Check for updates

JAMES | Journal of Advances in Modeling Earth Systems

RESEARCH ARTICLE
10.1029/2021MS002492

Rotating Shallow Water Flow Under Location Uncertainty With a Structure-Preserving Discretization

Rüdiger Brecht[1], Long Li[2], Werner Bauer[3], and Etienne Mémin[2]

[1]Department of Mathematics and Statistics, Memorial University of Newfoundland, St. John's, NL, Canada, [2]Inria/IRMAR, Campus universitaire de Beaulieu, Rennes, France, [3]Department of Mathematics, Imperial College London, London, UK

Key Points:

- A physically relevant stochastic parametrization of the shallow water model is introduced
- The proposed stochastic model conserves the total energy and motivates a structure preserving discretization
- This stochastic parametrization provides a good trade-off between model error representation and ensemble spread

Correspondence to:
R. Brecht,
rbrecht@mun.ca

Citation:
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*, e2021MS002492. https://doi.org/10.1029/2021MS002492

Received 4 FEB 2021
Accepted 11 OCT 2021

© 2021 The Authors. 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 We introduce a physically relevant stochastic representation of the rotating shallow water equations. The derivation relies mainly on a stochastic transport principle and on a decomposition of the fluid flow into a large-scale component and a noise term that models the unresolved flow components. As for the classical (deterministic) system, this scheme, referred to as modeling under location uncertainty (LU), conserves the global energy of any realization and provides the possibility to generate an ensemble of physically relevant random simulations with a good trade-off between the model error representation and the ensemble's spread. To maintain numerically the energy conservation feature, we combine an energy (in space) preserving discretization of the underlying deterministic model with approximations of the stochastic terms that are based on standard finite volume/difference operators. The LU derivation, built from the very same conservation principles as the usual geophysical models, together with the numerical scheme proposed can be directly used in existing dynamical cores of global numerical weather prediction models. The capabilities of the proposed framework is demonstrated for an inviscid test case on the f-plane and for a barotropically unstable jet on the sphere.

Plain Language Summary The motion of geophysical fluids on the globe needs to be modeled to get insights of tomorrow's weather. These forecasts must be precise enough while remaining computationally affordable. Ideally they should enable to estimate likely scenarios through an ensemble of physically relevant realizations, built from an accurate handling of the model errors that are inescapably introduced due to physical or numerical approximations. To address these issues, we advocate the use of a stochastic framework to represent the action of the many unresolved fast/small-scale processes on the resolved flow component. The derivation of the stochastic system, based on the usual conservation laws, is presented in detail and simulated with an adapted structure preserving numerical model to maintain numerically the nice properties of the stochastic setting inherited from a transport principle, namely: mass and energy conservation. The versatile nature of the stochastic derivation as well as of the proposed numerical scheme makes this framework suitable for existing dynamical cores of global numerical weather prediction models. Numerical results illustrate the energy conservation of the numerical model and the accuracy of large-scale stochastic simulations when compared to corresponding deterministic ones. The ability of the random dynamical system to represent model errors is also shown.

1. Introduction

Numerical simulations of the Earth's atmosphere and ocean play an important role in developing our understanding of weather forecasting. A major focus lies in determining the large-scale flow correctly, which is strongly related to the parameterizations of sub-grid processes (Frederiksen et al., 2013). The non-linear and non-local nature of the dynamics of geophysical fluid flows make the large-scale flow structures interact with the smaller components. Solving the Kolmogorov scales (Pope, 2000) of geophysical flows is today, and likely for a foreseeable future, completely out of reach. This is due, in the first place, to the formidable computational expense that would be necessary, but also to the complexity of the many fine-scale physical or bio-chemical processes involved. Truncating the fine scales and simply ignoring their actions is highly detrimental to a reliable simulation of the large-scale components of the flow. Yet, an accurate modeling of the fine-scale processes' effects is an excruciatingly difficult task and the idea of a stochastic modeling has strongly attracted the geophysical community since the seminal works of Hasselmann, 1976; and Leith, 1975. For several years, this interest has been strongly strengthened with the emergence of ensemble methods for probabilistic forecasting and data assimilation

19422466, 2021, 12, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021MS002492 by Cochrane France, Wiley Online Library on [25/10/2023]. 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

issues (Berner & Coauthors, 2017; Franzke et al., 2015; Gottwald et al., 2017; Majda et al., 2008; Palmer & Williams, 2008; Slingo & Palmer, 2011).

The schemes proposed so far rely on very different methodological concepts. Multiplicative random forcing and randomization of parameters based on early turbulence studies on energy backscattering (Leith, 1990; Mason & Thomson, 1992) have been proposed (Buizza et al., 1999; Porta, Mana & Zanna, 2014; Shutts, 2005). The ad hoc nature of these schemes makes a systematic stochastic derivation of any flow dynamical model or configuration difficult. In addition, the absence of an explicit energy balance of the noise term leads to an uncontrolled increase of variance that is potentially problematic. They consequently require a proper tuning of the large-scale subgrid model and of the noise amplitude to stabilize the system. The subgrid model is, however, not related to the noise term and the amplitude of the perturbations to apply is also difficult to specify on physical grounds. More importantly, even for low noise, an arbitrary random perturbation defined outside of the physical principles on which the system has been built upon may lead to strongly erroneous probability density functions of the system's dynamics (Chapron et al., 2018). Other schemes based on an averaging and homogenization theory have been proposed (Franzke et al., 2006; Franzke & Majda, 2006) in the wake of Majda et al. (1999) and extended through the Mori-Zwanzig formalism (see the review Gottwald et al., 2017 and references therein). Those techniques are well suited for the design of stochastic reduced order systems.

In this study, we propose to stick to a specific stochastic model, called modeling under *Location Uncertainty* (LU) derived by Mémin (2014), which emerges from a decomposition of the Lagrangian velocity into a smooth-in-time drift and a highly oscillating random term. Such a slow/fast or smooth/oscillating decomposition is reminiscent to the Lagrangian decomposition introduced in the seminal work of Andrews & McIntyre (1978), which is currently used for surface or internal waves studies (Kafiabad et al., 2021; Salmon, 2013; Xie & Vanneste, 2015; Young & Jelloul, 1997). A similar random decomposition is also at the center of the variational stochastic framework of Holm (2015). Like our setting this latter approach applies in a broader context and not only to wave solutions. Both frameworks rely on a stochastic transport principle, with (Holm, 2015) dedicated to Hamiltonian dynamical systems and defined from a circulation preserving constrained variational formulation, while Mémin, (2014) is general and built upon classical physical conservation laws.

This stochastic transport principle has been used as a fundamental tool to derive stochastic representations of large-scale geophysical dynamics (Bauer, Chandramouli, Chapron, et al., 2020; Bauer, Chandramouli, Li, et al., 2020; Chapron et al., 2018; Resseguier et al., 2017a, 2017b, 2017c) or to define large eddy simulation models of turbulent flows (Chandramouli et al., 2020; Kadri Harouna & Mémin, 2017). The LU framework relies on a stochastic representation of the Reynolds transport theorem (Kadri Harouna & Mémin, 2017; Mémin, 2014) which introduces naturally meaningful terms for turbulence studies.

It gathers a multiplicative random advection which is responsible for an energy backscattering, a subgrid diffusion operator describing the mixing of the large-scale flow component by the small-scale random component, and an effective advection which is attached to the small scales spatial inhomogeneity. This latter term has been shown to be reminiscent of a generalized Stokes drift component, hence designated as Itô-Stokes drift (Bauer, Chandramouli, Chapron, et al., 2020). Backscattering and diffusion are energetically in balance which leads hence to global energy conservation.

Recently, the LU formulation was shown to perform very well for oceanic quasi-geostrophic flow models (Bauer, Chandramouli, Chapron, et al., 2020; Bauer, Chandramouli, Li, et al., 2020; Resseguier et al., 2017b; 2017a). It was found to be more accurate in predicting the extreme events, in diagnosing the frontogenesis and filamentogenesis, in structuring the large-scale flow and in reproducing long-terms statistics. Besides, for a LU version of the Lorentz-63 model, derived from a Rayleigh-Bénard convection -in the very same way as the original model (Berge et al., 1987; Lorenz, 1963), it has been demonstrated that the LU setting was more effective in exploring the range of the strange attractor compared to classical models as well as to stochastic models built with ad hoc multiplicative forcings (Chapron et al., 2018).

In this work, the performance of the LU representation is assessed for the numerical simulation of the rotating shallow water (RSW) system, which can be considered as the first step toward developing global random numerical weather prediction and climate models. In particular, this is the first time that the LU formulation is implemented for the dynamics evolving on the sphere. The global energy conservation of the RSW-LU system for any realization, which is analytically demonstrated here, is a strong asset of the approach and this invariant feature

19422466, 2021, 12, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021MS002492 by Cochrane France, Wiley Online Library on [25/10/2023]. 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

should be numerically conserved as closely as possible. Global energy conservation is especially important for long-term climatic simulations. However, classical purely damping parameterizations do not take into account energy and momentum fluxes from the unresolved to the resolved scales. In climatic models, this is believed to be a source of important biases (Gugole & Franzke, 2019).

Hence, we propose to combine the discrete variational integrator for RSW fluids as introduced in Bauer & Gay-Balmaz (2019a) and Brecht et al. (2019) with the numerical LU setting in order to maintain this conservation property as well as all the transport invariants. The benefit of the proposed method that relies on a modular combination of a variational integrator with a (potentially different) discretization of the LU formulation is that it should be directly applicable to existing dynamical cores of numerical weather prediction models.

The derivation of the variational integrator is based on the variational discretization framework introduced by Pavlov et al. (2011) for incompressible fluids, expanded by Gawlik et al. (2011) to incompressible fluids with advected quantities. In various papers, this framework has been further extended, for instance Desbrun et al. (2014) incorporated rotating and stratified fluids of atmospheric and oceanic dynamics and Bauer and Gay-Balmaz, (2019b) introduced soundproof approximations of the Euler equations. Variational integrators are designed by first discretizing the given Lagrangian, and then by deriving a discrete system of associated Euler-Lagrange equations from the discretized Lagrangian (see, Marsden & West, 2001).

The advantage of this approach is that the resulting discrete system inherits several important properties of the underlying continuous system, notably a discrete version of Noether's theorem that guarantees the preservation of conserved quantities associated to the symmetries of the discrete Lagrangian (see, Hairer et al., 2006). Variational integrators also exhibit superior long-term stability properties, (cf. e.g., Leimkuhler & Reich, 2004). Therefore, they typically outperform traditional integrators if one is interested in long-time integration or the statistical properties of a given dynamical system. Our choice for an energy preserving rather than an enstrophy conserving scheme is based on the following considerations. As shown in Bauer, Chandramouli, Li, et al. (2020) for stochastic barotropic quasi-geostrophic models, using an energy conserving scheme for long-term predictions yields better results than using an enstrophy conserving one. Besides, because of the direct cascade of enstrophy to high wave numbers, often stabilization through enstrophy dissipation is introduced, even in initially enstrophy conserving schemes, (cf. Bonaventura & Ringler, 2005; McRae & Cotter, 2014; Ringler & Randall, 2002).

Apart from taking into account the unresolved processes, it is paramount in probabilistic ensemble forecasting to model the uncertainties along time (Resseguier et al., 2020). In particular, operational ensemble data assimilation methods rely classically on random perturbations of the initial conditions (PIC) together with an artificially carefully inflated variance (Anderson & Anderson, 1999) to increase the otherwise deficient ensemble forecasts' spread (Franzke et al., 2015; Gottwald & Harlim, 2013). Such inflation has the side effect of augmenting also the representation error of the ensemble members. In the present work, we compare the reliability of the ensemble spread of such a PIC model with our RSW-LU system, under the same noise amplitude, and show that the LU strategy yields a good trade-off between model error representation and ensemble spread.

The remainder of this paper is structured as follows. Section 2 describes the basic principles of the derivation of the rotating shallow water system in the LU formulation. Section 3 explains the numerical discretization of the stochastic dynamical system. Section 4 discusses the numerical results for an inviscid test case with homogeneous noise and a viscous test case with heterogeneous noise. In Section 5 we draw some conclusions and provide an outlook for future work. In the Appendices we demonstrate the energy conservation of the RSW–LU system, review some parameterizations of the noise and describe the discretization of the stochastic terms.

2. Rotating Shallow Water Equations Under Location Uncertainty

In this section, we first review the LU representation introduced by Mémin (2014), then we derive the rotating shallow water equations under LU, denoted as RSW–LU, following the classical strategy (Vallis, 2017). In particular, we demonstrate one important characteristic of the RSW–LU, namely that it preserves the total energy of the large-scale flow.

2.1. Location Uncertainty Principles

The LU formulation is based on a temporal-scale-separation assumption of the following stochastic flow:

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

where $\boldsymbol{X}$ is the Lagrangian displacement defined within the bounded domain $\Omega \subset \mathbb{R}^d$ $(d = 2 \text{ or } 3)$, $\boldsymbol{w}$ is the large-scale velocity that is both spatially and temporally correlated, and $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$ is a highly oscillating unresolved component (also called noise) term that is only correlated in space. The spatial structure of such noise is specified through a deterministic integral operator $\boldsymbol{\sigma}:(L^2(\Omega))^d \to (L^2(\Omega))^d$, acting on square integrable vector-valued functions $f \in (L^2(\Omega))^d$, with a bounded kernel $\check{\sigma}$ such that

$$\boldsymbol{\sigma}\,[\boldsymbol{f}](\boldsymbol{x}, t) = \int_\Omega \check{\boldsymbol{\sigma}}(\boldsymbol{x}, \boldsymbol{y}, t)\boldsymbol{f}(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}, \quad \forall\, \boldsymbol{f} \in (L^2(\Omega))^d. \tag{2}$$

The randomness of such a noise is driven by a functional Brownian motion $\boldsymbol{B}_t$ (Da Prato & Zabczyk, 2014). The fact that the kernel is bounded, implies that the resulting random flow $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$ is a centered (of null ensemble mean) Gaussian process with the well-defined *covariance tensor*:

$$\begin{aligned}\boldsymbol{Q}(\boldsymbol{x}, \boldsymbol{y}, t, s) &= \mathbb{E}\Big[(\boldsymbol{\sigma}(\boldsymbol{x}, t)\,\mathrm{d}\mathbf{B}_t)(\boldsymbol{\sigma}(\boldsymbol{y}, s)\,\mathrm{d}\boldsymbol{B}_s)^T\Big] \\ &= \delta(t - s)\,\mathrm{d}t \int_\Omega \check{\boldsymbol{\sigma}}(\boldsymbol{x}, \boldsymbol{z}, t)\check{\boldsymbol{\sigma}}^T(\boldsymbol{y}, \boldsymbol{z}, s)\,\mathrm{d}\boldsymbol{z},\end{aligned} \tag{3}$$

where $\mathbb{E}$ stands for the expectation, δ is the Kronecker symbol and $\bullet^T$ denotes matrix or vector transpose. The strength of the noise is measured by its *variance*, denoted here as $\boldsymbol{a}$, and which is given by the diagonal components of the covariance per unit of time:

$$\boldsymbol{a}(\boldsymbol{x}, t)\mathrm{d}t = \boldsymbol{Q}(\boldsymbol{x}, \boldsymbol{x}, t, t). \tag{4}$$

We remark that this variance tensor has the same unit as a diffusion tensor ($\mathrm{m}^2 \cdot \mathrm{s}^{-1}$) and that the density of the turbulent kinetic energy (TKE) can be specified through it by $\frac{1}{2}\mathrm{tr}(\boldsymbol{a})/\mathrm{d}t$.

The previous representation (Equation 2) is a general way to define the noise, but other formulations can be conveniently used in practice. In particular, the covariance operator per unit of time, $\boldsymbol{Q}/\mathrm{d}t$, admits an orthogonal eigenfunction basis $\{\Phi_n(\bullet, t)\}_{n\in\mathbb{N}}$ weighted by the eigenvalues $\Lambda_n \geq 0$ such that $\sum_{n\in\mathbb{N}} \Lambda_n < \infty$. Therefore, one may equivalently define the noise and its variance, based on the following spectral decomposition:

$$\boldsymbol{\sigma}(\boldsymbol{x}, t)\;\mathrm{d}\boldsymbol{B}_t = \sum_{n\in\mathbb{N}} \boldsymbol{\Phi}_n(\boldsymbol{x}, t)\;\mathrm{d}\beta_t^n, \qquad \boldsymbol{a}(\boldsymbol{x}, t) = \sum_{n\in\mathbb{N}} \boldsymbol{\Phi}_n(\boldsymbol{x}, t)\boldsymbol{\Phi}_n^T(\boldsymbol{x}, t), \tag{5}$$

where β^n denotes n independent and identically distributed (i.i.d.) one-dimensional standard Brownian motions. The specification of those basis functions from data driven empirical covariance matrices enables one to construct specific noises, informed either by numerical or observational data. This strategy will allow us to devise various forms of the noise in the following.

Remark 1 Decomposition (Equation 1) is a temporal decomposition and not a spatial decomposition as classically formulated through spatial filters and/or decimation operators in large-eddies simulation (LES) techniques. However, in the case of turbulent flows, time and spatial scales are related. As a matter of fact, in the inertial range, the turn-over time ratio for two different scales L and ℓ reads $\tau_L/\tau_\ell \propto (L/\ell)^{2/3}$ and provides a direct relation between time-scale coarsening and spatial-scale dilation. Unless specifically needed, in the following, we will thus refer to large/small or unresolved scales without differentiating between time or space scales. Note also that temporal filtering has already been used for the definition of oceanic models (Hecht et al., 2008) or large-eddies simulation approaches (Meneveau & Katz, 2000).

Remark 2 Decomposition (Equation 1) is written in terms of an Itô stochastic integral. This decomposition could have been written in the form of a Stratonovich integral as well. The calculus associated to this latter integral has the advantage of following the classical chain rule. However, the Stratonovich noise no longer has zero expectation. This leads thus to a problematic decomposition with velocity fluctuations of non null ensemble mean. For smooth enough integrands, it is possible to safely move from one form to the other. For interested readers, more insights on the difference of the two settings and their implications in stochastic oceanic modeling are provided in Bauer, Chandramouli, Chapron, et al. (2020).

19422466, 2021, 12, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021MS002492 by Cochrane France, Wiley Online Library on [25/10/2023]. 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, 2021, 12, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021MS002492 by Cochrane France, Wiley Online Library on [25/10/2023]. 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

Remark 3 The approach could be extended to express flows on arbitrary Riemannian manifolds. In that case it is easier to work directly with the Stratonovich formulation since it is invariant under the change of coordinates. As we consider here only flows that assume the shallow approximation, the considered representation of the equations in $\mathbb{R}^2$ and $\mathbb{R}^3$ is a very accurate approximation.

The core of the LU model representation is based on a stochastic Reynolds transport theorem (SRTT), introduced by Mémin (2014), which describes the rate of change of a random scalar q transported by the stochastic flow (Equation 1) within a flow volume $\mathcal{V}$. In particular, for incompressible unresolved flows, $\boldsymbol{\nabla}\cdot\boldsymbol{\sigma}=0$, the SRTT can be written as

$$\mathrm{d}_t\left(\int_{\mathcal{V}(t)} q(\boldsymbol{x},t)\ \mathrm{d}\boldsymbol{x}\right)=\int_{\mathcal{V}(t)}\left(\mathbb{D}_t q+q\boldsymbol{\nabla}\cdot(\boldsymbol{w}-\boldsymbol{w}_s)\right)\mathrm{d}\boldsymbol{x}, \tag{6a}$$

$$\mathbb{D}_t q=\mathrm{d}_t q+(\boldsymbol{w}-\boldsymbol{w}_s)\cdot\boldsymbol{\nabla}q\,\mathrm{d}t+\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}q-\frac{1}{2}\boldsymbol{\nabla}\cdot(\boldsymbol{a}\boldsymbol{\nabla}q)\,\mathrm{d}t, \tag{6b}$$

where $\mathrm{d}_t q(\boldsymbol{x},t)=q(\boldsymbol{x},t+\mathrm{d}t)-q(\boldsymbol{x},t)$ stands for the forward time-increment of q at a fixed point $\boldsymbol{x}$, $\mathbb{D}_t$ is introduced as the stochastic transport operator in Resseguier et al. (2017c) and $\boldsymbol{w}_s=\frac{1}{2}\boldsymbol{\nabla}\cdot\boldsymbol{a}$ is referred to as the Itô-Stokes drift (ISD) in Bauer, Chandramouli, Chapron, et al. (2020). The transport operator plays the role of the material derivative in the stochastic setting. The ISD is defined by the variance tensor divergence and embodies the effect of statistical inhomogeneity of the unresolved flow on the large-scale component. As shown in Bauer, Chandramouli, Chapron, et al. (2020), it can be considered as a generalization of the Stokes drift associated to waves propagation with the emergence of a similar vortex force and Coriolis correction. In the definition of the stochastic transport operator in Equation 6b, the last two terms describe, respectively, an energy backscattering from the unresolved scales to the large scales and an inhomogeneous diffusion of the large scales driven by the variance of the unresolved flow components. The diffusion term generalizes the Boussinesq eddy viscosity assumption (here with a matrix eddy viscosity). This term is, nevertheless, directly related to the noise form and not anymore defined by loose analogy with the molecular dissipation mechanism. The backscattering term corresponds to an energy source that is exactly compensated by the diffusion term (Resseguier et al., 2017c).

In particular, for an isochoric flow with $\boldsymbol{\nabla}\cdot(\boldsymbol{w}-\boldsymbol{w}_s)=0$, one may immediately deduce from Equation 6a the following transport equation of an extensive scalar:

$$\mathbb{D}_t q=0, \tag{7}$$

where the energy of such random scalar q is globally conserved, as shown in Resseguier et al. (2017c):

$$\mathrm{d}_t\left(\int_\Omega \frac{1}{2}q^2\ \mathrm{d}\boldsymbol{x}\right)=\underbrace{\left(\frac{1}{2}\int_\Omega q\boldsymbol{\nabla}\cdot(\boldsymbol{a}\boldsymbol{\nabla}q)\,\mathrm{d}\boldsymbol{x}\right.}_{\text{Energy loss by diffusion}}+\underbrace{\left.\frac{1}{2}\int_\Omega(\boldsymbol{\nabla}q)^T\boldsymbol{a}\boldsymbol{\nabla}q\,\mathrm{d}\boldsymbol{x}\right)}_{\text{Energy intake by noise}}\mathrm{d}t=0. \tag{8}$$

Indeed, this can be interpreted as a process where the energy brought by the noise is exactly counterbalanced by that dissipated from the diffusion term.

2.2. Derivation of RSW–LU

This section describes in detail the derivation of the RSW–LU system. This model enriches the formulation described in Mémin, (2014). Here it is fully stochastic and includes rotation to suit simulations of geophysical flows on a rotating frame.

The above SRTT (Equation 6a) and Newton's second principle allow us to derive the following (three-dimensional) stochastic equations of motions in a rotating frame (Bauer, Chandramouli, Chapron, et al., 2020):

19422466, 2021, 12, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021MS002492 by Cochrane France, Wiley Online Library on [25/10/2023]. 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

$$\textit{Horizontal momentum equation}: \tag{9a}$$
$$\mathbb{D}_t \boldsymbol{u} + \boldsymbol{f} \times (\boldsymbol{u}\ \mathrm{d}t + \boldsymbol{\sigma}_H \mathrm{d}\boldsymbol{B}_t) = -\frac{1}{\rho}\boldsymbol{\nabla}_H \left(p\ \mathrm{d}t + \mathrm{d}p_t^\sigma\right) + \nu\boldsymbol{\nabla}^2(\boldsymbol{u}\ \mathrm{d}t + \boldsymbol{\sigma}_H \mathrm{d}\boldsymbol{B}_t),$$

$$\textit{Vertical momentum equation}: \tag{9b}$$
$$\mathbb{D}_t w = -\frac{1}{\rho}\partial_z \left(p\ \mathrm{d}t + \mathrm{d}p_t^\sigma\right) - g\ \mathrm{d}t + \nu\nabla^2(w\ \mathrm{d}t + \sigma_z \mathrm{d}B_t),$$

$$\textit{Mass equation}: \tag{9c}$$
$$\mathbb{D}_t \rho = 0,$$

$$\textit{Continuity equations}: \tag{9d}$$
$$\boldsymbol{\nabla}_H \cdot (\boldsymbol{u} - \boldsymbol{u}_s) + \partial_z(w - w_s) = 0, \qquad \boldsymbol{\nabla}_H \cdot \boldsymbol{\sigma}_H \mathrm{d}\boldsymbol{B}_t + \partial_z \sigma_z \mathrm{d}B_t = 0,$$

where $\boldsymbol{u} = (u, v)^T$ (resp. $\boldsymbol{\sigma}_H \mathrm{d}\boldsymbol{B}_t$) and w (resp. $\sigma_z \mathrm{d}B_t$) are the horizontal and vertical components of the three-dimensional large-scale flow $\overrightarrow{w}$ (resp. the unresolved random flow $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$); $\boldsymbol{f} = (2\tilde{\Omega}\sin\Theta)\boldsymbol{k}$ is the Coriolis parameter varying in latitude Θ, with the Earth's angular rotation rate $\tilde{\Omega}$ and the vertical unit vector $\boldsymbol{k} = [0, 0, 1]^T$; ρ is the fluid density; $\nabla_H = [\partial_x, \partial_y]^T$ denotes the horizontal gradient; p and $\dot{p}_t^\sigma = \mathrm{d}p_t^\sigma/\mathrm{d}t$ (informal definition) are the time-smooth and time-uncorrelated components of the pressure field, respectively; g is the Earth's gravity value and ν is the kinematic viscosity. In the following, the molecular friction term is assumed to be negligible and dropped from the equations. Note that in our setting the continuity Equation 9d ensure volume conservation (Resseguier et al., 2017c) and mass conservation (Equation 9c).

In order to model the large-scale circulations in the atmosphere and ocean, the hydrostatic balance approximation is widely adopted (Vallis, 2017). We now specify the scaling for this balance in the LU framework. We first adimensionalize the basic variables as

$$(x, y) = \mathcal{L}\ (x', y'), \qquad \boldsymbol{u} = \mathcal{U}\ \boldsymbol{u}', \qquad t = \mathcal{T}\ t',\ \mathcal{T} = \mathcal{L}/\mathcal{U}, \qquad z = \alpha\mathcal{L}z',\ \alpha = \mathcal{H}/\mathcal{L}, \tag{10}$$

where the capital letters are used for the characteristic scales of variables and $\bullet'$ denotes adimensional variables. To scale properly the vertical velocity, we propose to adopt a sufficient incompressible condition (Resseguier et al., 2017b, 2017c) for the resolved component in Equation 9d, that is

$$\nabla_H \cdot \boldsymbol{u} + \partial_z w = 0, \qquad \nabla_H \cdot \boldsymbol{u}_s + \partial_z w_s = 0. \tag{11}$$

Note that the latter divergence-free condition on the ISD is usually considered for the classical Stokes drift (McWilliams et al., 2004) although being controversial (Mellor, 2016). The three-dimensional bolus velocity introduced in the eddy-induced-advection parametrization (Gent et al., 1995; Gent & McWilliams, 1990; Griffies, 1998) is also assumed to be incompressible in order to preserve the tracer's moments. In our case, the justification of this constraint is further strengthen by global energy conservation and a desirable bridge between the classical (global energy conserving) rotating shallow water system and its stochastic representation. Under the condition (Equation 11), a classical scaling of the vertical (resolved) velocity holds:

$$w = \alpha\ \mathcal{U}\ w'. \tag{12}$$

Apart from these classical scaling numbers, the horizontal component $\boldsymbol{a}_H$ of the variance/diffusion tensor $\boldsymbol{a}$, which characterizes the strength of the unresolved component, is scaled as

$$\boldsymbol{a}_H = \epsilon\ \mathcal{U}\mathcal{L}\ \boldsymbol{a}'_H, \qquad \boldsymbol{a} = \begin{pmatrix} \boldsymbol{a}_H & \boldsymbol{a}_{Hz} \\ \boldsymbol{a}_{Hz} & a_z \end{pmatrix}, \qquad \epsilon = \frac{\mathcal{T}_\sigma}{\mathcal{T}}\frac{\mathrm{EKE}}{\mathrm{MKE}}, \tag{13}$$

where the specific factor ϵ (Resseguier et al., 2017b) is defined as the ratio between the eddy kinetic energy (EKE) and the mean kinetic energy (MKE), multiplied by the ratio between the unresolved scale correlation time $\mathcal{T}_\sigma$ and the large-scale advection time. From the definitions (Equations 3 and 4), the scaling of the horizontal small-scale flow reduces to

$$\boldsymbol{\sigma}_H \mathrm{d}\boldsymbol{B}_t = \sqrt{\epsilon}\ \ \mathcal{L}\ \ (\boldsymbol{\sigma}_H \mathrm{d}\boldsymbol{B}_t)'. \tag{14}$$

In addition, we consider the following scaling between the vertical and horizontal components of the unresolved flow:

$$\frac{\sigma_z \mathrm{d}B_t}{\|\boldsymbol{\sigma}_H \mathrm{d}\boldsymbol{B}_t\|} \sim \alpha\ \ \delta, \qquad \text{i.e. } \sigma_z \mathrm{d}B_t = \sqrt{\epsilon}\ \ \delta\ \ \mathcal{H}\ \ (\sigma_z \mathrm{d}B_t)', \tag{15}$$

where δ is a small factor (Resseguier et al., 2017b). Again, from the definitions (Equations 3 and 4), the other components of the variance/diffusion tensor scale then as:

$$\boldsymbol{a}_{Hz} = \epsilon\ \ \delta\ \ \mathcal{U}\mathcal{H}\ \ \boldsymbol{a}'_{Hz}, \qquad a_z = \epsilon\ \ \delta^2\ \ \alpha\ \ \mathcal{U}\mathcal{H}\ \ a'_z, \qquad \text{i.e. } \frac{a_z}{\|\boldsymbol{a}_H\|} \sim \alpha^2\delta^2. \tag{16}$$

This relation provides a ratio between the vertical and horizontal eddy diffusivities. It is in practice quite small at large scale (Lévy et al., 2010, 2012).

Now, with $\boldsymbol{f} = 0$ and a constant density ρ_0, the horizontal momentum Equation 9a implies the following scalings of the rescaled pressures:

$$\tilde{p} = p/\rho_0 = \mathcal{U}^2\ \ \tilde{p}', \qquad \mathrm{d}\tilde{p}_t^\sigma = \mathrm{d}p_t^\sigma/\rho_0 = \sqrt{\epsilon}\ \ \mathcal{U}\mathcal{L}\ \ (\mathrm{d}\tilde{p}_t^\sigma)'. \tag{17}$$

Finally, substituting all the above scalings into Equation 9b, the adimensional vertical momentum is given by

$$\begin{aligned}
\alpha^2 \Big[& \mathrm{d}_t w' + (\boldsymbol{u}' \cdot \boldsymbol{\nabla}'_H w' + w' \partial'_z w')\, \mathrm{d}t' + \sqrt{\epsilon} \left((\boldsymbol{\sigma}_H \mathrm{d}\boldsymbol{B}_t)' \cdot \boldsymbol{\nabla}'_H w' + \delta\ (\sigma_z \mathrm{d}B_t)' \partial'_z w' \right) \\
& - \frac{\epsilon}{2} \Big((\boldsymbol{\nabla}'_H \cdot \boldsymbol{a}'_H + \delta\ \partial'_z \boldsymbol{a}'_{Hz}) \cdot \boldsymbol{\nabla}'_H w' + \delta\ \ (\boldsymbol{\nabla}'_H \cdot \boldsymbol{a}'_{Hz} + \delta\ \partial'_z a'_z) \partial'_z w' \\
& + \boldsymbol{\nabla}'_H \cdot (\boldsymbol{a}'_H \boldsymbol{\nabla}'_H w' + \delta\ \boldsymbol{a}'_{Hz} \partial'_z w') + \delta\ \ \partial'_z (\boldsymbol{a}'_{Hz} \boldsymbol{\nabla}'_H w' + \delta\ a'_z \partial'_z w') \Big)\ \ \mathrm{d}t' \Big] \\
= & -\partial'_z \left(\tilde{p}'\ \ \mathrm{d}t' + \sqrt{\epsilon}\ \ (\mathrm{d}\tilde{p}_t^\sigma)' \right) - \mathrm{d}t'/\mathrm{Fr}^2,
\end{aligned} \tag{18}$$

where $\mathrm{Fr} = \mathcal{U}/\sqrt{g\mathcal{H}}$ is the *Froude number*. Let us now make the following assumptions:

$$\alpha^2 \ll 1, \qquad \mathrm{Fr}^2 = \mathcal{O}(1), \qquad \epsilon = \mathcal{O}(1), \qquad \delta \ll 1. \tag{19}$$

The acceleration term on the left-hand side (LHS) of Equation 9b has now a lower order of magnitude than the RHS terms. Restoring the dimensions, the hydrostatic balance under moderate horizontal uncertainty and weak vertical uncertainty hence boils down to

$$\partial_z \left(p\ \ \mathrm{d}t + \mathrm{d}p_t^\sigma \right) = -\rho g\ \ \mathrm{d}t, \qquad \text{i.e. } \partial_z p = -\rho g,\ \ \partial_z \mathrm{d}p_t^\sigma = 0. \tag{20a}$$

We remark that the unique decomposition principle of a semimartingale process (Kunita, 1997) is used here to separate the bounded variation component (in terms of $\mathrm{d}t$) and the martingale part (in terms of $\mathrm{d}\boldsymbol{B}_t$ or $\mathrm{d}p_t^\sigma$). Integrating vertically these hydrostatic balances (Equation 20a) from 0 to z (see Figure 1), we have

$$p(x, y, z, t) = p_0(x, y, t) - \rho_0 g z, \qquad \mathrm{d}p_t^\sigma(x, y, z, t) = \mathrm{d}p_t^\sigma(x, y, 0, t), \tag{20b}$$

where p_0 denotes the pressure at the bottom of the basin ($z = 0$). Following (Vallis, 2017), we assume that the weight of the overlying fluid is negligible, that is $p(x, y, \eta, t) \approx 0$ with η the height of the free surface, leading to $p_0 = \rho_0 g \eta$. This allows us to rewrite Equation 20b such that for any $z \in [0, \eta]$ we have

$$p(x, y, z, t) = \rho_0 g\, (\eta(x, y, t) - z)\,. \tag{20c}$$

Subsequently, the pressure gradient force in the horizontal momentum Equation 9a reads

19422466, 2021, 12, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021MS002492 by Cochrane France, Wiley Online Library on [25/10/2023]. 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

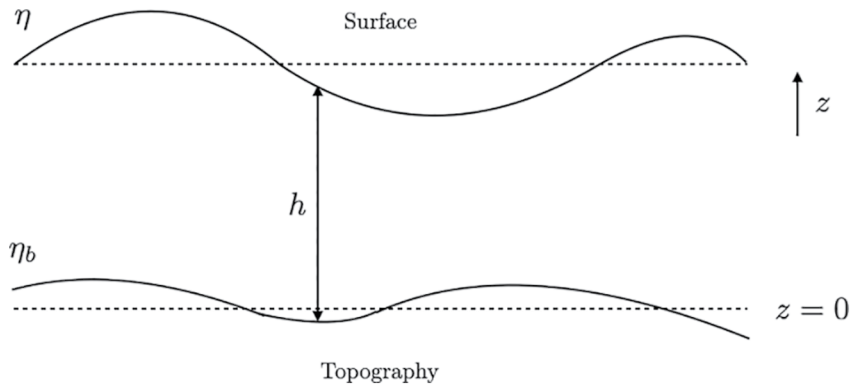


Figure 1. Illustration of a single-layered shallow water system (inspired by Vallis, 2017). h is the thickness of a water column, η is the height of the free surface and η_b is the height of the bottom topography. As a result, we have $h = \eta - \eta_b$.

$$-\frac{1}{\rho_0}\boldsymbol{\nabla}_H\left(p\ \mathrm{d}t + \mathrm{d}p_t^\sigma\right) = -g\boldsymbol{\nabla}_H\eta\,\mathrm{d}t\ \ -\frac{1}{\rho_0}\boldsymbol{\nabla}_H\mathrm{d}p_t^\sigma, \tag{20d}$$

which does not depend on z according to Equations 20b and 20c. Therefore, the acceleration terms on the LHS of Equation 9a cannot depend on z, and the shallow water momentum equation under weak vertical uncertainty ($\delta \ll 1$) can be written finally as

$$\mathbb{D}_t^H\boldsymbol{u} + \boldsymbol{f}\times\left(\boldsymbol{u}\ \mathrm{d}t + \boldsymbol{\sigma}_H\mathrm{d}\boldsymbol{B}_t\right) = -g\boldsymbol{\nabla}_H\eta\ \mathrm{d}t - \frac{1}{\rho_0}\boldsymbol{\nabla}_H\mathrm{d}p_t^\sigma, \tag{21a}$$

$$\mathbb{D}_t^H u = \mathrm{d}_t u + \left((\boldsymbol{u}-\boldsymbol{u}_s)\ \mathrm{d}t + \boldsymbol{\sigma}_H\mathrm{d}\boldsymbol{B}_t\right)\cdot\boldsymbol{\nabla}_H u - \frac{1}{2}\boldsymbol{\nabla}_H\cdot\left(\boldsymbol{a}_H\boldsymbol{\nabla}_H u\right)\ \mathrm{d}t, \tag{21b}$$

where $\boldsymbol{u}_s = \frac{1}{2}\boldsymbol{\nabla}_H\cdot\boldsymbol{a}_H$ is the two-dimensional ISD and $\mathbb{D}_t^H$ denotes the horizontal stochastic transport operator whose expression is recalled in Equation 21b for the u component. The relation between the unresolved flow component and the random pressure can be further specified by considering a scaling of the martingale part of the momentum equation:

$$\sqrt{\epsilon}\ \mathrm{d}_t\tilde{u}' + \sqrt{\epsilon}\ (\boldsymbol{\sigma}_H\mathrm{d}\boldsymbol{B}_t)'\cdot\boldsymbol{\nabla}_H' u' + \frac{\sqrt{\epsilon}}{\mathrm{R_o}}\ \boldsymbol{f}'\times(\boldsymbol{\sigma}_H\mathrm{d}\boldsymbol{B}_t)' = \sqrt{\epsilon}\ \boldsymbol{\nabla}_H'(\mathrm{d}p_t^\sigma)', \tag{22}$$

where $\mathrm{R_o} = \mathcal{U}/(f_0\mathcal{L})$ denotes the Rossby number with $\boldsymbol{f} = f_0\boldsymbol{f}'$, and $\tilde{u} = u - \mathbb{E}(u)$ stands for the martingale part of the horizontal velocity. We note that the scaling $\mathrm{d}_t\tilde{u} = \sqrt{\epsilon}\ \mathcal{U}\ \mathrm{d}_t\tilde{u}'$ is obtained from the variance of the martingale part of the vertical acceleration term (Equation 18) considering the hydrostatic balance (Equation 20a) and the continuity Equation 11. Therefore, for small Rossby number ($\mathrm{R_o} \leq 1$), the random Coriolis term counter-balances the random gradient pressure force:

$$\boldsymbol{f}\times\boldsymbol{\sigma}_H\mathrm{d}\boldsymbol{B}_t \approx -\frac{1}{\rho_0}\boldsymbol{\nabla}_H\mathrm{d}p_t^\sigma. \tag{23}$$

Besides, under weak vertical uncertainty, the dimensional continuity (Equations 11 and 9d) reduce to

$$\boldsymbol{\nabla}_H\cdot\boldsymbol{\sigma}_H\mathrm{d}\boldsymbol{B}_t = \boldsymbol{\nabla}_H\cdot\boldsymbol{u}_s = 0. \tag{24}$$

As a result, the vertical integration (from bottom topography η_b to free surface η) of the continuity Equation 9d become

$$(w-w_s)|_{z=\eta} - (w-w_s)|_{z=\eta_b} = -h\boldsymbol{\nabla}_H\cdot\boldsymbol{u}, \qquad \sigma\mathrm{d}B_t|_{z=\eta} - \sigma\mathrm{d}B_t|_{z=\eta_b} = 0, \tag{25a}$$

19422466, 2021, 12, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021MS002492 by Cochrane France, Wiley Online Library on [25/10/2023]. 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, 2021, 12, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021MS002492 by Cochrane France, Wiley Online Library on [25/10/2023]. 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 $h = \eta - \eta_b$ denotes the thickness of the water column (with a still bottom). On the other hand, a small vertical (Eulerian) displacement at the top and bottom of the fluid leads to a variation of the position of a particular fluid element (Vallis, 2017):

$$((w - w_s)\ \mathrm{d}t + \sigma \mathrm{d}\boldsymbol{B}_t)|_{z=\eta} = \mathbb{D}_t^H \eta, \qquad ((w - w_s)\ \mathrm{d}t + \sigma \mathrm{d}\boldsymbol{B}_t)|_{z=\eta_b} = \mathbb{D}_t^H \eta_b. \tag{25b}$$

Combining Equations 25a and 25b, we deduce the following stochastic mass equation:

$$\mathbb{D}_t^H h + h\boldsymbol{\nabla}_H \cdot \boldsymbol{u}\ \mathrm{d}t = 0. \tag{26}$$

Gathering all the elements derived so-far, we finally obtain the following RSW-LU system

$$(Conservation\ \ of\ \ momentum) \tag{27a}$$
$$\mathbb{D}_t \boldsymbol{u} + \boldsymbol{f} \times \boldsymbol{u}\ \ \mathrm{d}t = -g\boldsymbol{\nabla}\eta\ \mathrm{d}t,$$

$$(Conservation\ \ of\ \ mass) \tag{27b}$$
$$\boldsymbol{D}_t h + h\boldsymbol{\nabla} \cdot \boldsymbol{u}\ \mathrm{d}t = 0,$$

$$(Random\ \ balance) \tag{27c}$$
$$\boldsymbol{f} \times \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t = -\frac{1}{\rho}\boldsymbol{\nabla}\mathrm{d}p_t^{\sigma},$$

$$(Incompressible\ \ constraints) \tag{27d}$$
$$\boldsymbol{\nabla} \cdot \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t = 0, \qquad \boldsymbol{\nabla} \cdot \mathbf{u}_s = 0,$$

where the symbol H for all horizontal variables are dropped for readability reasons. In Appendix A it is shown that this stochastic system conserves the global energy:

$$\mathrm{d}_t \int_\Omega \frac{\rho}{2}\left(h|\boldsymbol{u}|^2 + gh^2\right)\ \mathrm{d}\boldsymbol{x} = 0. \tag{28}$$

It shares thus exactly the same energy conservation property as the deterministic one and beyond their formal resemblance this provides a strong physical link between the two systems. Moreover, it can be noticed that under a sufficiently weak (horizontal) uncertainty ($\boldsymbol{\sigma} \approx 0$), the system (Equations 27a, 27b, 27c, and 27d) reduces to the classical RSW system, in which the stochastic transport operator weighted by the unit of time, $\mathbb{D}_t/\mathrm{d}t$, reduces to the material derivative.

3. Structure-Preserving Discretization of RSW–LU

In order to perform numerical simulations of the RSW–LU (Equations 27a, 27b, 27c, and 27d) the noise term $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$ has to be a priori parametrized. Its shape is conveniently expressed through a spectral representation and a set of basis functions (Equation 5). In this work homogeneous as well as heterogeneous spatial structures have been used and the way they are defined is reviewed in Appendix B. The incompressible homogenous noise (see Appendix B1) is defined through a convolution kernel and is associated with Fourier modes orthogonal functions. It is easy to implement through fast Fourier transform (FFT). As shown in Section 4.1, this noise was in particular used to assess the numerical energy behavior of the discrete scheme. However, homogeneous noises, although carefully scaled from a known energy spectrum established at high resolution, fail to represent inhomogeneity effect encoded by spatially varying variance (the variance is constant and diagonal for homogeneous incompressible noise). This is detrimental to represent large scale effects shaped by the small-scale components in geophysical fluid dynamics. As a matter of fact as shown in Bauer, Chandramouli, Chapron, et al. (2020), heterogeneous noise shapes the large-scale flow in a way akin to the action of vortex force associated with the classical Stokes drift.

19422466, 2021, 12, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021MS002492 by Cochrane France, Wiley Online Library on [25/10/2023]. 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 this work, two different parameterizations of heterogeneous noise have been used and are described in Appendix B2. The former consists in calibrating empirical orthogonal basis functions (EOF) before the simulation (off-line) from available high-resolution simulation data while the latter consists in specifying the basis functions from the on-going (low resolution) simulation (i.e., on-line). The second basis functions do not depend on data and are time evolving whereas the first ones are data driven and stationary. A procedure based on dynamic mode decomposition (Schmid, 2010) to define the noise through evolving basis functions could have been as well used, as proposed by Gugole & Franzke (2019). Such a time evolving basis, learned from a high resolution simulation, are shown to perform better that stationary EOF based models. We will have the same type of conclusions for the non-stationary noise experimented here. In Section 4.2, both heterogeneous noises are adopted for identifying the barotropic instability of a mid-latitude jet.

In the following, we focus on an energy conserving (in space) approximation of the random dynamical system (RSW–LU). In this context, the spatial discretization allows us to mimic the balance between the global energy brought by the noise and the LU-diffusion (see Equation 8) at each time step, hence no additional numerical dissipation or energy increase is introduced into the system. Considering the definition of the stochastic transport operator $\mathbb{D}_t$ in Equation 6b, the RSW–LU system in Equations 27a and 27b can be explicitly written as

$$\mathrm{d}_t \boldsymbol{u} = (-\boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{u} - \boldsymbol{f}\times\boldsymbol{u} - g\boldsymbol{\nabla}\eta)\ \mathrm{d}t + \left(\frac{1}{2}\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\cdot(\boldsymbol{a}\boldsymbol{u})\ \mathrm{d}t - \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}\boldsymbol{u}\right), \tag{29a}$$

$$\mathrm{d}_t h = -\boldsymbol{\nabla}\cdot(\boldsymbol{u}h)\ \mathrm{d}t + \left(\frac{1}{2}\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\cdot(\boldsymbol{a}h)\ \mathrm{d}t - \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}h\right). \tag{29b}$$

We suggest to develop an approximation of the stochastic RSW–LU model (Equations 29a and 29b) by first discretizing the deterministic model underlying this system with a structure-preserving discretization method (that preserves energy in space) and, then, to approximate (with a potentially different discretization method) the stochastic terms. Here, we use for the former a variational discretization approach on a triangular C–grid while for the latter we apply a standard finite difference method. Note that for the methodology introduced in this manuscript, other spatially energy conserving discretizations rather than the suggested variational integrator could be used too. The *deterministic dynamical core* of our stochastic system results from simply setting $\boldsymbol{\sigma} \approx 0$ in the Equations 29a and 29b. To obtain the full discretized (in space and time) scheme for this stochastic system, we wrap the discrete stochastic terms around the deterministic core and combine this with an Euler–Marayama time scheme.

Introducing discretizations of the stochastic terms that do not necessarily share the same operators as the deterministic scheme has various advantages, as discussed in more detail in Section 3.2.1. For instance, such a well defined interface between these two model components minimizes the necessity to adapt the discretization schemes to each other which, in turn, would permit us to apply our method immediately to existing dynamical cores of global numerical weather prediction (NWP) models.

3.1. Discretization of Deterministic RSW Equations

As mentioned above, the deterministic model (or deterministic dynamical core) of the above stochastic system results from setting $\boldsymbol{\sigma} \approx 0$, which leads via (Equation 4) to $\boldsymbol{a} \approx 0$. Hence, Equations 29a and 29b reduce to the deterministic RSW equations

$$\mathrm{d}_t \boldsymbol{u} = \left(-(\boldsymbol{\nabla}\times\boldsymbol{u} + \boldsymbol{f})\times\boldsymbol{u} - \boldsymbol{\nabla}\left(\frac{1}{2}\boldsymbol{u}^2\right) - g\boldsymbol{\nabla}\eta\right)\ \mathrm{d}t, \qquad \mathrm{d}_t h = -\boldsymbol{\nabla}\cdot(\boldsymbol{u}h)\ \mathrm{d}t, \tag{30}$$

where we used the vector calculus identity $\boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{u} = (\boldsymbol{\nabla}\times\boldsymbol{u})\times\boldsymbol{u} + \frac{1}{2}\boldsymbol{u}^2$. Note that in the deterministic case $\mathrm{d}_t/\mathrm{d}t$ agrees (in the limit $\mathrm{d}t \to 0$) with the partial derivative $\partial/\partial t$.

3.1.1. Variational Discretizations

In the following we present an energy conserving (in space) approximation of these equations using a variational discretization approach. While details about the derivation can be found in Bauer & Gay-Balmaz (2019a); Brecht et al. (2019), here we only give the final, fully discrete scheme.

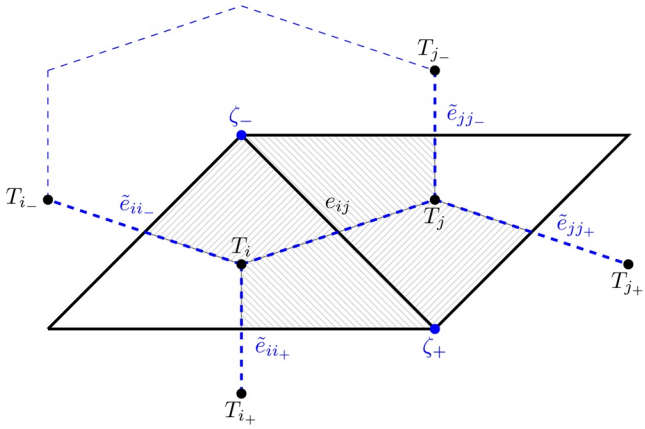


Figure 2. Notation and indexing conventions for the 2D simplicial mesh.

To do so, we start with introducing the mesh and some notation. The variational discretization of Equation 30 results in a scheme that corresponds to a C-grid staggering of the variables on a quasi uniform triangular grid with hexagonal/pentagonal dual mesh. Let N denote the number of triangles used to discretize the domain. As shown in Figure 2, we use the following notation: T denotes the primal triangle, ζ the dual hexagon/pentagon, $e_{ij} = T_i \cap T_j$ the primal edge and $\tilde{e}_{ij} = \zeta_+ \cap \zeta_-$ the associated dual edge. Furthermore, we have $\mathbf{n}_{ij}$ and $\mathbf{t}_{ij}$ as the normalized normal and tangential vector relative to edge e_{ij} at its midpoint. Moreover, D_i is the discrete water depth at the circumcenter of T_i, η_{b_i} the discrete bottom topography at the circumcenter of T_i, and $V_{ij} = (\mathbf{u} \cdot \mathbf{n})_{ij}$ the normal velocity at the triangle edge midpoints in the direction from triangle T_i to T_j. We denote $\bar{D}_{ij} = \frac{1}{2}(D_i + D_j)$ as the water depth averaged to the edge midpoints.

The variational discretization method does not require to define explicitly approximations of the differential operators because they directly result from the discrete variational principle. It turns out that on the given mesh, these operators agree with the following definitions of standard finite difference and finite volume operators:

$$\begin{aligned} (\mathrm{Grad}_n\ F)_{ij} &\triangleq \frac{F_{T_j} - F_{T_i}}{|\tilde{e}_{ij}|}, & (\mathrm{Div}\ V)_i &\triangleq \frac{1}{|T_i|} \sum_{k \in \{j, i_-, i_+\}} |e_{ik}| V_{ik}, \\ (\mathrm{Grad}_t\ F)_{ij} &\triangleq \frac{F_{\zeta_-} - F_{\zeta_+}}{|e_{ij}|}, & (\mathrm{Curl}\ V)_\zeta &\triangleq \frac{1}{|\zeta|} \sum_{\tilde{e}_{nm} \in \partial\zeta} |\tilde{e}_{nm}| V_{nm}, \end{aligned} \tag{31}$$

for the normal velocity V_{ij} and a scalar function F either sampled as F_{T_i} at the circumcenter of the triangle T_i or sampled as $F_{\zeta_\pm}$ at the center of the dual cell $\zeta_\pm$. The operators Grad_n and Grad_t correspond to the gradient in the normal and tangential direction, respectively, and Div to the divergence of a vector field:

$$(\nabla F)_{ij} \approx (\mathrm{Grad}_n\ F)\mathbf{n}_{ij} + (\mathrm{Grad}_t\ F)\mathbf{t}_{ij}, \tag{32}$$

$$(\nabla \cdot \mathbf{u})_i \approx (\mathrm{Div}\ V)_i, \tag{33}$$

$$(\nabla \times \mathbf{u})_\zeta \approx (\mathrm{Curl}\ V)_\zeta. \tag{34}$$

The last equation defines the discrete vorticity and for later use, we also discretize the potential vorticity as

$$\frac{\nabla \times \mathbf{u} + f}{h} \approx \frac{(\mathrm{Curl}\ V)_\zeta + f_\zeta}{D_\zeta}, \qquad D_\zeta = \sum_{\tilde{e}_{ij} \in \partial\zeta} \frac{|\zeta \cap T_i|}{|\zeta|} D_i. \tag{35}$$

3.1.2. Semi-Discrete RSW Scheme

With the above notation, the deterministic semi-discrete RSW equations read:

$$\mathrm{d}_t V_{ij} = \mathcal{L}^V_{ij}(V, D)\ \Delta t, \quad \text{for all edges } e_{ij}, \tag{36a}$$

$$\mathrm{d}_t D_i = \mathcal{L}^D_i(V, D)\ \Delta t, \quad \text{for all cells } T_i, \tag{36b}$$

where $\mathcal{L}^V_{ij}$ and $\mathcal{L}^D_i$ denote the deterministic spatial operators, and Δt stands for the discrete time step. The RHS of the momentum Equation 36a is given by

$$\mathcal{L}^V_{ij}(V, D) \triangleq -\mathrm{Adv}(V, D)_{ij} - \mathrm{K}(V)_{ij} - \mathrm{G}(D)_{ij}, \tag{37}$$

where Adv denotes the discretization of the advection term $(\boldsymbol{\nabla} \times \boldsymbol{u} + \boldsymbol{f}) \times \boldsymbol{u}$ of Equation 30, K the approximation of the gradient of the kinetic energy $\boldsymbol{\nabla}(\frac{1}{2}\boldsymbol{u}^2)$ and G of the gradient of the height field $g\nabla\eta$. Explicitly, the advection term is given by

19422466, 2021, 12, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021MS002492 by Cochrane France, Wiley Online Library on [25/10/2023]. 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

$$
\begin{aligned}
&Adv(V,D)_{ij} \triangleq \\
&-\frac{1}{\overline{D}_{ij}|\tilde{e}_{ij}|}\left((\mathrm{Curl}\ \ V)_{\zeta_-}+f_{\zeta_-}\right)\left(\frac{|\zeta_-\cap T_i|}{2|T_i|}\overline{D}_{ji_-}|e_{ii_-}|V_{ii_-}+\frac{|\zeta_-\cap\ T_j|}{2|T_j|}\overline{D}_{ij_-}|e_{jj_-}|V_{jj_-}\right) \\
&+\frac{1}{\overline{D}_{ij}|\tilde{e}_{ij}|}\left((\mathrm{Curl}\ \ V)_{\zeta_+}+f_{\zeta_+}\right)\left(\frac{|\zeta_+\cap T_i|}{2|T_i|}\overline{D}_{ji_+}|e_{ii_+}|V_{ii_+}+\frac{|\zeta_+\cap\ T_j|}{2|T_j|}\overline{D}_{ij_+}|e_{jj_+}|V_{jj_+}\right),
\end{aligned}
\tag{38}
$$

where $f_{\zeta_\pm}$ is the Coriolis term evaluated at the center of $\zeta_\pm$. Moreover, the two gradient terms read:

$$
K(V)_{ij} \triangleq \frac{1}{2}(\mathrm{Grad}_n\ \ F)_{ij}, \qquad F_{T_i}=\sum_{k\in\{j,i_-,i_+\}}\frac{|\tilde{e}_{ik}|\ \ |e_{ik}|(V_{ik})^2}{2|T_k|}, \tag{39}
$$

$$
G(D)_{ij} \triangleq g(\mathrm{Grad}_n\ \ (D+\eta_b))_{ij}. \tag{40}
$$

The RHS of the continuity equation (Equation 36b) is given by

$$
\mathcal{L}_i^D(V,D) \triangleq -\left(\mathrm{Div}\ \ (\overline{D}V)\right)_i, \tag{41}
$$

which approximates the divergence term $-\boldsymbol{\nabla}\cdot(\boldsymbol{u}h)$.

3.1.3. Time Scheme

For the time integrator we use a Crank-Nicolson-type scheme where we solve the system of fully discretized non-linear momentum and continuity equations by a fixed-point iterative method. The corresponding algorithm coincides for $\boldsymbol{\sigma}=0$ with the one given in Section 3.3.

3.2. Spatial Discretization of RSW–LU

The fully stochastic system has additional terms on the RHS of Equations 29a and 29b. With these terms the discrete equations read:

$$
\mathrm{d}_t V_{ij} = \mathcal{L}_{ij}^V(V,D)\ \ \Delta t + \Delta\mathcal{G}_{ij}^V, \tag{42a}
$$

$$
\mathrm{d}_t D_i = \mathcal{L}_i^D(V,D)\ \ \Delta t + \Delta\mathcal{G}_i^D, \tag{42b}
$$

where the stochastic LU-terms are given by

$$
\Delta\mathcal{G}_{ij}^V \triangleq \left(\frac{\Delta t}{2}\ \ (\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\cdot(\boldsymbol{a}\boldsymbol{u}))_{ij}-(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}\boldsymbol{u})_{ij}\right)\cdot\boldsymbol{n}_{ij}, \tag{42c}
$$

$$
\Delta\mathcal{G}_i^D \triangleq \frac{\Delta t}{2}\left(\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\cdot(\boldsymbol{a}D)\right)_i-(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}D)_i. \tag{42d}
$$

Note that the two terms within the large bracket in Equation 42c comprise two Cartesian components of a vector which is then projected onto the triangle edge's normal direction via $\boldsymbol{n}_{ij}$. The two terms in Equation 42d are scalar valued at the cell circumcenters i.

The parametrization of the noise described in Appendix B is formulated in Cartesian coordinates, because this allows using standard algorithms to calculate EOFs, for instance. Likewise, we represent the stochastic LU-terms in Cartesian coordinates but to connect both deterministic and stochastic terms, we will calculate the occurring differentials with operators as provided by the deterministic dynamical core (see interface description below). Therefore, we write the second term in Equation 42c as

$$
(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\nabla F)_{ij}=\sum_{l=1}^{2}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_{ij}^l(\nabla F)_{ij}^l, \tag{43}
$$

19422466, 2021, 12, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021MS002492 by Cochrane France, Wiley Online Library on [25/10/2023]. 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, 2021, 12, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021MS002492 by Cochrane France, Wiley Online Library on [25/10/2023]. 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 which $(\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t)_{ij}$ denotes the discrete noise vector with two Cartesian components, constructed as described in Appendix B and evaluated at the edge midpoint ij. The scalar function F is a placeholder for the Cartesian components of the velocity field $\boldsymbol{u} = (u^1, u^2)$. Likewise, the first term in Equation 42c can be written component-wise as

$$(\boldsymbol{\nabla} \cdot \boldsymbol{\nabla} \cdot (\boldsymbol{a} F))_{ij} = \sum_{k,l=1}^{2} \left(\partial_{x_k} \left(\partial_{x_l} (a_{kl} F) \right)_{ij} \right)_{ij}, \tag{44}$$

where a_{kl} denotes the matrix elements of the variance tensor which will be evaluated, similarly to the discrete noise vector, at the edge midpoints. For a concrete realization of the differentials on the RHS of both stochastic terms, we will use the gradient operator (Equation 32) as introduced next.

To calculate the terms in Equation 42d we also use the representations (Equations 43 and 44) for a scalar function $F = D$ describing the water depth. However, as our proposed procedure will result in terms at the edge midpoint ij, we have to average them to the cell centers i.

In the following, we will refer to this part of the code that generates the noise on a Cartesian mesh according to Appendix B as *noise generation module*.

3.2.1. Interface Between Dynamical Core and LU Terms

As mentioned above, the construction of the noise is done on a Cartesian mesh while the discretization of the deterministic dynamical core (variational RSW scheme, Section 3.1), corresponding to a triangular C-grid staggering, predicts the values for velocity normal to the triangle edges and for water depth at the triangle centers. We propose to exchange information between the noise generation module (see section above) and the dynamical core via the midpoints of the triangle edges where on such C-grid staggered discretizations the velocity values naturally reside. The technical details about how we realized such interface in our setup are given in Appendix C.

This modular approach with a well defined interface between these two model components has various advantages over directly implementing the noise terms on a triangular C-grid mesh as used by the dynamical core. Firstly, this approach allows us to easily explore various noise types, because using a Cartesian mesh for the latter permits the usage of standard algorithms for example FFT or singular value decomposition (SVD). In contrast, exploring these ideas directly on a triangular C-grid would significantly increase the implementation work. In fact, this manuscript also serves as a proof of concept study to show that such modular approach indeed works very well.

Moreover, the definition of an interface between the two model components should minimize (or maybe even avoid) the necessity of adapting the numerics of an existing deterministic core in order to incorporate the discrete stochastic LU-terms. This, in turn, should allow us to apply our method directly to existing dynamical cores of NWP models.

3.2.2. Computational Aspects

In addition to the deterministic scheme we have the terms $\Delta\mathcal{G}^V$ and $\Delta\mathcal{G}^D$ for the RSW–LU scheme (see, Equations 42c and 42d). Their discretization can be differentiated into:

1. The noise generation of $\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t$ and $\boldsymbol{a}$. The noise generation relies on generating a fixed number of pseudo-observations and carrying out a SVD to obtain the EOFs. The SVD can be carried out as an economy-size SVD which depends linearly on the number of triangles. Currently for LU on-line, EOFs are estimated at each time step, but less frequent estimations are also possible to save computational costs.
2. The computation of the divergence and gradient in Cartesian coordinates. The discretization of these operations are described in Appendix C, which results in matrix vector multiplications.

Here, we obtain the discretization of $\Delta\mathcal{G}^V$ and $\Delta\mathcal{G}^D$ using the interface, which is determined by the underlying discretization of the deterministic scheme. More specifically, we reformulate the differential operators in Cartesian coordinates with the local derivatives obtained from the deterministic scheme (see e.g., Equation C2). This results only in a few additional matrix vector multiplications.

Optimized standard methods for the noise generation on a Cartesian mesh are potentially more efficient than a direct (and not optimized) implementation on a triangular mesh. Besides the advantages mentioned above and given that the additional computational costs for interchanging the values via the interface consists of only a few matrix vector multiplications, we advocate our modular approach rather than a direct implementation.

19422466, 2021, 12, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021MS002492 by Cochrane France, Wiley Online Library on [25/10/2023]. 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

3.3. Temporal Discretization of RSW–LU

The iterated Crank-Nicolson method presented in Brecht et al. (2019) is adopted for the temporal discretization. Keeping the iterative solver and adding the LU terms results in an Euler-Maruyama scheme, which decrease the order of convergence of the deterministic iterative solver (see, Kloeden & Platen, 1992 for details).

To enhance readability, we denote V^t as the array over all edges e_{ij} of the velocity V_{ij} and D^t as the array over all cells T_i of the water depth D_i at time t. The governing algorithm reads:

For all simulations in this manuscript, we used a tolerance of 10^{-10} for simulation on the sphere. In all these cases, our suggested fixed point solver converges in less than 10 iterations.

4. Numerical Results

In this section, we first study the energy behavior of the numerical RSW–LU scheme introduced above for an inviscid test flow. Then, we show that for a viscous test case, the stochastic model captures more accurately the reference structure of the large-scale flow when compared to the deterministic model under the same coarse resolution. In addition, we demonstrate that the proposed RSW–LU system provides a more reliable ensemble forecast with larger spread, compared to a classical random model based on the perturbations of the initial conditions (PIC).

4.1. Inviscid Test Case–Energy Analysis

Algorithm 1: Time-stepping algorithm

Set iterative solver index $k = 0$ with initial guess at t:

$$V^*_{k=0} = V^t,$$
$$(D^*_{k=0}) = D^t + \Delta\mathcal{G}^D(D^t),$$

and compute $\Delta\mathcal{G}^V_{ij}(V^t)$.

while $\|V^*_{k+1} - V^*_k\| + \|D^*_{k+1} - D^*_k\| >$ *tolerance* **do**

$$\frac{D^*_{k+1} - D^t}{\Delta t} = -\frac{\mathrm{Div}\,(\overline{D^*_k}V^*_k) + \mathrm{Div}\,(\overline{D^t}V^t)}{2}$$
$$\frac{V^*_{k+1} - V^t}{\Delta t} = -\frac{\mathrm{Adv}(V^*_k, D^*_{k+1}) + \mathrm{Adv}(V^t, D^t)}{2} - \frac{\mathrm{K}(V^*_k) + \mathrm{K}(V^t)}{2} - \mathrm{G}(D^*_{k+1}) + \Delta\mathcal{G}^V_{ij}(V^t)$$

and set $k + 1 = k$.

end

This first test case consists of two co-rotating vortices on the f-plane. To illustrate the energy conservation of the spatial discretization of the RSW–LU system (Equations 27a–27d), we use the homogeneous stationary noise defined in Appendix B1 since the two incompressible constraints $\nabla \cdot \boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t = 0$ and $\boldsymbol{\nabla} \cdot \boldsymbol{\nabla} \cdot \boldsymbol{a} = 0$ in Equation 27d are naturally satisfied. Then, no extra steps are required to satisfy the incompressible constraints.

4.1.1. Initial Conditions

The simulations are performed on a rectangular double periodic domain $\Omega = [0, L_x] \times [0, L_y]$ with $L_x = 5000$ km and $L_y = 4330$ km, which is discretized into $N = 32768$ triangles. We use this resolution for both the deterministic and stochastic simulations. The large-scale flow is assumed to be under a geostrophic regime at the initial state, that is $f\boldsymbol{k} \times \boldsymbol{u} = -g\boldsymbol{\nabla} h$. We use an initial height field elevation (as e.g., in Bauer & Gay-Balmaz, 2019a) of the form

$$h(x, y, t = 0) = H_0 - H' \left(\exp\left(-\frac{{x'_1}^2 + {y'_1}^2}{2} \right) + \exp\left(-\frac{{x'_2}^2 + {y'_2}^2}{2} \right) - \frac{4\pi s_x s_y}{L_x L_y} \right), \tag{45a}$$

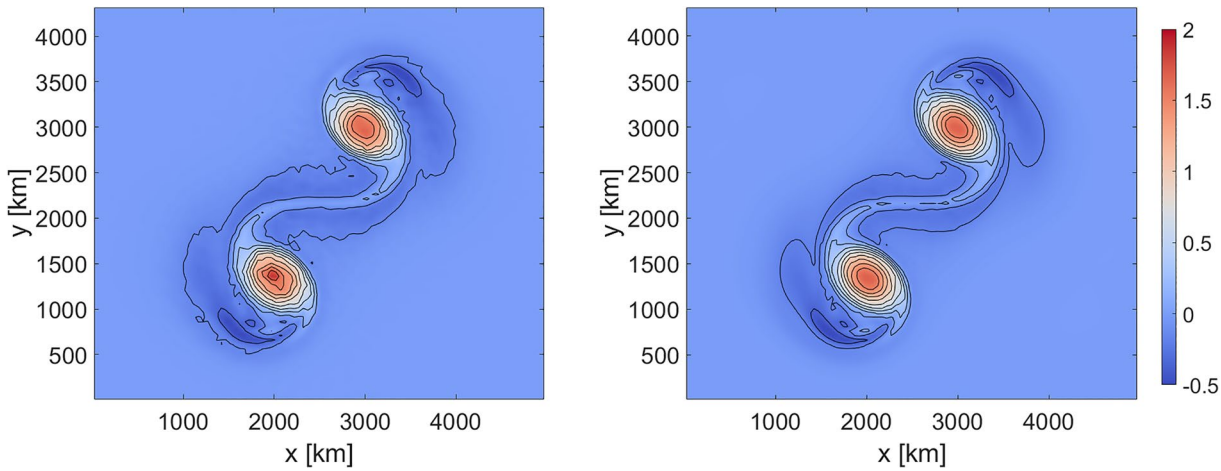


Figure 3. Contour plots of the potential vorticity fields after 2 days for (left) one realization of a LU simulation with homogeneous noise and (right) a deterministic run. The contour interval is 0.4 days^{-1} km^{-1}.

where the background height H_0 is set to 10 km, the magnitude of the small perturbed height H' is set to 75 m and the periodic extensions x_i', y_i' are given by

$$x_i' = \frac{L_x}{\pi s_x}\sin\left(\frac{\pi}{L_x}(x - x_{c_i})\right), \qquad y_i' = \frac{L_y}{\pi s_y}\sin\left(\frac{\pi}{L_y}(y - y_{c_i})\right), \qquad i = 1, 2 \tag{45b}$$

with the centers of the vertices located at $(x_{c_1}, y_{c_1}) = \frac{2}{5}(L_x, L_y)$, $(x_{c_2}, y_{c_2}) = \frac{3}{5}(L_x, L_y)$ with parameters $(s_x, s_y) = \frac{3}{40}(L_x, L_y)$.

To obtain the discrete initial water depth D_i, we sample the analytical function h at each cell center. Subsequently, the discrete geostrophic velocities at each triangle edge ij at the initial state can be deduced via

$$V_{ij} = -\frac{g}{f}(\mathrm{Grad}_t\ D)_{ij}, \tag{46}$$

where the Coriolis parameter f is set to 5.3108 days^{-1}. For the LU simulations, the magnitude of the homogeneous noise remains moderate with its constant variance a_0 set to be 169.1401 m$^2 \cdot$ s^{-1}.

4.1.2. Analysis of Energy Conservation

To analyze the energy conservation properties of our stochastic integrator, we use the above initial conditions to simulate the two co-rotating vortices for 2 days. In Figure 3, we show contour plots of the potential vorticity (as defined in Equation 35) fields of the deterministic and stochastic models. We observe that under the moderate noise with a_0 as chosen above, the large-scale structure of the stochastic system is similar to that of the deterministic run.

On the specific staggered grid as shown in Figure 2, the total energy of the shallow water Equations A1a and A1b, for both deterministic and stochastic case, is approximated by

$$\mathrm{E}(t) \approx \sum_{i=1}^{N} \frac{1}{2} D_i(t)|T_i| \sum_{k=j,i_-,i_+} \frac{1}{2|T_i|} h_{ik} f_{ik} (V_{ik}(t))^2 + \frac{1}{2} g (D_i(t))^2 |T_i|. \tag{47}$$

As shown in Bauer & Gay-Balmaz (2019a), the proposed discrete variational integrator (see Section 3.1) together with an iterative Crank-Nicolson time stepping method exhibits a 1st order convergence rate of the energy error with smaller time step size. This will allows us immediately to simply include the stochastic terms to result in an Euler-Maruyama type time integrator for stochastic systems (cf. Section 3.2).

In the present work, we consider the energy behavior of the deterministic scheme (i.e., the variational integrator) as reference, which is denoted as $\mathrm{E}_{\mathrm{REF}}(t)$ in the following. For the stochastic RSW model, the Euler-Maruyama time scheme might lead to a different behavior with respect to energy conservation when compared to the deterministic model. In order to quantify numerically the energy conservation of the RSW–LU, we propose to measure the relative errors between the mean stochastic energy, denoted as $\overline{\mathrm{E}}_{\mathrm{LU}}(t)$, and the reference $\mathrm{E}_{\mathrm{REF}}(t)$ by

19422466, 2021, 12, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021MS002492 by Cochrane France, Wiley Online Library on [25/10/2023]. 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, 2021, 12, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021MS002492 by Cochrane France, Wiley Online Library on [25/10/2023]. 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 1
Parameter List for Simulations of the Barotropic Instability

Parameters	Value	Description
(Θ_0, Θ_1)	$(2\pi, 5\pi)/14$ rad	Initial latitude limits
H_0	10.158 km	Background height
H'	120 m	Initial perturbation amplitude
R	6.371×10^3 km	Mean radius of earth
g	9.806 $m \cdot s^{-2}$	Gravity of earth
$\tilde{\Omega}$	7.292×10^{-5} s^{-1}	Angular rotation rate of earth
U_0	80 $m \cdot s^{-1}$	Maximum zonal velocity
μ_l	3.975×10^{14} $m^4 \cdot s^{-1}$	Fine-grid biharmonic viscosity
μ_L	3.199×10^{16} $m^4 \cdot s^{-1}$	Coarse-grid biharmonic viscosity
Δt_l	12 s	Fine-grid time step
Δt_L	50 s	Coarse-grid time step
N_l	3,27,680	Number of triangles for fine grid (60-km resolution)
N_L	20,480	Number of triangles for coarse grid (240-km resolution)

$\overline{\mathrm{E}}_{\mathrm{LU}}(t)/\mathrm{E}_{\mathrm{REF}}(t) - 1$, while using for both the same spatial resolution (see Table 1). This setup allows us to measure the influence of the stochastic terms on the energy conservation relative to the deterministic scheme. Figure 4 shows these relative errors for different time step sizes over a simulation time of 2 days. As we can confirm from the curves, taking successively smaller time steps

$\Delta t \in \left\{1.7361 \times 10^{-4}, 3.4722 \times 10^{-5}, 1.7361 \times 10^{-5}, 3.4722 \times 10^{-6}, 1.7361 \times 10^{-6}\right\}$ (in days^{-1}) results in smaller relative errors.

To determine more quantitatively the convergence rate of the stochastic scheme (relative to the reference) with respect to different time step sizes, we defined the following global (in space and time) error measure:

$$\epsilon(\mathrm{E}_{\mathrm{LU}}) = \frac{\|\mathrm{E}_{\mathrm{LU}}(t) - \mathrm{E}_{\mathrm{REF}}(t)\|_{L^2([0,\ T])}}{\|\mathrm{E}_{\mathrm{REF}}(t)\|_{L^2([0,\ T])}}, \tag{48}$$

where $\|f(t)\|_{L^2([0,\ T])} = \left(\int_0^T |f(t)|^2 \mathrm{d}t\right)^{1/2}$ and T is set to 2 days. We determine for an ensemble with 10 members such global errors in order to illustrate the convergence rate of each ensemble member and the spread between those rates. This spread is illustrated as blue shaded area in Figure 5. The area center is determined by the mean of the errors, and the dispersion of this area is given by one standard derivation (i.e. 68% confident interval of the ensemble of $\epsilon(\mathrm{E}_{\mathrm{LU}})$). Besides, the minimal and maximal values of the errors of the ensemble are represented by the vertical bar-plots. The blue line of Figure 5 shows that the convergence rate (w.r.t. various Δt) of the ensemble mean energy is of 1st order. This is consistent with the weak convergence rate of order $\mathcal{O}(\Delta t)$ of the Euler-Maruyama scheme, cf. Section 3.3.

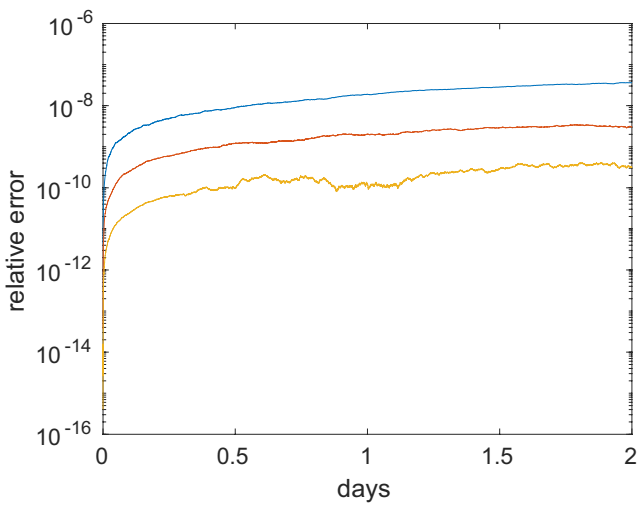


Figure 4. Evolution of the relative L_2 errors between the energy of the mean RSW–LU and the reference, using Δt (blue line), $\Delta t/10$ (red line) and $\Delta t/100$ (yellow line) respectively.

4.2. Viscous Test Case - Ensemble Prediction

Next, we want to show that our stochastic system better captures the structure of a large-scale flow than a comparable deterministic model. To this end, we use a viscous test case and heterogeneous noise.

The viscous test case we use is proposed by Galewsky et al. (2004) and it consists of a barotropically unstable jet at the mid-latitude on the sphere. This strongly non-linear flow will be destabilized by a small perturbation of the initial field, which induces decaying turbulence after a few days. However,

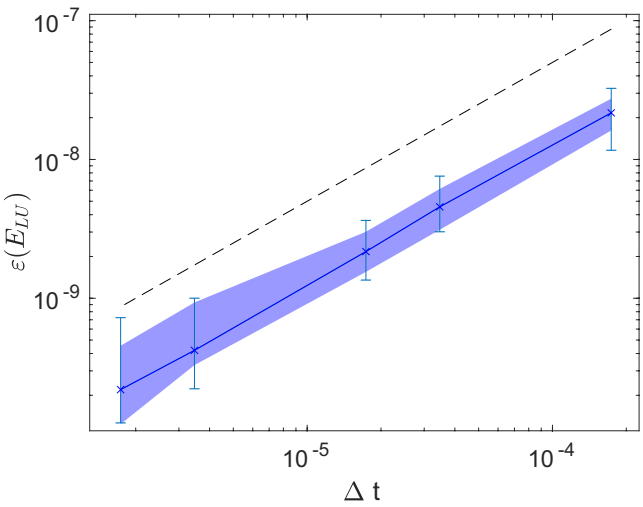


Figure 5. Convergence of the energy path of the RSW–LU to that of the reference w.r.t. time step sizes. The blue line shows the global errors of the ensemble mean energy, the blue area describes the 68% confident interval of the ensemble errors and the dashed line stands for the 1st order convergence rate.

the development of the barotropic instability in numerical simulations highly depends on accurately resolving the small-scale flow, which is particularly challenging for coarse-grid simulations. For the same reason, the performance of an ensemble forecast system in this test case is quite sensible to the numerical resolution. In the following, we demonstrate that the RSW–LU simulation on a coarse mesh under heterogeneous noises, provides better prediction of the barotropic instability compared to the deterministic coarse simulation, and produces more reliable ensemble spread than the classical PIC simulation.

4.2.1. Stabilization

The former test case in Section 4.1 consists of smooth enough fields such that no additional sub-grid dissipation is required. In contrast, the following test case consists of the evolution of decaying turbulence, in which sub-grid enstrophy will accumulate quickly, hence an efficient dissipation mechanism is needed, such as the biharmonic eddy viscosity (Galewsky et al., 2004) which is often used in atmospheric and oceanic flow models. Here, we include a biharmonic eddy viscosity with uniform coefficient μ (of unit m^4/s) in the momentum equation:

$$\mathrm{d}_t V = (-\mathrm{Adv}(V, D)_{ij} - \mathrm{K}(V)_{ij} - \mathrm{G}(D)_{ij} - \mu L(V)_{ij}) \quad \Delta t, \tag{49}$$

where:

$$L(V)_{ij} = (\mathrm{Grad}_n \ (\mathrm{Div} \ V)_{ij} - \mathrm{Grad}_t \ (\mathrm{Curl} \ V)_{ij})^2. \tag{50}$$

Although in the evolution Equation 29a the dissipative term is energetically exactly in balance with the random advection term, the supplementary biharmonic diffusion is needed here in this test case to drain the enstrophy pile-up. Using instead a dissipative discretization, in which numerical diffusion takes the role of such stabilization, might give stable simulations also without explicit diffusion but then we would lose control of the strength of the diffusion. Note that we used standard biharmonic dissipation, but there exist also energy conserving enstrophy dissipation methods, such as those introduced in McRae & Cotter (2014) or in Frank et al. (2003).

4.2.2. Initial Conditions

The values of the principle parameters for the simulations are specified in Table 1. Under the geostrophic regime, the initial zonal velocity and height is respectively given by

$$u(\Theta, t=0) = \frac{U_0}{e_n} \exp\left(\frac{1}{(\Theta-\Theta_0)(\Theta-\Theta_1)}\right), \quad \text{for } \Theta_0 < \Theta < \Theta_1, \tag{51a}$$

$$h(\Theta, t=0) = H_0 - \frac{R}{g}\int_\Theta u(\theta, t=0)\left(2\tilde{\Omega}\sin\theta + \frac{\tan\theta}{R}u(\theta, t=0)\right) \mathrm{d}\theta, \tag{51b}$$

where $e_n = \exp\left(-4/(\Theta_1-\Theta_0)^2\right)$ is used to rescale the jet magnitude to the maximal value U_0 at the jet's mid-point $\Theta = \pi/4$. As introduced by Galewsky et al. (2004), in order to initiate the barotropic instability, the following localized bump is included in the height field:

$$h'(\Upsilon, \Theta) = H' \cos \ \Theta \ \exp\left(-(3\Upsilon)^2 - \left(15(\frac{\pi}{4} - \Theta)\right)^2\right), \tag{51c}$$

where Υ denotes the longitude. Here, the Coriolis parameter is set to $f = 2 \times 7.292 \times 10^{-5}\sin(\Theta)$. Analogously to the previous inviscid test case, we then use these analytic functions (Equations 51a–51c) to sample the discrete velocity at the edge mid-points and the height field at the cell centers on the staggered mesh (See Figure 2).

For the LU simulations, we use the two heterogeneous noises described in Appendix B2, based on either the off-line learning of EOFs from the high-resolution simulation data, denoted as LU off-line, or on the on-line estimation of EOFs from the coarse-grid simulation, denoted as *LU on-line*. To allow for comparisons, the strength of

19422466, 2021, 12, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021MS002492 by Cochrane France, Wiley Online Library on [25/10/2023]. 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

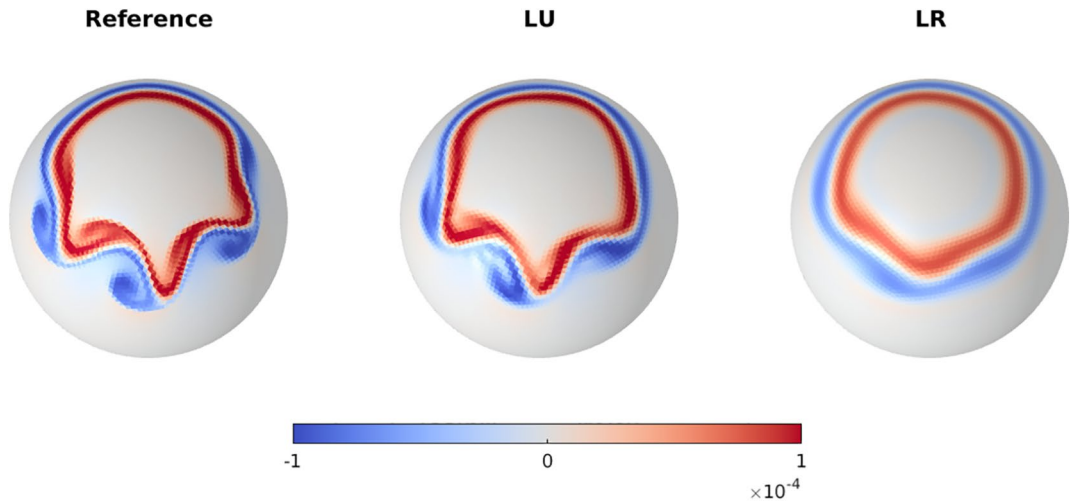


Figure 6. Snapshots of the vorticity field on the sphere for different models (with 20,480 triangles) after 5 days. From left to right: reference, ensemble mean of LU online and deterministic LR. For the simulations we use the parameters given in Table 1.

these two noises are imposed to be the same. The PIC stochastic model is obtained as follows: first, we perform ensemble simulations of the LU off-line and the LU on-line method over 1 day. Then, each realization of these ensemble runs is used as one initial random state and simulated for the remaining days using the deterministic scheme. We call the PIC simulation using the LU off-line method *PIC 1* and the PIC simulation obtained using the LU on-line method *PIC 2*. For each of these stochastic models, an ensemble run with 20 realizations is done.

Besides a deterministic coarse-grid simulation denoted as *LR*, a deterministic high resolution (HR) simulation is performed that provides us with a reference solution. For all coarse model runs (both deterministic and stochastic), the resolution and parameters given in Table 1 are fixed to be the same. Note that Table 1 states the resolutions and parameters used for these various simulations.

4.2.3. Prediction of Barotropic Instability

In this section, we compare the predictions of the barotropic instability for different coarse models to that provided by the HR reference simulation. The latter is obtained from the coarse-graining procedure through a bilinear interpolation of the high resolution snapshots.

In Figure 6, we illustrate snapshots of the vorticity fields on the sphere for the reference, LU and deterministic models after a simulation time of 5 days. We can clearly see that the LU ensemble mean better captures the large-scale structure of the reference flow than the deterministic simulation. To better distinguish the differences in the simulations, contour plots of the vorticity fields at day 4, 5 and 6, localized at the mid-latitude of the sphere, are given in Figure 7. From the evolution of the reference vorticity fields we observe that the barotropic instability of the mid-latitude jet starts to develop at day 4. Subsequently, more and more small-scale features emerge and the flow becomes turbulent. Furthermore, both LU on-line and LU off-line simulations exhibit the stretched out wave at day 5 in the same way as the reference does, and that some big vortices start to separate from the wave at day 6. On the other hand, these characteristics are not correctly captured in both PIC 1 and LR simulations. We remark that the results of the PIC 2 simulations are not included in Figure 7, since they behave quite similarly to the PIC 1 runs.

To physically interpret the above results, it is useful to analyze the energy spectra of the different models. From a basic knowledge of the two-dimensional turbulence theory (McWilliams, 2006), the potential enstrophy is transferred from the large scales to the small scales by the direct cascade, whereas the kinetic energy is transferred from the small scales to the large scales by the inverse cascade. However, introducing only a dissipation mechanism for coarse models often leads to an excessive decrease of the resolved kinetic energy (Arbic et al., 2013; Kjellsson & Zanna, 2017).

19422466, 2021, 12, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021MS002492 by Cochrane France, Wiley Online Library on [25/10/2023]. 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

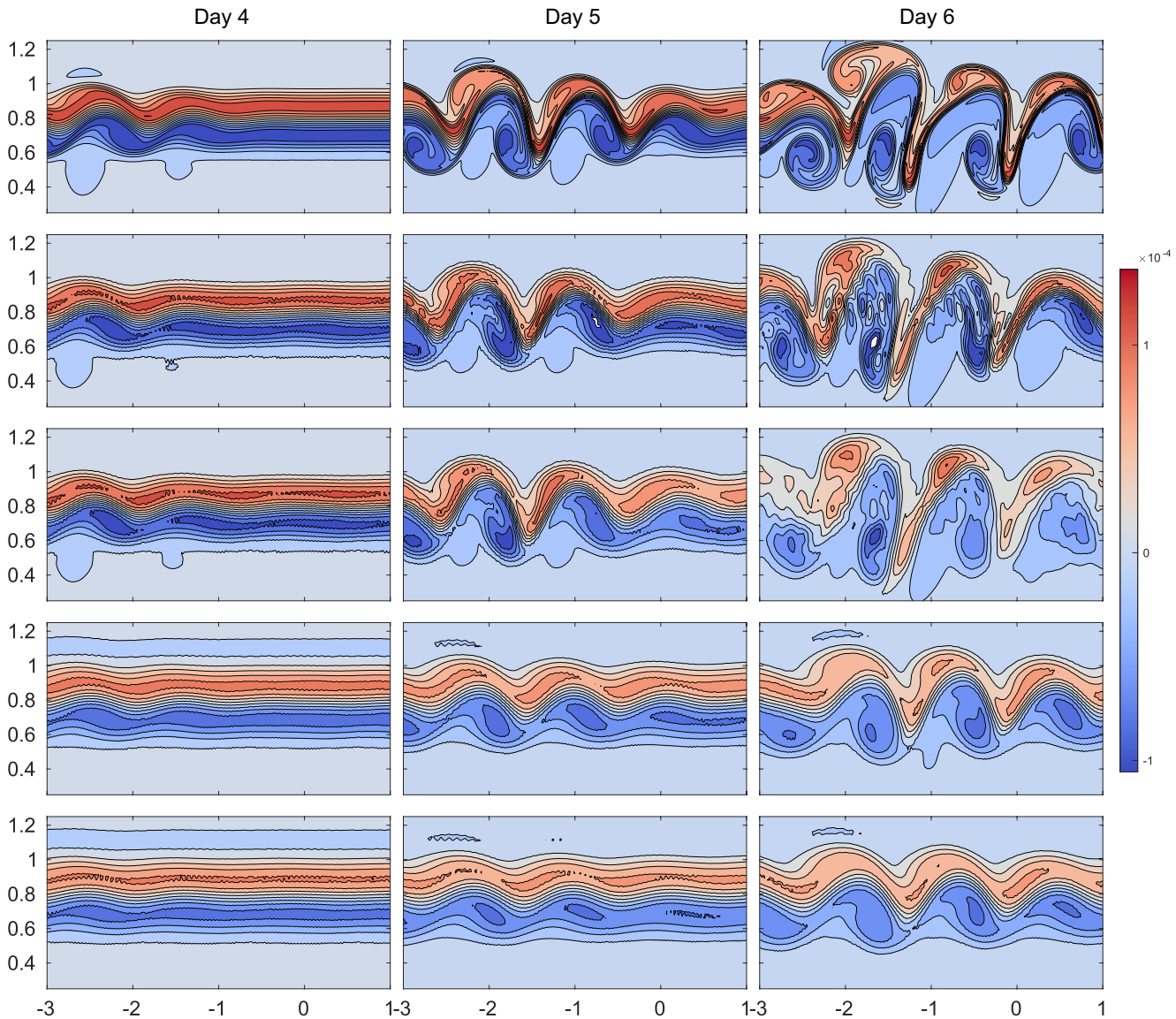


Figure 7. Comparison of the vorticity contour plots along the mid-latitude jet for different models (with 20,480 triangles) at day 4, 5 and 6 respectively. From top to bottom: reference, ensemble mean of LU on-line, ensemble mean of LU off-line, ensemble mean of PIC 1 and deterministic LR. The contour interval is fixed to 2×10^{-5} s^{-1}, the x-axis is longitude (in rad) and the y-axis is latitude (in rad). For the simulations we use the parameters given in Table 1.

In our test case, this kind of issue is present in both PIC and the LR simulations, where the small-scale energy and enstrophy are over-dissipated, as illustrated in Figure 8. On the other hand, introducing the non-linear convection by the noise, the LU dynamical systems bring higher turbulent energy and enstrophy to the small scales, which leads to a better structuring of the large-scale flow. For instance, the time evolutions of the ensemble mean of the energy and enstrophy spectra for both LU on-line and LU off-line simulations are much closer to that of the references. However, the LU off-line spectrum changes little over time between wavenumbers 10 and 40 because the a priori obtained EOFs impose at each time step large scale modes on those scales. This is a drawback from a stationary noise. Note that these spectra on the sphere are calculated using the method proposed by Aechtner et al. (2015): first, the energy and enstrophy is interpolated onto a Gaussian grid, then the spherical harmonics basis are used to compute the power spectral density.

4.2.4. Evaluation of Ensemble Forecasts

Once the ensembles have been produced by the random models, we measure the reliability of the ensemble forecast systems by some simple metrics. But before we do so, let us first demonstrate qualitatively the time evolution

19422466, 2021, 12, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021MS002492 by Cochrane France, Wiley Online Library on [25/10/2023]. 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

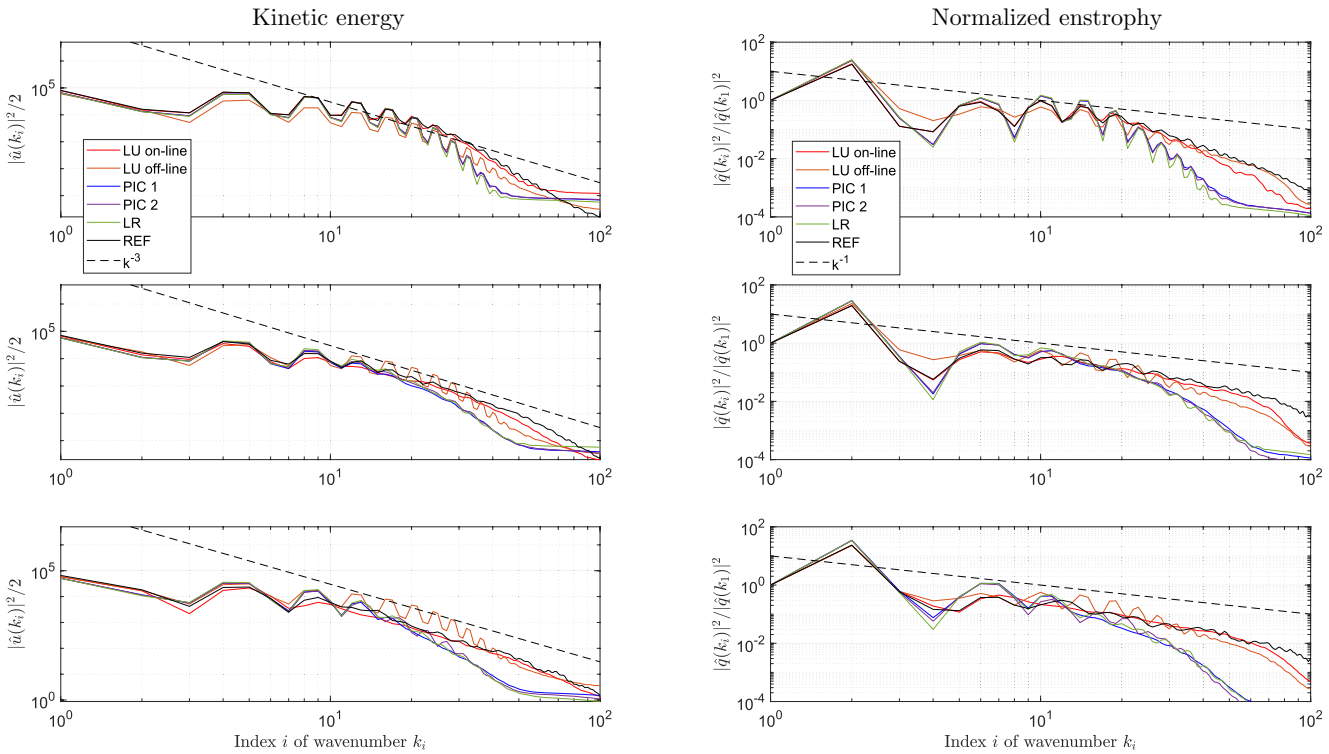


Figure 8. Comparison of the ensemble mean of the kinetic energy (left column) spectrums and the potential enstrophy (right column) spectrums for different models (with 20,480 triangles) at day 5 (1st row), 7 (2nd row) and 10 (3rd row) respectively. Note that the potential enstrophy is defined by the square of the potential vorticity and each potential enstrophy spectrum is normalized by its first value at the largest wavenumber. The dashed line is the k^{-3} (left column) and k^{-1} (right column) power law. These power laws for the RSW equations are discussed in (Chen et al., 2011; Ringler & Randall, 2002).

of each ensemble spread and compare it with the observation trajectory (obtained from the HR reference simulation). To determine the latter, we evaluate the local vorticity field of the reference at different grid points in the region of the mid-latitude jet. These points serve as observation points. The evolution of the spread of the ensemble forecast systems is then built by the 95% confident interval of its ensemble trajectories at each selected point.

In Figure 9 we compare the reference simulation and the simulations obtained from the off-line noise. To make the figure easier to read, only the off-line noise is shown since the on-line noise behaves in a similar way. As shown, for the six local points chosen along the longitude $\Upsilon = -1.53$ rad, the ensemble spreads of the LU off-line system are large enough to almost always include the observation trajectories, whereas the spreads of the PIC 1 system are quite small so that the observations are not always contained within the spread. For the latter, this will result in a wrong coupling of the measurement and the ensemble system, when performing data assimilation (Franzke et al., 2015; Gottwald & Harlim, 2013).

To quantify whether the ensemble spread of the forecast system represents the true uncertainty of the observations (obtained from the reference simulation), the rank histogram (Hamill, 2001; Talagrand et al., 1997) is widely adopted as a diagnostic tool. This approach checks where the verifying observation usually falls w.r.t. the ensemble forecast states which are arranged in an increasing order at each grid point. In an ensemble with perfect spread, each member represents an equally likely scenario, so the observation is equally likely to fall between any two members. To construct the rank histogram in our test case, we proceed as follows:

1. At every grid point $\boldsymbol{x}_i$, we rank the N_e vorticity values $\left\{q^{(j)}(\boldsymbol{x}_i)\right\}_{j=1,\ldots,N_e}$ of the ensemble from lowest to highest. This results in $N_e + 1$ possible bins which the observations can fall into, including the two extremes;
2. Identify which bin the observation vorticity $q^o(\boldsymbol{x}_i)$ falls into at each point $\boldsymbol{x}_i$;
3. Tally over all observations $\{q^o(\boldsymbol{x}_i)\}_{i=1,\ldots,N_o}$ to create a histogram of rank.

19422466, 2021, 12, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021MS002492 by Cochrane France, Wiley Online Library on [25/10/2023]. 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

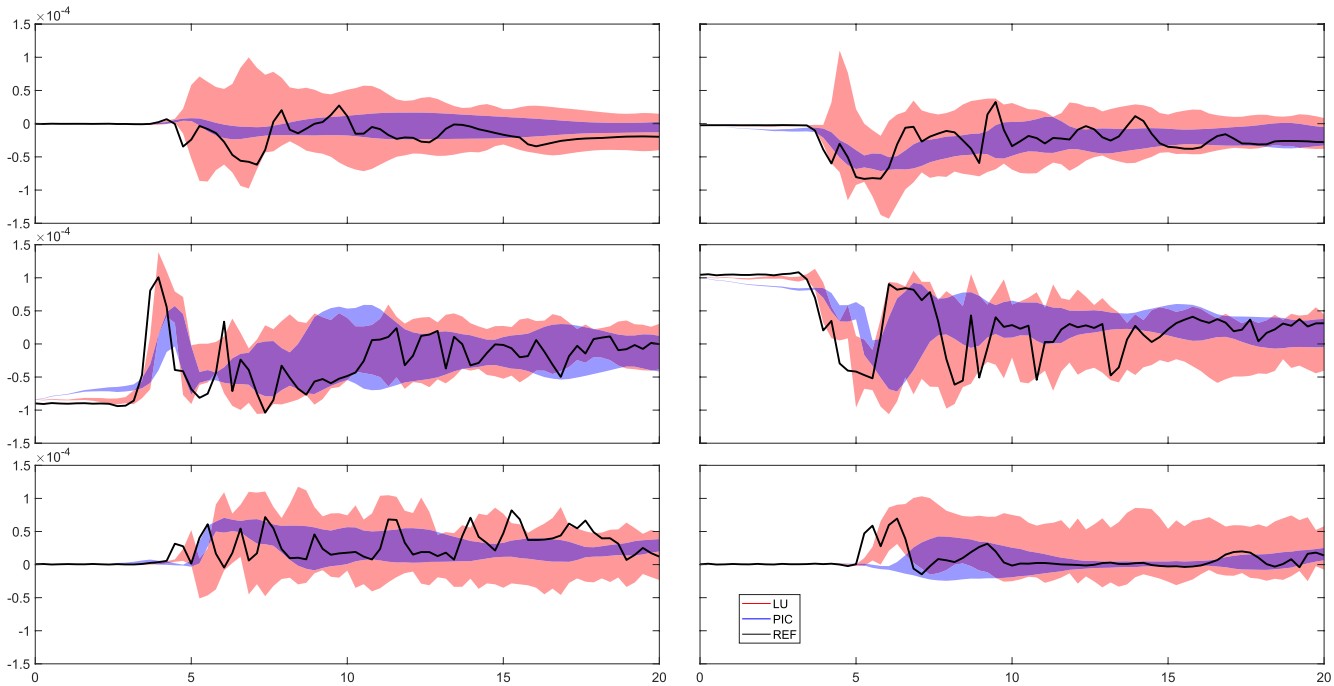


Figure 9. Comparison of the ensemble spread evolution over 20 days of the vorticity field for the LU-offline (red area) runs and the PIC-offline (blue area) runs, at six different locations $\Theta = (0.4, 0.56, 0.72, 0.88, 1.04, 1.2)$ rad along the longitude $\Upsilon = -1.53$ rad. The observation trajectories are shown by the black lines.

As shown in Figure 10, the histograms of both random models exhibit a U-shape for a few days in the beginning, while after a simulation time of about 10 days, the histograms of both LU on-line and LU off-line systems become mostly flat. A U-shape indicates that the ensemble spread is too small so that many observations are falling outside of the extremes of the ensemble while a dome-shape indicates the contrary. A flat histogram, in contrast, indicates that the ensemble members and observations are sampled from a common distribution. We observe that the LU off-line system performs slightly better than the LU on-line version. In contrast to these very good ensemble spreads, the histograms of both PIC 2 and PIC 1 systems remain in a U-shape during the entire simulation period which indicates that these systems do not accurately estimate the correct uncertainty around the observations.

It is important to notice that a flat rank histogram does not necessarily imply good forecasts, it only measures whether the observed probability distribution is well represented by the ensemble. To verify that a forecast is reliable, we need more criteria. One necessary criterion (Weigel, 2012) for a reliable ensemble forecast is that the mean squared error (MSE) of the ensemble matches the mean intra-ensemble variance (MEV), up to an ensemble size-dependent scaling factor, that is

$$\begin{aligned} \text{MSE}\ (t) &= \frac{1}{N_o}\sum_{i=1}^{N_o}\left(q^o - \hat{\mathbb{E}}[q]\right)^2(t, \boldsymbol{x}_i) \\ &\approx \left(\frac{N_e + 1}{N_e}\right)\frac{1}{N_o}\sum_{i=1}^{N_o}\widehat{\text{Var}}[q](t, \boldsymbol{x}_i) = \frac{N_e + 1}{N_e}\ \text{MEV}\ (t), \end{aligned} \tag{52}$$

where $\hat{\mathbb{E}}[q] = \frac{1}{N_e}\sum_{j=1}^{N_e} q^{(j)}$ and $\widehat{\text{Var}}[q] = \frac{1}{N_e - 1}\sum_{j=1}^{N_e}\left(q^{(j)} - \hat{\mathbb{E}}[q]\right)^2$ denote the empirical mean and the empirical variance, respectively.

In Figure 11, we compare the differences in time between the MSE and the MEV, normalized by the squared maximum of the initial vorticity, for the different random models from above. From these curves we can deduce that the LU off-line system exhibits the lowest errors during the entire simulation time of 20 days. In particular, during the first 10 days, these errors are significantly lower when compared to the other models, which can be explained by the fact that the LU off-line system incorporates data from the reference into the ensemble, which increases

19422466, 2021, 12, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021MS002492 by Cochrane France, Wiley Online Library on [25/10/2023]. 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

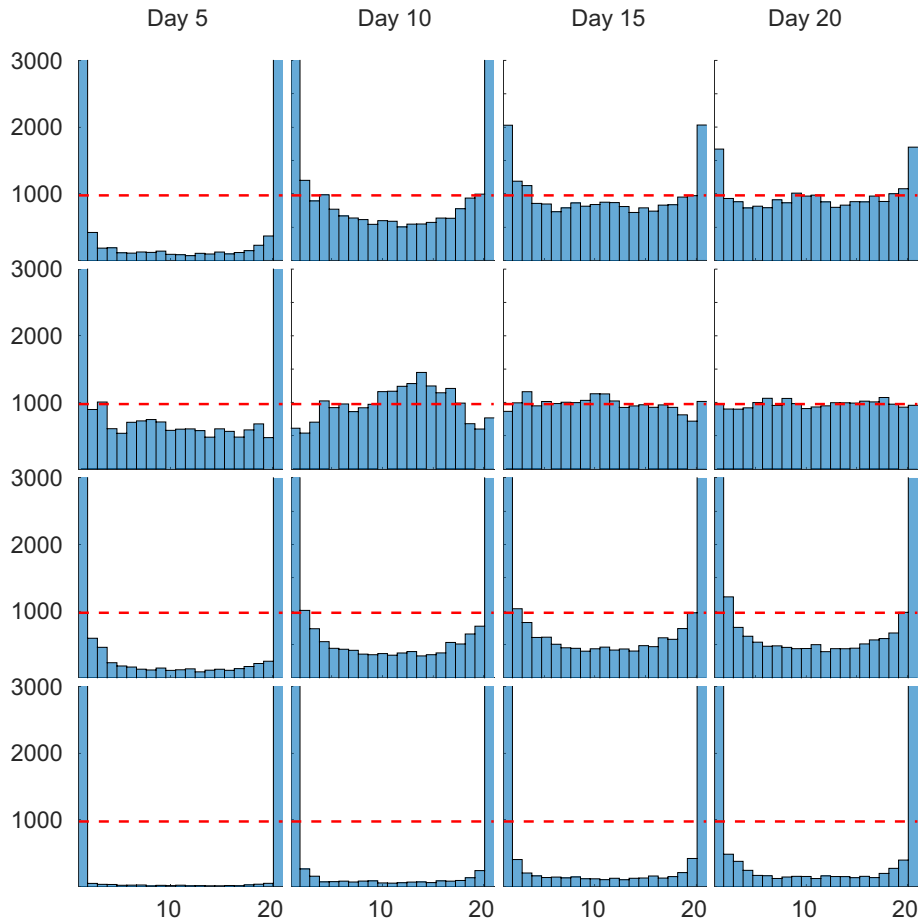


Figure 10. Comparison of the rank histograms for the LU on-line (1st row) runs, the LU off-line (2nd row) runs, the PIC 2 (3rd row) runs and PIC 1 (last row) runs, at day 5, 10, 15 and 20 respectively.

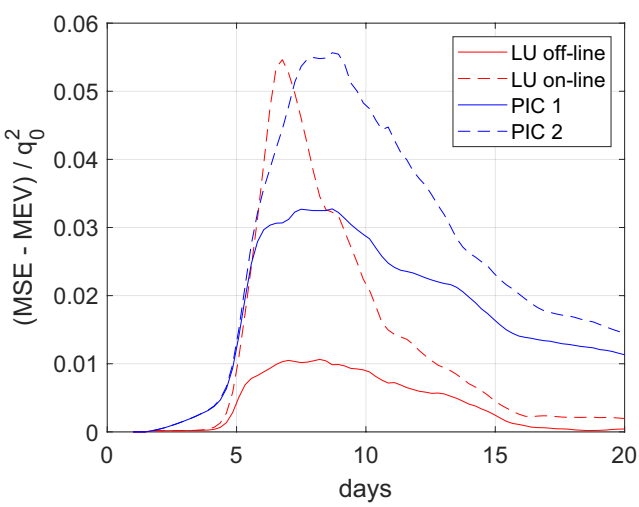


Figure 11. Comparison of the differences between the mean square error (MSE) and the mean ensemble variance (MEV) of the ensemble vorticity fields for the LU on-line (red dashed line) runs, the LU off-line (red solid line) runs, the PIC 2 (blue dashed line) runs and the PIC 1 (blue solid line) runs. Note that these differences are normalized by $q_0 = \|q(\Upsilon, \Theta, t = 0)\|_\infty$.

the reliability of the ensemble forecast. Although the errors between MSE and MEV of the LU on-line system is larger than the LU offline system from day 5 to day 10, they remain at low level from day 10 onwards, implying that the reliability of the former increases for longer simulation times. In contrast, both PIC 1 and PIC 2 systems show higher error values at most of the times and hence provide less reliable ensembles. We remark that other metrics, such as the continuous ranked probability score (Resseguier et al., 2020; Weigel, 2012), can also be used to measure a calibrated ensemble.

5. Conclusions

In this study, we introduced a stochastic version of the rotating shallow water equations under location uncertainty (RSW-LU). The derivation is based on a stochastic Reynolds transport theorem, where the fluid flow is decomposed into a large-scale component and a noise term modeling the unresolved small-scale flow. A benefit of this approach is that the total energy is conserved along time for any realization. In order to preserve this structure, we combined an energy (in space) preserving discretization of the underlying deterministic equations of this RSW–LU system with approximations of the stochastic terms that are based on standard finite volume/difference operators.

19422466, 2021, 12, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021MS002492 by Cochrane France, Wiley Online Library on [25/10/2023]. 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, 2021, 12, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021MS002492 by Cochrane France, Wiley Online Library on [25/10/2023]. 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

We could show for an f-plane test case that this approach leads for homogeneous noise to a discretization of the RSW–LU system that preserves (spatially) the total energy. Moreover, using inhomogeneous noise that well captures the impact of small scales on the large-scale flow, we demonstrated that for a barotropically unstable jet on the sphere our proposed RSW–LU model better predicts the development of the instabilities than a comparable deterministic model, while the ensemble spread of the RSW–LU system is more likely to contain the observations compared to an ensemble of deterministic simulations with perturbed initial conditions (PIC). We also showed that the RSW–LU forecast systems follows a common distribution of the observations and is more reliable than the PIC system.

Showing accurate ensemble spreads and reliable forecasting skills, we will next apply our developed RSW–LU system to data assimilation. We will also work toward discretizations of stochastic flow models in the framework of LU that preserve total energy both in space and time to which the present work provides a first step. Exploiting the modular approach of combining different discretizations for deterministic and stochastic terms, in future work we will explore the possibility to consistently extend existing atmospheric and ocean models with stochastic parametrizations. We remark that the stochastic approach proposed in this work could be extended to arbitrary Riemannian manifold. In this setting, it would be easier to first convert the Itô integrals to the Stratonovich representations (see Remark 2), and then transform the latter from Euclidean space to other subspaces of Riemannian manifold under diffeomorphism (Hsu, 2000). This application could be helpful for the deep atmosphere component of various global numerical weather prediction and climate models, where the domain significantly differs from Euclidean space.

Appendix A: Energy Conservation of RSW–LU

This appendix demonstrates the energy conservation of the RSW–LU system (Equations 27a–d). Let us recall that the density of the kinetic energy (KE) and of the potential energy (PE) of the large-scale flow in the shallow water system (Vallis, 2017) is, respectively, given by

$$\mathrm{KE} = \int_0^h \frac{\rho_0}{2}|\boldsymbol{u}|^2 \ \mathrm{d}z = \frac{\rho_0}{2} h|\boldsymbol{u}|^2, \tag{A1a}$$

$$\mathrm{PE} = \int_0^h \rho_0 g z \ \mathrm{d}z = \frac{\rho_0}{2} g h^2, \tag{A1b}$$

where $|\boldsymbol{u}|^2 = \boldsymbol{u}\cdot\boldsymbol{u}$ and we assume that $\rho_0 = 1$ and the bottom is flat, that is $\eta_b = 0$ for algebraic simplicity. In order to explain the conservation of energy more concisely, we adopt the following product rule of the stochastic transport operator as derived in Resseguier et al. (2017c). For scalar tracers f, g transported by the stochastic flow and incorporating smooth-in-time external forcings F, G, that is, $\mathbb{D}_t f = F \ \mathrm{d}t$ and $\mathbb{D}_t g = G \ \mathrm{d}t$, we have

$$\mathbb{D}_t(fg) = g\mathbb{D}_t f + f\mathbb{D}_t g. \tag{A2}$$

Applying this rule to the definition of PE (Equation A1b) and using the mass Equation 27b, the PE evolution reads

$$\mathbb{D}_t \mathrm{PE} = gh\mathbb{D}_t h = -gh^2\boldsymbol{\nabla}\cdot\boldsymbol{u} \ \mathrm{d}t = -2\mathrm{PE}\boldsymbol{\nabla}\cdot\boldsymbol{u}\,\mathrm{d}t. \tag{A3a}$$

Similarly, from both mass equation and momentum Equations 27a–27d, noting that $\boldsymbol{u}\cdot(\boldsymbol{f}\times\boldsymbol{u}) = 0$ and recalling that $\eta_b = 0$, we derive the evolution of KE (Equation A1a):

$$\begin{aligned}\mathbb{D}_t \mathrm{KE} \quad &= h\boldsymbol{u}\cdot\mathbb{D}_t\boldsymbol{u} + \frac{1}{2}|\boldsymbol{u}|^2\mathbb{D}_t h \\ &= -\frac{1}{2}\boldsymbol{u}\cdot\boldsymbol{\nabla}\left(gh^2\right) \ \mathrm{d}t - \frac{1}{2}h|\boldsymbol{u}|^2\boldsymbol{\nabla}\cdot\boldsymbol{u} \ \mathrm{d}t = -\left(\boldsymbol{u}\cdot\boldsymbol{\nabla}\mathrm{PE} + \mathrm{KE}\boldsymbol{\nabla}\cdot\boldsymbol{u}\right) \mathrm{d}t.\end{aligned} \tag{A3b}$$

Subsequently, we deduce the evolution of the total energy density $\mathrm{E} = \mathrm{KE} + \mathrm{PE}$,

$$\mathbb{D}_t \mathrm{E} = -\left(\boldsymbol{\nabla}\cdot(\boldsymbol{u} \ \mathrm{PE}) + \mathrm{E}\boldsymbol{\nabla}\cdot\boldsymbol{u}\right) \mathrm{d}t. \tag{A4a}$$

19422466, 2021, 12, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021MS002492 by Cochrane France, Wiley Online Library on [25/10/2023]. 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

Expanding the stochastic transport operator (Equation 6b), and including the incompressible constraints (Equation 27d), the previous equation can be re-written as

$$\mathrm{d}_t \mathrm{E} = -\boldsymbol{\nabla}\cdot\left(\left(\mathrm{E}\ (\boldsymbol{u}-\boldsymbol{u}_s)+\mathrm{PE}\ \boldsymbol{u}-\frac{1}{2}\boldsymbol{a}\boldsymbol{\nabla}\mathrm{E}\right)\ \mathrm{d}t+\mathrm{E}\ \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right). \tag{A4b}$$

Let us now assume some ideal boundary conditions for the resolved and unresolved components:

$$\boldsymbol{u}\cdot\boldsymbol{n}|_{\partial\Omega} = \boldsymbol{u}_s\cdot\boldsymbol{n}|_{\partial\Omega} = \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{n}|_{\partial\Omega} = 0, \tag{A5}$$

where $\partial\Omega$ denotes the boundary of the fluid domain Ω and $\boldsymbol{n}$ stands for the outward pointing unit normal. Combining Equations A4b and A5, one can show that the total energy (integration of energy density over domain) is invariant over time:

$$\mathrm{d}_t\int_\Omega \mathrm{E}(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x} = -\int_{\partial\Omega}\left(\left(\mathrm{E}\ (\boldsymbol{u}-\boldsymbol{u}_s)+\mathrm{PE}\ \boldsymbol{u}-\frac{1}{2}\boldsymbol{a}\boldsymbol{\nabla}\mathrm{E}\right)\ \mathrm{d}t+\mathrm{E}\ \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)\cdot\boldsymbol{n}\ \mathrm{d}l = 0, \tag{A6}$$

in which the following argument is used

$$\boldsymbol{n}\cdot(\boldsymbol{a}\boldsymbol{\nabla}\mathrm{E})\ \mathrm{d}t = \sum_{i,j=1,2} n_i(a_{i,j}\ \mathrm{d}t)\partial_{x_j}\mathrm{E} = \sum_{j=1,2}\mathbb{E}\left[\underbrace{\sum_{i=1,2} n_i(\sigma\mathrm{d}\boldsymbol{B}_t)^i}_{=\ 0\ \text{at}\ \partial\Omega}(\sigma\mathrm{d}\boldsymbol{B}_t)^j\right]\partial_{x_j}\mathrm{E}. \tag{A7}$$

Appendix B: Parameterizations of Noise

This section describes briefly some existing parametrization methods for the noise structure. For interested readers, more technical details can be found in Resseguier et al. (2020).

B1. Homogeneous Noise

From Definitions (Equations 2 and 4), a homogeneous noise means that its correlation operator $\vec{\sigma}$ is a convolution operator and the variance tensor $\vec{a}$ reduces to a constant matrix. To ensure the incompressible constraint (Equation 27d) of a two-dimensional noise Resseguier et al. (2017b) proposed an isotropic model defined through a random stream function

$$\boldsymbol{\sigma}(\boldsymbol{x})\ \mathrm{d}\boldsymbol{B}_t = \nabla^\perp(\breve{\varphi}\star\mathrm{d}B_t)(\boldsymbol{x}), \tag{B1}$$

where $\nabla^\perp = [-\partial_y,\partial_x]^T$ denotes the perpendicular gradient and $\breve{\varphi}\star\mathrm{d}B_t$ stands for the random stream function with a convolution kernel $\breve{\varphi}$ (and the symbol $\star$ denotes a convolution). Both isotropy and incompressibility of the noise result in a (constant) diagonal variance tensor $a_0\vec{\mathrm{I}}_2$ with the eddy-viscosity-like coefficient a_0 and the two-dimensional identity matrix $\vec{\mathrm{I}}_2$. For the current work, the divergence-free constraint of the ISD in Equation 27d is thus naturally satisfied. In practice, the convolution kernel $\breve{\varphi}$ is specified by three parameters: a fixed omni-directional spectrum slope s, a band-pass filter f_{BP} with support in the range of two wavenumbers κ_m and κ_M, and the coefficient a_0. In fact, the Fourier transform of the random stream function $\breve{\varphi}\star\mathrm{d}B_t$ can be defined as:

$$\widehat{\breve{\varphi}\star\mathrm{d}B_t}(\boldsymbol{k}) \overset{\triangle}{=} \frac{A}{\sqrt{\Delta t}} f_{\mathrm{BP}}(\|\boldsymbol{k}\|)\|\boldsymbol{k}\|^{-\alpha}\hat{\xi}_t(\boldsymbol{k})\ \text{ with }\ \alpha=(3+s)/2, \tag{B2}$$

where $\hat{\bullet}$ denotes the Fourier transform coefficient, ξ_t is a space-time white noise, and A is a constant to ensure $\mathbb{E}\|\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\|^2 = 2a_0\Delta t$ (see Equations 3 and 4) with Δt the size of one time stepping and $\mathbb{E}$ the expectation operator. In the simulations, the maximal wavenumber k_M of the noise can usually be chosen as the effective resolution cutoff, the minimal wavenumber can be set to $k_m = k_M/2$, and the theoretical spectrum slope of a two-dimensional flow is given by $s=-3$. Note that on the sphere homogeneous noise could be generated through spherical harmonics.

19422466, 2021, 12, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021MS002492 by Cochrane France, Wiley Online Library on [25/10/2023]. 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

B2. Heterogeneous Noise

In the following, two parameterizations of the heterogeneous noise are presented. These approaches result from the spectral decomposition (Equation 5) used to construct the EOFs of the covariance. However in practice, we work with the Eulerian velocity rather than with the Lagrangian displacement.

B2.1. Off-Line Learning of EOFs

The first method consists in calibrating EOFs from the off-line simulation data with the EOFs assumed to be time-independent. To this end, let us consider a set of velocity snapshots $\{\boldsymbol{u}_{\mathrm{o}}(\boldsymbol{x}, t_i)\}_{i=1,\dots,N_t}$, that have been a priori coarse-grained from high-dimensional data. Applying the singular value decomposition (SVD) for the fluctuations $\boldsymbol{u}'_{\mathrm{o}} = \boldsymbol{u}_{\mathrm{o}} - \bar{\boldsymbol{u}}_{\mathrm{o}}$ (where $\bar{\bullet}$ denotes a temporal average) enables us to build a set of EOFs $\left\{\boldsymbol{\phi}_i\right\}_{i=1,\dots,N_t}$. In addition, we suppose that the fluctuations of the large-scale flow live in a subspace spanned by $\left\{\boldsymbol{\phi}_i\right\}_{i=1,\dots,m-1}$ (with $m < N_t$) and that the small-scale random drift $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t/\Delta t$ lives in the complemented subspace spanned by $\left\{\boldsymbol{\phi}_i\right\}_{i=m,\dots,N_t}$ such that

$$\frac{1}{\Delta t}\boldsymbol{\sigma}(\boldsymbol{x})\ \mathrm{d}\mathbf{B}_t = \sum_{i=m}^{N_t} \sqrt{\lambda_i}\boldsymbol{\phi}_i(\boldsymbol{x})\xi_i, \qquad \frac{1}{\Delta t}\boldsymbol{a}(\boldsymbol{x}) = \sum_{i=m}^{N_t} \lambda_i \boldsymbol{\phi}_i(\boldsymbol{x})\boldsymbol{\phi}_i^T(\boldsymbol{x}), \tag{B3}$$

where λ_i is the eigenvalue associated to the spatial mode $\boldsymbol{\phi}_i$ and ξ_i is a standard Gaussian variable. In practice, there exists an open question in (Equation B3), that is how to adequately choose the "splitting mode" $\boldsymbol{\phi}_m$. Recently Bauer, Chandramouli, Li, et al. (2020) proposed to fix it by comparing the time-averaged energy spectrum of the observations and the one from a coarse-grid deterministic simulation.

B2.2. On-Line Learning of EOFs

The previously described data-driven calibriation of EOFs is a quite efficient procedure. However, such observation data are not always available Bauer, Chandramouli, Chapron, et al. (2020); Resseguier et al. (2020) proposed an alternative approach in which some local fluctuations, called *pseudo-observations* (PSO), are generated directly from a coarse-grid simulation. Then, the SVD is applied on those PSO to estimate a set of EOFs such that the noise associated with its variance tensor will be built in the same way as in (Equation B3). Finally, the magnitude of the noise and variance should be scaled down to smaller scales based on a similarity analysis.

The approach proposed first defines N_o PSO (denoted as $\boldsymbol{u}'$) at each grid point. For a given time t and a current coarse velocity $\boldsymbol{u}$, we build the PSO by sliding a local window of size $N_w \times N_w$ over the spatial grid (with N_w the grid number in one direction of the local window). We denote the spatial scale of the window by $L = N_w l$, where l is the smallest scale of the simulation. At every grid point $\boldsymbol{x}_{i,j}$, we list the N_w^2 velocity values contained in the window centered at that point:

$$\boldsymbol{I}(\boldsymbol{x}_{i,j}, t) \triangleq \left\{ \boldsymbol{u}(\boldsymbol{x}_{p,q}, t) || p - i| \leq \frac{N_w - 1}{2}, |q - j| \leq \frac{N_w - 1}{2} \right\}. \tag{B4}$$

Note that appropriate boundary conditions (replication, periodicity, etc.) are adopted when looking at a point on the border. Then, independently for each $n \in \{1, \dots, N_o\}$ and for each point $\boldsymbol{x}_{i,j}$, we set the value of the PSO $\boldsymbol{u}'(\boldsymbol{x}_{i,j}, t, n)$ by randomly choosing a value in the set $\boldsymbol{I}(\boldsymbol{x}_{i,j}, t)$. After this, we average over the realization index n to build an empirical covariance. Then, from the SVD we obtain a set of EOFs $\left\{\boldsymbol{\phi}_i^{(L)}\right\}_{i=1,\dots,N_o}$, and a spectral representation of the small-scale velocity:

$$\frac{1}{\Delta t}\boldsymbol{\sigma}^{(L)}(\boldsymbol{x}, t)\ \mathrm{d}\boldsymbol{B}_t = \sum_{i=1}^{N_o} \boldsymbol{\phi}_i^{(L)}(\boldsymbol{x}, t)\xi_i. \tag{B5a}$$

Since the PSO $\boldsymbol{u}'$ have been generated at a spatial scale of the window $L = N_w l$, they must be scaled down to the "simulation scale" l. In 3D, according to an auto-similarity assumption of the velocity fluctuations (Kadri Harouna & Mémin, 2017), the small-scale flow $\boldsymbol{\sigma}^{(l)}\mathrm{d}\boldsymbol{B}_t$ associated with its variance tensor $\boldsymbol{a}^{(l)}$ can be rescaled as

19422466, 2021, 12, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021MS002492 by Cochrane France, Wiley Online Library on [25/10/2023]. 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

$$\boldsymbol{\sigma}^{(l)}\mathrm{d}\boldsymbol{B}_t = \left(\frac{l}{L}\right)^{1/3} \boldsymbol{\sigma}^{(L)}\mathrm{d}\boldsymbol{B}_t, \qquad \boldsymbol{a}^{(l)} = \left(\frac{l}{L}\right)^{2/3} \boldsymbol{a}^{(L)}. \tag{B5b}$$

In our case, noting that the small-scale fluctuations are still 3D (even though the vertical component is not known), we keep the same scaling. As shown in Section 4.2, such flow-dependent noise has a good performance in long-term simulation, yet the drawback is that the computational costs are significantly higher compared to the previous off-line procedure, as the SVD is computed at each time step.

Appendix C: Discretization of LU Terms

Starting with a given predicted velocity vector with edge values V_{ij}, we first have to reconstruct the full velocity vector field from these normal values. We use the reconstruction of the vector field in the interior of each triangle proposed by Perot et al. (2006):

$$\boldsymbol{u}_i = \frac{1}{|T_i|} \sum_{k=j,l_,i_} |e_{ik}|(\boldsymbol{x}^{e_{ik}} - \boldsymbol{x}^{T_i})V_{ik}, \tag{C1}$$

where $\boldsymbol{x}^{e_{ik}}$ are the coordinates of the edge midpoint and $\boldsymbol{x}^{T_i}$ are the coordinates of the triangle circumcenter. By averaging values from neighboring triangles, we obtain the corresponding values at the edge midpoints or vertices (see Bauer, 2013 for details).

This reconstructed velocity vector field will be used to generate the noise as described in Appendix B. After the noise has been constructed on the Cartesian mesh, we evaluate the discrete noise vector $(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_{ij}$ and the discrete variance tensor $(\boldsymbol{a})_{ij}$ at the triangle edge midpoints. This information will then be used to calculate the LU noise terms in Equation 42c and 42d.

To calculate the derivatives in these stochastic terms, we use the normal and tangential gradient operators, that is, the gradient operator of Equation 32. To use it, we have to average values, for example, the term $(a_{kl}F)$, to cell centers and vertices and the resulting differential will be an expression located at the edge midpoint. In more detail, we can represent the partial derivative in Cartesian coordinates by

$$(\partial_{x_l}F)_{ij} = (\mathrm{Grad}_n \;\; F)n^l_{ij} + (\mathrm{Grad}_t \;\; F)t^l_{ij}, \qquad l = 1, 2. \tag{C2}$$

Concretely, to discretize (Equation 44), we first compute $(\partial_{x_l}(a_{kl}F))_{ij}$ using Equation C2. The subindex ij indicates that the resulting term is associated to the edge midpoint. To apply the second derivative in Equation 44, that is, $(\partial_{x_k}(\partial_{x_l}(a_{kl}F))_{ij})_{ij}$, we proceed analogously, that is, we first average the terms describing the first derivative to cells and vertices and then apply once more Equation C2. We proceed similarly to represent the term ∇F in Equation 43.

Data Availability Statement

The code to reproduce the results is available at https://github.com/RudigerBrecht/RSW-LU (Brecht et al., 2021). The scripts and data to reproduce the figures can be obtained from https://zenodo.org/record/5576233.

Acknowledgments
The authors acknowledge the support of the Mitacs Globalink Research Award and of the ERC EU project 856408-STUOD. The author Werner Bauer would like to acknowledge funding from NERC NE/R008795/1. Besides, we would like to thank Alexander Bihlo and Scott MacLachlan for helpful discussions and thank Matthias Achtner for providing code to compute the energy spectrum on the sphere. We also gratefully acknowledge the reviewers and the associate editor for their insightful comments and suggestions that helped us to significantly improve this manuscript.

References

Aechtner, M., Kevlahan, N. K.-R., & Dubos, T. (2015). A conservative adaptive wavelet method for the shallow-water equations on the sphere. *Quarterly Journal of the Royal Meteorological Society*, *141*(690), 1712–1726. https://doi.org/10.1002/qj.2473

Anderson, J., & Anderson, S. (1999). A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Monthly Weather Review*, *127*(12), 2741–2758. https://doi.org/10.1175/1520-0493(1999)127<2741:amciot>2.0.co;2

Andrews, D., & McIntyre, M. (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

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

Bauer, W. (2013). *Toward goal-oriented R-adaptive models in geophysical fluid dynamics using a generalized discretization approach* [PhD Thesis, Universität Hamburg]. Berichte zur Erdsystemforschung. https://doi.org/10.17617/2.1668145

Bauer, W., Chandramouli, P., Chapron, B., Li, L., & 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

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., & Gay-Balmaz, F. (2019a). Towards a geometric variational discretization of compressible fluids: The rotating shallow water equations. *Journal of Computational Dynamics*, *6*(1), 1–37. https://doi.org/10.3934/jcd.2019001

Bauer, W., & Gay-Balmaz, F. (2019b). Variational integrators for anelastic and pseudo-incompressible flows. *Journal of Geometric Mechanics*, *11*(4), 511–537. https://doi.org/10.3934/jgm.2019025

Berge, P., Pomeau, Y., & Vidal, C. (1987). *Order within chaos: Towards a deterministic approach to turbulence*. John Wiley & Sons.

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

Bonaventura, L., & Ringler, T. (2005). Analysis of discrete shallow-water models on geodesic delaunay grids with C-type staggering. *Monthly Weather Review*, *133*(8), 2351–2373. https://doi.org/10.1175/mwr2986.1

Brecht, R., Bauer, W., Bihlo, A., Gay-Balmaz, F., & MacLachlan, S. (2019). Variational integrator for the rotating shallow-water equations on the sphere. *Quarterly Journal of the Royal Meteorological Society*, *145*(720), 1070–1088. https://doi.org/10.1002/qj.3477

Brecht, R., Li, L., Bauer, W., & Mémin, E. (2021). Rudigerbrecht/rsw-lu*: First release*. Zenodo. https://doi.org/10.5281/zenodo.4884919

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

Chandramouli, P., Memin, 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

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

Chen, Q., Gunzburger, M., & Ringler, T. (2011). A scale-invariant formulation of the anticipated potential vorticity method. *Monthly Weather Review*, *139*(8), 2614–2629. https://doi.org/10.1175/mwr-d-10-05004.1

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

Desbrun, M., Gawlik, E., Gay-Balmaz, F., & Zeitlin, V. (2014). Variational discretization for rotating stratified fluids. *Discrete & Continuous Dynamical Systems-A*, *34*(2), 477.

Frank, J., Gottwald, G., & Reich, S. (2003). A Hamiltonian particle-mesh method for the rotating shallow-water equations. In M. Griebel, & M. Schweitzer (Eds.), *Meshless methods for partial differential equations* (Vol. 26, pp. 131–142). Springer. https://doi.org/10.1007/978-3-642-56103-0_10

Franzke, C., Majda, A., & Vanden-Eijnden, E. (2006). Low-order stochastic mode reduction for a realistic barotropic model climate. *Journal of the Atmospheric Sciences*, *62*(6), 1722–1757.

Franzke, C. E., & Majda, A. J. (2006). Low-order stochastic mode reduction for a prototype atmospheric GCM. *Journal of the Atmospheric Sciences*, *63*(2), 457–479. https://doi.org/10.1175/jas3633.1

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

Frederiksen, J. S., O'Kane, T. J., & Zidikheri, M. J. (2013). Subgrid modelling for geophysical flows. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *371*(1982), 20120166. https://doi.org/10.1098/rsta.2012.0166

Galewsky, J., Scott, R. K., & Polvani, L. M. (2004). An initial-value problem for testing numerical models of the global shallow-water equations. *Tellus A: Dynamic Meteorology and Oceanography*, *56*(5), 429–440. https://doi.org/10.3402/tellusa.v56i5.14436

Gawlik, E., Mullen, P., Pavlov, D., Marsden, J., & Desbrun, M. (2011). Geometric, variational discretization of continuum theories. *Physica D: Nonlinear Phenomena*, *240*(21), 1724–1760. https://doi.org/10.1016/j.physd.2011.07.011

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

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

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

Gottwald, G., & Harlim, J. (2013). The role of additive and multiplicative noise in filtering complex dynamical systems. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science*, *469*(2155), 20130096. https://doi.org/10.1098/rspa.2013.0096

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

Gugole, F., & Franzke, C. E. (2019). Numerical development and evaluation of an energy conserving conceptual stochastic climate model. *Mathematics of Climate and Weather Forecasting*, *5*(1), 45–64. https://doi.org/10.1515/mcwf-2019-0004

Hairer, E., Lubich, C., & Wanner, G. (2006). *Geometric numerical integration: structure-preserving algorithms for ordinary differential equations* (Vol. 31). Springer Science & Business Media.

Hamill, T. M. (2001). Interpretation of rank histograms for verifying ensemble forecasts. *Monthly Weather Review*, *129*, 550–560. https://doi.org/10.1175/1520-0493(2001)129<0550:iorhfv>2.0.co;2

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

Hecht, M., Holm, D., Petersen, M., & Wingate, B. (2008). Implementation of the Lans-alpha turbulence model in a primitive equation ocean model. *Journal of Computational Physics*, *27*(11), 5691–5716. https://doi.org/10.1016/j.jcp.2008.02.018

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

Hsu, E. (2002). *Stochastic analysis on manifolds*. American Mathematical Society.

Kadri Harouna, S., & 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

Kafiabad, H. A., Vanneste, J., & Young, W. R. (2021). Wave-averaged balance: A simple example. *Journal of Fluid Mechanics*, *911*, R1. https://doi.org/10.1017/jfm.2020.1032

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

Kloeden, P. E., & Platen, E. (1992). *Numerical solution of stochastic differential equations* (Vol. 23). Springer-Verlag Berlin Heidelberg.

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

Leimkuhler, B., & Reich, S. (2004). *Simulating Hamiltonian dynamics*. Cambridge University Press.

19422466, 2021, 12, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021MS002492 by Cochrane France, Wiley Online Library on [25/10/2023]. 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

Leith, C. (1975). Climate response and fluctuation dissipation. *Journal of the Atmospheric Sciences*, *32*(10), 2022–2026. https://doi.org/10.1175/1520-0469(1975)032<2022:crafd>2.0.co;2

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

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., & Eth'e, C. (2012). Grid degradation of submesoscale resolving ocean models: Benefits for offline passive tracer transport. *Ocean Modelling*, *48*(1–2), 1–9. https://doi.org/10.1016/j.ocemod.2012.02.004

Lorenz, E. (1963). Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, *73*(12), 130–141. https://doi.org/10.1175/1520-0469(1963)020<0130:dnf>2.0.co;2

Majda, A., Franzke, C., & Khouider, B. (2008). An applied mathematics perspective on stochastic modelling for climate. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, *366*(1875), 2427–2453. https://doi.org/10.1098/rsta.2008.0012

Majda, A., Timofeyev, I., & Eijnden, E. V. (1999). Models for stochastic climate prediction. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(26), 14687–14691.

Marsden, J. E., & West, M. (2001). Discrete mechanics and variational integrators. *Acta Numerica*, *10*(1), 357–514. https://doi.org/10.1017/s096249290100006x

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

McRae, A. T., & Cotter, C. (2014). Energy-and enstrophy-conserving schemes for the shallow-water equations, based on mimetic finite elements. *Quarterly Journal of the Royal Meteorological Society*, *140*(684), 2223–2234. https://doi.org/10.1002/qj.2291

McWilliams, J., Restrepo, J., & Lane, E. (2004). An asymptotic theory for the interaction of waves and currents in coastal waters. *Journal of Fluid Mechanics*, *511*, 135–178. https://doi.org/10.1017/s0022112004009358

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

Mellor, G. (2016). On theories dealing with the interaction of surface waves and ocean circulation. *Journal of Geophysical Research: Oceans*, *121*(7), 4474–4486. https://doi.org/10.1002/2016jc011768

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

Meneveau, C., & Katz, J. (2000). Scale-invariance and turbulence models for large-eddy simulation. *Annual Review of Fluid Mechanics*, *32*, 1–32. https://doi.org/10.1146/annurev.fluid.32.1.1

Palmer, T., & Williams, P. (2008). Theme issue 'stochastic physics and climate modelling. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *366*, 1875. https://doi.org/10.1098/rsta.2008.0059

Pavlov, D., Mullen, P., Tong, Y., Kanso, E., Marsden, J. E., & Desbrun, M. (2011). Structure-preserving discretization of incompressible fluids. *Physica D: Nonlinear Phenomena*, *240*(6), 443–458. https://doi.org/10.1016/j.physd.2010.10.012

Perot, J. B., Vidovic, D., & Wesseling, P. (2006). Mimetic reconstruction of vectors. In *Compatible spatial discretizations* (pp. 173–188). Springer.

Pope, S. (2000). *Turbulent flows*. Cambridge University Press.

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

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

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

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

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

Ringler, T., & Randall, D. (2002). A potential enstrophy and energy conserving numerical scheme for solution of the shallow-water equations on a geodesic grid. *Monthly Weather Review*, *130*(5), 1397–1410. https://doi.org/10.1175/1520-0493(2002)130<1397:apeaec>2.0.co;2

Salmon, R. (2013). An alternative view of generalized Lagrangian mean theory. *Journal of Fluid Mechanics*, *719*, 165–182. https://doi.org/10.1017/jfm.2012.638

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

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

Slingo, J., & Palmer, T. (2011). Uncertainty in weather and climate prediction. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *369*, 4751–4767. https://doi.org/10.1098/rsta.2011.0161

Talagrand, O., Vautard, R., & Strauss, B. (1997). *Evaluation of probabilistic prediction systems*. ECMWF.

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

Weigel, A. P. (2012). Ensemble forecasts. In *Forecast verification* (pp. 141–166). John Wiley and Sons, Ltd. https://doi.org/10.1002/9781119960003.ch8

Xie, J.-H., & Vanneste, J. (2015). A generalised-Lagrangian-mean model of the interactions between near-inertial waves and mean flow. *Journal of Fluid Mechanics*, *774*, 143–169. https://doi.org/10.1017/jfm.2015.251

Young, W., & Jelloul, M. B. (1997). Propagation of near-inertial oscillation through a geostrophic flow. *Journal of Marine Research*, *55*(4), 735–766. https://doi.org/10.1357/0022240973224283

19422466, 2021, 12, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021MS002492 by Cochrane France, Wiley Online Library on [25/10/2023]. 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