Rotating shallow water flow under location uncertainty with a structure-preserving discretization

Rüdiger Brecht, Long Li, Werner Bauer, Etienne Mémin

Rotating shallow water flow under location uncertainty with a structure-preserving discretization

Rüdiger Brecht[†], **Long Li**[§], **Werner Bauer**[‡], **Etienne Mémin**[§]

[†] *Department of Mathematics and Statistics, Memorial University of Newfoundland, St. John's (NL) A1C 5S7, Canada*

[§] *Inria/IRMAR, Campus universitaire de Beaulieu, Rennes, France*

[‡] *Imperial College London, Department of Mathematics, 180 Queens Gate, London SW7 2AZ, United Kingdom.*

Abstract

We introduce a new representation of the rotating shallow water equations based on a stochastic transport principle. The derivation relies on a decomposition of the fluid flow into a large-scale component and a noise term that models the unresolved small-scale flow. The total energy of such a random model is demonstrated to be preserved along time for any realization. To preserve this structure, 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. This way, our method can directly be used in existing dynamical cores of global numerical weather prediction and climate models. For an inviscid test case on the f-plane we use a homogenous noise and illustrate that the spatial part of the stochastic scheme preserves the total energy of the system. Moreover, using an inhomogenous noise, we show for a barotropically unstable jet on the sphere that the proposed random model better captures the structure of a large-scale flow than a comparable deterministic model.

Plain summary

The motion of geophysical fluids on the globe needs to be modelled to get some insights of tomorrow's climate. These forecasts must be precise enough while remaining computationally affordable. An ideal system should also deliver, across time, an accurate measurement of the uncertainties introduced through physical or numerical approximations. To address these issues, we use the rotating shallow water equations, which provide a simplified version of the dynamics, and a stochastic representation of the unresolved small-scale processes. The former is approximated with a structure preserving numerical model enabling the conservation of physical quantities such as mass and energy and the latter is modelled by the location uncertainty framework that relies on stochastic transport and has the great advantage to be energy conserving. Our method can directly be used in existing dynamical cores of global numerical weather prediction and climate models. Numerical results illustrate the energy conservation of the numerical model. Simulating a barotropically unstable jet on the sphere, we demonstrate that the random model better captures the structure of a large-scale flow than a comparable deterministic model. The random dynamical system is also shown to be associated with good uncertainty representation.

1 Introduction

Numerical simulations of the Earth's atmosphere and ocean plays 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 dynamical system make the large-scale flow structures interact with the smaller components. The computational expense for solving the Kolmogorov scales (Pope, 2000) of a geophysical flows is fare beyond reach today and likely in the future. Thus, the effect of unresolved scales has to be modeled or parametrized.

For several years, there is a growing interest in geophysical sciences to incorporate a stochastic representation (Franzke and Majda, 2006; Majda et al., 2008; Grooms and Majda, 2014; Gottwald et al., 2017) of the small-scale processes. In this study, we propose to stick to a specific stochastic model, the so-called *Location Uncertainty* (LU) derived by Mémin (2014), which emerges from a decomposition of the Lagrangian velocity into a time-smooth drift and a highly oscillating uncertainty term. Such random model allows us to develop by stochastic calculus a new stochastic transport operator (Resseguier et al., 2017a) for the extensive scalars. In particular, this transport operator involves a multiplicative random forcing, a heterogeneous diffusion and a corrected advection resulting from the inhomogeneity of the random flow. This stochastic transport principle has been used as a fundamental tool to derive stochastic representations of large-scale geophysical dynamics (Resseguier et al., 2017a; Chapron et al., 2018; Bauer et al., 2020a). In the present work, we use this mathematical principle together with some physical conservation laws to derive a stochastic version of the rotating shallow water (RSW) system. One strong property of this random model is that it preserves the total energy of the resolved flow in time for each realization.

Recently, the LU model performed very well in Resseguier et al. (2017b,c); Bauer et al. (2020a,b) when studying oceanic quasi-geostrophic flows. 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 the long-term statistics. Besides, Chapron et al. (2018) investigated the Lorentz-63 test case and demonstrated that the LU model was more effective in exploring the range of the strange attractors compared to classical models.

In this work, the performance of the LU model is assessed for the numerical simulation of the RSW system, which can be considered as the first step towards developing numerical random global climate models. This is the first time that the LU model is implemented for the dynamics evolving on the sphere.

We propose to combine the discrete variational integrator for RSW fluids as introduced in (Bauer and Gay-Balmaz, 2019) and (Brecht et al., 2019) with the numerical LU model in order to mimic the continuous conservation properties. 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 and 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. Therefore, they typically outperform traditional integrators if one is interested in long-time integration or the statistical properties of a given dynamical system. The benefit of the proposed method that relies on a combination of a variational integrator with a potentially differently approximated LU model is that it can directly be applied to existing dynamical cores of numerical weather prediction and climate models.

Apart from taking into account the unresolved processes, it is paramount in uncertainty quantification and ensemble forecasting to model the uncertainties along time (Resseguier et al., 2020). For a long time, operational weather forecast centres had relied on random perturbations

of initial conditions (PIC) to spread the ensemble forecasts. However, in the application of data assimilation to geophysical fluid dynamics, such PIC model is known to under-estimate the true uncertainty compared to the observations (Gottwald and Harlim, 2013; Franzke et al., 2015). Hence, an assimilation system is overconfident for such a random model. To overcome this issue, the covariance inflation method (Anderson and Anderson, 1999) is often adopted, in which the ensemble covariance is increased by a carefully tuned parameter. 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 strength of uncertainty.

The remainder of this paper is structured as follows. Section 2 describes the basic principles of the LU model and the derivation of the rotating shallow water system under LU associated with the energy conservation property. Section 3 explains the parameterizations of the uncertainty and 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. Finally, in Section 5 we draw some conclusions and provide an outlook for future work.

2 Rotating shallow water equations under location uncertainty

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

2.1 Location uncertainty principles

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

$$\mathrm{d}\boldsymbol{X}_t = \boldsymbol{w}(\boldsymbol{X}_t, t)\,\mathrm{d}t + \boldsymbol{\sigma}(\boldsymbol{X}_t, t)\,\mathrm{d}\boldsymbol{B}_t, \tag{2.1}$$

where $\boldsymbol{X}$ is the Lagrangian displacement defined within the bounded domain $\Omega \subset \mathbb{R}^d$ ($d = 2$ 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 the small-scale uncertainty (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 $\boldsymbol{f} \in (L^2(\Omega))^d$, with a bounded kernel $\breve{\boldsymbol{\sigma}}$ such that

$$\boldsymbol{\sigma}[\boldsymbol{f}](\boldsymbol{x}, t) = \int_\Omega \breve{\boldsymbol{\sigma}}(\boldsymbol{x}, \boldsymbol{y}, t)\boldsymbol{f}(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}, \qquad \forall \boldsymbol{f} \in (L^2(\Omega))^d. \tag{2.2}$$

The randomness of such noise is driven by the cylindrical $\mathbf{I}_d$-Wiener process $\boldsymbol{B}_t$ (Da Prato and Zabczyk, 2014). The fact that the kernel is bounded, *i.e.* $\sup_{(\boldsymbol{x},\boldsymbol{y})\in\Omega} |\breve{\boldsymbol{\sigma}}(\boldsymbol{x}, \boldsymbol{y})| < +\infty$, implies that the operator $\boldsymbol{\sigma}$ is Hilbert-Schmidt on $(L^2(\Omega))^d$. Therefore, the resulting small-scale flow $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$ is a centered (of null ensemble mean) Gaussian process with the following *covariance tensor*, denoted as $\boldsymbol{Q}$, being well-defined:

$$\begin{aligned}\boldsymbol{Q}(\boldsymbol{x}, \boldsymbol{y}, t, s) &= \mathbb{E}\Big[\big(\boldsymbol{\sigma}(\boldsymbol{x}, t)\,\mathrm{d}\boldsymbol{B}_t\big)\big(\boldsymbol{\sigma}(\boldsymbol{y}, s)\,\mathrm{d}\boldsymbol{B}_s\big)^T\Big] \\ &= \delta(t - s)\,\mathrm{d}t \int_\Omega \breve{\boldsymbol{\sigma}}(\boldsymbol{x}, \boldsymbol{z}, t)\breve{\boldsymbol{\sigma}}^T(\boldsymbol{y}, \boldsymbol{z}, s)\,\mathrm{d}\boldsymbol{z},\end{aligned} \tag{2.3}$$

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

$$\boldsymbol{a}(\boldsymbol{x}, t) \triangleq \boldsymbol{Q}(\boldsymbol{x}, \boldsymbol{x}, t, t)/\mathrm{d}t = \boldsymbol{\sigma}\boldsymbol{\sigma}^T(\boldsymbol{x}, t). \tag{2.4}$$

We remark that such variance tensor $\boldsymbol{a}$ 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 by $\frac{1}{2}\mathrm{tr}(\boldsymbol{a})/\mathrm{d}t$.

The previous representation (2.2) is a general way to define the noise in LU models. In particular, the fact that $\boldsymbol{\sigma}$ is Hilbert-Schmidt ensures that the covariance operator per unit of time, $\boldsymbol{Q}/\mathrm{d}t$, admits an orthogonal eigenfunction basis $\{\boldsymbol{\Phi}_n(\bullet, t)\}_{n\in\mathbb{N}}$ 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:

$$\boldsymbol{\sigma}(\boldsymbol{x}, t)\,\mathrm{d}\boldsymbol{B}_t = \sum_{n\in\mathbb{N}} \boldsymbol{\Phi}_n(\boldsymbol{x}, t)\,\mathrm{d}\beta_t^n, \tag{2.5a}$$

$$\boldsymbol{a}(\boldsymbol{x}, t) = \sum_{n\in\mathbb{N}} \boldsymbol{\Phi}_n(\boldsymbol{x}, t)\boldsymbol{\Phi}_n^T(\boldsymbol{x}, t), \tag{2.5b}$$

where β^n denotes n independent and identically distributed (i.i.d.) one-dimensional standard Brownian motions.

The core of LU models 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 (2.1) within a flow volume $\mathcal{V}$. In particular, for incompressible small-scale flows, $\boldsymbol{\nabla}\cdot\boldsymbol{\sigma} = 0$, the SRTT can be written as

$$\mathrm{d}_t\Big(\int_{\mathcal{V}(t)} q(\boldsymbol{x}, t)\,\mathrm{d}\boldsymbol{x}\Big) = \int_{\mathcal{V}(t)} \big(\mathbb{D}_t q + q\,\boldsymbol{\nabla}\cdot(\boldsymbol{w} - \boldsymbol{w}_s)\big)\,\mathrm{d}\boldsymbol{x}, \tag{2.6a}$$

$$\mathbb{D}_t q \triangleq \mathrm{d}_t q + (\boldsymbol{w} - \boldsymbol{w}_s)\cdot\boldsymbol{\nabla}\, q\,\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}\, q - \frac{1}{2}\,\boldsymbol{\nabla}\cdot(\boldsymbol{a}\boldsymbol{\nabla} q)\,\mathrm{d}t, \tag{2.6b}$$

$$\mathrm{d}_t q \triangleq q_{t+\mathrm{d}t} - q_t, \qquad \boldsymbol{w}_s \triangleq \frac{1}{2}\,\boldsymbol{\nabla}\cdot\boldsymbol{a}, \tag{2.6c}$$

in which the stochastic transport operator $\mathbb{D}_t$ (Resseguier et al., 2017a) and the Itô-Stokes drift (ISD) $\boldsymbol{w}_s$ (Bauer et al., 2020a) are included. The latter term arises from the effect of statistical inhomogeneity of the small-scale flow on the large-scale component, which can be considered as a generalization of the Stokes drift in ocean circulations. In the definition of the stochastic transport operator in (2.6b), the first term on the right-hand side (RHS), defined in (2.6c), stands for a forward time-increment of q at a fixed point $\boldsymbol{x}$, and the last two terms describe, respectively, a backscattering from the small-scales to the large-scales and an inhomogeneous diffusion at the small-scales. In particular, for an isochoric flow with $\boldsymbol{\nabla}\cdot(\boldsymbol{w} - \boldsymbol{w}_s) = 0$, one may immediately deduce from (2.6a) the following transport equation of an extensive scalar:

$$\mathbb{D}_t q = 0, \tag{2.7}$$

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

$$\mathrm{d}_t\Big(\int_\Omega \frac{1}{2}q^2\,\mathrm{d}\boldsymbol{x}\Big) = \Big(\underbrace{\frac{1}{2}\int_\Omega q\,\boldsymbol{\nabla}\cdot(\boldsymbol{a}\boldsymbol{\nabla} q)\,\mathrm{d}\boldsymbol{x}}_{\text{Energy loss by diffusion}} + \underbrace{\frac{1}{2}\int_\Omega (\boldsymbol{\nabla} q)^T\boldsymbol{a}\boldsymbol{\nabla} q\,\mathrm{d}\boldsymbol{x}}_{\text{Energy intake by noise}}\Big)\,\mathrm{d}t = 0. \tag{2.8}$$

Indeed, this can be interpreted as a process where the energy brought by the noise is exactly counter-balanced 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. We remark that a formulation of the shallow water equations under LU in a non-rotating frame is outlined by Mémin (2014), whereas the new model that we present in this work is fully stochastic and includes rotation such that it is suited for simulations of geophysical flows.

The above SRTT (2.6a) and Newton's second principle allow us to derive the following (three-dimensional) stochastic equations of motions in a rotating frame (Resseguier et al., 2017a; Bauer et al., 2020a):

Horizontal momentum equation :

$$\mathbb{D}_t \boldsymbol{u} + \boldsymbol{f} \times \big(\boldsymbol{u}\,\mathrm{d}t + \boldsymbol{\sigma}_H \mathrm{d}\boldsymbol{B}_t\big) = -\frac{1}{\rho}\boldsymbol{\nabla}_H\big(p\,\mathrm{d}t + \mathrm{d}p_t^\sigma\big) + \nu\nabla^2\big(\boldsymbol{u}\,\mathrm{d}t + \boldsymbol{\sigma}_H \mathrm{d}\boldsymbol{B}_t\big), \tag{2.9a}$$

Vertical momentum equation :

$$\mathbb{D}_t w = -\frac{1}{\rho}\partial_z\big(p\,\mathrm{d}t + \mathrm{d}p_t^\sigma\big) - g\,\mathrm{d}t + \nu\nabla^2\big(w\,\mathrm{d}t + \sigma_z \mathrm{d}B_t\big), \tag{2.9b}$$

Mass equation :

$$\mathbb{D}_t \rho = 0, \tag{2.9c}$$

Continuity equation :

$$\boldsymbol{\nabla}\cdot\big(\boldsymbol{w} - \boldsymbol{w}_s\big) = 0, \qquad \boldsymbol{\nabla}\cdot\boldsymbol{\sigma} = 0, \tag{2.9d}$$

where $\boldsymbol{u}$ (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 $\boldsymbol{w}$ (resp. the small-scale 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; $\boldsymbol{\nabla}_H = [\partial_x, \partial_y]^T$ denotes the horizontal gradient; p and $\dot{p}_t^\sigma \triangleq \mathrm{d}p_t^\sigma/\mathrm{d}t$ (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. For the following derivation of the shallow water equations we drop the viscous terms.

In order to model the large-scale circulations in the atmosphere and ocean, the hydrostatic balance approximation is widely adopted (Vallis, 2017). Under a small aspect ratio, $\mathcal{H}/\mathcal{L} \ll 1$ with $\mathcal{L}$ and $\mathcal{H}$ the horizontal and vertical scales of the motion, the acceleration term $\mathbb{D}_t w$ on the left-hand side (LHS) of Equation (2.9b) has a lower order of magnitude than the RHS terms, hence the vertical momentum equation reduces to

$$\partial_z\big(p\,\mathrm{d}t + \mathrm{d}p_t^\sigma\big) = -g\,\mathrm{d}t. \tag{2.10a}$$

According to the Doob's theorem – unique decomposition of a semimartingale process (Kunita, 1997), the previous equation is equivalent to

$$\partial_z p = -\rho g, \qquad \partial_z \mathrm{d}p_t^\sigma = 0. \tag{2.10b}$$

Integrating vertically these hydrostatic balances (2.10b) from 0 to z (see Figure 1) under a constant density ρ_0, we have

$$p(x, y, z, t) = p_0(x, y, t) - \rho_0 g z, \tag{2.10c}$$

$$\mathrm{d}p_t^\sigma(x, y, z, t) = \mathrm{d}p_t^\sigma(x, y, 0, t), \tag{2.10d}$$

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, *i.e.* $p(x, y, \eta, t) \approx 0$ with η the height of the free surface, leading to $p_0 = \rho_0 g \eta$. This allows us to rewrite Equation (2.10c) such that for any $z \in [0, \eta]$ we have

$$p(x, y, z, t) = \rho_0 g\big(\eta(x, y, t) - z\big). \tag{2.10e}$$

Subsequently, the pressure gradient forces in the horizontal momentum equation (2.9a) reduce to

$$-\frac{1}{\rho_0}\boldsymbol{\nabla}_H\big(p\,\mathrm{d}t+\mathrm{d}p_t^\sigma\big)=-g\boldsymbol{\nabla}_H\eta-\frac{1}{\rho_0}\boldsymbol{\nabla}_H\mathrm{d}p_t^\sigma, \tag{2.10f}$$

which do not depend on z according to Equations (2.10e) and (2.10d). Therefore, the acceleration terms on the LHS of Equation (2.9a) must not depend on z, hence the shallow water momentum equation can be written as

$$\mathbb{D}_t^H\boldsymbol{u}+\boldsymbol{f}\times\big(\boldsymbol{u}\,\mathrm{d}t+\boldsymbol{\sigma}_H\mathrm{d}\boldsymbol{B}_t\big)=-g\boldsymbol{\nabla}_H\eta\,\mathrm{d}t-\frac{1}{\rho_0}\boldsymbol{\nabla}_H\mathrm{d}p_t^\sigma, \tag{2.11a}$$

$$\mathbb{D}_t^H u\triangleq\mathrm{d}_t u+\big((\boldsymbol{u}-\boldsymbol{u}_s)\,\mathrm{d}t+\boldsymbol{\sigma}_H\mathrm{d}\boldsymbol{B}_t\big)\cdot\boldsymbol{\nabla}_H u-\frac{1}{2}\boldsymbol{\nabla}_H\cdot\big(\boldsymbol{a}_H\boldsymbol{\nabla}_H u\big)\,\mathrm{d}t, \tag{2.11b}$$

$$\boldsymbol{u}_s\triangleq\frac{1}{2}\boldsymbol{\nabla}_H\cdot\boldsymbol{a}_H,\qquad\boldsymbol{a}=\begin{pmatrix}\boldsymbol{a}_H & \boldsymbol{a}_{Hz}\\ \boldsymbol{a}_{Hz} & a_z\end{pmatrix}, \tag{2.11c}$$

where $\mathbb{D}_t^H$ is the horizontal stochastic transport operator, $\boldsymbol{u}_s$ is the two-dimensional ISD, $\boldsymbol{a}_H$, a_z and $\boldsymbol{a}_{Hz}$ are the horizontal, vertical and cross components of the three-dimensional variance tensor $\boldsymbol{a}$. Note that Equation (2.11a) is valid only when the cross component $\boldsymbol{a}_{Hz}$ is vertically independent, *i.e.* $\partial_z\boldsymbol{a}_{Hz}=0$. For instance, one may consider that the horizontal small-scale flow $\boldsymbol{\sigma}_H\mathrm{d}\boldsymbol{B}_t$ is spatially uncorrelated with the vertical small-scale flow $\sigma_z\mathrm{d}B_t$, *i.e.* $\boldsymbol{a}_{Hz}=0$.

In order to derive the shallow water mass equation, let us first integrate vertically the continuity equation (2.9d) from the bottom topography η_b to the free surface η (see Figure 1):

$$(w-w_s)\big|_{z=\eta}-(w-w_s)\big|_{z=\eta_b}=-h\boldsymbol{\nabla}_H\cdot(\boldsymbol{u}-\boldsymbol{u}_s), \tag{2.12a}$$

$$\sigma\mathrm{d}B_t\big|_{z=\eta}-\sigma\mathrm{d}B_t\big|_{z=\eta_b}=-h\boldsymbol{\nabla}_H\cdot\boldsymbol{\sigma}_H\mathrm{d}\boldsymbol{B}_t, \tag{2.12b}$$

where $h=\eta-\eta_b$ denotes the thickness of the water column. On the other hand, a small vertical (Eulerian) displacement at the top and the bottom of the fluid leads to a variation of the position of a particular fluid element (Vallis, 2017):

$$\big((w-w_s)\,\mathrm{d}t+\sigma\mathrm{d}B_t\big)\big|_{z=\eta}=\mathbb{D}_t^H\eta, \tag{2.12c}$$

$$\big((w-w_s)\,\mathrm{d}t+\sigma\mathrm{d}B_t\big)\big|_{z=\eta_b}=\mathbb{D}_t^H\eta_b. \tag{2.12d}$$

Combining Equations (2.12), we deduce the following stochastic mass equation:

$$\mathbb{D}_t^H h+h\boldsymbol{\nabla}_H\cdot\big((\boldsymbol{u}-\boldsymbol{u}_s)\,\mathrm{d}t+\boldsymbol{\sigma}_H\mathrm{d}\boldsymbol{B}_t\big)=0. \tag{2.13}$$

The above two equations (2.11a) and (2.13) constitute a general formulation of the RSW–LU system. Using different levels of noise strength in such a stochastic system allows us to model different physical regimes of the large-scale flow. To characterise these regimes, Resseguier et al. (2017b) introduced the following scaling number

$$\epsilon=\frac{\mathcal{T}_\sigma}{\mathcal{T}}\frac{TKE}{MKE}, \tag{2.14a}$$

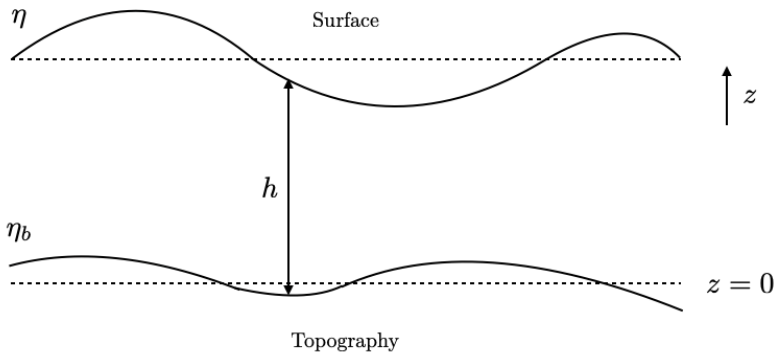


Figure 1. Illustration of a single-layered shallow water system (inspired by Vallis (2017)). h is the thickness of a water column, η is the height of the free surface and η_b is the height of the bottom topography. As a result, we have $h = \eta - \eta_b$.

where $\mathcal{T}$ and $\mathcal{T}_\sigma$ are the correlation time scales of the large-scale flow and the small-scale component, respectively. The mean kinetic energy scale (MKE) is given by $\mathcal{U}^2$ with $\mathcal{U} = \mathcal{L}/\mathcal{T}$ the typical velocity scale, and the turbulent kinetic energy scale (TKE) is defined by $\mathcal{A}/\mathcal{T}_\sigma$ with $\mathcal{A}$ the magnitude of the variance tensor $\boldsymbol{a}$. As such, the dimensional noise associated with its dimensional variance can be specified by

$$\boldsymbol{\sigma}_H \mathrm{d}\boldsymbol{B}_t = \sqrt{\epsilon}\,\mathcal{L}\,(\boldsymbol{\sigma}_H \mathrm{d}\boldsymbol{B}_t)', \qquad \boldsymbol{a} = \epsilon\,\mathcal{U}\mathcal{L}\,\boldsymbol{a}', \tag{2.14b}$$

where $\bullet'$ denotes adimensional variables. From expressions (2.14), one may easily conclude that the greater the scaling number ϵ, the stronger the noise $\boldsymbol{\sigma}_H \mathrm{d}\boldsymbol{B}_t$ (with higher variance $\boldsymbol{a}$). Furthermore, as interpreted in Resseguier et al. (2017c), a strong noise ($\epsilon \gg 1$) modifies the classical geostrophic equilibrium of the large-scale flow by including some correction terms to the isobaric velocities. In the present work, only moderate noise ($\epsilon \sim 1$) is adopted for the RSW–LU system. Under such assumption, the small-scale flow becomes approximately geostrophic and incompressible, *i.e.* $\boldsymbol{f} \times \boldsymbol{\sigma}_H \mathrm{d}\boldsymbol{B}_t \approx -\frac{1}{\rho_0}\boldsymbol{\nabla}_H \mathrm{d}p_t^\sigma$ and $\boldsymbol{\nabla}_H \cdot \boldsymbol{\sigma}_H \mathrm{d}\boldsymbol{B}_t = 0$. As a result, the RSW–LU system simplifies to

$$\mathbb{D}_t^H \boldsymbol{u} + \boldsymbol{f} \times \boldsymbol{u}\,\mathrm{d}t = -g\boldsymbol{\nabla}_H \eta\,\mathrm{d}t, \tag{2.15a}$$

$$\mathbb{D}_t^H h + h\boldsymbol{\nabla}_H \cdot (\boldsymbol{u} - \boldsymbol{u}_s)\,\mathrm{d}t = 0, \tag{2.15b}$$

$$\boldsymbol{\nabla}_H \cdot \boldsymbol{\sigma}_H \mathrm{d}\boldsymbol{B}_t = 0. \tag{2.15c}$$

We remark that an additional incompressible constraint must be imposed on the horizontal ISD, *i.e.* $\boldsymbol{\nabla}\cdot\boldsymbol{u}_s = 0$, so that the previous system preserves the total energy of the large-scale flow. This will be shown in the subsequent section. For the sake of readability, in the following we drop the symbol H for all horizontal variables.

2.3 Energy conservation of RSW–LU

This section demonstrates the energy conservation of the RSW–LU system (2.15). 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

$$\mathrm{KE} = \int_0^h \frac{\rho_0}{2}|\boldsymbol{u}|^2\,\mathrm{d}z = \frac{\rho_0}{2}h|\boldsymbol{u}|^2, \tag{2.16a}$$

$$\mathrm{PE} = \int_0^h \rho_0 g z \,\mathrm{d}z = \frac{\rho_0}{2} g h^2. \tag{2.16b}$$

The density of total energy is defined as the sum of them:

$$\mathrm{E} = \mathrm{KE} + \mathrm{PE} \tag{2.16c}$$

where $|\boldsymbol{u}|^2 = \boldsymbol{u} \cdot \boldsymbol{u}$ and we assume that $\rho_0 = 1$ and the bottom is flat, *i.e.* $\eta_b = 0$ for algebraic simplicity.

In order to explain the conservation of energy more concisely, we adopt an equivalent Stratonovich representation of the RSW–LU system (2.15), namely

$$\mathbb{D}_t \circ \boldsymbol{u} + \boldsymbol{f} \times \boldsymbol{u} \,\mathrm{d}t = -g \boldsymbol{\nabla} h \,\mathrm{d}t, \tag{2.17a}$$

$$\mathbb{D}_t \circ h + h \,\boldsymbol{\nabla} \cdot (\boldsymbol{u} - \boldsymbol{u}_s) \,\mathrm{d}t = 0, \tag{2.17b}$$

$$\boldsymbol{f} \times \boldsymbol{\sigma} \circ \mathrm{d}\boldsymbol{B}_t = -\boldsymbol{\nabla} \mathrm{d}_t \circ p^\sigma, \qquad \boldsymbol{\nabla} \cdot \boldsymbol{\sigma} \circ \mathrm{d}\boldsymbol{B}_t = 0, \tag{2.17c}$$

$$\mathbb{D}_t \circ u \triangleq \mathrm{d}_t \circ u + \big((\boldsymbol{u} - \boldsymbol{u}_s) \,\mathrm{d}t + \boldsymbol{\sigma} \circ \mathrm{d}\boldsymbol{B}_t \big) \cdot \boldsymbol{\nabla} \, u, \tag{2.17d}$$

where $\mathrm{d}_t \circ \triangleq u_{t+dt/2} - u_{t-dt/2}$ stands for a central time-increment based on the definition of Stratonovich integrals and $\mathbb{D}_t \circ$ denotes the stochastic transport operator under Stratonovich notations. We remark that the equivalence between the Itô form (2.11b) and the Stratonovich form (2.17d) are fully detailed in Appendix C of Bauer et al. (2020a). As shown by Kunita (1997), Stratonovich integrals are defined such that the chain rule and the integration-by-part formula of ordinary calculus holds. In particular, for two random tracers f and g, we have

$$\mathrm{d}_t \circ (fg) = f \mathrm{d}_t \circ g + g \mathrm{d}_t \circ f. \tag{2.18a}$$

Therefore, from the definition of the Stratonovich transport operator (2.17d), we deduce the following product rule:

$$\mathbb{D}_t \circ (fg) = g \mathbb{D}_t \circ f + f \mathbb{D}_t \circ g. \tag{2.18b}$$

Applying this rule on the definition of PE (2.16b) together with the mass equation (2.15b),

$$\mathbb{D}_t \circ \mathrm{PE} = g h \mathbb{D}_t \circ h = -g h^2 \,\boldsymbol{\nabla} \cdot (\boldsymbol{u} - \boldsymbol{u}_s) \,\mathrm{d}t, \tag{2.19a}$$

or

$$\mathbb{D}_t \circ \mathrm{PE} + 2\mathrm{PE} \,\boldsymbol{\nabla} \cdot (\boldsymbol{u} - \boldsymbol{u}_s) \,\mathrm{d}t = 0. \tag{2.19b}$$

Similarly, from both mass equation and momentum equation in (2.15), we derive the evolution of KE (2.16a):

$$\begin{aligned} \mathbb{D}_t \circ \mathrm{KE} &= h \boldsymbol{u} \cdot \mathbb{D}_t \boldsymbol{u} + \frac{1}{2} |\boldsymbol{u}|^2 \mathbb{D}_t \circ h \\ &= -\frac{1}{2} \boldsymbol{u} \cdot \boldsymbol{\nabla} \left(g h^2 \right) \mathrm{d}t - \frac{1}{2} h |\boldsymbol{u}|^2 \,\boldsymbol{\nabla} \cdot (\boldsymbol{u} - \boldsymbol{u}_s) \,\mathrm{d}t, \end{aligned} \tag{2.19c}$$

noting that $\boldsymbol{u}\cdot(\boldsymbol{f}\times\boldsymbol{u}\,\mathrm{d}t)=0$ and recalling that $\eta_b=0$, which yields

$$\mathbb{D}_t\circ\mathrm{KE}+\boldsymbol{u}\cdot\boldsymbol{\nabla}\,\mathrm{PE}\,\mathrm{d}t+\mathrm{KE}\,\boldsymbol{\nabla}\cdot(\boldsymbol{u}-\boldsymbol{u}_s)\,\mathrm{d}t=0. \tag{2.19d}$$

Subsequently, we deduce the evolution of the density of total energy:

$$\mathbb{D}_t\circ\mathrm{E}+\boldsymbol{\nabla}\cdot(\boldsymbol{u}\,\mathrm{PE})\,\mathrm{d}t-\mathrm{PE}\,\boldsymbol{\nabla}\cdot\boldsymbol{u}_s\,\mathrm{d}t+\mathrm{E}\,\boldsymbol{\nabla}\cdot(\boldsymbol{u}-\boldsymbol{u}_s)\,\mathrm{d}t=0. \tag{2.20a}$$

Expanding the Stratonovich transport operator (2.17d), the previous equation can be re-written as

$$\mathrm{d}_t\circ\mathrm{E}+\boldsymbol{\nabla}\cdot\big(\boldsymbol{F}_1\,\mathrm{d}t+\boldsymbol{F}_2\circ\mathrm{d}\boldsymbol{B}_t\big)=\mathrm{PE}\,\boldsymbol{\nabla}\cdot\boldsymbol{u}_s\,\mathrm{d}t, \tag{2.20b}$$

where $\boldsymbol{F}_1\triangleq(\boldsymbol{u}-\boldsymbol{u}_s)\,\mathrm{E}+\boldsymbol{u}\,\mathrm{PE}$ and $\boldsymbol{F}_2\circ\mathrm{d}\boldsymbol{B}_t\triangleq\mathrm{E}\,\boldsymbol{\sigma}\circ\mathrm{d}\boldsymbol{B}_t$ are the total energy flux due to the corrected large-scale drift $\boldsymbol{u}-\boldsymbol{u}_s$ and the noise component, respectively. The additional term $\mathrm{PE}\,\boldsymbol{\nabla}\cdot\boldsymbol{u}_s$ stands for sources or sinks of the potential energy due to the compressibility of the ISD. In particular, if we assume that the ISD is incompressible, *i.e.* $\boldsymbol{\nabla}\cdot\boldsymbol{u}_s=0$, the evolution of the energy density reduces to

$$\mathrm{d}_t\circ\mathrm{E}+\boldsymbol{\nabla}\cdot\big(\boldsymbol{F}_1\,\mathrm{d}t+\boldsymbol{F}_2\circ\mathrm{d}\boldsymbol{B}_t\big)=0. \tag{2.21a}$$

If the fluid domain has zero boundary conditions (e.g. the normal velocities vanish on each wall or there are no boundaries at all as on the sphere), then one can show that the total energy, $\overline{\mathrm{E}}\triangleq\int_\Omega\mathrm{E}(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}$, is invariant in time:

$$\mathrm{d}_t\circ\overline{\mathrm{E}}=\int_\Omega\mathrm{d}_t\circ\mathrm{E}\,\mathrm{d}\boldsymbol{x}=-\int_{\partial\Omega}(\boldsymbol{F}_1\,\mathrm{d}t+\boldsymbol{F}_2\circ\mathrm{d}\boldsymbol{B}_t)\cdot\boldsymbol{n}\,\mathrm{d}l=0, \tag{2.21b}$$

where $\partial\Omega$ and $\boldsymbol{n}$ denote the domain's boundaries and the unit normal vector, respectively.

In sum, in this work we propose the following RSW–LU system that preserves the global energy of the large-scale flow in time for any realization of a random noise:

Conservation of momentum :

$$\mathbb{D}_t\boldsymbol{u}+\boldsymbol{f}\times\boldsymbol{u}\,\mathrm{d}t=-g\boldsymbol{\nabla}\eta\,\mathrm{d}t, \tag{2.22a}$$

Conservation of mass :

$$\mathbb{D}_t h+h\,\boldsymbol{\nabla}\cdot\boldsymbol{u}\,\mathrm{d}t=0, \tag{2.22b}$$

Random geostrophic constraint :

$$\boldsymbol{f}\times\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t=-\frac{1}{\rho}\boldsymbol{\nabla}\mathrm{d}p_t^\sigma, \tag{2.22c}$$

Incompressible constraints :

$$\boldsymbol{\nabla}\cdot\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t=0,\qquad\boldsymbol{\nabla}\cdot\boldsymbol{u}_s=0, \tag{2.22d}$$

Conservation of energy :

$$\mathrm{d}_t\int_\Omega\frac{\rho}{2}\big(h|\boldsymbol{u}|^2+gh^2\big)\,\mathrm{d}\boldsymbol{x}=0. \tag{2.22e}$$

Note that for a sufficiently small noise ($\boldsymbol{\sigma}\approx0$), this system (2.22) 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 Discretization of RSW–LU and parametrization of noise

In order to perform a numerical simulation of the RSW–LU (2.22), the noise term $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$ and the variance tensor $\boldsymbol{a}$ have to be *a priori* parametrized. Then an adequate discretization in space-time have to be specified for solving the dynamical system. This section describes these two aspects.

3.1 Parameterizations of noise

In the following, we present two different kinds of spatial structure for the noise – homogeneous and heterogeneous. The first one is easy-to-implement, in particular when considering noise that respects the incompressible constraints (2.22d). We use such homogeneous noise to study the numerical energy behaviour, as shown in Section 4.1. On the other hand, because heterogeneous noise has more physical meaning, we will use the latter when studying realistic complex flows. As shown in Bauer et al. (2020a), heterogeneous noise induces a structuration of the large-scale flow through the inhomogeneity of the small-scale flow. In Section 4.2, such heterogeneous noise is adopted for identifying the barotropic instability of a mid-latitude jet.

3.1.1 Homogeneous noise

From the general definitions (2.2) and (2.4), a homogeneous noise means that its correlation operator $\boldsymbol{\sigma}$ is a convolution operator and the variance tensor $\boldsymbol{a}$ reduces to a constant matrix (independent of any position in the fluid domain). Furthermore, to ensure that a two-dimensional noise is incompressible, Resseguier et al. (2017b) proposed an isotropic model defined through a random stream function

$$\boldsymbol{\sigma}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{B}_t = \boldsymbol{\nabla}^{\perp}\big(\breve{\varphi} \star \mathrm{d}B_t\big)(\boldsymbol{x}), \tag{3.1}$$

where $\boldsymbol{\nabla}^{\perp} = [-\partial_y, \partial_x]^T$ 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). As shown in Resseguier et al. (2017b, 2020), both isotropy and incompressibility of the noise (3.1) result in a (constant) diagonal variance tensor $a_0\mathbf{I}_2$ with the eddy-viscosity-like coefficient a_0 and the two-dimensional identity matrix $\mathbf{I}_2$. In fact, the divergence-free constraint of the ISD in Equation (2.22d) is naturally satisfied (since $\boldsymbol{\nabla}\cdot\boldsymbol{u_s} = \boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\cdot(a_0\mathbf{I}_2) = 0$). As discussed at the end of Section 2.2, for the RSW–LU system (2.22) under *geostrophic noise*, $\boldsymbol{f}\times\boldsymbol{\sigma}_H\mathrm{d}\boldsymbol{B}_t \approx -\boldsymbol{\nabla}_H\mathrm{d}p_t^{\sigma}$, one can identify, for a constant Coriolis parameter f_0, the random pressure $\mathrm{d}p_t^{\sigma}$ with the proposed random stream function by $\mathrm{d}p_t^{\sigma} = \frac{1}{f_0}\breve{\varphi}\star\mathrm{d}B_t$.

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 an eddy-viscosity-like coefficient a_0. In fact, the Fourier transform of the random stream function $\breve{\varphi}\star\mathrm{d}B_t$ can be defined as:

$$\widehat{\breve{\varphi}\star\mathrm{d}B}_t(\boldsymbol{k}) \triangleq \frac{A}{\sqrt{\Delta t}} f_{\mathrm{BP}}\left(\|\boldsymbol{k}\|\right)\|\boldsymbol{k}\|^{-\alpha}\,\widehat{\xi}_t(\boldsymbol{k}) \text{ with } \alpha = (3+s)/2, \tag{3.2}$$

where $\widehat{\bullet}$ denotes the Fourier transform coefficient, ξ_t is a space-time white noise, and A is a constant to ensure $\mathbb{E}\left\|\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right\|^2 = 2a_0\Delta t$ (see Equations (2.3) and (2.4)) with Δt the size of one time stepping. 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 $s = -3$. The noise strength parameter a_0 will be specified in Section 4.1.

3.1.2 Heterogeneous noise

The homogeneous noise defined in Section 3.1.1 is quite simple to construct and to interpret, however, it lacks to represent physically important contributions of the small-scale to the large scale flow, which is crucial in order to accurately model realistic scenarios in geophysical fluid dynamics. For this reason, two parameterizations of the heterogeneous noise are presented in the following.

These approaches result from the spectral decomposition (2.5) used to construct the eigenfunction basis of the spatial covariance. In practice, we work with a finite set of Empirical

Orthogonal Functions (EOFs) of the small-scale Eulerian velocity rather than with the Lagrangian displacement. The first method for estimating the EOFs is an off-line procedure based on the Proper Orthogonal Decomposition (POD) technique of high-dimensional data in which the EOFs are assumed to be time-independent, whereas the second one is an on-line estimation from a coarse-grid simulation where the EOFs are time-dependent. As will be shown in Section 4.2, the former allows for incorporating data into the dynamical model and is more suitable for mid-term simulations, yet the latter is independent from observations and is more adequate for long-term simulations.

Off-line learning of EOFs 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 using a low-pass filter (such as the sharp spectral filter of Pope (2000) often used in large eddy simulations). Applying the snapshot POD procedure (Sirovich, 1987) for the fluctuations $\boldsymbol{u}'_{\mathrm{o}} = \boldsymbol{u}_{\mathrm{o}} - \overline{\boldsymbol{u}}_{\mathrm{o}}$ (where $\overline{\bullet}$ denotes a temporal average) enables us to build a set of EOFs $\{\boldsymbol{\phi}_i\}_{i=1,\dots,N_t}$. In addition, we suppose that the fluctuations of the large-scale flow live in a subspace spanned by $\{\boldsymbol{\phi}_i\}_{i=1,\dots,m-1}$ (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 $\{\boldsymbol{\phi}_i\}_{i=m,\dots,N_t}$ such that

$$\frac{1}{\Delta t}\boldsymbol{\sigma}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{B}_t = \sum_{i=m}^{N_t} \sqrt{\lambda_i}\boldsymbol{\phi}_i(\boldsymbol{x})\xi_i, \qquad \frac{1}{\Delta t}\boldsymbol{a}(\boldsymbol{x}) = \sum_{i=m}^{N_t} \lambda_i \boldsymbol{\phi}_i(\boldsymbol{x})\boldsymbol{\phi}_i^T(\boldsymbol{x}), \tag{3.3}$$

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 opening question in (3.3), that is how to adequately choose the "splitting mode" $\boldsymbol{\phi}_m$. Recently, Bauer et al. (2020b) proposed to fix it by comparing the time-averaged energy spectrum of the observations and the one from a coarse-grid deterministic simulation.

On-line learning of EOFs The previously described data-driven estimation of EOFs is a quite efficient procedure. However, such observation data, either from direct measurements or from high-dimensional simulations, are not always available. Therefore, Bauer et al. (2020a); 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 singular value decomposition (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 (3.3). Finally, the magnitude of the noise and variance should be scaled down to smaller scales based on a similarity analysis (Kadri Harouna and Mémin, 2017).

In the following, we describe in more details both the generation of PSO and the scaling technique. The approach proposed here 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:

$$I(\boldsymbol{x}_{i,j}, t) \triangleq \left\{ \boldsymbol{u}(\boldsymbol{x}_{p,q}, t) \middle| |p-i| \leq \frac{N_w - 1}{2}, |q-j| \leq \frac{N_w - 1}{2} \right\}. \tag{3.4}$$

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 $I(\boldsymbol{x}_{i,j}, t)$. After this, we average over the realization index n. Then, from the SVD we obtain a

set of EOFs $\{\boldsymbol{\phi}_i^{(L)}\}_{i=1,\dots,N_o}$, and a spectral representation of the small-scale velocity:

$$\frac{1}{\Delta t}\boldsymbol{\sigma}^{(L)}(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{B}_t = \sum_{i=1}^{N_o}\boldsymbol{\phi}_i^{(L)}(\boldsymbol{x},t)\xi_i. \tag{3.5a}$$

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. As such, the variance tensor $\boldsymbol{a}$ of the small-scale flow is rescaled according to a turbulence-power-law coefficient(Kadri Harouna and Mémin, 2017) such that

$$\boldsymbol{a}^{(l)} = \left(\frac{l}{L}\right)^{2/3}\boldsymbol{a}^{(L)}, \tag{3.5b}$$

where $\boldsymbol{a}^{(L)}$ and $\boldsymbol{a}^{(l)}$ are the variance tensors at the scales L and l respectively. Finally, the small-scale flow can be simulated at the "simulation scale" l as

$$\boldsymbol{\sigma}^{(l)}\mathrm{d}\boldsymbol{B}_t = \left(\frac{l}{L}\right)^{1/3}\boldsymbol{\sigma}^{(L)}\mathrm{d}\boldsymbol{B}_t. \tag{3.5c}$$

As will be 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.

3.2 Structure-preserving discretization of RSW–LU

In this subsection, we introduce an energy conserving (in space) approximation of the above derived stochastic system. Considering the definition of the stochastic transport operator $\mathbb{D}_t$ defined in (2.6b) with a time increment $\mathrm{d}_t q \triangleq q_{t+\mathrm{d}t} - q_t$ defined in (2.6c), the RSW–LU system in Eqn. (2.22a)–(2.22b) can be explicitly written as

$$\mathrm{d}_t\boldsymbol{u} = \Big(-\boldsymbol{u}\cdot\boldsymbol{\nabla}\,\boldsymbol{u} - \boldsymbol{f}\times\boldsymbol{u} - g\boldsymbol{\nabla}\eta\Big)\,\mathrm{d}t + \Big(\frac{1}{2}\,\boldsymbol{\nabla}\cdot\,\boldsymbol{\nabla}\cdot(\boldsymbol{a}\boldsymbol{u})\,\mathrm{d}t - \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}\,\boldsymbol{u}\Big), \tag{3.6a}$$

$$\mathrm{d}_t h = -\,\boldsymbol{\nabla}\cdot\,(\boldsymbol{u}h)\,\mathrm{d}t + \Big(\frac{1}{2}\,\boldsymbol{\nabla}\cdot\,\boldsymbol{\nabla}\cdot(\boldsymbol{a}h)\,\mathrm{d}t - \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}\,h\Big). \tag{3.6b}$$

We suggest to develop an approximation of the stochastic RSW–LU model (3.6a)–(3.6b) 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. The *deterministic dynamical core* of our stochastic system results from simply setting $\boldsymbol{\sigma} \approx 0$ in the equations (3.6a)–(3.6b). 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 not necessarily share the same operators as the deterministic scheme has the advantage that our method can be directly applied to existing dynamical cores of global numerical weather prediction (NWP) and climate models.

3.2.1 Structure-preserving discretization of the 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 (2.4) to $\boldsymbol{a} \approx 0$. Hence, Equations (3.6a)–(3.6b) reduce to the deterministic RSW equations

$$\mathrm{d}_t\boldsymbol{u} = \Big(-(\boldsymbol{\nabla}\times\boldsymbol{u}+\boldsymbol{f})\times\boldsymbol{u} - \boldsymbol{\nabla}(\frac{1}{2}\boldsymbol{u}^2) - g\boldsymbol{\nabla}\eta\Big)\,\mathrm{d}t, \qquad \mathrm{d}_t h = -\,\boldsymbol{\nabla}\cdot\,(\boldsymbol{u}h)\,\mathrm{d}t, \tag{3.7}$$

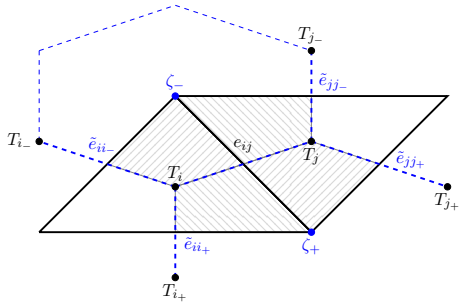


Figure 2. Notation and indexing conventions for the 2D simplicial mesh.

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 $\partial/\partial t$.

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 and Gay-Balmaz (2019); Brecht et al. (2019), here we only give the final, fully discrete scheme.

To do so, we start with introducing the mesh and some notation. The variational discretization of (3.7) 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 Fig. 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 circumcentre of T_i, η_{bi} the discrete bottom topography at the circumcentre 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 $\overline{D}_{ij} = \frac{1}{2}(D_i + D_j)$ 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:

$$\begin{aligned} (\mathrm{Grad}_n\ F)_{ij} &\triangleq \frac{F_{T_j} - F_{T_i}}{|\tilde{e}_{ij}|}, & (\mathrm{Div}\ V)_i &\triangleq \frac{1}{|T_i|} \sum_{k \in \{j, i_-, i_+\}} |e_{ik}| V_{ik}, \\ (\mathrm{Grad}_t\ F)_{ij} &\triangleq \frac{F_{\zeta_-} - F_{\zeta_+}}{|e_{ij}|}, & (\mathrm{Curl}\ V)_\zeta &\triangleq \frac{1}{|\zeta|} \sum_{\tilde{e}_{nm} \in \partial\zeta} |\tilde{e}_{nm}| V_{nm}, \end{aligned} \tag{3.8}$$

for the normal velocity V_{ij} and a scalar function F either sampled as F_{T_i} at the circumcentre of the triangle T_i or sampled as $F_{\zeta_\pm}$ at the centre 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:

$$(\nabla F)_{ij} \approx (\mathrm{Grad}_n\ F)\mathbf{n}_{ij} + (\mathrm{Grad}_t\ F)\mathbf{t}_{ij}, \tag{3.9}$$

$$(\nabla \cdot \mathbf{u})_i \approx (\mathrm{Div}\ V)_i, \tag{3.10}$$

$$(\nabla \times \mathbf{u})_\zeta \approx (\mathrm{Curl}\ V)_\zeta. \tag{3.11}$$

The last Equation (3.11) defines the discrete vorticity and for later use, we also discretize the

potential vorticity as

$$\frac{\nabla \times \mathbf{u} + f}{h} \approx \frac{(\text{Curl } V)_\zeta + f_\zeta}{D_\zeta}, \qquad D_\zeta = \sum_{\tilde{e}_{ij} \in \partial\zeta} \frac{|\zeta \cap T_i|}{|\zeta|} D_i. \tag{3.12}$$

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

$$\mathrm{d}_t V_{ij} = \mathcal{L}^V_{ij}(V, D)\,\Delta t, \quad \text{for all edges } e_{ij}, \tag{3.13a}$$

$$\mathrm{d}_t D_i = \mathcal{L}^D_i(V, D)\,\Delta t, \quad \text{for all cells } T_i, \tag{3.13b}$$

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 (3.13a) is given by

$$\mathcal{L}^V_{ij}(V, D) \triangleq -\text{Adv}(V, D)_{ij} - \text{K}(V)_{ij} - \text{G}(D)_{ij}, \tag{3.14}$$

where Adv denotes the discretization of the advection term $(\boldsymbol{\nabla} \times \boldsymbol{u} + \boldsymbol{f}) \times \boldsymbol{u}$ of (3.7), 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 $g\boldsymbol{\nabla}\eta$. Explicitly, the advection term is given by

$$\begin{aligned}
&\text{Adv}(V, D)_{ij} \triangleq \\
&- \frac{1}{\overline{D}_{ij}|\tilde{e}_{ij}|}\Big((\text{Curl } V)_{\zeta_-} + f_{\zeta_-}\Big)\left(\frac{|\zeta_- \cap T_i|}{2|T_i|}\overline{D}_{ji_-}|e_{ii_-}|V_{ii_-} + \frac{|\zeta_- \cap T_j|}{2|T_j|}\overline{D}_{ij_-}|e_{jj_-}|V_{jj_-}\right) \\
&+ \frac{1}{\overline{D}_{ij}|\tilde{e}_{ij}|}\Big((\text{Curl } V)_{\zeta_+} + f_{\zeta_+}\Big)\left(\frac{|\zeta_+ \cap T_i|}{2|T_i|}\overline{D}_{ji_+}|e_{ii_+}|V_{ii_+} + \frac{|\zeta_+ \cap T_j|}{2|T_j|}\overline{D}_{ij_+}|e_{jj_+}|V_{jj_+}\right),
\end{aligned} \tag{3.15}$$

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

$$\text{K}(V)_{ij} \triangleq \frac{1}{2}(\text{Grad}_n\ F)_{ij}, \qquad F_{T_i} = \sum_{k \in \{j, i_-, i_+\}} \frac{|\tilde{e}_{ik}|\ |e_{ik}|(V_{ik})^2}{2|T_k|}, \tag{3.16}$$

$$\text{G}(D)_{ij} \triangleq g(\text{Grad}_n\ (D + \eta_b))_{ij}. \tag{3.17}$$

The RHS of the continuity equation (3.13b) is given by

$$\mathcal{L}^D_i(V, D) \triangleq -\big(\text{Div } (\overline{D}V)\big)_i, \tag{3.18}$$

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

Stabilization. In addition, as often used in the simulations of large-scale atmospheric and oceanic flows, in order to stabilize the numerical solution (which will be important for the stochastic model), we include a biharmonic eddy viscosity with uniform coefficient μ (of unit m^4/s) in the momentum equation:

$$\mathrm{d}_t V = \Big(- \text{Adv}(V, D)_{ij} - \text{K}(V)_{ij} - \text{G}(D)_{ij} - \mu L(V)_{ij}\Big)\,\Delta t, \tag{3.19}$$

where:

$$L(V)_{ij} = \big(\text{Grad}_n(\text{Div } V)_{ij} - \text{Grad}_t(\text{Curl } V)_{ij}\big)^2. \tag{3.20}$$

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 $\boldsymbol{\sigma} = 0$ and $\mu = 0$ with the one given in Section 3.2.3.

3.2.2 Spatial discretization of RSW–LU

The fully stochastic system has additional terms on the RHS of Equations (3.6a) and (3.6b). With these terms the discrete equations read:

$$\mathrm{d}_t V_{ij} = \mathcal{L}^V_{ij}(V, D)\,\Delta t + \Delta\mathcal{G}^V_{ij}, \tag{3.21a}$$

$$\mathrm{d}_t D_i = \mathcal{L}^D_i(V, D)\,\Delta t + \Delta\mathcal{G}^D_i, \tag{3.21b}$$

where the stochastic LU-terms are given by

$$\Delta\mathcal{G}^V_{ij} \triangleq \Big(-\frac{\Delta t}{2}\big(\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\cdot(\boldsymbol{a}\boldsymbol{u})\big)_{ij} + (\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}\,\boldsymbol{u})_{ij}\Big)\cdot\boldsymbol{n}_{ij}, \tag{3.21c}$$

$$\Delta\mathcal{G}^D_i \triangleq -\frac{\Delta t}{2}\big(\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\cdot(\boldsymbol{a}D)\big)_i + (\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}\,D)_i. \tag{3.21d}$$

Note that the two terms within the large bracket in (3.21c) 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 (3.21d) are scalar valued at the cell circumcenters i.

The parametrization of the noise described in Section 3.1 is formulated in Cartesian coordinates, because this allows using standard algorithms to calculate e.g. EOFs and POD. 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 (3.21c) as

$$(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}\,F)_{ij} = \sum_{l=1}^{2}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)^l_{ij}(\nabla F)^l_{ij}, \tag{3.22}$$

in which $(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_{ij}$ denotes the discrete noise vector with two Cartesian components, constructed as described in Section 3.1 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 (3.21c) can be written component-wise as

$$(\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\cdot(\boldsymbol{a}F))_{ij} = \sum_{k,l=1}^{2}\Big(\partial_{x_k}\big(\partial_{x_l}(a_{kl}F)\big)_{ij}\Big)_{ij}, \tag{3.23}$$

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 (3.9) as introduced next.

To calculate the terms in (3.21d) we also use the representations (3.22) and (3.23) 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.

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, 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 and the dynamical core via the midpoints of the triangle edges where on such C-grid staggered discretizations the velocity values naturally reside.

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):

$$\boldsymbol{u}_i = \frac{1}{|T_i|} \sum_{k=j,i_-,i_+} |e_{ik}|(\boldsymbol{x}^{e_{ik}} - \boldsymbol{x}^{T_i})V_{ik}, \tag{3.24}$$

where $\boldsymbol{x}^{e_{ik}}$ are the coordinates of the edge midpoint and $\boldsymbol{x}^{T_i}$ are the coordinates of the triangle circumcentre. 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 Section 3.1. 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 (3.21c) and (3.21d).

To calculate the derivatives in these stochastic terms, we use the normal and tangential gradient operators, i.e. the gradient operator of (3.9). To use it, we have to average values, e.g. 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

$$(\partial_{x_l} F)_{ij} = (\mathrm{Grad}_n\ F)n^l_{ij} + (\mathrm{Grad}_t\ F)t^l_{ij}, \qquad l = 1, 2. \tag{3.25}$$

Concretely, to discretize (3.23), we first compute $(\partial_{x_l}(a_{kl}F))_{ij}$ using Equation (3.25). The subindex ij indicates that the resulting term is associated to the edge midpoint. To apply the second derivative in (3.23), i.e. $\left(\partial_{x_k}\left(\partial_{x_l}(a_{kl}F)\right)_{ij}\right)_{ij}$, we proceed analogously, i.e. we first average the terms describing the first derivative to cells and vertices and then apply once more Equation (3.25). We proceed similarly to represent the term ∇F in (3.22).

As mentioned above, the terms in (3.21d) are calculated similarly to (3.21c) with the only difference that the former have to be averaged to the cell centers after the proposed discretization procedure has been applied.

3.2.3 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 and 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:

Time-stepping algorithm

1. Start loop over $k = 0$ with initial guess at t : $V^*_{k=0} = V^t$ and $(D^*_{k=0})_i = D^t_i + \Delta\mathcal{G}^D_{ij}(D^t)$. Besides, we compute $\Delta\mathcal{G}^V_{ij}(V^t)$.

2. Update water depth D^*_{k+1} and velocity V^*_{k+1} using explicit equation:

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

and set $k+1=k$.

3. Stop loop if $\|V^*_{k+1}-V^*_k\|+\|D^*_{k+1}-D^*_k\|<$ tolerance.

For all simulations in this manuscript, we used a tolerance of 10^{-6} for simulations on the plane and 10^{-10} for simulation on the sphere.

This algorithm will be used in the next section to evolve the fluid flow in time.

4 Numerical results

In this section, we first study the energy behaviour of the numerical RSW–LU scheme from above for an inviscid test flow. Then, we show that for a viscous test flow, the stochastic model captures more accurately the referent 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 initial condition (PIC).

4.1 Inviscid test case – energy analysis

This first test case consists of two co-rotating vortices on the f-plane without viscosity (i.e. $\mu=0$). To illustrate the energy conservation of the spatial discretization of the RSW–LU system (2.22), we use the homogeneous stationary noise defined in Section 3.1.1 since the two incompressible constraints $\boldsymbol{\nabla\cdot\sigma}\mathrm{d}\boldsymbol{B}_t=0$ and $\boldsymbol{\nabla\cdot\nabla\cdot a}=0$ in (2.22d) are naturally satisfied. Then, no extra steps are required to satisfy the incompressible constraints.

Initial conditions

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

$$h\big(x,y,t=0\big)=H_0-H'\bigg(\exp\Big(-\frac{{x'_1}^2+{y'_1}^2}{2}\Big)+\exp\Big(-\frac{{x'_2}^2+{y'_2}^2}{2}\Big)-\frac{4\pi s_x s_y}{L_xL_y}\bigg),\tag{4.1a}$$

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 are given by

$$x'_i=\frac{L_x}{\pi s_x}\sin\big(\frac{\pi}{L_x}(x-x_{c_i})\big),\qquad y'_i=\frac{L_y}{\pi s_y}\sin\big(\frac{\pi}{L_y}(y-y_{c_i})\big),\qquad i=1,2\tag{4.1b}$$

with the centres 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

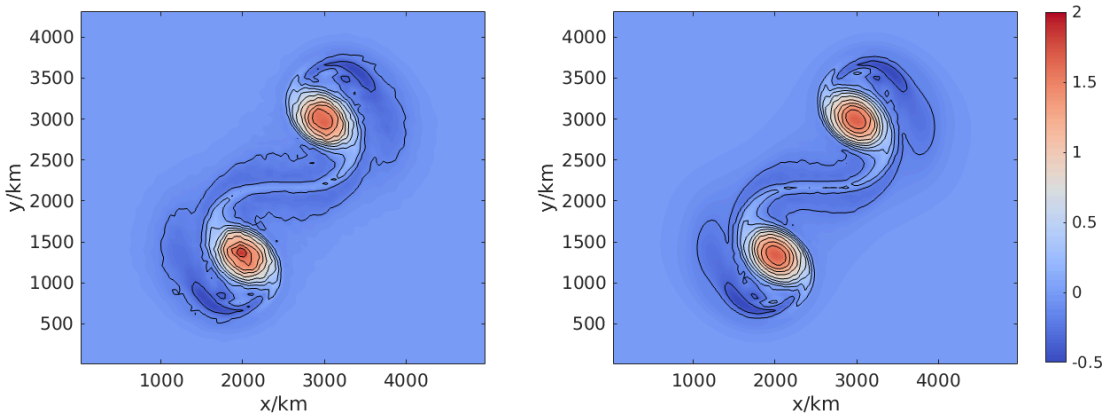


Figure 3. Contour plots of the potential vorticity fields after 2 days for (left) one realization of a LU simulation with homogeneous noise and (right) a deterministic run. The contour interval is 0.4 days^{-1} km^{-1}.

analytical function h at each cell centre. Subsequently, the discrete geostrophic velocities at each triangle edge ij at the initial state can be deduced via

$$V_{ij} = -\frac{g}{f}(\mathrm{Grad}_t\ D)_{ij}, \tag{4.2}$$

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·s^{-1}.

Analysis of energy conservation

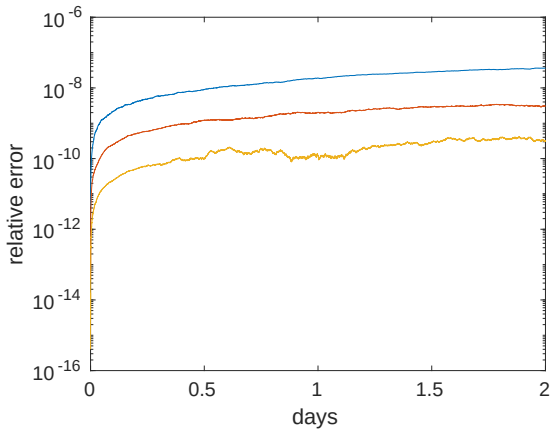
To analyze the energy conservation properties of our stochastic integrator, we use the above initial conditions to simulate the two co-rotating vortices for 2 days. In Figure 3, we show contour plots of the potential vorticity (as defined in (3.12)) fields of the deterministic and stochastic models. 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 (2.16) for both deterministic and stochastic case is approximated by

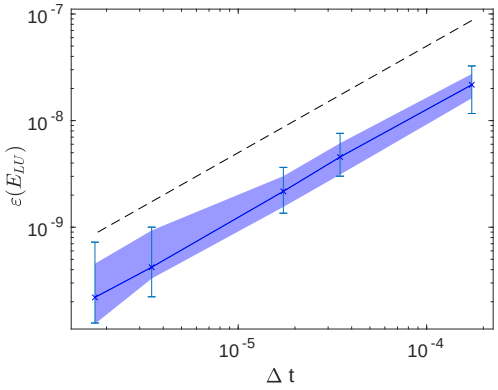
$$\mathrm{E}(t) \approx \sum_{i=1}^{N} \frac{1}{2} D_i(t)|T_i| \sum_{k=j,i_-,i_+} \frac{1}{2|T_i|} h_{ik} f_{ik} \big(V_{ik}(t)\big)^2 + \frac{1}{2} g \big(D_i(t)\big)^2 |T_i|. \tag{4.3}$$

As shown in Bauer and Gay-Balmaz (2019), the proposed discrete variational integrator (see Section 3.2.1) 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 3.2.2).

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



(a) 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.



(b) 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.

Figure 4. Analysis of the numerical energy conservation of the RSW–LU.

the mean stochastic energy, denoted as $\overline{\mathrm{E}}_{\mathrm{LU}}(t)$, and the reference $\mathrm{E}_{\mathrm{REF}}(t)$ by $\overline{\mathrm{E}}_{\mathrm{LU}}(t)/\mathrm{E}_{\mathrm{REF}}(t) - 1$. This setup allows us to measure the influence of the stochastic terms on the energy conservation relative to the determinitic scheme. Figure 4a shows these relative errors for different time step sizes over a simulation time of 2 days. As we can deduce from the curves, taking successively smaller time steps Δt 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:

$$\varepsilon(\mathrm{E}_{\mathrm{LU}}) \triangleq \frac{\|\mathrm{E}_{\mathrm{LU}}(t) - \mathrm{E}_{\mathrm{REF}}(t)\|_{L^2([0,T])}}{\|\mathrm{E}_{\mathrm{REF}}(t)\|_{L^2([0,T])}}, \tag{4.4}$$

where $\|f(t)\|_{L^2([0,T])} = (\int_0^T |f(t)|^2 \mathrm{d}t)^{1/2}$ 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 4b. The area centre 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 $\varepsilon(\mathrm{E}_{\mathrm{LU}})$). Besides, the minimal and maximal values of the errors of the ensemble are represented by the vertical bar-plots. The blue line of Figure 4b 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 3.2.3.

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, 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.

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

$$u(\Theta, t=0) = \frac{U_0}{e_n} \exp\Big(\frac{1}{(\Theta-\Theta_0)(\Theta-\Theta_1)}\Big), \qquad \text{for } \Theta_0 < \Theta < \Theta_1, \tag{4.5a}$$

$$h(\Theta, t=0) = H_0 - \frac{R}{g}\int_{\Theta} u(\theta, t=0)\Big(2\tilde{\Omega}\sin\theta + \frac{\tan\theta}{R}u(\theta, t=0)\Big)\,\mathrm{d}\theta, \tag{4.5b}$$

where $e_n = \exp\big(-4/(\Theta_1-\Theta_0)^2\big)$ 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:

$$h'(\Upsilon, \Theta) = H' \cos\Theta \exp\Big(-(3\Upsilon)^2 - \big(15(\frac{\pi}{4}-\Theta)\big)^2\Big), \tag{4.5c}$$

where Υ denotes the longitude. Analogously to the previous inviscid test case, we then use these analytic functions (4.5) to sample the discrete velocity at the edge mid-point and the height field at the cell centre on the staggered mesh (See Figure 2).

Parameters	Value	Description
(Θ_0, Θ_1)	$(2\pi, 5\pi)/14\,\mathrm{rad}$	Initial latitude limits
H_0	$10.158\,\mathrm{km}$	Background height
H'	$120\,\mathrm{m}$	Initial perturbation amplitude
R	$6.371 \times 10^3\,\mathrm{km}$	Mean radius of Earth
g	$9.806\,\mathrm{m}\cdot\mathrm{s}^{-2}$	Gravity of Earth
$\tilde{\Omega}$	$7.292 \times 10^{-5}\,\mathrm{s}^{-1}$	Angular rotation rate of Earth
U_0	$80\,\mathrm{m}\cdot\mathrm{s}^{-1}$	Maximum zonal velocity
μ_l	$3.975 \times 10^{14}\,\mathrm{m}^4\cdot\mathrm{s}^{-1}$	Fine-grid biharmonic viscosity
μ_L	$3.199 \times 10^{16}\,\mathrm{m}^4\cdot\mathrm{s}^{-1}$	Coarse-grid biharmonic viscosity
Δt_l	$12\,\mathrm{s}$	Fine-grid time step
Δt_L	$50\,\mathrm{s}$	Coarse-grid time step
N_l	327680	Number of triangles for fine grid
N_L	20480	Number of triangles for coarse grid

Table 1. Parameter list for simulations of the barotropic instability.

For the LU simulations, we use the two heterogeneous noises described in Section 3.1.2, 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 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 ensemble realization is used as one initial random state for the *PIC off-line* and the *PIC on-line* simulations, respectively. For each

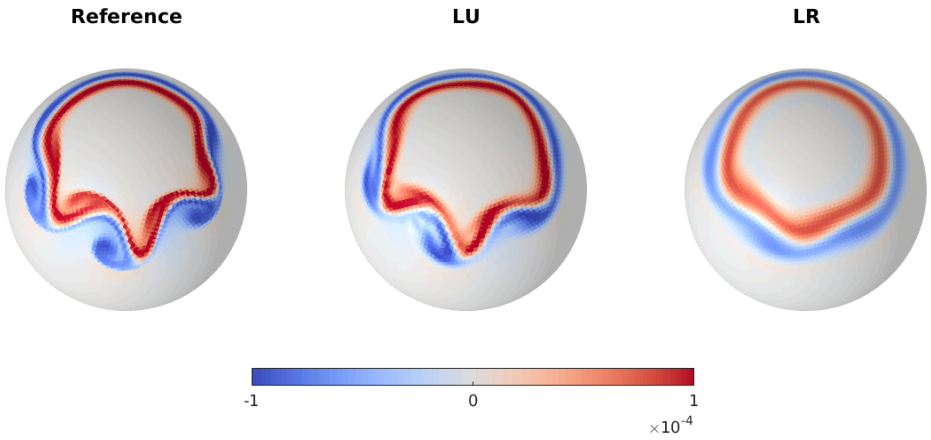


Figure 5. Snapshots of the vorticity field on the sphere for different models (with 20480 triangles) after 5 days. From left to right: reference, ensemble mean of LU online and deterministic LR.

stochastic model, an ensemble run with 20 realizations is done. Besides, a deterministic coarse-grid simulation, denoted as *LR*, is also performed. For all these coarse models, the biharmonic viscosity coefficient is fixed to be the same as given in Table 1.

Prediction of barotropic instability

In this section, we compare the predictions of the barotropic instability for different coarse models to that provided by the reference simulation. The latter is obtained from the coarse-graining procedure through a bilinear interpolation of the high-resolution snapshots. In Figure 5, 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 at that day 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 6. 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 off-line and LR simulations. We remark that the results of PIC on-line simulations are not include in Figure 6, since they behave quite similarly to the PIC off-line run.

To physically interpret the above results, it is useful to analyze the energy spectra of 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 and Zanna, 2017). 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 7. 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 better structuring

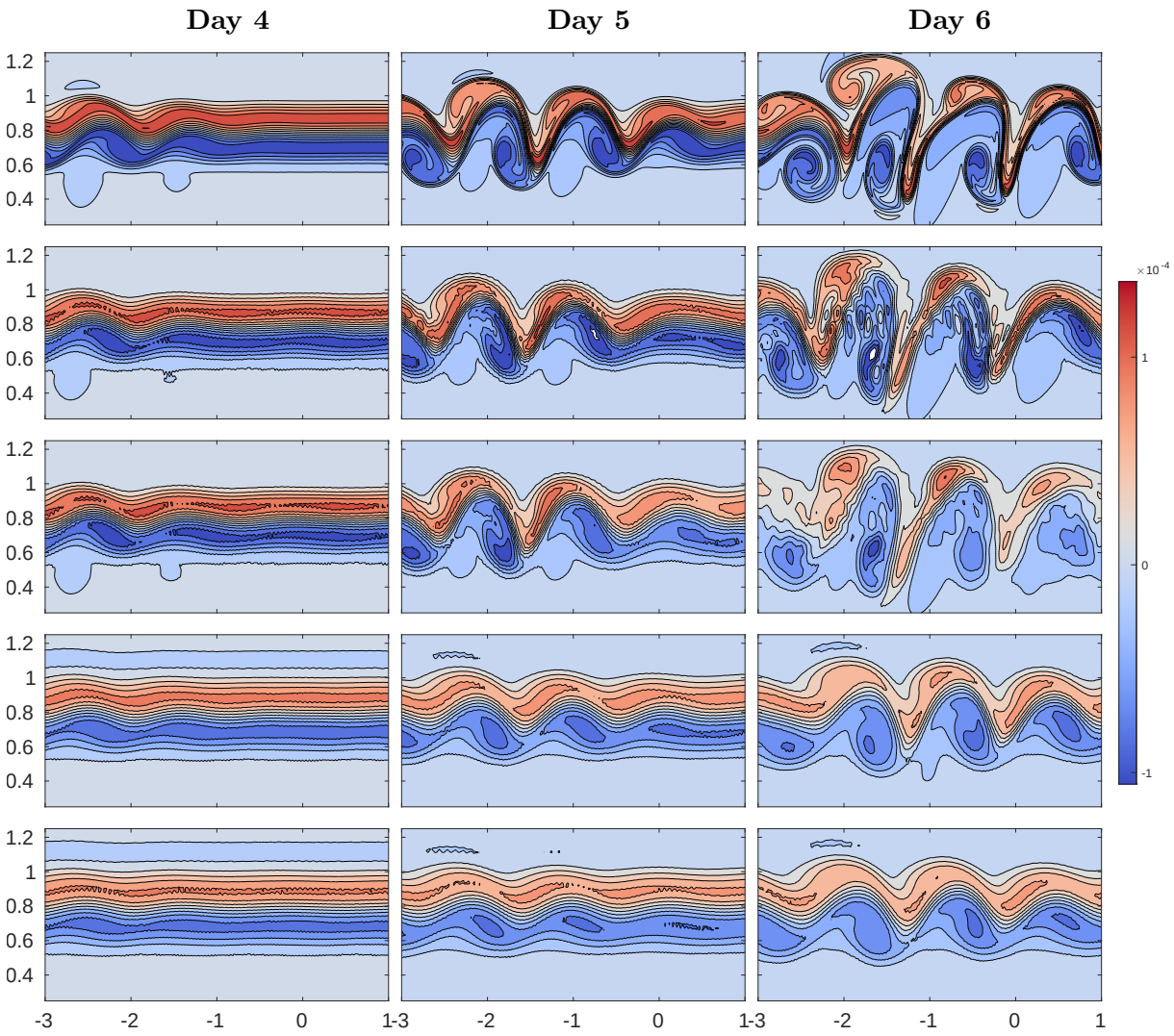


Figure 6. Comparison of the vorticity contour plots along the mid-latitude jet for different models (with 20480 triangles) at day 4, 5 and 6 respectively. From top to bottom: reference, ensemble mean of LU on-line, ensemble mean of LU off-line, ensemble mean of PIC off-line and deterministic LR. The contour interval is fixed to $2 \times 10^{-5}\,\mathrm{s}^{-1}$, the x-axis is longitude (in rad) and the y-axis is latitude (in rad).

of the large-scale flow. For instance, 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 at different days. 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.

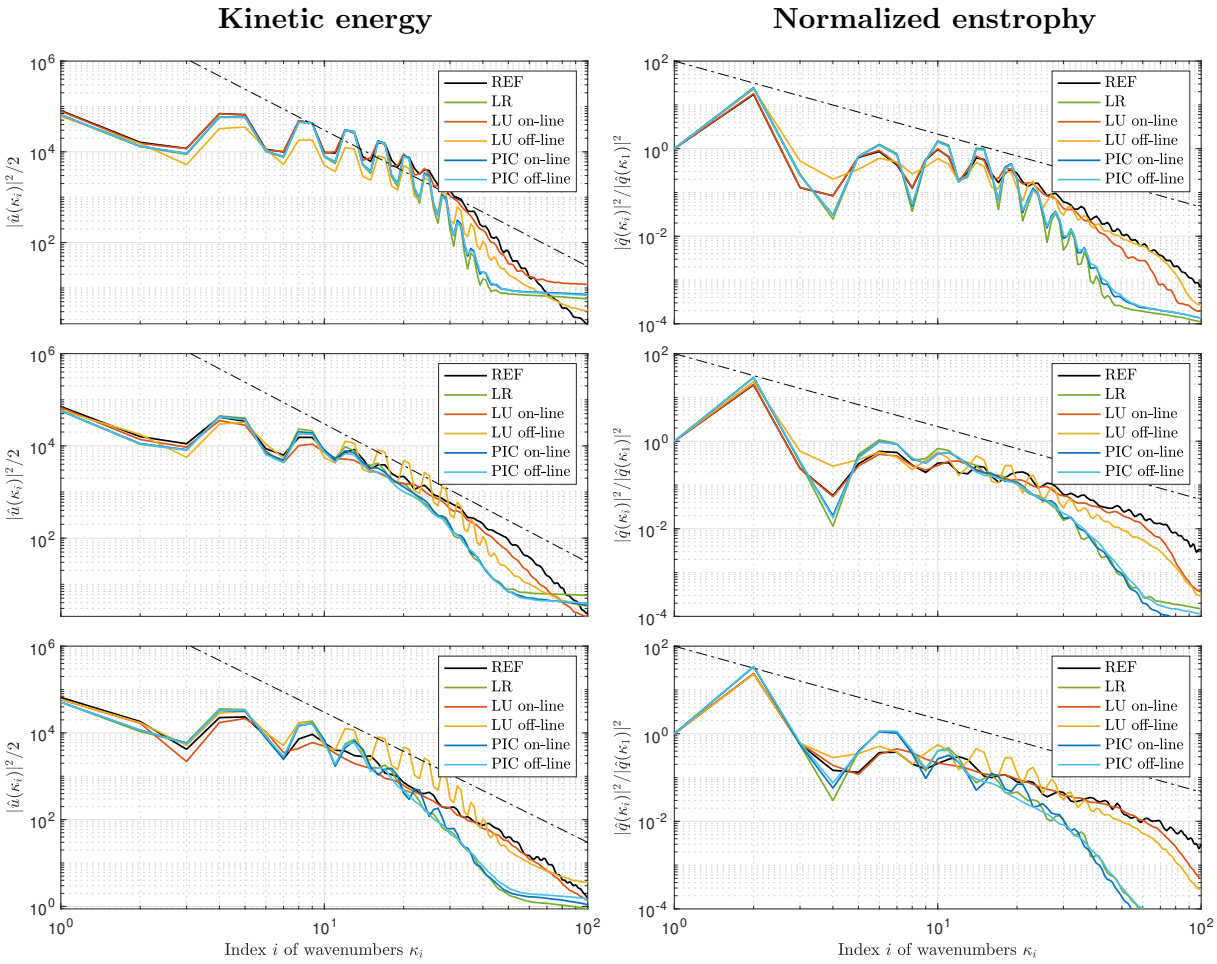


Figure 7. Comparison of the ensemble mean of the kinetic energy (left column) spectrums and the potential enstrophy (right column) spectrums for different models (with 20480 triangles) at day 5 (1st row), 7 (2nd row) and 10 (3rd row) respectively. Note that the potential enstrophy is defined by the square 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 $k^{-5/3}$ (right column) power law.

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 of each ensemble spread and compare it with the observation trajectory. 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 build by the 95% confident interval of its ensemble trajectories at each selected point. As shown in Figure 8, for the six local points chosen along the longitude $\Upsilon = -1.53\,\mathrm{rad}$, 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 off-line 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 (Gottwald and Harlim, 2013; Franzke et al., 2015).

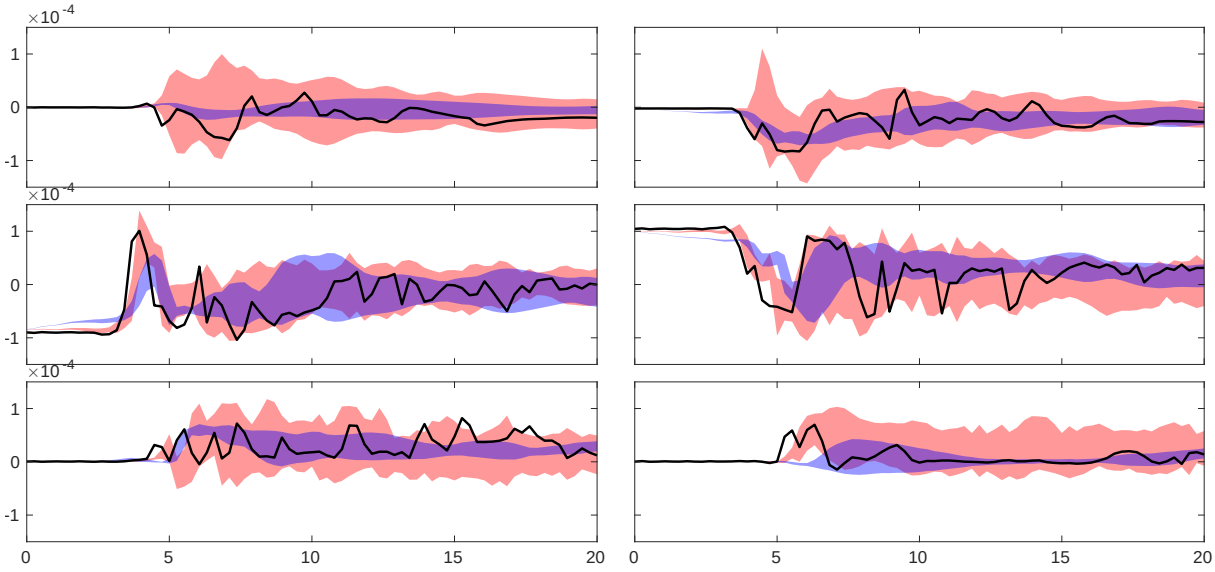


Figure 8. Comparison of the ensemble spread evolution over 20 days of the vorticity field for the LU-offline (red area) runs and the PIC-offline (blue area) runs, at six different locations $\Theta = (0.4, 0.56, 0.72, 0.88, 1.04, 1.2)$ rad along the longitude $\Upsilon = -1.53$ rad. The observation trajectories are shown by the black lines.

To quantify whether the ensemble spread of the forecast system represents the true uncertainty of the observations, the rank histogram (Talagrand et al., 1997; Hamill, 2001) 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 $\{q^{(j)}(\boldsymbol{x}_i)\}_{j=1,\dots,N_e}$ 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 $\{q^o(\boldsymbol{x}_i)\}_{i=1,\dots,N_o}$ to create a histogram of rank.

As shown in Figure 9, 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 on-line and PIC off-line 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

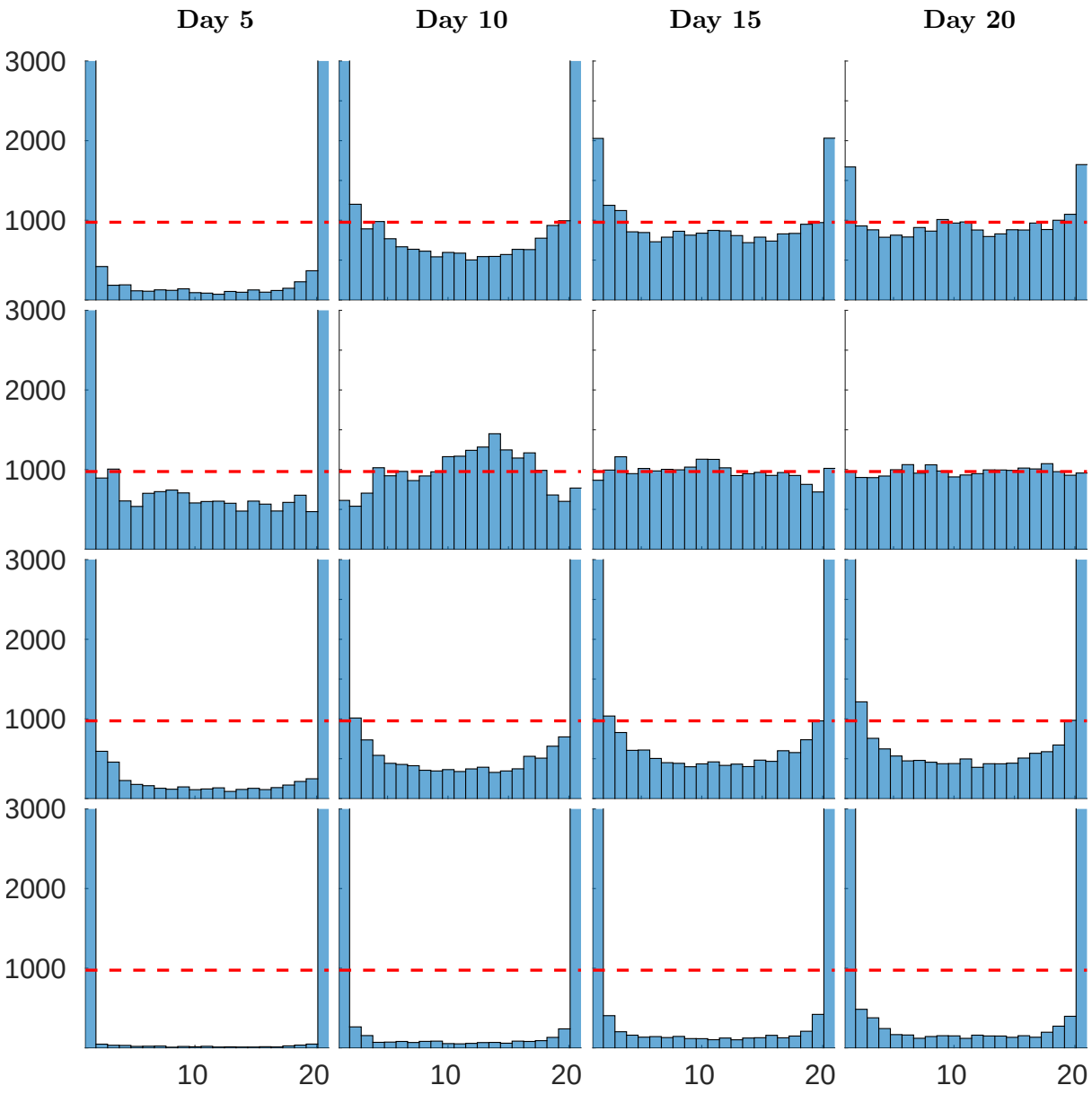


Figure 9. Comparison of the rank histograms for the LU on-line (1st row) runs, the LU off-line (2nd row) runs, the PIC on-line (3rd row) runs and PIC off-line (last row) runs, at day 5, 10, 15 and 20 respectively.

ensemble. To verify that a forecast is reliable, we need more criterions. 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, *i.e.*

$$\begin{aligned}\mathrm{MSE}\,(t) &= \frac{1}{N_o}\sum_{i=1}^{N_o}\left(q^o - \widehat{\mathbb{E}}[q]\right)^2(t,\boldsymbol{x}_i)\\ &\approx \left(\frac{N_e+1}{N_e}\right)\frac{1}{N_o}\sum_{i=1}^{N_o}\widehat{\mathrm{Var}}[q](t,\boldsymbol{x}_i) = \frac{N_e+1}{N_e}\,\mathrm{MEV}\,(t),\end{aligned} \tag{4.6}$$

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

In Figure 10, 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 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 off-line and PIC on-line 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.

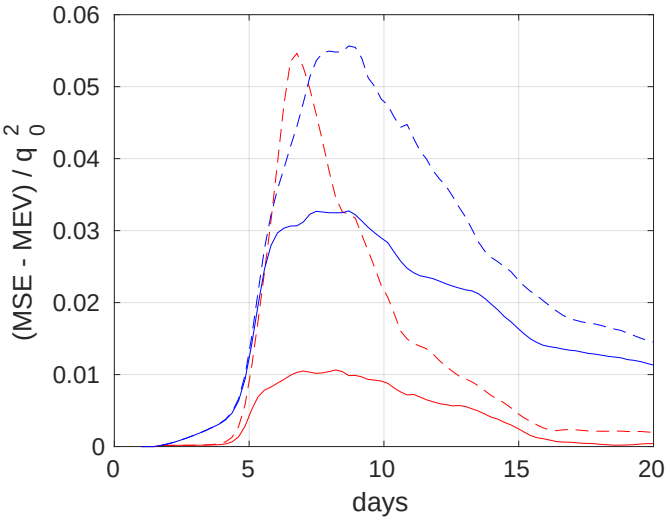


Figure 10. Comparison of the differences between the mean square error (MSE) and the mean ensemble variance (MEV) of the ensemble vorticity fields for the LU on-line (red dashed line) runs, the LU off-line (red solid line) runs, the PIC on-line (blue dashed line) runs and the PIC off-line (blue solid line) runs. Note that these differences are normalized by $q_0 = \|q(\Upsilon,\Theta,t=0)\|_\infty$.

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 modelling 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.

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 to 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 uncertainty quantification, we will next apply our developed RSW-LU system to data assimilation. We will also work towards 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.

Acknowledgments

The authors acknowledge the support of the Mitacs Globalink Research Award and of the ERC EU project 856408-STUOD. 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.

References

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

J. Anderson and S. Anderson. A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Monthly Weather Review*, 127(12):2741–2758, 1999.

B. K. Arbic, K. L. Polzin, R. B. Scott, J. G. Richman, and J. F. Shriver. On eddy viscosity, energy cascades, and the horizonal resolution of gridded stallite altimeter products. *Journal of Physical Oceanography*, 43(2):283–300, 2013.

W. Bauer and F. Gay-Balmaz. Towards a geometric variational discretization of compressible fluids: the rotating shallow water equations. *Journal of Computational Dynamics*, 6:1, 2019.

W. Bauer, P. Chandramouli, B. Chapron, L. Li, and E. Mémin. Deciphering the role of small-scale inhomogeneity on geophysical flow structuration: a stochastic approach. *Journal of Physical Oceanography*, 50(4):983–1003, 2020a.

W. Bauer, P. Chandramouli, L. Li, and E. Mémin. Stochastic representation of mesoscale eddy effects in coarse-resolution barotropic models. *Ocean Modelling*, 151:101646, 2020b.

Werner Bauer. *Toward goal-oriented R-adaptive models in geophysical fluid dynamics using a generalized discretization approach*. PhD thesis, Hamburg University Hamburg, 2013.

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

B. Chapron, P. Dérian, E. Mémin, and V. Resseguier. Large-scale flows under location uncertainty: a consistent stochastic framework. *Quarterly Journal of the Royal Meteorological Society*, 144(710):251–260, 2018.

G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2 edition, 2014.

C. E. Franzke and A. J. Majda. Low-order stochastic mode reduction for a prototype atmospheric GCM. *Journal of the Atmospheric Sciences*, 63(2):457–479, 2006.

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

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

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

G. Gottwald and J. Harlim. 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, 2013.

G. Gottwald, D. T. Crommelin, and C. E. Franzke. Stochastic climate theory. In *Nonlinear and Stochastic Climate Dynamics*, pages 209–240. Cambridge University Press, 2017.

I. Grooms and A. J. Majda. Stochastic superparameterization in quasigeostrophic turbulence. *Journal of Computational Physics*, 271:78–98, 2014.

Ernst Hairer, Christian Lubich, and Gerhard Wanner. *Geometric numerical integration: structure-preserving algorithms for ordinary differential equations*, volume 31. Springer Science & Business Media, 2006.

T. M. Hamill. Interpretation of rank histograms for verifying ensemble forecasts. *Monthly Weather Review*, 129:550–560, 2001.

S. Kadri Harouna and E. Mémin. Stochastic representation of the Reynolds transport theorem: revisiting large-scale modeling. *Computers and Fluids*, 156:456–469, 2017.

J. Kjellsson and L. Zanna. The impact of horizontal resolution on energy transfers in global ocean models. *Fluids*, 2(3):45, 2017.

P. E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*, volume 23. Springer-Verlag Berlin Heidelberg, 1992.

H. Kunita. *Stochastic flows and stochastic differential equations*, volume 24 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 1997.

A. Majda, C. Franzke, and B. Khouider. 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, 2008.

Jerrold E Marsden and Matthew West. Discrete mechanics and variational integrators. *Acta Numerica*, 10(1):357–514, 2001.

J. C. McWilliams. *Fundamentals of Geophysical Fluid Dynamics*. Cambridge University Press, 2006.

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

J Blair Perot, Dragan Vidovic, and Pieter Wesseling. Mimetic reconstruction of vectors. In *Compatible Spatial Discretizations*, pages 173–188. Springer, 2006.

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

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

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

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

V. Resseguier, L. Li, G. Jouan, P. Derian, E. Mémin, and B. Chapron. New trends in ensemble forecast strategy: uncertainty quantification for coarse-grid computational fluid dynamics. *Archives of Computational Methods in Engineering*, pages 1886–1784, 2020.

L. Sirovich. Turbulence and the dynamics of coherent structures, part I: Coherent structures. *Quarterly of Applied Mathematics*, 45(3):561–571, 1987.

O. Talagrand, R. Vautard, and B. Strauss. Evaluation of probabilistic prediction systems. Workshop on Predictability, ECMWF, 1997.

G. K. Vallis. *Atmospheric and oceanic fluid dynamics: fundamentals and large-scale circulation*. Cambridge University Press, 2 edition, 2017.

A. P. Weigel. Ensemble forecasts. In *Forecast Verification*, chapter 8, pages 141–166. John Wiley and Sons, Ltd, 2012.