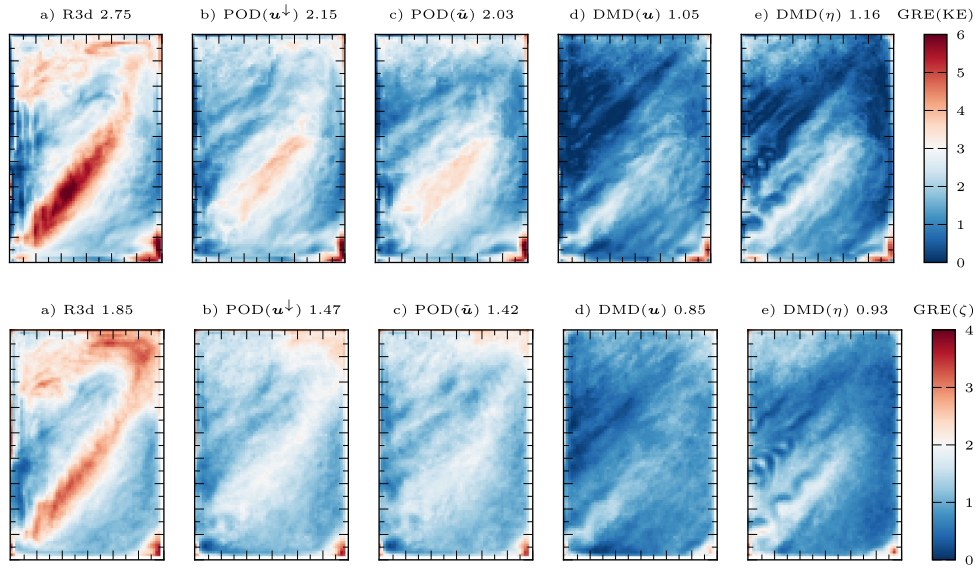


Figure 8. Gaussian relative entropy (GRE) maps of kinetic energy (top) and vorticity (bottom) for different eddy–permitting models. The value displayed above each plot represents the corresponding domain–averaged GRE. A lower GRE indicates a better reconstruction. In the color scheme, blue denotes better agreement, while red indicates higher discrepancies.

4.2.5. Summary

Within the eddy–permitting regime, all tested stochastic models successfully reproduce the large–scale flow structure and significantly enhance intrinsic small–scale variability compared to the benchmark deterministic model. The best results are achieved with DMD–based noise models. In particular, the DMD noise derived from SSH data, when combined with VPP noise, appears especially promising, as it opens opportunities for applications using satellite observational data. This noise exhibits greater variability than the noise generated through the DMD procedure applied to velocity fields.

4.3. Eddy–Resolving Regime

The formulation of the proposed stochastic model is fairly general and thus its application is not limited to large scale simulation. In this section we apply it to finer resolution simulations. The targeted resolution is now that of a $1/9$ of degree, corresponding to a horizontal resolution of 11.9 km. In this case, large scale features are already resolved, even if their long–term behavior is not as stable as in the reference simulation (3.9 km). The objective of this section is to evaluate whether the proposed methodology can help stabilize long–term statistics and enhance short–term variability in the simulation.

Given the large size of the simulation, on–line models are preferred to avoid burdening it with heavy data input. However, we conduct simulations using both on–line and hybrid models to assess their respective advantages. For off–line models, special care must be taken in the construction of noise to avoid over–representing motion scales already present in the R9 simulation. The dynamic component of the Girsanov drift will be modeled using DMD, based on a spatial high–pass filtered field, with the slow component selected through frequency separation. The noise component will be modeled using either off–line methods, such as DMD and POD (from high–pass filtered fields, spatially for DMD and both spatially and temporally for POD), or on–line methods.

Five experiments are conducted, as summarized in Table 3. Proper Orthogonal Decomposition, based on Equation 26, is used for POD($\boldsymbol{u}_{\mathcal{I}}$) and POD$^{\flat}$($\boldsymbol{u}_{\mathcal{I}}$), Dynamical Mode Decomposition is employed in DMD(η) and VPP is used in $\mathcal{G} * \mathrm{d}\boldsymbol{B}_t^{\mathrm{H}}$ and $\mathcal{G}^{\flat} * \mathrm{d}\boldsymbol{B}_t^{\mathrm{H}}$. The simulations denoted with a flat symbol $\flat$ employ a stationary Girsanov drift equal to Equation 21. Again, for all these noises, with only exception of the VPP component in R9ssh, both Girsanov and noise components are projected onto isopycnal surfaces.

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 3
Summary of Noise Models Employed in Eddy–Resolving Regime

		Girsanov drift			Noise		
Name	Bias	Technique	Field	Filtering	Technique	Field	Filtering
POD($\boldsymbol{u}_I$)	–	–	–	–	POD (31)	$\boldsymbol{u}$	High–pass (time)
POD$^\flat$($\boldsymbol{u}_I$)	$\boldsymbol{u}^\flat$ (26)	–	–	–	POD (31)	$\boldsymbol{u}$	High–pass (time)
$\mathcal{G} * \mathrm{d}\boldsymbol{B}_t^{\mathrm{H}}$	–	–	–	–	VPP (42)	w	Box–Filter
$\mathcal{G}^\flat * \mathrm{d}\boldsymbol{B}_t^{\mathrm{H}}$	$\boldsymbol{u}^\flat$ (26)	–	–	–	VPP (42)	w	Box–Filter
DMD(η)	$\boldsymbol{u}^\flat$ (46)	DMD (46)	η	Low–pass (space)	VPP (42)	w	Box–Filter

We compare the performance of different eddy–resolving models using diagnostics similar to those applied in the eddy–permitting case, except for the kinetic energy time series, as the R9 simulations exhibit nearly identical energy levels.

4.3.1. Qualitative Analyses

The sequence of snapshots in Figure 9 clearly illustrates that the model–driven VPP noise enhances the intrinsic flow variability, whereas the POD–based data–driven noise models do not show significant improvement over the benchmark deterministic run on the R9 mesh. Notably, the hybrid DMD(η) method retains the benefits of VPP noise.

The 10–years averaged fields in Figure 10 provide an initial overview of model performance. At first glance, the POD($\boldsymbol{u}_I$) appears to stabilize the jet current most effectively, as the average current expands further into the inner domain. Its *biased* counterpart, POD$^\flat$($\boldsymbol{u}_I$), preserves the jet current's length while slightly shifting the detachment point toward the reference detachment point of R27. This effect, a "southward" displacement of the detachment

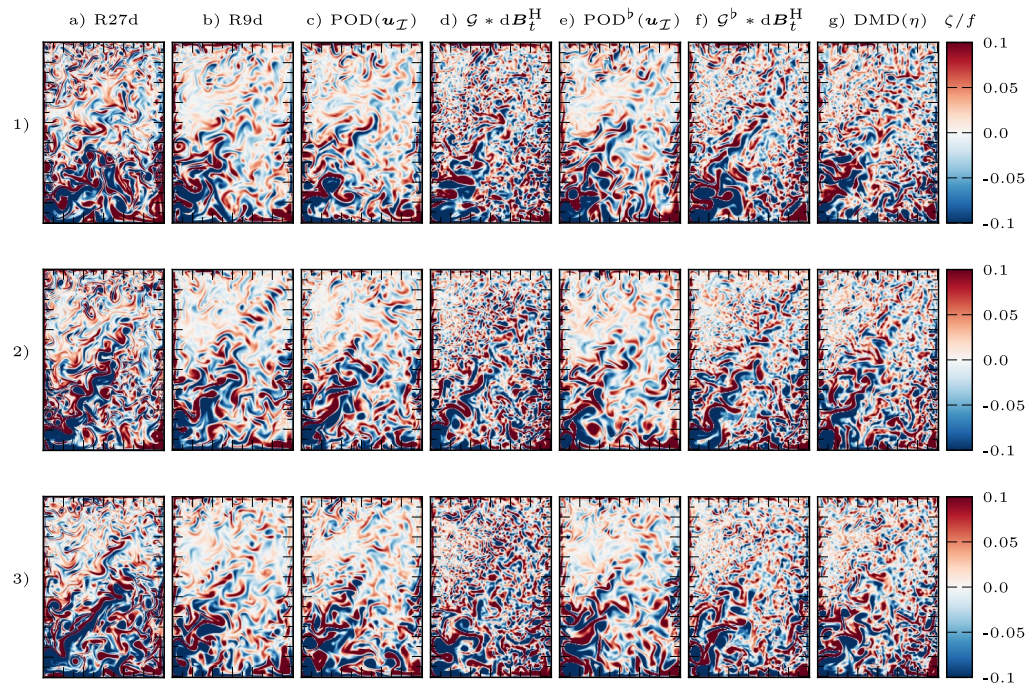


Figure 9. A sequence of snapshots of relative vorticity (ζ/f) maps for the five stochastic cases and the two deterministic control cases. Each row corresponds to a different time instant, chosen to be well representative of the instantaneous behavior: the first row shows one of the best snapshots for POD($\boldsymbol{u}_I$), the second shows one of the best snapshots for $\mathcal{G} * \mathrm{d}\boldsymbol{B}_t^{\mathrm{H}}$, the third shows one of the best snapshots for DMD(η). Once again, we have chosen to show all models at each instant to show that the stochastic models are consistent in their temporal evolution.

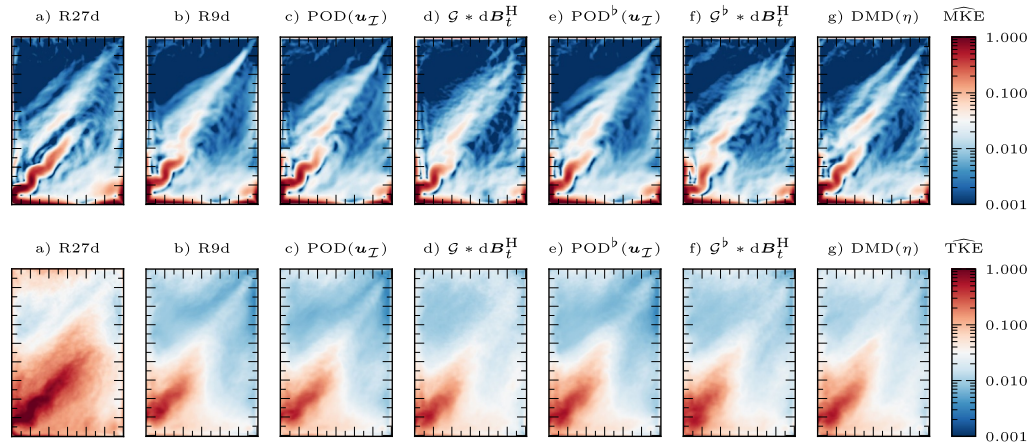


Figure 10. Comparison of the 10–years averaged mean kinetic energy ($\overline{\mathrm{MKE}}^{10\mathrm{Y}}$, top) and turbulent kinetic energy ($\overline{\mathrm{TKE}}^{10\mathrm{Y}}$, bottom) for the different models: (a) The deterministic high–resolution reference R27d; (b) the deterministic benchmark R9d; (c) POD($\boldsymbol{u}_\mathcal{I}$), constructed using an ensemble of temporally high–passed velocity fields; (d) $\mathcal{G} * \mathrm{d}\boldsymbol{B}_t^{\mathrm{H}}$, constructed using a Vertical Profile Prescription (VPP) noise; (e) POD$^\flat$($\boldsymbol{u}_\mathcal{I}$), constructed similarly to Proper Orthogonal Decomposition ($\boldsymbol{u}_\mathcal{I}$) but including a stationary bias; (f) $\mathcal{G}^\flat * \mathrm{d}\boldsymbol{B}_t^{\mathrm{H}}$, constructed similarly to $\mathcal{G} * \mathrm{d}\boldsymbol{B}_t^{\mathrm{H}}$ but including a stationary bias; and (g) Dynamic Mode Decomposition (η), constructed using sea surface height and VPP. The kinetic energy fields are normalized with respect to the maximum energy value of the high–resolution deterministic simulation (R27d), such that $\hat{f} = f/\|f_{\mathrm{R27d}}\|_\infty$.

point, is observed in all models that include the stationary bias (Equation 21) and it is particularly noticeable for $\mathcal{G}^\flat * \mathrm{d}\boldsymbol{B}_t^{\mathrm{H}}$, where the first meander of the jet current nearly coincides with that of the high–resolution reference simulation. However, in this experiment, the jet current length is not improved but rather shortened. Notably, in DMD(η), the length of the jet current is enhanced, its detachment point is coherent with the high–resolution simulation. The behavior of TKE is similar, the stochastic models (with the exception of $\mathcal{G}^\flat * \mathrm{d}\boldsymbol{B}_t^{\mathrm{H}}$) localize better the energy fluctuations around the main jet current and enhances its extension.

4.3.2. Spectral Analyses

Figure 11 presents a comparison of the spatial and temporal spectra across different models, displaying only those that exhibit distinguishable variations. The spatial TKE spectrum reveals that the VPP–based noise models induce the most substantial enhancement in energy density at mesoscales, particularly within the 30–100 km range. At wavelengths greater than 100 km, the overall increase in TKE becomes less pronounced across all models. These differences in TKE distribution have direct implications for total KE (not shown), leading to a similar trend. However, the effect is less pronounced in total KE, as all R9 simulations exhibit very similar MKE spectra.

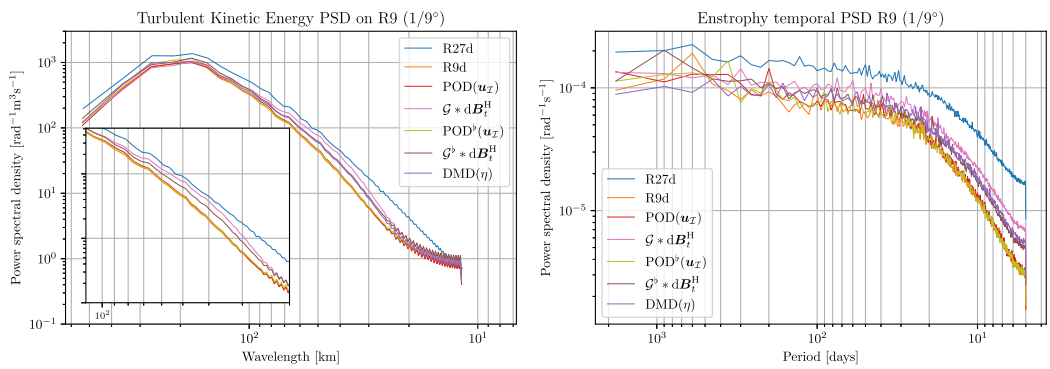


Figure 11. Comparison of the 10–years averaged turbulent kinetic energy $\left(\overline{\mathrm{TKE}}^{10\mathrm{Y}}\right)$ spatial power spectral density (PSD) (left) and enstrophy temporal PSD (right) for different eddy–resolving models with the high–resolution reference.

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

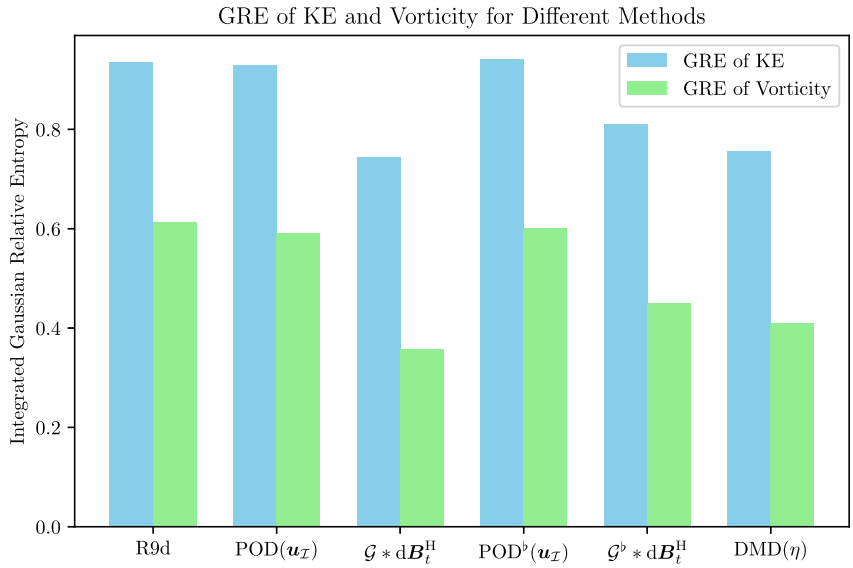


Figure 12. Comparison of the domain–averaged Gaussian Relative Entropy for kinetic energy (KE) and vorticity across different eddy–resolving models.

The temporal enstrophy spectrum in Figure 11 provides additional insight into the effects of different noise models on vorticity fluctuations. The results indicate that both POD($\boldsymbol{u}_I$) and POD$^\flat$($\boldsymbol{u}_I$) do not deviate significantly from the deterministic baseline model at scales below 100 days, despite noticeable differences in the average fields. In contrast, VPP–based noises, such as $\mathcal{G} * \mathrm{d}\boldsymbol{B}_t^{\mathrm{H}}$, $\mathcal{G}^\flat * \mathrm{d}\boldsymbol{B}_t^{\mathrm{H}}$ and DMD(η) introduce noticeably higher enstrophy than the other models, particularly at faster scales. This increased variability reflects the ability of these noise formulations to energize high–frequency fluctuations.

4.3.3. Statistical Analysis

The effectiveness of the models is statistically evaluated using the GRE metric (Equation 44).

Unlike in the eddy–permitting case, the local differences in GRE maps for different eddy–resolving models are less discernible. Therefore, we present only the domain–averaged global GRE for kinetic energy and vorticity across different models, as summarized in the bar plots in Figure 12.

The results indicate that the two POD–based models do not surpass the global GRE scores of the baseline deterministic model, particularly in the case of POD$^\flat$($\boldsymbol{u}_I$) model. Here, the stationary bias imposed by Equation 21 appears to be detrimental, potentially leading to a decline in performance. Conversely, the VPP–based noise models yield significant improvements in GRE scores, aligning with observations from the temporal enstrophy spectra and further reinforcing the differences in variability representation. Notably, the additional benefit of DMD(η) is further emphasized by this criterion.

4.3.4. Summary

In the eddy–resolving regime, we observe improvements across several quality criteria for the different noise types employed. However, these enhancements are notably less significant than those observed in the eddy–permitting regime. For data–driven noise models, this difference can be attributed to the greater scale separation in the eddy–permitting case, where the noise was specified using R27 data, compared to the eddy–resolving case, which relied on the same resolution data. Introducing higher–resolution data could potentially yield similar benefits to those observed at the $1/3^\circ$ resolution. The combination of model–driven VPP noise with DMD decomposition applied to the SSH variable appears particularly promising and will be the subject of future investigations.

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

5. Conclusion

This paper has investigated the primary benefits of implementing the LU framework in an operational ocean model. The test cases included both an eddy–permitting simulation (1/3°) and an eddy–resolving simulation (1/9°). In the eddy–permitting case, all selected metrics demonstrate a significant improvement in the representation of large–scale features and an enhancement of variability. Clearly, the choice of noise model plays a crucial role: Off–line data–driven models are preferred over on–line data–driven models, as the simulation's base state lacks sufficient energetic structures. These structures must be reinforced through higher–resolution simulations or other data sources, such as satellite observations. In this context, satellite altimeter data were prototyped in the DMD(η) noise model, where the sea surface elevation η from higher–resolution simulations serves as a proxy for real altimeter data.

Evaluating the eddy–resolving simulation is more challenging. This simulation already resolves a substantial portion of the large–scale energy, necessitating careful calibration when applying off–line data. Furthermore, it is difficult to test physics–oriented parametrizations in such an idealized configuration. Online noise models, like the proposed VPP method, enhance simulation variability, and combining off–line and on–line noise models allows for the inclusion of various physical processes in noise generation. The very weak improvement of the stationary bias, when not detrimental, suggests that the stationary mode of the high resolution simulation is not well estimated and, to work with a resolution such as 1/9°, a better converged statistics is necessary. Let us also point out that the discrepancy between the reference data (R27) and R9 is significantly reduced compared to R3. The role of the bias might be significantly improved with higher resolution data. The eddy–permitting test case is in addition useful and necessary for assessing the scalability of noise generation methods and their application, enabling us to design and test new strategies before applying them to realistic configurations.

Gaussian random fields with a prescribed vertical profile have been shown to yield promising results when coupled with SSH data. The vertical profile is typically determined from the resolved vertical velocity component. An alternative and promising approach involves introducing a dynamic vertical profile obtained from DMD and high–resolution simulations. Incorporating noise derived from SSH data, along with dynamically evolving vertical profiles, could enable eddy–resolving simulations that integrate elements from higher–resolution models and observational data. Additionally, a potential improvement lies in the development of noise models that capture the effects of internal and surface waves (Mémin et al., 2024), providing a more accurate representation of wave–current interactions. All these directions will be explored in future works, both in a low–resolution global circulation configuration and in a high–resolution regional configuration of the Mediterranean basin.

Appendix A: Remarks on the Noise Properties

Equation (1) represents the differential formulation of equation

$$\boldsymbol{X}_t = \boldsymbol{X}_0 + \int_0^t \boldsymbol{v}(\boldsymbol{X}_s, s)\,\mathrm{d}s + \int_0^t \int_S \check{\boldsymbol{\sigma}}(\boldsymbol{X}_t, \boldsymbol{y}, s)\,\mathrm{d}\boldsymbol{B}_s\,(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}. \tag{A1}$$

To specify the characteristics of the martingale term $\boldsymbol{\sigma}_t \mathrm{d}\boldsymbol{B}_t$, let H be the Hilbert space $H = L^2\left(S, \mathbb{R}^3\right)$, the space of square integrable functions over S with value in $\mathbb{R}^3$, with inner product $(\boldsymbol{f}, \boldsymbol{g})_H = \int_S \left(\boldsymbol{f}^\dagger \boldsymbol{g}\right) \mathrm{d}\boldsymbol{x}$ (with the upperscript † denoting transposed complex conjugation, making the operator linear only with respect to the second argument) and induced norm $\|\boldsymbol{f}\|_H = (\boldsymbol{f}, \boldsymbol{f})_H^{1/2}$, and let $T < +\infty$ be a finite time. In this context, $\{\boldsymbol{B}_t\}_{0 \le t \le T}$ is a cylindrical Wiener process defined on H as (Da Prato & Zabczyk, 2014):

$$\boldsymbol{B}_t = \sum_{n \in \mathbb{N}} \hat{\beta}^n \boldsymbol{e}_n, \tag{A2}$$

where $(\boldsymbol{e}_n)_{n \in \mathbb{N}}$ is a Hilbertian orthonormal basis of the space H and $\left(\hat{\beta}_n\right)_{n \in \mathbb{N}}$ is a sequence of independent standard Brownian motions.

The spatial correlation of the unresolved flow component is represented by the integral operator $\boldsymbol{\sigma}_t$, defined through a matrix kernel $\check{\boldsymbol{\sigma}} = \left(\check{\boldsymbol{\sigma}}_{ij}\right)_{1 \le i,j \le 3}$, interpreted as follows:

$$(\sigma_t f)^i(\boldsymbol{x}) = \int_S \breve{\sigma}_{ij}(\boldsymbol{x},\boldsymbol{y},t) f^j(\boldsymbol{y}) \mathrm{d}\boldsymbol{y}, \quad i = 1,2,3, \tag{A3}$$

where the Einstein summation convention for repeated indices is employed. The integral kernel $\breve{\boldsymbol{\sigma}}$ is assumed to be symmetric, positive definite and bounded in space and time. It follows that the integral operator $\boldsymbol{\sigma}_t$ is Hilbert–Schmidt, compact, self–adjoint, positive definite and thus, by Mercer's theorem, it admits a set of orthonormal eigenfunctions $\{\boldsymbol{\phi}_n(\cdot,t), n \in \mathbb{N}\}$ with the (strictly) positive eigenvalues $\lambda_n(t)$ decreasing toward zero and satisfying $\sum_{n\in\mathbb{N}}\lambda_n(t) < +\infty, \forall\ t \in [0,T]$. For random correlation operator $\boldsymbol{\sigma}_t$, the eigenvalues are random and in that case we will assume that $\mathbb{E}\left[\sum_{n\in\mathbb{N}}\lambda_n(t)\right] < \infty,\ \forall\ t \in [0,T]$, implying that the operator–valued process $\{\boldsymbol{\sigma}_t(\cdot)\}_{0\le t\le T}$ is stochastically integrable. As such, the spectral representation of the noise converges almost surely in H and in the sense of $L^p(\Omega, H)$, where $p \ge 1$ and Ω denotes the sample space associated with a probability space.

Associated with the correlation operator $\boldsymbol{\sigma}_t$, we define a matrix tensor $\boldsymbol{a} = (a_{ij})_{1\le i,j\le 3}$:

$$a_{ij}(\boldsymbol{x},t)\ \mathrm{d}t = \mathrm{d}\left\langle \int_0^{\cdot} (\sigma_s\ \mathrm{d}B_s)^i, \int_0^{\cdot} (\sigma_s\ \mathrm{d}B_s)^j \right\rangle_t = \int_S \breve{\sigma}_{ik}(\boldsymbol{x},\boldsymbol{y},t)\, \breve{\sigma}_{kj}(\boldsymbol{y},\boldsymbol{x},t)\ \mathrm{d}\boldsymbol{y}\ \mathrm{d}t. \tag{A4}$$

The quadratic variation—see the appendix of Bauer, Chandramouli, Li, Mémin, et al. (2020) for a brief recap—is a bounded variation process when the correlation operator, $\boldsymbol{\sigma}_t$, is random.

As introduced, the noise term $\boldsymbol{\sigma}_t\ \mathrm{d}\boldsymbol{B}_t$ is a zero–mean Wiener process of covariance tensor $\boldsymbol{Q}$, that is $\boldsymbol{\sigma}_t\ \mathrm{d}\boldsymbol{B}_t \sim \mathcal{N}(0,\boldsymbol{Q})$, where $\boldsymbol{Q}$ is the two–points two–times correlation tensor extending the one–point one–time correlation tensor $\boldsymbol{a}$, defined on a filtered probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P}\right)$, structured by a quadruple composed of a set, a sigma algebra, a filtration (i.e.., a nondecreasing family of sub–algebra) and a probability distribution.

Let $(\mathcal{F}_t)_{t\in[0,T]}$ be the filtration associated with $(\boldsymbol{B}_t)_{t\in[0,T]}$. A non–centered Wiener process shifted by a random process $\left(\tilde{\boldsymbol{Y}}_t\right)_{t\in[0,T]}$ can be defined as:

$$\tilde{\boldsymbol{B}}_t = \boldsymbol{B}_t + \int_0^t \tilde{\boldsymbol{Y}}_s\ \mathrm{d}s. \tag{A5}$$

Under good properties of $\left(\tilde{\boldsymbol{Y}}_t\right)_{t\in[0,T]}$ ($\mathcal{F}_t$–measurability, almost sure L^2–integrability and Novikov condition) there exists a measure $\mathbb{Q}$ such that:

- $\left(\tilde{\boldsymbol{B}}_t\right)_{t\in[0,T]}$ is a $\mathbb{Q}$–Wiener process;
- The Radon–Nikodym derivative (which is well defined due to Novikov condition (see e.g. Brémaud, 2020; Da Prato & Zabczyk, 2014) writes:

$$\frac{\mathrm{d}\,\mathbb{Q}}{\mathrm{d}\,\mathbb{P}} = \mathbb{E}\left[-\frac{1}{2}\left(\int_0^t \left(\tilde{\boldsymbol{Y}}_s,\ \mathrm{d}\boldsymbol{B}_s\right)_H + \int_0^t \|\tilde{\boldsymbol{Y}}_s\|_H^2\ \mathrm{d}s\right)\right]. \tag{A6}$$

With the non–centered (under $\mathbb{P}$) random process $\tilde{\boldsymbol{B}}_t$, we can rewrite the equations with respect to $\tilde{\boldsymbol{B}}_t$ as

$$\boldsymbol{\sigma}_t\ \mathrm{d}\boldsymbol{B}_t = \boldsymbol{\sigma}_t\ \mathrm{d}\tilde{\boldsymbol{B}}_t - \boldsymbol{\sigma}_t\ \tilde{\boldsymbol{Y}}_t\ \mathrm{d}t. \tag{A7}$$

The Lagrangian displacement under LU can thus be written as

$$\mathrm{d}\boldsymbol{X}_t = \left[\boldsymbol{v}(\boldsymbol{X}_t,t) - \boldsymbol{\sigma}_t\ \tilde{\boldsymbol{Y}}_t(\boldsymbol{X}_t)\right]\mathrm{d}t + \boldsymbol{\sigma}_t\ \mathrm{d}\tilde{\boldsymbol{B}}_t(\boldsymbol{X}_t). \tag{A8}$$

Under $\mathbb{Q}$ the Wiener process $\tilde{\boldsymbol{B}}_t$ is centered thus the writing of $\boldsymbol{X}_t$ has the same form as (Equation A1) but under a new measure.

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Appendix B: Stochastic Boussinesq Equations

Density variations are often induced by temperature or chemicals (as salinity), but these effects can be often regarded as small. To deal with the results of these changes without using the compressible Navier–Stokes equations, the so called *Boussinesq approximation* is employed. This approximation assumes that the density field involves just small changes around a reference ρ_0, so that the density field can be written as

$$\rho(\boldsymbol{x},t) = \rho_0 + \delta\rho(\boldsymbol{x},t) = \rho_0(1 + \epsilon\ \delta\hat{\rho}(\boldsymbol{x},t)), \tag{B1}$$

where $\delta\hat{\rho}(\boldsymbol{x},t)$ is a non–dimensional variable and $\epsilon \ll 1$ is a small constant parameter. Conservation of mass in the LU framework can be stated as

$$\mathrm{d}\int_{V_t} \rho(\boldsymbol{x},t)\ \mathrm{d}\boldsymbol{x} = 0 \tag{B2}$$

so that after inserting the Boussinesq approximation (Equation B1) one finds (omitting spatial and temporal dependencies and dividing both sides by the constant reference density ρ_0)

$$\mathrm{d}\int_{V_t} (1 + \epsilon\ \delta\hat{\rho})\ \mathrm{d}\boldsymbol{x} = 0. \tag{B3}$$

Upon applying the stochastic Reynolds transport theorem (and canceling all the differentiations of the constant term) we obtain

$$\int_{V_t} \mathrm{d}_t\ \epsilon\delta\hat{\rho} + \nabla\cdot\{(1 + \epsilon\ \delta\hat{\rho})\,[(\boldsymbol{v} - \boldsymbol{v}^s)\ \mathrm{d}t + \boldsymbol{\sigma}_t\ \mathrm{d}\boldsymbol{B}_t]\} - \epsilon\frac{1}{2}\nabla\cdot(\boldsymbol{a}\nabla\ \delta\hat{\rho})\ \mathrm{d}t\ \mathrm{d}\boldsymbol{x} = 0, \tag{B4}$$

which corresponds to the formulation

$$\int_{V_t} \nabla\cdot[(\boldsymbol{v} - \boldsymbol{v}^s)\ \mathrm{d}t + \boldsymbol{\sigma}_t\mathrm{d}\boldsymbol{B}_t]\ \mathrm{d}\boldsymbol{x} + \epsilon\left[\mathrm{d}\int_{V_t} \delta\hat{\rho}\ \mathrm{d}\boldsymbol{x}\right] = 0. \tag{B5}$$

The first order term in this equation corresponds to an isochoric constraint. If ϵ is small enough, the incompressibility condition predominates. This of course does not imply that the fluid is itself incompressible, but rather that mathematically at the limit of small compressibility, the flow can be considered as isochoric. At lowest order, the divergence–free condition

$$\nabla\cdot[(\boldsymbol{v} - \boldsymbol{v}^s)\ \mathrm{d}t + \boldsymbol{\sigma}_t\mathrm{d}\boldsymbol{B}_t] = 0, \tag{B6}$$

is hence found. Separating the scales of motion (in terms of martingale and bounded variation terms), we get that the noise term is incompressible by itself, with

$$\nabla\cdot\boldsymbol{\sigma}_t\mathrm{d}\boldsymbol{B}_t = 0, \tag{B7}$$

implying that $\nabla\cdot\boldsymbol{\sigma}_t = 0$, so that the modified advection reduces to $\boldsymbol{v}^\star = \boldsymbol{v} - \boldsymbol{v}^s$. The transport expressions of a scalar θ are provided directly by the stochastic Reynolds transport theorem where the term $\boldsymbol{\sigma}^{\mathrm{T}}(\nabla\cdot\boldsymbol{\sigma}_t)$ cancels, meaning that the two forms

$$\mathrm{d}_t\theta + \nabla\cdot[\theta(\boldsymbol{v} - \boldsymbol{v}^s)\ \mathrm{d}t + \theta\boldsymbol{\sigma}_t\mathrm{d}\boldsymbol{B}_t] - \frac{1}{2}\nabla\cdot(\boldsymbol{a}\nabla\theta)\ \mathrm{d}t + \nabla\cdot(\boldsymbol{\sigma}_t\boldsymbol{\Theta}_\sigma)\ \mathrm{d}t = \Theta_t\ \mathrm{d}t + \boldsymbol{\Theta}_\sigma\cdot\mathrm{d}\boldsymbol{B}_t,$$

$$\mathrm{d}_t\theta + [(\boldsymbol{v} - \boldsymbol{v}^s)\ \mathrm{d}t + \boldsymbol{\sigma}_t\mathrm{d}\boldsymbol{B}_t]\cdot\nabla\theta - \frac{1}{2}\nabla\cdot(\boldsymbol{a}\nabla\theta)\ \mathrm{d}t + \nabla\cdot(\boldsymbol{\sigma}_t\boldsymbol{\Theta}_\sigma)\ \mathrm{d}t = \Theta_t\ \mathrm{d}t + \boldsymbol{\Theta}_\sigma\cdot\mathrm{d}\boldsymbol{B}_t,$$

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

are interchangeable. Moreover, the continuity equation states that at lowest order there is no dynamics for the density. We stress that all the previous computations should have been performed in non–dimensional variables. However, all terms are homogeneous in dimensions so the dimensional terms factorize without exposing particular relations (contrary to what happens for momentum equations) and thus the notation has been kept dimensional for the sake of notation clarity. Conservation of momentum can be derived starting from an integral balance between the variation of momentum, the gravitational field acting on the fluid and the divergence of the stress tensor for a Newtonian fluid, neglecting terms involving molecular diffusion (which is standard practice in large scale dynamics) and noise compressibility, resulting in:

$$\mathrm{d}\int_{V_t}\rho\boldsymbol{v}\,\mathrm{d}\boldsymbol{x}=-\int_{V_t}\left[\rho g\boldsymbol{e}_3\,\mathrm{d}t+\nabla\mathrm{d}\mathrm{P}_t\right]\mathrm{d}\boldsymbol{x}, \tag{B8}$$

where $\boldsymbol{e}_3$ stands for the vertical unit vector. It has to be highlighted that, in order for the previous equation to be consistent in its time–scale separation, the pressure dP_t must be coherent with the scales of motion of the process, so that the right and left hand sides must match in both slow scale terms and fast scale terms. The total pressure must thus be composed of a bounded variation process and a martingale, that is

$$\mathrm{dP}_t=p\,\mathrm{d}t+\mathrm{d}p_t^{\sigma}, \tag{B9}$$

where the martingale term $\mathrm{d}p_t^{\sigma}$ is referred to as *stochastic pressure* and is defined as a Lagrange multiplier associated to the divergence–free condition of the noise. It takes a similar integral form as the noise:

$$\mathrm{d}p_t^{\sigma}(\boldsymbol{x})=\int_S\vartheta_i(\boldsymbol{x},\boldsymbol{y},t)\,\mathrm{d}B_t^i(\boldsymbol{y})\mathrm{d}\boldsymbol{y}, \tag{B10}$$

and is a zero–mean turbulent pressure related to the small–scale velocity component. The term $\boldsymbol{\vartheta}$ is the martingale pressure kernel and represents the spatial correlation of the pressure fluctuations induced by the small–scale velocity, and it will be responsible for the term $\nabla\cdot[(\nabla\boldsymbol{\vartheta})\boldsymbol{\sigma}_t]\,\mathrm{d}t$ in Equation B26. In practice, this term is often neglected, recovering the models described in Resseguier et al. (2017). Splitting the total pressure into its scale separated form reads

$$\mathrm{d}\int_{V_t}\rho\boldsymbol{v}\,\mathrm{d}\boldsymbol{x}=-\int_{V_t}\left[\left(\nabla p+\rho g\boldsymbol{e}_z\right)\mathrm{d}t+\nabla\mathrm{d}p_t^{\sigma}\right]\mathrm{d}\boldsymbol{x}. \tag{B11}$$

Inserting (Equation B1) one finds

$$\mathrm{d}\int_{V_t}\left(\rho_0+\delta\rho\right)\boldsymbol{v}\,\mathrm{d}\boldsymbol{x}=-\int_{V_t}\left[\left(\nabla p+\left(\rho_0+\delta\rho\right)g\boldsymbol{e}_z\right)\mathrm{d}t+\nabla\mathrm{d}p_t^{\sigma}\right]\mathrm{d}\boldsymbol{x}, \tag{B12}$$

that can be rearranged conveniently as

$$\mathrm{d}\int_{V_t}\boldsymbol{v}\,\mathrm{d}\boldsymbol{x}+\mathrm{d}\int_{V_t}\frac{\delta\rho}{\rho_0}\boldsymbol{v}\,\mathrm{d}\boldsymbol{x}=-\int_{V_t}\frac{1}{\rho_0}\nabla\left[p\,\mathrm{d}t+\mathrm{d}p_t^{\sigma}\right]\mathrm{d}\boldsymbol{x}-\int_{V_t}g\boldsymbol{e}_z\,\mathrm{d}t\,\mathrm{d}\boldsymbol{x}-\int_{V_t}\frac{\delta\rho}{\rho_0}g\boldsymbol{e}_z\,\mathrm{d}t\,\mathrm{d}\boldsymbol{x}.$$

To understand the leading orders of the momentum equation one must employ an asymptotic analysis, so the non–dimensional variables

$$\boldsymbol{x}=L\hat{\boldsymbol{x}},\qquad\boldsymbol{v}=U\hat{\boldsymbol{v}},\qquad t=\frac{L}{U}\hat{t},\qquad p=P\hat{p},\qquad g=\frac{P}{\rho_0 L}\hat{g}, \tag{B13}$$

are introduced, with the definition $\frac{\delta\rho}{\rho_0}=\epsilon\hat{\rho}_1$, so that the non–dimensional integral momentum balance reads

$$\mathrm{d}\int_{\hat{V}_t}\hat{\boldsymbol{v}}\,\mathrm{d}\hat{\boldsymbol{x}}+\epsilon\mathrm{d}\int_{\hat{V}_t}\hat{\rho}_1\,\hat{\boldsymbol{v}}\,\mathrm{d}\hat{\boldsymbol{x}}=-\frac{1}{\rho_0 UL}\int_{\hat{V}_t}\hat{\nabla}\left[\frac{PL}{U}\hat{p}\,\mathrm{d}\hat{t}+P^{\sigma}\mathrm{d}\hat{p}_t^{\sigma}\right]+\frac{PL}{U}\hat{\boldsymbol{g}}\boldsymbol{e}_3\,\mathrm{d}\hat{t}+\epsilon\frac{PL}{U}\hat{\rho}_1\hat{\boldsymbol{g}}\boldsymbol{e}_3\,\mathrm{d}\hat{t}\,\mathrm{d}\hat{\boldsymbol{x}}.$$

The scaling P^{σ} for the martingale part of the pressure has been also introduced. Expanding each variable as an asymptotic with ϵ taken as a smallness parameter one has

$$\mathrm{d}\int_{\hat{V}_t}\hat{\boldsymbol{v}}_0\,\mathrm{d}\hat{\boldsymbol{x}}+\epsilon\mathrm{d}\int_{\hat{V}_t}\hat{\rho}_1\,\hat{\boldsymbol{v}}_0\,\mathrm{d}\hat{\boldsymbol{x}}=-\frac{1}{\rho_0 UL}\int_{\hat{V}_t}\left[\frac{PL}{U}\left[\hat{\nabla}\left(\hat{p}_0+\epsilon\hat{p}_1\right)+\hat{g}\right]\mathrm{d}\hat{t}\right.$$
$$\left.+\hat{\nabla}P_0^{\sigma}\mathrm{d}\hat{p}_t^{\sigma}+\epsilon\frac{PL}{U}\hat{\rho}_1\hat{\boldsymbol{g}}\boldsymbol{e}_3\,\mathrm{d}\hat{t}\right]\mathrm{d}\hat{\boldsymbol{x}}.$$

Defining $P=\rho^{\star}U^2$ with $\rho^{\star}$ such that $\rho^{\star}=\frac{\rho_0}{\epsilon}$ provides

$$\mathrm{d}\int_{\hat{V}_t}\hat{\boldsymbol{v}}\,\mathrm{d}\hat{\boldsymbol{x}}+\epsilon\mathrm{d}\int_{\hat{V}_t}\hat{\rho}_1\,\hat{\boldsymbol{v}}\,\mathrm{d}\hat{\boldsymbol{x}}=\frac{1}{\epsilon}\int_{\hat{V}_t}\left[\hat{\nabla}\hat{p}_0+\hat{\boldsymbol{g}}\boldsymbol{e}_3\right]\mathrm{d}\hat{\boldsymbol{x}}-\int_{\hat{V}_t}\left[\hat{\nabla}\hat{p}_1+\hat{\rho}_1\hat{\boldsymbol{g}}\boldsymbol{e}_3\right]\mathrm{d}\hat{t}\,\mathrm{d}\hat{\boldsymbol{x}}$$
$$-\frac{1}{\rho_0 UL}\int_{\hat{V}_t}\hat{\nabla}P^{\sigma}\mathrm{d}\hat{p}_t^{\sigma}\,\mathrm{d}\hat{\boldsymbol{x}}.$$

This choice is stating that the dynamics is driven by density variations rather than the mean state. Consequently, the reference density is defined as the mean state density ρ_0 amplified by the ratio between the mean state and the variation, $\frac{\rho_0}{\delta\rho}=\frac{1}{\epsilon}$. In this equation, the leading order is $\mathcal{O}(1/\epsilon)$ and is described by a hydrostatic distribution of the pressure

$$\int_{\hat{V}_t}\nabla\hat{p}_0\,\mathrm{d}\hat{\boldsymbol{x}}=-\int_{\hat{V}_t}\hat{\boldsymbol{g}}\boldsymbol{e}_3\,\mathrm{d}\hat{\boldsymbol{x}}. \tag{B14}$$

Once dimensional variables are replaced by non–dimensional variables into Equation ($B14$), the hydrostatic balance reads

$$\nabla p_0=-\rho_0 g\boldsymbol{e}_z, \tag{B15}$$

and so pressure has a zero–th order component that is only depending on the vertical coordinate and on the mean state density:

$$p_0(z)=-\rho_0 gz. \tag{B16}$$

The separation of the density into two components, the background constant density and the deviation from it, corresponds thus to a division of the pressure between a component in hydrostatic balance and the fluctuation from this reference,

$$\rho(\boldsymbol{x},t)=\rho_0+\delta\rho(\boldsymbol{x},t),\qquad p(\boldsymbol{x},t)=-\rho_0 gz+p'(\boldsymbol{x},t), \tag{B17}$$

with this latter fluctuation that can be identified with the first order component of the pressure

$$p'(\boldsymbol{x},t)=p(\boldsymbol{x},t)+\rho_0 gz. \tag{B18}$$

The following order, $\mathcal{O}(\epsilon)$, reads

$$\mathrm{d}\int_{\hat{V}_t}\hat{\boldsymbol{v}}\,\mathrm{d}\hat{\boldsymbol{x}}=-\int_{\hat{V}_t}\left[\hat{\nabla}\hat{p}_1+\hat{\rho}_1\hat{\boldsymbol{g}}\boldsymbol{e}_3\right]\mathrm{d}\hat{t}\,\mathrm{d}\hat{\boldsymbol{x}}-\frac{1}{\rho_0 UL}\int_{\hat{V}_t}\hat{\nabla}P^{\sigma}\mathrm{d}\hat{p}_t^{\sigma}, \tag{B19}$$

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

so the stochastic Reynolds transport theorem can be applied to this equation, providing

$$\begin{aligned} \mathrm{d}_t \boldsymbol{v} + \nabla \cdot [\boldsymbol{v} \otimes (\boldsymbol{v} - \boldsymbol{v}^s)\, \mathrm{d}t + \boldsymbol{v} \otimes \boldsymbol{\sigma}_t \mathrm{d}\boldsymbol{B}_t] - \frac{1}{2}\nabla \cdot (\boldsymbol{a}\nabla \boldsymbol{v})\, \mathrm{d}t - \nabla \cdot [(\nabla \boldsymbol{\vartheta})\boldsymbol{\sigma}_t]\, \mathrm{d}t \\ = -\frac{1}{\rho_0}\nabla\left[p'\, \mathrm{d}t + \mathrm{d}p_t^\sigma\right] - \frac{\delta\rho}{\rho_0} g \boldsymbol{e}_3\, \mathrm{d}t. \end{aligned} \tag{B20}$$

The last term is the *buoyancy*

$$\boldsymbol{b}(\boldsymbol{x}, t) = -g\frac{\delta\rho(\boldsymbol{x}, t)}{\rho_0}\boldsymbol{e}_3, \tag{B21}$$

representing the upward (or downward) force associated with the density anomaly $\delta\rho$. Within the hydrostatic setting presented in Equation 13 the effects of the correlation term between martingale pressure and small–scale velocity $\nabla \cdot [(\nabla \boldsymbol{\vartheta})\boldsymbol{\sigma}_t]\, \mathrm{d}t$ can be neglected with the following scaling argument: introducing the length scale of the noise L^σ and the temporal scale of the noise T^σ, one has the following scalings

$$\boldsymbol{\sigma}_t \mathrm{d}\boldsymbol{B}_t = L^\sigma\, \hat{\boldsymbol{\sigma}}_t \mathrm{d}\hat{\boldsymbol{B}}_t, \quad \boldsymbol{B}_t = \sqrt{T^\sigma}\, \hat{\boldsymbol{B}}_t, \quad \boldsymbol{\sigma}_t = \frac{L^\sigma}{\sqrt{T^\sigma}}\, \hat{\boldsymbol{\sigma}}_t. \tag{B22}$$

The scaling of the pressure kernel, can be estimated by equating the scaling of general definition of the stochastic pressure (Equation B10)

$$\mathrm{d}p_t^\sigma = \int_\Omega \vartheta_i(\boldsymbol{x}, \boldsymbol{y}, t)\, \mathrm{dB}_t^i(\boldsymbol{y})\mathrm{d}\boldsymbol{y} = \Theta\sqrt{T^\sigma}\int_\Omega \hat{\vartheta}_i(\boldsymbol{x}, \boldsymbol{y}, t)\, \mathrm{d}\hat{\mathrm{B}}_t^i(\boldsymbol{y})\mathrm{d}\boldsymbol{y}, \tag{B23}$$

with the scaling of its hydrostatic counterpart:

$$\mathrm{d}p_t^\sigma = \int_0^z \boldsymbol{\sigma}_t \mathrm{d}\boldsymbol{B}_t \cdot \nabla w = L^\sigma \frac{H^2}{L^2} U \int_0^{\hat{z}} \hat{\boldsymbol{\sigma}}_t \mathrm{d}\hat{\boldsymbol{B}}_t \cdot \nabla \hat{w}. \tag{B24}$$

This renders $\Theta = \frac{L^\sigma}{\sqrt{T^\sigma}}\frac{H^2}{L^2}U$ and thus $\nabla \cdot [(\nabla \boldsymbol{\vartheta})\boldsymbol{\sigma}_t]\, \mathrm{d}t \sim \frac{1}{L}\left[\frac{L^\sigma}{\sqrt{T^\sigma}}\frac{H}{L}\right]^2 \ll 1$ for a moderate noise $\left(L^\sigma/\sqrt{T^\sigma} = \mathcal{O}(1)\right)$. Within the hydrostatic hypothesis it can be safely neglected.

Using the notation of the stochastic transport operator (Equation 10), the momentum of the Stochastic Boussinesq equations can be compactly written as

$$\mathbb{D}_t \boldsymbol{v} = -\frac{1}{\rho_0}\nabla\left[p'\, \mathrm{d}t + \mathrm{d}p_t^\sigma\right] - \boldsymbol{b}\, \mathrm{d}t. \tag{B25}$$

Considering now the terms of order $\mathcal{O}(\epsilon)$ of the conservation of mass statement, we get for density variation $\delta\rho$ the conservation equation,

$$\mathrm{d}_t \delta\hat{\rho} + \nabla \cdot [\delta\hat{\rho}(\boldsymbol{v} - \boldsymbol{v}^s)\, \mathrm{d}t + \delta\hat{\rho}\, \boldsymbol{\sigma}_t \mathrm{d}\boldsymbol{B}_t] - \frac{1}{2}\nabla \cdot (\boldsymbol{a}\nabla\, \delta\hat{\rho})\, \mathrm{d}t = 0, \tag{B26}$$

which after multiplication by $\frac{g}{\rho_0}$ provides an equation for transport of buoyancy:

$$\mathbb{D}_t\, b = 0. \tag{B27}$$

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

B1. Effects of Rotation on the Stochastic Material Derivative

To take into account the effect of a frame in deterministic rotation around a given axis, we can still use the Frenet–Serret formulas:

$$\mathrm{d}\,\boldsymbol{C} = |C|\;|\Omega\,\mathrm{d}t|\,(\sin\theta)\,\boldsymbol{m}, \tag{B28}$$

where $\boldsymbol{m}$ is direction of the circular motion of the point C connected with the center by the vector $\boldsymbol{C}$. For a better understanding of this formula, refer to Chapter 2 of Vallis (2017). This formula can be written in vector notation as $\mathrm{d}\boldsymbol{C} = \boldsymbol{\Omega}\,\mathrm{d}t \times \boldsymbol{C}$. For convenience, the translation of the non–inertial system is a priori neglected and the same considerations as in the deterministic case hold in this case too. The procedure mimics the deterministic one almost verbatim, showing however some subtleties. Within the LU model, it is important to understand that $t \mapsto \boldsymbol{e}_i(t)$ is a bounded variation process, its covariation with any process is null, so defining $\boldsymbol{v}$ as the (stochastic) velocity in the inertial frame and $\boldsymbol{v}'$ its counterpart in the rotating frame, with the application of the Frenet–Serret formula one retrieves the usual relation

$$\begin{aligned}\mathrm{d}_I\boldsymbol{v} &= \mathrm{d}_I\boldsymbol{v}' = \mathrm{d}_I\left(v_1'\boldsymbol{e}_1' + v_2'\boldsymbol{e}_2' + v_3'\boldsymbol{e}_3'\right) = \left(\mathrm{d}_I v_i'\right)\boldsymbol{e}_i' + v_i'\,\mathrm{d}_I\boldsymbol{e}_i' \\ &= \mathrm{d}_R\boldsymbol{v}' + (\boldsymbol{\Omega}\,\mathrm{d}t \times \boldsymbol{v}'),\end{aligned}$$

where the subscript I indicates the inertial frame and R the rotating frame. Starting from Equation 1 a total velocity can then be formally defined as

$$\boldsymbol{V}(\boldsymbol{x},t) = \boldsymbol{v}(\boldsymbol{x},t) + \boldsymbol{\sigma}_t\dot{\boldsymbol{B}}_t(\boldsymbol{x}), \tag{B29}$$

where the uncertainty term $\mathrm{d}\boldsymbol{B}_t$ is not differentiable in time in the classical sense and so its derivative $\dot{\boldsymbol{B}}_t$ must be thought in terms of weak derivatives, that is

$$\int_T \varphi(s)\dot{\boldsymbol{B}}_s(\boldsymbol{x})\,\mathrm{d}s = -\int_T \dot{\varphi}(s)\boldsymbol{B}_s(\boldsymbol{x})\,\mathrm{d}s, \qquad \forall\,\varphi \in C_0^\infty(T). \tag{B30}$$

From the physical point of view, $\boldsymbol{v}$ is a smooth–in–time random velocity, while $\boldsymbol{\sigma}_t\dot{\mathbf{B}}_t$ represents the noise: A generalized stochastic process that has to be defined in the space of tempered distributions. We assume that this term, living at a much faster scale than $\boldsymbol{B}_t$ is negligible or compensated by high fluctuations of the forcings. Applying Frenet–Serret formula Equation B28–B29.

$$\begin{aligned}\mathrm{d}_I\int_{V_t}\rho\boldsymbol{V}\,\mathrm{d}\boldsymbol{x} &= \mathrm{d}_R\int_{V_t}\rho\boldsymbol{V}'\,\mathrm{d}\boldsymbol{x} + \boldsymbol{\Omega}\,\mathrm{d}t\times\left(\int_{V_t}\rho\boldsymbol{V}'\,\mathrm{d}\boldsymbol{x}\right), \\ \mathrm{d}_I\int_{V_t}\rho\left(\boldsymbol{v}+\boldsymbol{\sigma}\dot{\boldsymbol{B}}_t\right)\mathrm{d}\boldsymbol{x} &= \mathrm{d}_R\int_{V_t}\rho\left(\boldsymbol{v}'+\boldsymbol{\sigma}'\dot{\boldsymbol{B}}_t'\right)\mathrm{d}\boldsymbol{x} + \boldsymbol{\Omega}\,\mathrm{d}t\times\int_{V_t}\rho\left(\boldsymbol{v}'+\boldsymbol{\sigma}'\dot{\boldsymbol{B}}_t'\right)\mathrm{d}\boldsymbol{x},\end{aligned}$$

and interpreting the derivatives in a distributional sense, one finds the relation

$$\begin{aligned}\int_T\varphi(s)\left[\mathrm{d}_I\int_{V_s}\rho\boldsymbol{v}\,\mathrm{d}\boldsymbol{x}\right]\mathrm{d}s = &\int_T\varphi(s)\left[\mathrm{d}_R\int_{V_s}\rho\boldsymbol{v}'\,\mathrm{d}\boldsymbol{x} + \boldsymbol{\Omega}\times\int_{V_s}\rho\left(\boldsymbol{v}'\,\mathrm{d}t + \boldsymbol{\sigma}'\mathrm{d}\boldsymbol{B}_t'\right)\mathrm{d}\boldsymbol{x}\right]\mathrm{d}s \\ &-\int_T\dot{\varphi}(s)\int_{V_s}\rho\boldsymbol{\sigma}'\,\mathrm{d}\boldsymbol{B}_s'\,\mathrm{d}\boldsymbol{x},\end{aligned}$$

that can be equated to the source of momentum variations, as in Appendix B to obtain the Rotating Boussinesq equations. The acceleration term can be written in its relevant terms as:

$$\mathbb{D}_t^I\boldsymbol{v} = \mathbb{D}_t^R\boldsymbol{v}' + 2\boldsymbol{\Omega}\times\left(\boldsymbol{v}'\mathrm{d}t + \boldsymbol{\sigma}_t'\mathrm{d}\boldsymbol{B}_t'\right). \tag{B31}$$

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

If the stochastic model in flux form has to be ported to a global configuration, the effect of the change of variable has to be assessed. In particular, metrics terms can be computed with procedure only slightly deviating from the deterministic one. The displacements along $\boldsymbol{i}$ and $\boldsymbol{j}$ (the only two displacements that cause a change in the frame) are.

$$r\,\mathrm{d}\theta = \mathrm{dX}_j, \tag{B32}$$

$$\mathrm{d}\lambda\; r\cos\theta = \mathrm{dX}_i, \tag{B33}$$

where the first, inducing a variation in latitude $\mathrm{d}\theta$, is a negative rotation around the $\boldsymbol{i}$ axis, while the second is a rotation around the axis of the earth, as it induces a variation in longitude $\mathrm{d}\lambda$. This rotation around earth axis can be written as

$$\boldsymbol{\Omega}\,\mathrm{d}t = \Omega\,\mathrm{d}t(\boldsymbol{j}\cos\theta + \boldsymbol{k}\sin\theta), \tag{B34}$$

The infinitesimal displacement $\Omega\,\mathrm{d}t$ can be regarded as the rotation induced by the zonal motion, hence $\Omega\,\mathrm{d}t = \mathrm{dX}_i/r\cos\theta$ The variations of the vectors $\boldsymbol{i}, \boldsymbol{j}$ and $\boldsymbol{k}$ can then be defined as.

$$\mathrm{d}\boldsymbol{i} = \boldsymbol{\Omega}_f \mathrm{d}t \times \boldsymbol{i} = \frac{\mathrm{dX}_i \tan\theta}{r}\boldsymbol{j} + \frac{\mathrm{dX}_i}{r}\boldsymbol{k}, \tag{B35}$$

$$\mathrm{d}\boldsymbol{j} = \boldsymbol{\Omega}_f \mathrm{d}t \times \boldsymbol{j} = -\frac{\mathrm{dX}_i \tan\theta}{r}\boldsymbol{i} - \frac{\mathrm{dX}_j}{r}\boldsymbol{k}, \tag{B36}$$

$$\mathrm{d}\boldsymbol{k} = \boldsymbol{\Omega}_f \mathrm{d}t \times \boldsymbol{k} = \frac{\mathrm{dX}_i}{r}\boldsymbol{i} + \frac{\mathrm{dX}_j}{r}\boldsymbol{j}. \tag{B37}$$

Figure B1

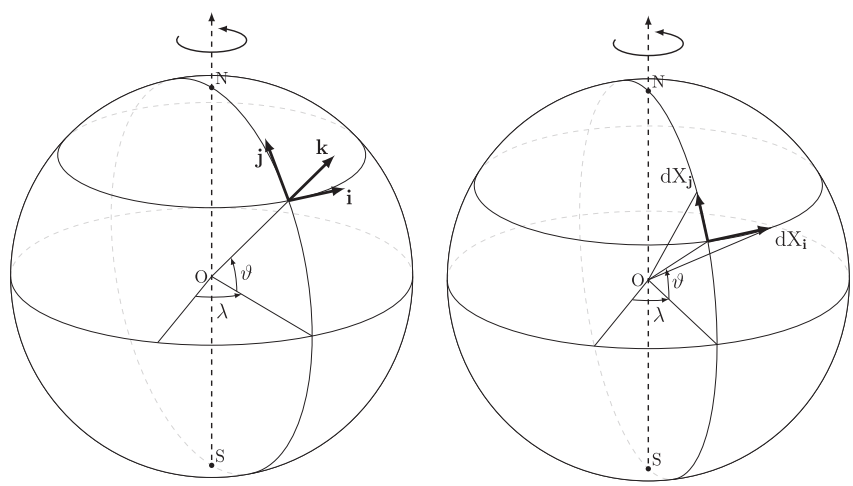


Figure B1. Representation of the geographical coordinate system and associated tangent coordinate system.

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Appendix C: Filtering Procedures Explained

In this work, we exploit data at different resolutions. To achieve this without incurring aliasing, several filters and grid–transition operators are defined. The downsampling operator is defined as

$$\boldsymbol{u}^{\downarrow_{\mathrm{Rn}}}(\boldsymbol{x}_i,t)=\int_{\mathbb{R}}\boldsymbol{u}(\boldsymbol{\varsigma},t)\delta(\boldsymbol{x}_i-\boldsymbol{\varsigma})\,\mathrm{d}\boldsymbol{\varsigma},\quad \forall\,\boldsymbol{x}_i\in\mathrm{Rn}. \tag{C1}$$

Its effect is to select the value of $\boldsymbol{u}$ at the location $\boldsymbol{x}_i$ on a predefined low–resolution grid Rn (i.e.., R3 or R9 in this work). The spatial filter is defined as

$$\overline{\boldsymbol{u}}^{\,c_x}(\boldsymbol{x},t)=\int_{\mathcal{S}(g_{c_x})}\boldsymbol{u}(\boldsymbol{\xi},t)\,g_{c_x}(\boldsymbol{x},\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}, \tag{C2}$$

where g_{c_x} is a spatial low–pass filter kernel of a given length scale c_x. The support of g_{c_x} is denoted by $\mathcal{S}(g_{c_x})$. Correspondingly, the spatial high–pass filter is given by

$$\tilde{\boldsymbol{u}}^{c_x}(\boldsymbol{x},t)=\left[\boldsymbol{u}-\overline{\boldsymbol{u}}^{\,c_x}\right](\boldsymbol{x},t)=\boldsymbol{u}(\boldsymbol{x},t)-\int_{\mathcal{S}(g_{c_x})}\boldsymbol{u}(\boldsymbol{\xi},t)\,g_{c_x}(\boldsymbol{x},\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}. \tag{C3}$$

The low–pass filter, used in combination with the downsampling operator is defined as

$$\overline{\boldsymbol{u}}^{\,c_x\downarrow_{\mathrm{Rn}}}(\boldsymbol{x}_i,t)=\int_{\mathbb{R}}\int_{\mathcal{S}(g_{c_x})}\boldsymbol{u}(\boldsymbol{\xi},t)\,g_{c_x}(\boldsymbol{\varsigma},\boldsymbol{\xi})\,\delta\left(\boldsymbol{x}_i-\boldsymbol{\varsigma}\right)\mathrm{d}\boldsymbol{\xi}\,\mathrm{d}\boldsymbol{\varsigma},\quad \forall\,\boldsymbol{x}_i\in\mathrm{Rn}, \tag{C4}$$

while the high–pass filter, down–sampled, is expressed as

$$\tilde{\boldsymbol{u}}^{c_x\downarrow_{\mathrm{Rn}}}(\boldsymbol{x}_i,t)=\overline{\boldsymbol{u}-\overline{\boldsymbol{u}}^{\,c_x}}^{\,c_x\downarrow_{\mathrm{Rn}}}=\int_{\mathbb{R}}\int_{\mathcal{S}(g_{c_x})}\left[\boldsymbol{u}-\overline{\boldsymbol{u}}^{\,c_x}\right](\boldsymbol{\xi},t)\,g_{c_x}(\boldsymbol{\varsigma},\boldsymbol{\xi})\delta(\boldsymbol{x}_i-\boldsymbol{\varsigma})\,\mathrm{d}\boldsymbol{\xi}\,\mathrm{d}\boldsymbol{\varsigma} \tag{C5}$$

for all $\boldsymbol{x}_i\in\mathrm{Rn}$. Notice that for a Gaussian low–pass filter, this filter corresponds to the application of Difference of Gaussians (DOG) together with a subsampling. The temporal filter is represented as

$$\overline{\boldsymbol{u}}^{\,b,\epsilon}(\boldsymbol{x},t)=\int_{\mathcal{S}(f)}\boldsymbol{u}(\boldsymbol{x},\tau)f_{\epsilon}(\tau,t)\,\mathrm{d}\tau, \tag{C6}$$

where f is a 5th order (low pass) Butterworth windowing function with cutoff frequency ϵ. The associated high–pass filter is

$$\tilde{\boldsymbol{u}}^{\,b,\epsilon}(\boldsymbol{x},t)=\left[\boldsymbol{u}-\overline{\boldsymbol{u}}^{\,b,\epsilon}\right](\boldsymbol{x},t)=\boldsymbol{u}(\boldsymbol{x},t)-\int_{\mathcal{S}(f)}\boldsymbol{u}(\boldsymbol{x},\tau)f_{\epsilon}(\tau,t)\,\mathrm{d}\tau. \tag{C7}$$

When associated to the spatial filter and downsampling, the two temporal filters reads

$$\overline{\overline{\boldsymbol{u}}^{\,b,\epsilon}}^{\,c_x\downarrow_{\mathrm{Rn}}}(\boldsymbol{x}_i,t)=\int_{\mathbb{R}}\int_{\mathcal{S}(g_{c_x})}\overline{\boldsymbol{u}}^{\,b,\epsilon}(\boldsymbol{\xi},t)\,g_{c_x}(\boldsymbol{\varsigma},\boldsymbol{\xi})\delta(\boldsymbol{x}_i-\boldsymbol{\varsigma})\,\mathrm{d}\boldsymbol{\xi}\,\mathrm{d}\boldsymbol{\varsigma},\quad \forall\,\boldsymbol{x}_i\in\mathrm{Rn}, \tag{C8}$$

and

$$\overline{\tilde{\boldsymbol{u}}^{\,b,\epsilon}}^{\,c_x\downarrow_{\mathrm{Rn}}}(\boldsymbol{x}_i,t)=\int_{\mathbb{R}}\int_{\mathcal{S}(g_{c_x})}\left[\boldsymbol{u}-\overline{\boldsymbol{u}}^{\,b,\epsilon}\right](\boldsymbol{\xi},t)\,g_{c_x}(\boldsymbol{\varsigma},\boldsymbol{\xi})\delta(\boldsymbol{x}_i-\boldsymbol{\varsigma})\,\mathrm{d}\tau\,\mathrm{d}\boldsymbol{\xi}\,\mathrm{d}\boldsymbol{\varsigma},\quad \forall\,\boldsymbol{x}_i\in\mathrm{Rn}. \tag{C9}$$

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Appendix D: Milstein–Type Scheme for Transport Noise

In this section, we describe the time–stepping scheme in the stochastic case, focusing on the noise advection and diffusion terms in the stochastic transport equation for a scalar quantity q (i.e., the resolved advection velocity, $\boldsymbol{u}^*$, is here assumed to be null without loss of generalization). In compact form, the equation is expressed as follows:

$$\mathrm{d}_t q = f(q,t)\,\mathrm{d}t + \sum_{i=1}^{m} g^i(q,t)\,\mathrm{d}\beta_t^i, \tag{D1}$$

$$g^i(q,t) = -\boldsymbol{\varphi}_t^i \cdot \nabla q, \tag{D2}$$

$$f(q,t) = \frac{1}{2}\nabla\cdot(\boldsymbol{a}_t \nabla q) = \frac{1}{2}\sum_{i=1}^{m} g^i \circ g^i(q,t), \tag{D3}$$

where $\left(\boldsymbol{\varphi}_t^i\right)_{i=1,\dots,m}$ represents a set of non–divergent $\left(\nabla\cdot\boldsymbol{\varphi}_t^i = 0\right)$ basis functions at time t, and $\left(\beta^i\right)_{i=1,\dots,m}$ is a sequence of independent standard Brownian motions. The last equality in (Equation D3) derives from the definition of the diffusion tensor, $\boldsymbol{a}_t = \sum_{i=1}^{m}\boldsymbol{\varphi}_t^i\left(\boldsymbol{\varphi}_t^i\right)^{\mathrm{T}}$, along with the divergence–free property of $\boldsymbol{\varphi}_t^i$. We now consider the semi–discrete form of this equation:

$$\mathrm{d}q_h(t) = f_h\left(q_h(t),t\right)\mathrm{d}t + \sum_{i=1}^{m} g_h^i\left(q_h(t),t\right)\mathrm{d}\beta_t^i. \tag{D4}$$

Here, the subscript h denotes quantities or operators that have been discretized in space. Applying the functional extension of the Itô formula (Cont & Fournie, 2010) to both f_h and g_h^i, we obtain:

$$\begin{aligned} q_h(t) = q_h(s) &+ \int_s^t f_h\left(q_h(r),r\right)\mathrm{d}r + \int_s^t \sum_{i=1}^{m} g_h^i\left(q_h(r),r\right)\mathrm{d}\beta_r^i \\ &+ \int_s^t \sum_{j=1}^{m} \dot{g}_h^j\left(\int_s^r \sum_{i=1}^{m} g_h^i\left(q_h(p),p\right)\mathrm{d}\beta_p^i, r\right)\mathrm{d}\beta_r^j + R, \end{aligned} \tag{D5}$$

where R denotes higher–order terms in the Milstein scheme, and $\dot{g}_h^i$ denotes the derivative of g_h^i with respect to function q_h. From Equation D2, we find that $\dot{g}_h^i = -\boldsymbol{\varphi}^i\cdot\nabla = g_h^i$. Consequently, we obtain the following first–order approximation of the above integral problem:

$$\begin{aligned} q_h^{n+1} = q_h^n &+ f_h\left(q_h^n,t_n\right)\Delta t + \sum_{i=1}^{m} g_h^i\left(q_h^n,t_n\right)\Delta\beta^i \\ &+ \sum_{i,j=1}^{m} g_h^j \circ g_h^i\left(q_h^n,t_n\right)\underbrace{\int_{t_n}^{t_{n+1}}\int_{t_n}^{r}\mathrm{d}\beta_p^i\mathrm{d}\beta_r^j}_{:=I_{ij}}, \end{aligned} \tag{D6}$$

where q^n is the value of q at time t_n, $\Delta t = t_{n+1} - t_n$ is the time step, and $\Delta\beta \sim \sqrt{\Delta t}\,\mathcal{N}(0,1)$ represents the independent Brownian motion increments. The double integral I_{ij} can be decomposed into a symmetric and an anti–symmetric part:

$$I_{ij} = \frac{1}{2}\underbrace{\left(I_{ij}+I_{ji}\right)}_{:=S_{ij}} + \frac{1}{2}\underbrace{\left(I_{ij}-I_{ji}\right)}_{:=A_{ij}}, \tag{D7}$$

where the symmetric component expands as follows:

$$
\begin{aligned}
S_{ij} &= \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} \mathrm{d}\beta_r^i \mathrm{d}\beta_s^j + \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} \mathrm{d}\beta_r^j \mathrm{d}\beta_s^i \\
&= \int_{t_n}^{t_{n+1}} \beta_s^i \, \mathrm{d}\beta_s^j + \int_{t_n}^{t_{n+1}} \beta_s^j \, \mathrm{d}\beta_s^i - \int_{t_n}^{t_{n+1}} \beta_{t_n}^i \, \mathrm{d}\beta_s^j - \int_{t_n}^{t_{n+1}} \beta_{t_n}^j \, \mathrm{d}\beta_s^i \\
&= \beta_{t_{n+1}}^i \beta_{t_{n+1}}^j - \beta_{t_n}^i \beta_{t_n}^j - \delta_{ij}\Delta t - \beta_{t_n}^i \beta_{t_{n+1}}^j + \beta_{t_n}^i \beta_{t_n}^j - \beta_{t_{n+1}}^i \beta_{t_n}^j + \beta_{t_n}^i \beta_{t_n}^j \\
&= \underbrace{\left(\beta_{t_{n+1}}^j - \beta_{t_n}^j\right)}_{\Delta\beta^j} \underbrace{\left(\beta_{t_{n+1}}^i - \beta_{t_n}^i\right)}_{\Delta\beta^i} - \delta_{ij}\Delta t
\end{aligned}
\tag{D8}
$$

using the Itô's–integration–by–part formula (Kunita, 1997) for two independent Brownian motions. Therefore, I_{ij} can be rewritten as:

$$
I_{ij} = \frac{1}{2}\left(\Delta\beta^i \Delta\beta^j - \delta_{ij}\Delta t + A_{ij}\right). \tag{D9}
$$

Substituting this into (Equation D6) and recalling (Equation D3), we derive the Milstein scheme for the semi–discrete problem Equation D4:

$$
q_h^{n+1} = q_h^n + \sum_{i=1}^{m} g_h^i\left(q_h^n, t_n\right) \Delta\beta^i + \frac{1}{2}\sum_{i,j=1}^{m} g_h^j \circ g_h^i\left(q_h^n, t_n\right)\left(\Delta\beta^j \Delta\beta^i + A_{ij}\right). \tag{D10}
$$

Accordingly, the diffusion term in the continuous formulation is precisely balanced by the diagonal component of the higher–order double integral. The anti–symmetric component, A_{ij}, known as the Lévy area process, is notably challenging to simulate. However, as numerically supported in recent studies Fiorini et al. (2023), Boulvard and Mémin (2023), for LU types dynamics, the Milstein scheme, without Lévy area, significantly reduces approximation error relative to the Euler scheme, though the strong convergence order remains the same. Consequently, we choose to disregard A_{ij} in practical applications. Additionally, as illustrated in Mémin et al. (2024), for linear waves, the proposed scheme enables to implement effectively the energy balance between the noise and the diffusion through an implicit discretization of the anisotropic diffusion operator involved (via the iterative advection term).

With the discrete noise term given by $\boldsymbol{\sigma}^n \Delta\boldsymbol{B} = \sum_{i=1}^{m} \boldsymbol{\varphi}_{t_n}^i \, \Delta\beta^i$, generated at each time step, the previous scheme can be expressed in a more compact form as follows:

$$
q_h^{n+1} = q_h^n - \left(\boldsymbol{\sigma}^n \Delta\boldsymbol{B} \cdot \nabla q^n\right)_h + \frac{1}{2}\left(\boldsymbol{\sigma}^n \Delta\boldsymbol{B} \cdot \nabla\left(\boldsymbol{\sigma}^n \Delta\boldsymbol{B} \cdot \nabla q^n\right)\right)_h. \tag{D11}
$$

Incorporating transport noise into the Leapfrog scheme used in the current version of NEMO, we propose the following multi–scale approach:

$$
\begin{aligned}
\tilde{q}_h &= q_h^{n-1} + F_h\left(q_h^n, t_n\right)\Delta t \\
q_h^{n+1} &= \tilde{q}_h + F_h\left(q_h^n, t_n\right)\Delta t - \left(\boldsymbol{\sigma}^n \Delta\boldsymbol{B} \cdot \nabla q^n\right)_h + \frac{1}{2}\left(\boldsymbol{\sigma}^n \Delta\boldsymbol{B} \cdot \nabla\left(\boldsymbol{\sigma}^n \Delta\boldsymbol{B} \cdot \nabla q^n\right)\right)_h.
\end{aligned}
\tag{D12}
$$

In this scheme, F gathers the classical resolved advection together with the advection of q by the Itô–Stokes drift $\boldsymbol{v}_s = \frac{1}{2}\nabla \cdot \boldsymbol{a}$ and the Girsanov drift $\boldsymbol{\sigma}_t \tilde{\boldsymbol{Y}}_t$ in the stochastic formulation. This approach is also adaptable to other schemes, such as Adams–Bashforth or Runge–Kutta, widely used in OGCMs, where transport noise would be included in the final correction.

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Data Availability Statement

The proposed implementation of the method is available at https://github.com/ftucciarone/LocationUncertainty (branch `JAMES-2024`). The frozen version of the code, complemented with the data necessary to reproduce the results presented, is publicly available on https://doi.org/10.5281/zenodo.15632926 (Tucciarone et al., 2025).

Acknowledgments
The authors gratefully acknowledge the three anonymous reviewers for their insightful comments and suggestions, which helped to significantly improve the manuscript, and the support of the ERC EU project 856408–STUOD, which enabled the research presented in this paper.

References

Andrews, D. G., & Mcintyre, M. E. (1978). An exact theory of nonlinear waves on a Lagrangian-mean flow. *Journal of Fluid Mechanics*, *89*(4), 609–646. https://doi.org/10.1017/s0022112078002773

Auclair, F., Estournel, C., Floor, J. W., Herrmann, M., Nguyen, C., & Marsaleix, P. (2011). A non-hydrostatic algorithm for free-surface ocean modelling. *Ocean Modelling*, *36*(1–2), 49–70. https://doi.org/10.1016/j.ocemod.2010.09.006

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

Bagaeva, E., Danilov, S., Oliver, M., & Juricke, S. (2024). Advancing eddy parameterizations: Dynamic energy backscatter and the role of subgrid energy advection and stochastic forcing. *Journal of Advances in Modeling Earth Systems*, *16*(4), e2023MS003972. https://doi.org/10.1029/2023ms003972

Bardina, J., Ferziger, J., & Reynolds, W. (1980). Improved subgrid-scale models for large-eddy simulation. In *13th fluid and plasmadynamics conference* (p. 1357).

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

Bauer, W., Chandramouli, P., Li, L., Mémin, E., & Mémin, E. (2020). 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

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

Berner, J., Achatz, U., Batté, L., Bengtsson, L., de la Cámara, A., Christensen, H. M., et al. (2017). Stochastic parameterization: Toward a new view of weather and climate models. *Bulletin of the American Meteorological Society*, *98*(3), 565–588. https://doi.org/10.1175/BAMS-D-15-00268.1

Borisenko, A. I., Tarapov, I. E., & Silverman, R. A. (1979). *Vector and tensor analysis with applications*. Dover Publications.

Bougeault, P., & Lacarrère, P. (1989). Parametrization of orography-induced turbulence in a mesobeta-scale model. *Monthly Weather Review*, *117*(8), 1872–1890.

Boulvard, P., & Mémin, E. (2023). Diagnostic of the Lévy area for geophysical flow models in view of defining high order stochastic discrete-time schemes. *Foundations of Data Science*, *6*(1), 1–21. https://doi.org/10.3934/fods.2023011

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

Brémaud, P. (2020). Probability theory and stochastic processes. *Universitext*, 421–437. https://doi.org/10.1201/9781003006626-25

Buizza, R., Miller, M., & Palmer, T. (1999). Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Quarterly Journal Royal Meteorological Society*, *125*(560), 2887–2908. https://doi.org/10.1256/smsqj.56005

Callies, J., Ferrari, R., Klymak, J. M., & Gula, J. (2015). Seasonality in submesoscale turbulence. *Nature Communications*, *6*(1), 6862. https://doi.org/10.1038/ncomms7862

Capet, X., McWilliams, J. C., Molemaker, M. J., & Shchepetkin, A. F. (2008). Mesoscale to submesoscale transition in the California current system. Part I: Flow structure, eddy flux, and observational tests. *Journal of Physical Oceanography*, *38*(1), 29–43. https://doi.org/10.1175/2007jpo3671.1

Chandramouli, P., Mémin, E., & Heitz, D. (2020). 4D large scale variational data assimilation of a turbulent flow with a dynamics error model. *Journal of Computational Physics*, *412*, 109446. https://doi.org/10.1016/j.jcp.2020.109446

Chandramouli, P., Mémin, E., Heitz, D., & Laizet, S. (2018). Coarse large-eddy simulations in a transitional wake flow with flow models under location uncertainty. *Computers and Fluids*, *168*, 170–189. https://doi.org/10.1016/j.compfluid.2018.04.001

Chapron, B., Dérian, P., Mémin, E., & Resseguier, V. (2018). Large-scale flows under location uncertainty: A consistent stochastic framework. *Quarterly Journal of the Royal Meteorological Society*, *144*(710), 251–260. https://doi.org/10.1002/qj.3198

Charney, J. G. (1951). Dynamic forecasting by numerical process. *Compendium of Meteorology: Prepared under the Direction of the Committee on the Compendium of Meteorology*, 470–482. https://doi.org/10.1007/978-1-940033-70-9_40

Cintolesi, C., & Mémin, E. (2020). Stochastic modelling of turbulent flows for numerical simulations. *Fluid*, *5*(3), 108. https://doi.org/10.3390/fluids5030108

Cont, R., & Fournie, D. (2010). A functional extension of the Ito formula. *Comptes Rendus Mathematique*, *348*(1), 57–61. https://doi.org/10.1016/j.crma.2009.11.013

Craik, A., & Leibovich, S. (1976). 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.

Debussche, A., Hug, B., & Mémin, E. (2023). A consistent stochastic large-scale representation of the Navier–Stokes equations. *Journal of Mathematical Fluid Mechanics*, *25*(1), 19. https://doi.org/10.1007/s00021-023-00764-0

Debussche, A., & Mémin, E. (2024). Variational principles for fully coupled stochastic fluid dynamics across scales. *arXiv*.

Debussche, A., Mémin, E., & Moneyron, A. (2025). *Some properties of a non-hydrostatic stochastic oceanic primitive equations model* (Vol. III, pp. 161–182). Springer International Publishing. https://doi.org/10.1007/978-3-031-70660-8_8

Fiorini, C., Boulvard, P., & Li, L., & Mémin. (2023). A two-step numerical scheme in time for surface quasi geostrophic equations under location uncertainty. In *Stochastic transport in upper ocean dynamics* (pp. 57–67). Springer International Publishing.

Fox-Kemper, B., Adcroft, A., Böning, C. W., Chassignet, E. P., Curchitser, E., Danabasoglu, G., et al. (2019). Challenges and prospects in ocean circulation models. *Frontiers in Marine Science*, *6*, 65. https://doi.org/10.3389/fmars.2019.00065

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

Gottwald, G., Crommelin, D., & Franzke, C. (2017). Stochastic climate theory. In *Nonlinear and stochastic climate dynamics* (pp. 209–240). Cambridge University Press.

Griffies, S. M., & Treguier, A.-M. (2013). *Ocean circulation models and modeling*. (Vol. 103, pp. 521–551). Academic Press. https://doi.org/10.1016/b978-0-12-391851-2.00020-9

Grooms, I., Majda, A., & Smith, S. (2014). 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

Grooms, I., & Majda, A. J. (2013). Efficient stochastic superparameterization for geophysical turbulence. *Proceedings of the National Academy of Sciences* (Vol. 110(12), 4464–4469). https://doi.org/10.1073/pnas.1302548110

Hannachi, A., Jolliffe, I. T., & Stephenson, D. B. (2007). Empirical orthogonal functions and related techniques in atmospheric science: A review. *International Journal of Climatology*, *27*(9), 1119–1152. https://doi.org/10.1002/joc.1499

Harouna, S. K., & Mémin, E. (2017). Stochastic representation of the Reynolds transport theorem: Revisiting large-scale modeling. *Computers and Fluids*, *156*, 456–469. https://doi.org/10.1016/j.compfluid.2017.08.017

Hasselmann, K. (1976). Stochastic climate models Part I. Theory. *Tellus*, *28*(6), 473–485. https://doi.org/10.3402/tellusa.v28i6.11316

Holm, D. D. (2015). Variational principles for stochastic fluid dynamics. *Proceedings of the royal society A: Mathematical, physical and engineering sciences* (Vol. 471(2176), 20140963). https://doi.org/10.1098/rspa.2014.0963

Holmes, P., Lumley, J. L., & Berkooz, G. (1996). *Turbulence, coherent structures, dynamical systems, and symmetry*. Cambridge University Press.

Hurlburt, H. E., & Hogan, P. J. (2000). Impact of 1/8° to 1/64° resolution on Gulf Stream model–data comparisons in basin-scale subtropical Atlantic Ocean models. *Dynamics of Atmospheres and Oceans*, *32*(3–4), 283–329. https://doi.org/10.1016/s0377-0265(00)00050-6

Ilıcak, M., Adcroft, A. J., Griffies, S. M., & Hallberg, R. W. (2012). Spurious dianeutral mixing and the role of momentum closure. *Ocean Modelling*, *45–46*, 0–58. https://doi.org/10.1016/j.ocemod.2011.10.003

Jamet, Q., Mémin, E., Dumas, F., Li, L., & Garreau, P. (2024). Toward a stochastic parameterization for oceanic deep convection. In *Stochastic transport in upper ocean dynamics II* (pp. 143–157). Springer International Publishing.

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

Kitsios, V., Frederiksen, J. S., & O'Kane, T. J. (2023). Subgrid parameterization of eddy, meanfield and topographic interactions in simulations of an idealized Antarctic Circumpolar Current. *Journal of Advances in Modeling Earth Systems*, *15*(5), e2022MS003412. https://doi.org/10.1029/2022MS003412

Klingbeil, K., & Burchard, H. (2013). Implementation of a direct nonhydrostatic pressure gradient discretisation into a layered ocean model. *Ocean Modelling*, *65*, 64–77. https://doi.org/10.1016/j.ocemod.2013.02.002

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

Kutz, J. N., Brunton, S. L., Brunton, B. W., & Proctor, J. L. (2016). *Dynamic mode decomposition*. Society for Industrial and Applied Mathematics.

Leith, C. (1990). Stochastic backscatter in a subgrid-scale model: Plane shear mixing layer. *Physics of Fluids*, *2*(3), 1521–1530. https://doi.org/10.1063/1.857779

Lemarié, F., Debreu, L., Madec, G., Demange, J., Molines, J., & Honnorat, M. (2015). Stability constraints for oceanic numerical models: Implications for the formulation of time and space discretizations. *Ocean Modelling*, *92*, 124–148. https://doi.org/10.1016/j.ocemod.2015.06.006

Leutbecher, M., Lock, S.-J., Ollinaho, P., Lang, S. T. K., Balsamo, G., Bechtold, P., et al. (2017). Stochastic representations of model uncertainties at ECMWF: State of the art and future vision. *Quarterly Journal of the Royal Meteorological Society*, *143*(707), 2315–2339. https://doi.org/10.1002/qj.3094

Lévy, M., Klein, P., Tréguier, A.-M., Iovino, D., Madec, G., Masson, S., & Takahashi, K. (2010). Modifications of gyre circulation by sub-mesoscale physics. *Ocean Modelling*, *34*(1–2), 1–15. https://doi.org/10.1016/j.ocemod.2010.04.001

Lévy, M., Resplandy, L., Klein, P., Capet, X., Iovino, D., & Ethé, C. (2012). Grid degradation of submesoscale resolving ocean models: Benefits for offline passive tracer transport. *Ocean Modelling*, *48*, 1–9. https://doi.org/10.1016/j.ocemod.2012.02.004

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*(2), e2022MS003297. https://doi.org/10.1029/2022ms003297

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

Lilly, D. (1992). A proposed modification of the Germano subgrid-scale closure. *Physics of Fluids*, *3*, 2746–2757.

Lucarini, V., & Chekroun, M. D. (2023). Theoretical tools for understanding the climate crisis from Hasselmann's programme and beyond. *Nature Reviews Physics*, *12*, 744–765. https://doi.org/10.1038/s42254-023-00650-8

Madec, G., Bourdallé-Badie, R., Chanut, J., Clementi, E., Coward, A., Ethé, C., et al. (2019). *Nemo ocean engine*. Zenodo. (Add SI3 and TOP reference manuals). https://doi.org/10.5281/zenodo.3878122

Majda, A., Timofeyev, I., & Eijnden, E. V. (1999). Models for stochastic climate prediction. *PNAS*, *96*(26), 14687–14691. https://doi.org/10.1073/pnas.96.26.14687

Maltrud, M. E., & McClean, J. L. (2005). An eddy resolving global 1/10° ocean simulation. *Ocean Modelling*, *8*(1–2), 31–54. https://doi.org/10.1016/j.ocemod.2003.12.001

Marsaleix, P., Michaud, H., & Estournel, C. (2019). 3D phase-resolved wave modelling with a non-hydrostatic ocean circulation model. *Ocean Modelling*, *136*, 28–50. https://doi.org/10.1016/j.ocemod.2019.02.002

Mason, P., & Thomson, D. (1992). Stochastic backscatter in large-eddy simulations of boundary layers. *Journal of Fluid Mechanics*, *242*, 51–78. https://doi.org/10.1017/s0022112092002271

McWilliams, J., Sullivan, P., & Moeng, C.-H. (1997). Langmuir turbulence in the ocean. *Journal of Fluid Mechanics*, *334*, 1–30. https://doi.org/10.1017/s0022112096004375

McWilliams, J. C. (2016). Submesoscale currents in the ocean. *Proceedings of the royal society A: Mathematical, physical and engineering sciences* (Vol. 472(2189), 1–32). https://doi.org/10.1098/rspa.2016.0117

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

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Mémin, E., Li, L., Lahaye, N., Tissot, G., & Chapron, B. (2024). Linear wave solutions of a stochastic shallow water model. In *Stochastic transport in upper ocean dynamics ii* (pp. 223–245). Springer International Publishing.

O'Kane, T. J., Fiedler, R., Collier, M. A., & Kitsios, V. (2023). Ocean model response to stochastically perturbed momentum fluxes. *Journal of Climate*, *36*(6), 1895–1922. https://doi.org/10.1175/jcli-d-21-0796.1

Pinier, B., Mémin, E., Laizet, S., & Lewandowski, R. (2019). Stochastic flow approach to model the mean velocity profile of wall-bounded flows. *Physical Review E*, *99*(6), 063101. https://doi.org/10.1103/physreve.99.063101

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. (2021). New trends in ensemble forecast strategy: Uncertainty quantification for coarse-grid computational fluid dynamics. *Archives of Computational Methods in Engineering*, *28*(1), 215–261. https://doi.org/10.1007/s11831-020-09437-x

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

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

Resseguier, V., Mémin, E., Heitz, D., & Chapron, B. (2017). Stochastic modelling and diffusion modes for proper orthogonal decomposition models and small-scale flow analysis. *Journal of Fluid Mechanics*, *826*, 888–917. https://doi.org/10.1017/jfm.2017.467

Sanderson, B. G. (1998). Order and resolution for computational ocean dynamics. *Journal of Physical Oceanography*, *28*(6), 1271–1286. https://doi.org/10.1175/1520-0485(1998)028<1271:oarfco>2.0.co;2

Sasaki, Y. N., Minobe, S., Schneider, N., Kagimoto, T., Nonaka, M., & Sasaki, H. (2008). Decadal sea level variability in the South Pacific in a global eddy-resolving ocean model hindcast. *Journal of Physical Oceanography*, *38*(8), 1731–1747. https://doi.org/10.1175/2007jpo3915.1

Shchepetkin, A. F., & McWilliams, J. C. (2009). Computational kernel algorithms for fine-scale, multiprocess, longtime oceanic simulations. In *Handbook of Numerical Analysis* (Vol. 14, pp. 121–183). Elsevier. https://doi.org/10.1016/s1570-8659(08)01202-0

Shutts, G. (2005). A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Quarterly Journal of the Royal Meteorological Society*, *612*, 3079–3102. https://doi.org/10.1256/qj.04.106

Siegel, A., Weiss, J. B., Toomre, J., McWilliams, J. C., Berloff, P. S., & Yavneh, I. (2001). Eddies and vortices in ocean basin dynamics. *Geophysical Research Letters*, *28*(16), 3183–3186. https://doi.org/10.1029/1999gl011246

Smagorinsky, J. (1963). General circulation experiments with the primitive equation: I. The basic experiment. *Monthly Weather Review*, *91*, 99–165.

Soufflet, Y., Marchesiello, P., Lemarié, F., Jouanno, J., Capet, X., Debreu, L., & Benshila, R. (2016). On effective resolution in ocean models. *Ocean Modelling*, *98*, 36–50. https://doi.org/10.1016/j.ocemod.2015.12.004

Storto, A., & Andriopoulos, P. (2021). A new stochastic ocean physics package and its application to hybrid-covariance data assimilation. *Quarterly Journal of the Royal Meteorological Society*, *147*(736), 1691–1725. https://doi.org/10.1002/qj.3990

Thuburn, J., Kent, J., & Wood, N. (2014). Cascades, backscatter and conservation in numerical models of two-dimensional turbulence. *Quarterly Journal of the Royal Meteorological Society*, *140*(679), 626–638. https://doi.org/10.1002/qj.2166

Tissot, G., Mémin, E., & Jamet, Q. (2024). Stochastic compressible Navier–Stokes equations under location uncertainty. In *Stochastic transport in upper ocean dynamics II* (pp. 239–319). Springer International Publishing.

Tucciarone, F. L., Li, L., Mémin, E., & Chandramouli, P. (2025). Derivation and assessment of a stochastic hydrostatic primitive equations model [Dataset]. *Zenodo*. https://doi.org/10.5281/zenodo.15632927

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

Tucciarone, F. L., Mémin, E., & Li, L. (2024). Data driven stochastic primitive equations with dynamic modes decomposition. In *Stochastic transport in upper ocean dynamics II* (pp. 321–336). Springer International Publishing.

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

19422466, 2025, 7, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024MS004783 by Test, Wiley Online Library on [13/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License