Scale-Aware Deterministic and Stochastic Parametrizations of Eddy-Mean Flow Interaction

Laure Zanna[a,*], PierGianLuca Porta Mana, James Anstey, Tomos David, Thomas Bolton[a]

[a]*Atmospheric, Oceanic and Planetary Physics, Department of Physics, University of Oxford, Oxford, OX1 3PU, UK*

Keywords:

Eddy-permitting models, Mesoscale eddies, Stochastic parametrization, Energy backscatter, Upgradient fluxes

ABSTRACT

The role of mesoscale eddies is crucial for the ocean circulation and its energy budget. The sub-grid scale eddy variability needs to be parametrized in ocean models, even at so-called eddy permitting resolutions. Porta Mana and Zanna (2014) propose an eddy parametrization based on a non-Newtonian stress which depends on the partially resolved scales and their variability. In the present study, we test two versions of the parametrization, one deterministic and one stochastic, at coarse and eddy-permitting resolutions in a double gyre quasi-geostrophic model. The parametrization leads to drastic improvements in the mean state and variability of the ocean state, namely in the jet rectification and the kinetic-energy spectra as a function of wavenumber and frequency for eddy permitting models. The parametrization also appears to have a stabilizing effect on the model, especially the stochastic version. The parametrization possesses attractive features for implementation in global models: very little computational cost, it is flow aware and uses the properties of the underlying flow. The deterministic coefficient is scale-aware, while the stochastic parameter is scale- and flow-aware with dependence on resolution, stratification and wind forcing.

*Corresponding author. Tel.: +44 1865 272925
Email address: laure.zanna@physics.ox.ac.uk (Laure Zanna)

1. Introduction

Ocean mesoscale eddies, with scales of 10–100 kilometres, are turbulent features in the ocean derived from barotropic and baroclinic instabilities, and are strongly influenced by wind forcing and stratification. Eddies play a key role in ocean circulation, including tracer transport, mixing and stirring, and actively participate in energy transfer between scales. The mesoscale eddy energy is particularly enhanced in the vicinity of western boundary currents and their extension (e.g. Gulf Stream and Kuroshio), and in the Southern Ocean. Eddies are crucial in the feedback of energy to the large-scale flow (e.g., Scott and Arbic 2007) and in maintaining the jet extension via upgradient momentum fluxes leading to sharpening of gradients (Greatbatch et al. 2010).

Climate models from the Coupled Model Intercomparison Project (CMIP) archive (Taylor et al. 2012) used for the last Intergovernmental Panel on Climate Change (IPCC 2013) have too coarse horizontal resolution to resolve these eddies. The effect of eddies on the large scale is parametrized in such coarse resolution models using the Gent-McWilliams parametrization (Gent and McWilliams 1990; Gent et al. 1995). The parametrization has shown great success in reducing spurious convective instabilities in coarse-resolution models. The parametrization mimics the effects of baroclinic instability, converting available potential energy into kinetic energy, and acts on buoyancy and passive tracers, but neglects eddy Reynolds stresses and sub-grid scale fluctuations. The horizontal resolution of the most recent generation of global climate models has increased to a scale close to the Rossby radius of deformation. These models, often called eddy-permitting, are therefore starting to successfully capture some of the mesoscale eddy behaviour, especially at low- and mid- latitudes. However, eddy-permitting models remain unsuccessful at resolving the full mesoscale eddy field (Gnanadesikan and Hallberg 2000; Hallberg 2013) and its interaction with the large scales, and might not be able to do so in the near feature (Fox-Kemper et al. 2014). Therefore parametrizing sub-grid eddies, especially in eddy-permitting models, remains an important topic of research, as the previous generation of parametrizations, de-

rived for coarse-resolution models, might not be able to successfully mimic the effects of the unresolved scales on the large-scale flow.

Sub-grid parametrization at eddy-permitting resolution is necessary not only to represent the unresolved scales but also to maintain numerical stability. Numerical dissipation is often achieved using Laplacian viscosity (or diffusion) with too large coefficients, or using hyperviscous parametrization (Holloway 1992; Frisch et al. 2008) or biharmonic closure (Smagorinsky 1963; Leith 1990; Griffies and Hallberg 2000) which dissipates enstrophy at the grid scale near the deformation radius and scales with model resolution. However, recent studies have shown that hyperviscosity, in addition to representing a direct enstrophy cascade (Bachman et al. 2016), spuriously dissipates energy at small scales (Arbic et al. 2007; Jansen and Held 2014). Parametrization of sub-grid scale eddies for eddy-permitting regimes are therefore needed to either correct the spurious loss of energy resulting from the use of hyperviscosity (including modified hyperviscosity; Fox-Kemper and Menemenlis 2008), or to replace hyperviscosity altogether. The aim of our paper is to introduce an eddy parametrization, derived for eddy-permitting models, that makes use of the resolved variability, mimics the behaviour of Reynolds stresses such as sharpening ocean jets, scales with resolution and the flow, and feeds back energy lost due to viscosity.

Jansen and Held (2014) propose to re-inject the energy lost at small scales using a negative viscosity determined by an energy equation following Eden (2010). Filtering of the velocities, as done for example in the Lagrangian-averaged Navier-Stokes-α model (Holm and Wingate 2005; Holm and Nadiga 2003), or the nonlinear gradient approximation (Nadiga and Bouchet 2011) have shown promising results (see PMZ14 and Anstey and Zanna (2016) for comparisons between our proposed schemes and these studies). However, recent studies (Graham and Ringler 2013) highlighted that these parametrizations can lead to a build-up of enstrophy at small scales and to numerical instability. Other approaches at eddy-permitting resolutions have argued for the use of a stochastic term for upgradient momentum fluxes and energy backscatter in spectral models (Kraichnan 1976; Frederiksen and Davies 1997;

Duan and Nadiga 2007; Nadiga 2008; Kitsios et al. 2012; Grooms and Majda 2013). The sub-grid forcing is generally constrained by an energy spectrum. In quasi-geostrophic models the need for upgradient momentum closures based on a stochastic model was also pointed out (Berloff 2005b, 2015, 2016). However, all approaches require some a priori knowledge of sub-grid eddy statistics.

Here we implement a parametrization proposed by Porta Mana and Zanna (2014, referred to as PMZ14). In PMZ14 we diagnosed a relationship between the missing eddy forcing and a non-Newtonian stress divergence (Ericksen 1956; Rivlin 1957). The missing forcing is defined as the PV eddy flux divergence resulting from a high-resolution eddy resolving model compared to an eddy permitting model. The non-Newtonian stress divergence depends on the Lagrangian rate of change of the potential vorticity (PV) gradient and its local deformation. The relationship between the missing eddy forcing and a non-Newtonian stress divergence was inspired by general principles of potential vorticity conservation, frame-invariance, differential memory (Truesdell and Noll 2004) and symmetry properties of the stress tensor (Bachman and Fox-Kemper 2013). The relationship, more intuitively, is based on an argument that in eddy-permitting models the rate of strain, eddy shape and orientation, and the PV gradient can be used to mimic the evolution of the eddy PV forcing (Nadiga 2008; Anstey and Zanna 2016). The work argued that the parametrization could be efficient in a deterministic mode, with a coefficient for the parametrization that scales with model resolution. In addition, a stochastic parametrization was also presented, with a stochastic forcing term whose probability is conditional on the non-Newtonian forcing, wind forcing, stratification, and model resolution.

This paper is structured as follows. In Section 2 we briefly present the quasi-geostrophic model used in the current study. In Section 3 we discuss two implementations of the parametrization, one deterministic and one stochastic. In Section 4 we present the results of the two different implementations for the mean flow and variability. Section 5 is a discussion of the impact of the parametrized forcing on the momentum, energy and enstrophy budgets

and presents ways forward for implementation in primitive-equations models. We briefly conclude in section 6.

2. Model Setup

The model used in the present study, PEQUOD, solves the forced dissipative baroclinic quasi-geostrophic (QG) potential vorticity (PV) equation on a beta plane in a square basin (e.g., Berloff 2005b,a). The main setup is similar to the one used in Porta Mana and Zanna (2014, PMZ14). The model is composed of three isopycnal layers with thicknesses H_m (with $m = 1, 2, 3$ for the upper, middle and bottom layer, respectively). For each layer m, the prognostic equation solved for the potential vorticity q is given by

$$\frac{\mathrm{D}q_m}{\mathrm{D}t} = \frac{\partial q_m}{\partial t} + \mathbf{u}_m \cdot \nabla q_m = \mathcal{D}_m + F_m^{\mathrm{wind}} + F_m^{\mathrm{eddy}}, \tag{1}$$

with

$$q_m = \nabla^2 \psi_m + \beta y + \frac{\partial}{\partial z}\left(\frac{f_0^2}{N^2}\frac{\partial \psi_m}{\partial z}\right). \tag{2}$$

The planetary vorticity is $f = f_0 + \beta y$, $\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right)$ is the horizontal gradient, N is the Brunt-Väisälä frequency of the mean density stratification and ψ is the streamfunction derived from the non-divergent velocity such that $\mathbf{u}_m = \left(-\frac{\partial \psi_m}{\partial y}, \frac{\partial \psi_m}{\partial x}\right)$.

The dissipation term is $\mathcal{D}_m = -r\nabla^2\psi\delta_{m,3} - \nu\nabla^6\psi_m$, where $\delta_{m,i}$ is the Kronecker delta function. The first term parametrizes the presence of a bottom Ekman layer with a bottom drag coefficient r. The second term is a horizontal biharmonic viscosity term, with viscosity coefficient ν, which scale-selectively dissipates enstrophy near the grid-scale. Note that PMZ14 used a Laplacian viscosity term rather than a biharmonic term for high-resolution and eddy-permitting runs (but not for coarse resolution runs). The use of hyperviscosity in the present study is to allow a setup that mimics current eddy-permitting ocean model setups, and ensures small-scale dissipation and numerical stability (we struggled to keep the model stable when using the deterministic parametrization; see Sec. 3). The hyperviscous

term was calculated at the previous timestep for practical reasons (see Section 3b).

The forcing F_m, applied to the upper layer, is the curl of the wind stress τ:

$$F_m^{\text{wind}}(x,y) = \frac{(\nabla \times \tau)_z}{\rho_0 \; H_1}\delta_{m,1}, \tag{3}$$

where ρ_0 is the reference density. The wind stress curl profile is identical to PMZ14 and spins up two gyres separated by a strong meandering jet emanating from the western boundary.

The term F_m^{eddy} is the eddy parametrization, which can take a deterministic or a stochastic form. We use different model configurations defined as follows:

i. The "truth": a high-resolution run with 7.5 km horizontal resolution. The eddy forcing term F_m^{eddy}, in Eq. 1, is set to 0.

ii. Low-resolution unparametrized runs: runs at eddy-permitting resolution with horizontal grid-spacing of 30 km and 60 km; and a coarse-resolution run at resolution of 120 km. No parametrization of eddy forcing is included, i.e. F_m^{eddy} is again set to 0.

iii. Low-resolution parametrized runs: Same as in (2), except for F_m^{eddy} being non-zero. The term F_m^{eddy} has a spatial and temporal dependence on the flow, which can be deterministic or stochastic as defined in Section 3.

The length of the integrations from rest to statistically steady state is 410 years. All simulations presented are numerically converged and are solved using centered-leapfrog with RAW filter (Williams 2009) and a modified Arakawa advective scheme (Arakawa 1966). Additional experiments using different numerical schemes led to similar results to the ones presented in the following sections. We output the data daily (as snapshots), with all steady states and variance calculations done using the last 200 years of model integration. The common model parameters for all runs are defined in Table 1 while the differing model parameters for various resolutions are listed in Table 2.

Figure 1 shows the statistically steady streamfunction ψ for the high resolution truth (panel a) and the 30 km-resolution unparametrized model (panel b). A few differences between the runs include the absence of a strong and narrow eastward jet at low resolution

compared to the truth; and a too far south separation point at the western boundary. The eddy-permitting simulations at 30 km and 60 km (the latter not shown) generate some eddies via barotropic and baroclinic instabilities but with fewer filaments and weaker turbulence compared to the eddy resolving run (cf Fig. 2a and b). The potential vorticity snapshot at low resolution also hints at some numerical instability near the jet, due to the appearance of sharp features of alternating sign with spatial scale of the model gridscale (30km). The differences in the simulations arise as a result of the small Reynolds number at low resolution, due to the increased horizontal grid box size and viscosity.

3. Parametrization of sub-grid mesoscale eddies

In this section we introduce the parametrization proposed by PMZ14 and discuss the deterministic and stochastic implementations in the baroclinic QG model of Section 2. In the remainder of the paper we omit the layer subscript m for conciseness.

a. PMZ14 parametrization

In PMZ14 we postulate that the divergence of the sub-grid eddy PV stress in QG (S^*) reflecting the missing forcing due to truncated nonlinear advection and increased dissipation between the truth and a low-resolution run can be well approximated by

$$F^{\text{eddy}} = \kappa \nabla \cdot \nabla \frac{\mathrm{D}q}{\mathrm{D}t}, \tag{4}$$

where κ is a scalar, independent of space or time. The value of κ is estimated by coarse-graining high-resolution simulations onto a coarse resolution grid. Unlike common closures, the parameter κ is not a diffusivity or viscosity coefficient, as it has units of length squared. As discussed in PMZ14, the proposed closure, obtained by imposing several mathematical and physical constraints such as frame-invariance and memory, can be expressed as a non-

Newtonian second order Rivlin-Ericksen-like stress (Ericksen 1956; Rivlin 1957) using

$$\frac{\mathrm{D}}{\mathrm{D}t}\nabla q + \nabla \mathbf{u}^T \cdot \nabla q = \nabla \frac{\mathrm{D}q}{\mathrm{D}t}. \tag{5}$$

where $\nabla \mathbf{u}$ is a rank-2 tensor (i.e., a matrix) given by

$$\nabla \mathbf{u}^T = \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial v}{\partial x} \\ \frac{\partial u}{\partial y} & \frac{\partial v}{\partial x} \end{bmatrix}. \tag{6}$$

Qualitatively, the parametrization can amplify or weaken existing gradients of PV. The parametrization is mainly applicable at eddy-permitting resolution in which instabilities can create meanders and deformations of the flow field (Fig. 2b) which are required for the parametrization to perform most effectively.

The robustness of the relationship between the diagnosed eddy forcing (S^*) and the proposed parametrization is thoroughly quantified in PMZ14 by constructing probability distribution functions (PDFs) and conditional probability distribution functions (cPDFs)[1]. The diagnosed eddy forcing PDFs conditional on the expression (4) are used to assess the degree to which the "true" eddy forcing correlates with $\nabla^2 \frac{\mathrm{D}q}{\mathrm{D}t}$ for each layer. PMZ14 find, using the conditional PDFs $\mathrm{P}(S^* | \nabla^2 \frac{\mathrm{D}q}{\mathrm{D}t})$, a very strong linear correlation between the two quantities, such that the mean of the PDFs of the eddy forcing term conditional on $\nabla^2 \frac{\mathrm{D}q}{\mathrm{D}t}$ can be used as a deterministic parametrization using a constant coefficient κ, where κ is negative with no spatio-temporal dependence.[2]

The correlation can exhibit large deviations from its mean for large values of eddy PV forcing (i.e., the PDFs are not delta functions). Excursions in the eddy forcing can reach more than 40% of its mean value. A single constant coefficient, or a single value of the parametrized eddy forcing, therefore cannot adequately represent the behaviour of highly turbulent eddying regions. Large values of eddy forcing parametrization, $\nabla^2 \frac{\mathrm{D}q}{\mathrm{D}t}$, describe

[1] Similar relationship are found if the truth is run at resolution of 2.5km rather than at 7.5km.

[2] A more thorough investigation of κ as a tensor rather than a scalar, taking into account the anisotropy of the stress is left for a follow-up study.

large growth rate of nonlinear instabilities developing in the flow. Such nonlinear processes can exhibit large fluctuations which are reflected in skewed PDFs, as found in PMZ14. Therefore a statistical (stochastic) parametrization of ocean eddies might be more appropriate when attempting to capture the effect of eddies in highly turbulent jet regions. In the following two sections we describe the implementation of a deterministic and stochastic version of the eddy parametrization in the QG model.

b. Deterministic Parametrization

Assuming that a constant (and negative) value for κ can be used in (4), the QG model equation with the deterministic eddy parametrization is given by

$$\frac{\mathrm{D}q}{\mathrm{D}t} \equiv \frac{\partial q}{\partial t} + \nabla \cdot (\mathbf{u}q) = \kappa \nabla^2 \frac{\mathrm{D}q}{\mathrm{D}t} - \nu \nabla^6 \psi + F^{\mathrm{wind}}. \tag{7}$$

The parametrized equation can be rewritten as

$$\left(1 - \kappa \nabla^2\right) \frac{\mathrm{D}q}{\mathrm{D}t} = -\nu \nabla^6 \psi + F^{\mathrm{wind}}, \tag{8}$$

or

$$\frac{\mathrm{D}q}{\mathrm{D}t} = \left(1 - \kappa \nabla^2\right)^{-1} \left(-\nu \nabla^6 \psi + F^{\mathrm{wind}}\right). \tag{9}$$

Using coarse-grained high-resolution experiments, PMZ14 observe that the value of κ, which was found to be negative, only scales with the grid size of the coarse-resolution model. The operator acting on the PV tendency and advective terms, $(1 - \kappa \nabla^2)$, behaves as a roughener of total forcing in the PV equation. This operator could be viewed as a filter acting to replace the scales of motion that are truncated by the coarse grid scale of the model, with the strength of the filter being determined by the value of κ.

Due to the singularity present at wavenumbers $K^2 = -1/\kappa$, inverting the operator $(1 - \kappa \nabla^2)$ to solve Eq. 9 is not feasible. Instead, we numerically solve Eq. 7 by evaluating $\kappa \nabla^2 \frac{\mathrm{D}q}{\mathrm{D}t}$ at the previous timestep, so that the parametrization is implemented as

$$\frac{\partial q}{\partial t} + \mathbf{u} \cdot \nabla q = \kappa \nabla^2 \frac{\mathrm{D}q}{\mathrm{D}t}(t - \Delta t) - \nu \nabla^6 \psi + F^{\mathrm{wind}}, \tag{10}$$

where Δt is the model timestep increment (the hyperviscous term is also commonly calculated at the previous time step for numerical stability) and other terms are calculated at time t. The implementation of the deterministic parametrization comes with the same computational cost as biharmonic viscosity. The implementation only requires saving the Lagrangian (material) derivative of potential vorticity at the previous timestep and calculating its Laplacian.

Using a large range of numerical simulations under different forcing and dissipation and different geometric configurations, PMZ14 estimate the mean κ value as $-(\alpha\Delta x)^2$, where $\alpha \approx 0.45$. To avoid the singularity present at wavenumbers $K^2 \approx -1/(\alpha\Delta x)^2$, we adjust the value of κ such that $|\kappa| = (0.31\Delta x)^2$ in order to ensure numerical stability (see Sect. 5a for further discussion).

c. Stochastic parametrization

Even if the deterministic parametrization shows a useful relationship between the eddy properties and the large scale flow, it remains difficult to predict the behaviour of turbulent flow with a single deterministic relationship based on the resolved scales. The turbulent nature of the flow field suggests the use of a non-deterministic closure to describe the eddy-mean flow interaction. PMZ14 showed that probability distribution functions of the eddy forcing conditional on the large scale flow, $\nabla^2 \frac{\mathrm{D}q}{\mathrm{D}t}$, can be reconstructed given solely information from the coarse-resolution QG model. One main characteristic is that the standard deviation and the higher-order moments of the PDF increase with the Reynolds number and decrease with the coarse-graining grid box size.

In the eddy-permitting range considered here, the standard deviation can be expressed to a very good approximation as $\sigma = \gamma \frac{\tau_0}{\Delta x H \rho_0}$, with $\gamma = 7.9 \times 10^{-4}$ (a non-dimensional parameter). The standard deviation scales linearly with the wind stress and exhibits an inverse dependence on the resolution. The standard deviation also scales linearly with the PV variance, such that $\sigma = \gamma_\mathrm{q} \frac{\sigma_q^2 H}{\Delta x}$ with $\gamma_\mathrm{q} = 1.7$ (also non-dimensional) and is valid whether we

consider wind or buoyancy forcing (see PMZ14 for sensitivity to forcing, resolution and model configuration). Note that using coarse-graining diagnostics, we find that the maximum value of σ is always proportional to the mean value of the eddy forcing, with a scale-independent proportionality constant, such that $\sigma \propto \Delta x^2$ (not shown). Therefore, as Δx goes to zero, the mean and standard deviations also tend to zero. The standardized skewness and kurtosis (i.e., skewness and kurtosis scaled by the standard deviation) were shown to be roughly independent of model parameters and found to be $O(1)$, with values $\mu_3 \approx 0.61$ and $\mu_4 \approx 1.4$, respectively. The resolved quantities therefore fully determine the standard deviation, skewness, and kurtosis of the PDFs.

For each run, we use the relationships found in PMZ14 and described above to reconstruct the PDFs, based on a maximum entropy procedure (Mead and Papanicolaou 1984), and use them as the basis for a stochastic parametrization. The implementation of the stochastic parametrization can thus be written as

$$\frac{\partial q}{\partial t} + \mathbf{u} \cdot \nabla q = \mathcal{F}\left[\nabla^2 \frac{\mathrm{D}q}{\mathrm{D}t}(t - \Delta t), \Delta x, H, \tau_0\right] - \nu \nabla^6 \psi + F^{\mathrm{wind}}, \tag{11}$$

where the stochastic term $\mathcal{F}$ is sampled according to the reconstructed conditional PDFs, with spatial and temporal scales, Δx_{stoc} and Δt_{stoc}, respectively. The scales determine intervals at which the stochastic term $\mathcal{F}$ is selected before being recalculated. In all the runs presented here, we diagnose the value of $\nabla^2 \frac{\mathrm{D}q}{\mathrm{D}t}(t - \Delta t)$ at every grid box ($\Delta x_{\mathrm{stoc}} = \Delta x$), then sample $\mathcal{F}$ from the corresponding conditional PDFs $\mathrm{P}(S^* | \nabla^2 \frac{\mathrm{D}q}{\mathrm{D}t})$; the value of $\mathcal{F}$ is kept for $\Delta t_{\mathrm{stoc}} = 1$ day. The timescale is chosen by using the coarse-grained information which showed a decorrelation time-scale of about 1 day for the eddy forcing term. The results deteriorate if the timescale is too short (on the order of a few timesteps), which turns the parametrization into a simple AR1 process and is not an accurate representation of the eddy forcing (see Berloff 2005b). It is important to reiterate that the stochastic parametrization is not equivalent to simply adding random fluctuations: the stochastic term is sampled from a PDFs given a value of the local $\nabla^2 \frac{\mathrm{D}q}{\mathrm{D}t}(t - \Delta t)$; it is therefore strongly dependent on the resolved flow field. For this implementation there is no rescaling of the mean or any other

moments of the PDFs. All results presented in the paper are using the above implementation. The computational cost associated with this stochastic parametrization is about 10% more compared to the deterministic parametrization. The cost can be reduced by changing the spatial and temporal decorrelation of the stochastic term ($\Delta x_{\rm stoc}$ and $\Delta t_{\rm stoc}$, respectively) or by slightly changing the implementation.

A slightly faster implementation can be introduced. The following implementation, similarly to the deterministic runs, requires tuning of the mean of the PDFs to ensure numerical stability by avoiding the singularity discussed in Sect.3b. This additional implementation is briefly outlined below, however readers can skip directly to Sect. 4 for the results of the simulations described above. We can decompose the eddy parametrization into a deterministic and stochastic term such that

$$F^{\rm eddy} = \kappa\nabla^2\frac{{\rm D}q}{{\rm D}t}(t-\Delta t) + s^* \times \sigma\Big[\nabla^2\frac{{\rm D}q}{{\rm D}t}(t-\Delta t)\Big].$$

The deterministic part, $\kappa\nabla^2\frac{{\rm D}q}{{\rm D}t}(t-\Delta t)$, is calculated as described in Sect. 3b using $|\kappa| = (0.31\Delta x)^2$. The stochastic term is determined as follows. The standard deviation σ of the cPDFs scales with $\nabla^2\frac{{\rm D}q}{{\rm D}t}(t-\Delta t)$, and skewness and kurtosis of the cPDFs are standardized moments such that the shape of the cPDFs is self-similar. We can therefore select a value s^* from a single rescaled conditional PDF with zero mean, standard deviation of one, and constant moments μ_3, μ_4 such that ${\rm P}(s^*|\,\mu = 0, \sigma = 1, |\mu_3|, \mu_4)$. The rescaled cPDF is independent of $\nabla^2\frac{{\rm D}q}{{\rm D}t}$. We can then rescale s^* by the standard deviation of $\nabla^2\frac{{\rm D}q}{{\rm D}t}(t-\Delta t)$, $\sigma\Big[\nabla^2\frac{{\rm D}q}{{\rm D}t}(t-\Delta t)\Big]$. With this implementation, there is no need to first check the value of $\nabla^2\frac{{\rm D}q}{{\rm D}t}(t-\Delta t)$ and then choose from a specific conditional PDF. The results of this stochastic implementation are slightly closer to the deterministic simulations than to the stochastic simulations described using Eq. 11, possibly due to the reduction in κ necessary for numerical stability.

4. Results

In this section, we analyse and compare the results of deterministic and stochastic implementations at different horizontal resolutions, with a focus on runs performed at 30 km resolution.

Figure 1 shows the statistically steady streamfunction ψ for the high resolution truth (7.5 km horizontal resolution) and for the unparametrized and parametrized eddy permitting runs at 30 km horizontal resolution. The parametrized runs show clear improvement in the steady state streamfunction for the deterministic (Fig. 1c) and stochastic (Fig. 1d) implementations, compared to the unparametrized run at the same resolution (Fig. 1b). The steady state streamfunction for the parametrized simulations are closer to the truth (Fig. 1a) than the unparametrized run. The strength of the subtropical gyre is marginally increased in the parametrized runs to 14.71 Sv and 15.22 Sv, compared to 14.48 Sv in the unparametrized run and 15.54 Sv in the high resolution run. The strength of the subpolar gyre is significantly improved from -17.40 Sv in the unparametrized run, to -18.92 Sv and -21.6 Sv in deterministic and stochastic runs, compared to -24.75 Sv in the high resolution run. Another feature is the improvement of the separation point of the jet in the parametrized runs. In the unparametrized run, the jet separates from the western boundary 300 km further south than in the high resolution. In the deterministic and the stochastic runs, the separation point is located only 210 km and 60 km south of the separation point in the high resolution run, respectively.

Snapshots of potential vorticity for the different model runs, taken after 250 years of model run, are shown in Fig. 2, only for illustrative purposes. The unparametrized run shows the presence of weak turbulence and also some grid scale numerical noise (panel b). However the high resolution (panel a) and parametrized runs (panels c-d) exhibit a lot more turbulence and filamentation. The introduction of the parametrization results in the enhancement of gradient of potential vorticity but also attenuates the numerical noise present in the unparametrized simulation (see Sect 5 for discussion).

Figure 3 shows the steady state kinetic energy as a function of latitude at $x = 120$km and $x = 300$km, further highlighting the presence of a strong jet detaching from the western boundary in the high-resolution simulation (black). The zonal eddy momentum stress divergence, $\nabla \cdot \overline{u'\mathbf{u}'}$, is shown in panels (c)-(d) for the same longitudes as panels (a)-(b), where primes are anomalies from the time mean denoted by an overline (Waterman and Jayne 2012). In the unparametrized simulation (blue), the jet is weaker and broader than in the high resolution simulation. The deterministic (red) and stochastic (grey) parametrizations lead to a sharpening of the jet and an increase in the zonal velocity (hence in the kinetic energy) in the core of the jet. The jet sharpening is a direct consequence of the parametrization, which enhances the eddy momentum stress divergence. Another notable impact of the parametrization is to inject energy back into the large-scale flow, as shown in Fig. 4. The introduction of the eddy closure vastly improves the spectrum of kinetic energy over all wavenumbers by reducing the amount of energy dissipated at small scales and allowing energy spuriously lost to be backscattered.

This can further be explored by considering the turbulent energy budget (Arbic et al. 2007; Larichev and Held 1995; Scott and Arbic 2007; Straub and Nadiga 2014). As shown previously by these authors, the spectral budget for total energy in the QG model can be expressed as the sum of the contribution in Fourier space from forcing, bottom friction, hyperviscosity and from the redistribution of energy across scales from the spectral transfer of kinetic and available potential energy (APE). Figure 5 shows the statistically steady state budget of the different model simulations. For the high-resolution run at a resolution of 7.5km (panel a), the spectral characteristics are reminiscent of other QG and primitive equation models, and of baroclinic turbulence (Charney 1971; Rhines 1977; Salmon 1978). Forcing and dissipation mostly balance each other, and the spectral transfer of kinetic and APE summing up to zero for all wavenumbers. The transfer of kinetic energy has a large negative lobe (sink of kinetic energy) at higher wavenumbers, and a large positive lobe (source of KE) at smaller wavenumbers. Integrating this transfer (starting from large wavenumbers) would

produce a single, large negative lobe representing a net inverse transfer of energy to larger scales. APE is extracted at large scales and transferred down-scale toward the deformation scale. At the deformation scale, energy is converted to kinetic energy and this kinetic energy is transferred to large scales. Our parametrization is attempting to reproduce this energetic behaviour. There is no transfer of kinetic energy toward small scale, therefore as expected the term due to hyperviscosity is very small at a resolution of 7.5km.

At a resolution of 30km (Fig. 5b), without an eddy parametrization, our results are similar to Scott and Arbic (2007); Hallberg (2013); Jansen and Held (2014) with energy being lost near the grid scale (roughly the deformation scale) due to hyperviscosity, and therefore only part of the energy being fluxed towards small scale as APE is then fluxed upscale as kinetic energy. Both the sink of kinetic energy at higher wavenumbers and the source at lower wavenumbers are significantly smaller compared to the high-resolution case, thus leading to a reduced inverse transfer of kinetic energy to larger scales. As a consequence, there is less kinetic energy at large scales and a reduction of the APE being extracted. When the deterministic parametrization is introduced (Fig. 5c), energy is returned mostly at scales only somewhat larger than the deformation wavenumber as seeing by the increase in the size of both the positive and negative lobes in spectral transfer. The most dramatic improvements are when the stochastic parametrization is introduced (Fig. 5d), a larger portion of kinetic energy is being fluxed back up to the larger scales and more closely matches the kinetic energy transfer of the high-resolution simulation. These results are is in agreement with our other diagnostics.

Figure 6 summarizes the impact of the deterministic and stochastic parametrizations at different resolutions. The mean-flow error is vastly reduced in the vicinity of the western boundary current and its extension (cf. panels c-d with b). The deterministic parametrization leads to the creation of a jet extension, thereby reducing the error in the streamfunction away from the recirculation region. However, the introduction of the stochasticity vastly improves the recirculation gyres and leads to an error of less than 1 Sv at any location in

the basin. At eddy permitting resolutions (30 km and 60 km), the deterministic version of the parametrization leads to about 50% error reduction, and the stochastic component to a further 20 to 40%. At coarser resolution (120 km), there is a very small improvement in the mean flow when the deterministic parametrization is implemented, and there is no added value to the implementation of stochastic parametrization when considering the mean flow error.

Despite the limited number of tests at different resolutions, it appears that there is a cut-off at which stochasticity impacts upon the mean flow compared to the deterministic parametrization, which is roughly equal to the Rossby radius of deformation. Further tests will be required to validate the critical or cut-off resolution. Additionally, the reduction in mean flow error at coarse resolution (120 km) is rather small which is in agreement with our initial assumption that the parametrization will be successful at eddy-permitting resolution by reinforcing the existing gradients.

Figures 7 and 8 show that not only the mean flow is improved but also its spatial and temporal variability. The absence of a meandering jet in the low resolution run without any eddy closure is reflected by the loss of variance in the center of the domain (Fig. 7). The deterministic parametrization is shown to improve the variance significantly, especially at low frequency (Fig. 8). The stochastic parametrization has a positive impact at high-frequency, as expected by the introduction of random fluctuations, but also at low frequency. Therefore the high-frequency variability of mesoscale eddies can have a significant impact on modulating low frequency variability (Berloff et al. 2007).

Figure 9 shows the error in variance for the different runs: deterministic, and stochastic with different moments for the reconstruction of the cPDFs. The ability of the stochastic parametrization to reinforce or introduce variance is clear at all resolutions, even at very coarse resolution. There is some benefit to include non-zero skewness and kurtosis of the cPDFs. For the run at 120 km horizontal resolution, the injection of non-Gaussian stochastic forcing leads to a non-zero contribution to the variance. The deterministic parametrization,

on the other hand, increases the variance of the eddy permitting runs only, with no visible impact for the coarse resolution run.

5. Discussion

In the previous section, we showed that the parametrization improves the energy spectrum as a function of frequency and wavenumber, and the jet strength and its separation point, while keeping the model numerically stable. We now discuss several aspects of vorticity, energy and enstrophy budget to provide some insights concerning the results described above. We also propose an implementation for a primitive equation model based on our results. Most aspects are discussed in the context of the deterministic parametrization but the stochasticity is discussed as well since its inclusion is beneficial to the representation of the mean and variance.

a. Vorticity forcing and Stability Criteria

Consider the parametrized potential vorticity equation in spectral space by taking the Fourier Transform of (Eq. 7), we obtain (similarly to Eq. 12 in PMZ14)

$$\begin{aligned} \widehat{\frac{\mathrm{D}q}{\mathrm{D}t}} &= |\kappa|\mathbf{K}^2\widehat{\frac{\mathrm{D}q}{\mathrm{D}t}} + \mathbf{K}^6\nu\widehat{\psi} + \widehat{F}^{\mathrm{wind}}, \\ &= \frac{1}{1-|\kappa|\mathbf{K}^2}\left(\mathbf{K}^6\nu\widehat{\psi} + \widehat{F}^{\mathrm{wind}}\right), \end{aligned} \tag{12}$$

where $\kappa < 0$, i.e. $\kappa = -|\kappa|$; $\widehat{f}(\mathbf{K},t)$ is the spatial Fourier transform of $f(\mathbf{x},t)$, $\mathbf{x}$ is the position vector, $\mathbf{K} = (k,l)$ is the total 2D wavenumber with modulus $K = |\mathbf{K}| = \sqrt{k^2+l^2}$, and $\widehat{\nabla f} = \mathrm{i}\mathbf{K}\widehat{f}$. The argument $(\mathbf{K},t)$ is omitted for convenience. As the total wavenumber K increases (i.e., the wavelength decreases), the amplitude of the Fourier transform of PV forcing increases (see Eq. 12 and Fig. 9 of PMZ14).

The value of κ in the simulations required some tuning to keep the model numerically stable and this is in part a direct consequence of Eq. 12. The amplitude of the Fourier

Transform of PV forcing has a singularity at $\mathbf{K}^2 = \frac{1}{|\kappa|}$, for a length scale of $\sqrt{|\kappa|}$. However, the maximum wavenumber the model can resolve is equal to $K_{max} = 2\pi/2\Delta x = \pi/\Delta x$. Therefore avoiding the singularity requires that $1 - |\kappa| K_{max}^2 > 0$, which using $|\kappa| = (\alpha \Delta x)^2$ leads to the condition $(\alpha \Delta x)^2 < K_{max}^2 = \Delta x^2/\pi^2$. The stability condition is thus $\alpha < 1/\pi = 0.318$, justifying the choice of $\alpha = 0.31$ for numerical stability (rather than the diagnosed value of roughly 0.45 from PMZ14).

The tuning of κ is also sensitive to the sub-grid viscous dissipation, and the likely build-up of small scale enstrophy (see below). For a Laplacian viscous term (as used in PMZ14), the maximum possible values of κ for the deterministic and stochastic implementations are $\kappa = -(0.27\Delta x)^2$ and $\kappa = -(0.29\Delta x)^2$, respectively. One can increase κ if the viscosity coefficient is also increased, but the influence of the parametrization is then damped by the dissipation, making it therefore inefficient to represent up-gradient momentum fluxes. Numerical stability is improved by raising the power of the Laplacian operator, which increases the scale-selective behaviour of the dissipative term, allowing for a stronger forcing with larger values of $|\kappa|$. For the stochastic implementations, the effective value of parametrization (henceforth of $|\kappa|$) is always larger than that of the deterministic case. This reflects the fact that stochastic fluctuations can have a stabilizing effect on the model solution. As pointed out in Palmer (2012), the tuning of a stochastic parametrization should not be done starting from the deterministic version.

b. Relationship of PMZ14 to hyperviscosity

We can approximate the effective viscosity term, i.e. the combination of the biharmonic viscosity and the RE parametrization, in Eq. 12 as

$$\begin{aligned} \widehat{F}_{\nu,subgrid} &= \frac{1}{1 - \mathbf{K}^2|\kappa|} \left(\mathbf{K}^6 \nu \widehat{\psi} \right) \\ &= \left(1 + |\kappa|\mathbf{K}^2 + \left(|\kappa|\mathbf{K}^2\right)^2 + \left(|\kappa|\mathbf{K}^2\right)^3 + ... \right) \left(\mathbf{K}^6 \nu \widehat{\psi} \right) \\ &= \mathbf{K}^6 \nu \widehat{\psi} \sum_{n=0}^{\infty} (|\kappa|\mathbf{K}^2)^n \end{aligned} \tag{13}$$

using a Taylor expansion $1/(1-x) = 1 + x + x^2 + x^3 + \cdots$, valid for $|x| < 1$. The subgrid effective dissipation differs in its dependence on $\mathbf{K}$ from that in the unparametrized equation. The infinite series for the effective viscosity corresponds to a dissipative behaviour on the Fourier amplitude of $\mathrm{D}q/\mathrm{D}t$, reminiscent of hyperdiffusive closures (e.g. $\nu\nabla^6\psi$, $\nu\nabla^8\psi$ ...), hence removing energy and enstrophy at small scales. Recall from Eq. 12 that avoiding the singularity requires the value of κ to be chosen so that $|\kappa|\mathbf{K}^2 < 1$ for any wavenumber $\mathbf{K}$ resolved by the model. We therefore have $(|\kappa|\mathbf{K}^2)^n < 1$ for $n \geq 0$, and hence the magnitude of the terms in the series of Eq. (13) decreases with increasing n. The effective viscosity in spectral space thus corresponds to a series of hyperviscous terms of increasing order, where the first and largest term is the explicitly specified hyperviscosity (as defined in Sec. 2) and the subsequent terms could be characterized as higher-order hyperviscous corrections.

Hyperviscous closures are motivated by the desire to enhance dissipation at very small scales while retaining as much resolved turbulence as possible in the slightly larger scales, i.e. to create a sharper transition between regions of spectral space in which advective and dissipative behaviour dominates (Holloway 1992). The addition of higher-order hyperviscous corrections (in spectral space) should further enhance the dissipation at the smallest resolved scales, with increasingly larger effect as the scale decreases. The effect of the parametrization can be interpreted as modifying the scale-selectivity of the dissipative term. Unlike hyperviscosity, the parametrization (Eq. 5) involves nonlinear terms and a temporal derivative. Another interpretation, given in PMZ14 considering the terms in Eq. 5, is to view the parametrization as acting to enhance or weakens flow parcels depending on their "history".

As discussed in Sec. 3, a practical workaround to implement the parametrization is to use $\frac{\mathrm{D}q}{\mathrm{D}t}$ at the previous timestep, so that the timestepping can remain explicit. The parametrized model equation, (1), can be written as

$$\begin{aligned} \frac{\mathrm{D}q}{\mathrm{D}t}(t) &= \mathcal{D}(t) + F^{\mathrm{wind}}(t) + F^{\mathrm{eddy}}(t-\Delta t), \\ &= \mathcal{D}(t) + F^{\mathrm{wind}}(t) + \kappa\nabla^2\frac{\mathrm{D}q}{\mathrm{D}t}(t-\Delta t). \end{aligned} \tag{14}$$

Applying (14) recursively we obtain

$$\begin{aligned}\frac{\mathrm{D}q}{\mathrm{D}t}(t) &= \mathcal{D}(t) + F^{\mathrm{wind}}(t) + \kappa\nabla^2\left[\mathcal{D}(t-\Delta t) + F^{\mathrm{wind}}(t-\Delta t) + \kappa\nabla^2\frac{\mathrm{D}q}{\mathrm{D}t}(t-2\Delta t)\right] \\ &= \mathcal{D}(t) + F^{\mathrm{wind}}(t) + \kappa\nabla^2\left[\mathcal{D}(t-\Delta t) + F^{\mathrm{wind}}(t-\Delta t)\right] \\ &\quad + \kappa^2\nabla^4\left[\mathcal{D}(t-2\Delta t) + F^{\mathrm{wind}}(t-2\Delta t) + \kappa\nabla^2\frac{\mathrm{D}q}{\mathrm{D}t}(t-3\Delta t)\right] = \ldots\end{aligned}$$

leading to

$$\frac{\mathrm{D}q}{\mathrm{D}t}(t) = \sum_{n=0}^{m}(\kappa\nabla^2)^n\left[\mathcal{D}(t-n\Delta t) + F^{\mathrm{wind}}(t-n\Delta t)\right], \tag{15}$$

where for $n = 0$ we set $(\kappa\nabla^2)^0 \equiv 1$. The total forcing of potential vorticity at time t involves successive applications of the operator $\kappa\nabla^2$ to the wind forcing F^{wind} and dissipation $\mathcal{D}$ (as defined in Sec. 2) at previous times.

We can neglect terms involving $\nabla^{2n}F^{\mathrm{wind}}$, for $n \geq 1$, assuming that the wind forcing is large-scale and constant in time, so that such higher order Laplacians are increasingly small. The evolution equation (15) can then be written as

$$\begin{aligned}\frac{\mathrm{D}q}{\mathrm{D}t}(t) &= F^{\mathrm{wind}}(t) - \sum_{n=0}^{m}(\kappa\nabla^2)^n\nu\nabla^6\psi(t-n\Delta t) \\ &= F^{\mathrm{wind}}(t) - \nu\nabla^6[\psi(t) + \kappa\nabla^2\psi(t-\Delta t) + \kappa^2\nabla^4\psi(t-2\Delta t) + \cdots] \\ &= F^{\mathrm{wind}}(t) - \nu\nabla^6[\psi(t) - |\kappa|\nabla^2\psi(t-\Delta t) + |\kappa|^2\nabla^4\psi(t-2\Delta t) + \cdots],\end{aligned} \tag{16}$$

where the bottom drag contribution to $\mathcal{D}$ has been ignored, and the last line uses the fact that $\kappa < 0$. In spectral space we then have

$$\widehat{\frac{\mathrm{D}q}{\mathrm{D}t}} = \widehat{F}^{\mathrm{wind}} + \nu\mathbf{K}^6\left[\widehat{\psi}(t) + |\kappa|\mathbf{K}^2\widehat{\psi}(t-\Delta t) + |\kappa|^2\mathbf{K}^4\widehat{\psi}(t-2\Delta t) + \cdots\right] \tag{17}$$

Since $\Delta t \neq 0$, the relative signs of $\widehat{\psi}(t)$, $\widehat{\psi}(t-\Delta t)$, $\widehat{\psi}(t-2\Delta t)$... can differ, and the term need not be strictly dissipative. This contrasts with the behaviour expected when $\mathrm{D}q/\mathrm{D}t$ is evaluated at the present timestep ($\Delta t = 0$), as the different terms would act as hyperviscosity (see Eq. 13). However, $\widehat{\psi}(t)$ in most cases would have the same sign as $\widehat{\psi}(t-\Delta t)$, assuming an AR1 process with memory of order the eddy timescale, such that Eq.17 does behave as a dissipative term.

An infinite series of terms is not warranted for implementation in a numerical model nor might be valid for all scales relevant here (indeed, our attempts of using a selected number of terms showed that the simulations did not converge). However, the series did provide some insights. We interpret the infinite series of hyperviscous terms as a change in the scale-selective dissipation of the flow, modifying the overall dissipation and the interaction between scales. However, improvements in the simulation are not solely due to changes in the overall dissipation, but also via an upscale energy transfer.

c. Energy and Enstrophy

Since the energy is spuriously being dissipated at small-scales in non-eddy resolving models, inhibiting energy backscatter from small to large scales (Jansen and Held 2014), this energy needs to be reinjected on average by our parametrization. The energy tendency due to the parametrization is given by

$$\begin{aligned} \frac{\partial E}{\partial t} &\propto -\psi F^{\text{eddy}} \\ &= \psi|\kappa|\nabla^2\frac{\mathrm{D}q}{\mathrm{D}t}, \end{aligned} \tag{18}$$

where E is the sum of the kinetic energy, $\frac{1}{2}(\nabla\psi)^2$, and APE $\frac{f^2}{2N^2}(\partial_z\psi)^2$ for all three layers. The energy tendency due to the parametrization is then the product of ψ and the potential vorticity eddy forcing, $|\kappa|\nabla^2\frac{\mathrm{D}q}{\mathrm{D}t}$. The nonlinear energy tendency term due to the parametrization can be further interpreted as follows.

Suppose that q following a parcel is increasing, $\frac{\mathrm{D}q}{\mathrm{D}t} > 0$, and that $\nabla^2\frac{\mathrm{D}q}{\mathrm{D}t} < 0$. If this occurs on a cyclonic eddy, i.e. $\psi < 0$, then $\psi\nabla^2\frac{\mathrm{D}q}{\mathrm{D}t} > 0$. The parametrized energy term $|\kappa|\psi\nabla^2\frac{\mathrm{D}q}{\mathrm{D}t}$ is thus positive, yielding an energy source. Correspondingly, $\nabla^2\frac{\mathrm{D}q}{\mathrm{D}t} > 0$ and $\psi > 0$ yield an energy source for an anticyclonic eddy. This suggests that the effect of the closure is to increase the kinetic energy of the flow when the potential vorticity tendency on a parcel has the same sign as the relative vorticity of the parcel. Assuming that the effect of the eddies is small compared to the wind forcing and dissipation, leading to $\frac{\mathrm{D}q}{\mathrm{D}t} \approx F^{\text{wind}} + \mathcal{D}$, the effect

of the eddies will be to amplify the energy input of the net forcing, i.e. the residual of the wind forcing plus the dissipation (with no effect if these are zero). Likewise, if $\nabla^2 \frac{\mathrm{D}q}{\mathrm{D}t}$ has opposite sign to the vorticity, there will be an energy sink, suggesting the parameterisation should act to amplify the dissipation of energy. Our parametrization therefore requires some input of energy from the wind (or buoyancy) forcing, in addition to the presence of small scale deformation (as discussed earlier). The input of energy will then modify the velocity, and the nonlinear advective term, leading to an indirect energy transfer between scales. We argued above that the effect of the closure on the dissipation resembles the effects of a hyperdiffusion on q. This suggests that it is dissipative for q under certain assumptions. The fact that it can behave as an energy source, but also dissipate enstrophy, is the desired qualitative behaviour for a parametrization of QG turbulence (Charney 1971). To further understand the contribution of the parametrization to the enstrophy budget, consider the enstrophy tendency due to the parametrization:

$$\frac{\partial G}{\partial t} \propto q\kappa\nabla^2 \frac{\mathrm{D}q}{\mathrm{D}t}, \tag{19}$$

where $G = \frac{1}{2}q^2$ is the enstrophy. Using the identity

$$\begin{aligned} \nabla \cdot \left(q \nabla \frac{\mathrm{D}q}{\mathrm{D}t} \right) &= \nabla \frac{\mathrm{D}q}{\mathrm{D}t} \cdot \nabla q + q \nabla \cdot \nabla \frac{\mathrm{D}q}{\mathrm{D}t} \\ &= \nabla \frac{\mathrm{D}q}{\mathrm{D}t} \cdot \nabla q + q \nabla^2 \frac{\mathrm{D}q}{\mathrm{D}t} \, , \end{aligned}$$

the enstrophy tendency (19) can be expressed as

$$\frac{\partial G}{\partial t} \propto \kappa \nabla \cdot \left(q \nabla \frac{\mathrm{D}q}{\mathrm{D}t} \right) - \kappa \nabla \frac{\mathrm{D}q}{\mathrm{D}t} \cdot \nabla q. \tag{20}$$

The first term on the right hand side of Eq. 19 does not contribute to the global enstrophy budget. The second term can act as a sink or source of enstrophy depending on the relative orientation of $\nabla \frac{\mathrm{D}q}{\mathrm{D}t}$ with ∇q. In a location where $\frac{\mathrm{D}q}{\mathrm{D}t}$ is tending to amplify q anomalies, $\nabla \frac{\mathrm{D}q}{\mathrm{D}t}$ will tend to point in a similar direction as ∇q so that $\nabla \frac{\mathrm{D}q}{\mathrm{D}t} \cdot \nabla q > 0$, yielding an enstrophy source, and similarly an enstrophy sink if the tendency is dissipative, meaning that $\frac{\mathrm{D}q}{\mathrm{D}t}$ is damping the q anomalies. Hence $\kappa < 0$ is required for self-consistency, and

the parametrization can provide both a source and sink of enstrophy. This demonstrates that both locally and globally enstrophy can build up in the model, further affecting the numerical stability of the model. We earlier highlighted that some tuning of κ was necessary to maintain numerical stability and that the value obtained (and thereby our results) were sensitive to the viscous term used ($\nabla^4\psi$ vs. $\nabla^6\psi$).

The improvement seen in the numerical simulations presented in this work are therefore due to a careful combination of the sub-grid eddy dissipation (here $\nu\nabla^6\psi$) and the eddy parametrization ($\kappa\nabla^2\frac{\mathrm{D}q}{\mathrm{D}t}$). For the parametrization to be effective, we must ensure that the build-up of enstrophy at the grid-scale does not overwhelm the solution and that the nonlinear upscale energy transfer is not inhibited by the viscous terms.

The stochastic backscatter is more effective than the deterministic nonlinear backscatter. It is difficult to disentangle the reasons for the differences. However, we hypothesize that it is the result of a combination of several factors. As described previously, the noise has a stabilizing impact on the model simulations. The nonlinearity of the model equation might be pushing the stochastic model into a new (and more stable) regime that the unparametrized or parametrized deterministic models cannot attain. This can be a consequence of the sampling of the tails of the PDFs of eddy forcing (with non-zero kurtosis), which the deterministic parametrization cannot sample adequately. The infrequent but extreme values of eddy forcing can have a significant impact on the scale interaction in turbulent flows, as illustrated here. Furthermore, the probabilistic forcing might avoid the frequent forcing at the singularity, stabilizing the model.

d. Implementation into a primitive equation model

The implementation of our parametrization is tied to a QG potential vorticity equation. Current ocean climate models do carry momentum, temperature and salinity as prognostic variables but not potential vorticity which is only diagnosed. We propose two possible ways forward to implement the parametrization into a primitive equation model. The proposed

implementations below are defined by imposing a total QG potential vorticity forcing as given by Eq. 4. Each implementation provides advantages but also caveats.

Two-dimensional implementation: The parametrization can be implemented into the momentum equations, without adding a contribution into the buoyancy equation, if we assume that the eddy forcing is purely 2D. This does not mean that the buoyancy forcing is neglected but rather that it projects directly on the momentum. The u and v momentum tendencies due to the eddy parametrization can be expressed as follows

$$\begin{aligned} \frac{\mathrm{D}u}{\mathrm{D}t} &= -\kappa \frac{\partial}{\partial y}\frac{\mathrm{D}q}{\mathrm{D}t} \\ \frac{\mathrm{D}v}{\mathrm{D}t} &= \kappa \frac{\partial}{\partial x}\frac{\mathrm{D}q}{\mathrm{D}t}, \end{aligned} \tag{21}$$

where q is the QG potential vorticity as defined in Eq. 2.

The curl of the momentum tendency then leads to the following vorticity forcing

$$\frac{\partial}{\partial x}\frac{\mathrm{D}v}{\mathrm{D}t} - \frac{\partial}{\partial y}\frac{\mathrm{D}u}{\mathrm{D}t} = \kappa \left[\frac{\partial^2}{\partial x^2}\frac{\mathrm{D}q}{\mathrm{D}t} + \frac{\partial^2}{\partial y^2}\frac{\mathrm{D}q}{\mathrm{D}t} \right] = \kappa \nabla^2 \frac{\mathrm{D}q}{\mathrm{D}t}. \tag{22}$$

The newly calculated velocity components from Eq. 21 are the residual velocity components which include the resolved and eddy-parametrized contribution. The residual velocity can then be used to advect active and passive tracers in ocean climate models. Online calculations of the potential vorticity and its Lagrangian tendency are therefore necessary for the implementation in a primitive equation model. The stochastic component can be used as described before, however one might wish to diagnose the PDFs of $\frac{\partial}{\partial y}\frac{\mathrm{D}q}{\mathrm{D}t}$ and $\frac{\partial}{\partial x}\frac{\mathrm{D}q}{\mathrm{D}t}$, rather than $\nabla^2 \frac{\mathrm{D}q}{\mathrm{D}t}$. Other possibilities includes keeping only the relative vorticity in Eq. 21 (namely, $-\kappa \frac{\partial}{\partial y}\frac{\mathrm{D}\zeta}{\mathrm{D}t}$ and $\kappa \frac{\partial}{\partial x}\frac{\mathrm{D}\zeta}{\mathrm{D}t}$) for a slightly simpler deterministic implementation.

Three-dimensional implementation: We define the quasi-geostrophic potential vorticity q as

$$q = q_u + q_b \tag{23}$$

where q_u is the relative and planetary vorticity and q_b is the stretching vorticity such that

$$q_u = \nabla^2\psi + \beta y \tag{24}$$

$$q_b = \frac{\partial}{\partial z}\left(\frac{f_0^2}{N^2}\frac{\partial \psi}{\partial z}\right). \tag{25}$$

Let us define F, G and B as the forcing terms due to the parametrization on the right hand side of the zonal momentum, meridional momentum and buoyancy equations, respectively. The QG PV tendency due to eddy parametrization via the forcing terms F, G and B can be estimated by deriving an equation for QG PV and is therefore

$$\frac{\mathrm{D}q}{\mathrm{D}t} = \frac{\partial G}{\partial x} - \frac{\partial F}{\partial y} + \frac{\partial}{\partial z}\left(\frac{f_0}{N^2}B\right). \tag{26}$$

Let

$$F = -\kappa\frac{\partial}{\partial y}\frac{\mathrm{D}q_u}{Dt} - (\kappa - \kappa_b)\frac{\partial}{\partial y}\frac{\mathrm{D}q_b}{Dt} \tag{27}$$

$$G = \kappa\frac{\partial}{\partial x}\frac{\mathrm{D}q_u}{Dt} + (\kappa - \kappa_b)\frac{\partial}{\partial x}\frac{\mathrm{D}q_b}{Dt} \tag{28}$$

$$B = \kappa_b\nabla^2\frac{\mathrm{D}b}{\mathrm{D}t} \tag{29}$$

where $b = f_0\psi_z$ is the buoyancy which can be shown to yield the following QG PV forcing

$$\frac{\mathrm{D}q}{\mathrm{D}t} = \kappa\nabla^2\frac{\mathrm{D}q}{\mathrm{D}t}, \tag{30}$$

i.e. the PZM14 closure. The arbitrary parameter κ_b has been introduced, which partitions the buoyancy component of the forcing between the momentum and buoyancy equations. Choosing $\kappa_b = 0$ gives the closure suggested in Eq. 21. Choosing $\kappa_b = \kappa$ puts all of the forcing associated with q_b into B, i.e. into the buoyancy equation. To avoid spurious diapycnal mixing, the implementation should be performed on isopycnal surfaces, which would also provide a close match to the QG PV implementation.

The implementation of the PZM14 closure proposed in Eq. 27 is equivalent to the proposed implementation of Eden (2010) (note the parametrizations are not equivalent, only their implementations in primitive equation models). Eden (2010) proposes PV diffusion

closure as well as a buoyancy diffusion closure. Combining both PV and buoyancy closures yields an expression for the eddy momentum flux forcing term that appears in the zonal-momentum equation (in a zonal-mean channel model). Similarly to our approach in Eq. 21 and Eq. 27, Eden (2010) assumes a specific forcing form, showing that it gives the desired forcing in the PV equation. Eden (2010) finds that an additional term, the "gauge term", is required to satisfy the momentum constraint. For the PMZ14 closure, the gauge term can be any contribution to F and G that vanishes when $\frac{\partial G}{\partial x} - \frac{\partial F}{\partial y}$ is computed.

Our simulations do not have a direct buoyancy forcing, therefore it is difficult to argue which implementation would lead to more physically based results. Furthermore, it is difficult to decide how to choose κ_b sensibly. One could attempt to diagnose it from high-resolution simulation model, as in PMZ14. For example, diagnostic results from the QG high-resolution model indicate that momentum and buoyancy contributions to $\nabla^2 Dq/Dt$ are of similar magnitude (not shown), suggesting that $\kappa_b = \kappa$ would a physically reasonable choice. One could use high-resolution primitive equation models to further diagnose the different contributions of momentum and buoyancy forcing and validate the QG results, however this is beyond the scope of the present study.

Both implementations in QG, deterministic and stochastic, do show some sensitivity to the timestep being used, as any numerical simulation would. Yet, the sensitivity to the timestep in the current parametrization might be more pronounced since the implementation requires the knowledge of the time-tendency of nonlinear quantities. In the present work, we choose to keep the timestep of the high resolution truth and the low-resolution 30km runs identical. This was done to isolate the role of timestepping in the stability of the simulations. Runs with increased timesteps up to 4 times the present values lead to results identical to the ones presented here. As for the QG implementation, the Lagrangian derivative of potential vorticity is required in primitive equation, therefore there might be some sensitivity to the timestep chosen to integrate the model which will need to be investigated.

Another (related) approach for an implementation in primitive equation is to use the

different parts of the non-Newtonian Rivlin-Ericksen stress directly from the momentum equations, rather than the analogy from PMZ14. Anstey and Zanna (2016) have shown that the deformation part of the Rivlin-Ericksen stress in the momentum equation can dissipate enstrophy, conserve energy while mimicking very closely the effect of eddies onto the mean flow.

6. Summary

We presented a new parametrization of eddy-mean flow interaction for use in eddy permitting models. The closure, based on a form resembling a Rivlin-Ericksen stress which includes a deformation and a memory term of potential vorticity gradient, can be implemented as a deterministic or stochastic parametrization. When either formulation of the parametrization are implemented, we obtained a drastic improvement in the mean state and the variability of the low resolution models over all wavenumbers and frequencies.

The parametrization requires only resolved variables from the low resolution model, namely the Laplacian of the Lagrangian tendency of potential vorticity, $\nabla^2 \frac{\mathrm{D}q}{\mathrm{D}t}$. For the deterministic parametrization, the parameter (which has dimensions of length squared) depends only the low resolution grid size. For the stochastic parametrization, in addition to the low resolution grid box size, the maximum strength of the wind stress (or the variance in PV under any forcing - mechanical or thermodynamical) and the local stratification are necessary input to estimate the contribution of the eddy forcing. The deterministic and stochastic parametrizations are therefore scale-aware, and flow-aware due to the necessary evaluation of the local $\nabla^2 \frac{\mathrm{D}q}{\mathrm{D}t}$.

The present results indicate there is a cut-off resolution at which the deterministic parametrization and the stochasticity do not impact the mean flow. The cut-off is around the Rossby radius of deformation, but further tests are required to ascertain this value. For horizontal resolutions smaller than the cut-off value, the model produces instabilities gen-

643 erating barotropic and baroclinic eddies. The effects of these eddies can be enhanced by
644 the parametrization, affecting local and global momentum, energy and enstrophy budgets.
645 The parametrization manages to overcome dissipation, especially in regions where the eddy
646 forcing has a strong impact on the larger scale flow.

647 The parametrization captures some key ingredients of geotrosphic turbulence such as jet
648 sharpening and upgradient momentum fluxes, energy backscatter and enstrophy dissipation.
649 It requires only spatial and temporal derivatives already computed by the model. Some
650 numerical stability criteria must be respected and would need to be tested in more complex
651 models. However in QG, the parametrization also appears to have a stabilizing effect on
652 the model. It might be necessary to revisit some of the initial assumption when consider-
653 ing primitive equations such as decomposing the forcing into momentum and buoyancy as
654 described in the discussion, or combine the parametrization with an energy or an enstrophy
655 equation (Marshall and Adcroft 2010; Marshall et al. 2012).

656 The stochastic backscatter is shown to be a more efficient and a more stable eddy
657 parametrization than its deterministic counterpart. Stochastic parametrizations for convec-
658 tion have been fairly routine in atmospheric models (Raisanen et al. 2004; Plant and Craig
659 2008), especially in the grey zone (analagous to the horizontal resolution cut-off discussed
660 above for eddy-mean interaction). Stochastic parametrizations in primitive equation ocean
661 models, mainly at coarse- non-eddying resolution, are slowly being implemented showing
662 various degrees of success (e.g., Brankart 2013; Andrejczuk et al. 2016; Williams et al. 2016;
663 Grooms 2016; Juricke et al. 2017). The encouraging results presented here reinforces the need
664 for developing and implementing scale- and flow-aware stochastic ocean parametrization in
665 ocean climate models.

Acknowledgements

LZ thanks Eli Tziperman, Jim McWilliams, Tapio Schneider, David Randall and David Neelin for useful discussions and suggestions. We thank the reviewers and the editor for their comments which have helped improved the manuscript. This work was funded by the John Fell Oxford University Press (OUP) Research Fund and NERC grant NE/K013548/1.

References

Andrejczuk, M., Cooper, F. C., Juricke, S., Palmer, T. N., Weisheimer, A., Zanna, L., 2016. Oceanic stochastic parameterizations in a seasonal forecast system. Monthly Weather Review 144 (5), 1867–1875.

Anstey, J., Zanna, L., 2016. Deformation-based parametrization of mesoscale eddies. Oc. Modell. Submitted, https://www.dropbox.com/s/ocjkroio1y96p0v/AnsteyZannasubmitted.pdf?dl=0.

Arakawa, A., 1966. Computational design for long term numerical integration of the equations of fluid motion: Two dimensional incompressible flow. Part 1. Computational Physics 1, 119–143.

Arbic, B. K., Flierl, G. R., Scott, R. B., 2007. Cascade inequalities for forced-dissipated geostrophic turbulence. J. Phys. Oceanogr. 37 (6), 1470–1487.

Bachman, S., Fox-Kemper, B., 2013. Eddy parameterization challenge suite I: Eady spindown. Oc. Modell. 64, 12–28.

Bachman, S. D., Fox-Kemper, B., Pearson, B., 2016. A scale-aware subgrid model for quasigeostrophic turbulence. Ocean Modell. Submitted.

Berloff, P. S., 2005a. On dynamically consistent eddy fluxes. Dyn. Atmos. Oceans 38 (3–4), 123–146.

Berloff, P. S., 2005b. Random-forcing model of the mesoscale oceanic eddies. J. Fluid Mech. 529, 71–95.

Berloff, P. S., 2015. Dynamically consistent parameterization of mesoscale eddies. Part I: Simple model. Oc. Modell. 87, 1–19.

Berloff, P. S., 2016. Dynamically consistent parameterization of mesoscale eddies. Part II: Eddy fluxes and diffusivity from transient impulses. Fluids 1 (22), doi:10.3390/fluids1030022.

Berloff, P. S., Hogg, A. M. C., Dewar, W., 2007. The turbulent oscillator: A mechanism of low-frequency variability of the wind-driven ocean gyres. J. Phys. Oceanogr. 37 (9), 2363–2386.

Brankart, J. M., 2013. Impact of uncertainties in the horizontal density gradient upon low resolution global ocean modelling. Oc. Modell. 66, 64–76.

Charney, J., 1971. Geostrophic Turbulence. J. Atmos. Sci. 28 (6), 1087.

Duan, J., Nadiga, B. T., 2007. Stochastic parameterization for large eddy simulation of geophysical flows. Proc. Am. Math. Soc. 135 (4), 1187–1196.

Eden, C., 2010. Parameterising meso-scale eddy momentum fluxes based on potential vorticity mixing and a gauge term. Ocean Modell. 32 (1–2), 58–71.

Ericksen, J. L., 1956. Stress deformation relations for solids. Can. J. Phys. 34 (2), 226–227.

Fox-Kemper, B., Bachman, S., Pearson, B., Reckinger, S., July 2014. Principles and advances in subgrid modeling for eddy-rich simulations. CLIVAR Exchanges 19 (2), 42–46.

Fox-Kemper, B., Menemenlis, D., 2008. Can large eddy simulation techniques improve mesoscale-rich ocean models? In: Hecht, M., Hasumi, H. (Eds.), Ocean Modeling in an Eddying Regime. Vol. 177. AGU Geophysical Monograph Series, pp. 319–338.

Frederiksen, J. S., Davies, A. G., 1997. Eddy viscosity and stochastic backscatter parameterizations on the sphere for atmospheric circulation models. J. Atmos. Sci. 54 (20), 2475–2492.

Frisch, U., Kurien, S., Pandit, R., Pauls, W., Ray, S. S., Wirth, A., Zhu, J.-Z., 2008. Hyperviscosity, Galerkin truncation, and bottlenecks in turbulence. Phys. Rev. Lett. 101 (14), 144501, arXiv:0803.4269.

Gent, P. R., McWilliams, J. C., 1990. Isopycnal mixing in ocean circulation models. J. Phys. Oceanogr. 20 (1), 150–155.

719 Gent, P. R., Willebrand, J., McDougall, T. J., McWilliams, J. C., 1995. Parameterizing
720 eddy-induced tracer transports in ocean circulation models. J. Phys. Oceanogr. 25 (4),
721 463–474.

722 Gnanadesikan, A., Hallberg, R., 2000. On the relationship of the Circumpolar Current to
723 Southern Hemisphere winds in coarse-resolution ocean models. J. Phys. Oceanogr. 30 (8),
724 2013–2034.

725 Graham, J. P., Ringler, T., 2013. A framework for the evaluation of turbulence closures used
726 in mesoscale ocean large-eddy simulations.

727 Greatbatch, R. J., Zhai, X., Claus, M., Czeschel, L., Rath, W., 2010. Transport driven by
728 eddy momentum fluxes in the Gulf Stream Extension region. Geophys. Res. Lett. 37.

729 Griffies, S. M., Hallberg, R. W., 2000. Biharmonic friction with a Smagorinsky viscosity for
730 use in large-scale eddy-permitting ocean models. Mon. Weath. Rev. 128, 2935–2946.

731 Grooms, I., 2016. A Gaussian-product stochastic Gent-McWilliams parameterization. Ocean
732 Modell. 106, 27–43.

733 Grooms, I., Majda, A. J., 2013. Efficient stochastic superparameterization for geophysical
734 turbulence. Proc. Natl. Acad. Sci. (USA) 110 (12), 4464–4469.

735 Hallberg, R., 2013. Using a resolution function to regulate parameterizations of oceanic
736 mesoscale eddy effects. Oc. Modell. 72, 92–103.

737 Holloway, G., 1992. Representing topographic stress for large-scale ocean models. J. Phys.
738 Oceanogr. 22 (9), 1033–1046.

739 Holm, D. D., Nadiga, B. T., 2003. Modeling mesoscale turbulence in the barotropic double-
740 gyre circulation. J. Phys. Oceanogr. 33 (11), 2355–2365.

741 Holm, D. D., Wingate, B. A., 2005. Baroclinic instabilities of the two-layer quasigeostrophic
742 alpha model. J. Phys. Oceanogr. 35 (7), 1287–1296.

743 IPCC, 2013. Summary for Policymakers. In: Stocker, T. F., Qin, D., Plattner, G.-K., Tignor,
744 M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P. M. (Eds.),
745 Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the
746 Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge
747 University Press, Cambridge, United Kingdom and New York, NY, USA.

748 Jansen, M. F., Held, I. M., 2014. Parameterizing subgrid-scale eddy effects using energetically
749 consistent backscatter. Oc. Modell. 80, 36–48.

750 Juricke, S., Palmer, T. N., Zanna, L., 2017. Stochastic sub-grid scale ocean mixing: Impacts
751 on low frequency variability. J. Climate Submitted.

752 Kitsios, V., Frederiksen, J. S., Zidikheri, M. J., 2012. Subgrid parameterisation of the eddy-
753 meanfield interactions in a baroclinic quasi-geostrophic ocean. ANZIAM J. 54, C459–C475.

754 Kraichnan, R. H., 1976. Eddy viscosity in two and three dimensions. J. Atmos. Sci. 33 (8),
755 1521–1536.

756 Larichev, V. D., Held, I. M., 1995. Eddy amplitudes and fluxes in a homogeneous model of
757 fully developed baroclinic instability. J. Phys. Oceanogr. 25 (10), 2285–2297.

758 Leith, C. E., 1990. Stochastic backscatter in a subgridscale model: Plane shear mixing layer.
759 Phys. Fluids A 2 (3), 297.

760 Marshall, D. P., Adcroft, A. J., 2010. Parameterization of ocean eddies: Potential vorticity
761 mixing, energetics and Arnold's first stability theorem. Oc. Modell. 32 (3-4, SI), 188–204.

762 Marshall, D. P., Maddison, J. R., Berloff, P. S., 2012. A framework for parameterizing eddy
763 potential vorticity fluxes. J. Phys. Oceanogr. 42 (4), 539–557.

764 Mead, L. R., Papanicolaou, N., 1984. Maximum entropy in the problem of moments. J.
765 Math. Phys. 25 (8), 2404–2417.

766 Nadiga, B. T., 2008. Orientation of eddy fluxes in geostrophic turbulence. Phil. Trans. R.
767 Soc. A 366 (1875), 2491–2510.

768 Nadiga, B. T., Bouchet, F., 2011. The equivalence of the Lagrangian-averaged Navier-Stokes-
769 α model and the rational large eddy simulation model in two dimensions. Phys. Fluids
770 23 (9), 095–105.

771 Palmer, T. N., 2012. Towards the probabilistic Earth-system simulator: a vision for the
772 future of climate and weather prediction. Q. J. R. Meteorol. Soc. 138 (665, B), 841–861.

773 Plant, R. S., Craig, G. C., 2008. A stochastic parameterization for deep convection based on
774 equilibrium statistics. J. Atmos. Sci. 65 (1), 87–105.

775 Porta Mana, P., Zanna, L., 2014. Toward a stochastic parameterization of ocean mesoscale
776 eddies. Oc. Modell. 79 (0), 1 – 20.

777 Raisanen, P., Barker, H. W., Khairoutdinov, M. F., Li, J., Randall, D. A., 2004. Stochastic
778 generation of subgrid-scale cloudy columns for large-scale models. Q. J. R. Meteorol. Soc.
779 130 (601), 2047–2067.

780 Rhines, P. B., 1977. The dynamics of unsteady currents. In: Goldberg, E. D., McCave, I. N.,
781 O'Brien, J. J., Steele, J. H. (Eds.), The Sea. Vol. 6: Marine Modeling. Wiley, New York,
782 Ch. 7, pp. 189–318.

783 Rivlin, R. S., 1957. The relation between the flow of non-Newtonian fluids and turbulent
784 Newtonian fluids. Q. Appl. Math. 15 (2), 212–215.

785 Salmon, R., 1978. Two-layer quasi-geostrophic turbulence in a simple special case. Geophys.
786 Astrophys. Fluid Dyn. 10 (1), 25–52.

787 Scott, R. B., Arbic, B. K., 2007. Spectral energy fluxes in geostrophic turbulence: Implica-
788 tions for ocean energetics. J. Phys. Oceanogr. 37 (3), 673–688.

789 Smagorinsky, J., 1963. General circulation experiments with the primitive equations, i, the
790 basic experiment. Mon. Weath. Rev. 91, 99–164.

791 Straub, D. N., Nadiga, B. T., 2014. Energy fluxes in the quasigeostrophic double gyre prob-
792 lem. J. Phys. Oceanogr. 44 (6), 1505–1522.

793 Taylor, K. E., Stouffer, R. J., Meehl, G. A., 2012. An {O}verview of {CMIP}5 and the
794 experiment design. Bull. Amer. Meteor. Soc. 93, 485–498.

795 Truesdell, III, C. A., Noll, W., 2004. The Non-Linear Field Theories of Mechanics, 3rd
796 Edition. Springer, Berlin, ed. by Stuart S. Antman; first publ. 1965.

797 Waterman, S., Jayne, S. R., 2012. Eddy-driven recirculations from a localized transient
798 forcing. J. Phys. Oceanogr. 42, 430–447.

799 Williams, P. D., 2009. A Proposed Modification to the Robert-Asselin Time Filter. Mon.
800 Weath. Rev. 137 (8), 2538–2546.

801 Williams, P. D., Howe, N. J., Gregory, J. M., Smith, R. S., Joshi, M. M., 2016. Improved
802 climate simulations through a stochastic parameterization of ocean eddies. J. Climate 29,
803 8763–8781.

List of Tables

1 Common Model parameters 37

2 Differing Model parameters 38

TABLE 1. Common Model parameters

	Parameter	Value
L	Basin width	3840 km
L_{Ro}*	Rossby Radii of Deformation	$(40, 23)$ km
H_1, H_2, H_3	Layer thicknesses	$(250\text{ m}, 750\text{ m}, 3000\text{ m})$
f_0	Planetary vorticity at mid-y	10^{-4} s^{-1}
β	$\mathrm{d}f/\mathrm{d}y$	$2 \times 10^{-11}\text{ m}^{-1}\text{s}^{-1}$
g	gravity	9.8 ms^{-2}
g'	Reduced gravity	$(0.034, 0.018)\text{ ms}^{-2}$
r	Bottom Drag	$4 \times 10^{-8}\text{ s}^{-1}$
τ_0	Wind stress	0.8 N/m^2
ρ_0	Reference Density	10^3kg/m^3
N	Brunt-Väisälä frequency at interfaces	$(6.82 \times 10^{-3}\text{ s}^{-1}, 2.56 \times 10^{-3}\text{ s}^{-1})$

* $1/L_{\mathrm{Ro}}^2 = \lambda$ where λ are eigenvalues of the equation $\partial_z\left(\frac{f_0^2}{N^2}\partial_z\psi\right) = \lambda\psi$.

TABLE 2. Differing Model parameters

Resolution Δx [km]	Timestep Δt [s]	Viscosity] ν [$m^4 s^{-1}$]
7.5	600	1.2×10^9
30	600	5.4×10^{10}
60	1400	2.0×10^{11}

Δx [km]	Δt [s]	ν [$m^2 s^{-1}$]
120	2000	200

List of Figures

1 Statistically steady state streamfunction ψ in the upper layer for the model runs at horizontal resolutions of (a) 7.5 km (eddy resolving model), (b) 30 km (eddy permitting model) with $F_m^{\text{eddy}} = 0$; (c) 30 km (eddy permitting model) with the deterministic parametrization, $F_m^{\text{eddy}} \neq 0$; (d) 30 km (eddy permitting model) with the stochastic parametrization, $F_m^{\text{eddy}} \neq 0$. 41

2 Snapshots of potential vorticity, q, in the upper layer for the model runs at horizontal resolutions of (a) 7.5 km (eddy resolving model), (b) 30 km (eddy permitting model) with $F_m^{\text{eddy}} = 0$; (c) 30 km (eddy permitting model) with the deterministic parametrization, $F_m^{\text{eddy}} \neq 0$; (d) 30 km (eddy permitting model) with the stochastic parametrization, $F_m^{\text{eddy}} \neq 0$. 42

3 Kinetic energy as a function of latitude at longitudes (a) $x = 120$km and (b) $x = 300$km for the following runs with horizontal resolution: 7.5 km (black), 30 km unparametrized (blue), 30 km with deterministic parametrization (red), 30 km with stochastic parametrization (grey). Eddy momentum stress divergence as a function of latitude at (c) $x = 120$km and (d) $x = 300$km, same legend as for (a)- (b). 43

4 Kinetic energy as a function of wavenumbers for the following runs with horizontal resolution: 7.5 km (black), 30 km unparametrized (blue), 30 km with deterministic parametrization (red), 30 km with stochastic parametrization (grey). 44

5 Spectral budget of total energy as a function of horizontal wavenumbers K for the following runs with horizontal resolution: (a) 7.5 km, (b) 30 km unparametrized, (c) 30 km with deterministic parametrization, (d) 30 km with stochastic parametrization. The different terms in the spectral energy budget are due to forcing (blue), dissipation by bottom friction (red), hyperviscosity (green), and spectral transfer of kinetic (violet) and available potential energy (yellow). All terms are normalised by the total energy and by dK/K. The black vertical lines denote the deformation scale. 45

6 Mean Error relative to the high resolution truth: (a) Root-mean-square error, defined as the L2 norm of the low resolution runs with respect to the high resolution truth, as a function of resolution; Difference between the time-mean streamfunction of the high-resolution simulation and the 30km low-resolution (b) unparametrized, (c) with deterministic parametrization, (d) with stochastic parametrization. 46

7 Standard Deviation of the Eddy Kinetic Energy for the following runs with horizontal resolution: (a) 7.5 km, (b) 30 km unparametrized, (c) 30 km with deterministic parametrization, (d) 30 km with stochastic parametrization. 47

8 Kinetic energy as a function of frequency for the following runs with horizontal resolution: 7.5 km (black), 30 km unparametrized (blue), 30 km with deterministic parametrization (red), 30 km with stochastic parametrization (grey). 48

9 Error in variance as a function of resolution for the different low-resolution runs: unparametrized (blue), deterministic parametrization (red), stochastic parametrization (grey), Gaussian stochastic parametrization (black, assuming skewness and kurtosis are negligible). 49

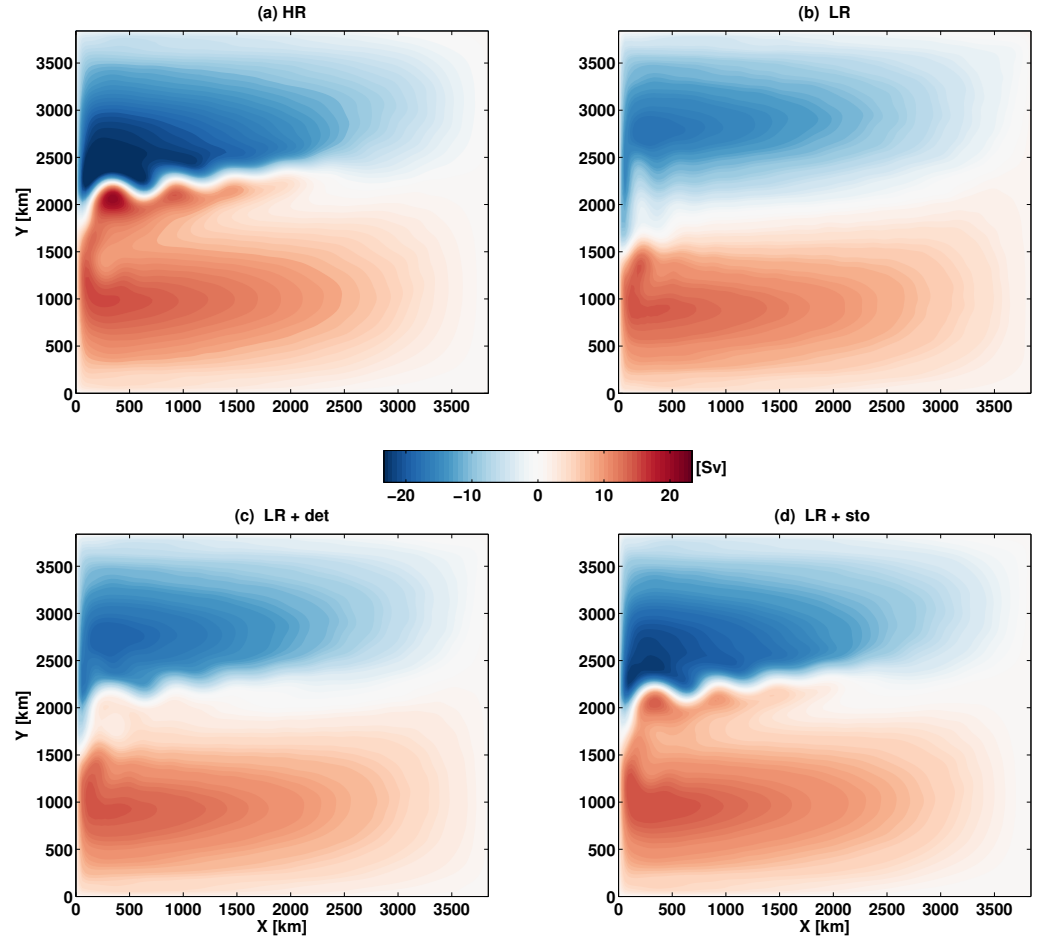

FIG. 1. Statistically steady state streamfunction ψ in the upper layer for the model runs at horizontal resolutions of (a) 7.5 km (eddy resolving model), (b) 30 km (eddy permitting model) with $F_m^{\text{eddy}} = 0$; (c) 30 km (eddy permitting model) with the deterministic parametrization, $F_m^{\text{eddy}} \neq 0$; (d) 30 km (eddy permitting model) with the stochastic parametrization, $F_m^{\text{eddy}} \neq 0$.

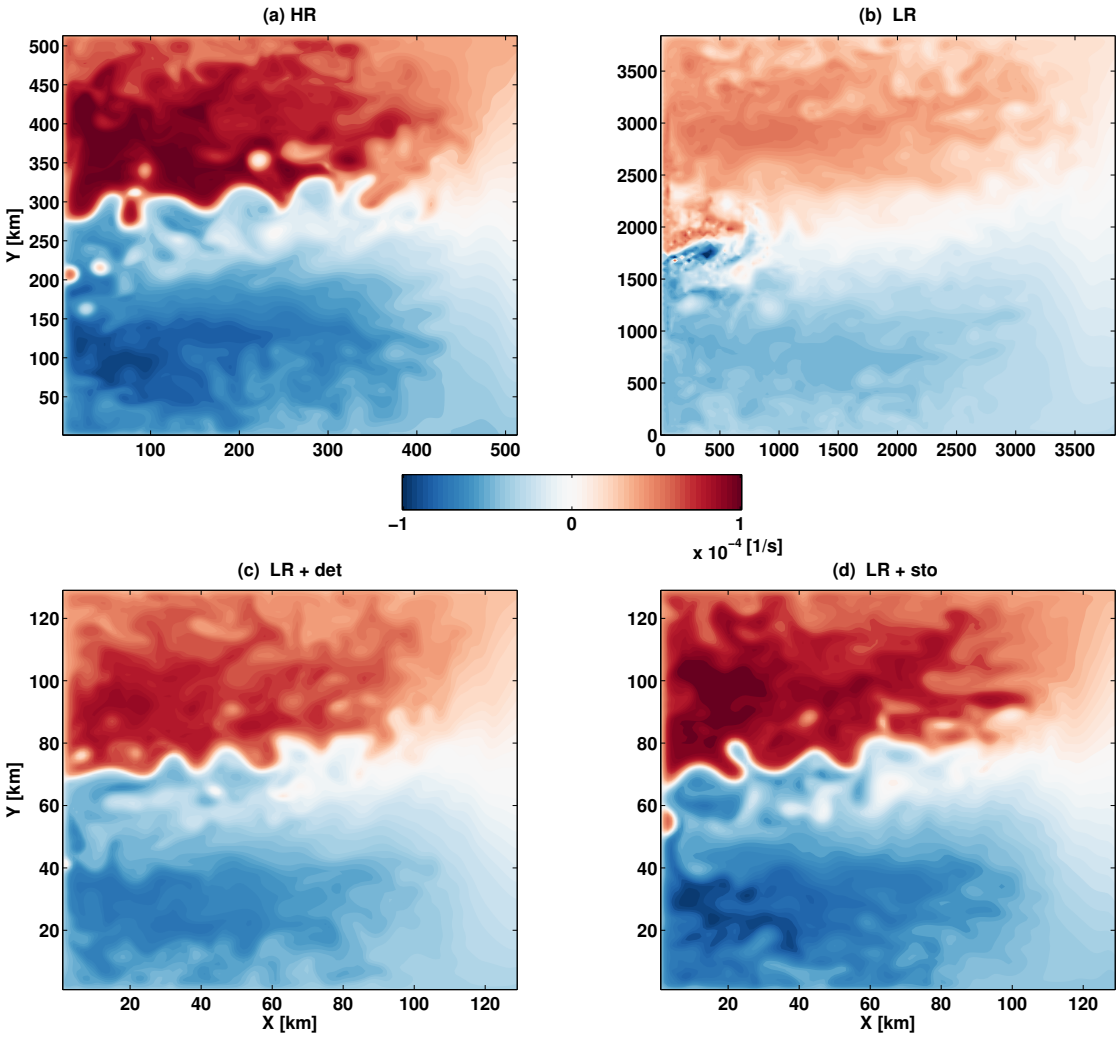

FIG. 2. Snapshots of potential vorticity, q, in the upper layer for the model runs at horizontal resolutions of (a) 7.5 km (eddy resolving model), (b) 30 km (eddy permitting model) with $F_m^{\text{eddy}} = 0$; (c) 30 km (eddy permitting model) with the deterministic parametrization, $F_m^{\text{eddy}} \neq 0$; (d) 30 km (eddy permitting model) with the stochastic parametrization, $F_m^{\text{eddy}} \neq 0$.

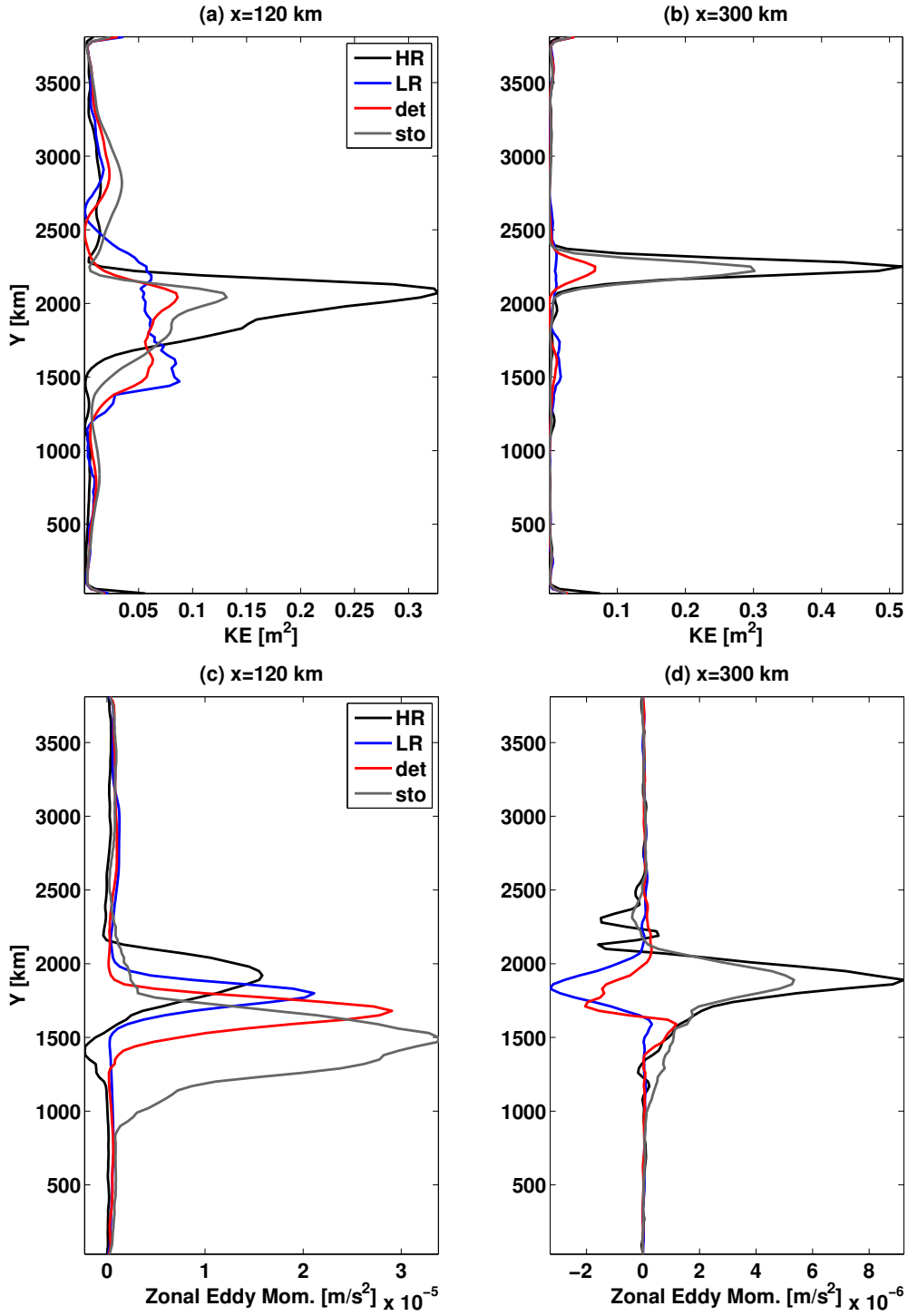

FIG. 3. Kinetic energy as a function of latitude at longitudes (a) $x = 120$km and (b) $x = 300$km for the following runs with horizontal resolution: 7.5 km (black), 30 km unparametrized (blue), 30 km with deterministic parametrization (red), 30 km with stochastic parametrization (grey). Eddy momentum stress divergence as a function of latitude at (c) $x = 120$km and (d) $x = 300$km, same legend as for (a)- (b).

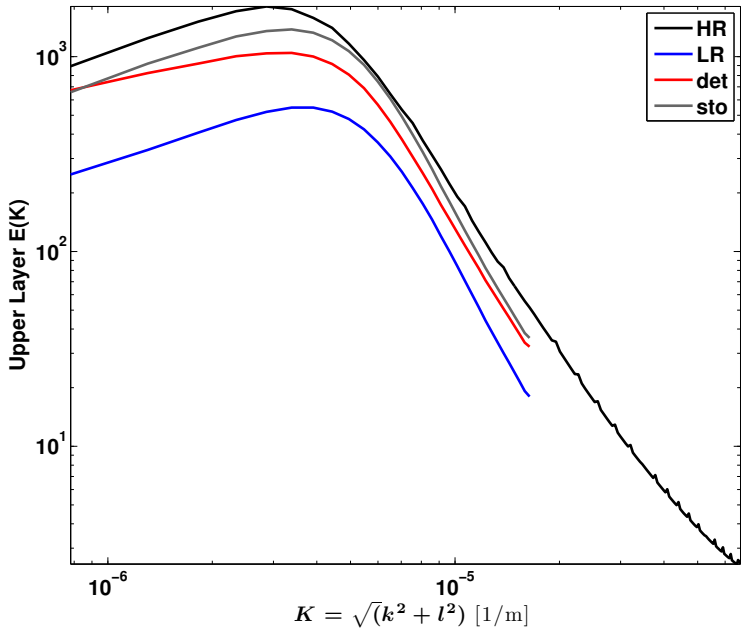

FIG. 4. Kinetic energy as a function of wavenumbers for the following runs with horizontal resolution: 7.5 km (black), 30 km unparametrized (blue), 30 km with deterministic parametrization (red), 30 km with stochastic parametrization (grey).

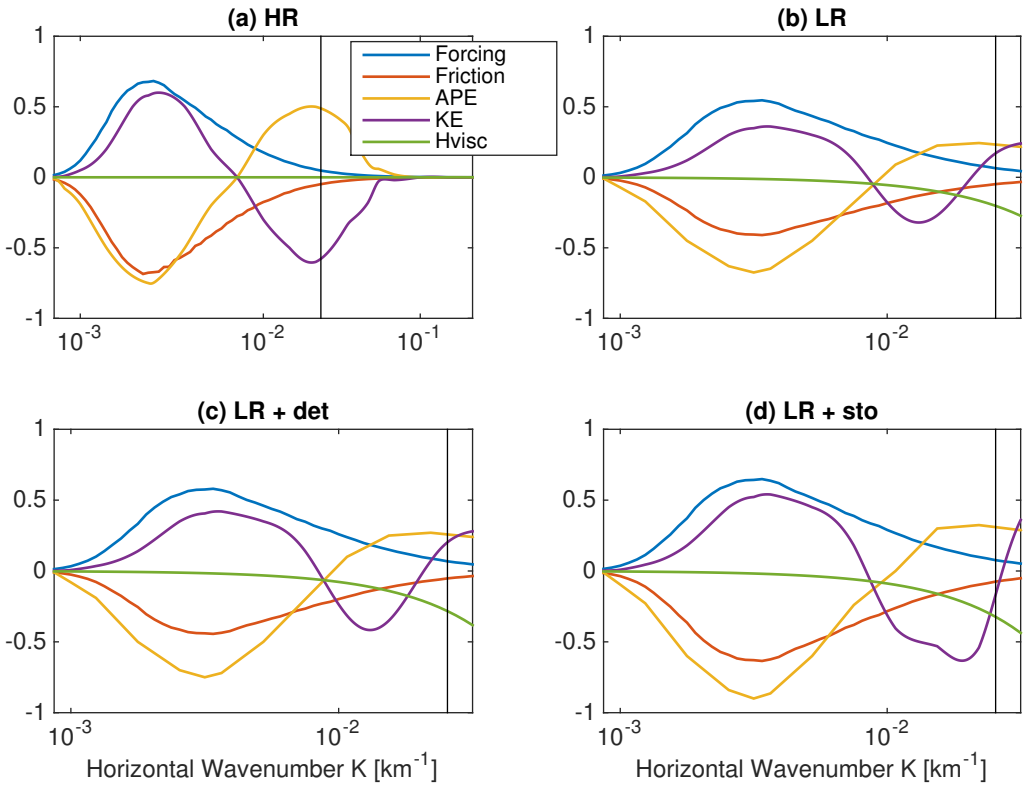

FIG. 5. Spectral budget of total energy as a function of horizontal wavenumbers K for the following runs with horizontal resolution: (a) 7.5 km, (b) 30 km unparametrized, (c) 30 km with deterministic parametrization, (d) 30 km with stochastic parametrization. The different terms in the spectral energy budget are due to forcing (blue), dissipation by bottom friction (red), hyperviscosity (green), and spectral transfer of kinetic (violet) and available potential energy (yellow). All terms are normalised by the total energy and by dK/K. The black vertical lines denote the deformation scale.

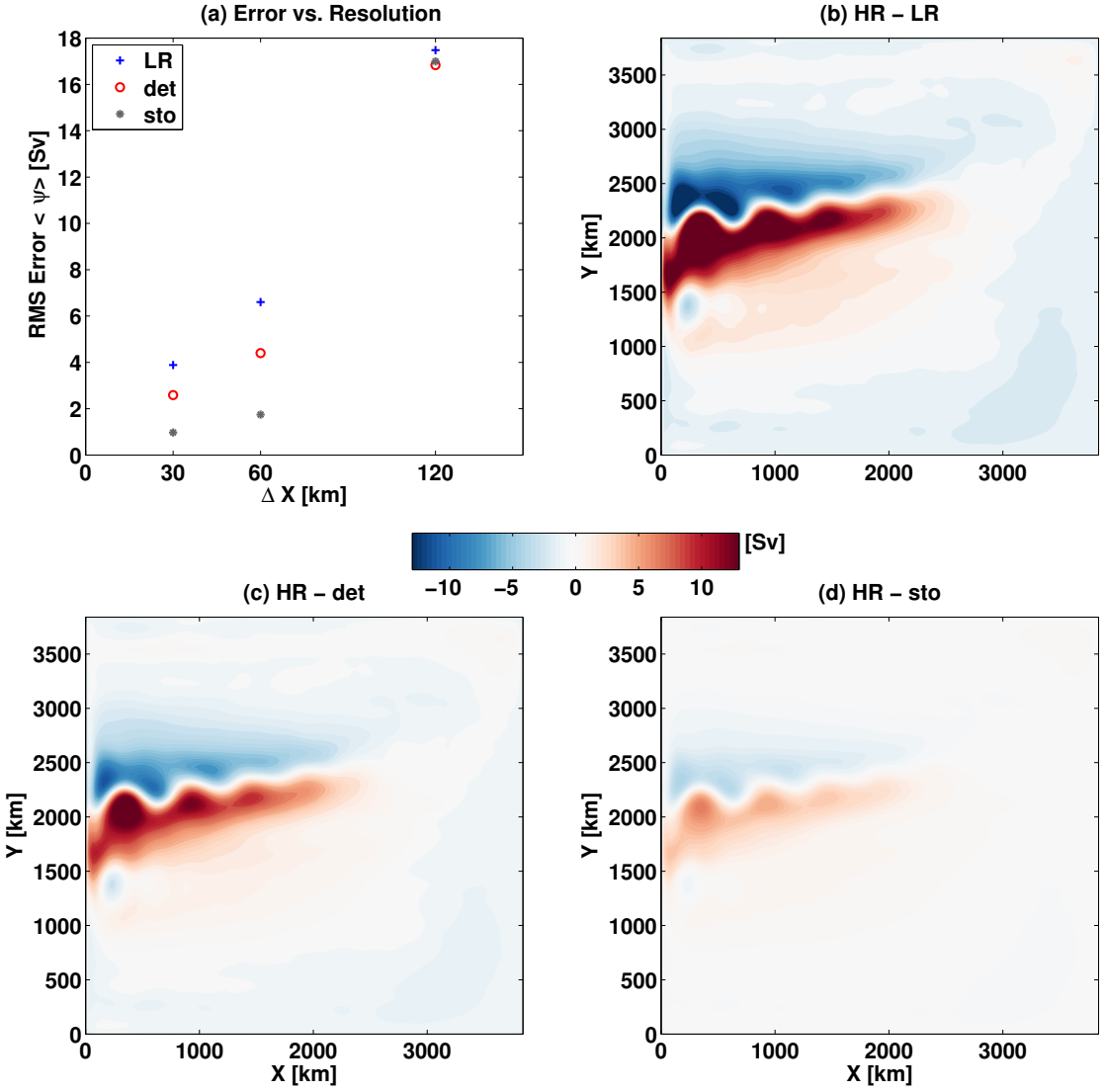

FIG. 6. Mean Error relative to the high resolution truth: (a) Root-mean-square error, defined as the L2 norm of the low resolution runs with respect to the high resolution truth, as a function of resolution; Difference between the time-mean streamfunction of the high-resolution simulation and the 30km low-resolution (b) unparametrized, (c) with deterministic parametrization, (d) with stochastic parametrization.

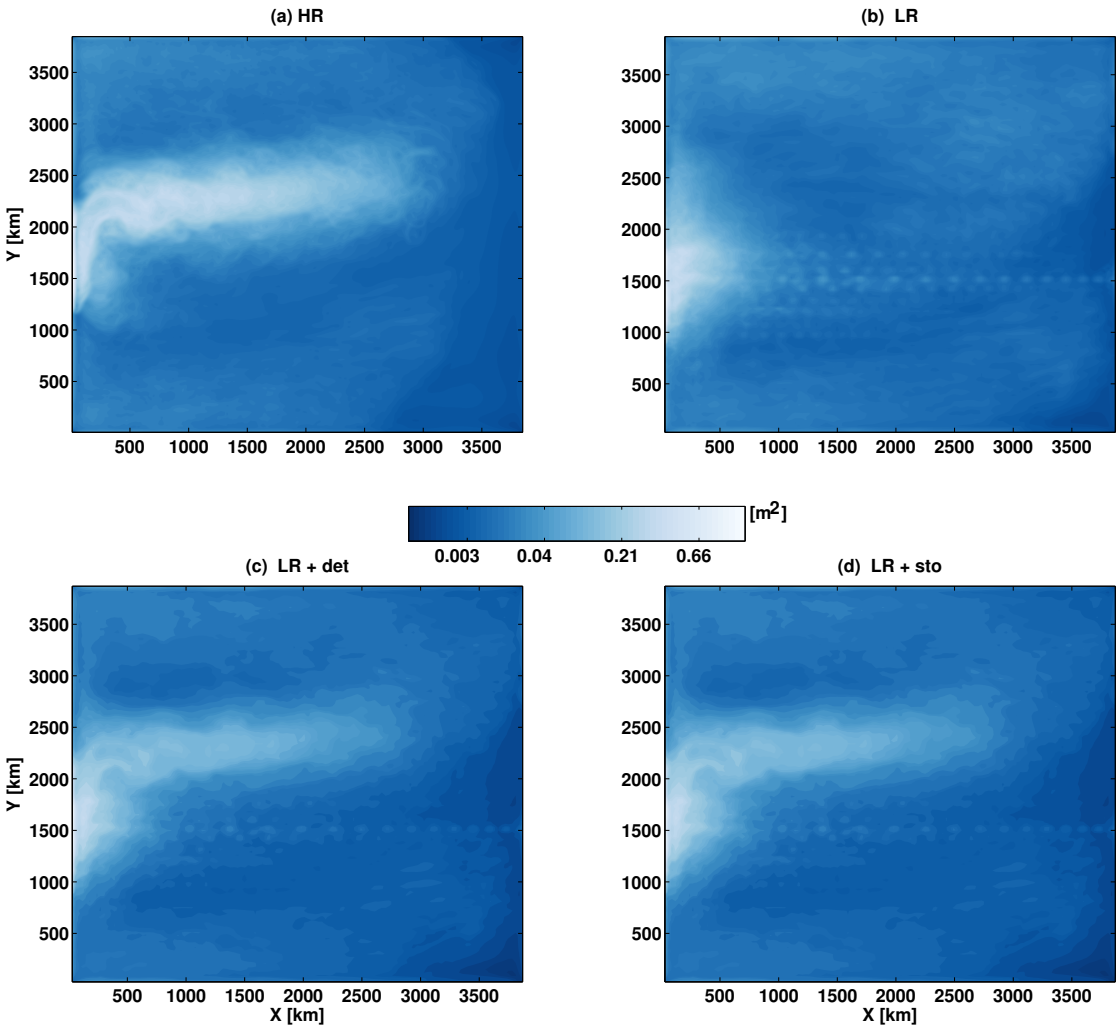

FIG. 7. Standard Deviation of the Eddy Kinetic Energy for the following runs with horizontal resolution: (a) 7.5 km, (b) 30 km unparametrized, (c) 30 km with deterministic parametrization, (d) 30 km with stochastic parametrization.

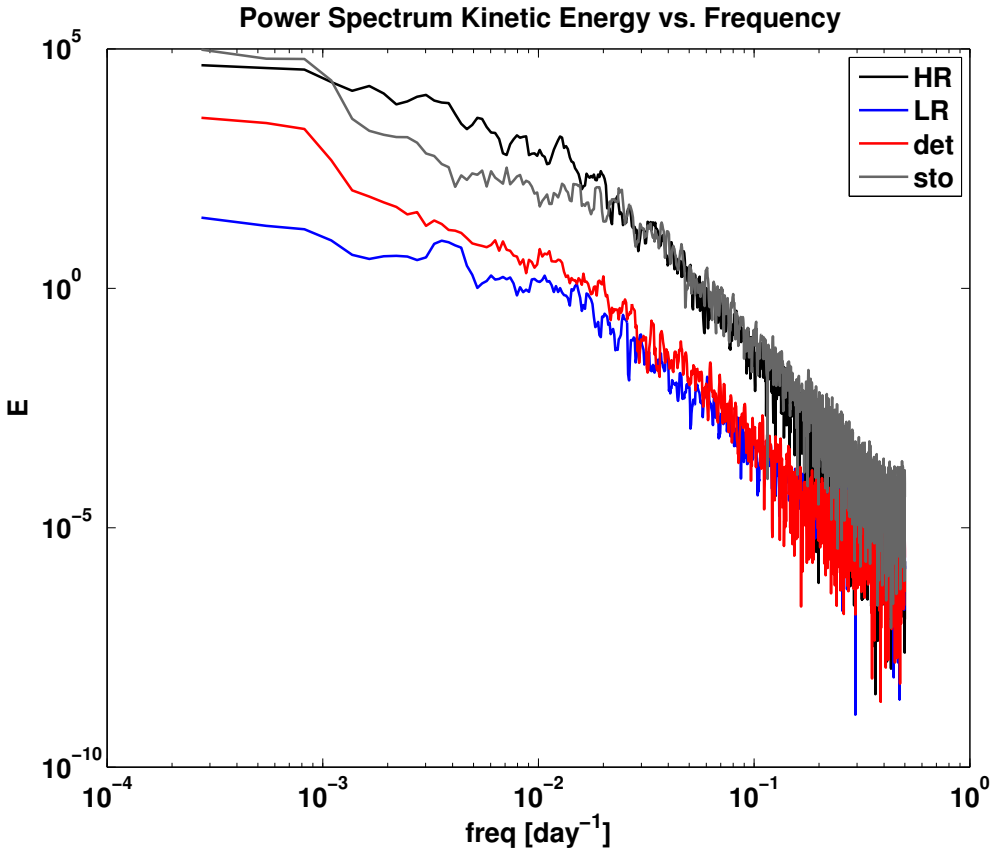

FIG. 8. Kinetic energy as a function of frequency for the following runs with horizontal resolution: 7.5 km (black), 30 km unparametrized (blue), 30 km with deterministic parametrization (red), 30 km with stochastic parametrization (grey).

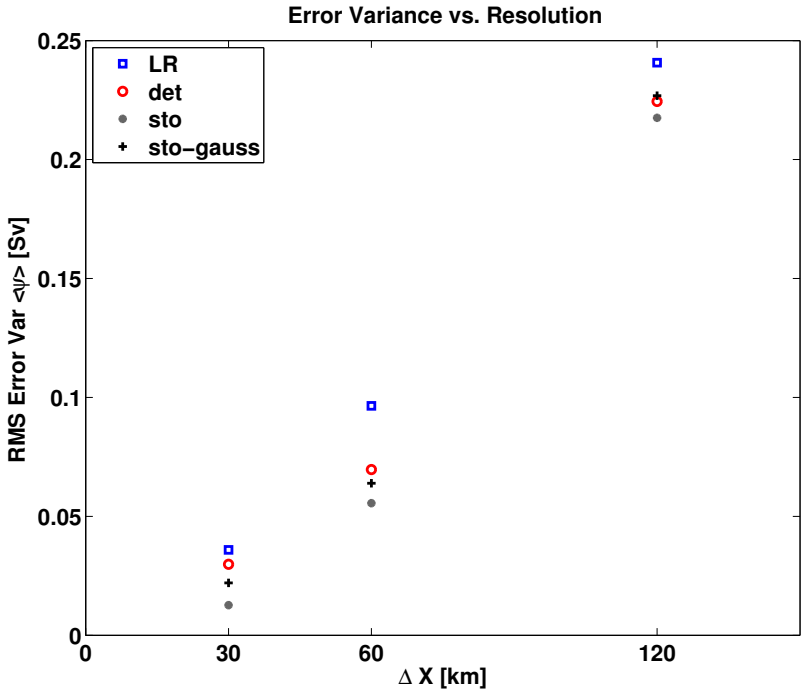

FIG. 9. Error in variance as a function of resolution for the different low-resolution runs: unparametrized (blue), deterministic parametrization (red), stochastic parametrization (grey), Gaussian stochastic parametrization (black, assuming skewness and kurtosis are negligible).