Ocean Modelling 79 (2014) 1–20

Contents lists available at ScienceDirect

Ocean Modelling

journal homepage: www.elsevier.com/locate/ocemod

Toward a stochastic parameterization of ocean mesoscale eddies

PierGianLuca Porta Mana, Laure Zanna *

Atmospheric, Oceanic and Planetary Physics, Department of Physics, University of Oxford, Oxford OX1 3PU, UK

ARTICLE INFO

Article history:
Received 22 October 2013
Received in revised form 2 April 2014
Accepted 7 April 2014
Available online 26 April 2014

Keywords:
Mesoscale eddies
Stochastic parametrization
Sub-grid scale fluctuations
Eddy backscatter

ABSTRACT

A stochastic parameterization of ocean mesoscale eddies is constructed in order to account for the fluctuations in subgrid transport and to represent upscale turbulent cascades. Eddy-resolving simulations to derive the parameterization are performed in a quasi-geostrophic (QG) model in a double-gyre configuration. The coarse-graining of the high-resolution model is giving rise to an eddy source term which represents the turbulent Reynolds stresses. The eddy source term, its mean and fluctuations are analyzed as function of the resolved scales and external parameters.

A functional form of the resolved scales, based on a representation of turbulence as a non-Newtonian viscoelastic medium and including the rate of strain, is used to describe the eddy source term mean, variance and decorrelation timescale. Probability density functions (PDFs) of the eddy source term conditional on the resolved scales are then calculated, capturing the fluctuations associated with mesoscale eddies and their impact on the mean flow. Scalings for the mean, standard deviation, skewness, and kurtosis of the conditional PDFs are provided as function of the grid size, forcing, and stratification of the coarse-resolution model.

In light of these scalings, no preliminary high-resolution (QG) model runs are necessary to diagnose the subgrid forcing and the implementation of a stochastic closure based on the conditional PDFs requires in principle very little tuning.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

The ocean contains a vigorous geostrophic mesoscale eddy field with spatial scales of approximately 100 km at mid-latitudes, evolving over time scales from weeks to months. In the western boundary currents (such as the Gulf Stream and the Kuroshio) and their extensions and in the Antarctic Circumpolar Current (ACC) in the Southern Ocean, eddies are the primary means by which heat is transported poleward (e.g., Danabasoglu et al., 1994; Marshall, 1997) and have an order one effect on the mean stratification (Henning and Vallis, 2004; Cessi and Fantini, 2004). While most studies concentrate on the time-mean effect of the eddies on the mean flow, it has been shown that the transient fluctuations are several orders of magnitude larger than the mean (Berloff et al., 2007a; Li and von Storch, 2013). Eddies can be key in maintaining or rectifying the meandering jet (Berloff, 2005b; Waterman and Jayne, 2012) therefore not only acting as a source of dissipation but also as a driving force through nonlinear eddy-eddy and eddy-mean flow interactions. Neglecting eddy variability can possibly lead to large errors in the climatology of the dominant scales. For example, the mean path and variability of western boundary currents, which serve as boundary conditions for the atmosphere, have a crucial impact on the dynamics of the main atmospheric patterns (e.g., Kirtman and Vecchi, 2011). Therefore, mesoscale eddies need to be either resolved, or understood and parameterized, before one can have confidence in interannual and decadal climate predictions. To properly simulate the mesoscale eddy field and associated turbulent cascades of tracers (active and passive) in numerical simulations, grid spacings of roughly 10 km and smaller are necessary. It is therefore unlikely that ocean climate models will systematically resolve the Rossby radius of deformation especially for long integrations and large ensembles, and the effect of eddies needs to be parametrized instead.

The traditional approach of parameterizing ocean mesoscale eddies is based on mimicking the effect of eddy fluxes on the mean flow using schemes based on several key assumptions including: (1) the eddy fluxes of tracers are directed down the mean gradients; (2) a single value of the eddy diffusivity is associated with a given value of the resolved scales and/or latitude and depth. For a potential vorticity closure, these assumptions lead to homogenization of conserved tracers within closed geostrophic contours (Rhines and Young, 1982a). Those assumptions are also the basis of the Gent and McWilliams (1990) parameterization (GM hereafter) leading to a flattening of isopycnals. The parameterization forces the mixing of water-mass properties along isopycnals.

* Corresponding author. Tel.: +44 1865 272925.
E-mail address: zanna@atm.ox.ac.uk (L. Zanna).

http://dx.doi.org/10.1016/j.ocemod.2014.04.002
1463-5003/© 2014 Elsevier Ltd. All rights reserved.

Eden and Greatbatch (2008a,b), stressing the importance of diapycnal mixing by mesoscale eddies, propose a parametrization in which the GM diffusivity coefficient is determined by the eddy kinetic energy (EKE), which is in turn evolved through a prognostic equation. This parametrization leads to more reasonable values of the diagnosed GM diffusivity in eddy-resolving models but presents some drawbacks in non-eddy-resolving models due in part to the lack of EKE. Turbulent diffusion of potential vorticity (PV) in the interior accompanied by buoyancy diffusion in surface and bottom layers has also been proposed (Treguier et al., 1997) as an alternative to GM. Yet, many studies using eddy-resolving models have shown that downgradient eddy closures in terms of thickness or potential vorticity only show limited skill (e.g., Roberts and Marshall, 2000) and a few up-gradient schemes have been discussed (e.g., Berloff, 2005b; Eden, 2010). For example, Wilson and Williams (2006) propose that localized regions of down- or up-gradient eddy transfer of tracers are associated with a characteristic life cycle of ocean eddies. At steady state, there are regions where the eddy fluxes are locally controlled by the advection of background variance by the total flow, making it difficult to parameterize the detailed direction of the eddy fluxes as solely being downgradient. Despite the tremendous improvements of the flow after the implementation of the traditional deterministic downgradient parameterizations, various processes and their variability remain poorly represented, e.g., up-gradient momentum fluxes leading to jet formation, sub-grid scale transport or high-frequency variability. Lastly, the parameterizations were derived for coarse-resolution models and their use in eddy-permitting models remains questionable. The parameterized eddy fluxes do not match the instantaneous diagnosed fluxes, neither the statistical nature of errors in parameterized fluxes nor their impact on ocean and climate simulations is well quantified.

An alternative approach to the deterministic parameterization is a stochastic or statistical closure (e.g., Kraichnan, 1959; Herring and Kraichnan, 1972; Frederiksen and Davies, 1997; Palmer, 2001; Williams et al., 2004; Holloway, 2004; Neelin et al., 2008; Dorrestijn et al., 2013; Grooms and Majda, 2013) which offers the advantage of an explicit representation of the inherent uncertainty associated with the parametrized processes and allows for the representation of sub-grid scale transport associated with eddies. For example, stochastic backscatter as a means to excite upscale energy cascade has been explored in the hierarchy of the numerical atmospheric models including operational center models (e.g., Frederiksen and Davies, 1997; Shutts, 2005; Bowler et al., 2009). The stochastic framework for subgrid-scale parameterizations has shown some rather promising results (e.g., Palmer, 2012). For example, Buizza et al. (1999) proposed adding a form of multiplicative noise to the equations of motion in the atmosphere, where the multiplicative factor was applied to the ensemble-mean parameterized tendency. This representation of 'missing physics' as a stochastic term led to several improvements over the past decades such as improving the skill of weather and seasonal forecasts. Frederiksen and colleagues (Frederiksen, 1999; O'Kane and Frederiksen, 2008; Zidikheri and Frederiksen, 2010; Frederiksen et al., 2012; Kitsios et al., 2012) combine eddy viscosity and a stochastic backscatter to represent the unresolved scales in idealized models of the atmosphere and ocean. The approach is tested in spectral space by forcing a truncated version of the model, following early ideas of Kraichnan (1959). The kinetic energy spectrum of the parametrized truncated model is shown to improve compared to the unparameterized case. Berloff (2005b)'s approach attempts to mimic the localized forcing as a temporally and spatially correlated stochastic process producing jet rectification. Brankart (2013) introduces a stochastic model to reproduce the density fluctuations in the equation of state for an ocean GCM improving the large-scale flow.

When building a stochastic parameterization, similarly to the deterministic case, one needs to estimate the parameters associated with the stochastic forcing such as its temporal and spatial correlations. For example, Penland and Sardeshmukh (1995) and Zanna (2012) consider non-linear sub-grid scales processes as white Gaussian noise in examining inverse models of sea-surface temperature fluctuations in the Pacific and Atlantic, respectively. Farrell and Ioannou (2003, 2007) consider Gaussian white noise to represent the unresolved scales arising from atmospheric jet instability. However, the use of Gaussian distributions as forcing of non-Gaussian systems has been shown to lead to quantitative discrepancies in the system's response (Franzke et al., 2007; Ring and Plumb, 2008; Cooper and Haynes, 2011). Gaussian distributions, augmented with non-Gaussian moments, are considered for a stochastic backscatter process and implemented in various studies (Frederiksen and Davies, 1997; O'Kane and Frederiksen, 2008).

Statistical properties of the surface mesoscale eddy field can be analyzed using Altimetry data (Hughes et al., 2010) or floats (LaCasce, 2008). While the observing system is improving, ocean observations remain sparse, and it is difficult to derive quantitative models for temporal and spatial mesoscale eddy forcing solely from observed data. An alternative approach, similar to Shutts and Palmer (2007), is to use high-resolution models that resolve mesoscale eddies as estimates of truth and conditional probability distribution functions as the base of our stochastic parameterization (e.g., Khouider et al., 2003; Crommelin and Vanden-Eijnden, 2008). All parameters and relationships are assessed by using the data from high-resolution models to inform the choices of the parameterization.

In the present study, we perform coarse-grained budgets of high-resolution potential vorticity using a coarse-graining length scale representative of current climate model grids. In Section 2 and 3, we present the coarse graining procedure, based on integrations of a quasi-geostrophic model in a double gyre configuration then analyze the structure and strength of the mean and fluctuations of the coarse-grained mesoscale eddy forcing. In Section 4, using estimates of the strength of these fluctuations as a function of the ensemble-mean parameterized tendency, we look for a functional relationship between the transient mesoscale eddy forcing and the resolved variables for a possible (deterministic) parameterization. The physical interpretation of the functional form is discussed in the context of the theory of the sub-grid scales. In Section 5, we analyze the relationship between the transient eddy forcing and the large flow using conditional probability distribution functions in addition to the dependence on resolution, stratification, forcing and sub-grid scale viscosity. We conclude in Section 6.

2. Model description and coarse graining

2.1. 3D quasi-geostrophic potential vorticity model

The model used in the present study, PEQUOD, solves the three-dimensional baroclinic quasi-geostrophic (QG) potential vorticity equation in the presence of forcing and dissipation on a beta plane in a square basin with size $L = 3840$ km (e.g., Berloff, 2005b,a). The potential vorticity (PV) q is given by

$$q = \nabla^2\psi + \beta y + \frac{\partial}{\partial z}\left(\frac{f_0^2}{N^2}\frac{\partial \psi}{\partial z}\right), \tag{1}$$

where $f = f_0 + \beta y$ is the planetary vorticity, $\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right)$ the horizontal gradient, $N = \left(-\frac{g}{\rho}\frac{d\rho}{dz}\right)^{1/2}$ the Brunt–Väisälä buoyancy frequency of the mean density stratification and ψ the

Table 1
Model parameters.

	Parameter	High-resolution model	Coarse-resolution model
Δx	Resolution	7.5 km	30 km
ν	Viscosity	$\nu = 50\ \mathrm{m^2\ s^{-1}}$	$\tilde{\nu} = 200\ \mathrm{m^2\ s^{-1}}$
L_{Ro}*	Rossby Radii of deformation	(40,23) km	(40,23) km
U	Velocity scale, $\sqrt{EKE}$	$0.21\ \mathrm{m\ s^{-1}}$	$0.12\ \mathrm{m\ s^{-1}}$
$\mathrm{Re} = \frac{UL_{\mathrm{Ro}}}{\nu}$	Reynolds number	174	24
f_0	Planetary vorticity at mid-y	$10^{-4}\ \mathrm{s^{-1}}$	$10^{-4}\ \mathrm{s^{-1}}$
β	$\mathrm{d}f/\mathrm{d}y$	$2\times10^{-11}\ \mathrm{m^{-1}\ s^{-1}}$	$2\times10^{-11}\ \mathrm{m^{-1}\ s^{-1}}$
g	gravity	$9.8\ \mathrm{m\ s^{-2}}$	$9.8\ \mathrm{m\ s^{-2}}$
g'	Reduced gravity	$(0.034, 0.018)\ \mathrm{m\ s^{-2}}$	$(0.034, 0.018)\ \mathrm{m\ s^{-2}}$
r	Bottom drag	$4\times10^{-8}\ \mathrm{s^{-1}}$	$4\times10^{-8}\ \mathrm{s^{-1}}$
τ_0	Wind stress	$0.3\ \mathrm{N/m^2}$	$0.3\ \mathrm{N/m^2}$
ρ_0	Reference density	$10^3\mathrm{kg/m^3}$	$10^3\ \mathrm{kg/m^3}$

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

streamfunction derived from the non-divergent velocity $\mathbf{u}$, such that $\mathbf{u} = \left(-\frac{\partial\psi}{\partial y}, \frac{\partial\psi}{\partial x}\right)$. While the diapycnal velocity w vanishes, the vertical extent of the fluid is included. The model is composed of three isopycnal layers with thicknesses H_m with $m = 1, 2, 3$ ($m = 1$ upper, $m = 2$ middle and $m = 3$ bottom).

The prognostic equation solved for each layer m is given by

$$\frac{Dq}{Dt} = \frac{\partial q}{\partial t} + \mathbf{u}\cdot\nabla q = \mathcal{D} + F. \tag{2}$$

The dissipation $\mathcal{D} = \nu\nabla^4\psi - r\nabla^2\psi\delta_{m,3}$ where the first term is a fourth-order term equivalent to Laplacian viscosity, with viscosity coefficient ν, acting to dissipate small scale enstrophy (q^2). The second order term, present only in the bottom layer ($m = 3$), parametrizes the presence of an Ekman layer with a bottom drag coefficient r.

The forcing F applied to the upper layer ($m = 1$) is the curl of the wind stress τ:

$$F(x,y) = \frac{(\nabla\times\tau)_z}{\rho_0 H_1}\delta_{m,1} = \begin{cases} -\tau_0\frac{0.92\pi}{L\rho_0 H_1}\sin\frac{\pi y}{g(x)}, & y \leqslant g(x), \\ \tau_0\frac{2\pi}{0.9L\rho_0 H_1}\sin\frac{\pi[2y-g(x)]}{L-g(x)}, & y > g(x), \end{cases} \tag{3}$$

where $g(x) = L/2 + 0.2(x - L/2)$ and ρ_0 is the reference density.

In the present study we use primarily three different model outputs:

- The first set of data is the output of the QG model at an eddy-resolving horizontal resolution of 7.5 km (see Table 1), wind strength $\tau_0 = 0.3\ \mathrm{N/m^2}$, viscosity coefficient $\nu = 50\ \mathrm{m^2/s}$, and layer thicknesses $(H_1, H_2, H_3) = (250\ \mathrm{m}, 750\ \mathrm{m}, 3000\ \mathrm{m})$. The density differences across the layer interfaces are $2.3688\ \mathrm{kg/m^3}$ and $1.2538\ \mathrm{kg/m^3}$ such that the associated Brunt–Väisälä frequency are $N = 6.82\times10^{-3}\ \mathrm{s^{-1}}$ and $2.56\times10^{-3}\ \mathrm{s^{-1}}$.
- The second set of data is the coarse-grained (spatially averaged) output of the high-resolution run, down to an eddy-permitting resolution of 30 km (see Section 2.2). This coarse-grained output represents the 'truth' - defined as the output that a coarse-resolution model with horizontal resolution of 30 km and an adequate eddy parameterization should attempt to reproduce.
- The third set of data is the output of the QG model at an eddy-permitting resolution of 30 km, with the same parameters as the high-resolution model except for an increased viscosity coefficient of $\tilde{\nu} = 200\ \mathrm{m^2/s}$. The output of the unparametrized coarse-resolution model is mainly used as a reference solution in order to analyze the discrepancies between this run and the high-resolution model output.

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 which conserves energy but not enstrophy (Arakawa, 1966). Additional experiments using different numerical schemes led to similar statistics to those presented in the following sections. All parameters for these runs are given in Table 1. Fig. 1 shows the statistically steady state streamfunction ψ in the upper layer for the eddy resolving run (panel a) and the coarse resolution model (panel b). The most striking difference between the runs is the absence of a strong eastward flowing jet in the coarse-resolution model mainly due to the increased viscosity and the lower Reynolds number.

2.2. Coarse-graining of the high-resolution model: equations and output

The goal of our study is to deduce a parameterization of ocean mesoscale eddies using the output of a high-resolution model at 7.5 km as the truth by coarse-graining its equations and output to mimic the output and equations of a coarser-resolution model at a resolution of 30 km (Berloff, 2005b,a; Duan and Nadiga, 2007). Let us consider the QG potential vorticity equation (Eq. (2)) at each grid-box i of the high-resolution model within a coarse-resolution grid box n given by[1]

$$\frac{\partial q_{n,i}}{\partial t} = -\nabla\cdot(\mathbf{u}_{n,i}q_{n,i}) + F_{n,i} + \nu\nabla^4\psi_{n,i}. \tag{4}$$

Eq. (4) is coarse-grained for each layer by convolving the high-resolution output with a constant window function of width 30 km as schematically illustrated in Fig. 2(a) (Murdoch and Bedeaux, 1994; Khouider et al., 2003; Shutts and Palmer, 2007), equivalent to a weighted average over all fine-grid cells within a coarse-grid cell of 30 km. Each coarse-grid cell contains a given number of fine grid boxes with integers n and i indexing the coarse- and fine-grid cell numbers, respectively. The value of any physical variable, e.g., ψ, in the ith fine-grid cell lying within the nth coarse-grid cell is denoted by $\psi_{n,i}$ and represents a grid-point value in the eddy-resolving model at $\Delta x = 7.5$ km horizontal resolution. The coarse-grained field $\overline{\psi}_n$, denoted with an overbar, is given by

$$\overline{\psi}_n = \sum_{i=1}^{I} W_{n,i}\psi_{n,i}, \tag{5}$$

[1] The bottom dissipation is omitted only for simplicity but is still part of the coarse-graining procedure.

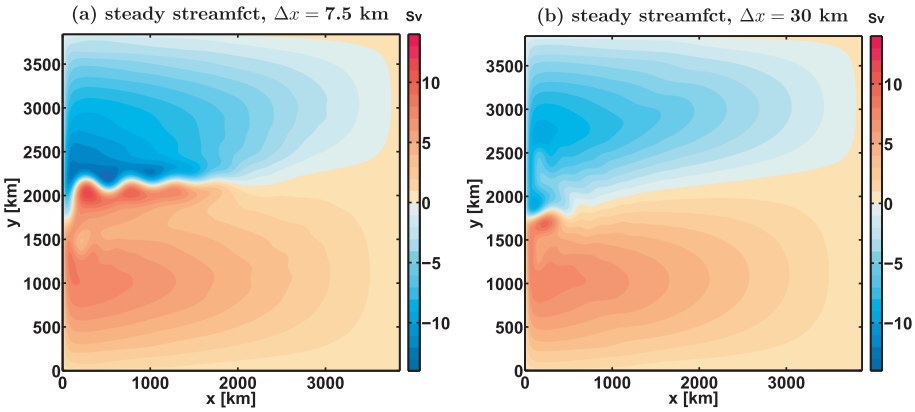

Fig. 1. Statistically steady state streamfunction ψ for the model runs at horizontal resolutions of (a) 7.5 km (eddy resolving model) and (b) 30 km (eddy permitting model).

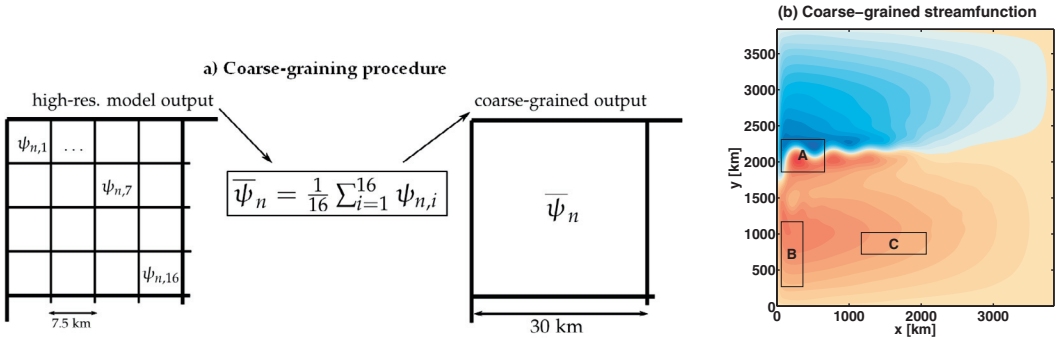

Fig. 2. (a) Schematic of the coarse-graining (spatial Reynolds averaging) procedure. (b) Statistically steady state streamfunction for the high-resolution model output coarse-grained to 30 km highlighting dynamically different regions for analysis purposes.

where $W_{n,i}$ are the equal and constant weights (except near the boundaries) assigned to the different fine-grid cells such that $W_{n,i} = 1/I$, where $I = 16$ is the number of fine-grid cells within a coarse-grid cell of 30 km. Note that the coarse-graining, hence the overbar, applies only to horizontal variables and operators. No coarse-graining is applied to the temporal and vertical variables (hence no overbar is used for vertical and temporal derivatives).

Using Eq. (5), the coarse-graining of Eq. (4) leads to $\frac{\partial \bar{q}_n}{\partial t} = -\overline{\nabla\cdot(\mathbf{u}_n q_n)} + \nu\overline{\nabla^4\psi}_n + \overline{F}_n$. Assuming that the viscosity in the high-resolution and low-resolution models will differ, we introduce a low-resolution dissipation coefficient $\tilde{\nu}$. By adding the coarse-resolution advective and dissipation terms, $\overline{\nabla}\cdot(\overline{\mathbf{u}}_n\bar{q}_n) + \tilde{\nu}\overline{\nabla}^4\overline{\psi}_n$, to both sides of the equation we obtain an equation for the coarse-grained output given by

$$\frac{\partial \bar{q}_n}{\partial t} + \overline{\nabla}\cdot(\overline{\mathbf{u}}_n\bar{q}_n) = \left[\overline{\nabla}\cdot(\overline{\mathbf{u}}_n\bar{q}_n) - \overline{\nabla\cdot(\mathbf{u}_n q_n)} + \nu\overline{\nabla^4\psi}_n - \tilde{\nu}\overline{\nabla}^4\overline{\psi}_n\right] + \tilde{\nu}\overline{\nabla}^4\overline{\psi}_n + \overline{F}_n, \tag{6}$$

where $\overline{\nabla}$, is the gradient operator acting on the low-resolution fields.

For convenience, let us drop the subscript n and rewrite

$$\frac{\partial \bar{q}}{\partial t} + \overline{\nabla}\cdot(\overline{\mathbf{u}}\bar{q}) = S^* + \tilde{\nu}\overline{\nabla}^4\overline{\psi} + \overline{F}, \tag{7}$$

where the eddy source term S^* is given by

$$S^* = \overline{\nabla}\cdot(\overline{\mathbf{u}}\bar{q}) - \overline{\nabla\cdot(\mathbf{u}q)} + \nu\overline{\nabla^4\psi} - \tilde{\nu}\overline{\nabla}^4\overline{\psi}. \tag{8}$$

The coarse-graining procedure gives rise to an additional 'source' (or forcing) term in the QG equation – the divergence of a Reynolds stress. The eddy source term S^* is composed of the coarse-grained high-resolution advection $\overline{\nabla\cdot(\mathbf{u}q)}$; the coarse-resolution advection $\overline{\nabla}\cdot(\overline{\mathbf{u}}\bar{q})$; the coarse-grained high-resolution viscosity $\nu\overline{\nabla^4\psi}$, and the coarse-resolution viscosity $\tilde{\nu}\overline{\nabla}^4\overline{\psi}$. The eddy source term reflects mainly the difference between the high-resolution advection and the coarse-resolution advection and therefore represents the transient mesoscale eddies and their interaction with the large-scale flow (Berloff, 2005b,a; Duan and Nadiga, 2007). Further details on the coarse-graining methodology are given in Appendix A, including the treatment of boundary conditions and the non-commutativity of the Laplacian operator which introduces a correction to the eddy source term.[2]

By coarse-graining the high-resolution model output, we can compare some of the general properties of the high-resolution, coarse-grained and coarse-resolution data. The steady state coarse-grained streamfunction field $\overline{\psi}$, is shown in Fig. 2(b). The steady state coarse-grained output presents the same large-scale features as the high-resolution output, such as the eastward-flowing jet but with features at scales smaller than 30 km being smoothed out. However, the high-resolution and coarse-grained output have markedly different kinetic energy power spectra as a function of total wavenumber (Fig. 3(a)). The power spectrum of the coarse-grained output increasingly diverges from the high-resolution output as the wavenumber increases. In the $(200\text{ km})^{-1}$ to $(60\text{ km})^{-1}$ range, the high-resolution kinetic energy power spectra obeys a power-law of $k^{-2.6}$ while the spectra of the

[2] The correction term arising from the discretization of the potential vorticity is $\partial_t(\overline{\nabla^2\psi} - \nabla^2\overline{\psi})$.

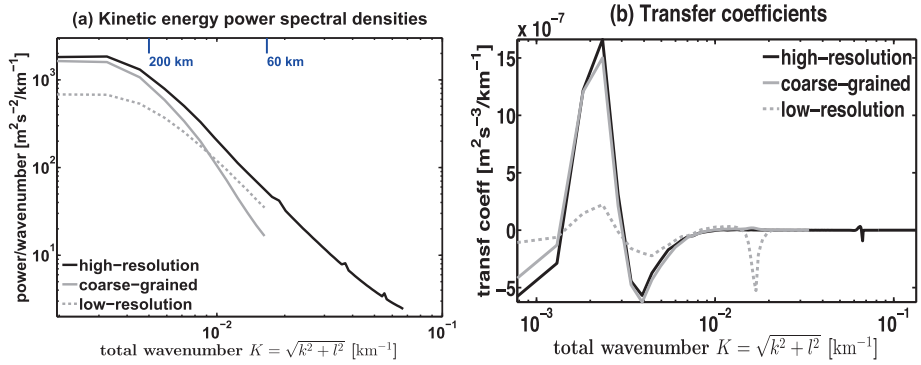

Fig. 3. Kinetic energy (a) power spectral density and (b) transfer coefficients for the high-resolution (black), coarse-grained (grey), and low-resolution (dashed grey) models as function of total wavenumber K. The total wavenumber is given by $K = \sqrt{k^2 + l^2}$, where k and l are the zonal and meridional wavenumbers, respectively.

coarse-grained output follows a much steeper power-law of $k^{-3.6}$. The coarse-graining procedure acts as a smoother at large wavenumbers. The power spectrum of the coarse-resolution (30 km) run has a much lower kinetic energy at all wavelengths compared to the high-resolution run but its power spectrum follows a noticeably shallower power-law of $k^{-2.3}$ in the range $(200\,\text{km})^{-1}$ to $(60\,\text{km})^{-1}$ compared to the coarse-grained output. A steeper spectrum for the low resolution model can be achieved by increasing the viscosity coefficient. The transfer coefficients (Qiu et al., 2008; Scott and Arbic, 2007) which give a measure of the transfer of kinetic energy between spatial scales are shown in Fig. 3(b). Despite the different kinetic energy spectrum of the coarse-grained model output compared to the high-resolution model spectrum, the transfer of energy between the different scales is preserved (black and grey curves), unlike for the coarse resolution model (dashed-grey). The main attempt will be to parametrize the eddy source term as function of the resolved scales such that the energy transfer can be represented in the coarse-resolution model (Berloff et al., 2007b).

3. Properties of the eddy source term

3.1. Spatio-temporal features of the eddy source

The coarse-grained eddy source term and its components are calculated from the high-resolution output which is saved every hour for the total length of the integration. Most of the analysis is be done by using 10 years of daily output of S^* which we have tested and found sufficient to capture the behavior of S^* keeping the size of the output manageable.

Fig. 4(a) shows the spatial pattern of the eddy source term S^* averaged over 10 years in the upper layer. The eddy source S^* has a complex spatial structure with alternating signs in the meridional and zonal directions, and strongest amplitude in the vicinity of the western-boundary currents and the meandering jet. A similar pattern is observed in the other layers but with a much reduced amplitude as the eddies and the jet are surface intensified. The along and across-jet succession of alternating signs pattern of the eddy source term evolves following the meanders of the jet (Fig. 4). A positive (negative) source term acts to create (destroy) potential vorticity. Therefore the action of S^* is to redistribute PV in the region of the jet, leading to alternating regions of increased and decreased PV gradient. For simplicity, the eddy source term (Eq. (8)) can be approximated by the divergence of the eddy flux $\mathcal{F} = -\overline{\mathbf{u}\bar{q}} + \bar{\mathbf{u}}\bar{q}$, such that $S^* \approx \overline{\nabla} \cdot \mathcal{F}$. An effective non-divergent eddy-induced velocity $\bar{\mathbf{u}}^*$ can be obtained such that $S^* \approx \overline{\nabla} \cdot (\bar{\mathbf{u}}^* \mathcal{H})$ where $\mathcal{H}$ depends on the components of $\overline{\nabla}\bar{q}$ and $\mathcal{F}$ (Vallis, 2006, pp 311–312, 428). The time-mean zonal component of the eddy-induced advective velocity $\bar{\mathbf{u}}^*$, shown in Fig. 4(b), accelerates the core of the jet and decelerates its flanks, therefore acting to sharpen the jet (Berloff et al., 2007a; Waterman and Jayne, 2011; Waterman and Hoskins, 2013). The velocity just north and south of the core of the jet can potentially explain the transport of eddies being shed off the jet and participating in the recirculation (merging with the jet itself). The meridional component (not shown) acts to intensify the western boundary currents.

Snapshots of S^* (panel a) and the terms defining it (panels b–d) for the upper layer are shown in Fig. 5. Figs. 5(b) and (c) show that the nonlinear terms related to the coarse-resolution advection

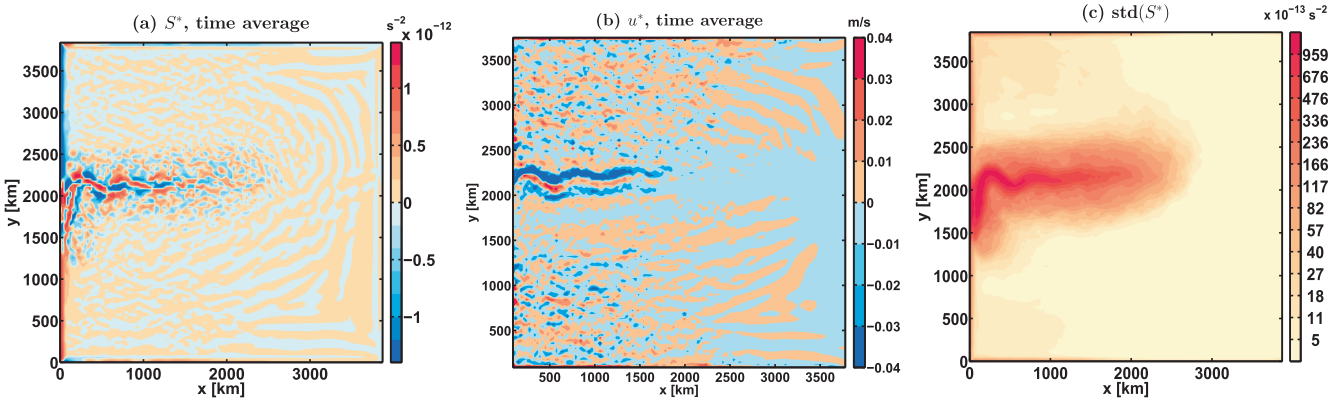

Fig. 4. (a) Ten-year averaged coarse-grained output in the upper layer of the eddy source term S^* defined in Eq. (8). An inverse-hyperbolic-sine color-scale (Johnson, 1949; Burbidge et al., 1988) is used owing to the great variability in magnitude of S^*, (b) Zonal component of the effective advective velocity $\bar{\mathbf{u}}^*$ provided by the source term S^*, and (c) Standard deviation of the eddy source term, $\sqrt{\langle S^{*\prime 2} \rangle}$.

$\overline{\nabla}\cdot(\overline{\mathbf{u}}\overline{q})$ and the coarse-grained high-resolution advection $\overline{\nabla\cdot(\mathbf{u}q)}$ are dominant. Each term is up to one order of magnitude larger than the source term itself in the vicinity of the jet and the source term is given by the interference between the two patterns of advection. This result is expected since the eddies are derived from these nonlinear terms and their difference represents the impact of the unresolved subgrid-scales onto the resolved flow. The small-scale dissipative terms $\nu\overline{\nabla^4\psi}$ and $\tilde{\nu}\overline{\nabla}^4\overline{\psi}$ tend to keep the same sign above and below the jet (not shown) and their sum is at least one order of magnitude smaller than the source term. Fig. 5(d) shows the contribution of the viscous terms and of the correction term, owing to the lack of commutativity between coarse-graining and spatial derivatives (this term has very little importance physically but is included for completeness).

3.2. Fluctuations and probability distribution functions of the eddy source term

A measure of the temporal vacillations of the eddy source term at each location can be described by its averaged temporal variance. Let us denote $\langle X\rangle$ as the temporal or ensemble average of a variable X and $X' = X - \langle X\rangle$ as the anomaly with respect to the average. The averaged temporal variance of the eddy source term is defined by $\langle S^{*\prime 2}\rangle = \langle(S^* - \langle S^*\rangle)^2\rangle$ and is shown in Fig. 4(c). Clearly, the western boundary currents and the jet exhibit high variance. The eddy source term exhibits spatial and temporal correlations with timescales of several days in the regions with large eddy activity, mostly the meandering jet and western boundary (Berloff, 2005b). The importance of the spatial and the temporal patterns and correlations of S^* have been analyzed and discussed by Berloff (2005b,a) and Berloff et al. (2007a) in QG models and Li and von Storch (2013) in an ocean GCM. Those studies point out that the interaction between the steady eddy source term $\langle S^*\rangle$ and its fluctuation $S^{*\prime}$ leads to a rectification of the eastward jet and a weakening of the gyres. Specifically, $\langle S^*\rangle$ is mainly responsible for the weakening of the gyres and of the western-boundary current(s), and for the meandering of the jet, while jet rectification is enhanced by $S^{*\prime}$. The effect of the eddy source term, $S^* = \langle S^*\rangle + S^{*\prime}$, is more than just a simple superposition of the effects of mean and the anomalies, owing to the non-linearity of the system. Therefore the mean eddy source term $\langle S^*\rangle$ and its temporal fluctuations $S^{*\prime}$ are essential features that ought to be reproduced in a parametrization of a coarse-resolution model in order to approximate the high-resolution flow.

The probability density functions (PDFs) for the eddy source term S^*, $\mathrm{P}(S^*)$, calculated after the flow has reached a statistically steady state using binning and histograms, are shown in Fig. 6. We consider three different regions in the basin – A: the jet, B: near maximum of the gyre strength, C: the interior, identified in Fig. 2(b). Dynamically different regions have vastly different shapes and widths of PDFs as expected from Fig. 4. All PDFs are centered around zero and shaped as delta functions in quiescent regions (regions A and B). In the core of the jet (region C), the PDF calculated clearly has a non-Gaussian shape, typical signature of intermittency, with exponential tails (Majda and Kramer, 1999; Bourlioux and Majda, 2002). The shape of the basin-wide PDF away from the origin is fairly similar to the PDF calculated in the jet region but its amplitude is reduced as we are sampling the regions of low S^* more often than when considering the jet region alone. Near the origin it exhibits a narrow peak observed in regions for which the eddy source term is weak.

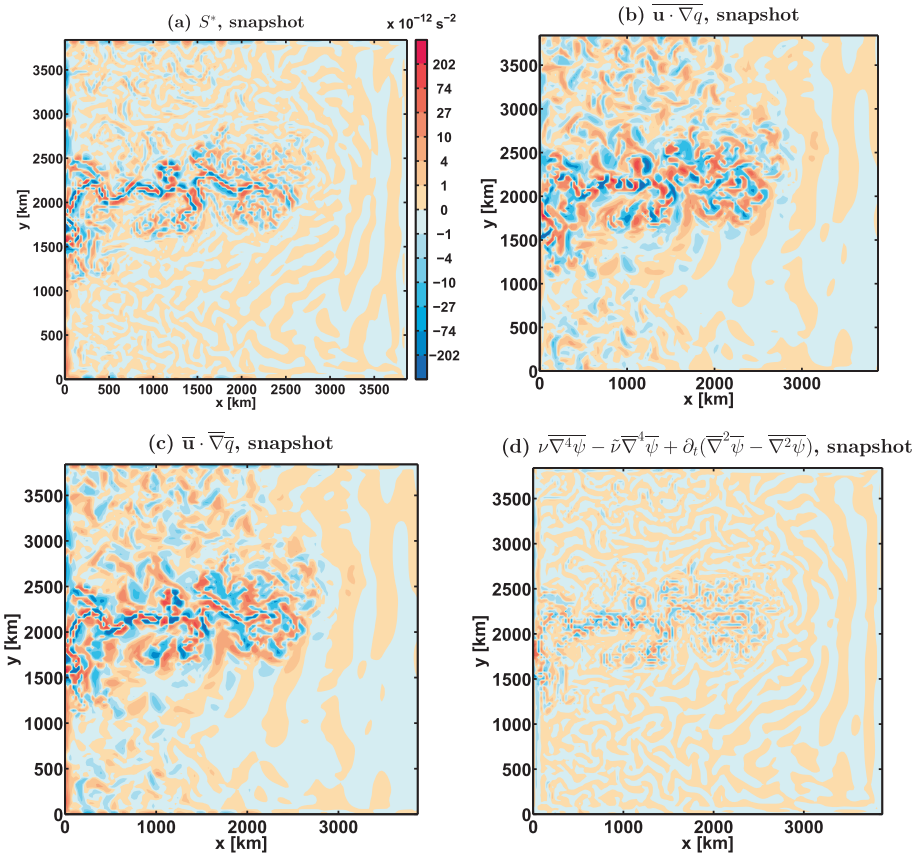

Fig. 5. Snapshots of (a) coarse-grained output of the eddy source term S^* and its individual components, (b) coarse-grained high-resolution advection, (c) coarse-resolution advection, and (d) viscous terms and discretization-correction term (the latter described in details in Appendix A).

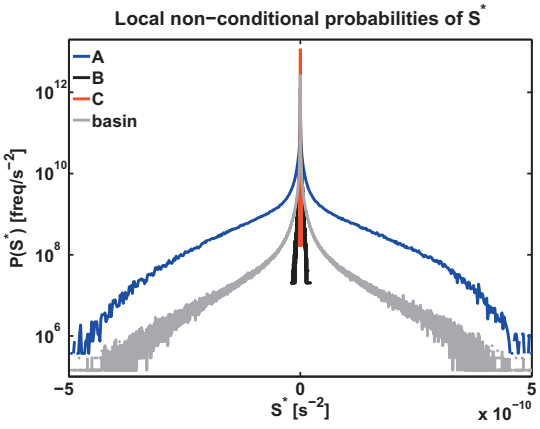

Fig. 6. Probability density for the eddy source term S^*, $P(S^*)$, in three dynamically different regions defined in Fig. 2(b), (A) the jet (blue), (B) near maximum of the gyre strength (black), and (C) the interior (red). All regions are such that the same number of points is used to calculate the PDFs. Also shown is the PDF for the entire basin (grey). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.3. Building the parameterization

It was postulated that a stochastic approach could better approximate the effect of the transient eddies on the mean flow, potentially capturing the upscale energy cascade and turbulence backscatter processes (Leith, 1990; Mason and Thomson, 1992) while representing the subgrid scale variability ignored in deterministic downgradient parameterizations (Frederiksen and Davies, 1997; Berloff, 2005b; Kitsios et al., 2013).

If we were to use the probability distribution $P(S^*)$ at all points and times as a stochastic parametrization, we would potentially lose the spatial and temporal correlations of the source term with the resolved scales, and therefore some essential features of dynamically different regions. It is possible to introduce the spatio-temporal correlations in the stochastic parameterization by explicitly making the probability distribution conditional on the location/coordinates $(\bar{x},\bar{y},z)$, i.e. $P(S^*|(\bar{x},\bar{y},z))$. This approach would be equivalent to that taken by Berloff (2005b) and Berloff et al. (2007a) in which the temporal and spatial correlations of the stochastic forcing are incorporated based on the results of the high-resolution runs. Making the PDFs with a spatial and/or temporal explicit dependence would involve the search of functions to approximate the spatio-temporal variability of source term (due to the different PDFs in different dynamical regions) and the fitting of a large number of parameters, leading to likely costly computations. Moreover, it would be heavily dependent on the specific model set-up, potentially requiring the run of a high-resolution model for each new set-up to diagnose $P(S^*|(\bar{x},\bar{y},z))$ and finally it would break the specific frame-invariance[3] of the QG equations under a different frame of reference.

A different approach to reproduce the spatial and temporal correlations in the statistical variations of the eddy source term S^* is to find a functional $\mathcal{R}(\bar{q})$ of the resolved-scale prognostic variable $\bar{q}$ that has a high correlation with the diagnosed eddy source term. Once a suitable functional $\mathcal{R}(\bar{q})$ is found, one can use a *conditional* probability density function (cPDF) expressing the statistical variations of S^* at a single point $(\bar{x},\bar{y},z)$ as a function of the value of the resolved functional $\mathcal{R}(\bar{q}(\bar{x},\bar{y},z,t))$ at that point, $P(S^*|\mathcal{R}(\bar{q}))$. In other words, the probability distribution of the variable S^* is updated after observing the realization of the conditional variable $\mathcal{R}(\bar{q})$, namely after the information that the variable $\mathcal{R}(\bar{q})$ has taken a particular value r. The updated probability distribution of S^* is then called the conditional probability distribution of S^* given $\mathcal{R}(\bar{q}) = r, P(S^*|\mathcal{R}(\bar{q}))$.[4] We refer to $\mathcal{R}(\bar{q})$ as the conditional variable. The statistical variations of the eddy source term, S^*, thus become functions of the coordinates only implicitly, via the coarse-scale functional $\mathcal{R}(\bar{q}(\bar{x},\bar{y},z,t))$, decreasing the number of parameters needed compared to a direct fitting for the reconstruction of $P(S^*|(\bar{x},\bar{y},z))$. Moreover, if the conditional variable $\mathcal{R}(\bar{q})$ is frame-invariant, the probability $P(S^*|\mathcal{R}(\bar{q}))$ is also frame-invariant by construction.

In the present approach, we wish to predict the mean and variance of the eddy source term and therefore we require a high correlation, but not necessarily a linear dependence, between S^* and $\mathcal{R}(\bar{q})$ (Section 4). The conditional PDFs allow us to statistically predict a value of the eddy source term S^* knowing the value of $\mathcal{R}(\bar{q})$. The statistical relationships between S^* and $\mathcal{R}(\bar{q})$ are then analyzed as function of coarse-resolution parameters and forcing (Section 5).

4. Relationship between the eddy source term and the resolved coarse-grained flow

To construct the conditional PDFs, $P(S^*|\mathcal{R}(\bar{q}))$, we impose mathematical and physical constraints to find an appropriate relationship between the eddy source term S^* and some functional $\mathcal{R}(\bar{q})$ of the resolved-scale prognostic variable $\bar{q}$.

4.1. A conditional variable linked to fluids of second grade and higher order strain rate

The conditional variable $\mathcal{R}(\bar{q})$ should be related to the turbulent Reynolds stress which appears as the eddy source term after the coarse-graining of the PV equation. We therefore look for conditional variables which are the divergence of a stress and have high correlations with S^*. We impose frame-invariance which is obtained by construction as long as the parametrization depends on PV solely through scalar functions of $\bar{q}$, the Lagrangian derivative D/Dt and the operator $\overline{\nabla}$.

A linear dissipative conditional variable of the form $\mathcal{R}(\bar{q}) = \overline{\nabla}^2\bar{q}$ leading to PV mixing (Rhines and Young, 1982a; Treguier et al., 1997) shows poor correlation with the eddy source term, which is not surprising as turbulence is not only dissipative. The order of the Laplacian can be increased leading to hyperviscous functionals, such as $\overline{\nabla}^4\bar{q}$ and $\overline{\nabla}^6\bar{q}$, which have been proposed as a scale-selective dissipation concentrated near the grid-scale (Macvean, 1983; Graham and Ringler, 2013; Kitsios et al., 2013). The correlation between the source term and the hyperviscous terms increases as the order of the Laplacian gets higher but still remains very weak with large spatial variations. Various other forms of the conditional variable such as (Gent and McWilliams, 1990, GM hereafter) or barotropic and baroclinic shears were used, all leading to poor correlations or widely spatially and temporally varying correlation coefficients. Some traditional closures and their correlations with the eddy source term are presented in Appendix B.

An alternative to the linear dissipative terms is to incorporate spatial and temporal information of the eddy-mean flow interaction using the Lagrangian derivative of the strain rate (the Eulerian time-derivative is ruled out as it breaks frame-invariance) and/or

[3] Definition of frame-invariance: The equations of motion or the quantitative representation of physical phenomena remain unchanged when the phenomena are observed under different conditions (e.g., when the observer is situated in an accelerating or a rotating frame of reference). The invariance refers to both space and time changes in the observer's frame of reference.

[4] The procedure to calculate the conditional probability distribution functions is explained in Section 5.1.

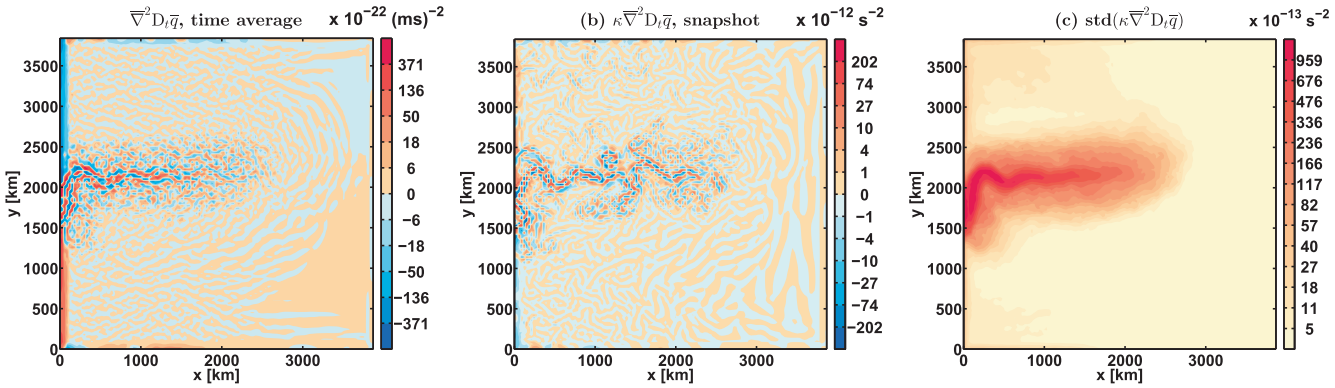

Fig. 7. (a) Ten-year average of the conditional variable, $\overline{\nabla}^2 \frac{\mathrm{D}\bar{q}}{\mathrm{D}t} = \overline{\nabla}^2 \mathrm{D}_t\bar{q}$, (b) snapshot of the conditional variable $\kappa\overline{\nabla}^2 \mathrm{D}_t\bar{q}$ with $\kappa = -(15\ \mathrm{km})^2$, and (c) standard deviation of $\kappa\overline{\nabla}^2 \mathrm{D}_t\bar{q}$.

non-linear terms.[5] The Lagrangian derivative introduces spatio-temporal information as it involves a dependence on the near history of the water parcel. A choice of $\boldsymbol{\mathcal{R}}(\bar{q})$ which takes the form of the (horizontal) divergence of a stress $\overline{\nabla}\cdot$(vector), includes the Lagrangian derivative, and does not break the frame invariance of the QG equation is given by

$$\boldsymbol{\mathcal{R}}(\bar{q}) = \overline{\nabla}\cdot\overline{\nabla}\frac{\mathrm{D}\bar{q}}{\mathrm{D}t}. \tag{9}$$

Fluid theories for which the stress includes a time derivative of strain-rate tensor are referred to fluids of second grade or 'Rivlin–Ericksen fluids' (Rivlin and Ericksen, 1955; Dunn and Fosdick, 1974; Truesdell and Rajagopal, 2009). The Rivlin–Ericksen stress is defined as the higher rate of change of deformation of the fluid relative to the linear Newtonian viscous term (fluids of first grade) and was derived to model the appearance of normal stresses in non-Newtonian fluid problems. The Rivlin–Ericksen stress in the momentum equation is given by $2\alpha_1(\frac{\mathrm{D}\boldsymbol{\mathcal{A}}}{\mathrm{D}t} + \overline{\nabla}\bar{\mathbf{u}}^{\mathrm{T}}\boldsymbol{\mathcal{A}} + \boldsymbol{\mathcal{A}}\overline{\nabla}\bar{\mathbf{u}}) + 4\alpha_2\boldsymbol{\mathcal{A}}^2$, where $\boldsymbol{\mathcal{A}} = \frac{1}{2}(\overline{\nabla}_3\bar{\mathbf{u}} + \overline{\nabla}_3\bar{\mathbf{u}}^{\mathrm{T}})$, $\overline{\nabla}_3$ is the 3D gradient operator, and α_1, α_2 are normal-stress moduli and are properties of the material. The exact derivation of the Rivlin Ericksen stress for the quasi-geostrophic approximation can be found by using a Rossby-number power expansion of the constitutive momentum equations for fluids of second grade; however it is beyond the scope of this work. We opt for a simple analogy to explain the relationship between $\boldsymbol{\mathcal{R}}(\bar{q})$ and the Rivlin–Ericsken stress in the QG potential vorticity equation. If the Rivlin–Ericksen stress was written in terms of potential vorticity instead of velocity and the non-linear term was negligible, the expression for the Rivlin–Ericksen stress would then take the form $\alpha\left(\frac{\mathrm{D}\overline{\nabla}\bar{q}}{\mathrm{D}t} + \overline{\nabla}\bar{\mathbf{u}}^{\mathrm{T}}\nabla\bar{q}\right) = \alpha\overline{\nabla}\frac{\mathrm{D}\bar{q}}{\mathrm{D}t}$. The latter expression $(\overline{\nabla}\frac{\mathrm{D}\bar{q}}{\mathrm{D}t})$ being the stress used for our conditional variable defined in Eq. (9).

The form of the conditional variable $\overline{\nabla}\cdot\overline{\nabla}\frac{\mathrm{D}\bar{q}}{\mathrm{D}t}$ is therefore related to second grade fluids and the Rivlin–Ericksen stress. The analogy between second grade (non-Newtonian) fluids and Newtonian turbulent flows is not new and was based initially on the appearance of secondary motions (recirculation) and normal-stresses acting on the flow in both instances (Rivlin, 1957). Since Rivlin, 1957, a large number of idealised studies have considered describing the dynamics of turbulent shear flows by proposing a non-Newtonian model for the apparent stress, induced by the fine-scale turbulence, on the large scale motion that required a viscosity dependent on the rate of strain (Liepmann, 1962; Crow, 1967; Lumley, 1970; Speziale, 1987).[6] Continuum models for second grade fluids are presently used in different applications such as glacier-ice dynamics and geophysical fluid dynamics (e.g., McTigue et al., 1985; Holm and Nadiga, 2003; Holm and Wingate, 2005; Riesen et al., 2010). For example, Foiaş et al. (2001) showed that the Euler α-model, used as a nonlinear dispersion model to represent the interaction between the unresolved and resolved scales, is mathematically identical to the inviscid second grade fluid equations. Further discussion to address the similarities between the QG version of Lagrange-averaged Navier–Stokes-α (LANS-α) model and our conditional variable are discussed in Appendix C. In the same section, we describe the conditional variable as a combination of the strain rate $\overline{\nabla}\bar{\mathbf{u}}^{\mathrm{T}}\nabla\bar{q}$ and the Lagrangian rate of change of the PV gradient $\frac{\mathrm{D}\overline{\nabla}\bar{q}}{\mathrm{D}t}$. The decomposition is described, similarly to the Okubo–Weiss description (Okubo, 1970; Weiss, 1991) of the stress related to the normal and shear components of strain and the relative vorticity, as acting to rectify the flow by strengthening and rotating the PV gradient.

4.2. Spatio-temporal features of the conditional variable

A ten-year average of the conditional variable $\overline{\nabla}^2\frac{\mathrm{D}\bar{q}}{\mathrm{D}t}$ is shown in Fig. 7(a). The similarities between the time-averaged $\boldsymbol{\mathcal{R}}(\bar{q}) = \overline{\nabla}^2\frac{\mathrm{D}\bar{q}}{\mathrm{D}t}$ and eddy source term (shown in Fig. 4(a)) are striking; however, the conditional variable is not as smooth as the eddy source term. Given that the time-averaging procedure can reduce the potential correlation between the eddy source and $\overline{\nabla}^2\frac{\mathrm{D}\bar{q}}{\mathrm{D}t}$, a snapshot of the spatial pattern of the conditional variable is shown in Fig. 7(b). The conditional variable is multiplied by a constant coefficient $\kappa < 0$. When compared with the snapshot of the eddy source term S^* in Fig. 5(a), we again see strong similarities between the two patterns although near the western boundary some discrepancies are apparent. The similarities between the S^* and $\boldsymbol{\mathcal{R}}(\bar{q}) = \overline{\nabla}^2\frac{\mathrm{D}\bar{q}}{\mathrm{D}t}$ are further confirmed by the very high and positive spatial Pearson correlation of the eddy source term with $-\overline{\nabla}^2\frac{\mathrm{D}\bar{q}}{\mathrm{D}t}$ shown in Fig. 8(a), except near the boundaries where we observe a strong anti-correlation. The differences between the regions near the boundaries and the ocean interior can be attributed either to the altered coarse-graining weights near the walls (Fig. A.16) or the complex dynamics involved near the boundaries which is not captured by the conditional variable, $\overline{\nabla}^2\frac{\mathrm{D}\bar{q}}{\mathrm{D}t}$, alone. The variance of the

[5] As noted by Speziale (1987), if the turbulent stresses are represented by a non-linear quadratic function of the strain rate with dimensions $[T^{-2}]$ where T is time, then dimensional consistency requires us to include a time derivative of the strain rate which has the same dimensions.

[6] Those studies described the behavior of turbulence by regarding it as a viscoelastic medium acting on a mean flow field. Moreover, note that the Rivlin Ericksen description of the normal-stress and visco-elasticity in polymers takes into account the effects produced by the preferential orientations of the polymer molecules under shear. Rivlin (1957) pointed out that eddies in a turbulent flow also exhibit preferential orientations under shear (Marshall et al., 2012), leading to further analogy between Newtonian turbulent and non-Newtonian visco-elastic fluids.

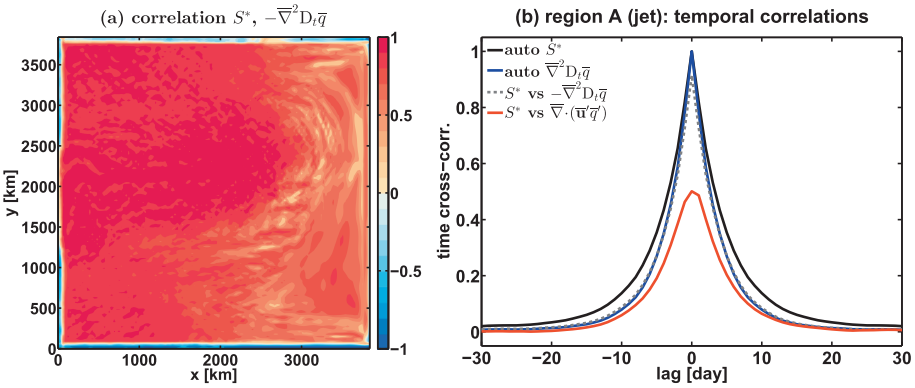

Fig. 8. (a) Time-averaged local Pearson correlation coefficients between S^* and $-\overline{\nabla}^2 \frac{\mathrm{D}\bar{q}}{\mathrm{D}t}$ (the correlation coefficient at each grid-box varies by less than 5% over 10 years) and (b) auto-correlation of the eddy source term (black) and the conditional variable (blue). Also on the same panel decorrelation time-scales of the eddy source term with the conditional variable (dashed grey) and with the divergence of the eddy fluxes, defined using Reynolds temporal averages (red). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

conditional variable (Fig. 7(c)), similarly to the eddy source term variance (Fig. 4(c)), is concentrated in the jet region. However again the fluctuations near the boundaries are not well captured.

Our search for a conditional variable is invoking a local and instantaneous relationship between the eddy source and the conditional variable (both calculated at the same location and time-step). Various functional forms considering remote and delayed variables, while still enforcing the basic mathematical constraints such as frame-invariance and stress-divergence, were tested but showed no better correlation than the one demonstrated in Fig. 8(a). Our conditional variable is still non-local to some extent as it uses information from the Lagrangian derivative of $\bar{q}$ (following the fluid parcel), and from neighboring points owing to the Laplacian operator. The claim of non-locality can be further verified when implementing the functional $\mathcal{R}(\bar{q}) = \overline{\nabla}^2 \frac{\mathrm{D}\bar{q}}{\mathrm{D}t}$ as a parameterization in a coarse resolution version of the model. Information from further remote locations or from the propagation of Rossby waves (for example) is probably not well represented by this conditional variable, nor are boundary layer processes.

By introducing the Lagrangian derivative for the expression of the conditional variable, we implicitly assume that the turbulent processes are Markovian with infinitesimal memory.[7] Kraichnan (1959) and Frederiksen (1999) have shown that subgrid scale turbulence in spectral space is non-Markovian with finite-time correlation equivalent to the decorrelation time (memory) between the subgrid tendency and the resolved field (Kitsios et al., 2012). The implicit Markovian approximation is used in our study for simplicity. However the decorrelation time of the eddy source term is well captured by the large-scale conditional variable chosen, $-\overline{\nabla}^2 \frac{\mathrm{D}\bar{q}}{\mathrm{D}t}$. Therefore we feel that the appropriate time-scales between the subgrid scales and the large scales are adequately represented and that the implicit Markovian approximation does not ignore the finite-time decorrelation necessary to represent the subgrid scale turbulence. For example, Fig. 8(b) shows the autocorrelation of the eddy source term in the jet region (black curve) in addition to the autocorrelation of the conditional variable (blue curve). The autocorrelation of the conditional variable decays slightly faster than that of the eddy source term but the overall behavior is well captured. Furthermore, the decorrelation time of the eddy source term with the conditional variable (dashed grey) is several days and higher than the correlation of the eddy source term with the temporal Reynolds averaged eddy fluxes (red) so that the temporal correlations are not ignored.

[7] A process is defined as a Markovian process if its future state at time $t + \delta t$ can be evaluated solely from its present state at time t and its most recent history at time $t - \delta t$.

4.3. Conditional variable as forcing of PV tendency

As mentioned in Section 4.2, the conditional variable, $\overline{\nabla}^2 \frac{\mathrm{D}\bar{q}}{\mathrm{D}t}$, is multiplied by a constant coefficient $\kappa < 0$ which is required to reproduce to appropriate sign of the eddy source term in the ocean interior. Since the spatial pattern and temporal variability of the source term is extremely well represented by $S_p = \kappa\overline{\nabla}^2 \frac{\mathrm{D}\bar{q}}{\mathrm{D}t}$, with a constant and negative κ over the entire domain, S_p could be viewed as a local deterministic parameterization and therefore replace the source term in Eq. (7). A coarse-resolution parameterized model would then take the form

$$\mathrm{D}_t\bar{q} = \frac{\partial\bar{q}}{\partial t} + \overline{\nabla}\cdot(\overline{\mathbf{u}}\bar{q}) = \kappa\overline{\nabla}^2\mathrm{D}_t\bar{q} + \tilde{\nu}\overline{\nabla}^4\overline{\psi} + \overline{F}, \tag{10}$$

leading to

$$\left(1 - \kappa\overline{\nabla}^2\right)\mathrm{D}_t\bar{q} = \tilde{\nu}\overline{\nabla}^4\overline{\psi} + \overline{F}. \tag{11}$$

Since κ is negative, the operator acting on the PV tendency and advective terms behaves as a "rougher" (instead of a smoother if κ was positive) increasing total forcing in the PV equation at small scales.

To further test and understand the meaning of the negative sign of κ, let us consider the parametrized PV equation in spectral space. The spatial Fourier transform of any function $f(\mathbf{x}, t)$ is denoted by $\widehat{f}(\mathbf{K}, t)$, where $\mathbf{x}$ is the position vector and $\mathbf{K}$ is the total wavenumber such that $K = |\,\mathbf{K}\,| = \sqrt{k^2 + l^2}$. The Fourier transform of Eq. (10) is therefore given by

$$\widehat{\mathrm{D}_t\bar{q}} = 16\pi^4\tilde{\nu}\mathbf{K}^4\widehat{\bar{\psi}} + \widehat{\bar{F}} - 4\pi^2K^2\kappa\widehat{\mathrm{D}_t\bar{q}}, \tag{12}$$

where we use $\widehat{\overline{\nabla}f} = 2\pi\mathrm{i}\mathbf{K}\widehat{f}$ and omit the argument $(\mathbf{K}, t)$ for convenience.

By considering $\kappa < 0$ and introducing the (non-dimensional) normalized wavenumber $\tilde{K} = K\sqrt{|\kappa|}$, we obtain

$$\begin{aligned}\widehat{\mathrm{D}_t\bar{q}} &= 16\pi^4\tilde{\nu}\mathbf{K}^4\widehat{\bar{\psi}} + \widehat{\bar{F}} + 4\pi^2K^2|\kappa|\widehat{\mathrm{D}_t\bar{q}},\\ &= 16\pi^4\tilde{\nu}\mathbf{K}^4\widehat{\bar{\psi}} + \widehat{\bar{F}} + 4\pi^2\tilde{K}^2\widehat{\mathrm{D}_t\bar{q}},\\ &= \frac{1}{1 - 4\pi^2\tilde{K}^2}\left[16\pi^4\tilde{\nu}\mathbf{K}^4\widehat{\bar{\psi}} + \widehat{\bar{F}}\right].\end{aligned} \tag{13}$$

If $\kappa = 0$, we recover the simple Fourier transform of the unparametrized coarse resolution model equation. However for the non-vanishing negative κ, the amplitude of the modified forcing in spectral space, $\left(1 - 4\pi^2\tilde{K}^2\right)^{-1}$, is increasing as the normalized

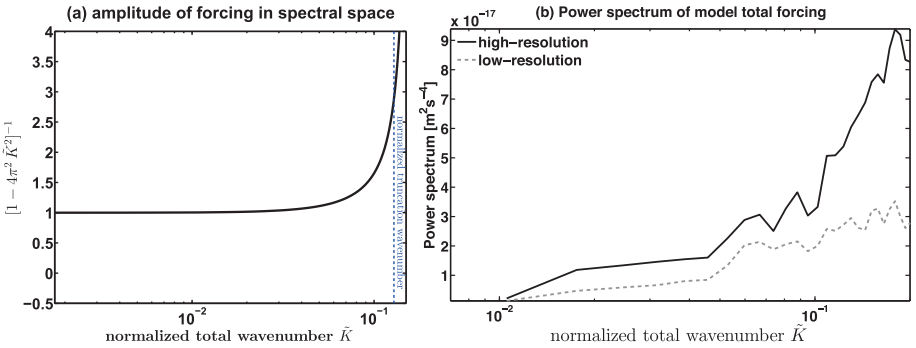

Fig. 9. Eddy forcing in spectral space: (a) amplitude of the total forcing, $(1-4\pi^2\widetilde{K}^2)^{-1}$, as function of the effective total wavenumber $\widetilde{K}$ and (b) power spectrum of right-hand side of Eq. (2) from numerical experiments at high and low resolutions as function of $\widetilde{K}$.

wavenumber $\widetilde{K}$ increases, with an asymptote at $K=\left(2\pi\sqrt{|\kappa|}\right)^{-1}$ which is of similar scale to our grid-box size (see Section 5.3). Therefore as postulated earlier, the total forcing (wind and dissipation) is increased by our eddy parametrization at small spatial scales (high wavenumbers) as shown in Fig. 9(a). Since we expect dissipation to be predominant at high wavenumbers (while wind forcing acts at large scales), the net effect of our parametrization is in general to increase dissipation at small scales, similarly to Kraichnan (1976) and other parameterizations derived in spectral space by Frederiksen and colleagues (e.g., Frederiksen and Davies, 1997; Frederiksen and Kepert, 2006; Zidikheri and Frederiksen, 2010).

The injection of eddy forcing at high wavenumbers in the PV equation can also be shown in a numerical experiment in which we evaluate the Fourier Transform of the right-hand side of the QGPV equation (Eq. (2)) for the high and coarse resolution models over one day. As initial condition, we use a snapshot of the high-resolution model after reaching its statistical steady state with a fully developed jet. After the initialization, the coarse resolution model gradually looses the jet while the high-resolution maintains it. The power spectrum (absolute value squared of the Fourier transform) of the right-hand side of the coarse and high resolution models are shown in Fig. 9(b). The "forcing" in spectral space of the high resolution model is higher at high wavenumbers than that of the coarse resolution model - the latter incapable of maintaining the jet, perhaps owing to the weak forcing at high wavenumbers.

5. cPDFs for the RE parameterization: forcing, resolution and stratification

5.1. Evaluation of conditional PDFs

To further investigate the relationship between the eddy source term and the conditional variable, we turn to the conditional PDFs described in Section 3.3 as the basis for our stochastic parametrization. Using $\mathcal{R}(\bar{q})=\overline{\nabla}^2\frac{\mathrm{D}q}{\mathrm{D}t}$, we construct the discretized versions of the conditional PDFs from the coarse-grained output as follows: (1) compute the extrema of $\mathcal{R}(\bar{q})$ over all points and times; (2) select 100 bins equally divided between the extrema of $\mathcal{R}(\bar{q}),(\Delta_1\mathcal{R}(\bar{q}),\ldots c,\Delta_{100}\mathcal{R}(\bar{q}))$; (3) Given a bin $\Delta_j\mathcal{R}(\bar{q})$, diagnose the distribution (quantiles, moments,...) of the values of the source term at all points and times at which $\mathcal{R}(\bar{q})$ assumes values within the bin $\Delta_j\mathcal{R}(\bar{q})$; (4) the collection of these 100 PDFs, one for each bin of the conditional variable $\mathcal{R}(\bar{q})$, constitutes the cPDFs $P(S^*|\mathcal{R}(\bar{q}))$. Therefore for each value of $\mathcal{R}(\bar{q})$ (or bin $\Delta_j\mathcal{R}(\bar{q})$) we obtain a statistical distribution for values of the eddy source term S^*. Given the large amount of information contained in the cPDFs, we collapse their representation onto a 2D plot to capture the essential information. Fig. 10(a) shows the cPDFs of the eddy source term S^* conditional on the variable $\overline{\nabla}^2\mathrm{D}_t\bar{q}$ for the upper layer of the model, represented as a 2D grey-scale contour plot for S^* (y-axis) and $\mathcal{R}(\bar{q})$ (x-axis), the contours being 5% quantile ranges of the cPDF, with their mean and standard deviation in red and blue, respectively. The median of the cPDFs is the black region and does not always correspond to the mean of the cPDFs, meaning that the PDFs can be skewed as shown in Fig. 10(b).

The cPDFs for the second and third layers are similar to Fig. 10(a) differing only in the magnitude of the eddy source term and the conditional variable as both eddy activity and the mean flow are diminished below the surface. There is a clear monotonic correlation between the source term and the conditional variable: as the magnitude of the conditional variable increases, so does the eddy source term magnitude, confirming the high correlation seen in Fig. 8(a). There are also large deviations from the mean, which increase with the magnitude of the conditional variable, as expected. For example, the spread in the standard deviation is about 20% of the mean value for non-vanishing values of the conditional variable. Therefore the distributions' mean alone cannot entirely represent the variability and fluctuations associated with the eddy-eddy and eddy-mean flow interactions.[8]

Similarly to the fact that PDFs themselves are regionally dependent so are the conditional PDFs – however the main reason to use the conditional PDFs is so that the eddy source term is captured by the large scale variable (which is implicitly a function of space and time). Therefore, near the origin of Fig. 10, we are mostly sampling the quiescent regions of the basin with low eddy activity (low values of the eddy source term and conditional variable) and less frequently regions near the jet. Away from the origin, for high values of the conditional variable, we are sampling exclusively regions with high eddy activity (vicinity of the jet). The spread in the PDFs is therefore not due to the inhomogeneities or the non-locality of the cPDFs. The main features not captured by the conditional variable and the cPDFs are (1) the grid cells near the viscous boundary layers (including the western boundary) which are anti-correlated with the conditional variable as shown in Fig. 8(a) and (2) the very small values of the eddy source term.

5.2. Scalings: sensitivity experiments

We wish to provide a scaling for explaining and predicting the moments of the cPDFs. The scaling for the cPDFs would then provide the basis for their reconstruction without the need to rerun

[8] Eddy-eddy interactions are defined as the interactions between subgrid eddies and transient resolved eddies but also as the subgrid–subgrid eddies interactions; eddy-mean flow interactions are defined as the interactions between subgrid-eddies and the resolved mean flow.

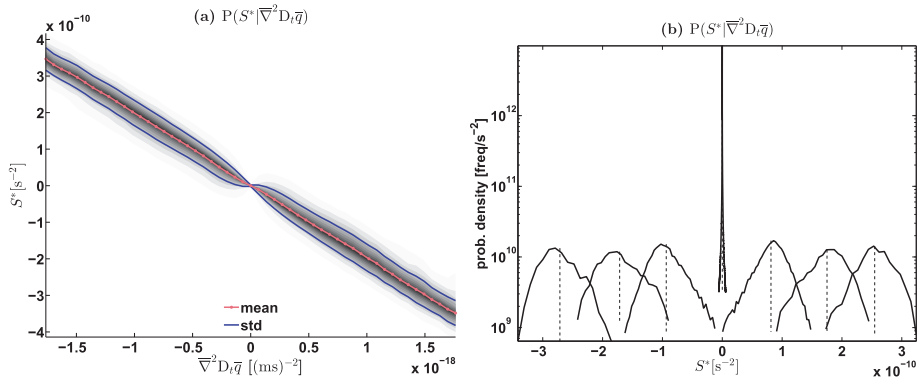

Fig. 10. (a) Conditional Probability Density Functions (cPDFs) for the eddy source term S^* conditional on $\overline{\nabla}^2 \frac{D\bar{q}}{Dt}$ for the upper layer of the 30 km coarse-grained data. The grey shading shows 5% quantiles increments of the conditional PDFs, the red and blue lines are the mean and standard deviation, respectively and (b) examples of cPDFs from panel (a) for selected values of the conditional variable $\overline{\nabla}^2 D_t\bar{q}$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the high resolution model. The mean is key in a deterministic parameterization and the fluctuations arising from the subgrid-scale processes can be described by using the variance (or standard deviation) and the high order moments and are essential to guide the stochastic part of the parameterization. For example, Kitsios et al. (2012) and Kitsios et al. (2013) show that the drain viscosity coefficient and backscatter eddy viscosity coefficient (governing the variance of the backscatter noise) proposed by Frederiksen and Davies (1997) and Frederiksen (1999) follow scaling laws which depend on resolution (truncation wavenumber), enstrophy flux, Rossby wavenumber and energy range.

Several sensitivity experiments are carried out to identify how the resolution, Reynolds number (wind forcing and dissipation), stratification, subgrid scale parameterization of viscosity and model configuration affect the characteristics of the cPDFs (in QG). A total number of 12 numerical experiments are performed in addition to the reference double gyre QG model run defined in Section 2.1. The same configuration (at a resolution of 7.5 km) is run using different Reynolds numbers, by changing both the wind strength τ_0 and the viscosity coefficient ν. The set of chosen values for eight additional runs with varying Reynolds numbers are $\left(\tau_0[\mathrm{N/m^2}], \nu[\mathrm{m^2/s}]\right) = [(0.3, 75); (0.3, 100); (0.6, 50); (0.6, 75); (0.6, 100); (0.8, 100); (0.8, 150); (0.9, 100)]$. The stratification (or equivalently the layer thicknesses) is modified compared to the original run with $\left(\tau_0[\mathrm{N/m^2}], \nu[\mathrm{m^2/s}]\right) = (0.3, 50)$, a set of 2 additional runs with the following layer thicknesses were performed (H_1, H_2, H_3) [m] $= [(200, 1200, 2600), (300, 800, 3100)]$. We perform one additional sensitivity experiment in which the viscous dissipation is parametrized as a bottom-friction-like viscosity given by $-r_0\nabla^2\psi$. Lastly, we change the configuration of the QG model from a closed basin to a reentering zonal channel setup of meridional $\times$ longitudinal extent of 3840 km $\times$ 7680 km forced by buoyancy flux (instead of wind). Note however that in QG, buoyancy and wind are ultimately equivalent as forcing to the PV equation (but project differently onto the barotropic mode).

The outputs of these high-resolution runs are thereafter coarse-grained, as explained in Section 2.2 and Appendix A, to resolutions of 30 km (eddy-permitting), and 60 km and 120 km (non-eddy-resolving). This variety of numerical experiments, while non-exhaustive, still allows us to tackle different issues potentially arising regarding the dependence of our scalings (in QG) on the coarse-graining size and resolution, model setup, Reynolds number, forcing and subgrid parameterization.

We show that the main advantage of using a parametrization based on the cPDF $P(S^*|\overline{\nabla}^2 D_t\bar{q})$ with the conditional variable $\overline{\nabla}^2 D_t\bar{q}$ is that we actually do not need to run a preliminary high-resolution run for diagnostics purposes, because the mean, spread, and particular shape of the cPDF $P(S^*|\overline{\nabla}^2 D_t\bar{q})$ can be simply deduced from coarse-resolution model, namely forcing, resolution and stratification.

5.3. Mean

The almost linear relationship between the mean μ of cPDFs of the eddy source term S^*, which represents mean spatio-temporal pattern of S^*, and the conditional variable $\mathcal{R}(\bar{q}) = \overline{\nabla}^2 \frac{D\bar{q}}{Dt}$ can be assumed to take the form $\mu = \kappa\overline{\nabla}^2 D_t\bar{q}$. The proportionality coefficient κ corresponds to the constant slope of the mean of the cPDFs of S^* with respect to the conditional variable $\overline{\nabla}^2 D_t\bar{q}$ (red curve in Fig. 10) such that $\kappa < 0$ (as expected). The linear dependence strongly suggests that a non-Newtonian second-grade model underlines the main properties of the fluid at coarse scales, as discussed in Section 4.1. The cPDFs for a run with higher Re number (wind forcing amplitude of $\tau_0 = 0.8\mathrm{N/m^2}$ and viscosity coefficient $\nu = 100\mathrm{m^2/s}$) are shown in Fig. 11(a), indicating that the linear relationship between the mean of the cPDFs and the conditional variable holds for a different strength of external forcing and sub-grid scale dissipation. The linear relationship holds for all coarse graining sizes ($\Delta x = 30, 60, 120$ km) and for all other model runs including runs in which the parameterization of viscous dissipation is altered or when the model is configured as a periodic channel with buoyancy forcing as shown in Fig. 11(b) and (c).

It is found that the value of κ is independent of the model run or its setup but scales with the coarse-graining size. Unlike a non-Newtonian fluid in which the parameters for the Rivlin–Ericsken stress are determined by the properties of the material, κ is simply determined by the scale of the grid-box. As discussed in Section 4.3, the parameterization acts at high wavenumber and close to grid-box size. From a dimensional relation, we assume that the influence of the eddies onto the mean flow is such that external forcing and dissipation can to first order be neglected (a valid assumption in the vicinity of the meandering jet) and assume that the effect of the eddies is well approximated by the mean of the cPDFs $\mu = \kappa\overline{\nabla}^2 \frac{D\bar{q}}{Dt}$ such that

$$\left[\frac{D\bar{q}}{Dt}\right] \approx [\kappa]\left[\overline{\nabla}^2 \frac{D\bar{q}}{Dt}\right], \tag{14}$$

leading to $\kappa \approx O(\Delta x^2)$. We find that the proportionality coefficient κ increases with resolution or more precisely with the square of the resolution length Δx. The results are reminiscent of the (Smagorinsky, 1963) parameterization in which the eddy viscosity dependent on the square of the typical length scale, often assumed to be the grid-spacing. For numerical experiments, a least square fit gives $\kappa = -(\alpha_{\mathrm{fit}}\Delta x)^2$ with $\alpha_{\mathrm{fit}} = 0.45$. Fig. 11(d) shows κ diagnosed

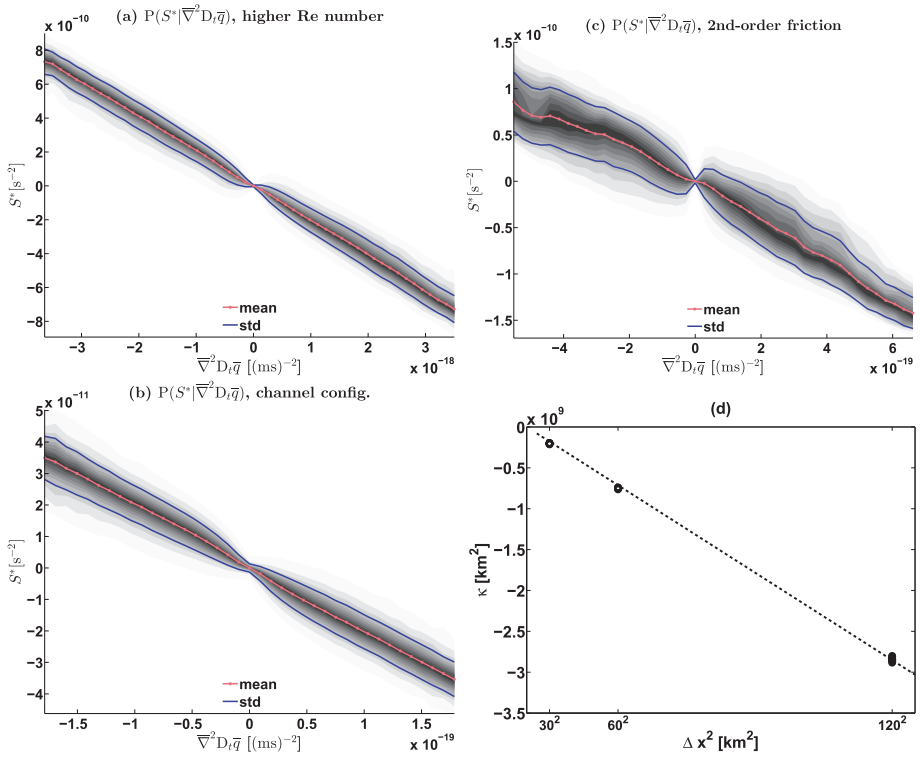

Fig. 11. cPDFs for the eddy source term S^* conditioned on $\overline{\nabla}^2 \frac{D\bar{q}}{Dt}$ for the upper layer, coarse-graining of 30 km from configurations with (a) strong wind forcing equal to 0.8 N/m^2 and viscosity 100 m/s^2, (b) viscosity parametrized as bottom-friction $-r_0\nabla^2\psi$, (c) periodic-channel geometry forced with buoyancy, and (d) linear relationship between the slope of the mean of the cPDFs κ and the square of the coarse-grained resolution Δx^2. The proportionality coefficient is around $\alpha_{fit} = -(0.45)^2$ (non-dimensional). Each circle represents the value of κ for at least 11 of the runs for each coarse-graining resolution Δx; note that many circles overlap.

from all our sets of experiments vs. the respective resolutions (squared) of the coarse-graining. The relationship $\kappa = -(\alpha_{\text{fit}}\Delta x)^2 \approx -\left(\frac{\Delta x}{2}\right)^2$ holds for all model runs. Additional coarse-graining studies, using for example, Gaussian weight functions with different widths Σ, show that the coefficient $\sqrt{|\kappa|}$ is actually proportional to the width of the weight function used in the coarse-graining procedure. This width sets the effective length scale over which the fields are coarse-grained, so $\sqrt{|\kappa|}$ reflects this effective scale rather than the model-grid size Δx itself hence the presence of α_{fit}.[9]

5.4. Standard deviation

The distributions' mean alone cannot represent the variability and extreme fluctuations of the eddy-eddy and eddy-mean flow interactions. The standard deviation σ of the cPDFs gives a measure of the amplitude of the fluctuations of S^* and therefore a first measure of the statistical/stochastic spread of the parameterization and by extension the stochasticity of the eddy-eddy and eddy-mean flow interactions.

Fig. 10(a) shows that the standard deviation σ (in blue) starts from an almost vanishing value corresponding to a Dirac-delta distribution before increasing as $(\overline{\nabla}^2 D_t\bar{q})^n$ with $1/2 < n \lesssim 1$ and asymptotically becoming almost constant. The initial increase is a natural consequence of the increased turbulent fluctuations owing to subgrid eddy activity accompanying the increase in the magnitude of the conditional variable (reflecting the straining/shearing of the flow). However, the asymptotic behavior for high values of the eddy source term suggests that fluctuations saturate to some extent. The asymptotic saturated behavior is to be contrasted with the linear (or slightly quadratic) dependence of the standard deviation on the mean, with no saturation, displayed in Shutts and Palmer (2007) for the convective heating rate coarse-grained in a numerical weather prediction model. The stochastic fluctuations quantified by the standard deviation can be viewed as multiplicative noise for small values of the conditional variable but as additive noise as the magnitude of the eddy source and the conditional variable increase.

An increase in wind forcing should lead to an increase in turbulent activity and therefore in the magnitude of the eddy source term and of its fluctuations. We therefore expect the (asymptotic) standard deviation σ, which is a measure of these fluctuations, to be proportional to the external wind forcing or equivalently to the fluctuations in potential vorticity. All runs with different values of wind forcing show that the values of the eddy source term and of the conditional variables increase as the wind forcing increases and so does the variance (Figs. 10(a) and 11(a)). Therefore our conditional variable is capable of capturing the increase of the eddy source term and associated mean flow due to the increase of the wind. It is likely that eddy saturation (Straub, 1993; Meredith and Hogg, 2006; Munday et al., 2013) will be implicitly taken into account by the conditional variable and would be revealed by the distribution and variability of the conditional variable with respect to wind forcing.

Since the fluctuations are a reflection of subgrid-scale variability, they should be statistically damped as the number of fine-resolution grids corresponding to a coarse grid increases, so an inverse proportionality with resolution is also expected. For example, Hu and Pierrehumbert (2001) find an inverse proportionality to grid spacing in the fluctuations of passive tracer distributions, so do Shutts and Palmer (2007) for the PDFs of the convective heating rate. Lastly, the eddy source term and its fluctuations are related

[9] Coarse-graining with a Gaussian weight function of standard deviation Σ is equivalent to coarse-graining in spectral space with a Gaussian weight function of standard deviation $1/\Sigma$.

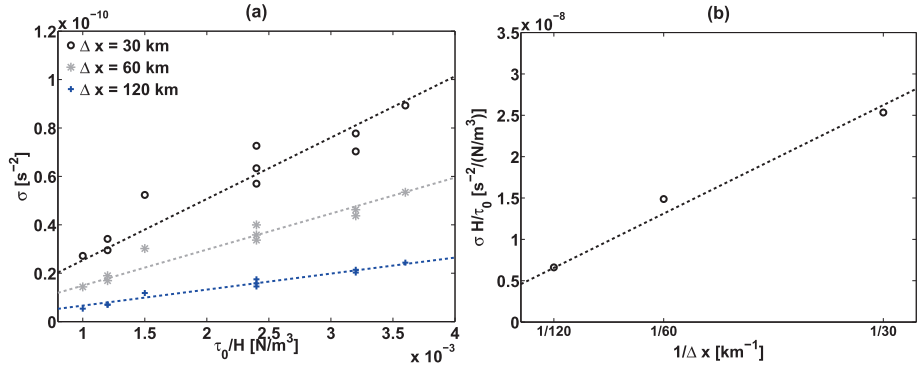

Fig. 12. (a) Linear relationship between the standard deviation σ and wind forcing τ_0 divided by layer thickness H for different resolutions. The proportionality coefficients are $(2.5, 1.5, 0.66) \times 10^{-8}\mathrm{m}^3/(\mathrm{Ns}^2)$ at resolutions $(30, 60, 120)$ km. (b) Inverse relationship between the standard deviation σ rescaled by the corresponding wind forcing per layer thickness τ_0/H, and the resolution Δx. The inverse proportionality coefficient, normalized by the density of water, is around 7.9×10^{-4} (non-dimensional).

to barotropic and baroclinic instabilities therefore we expect the shear, stratification and/or the Rossby radius of deformation to play a role in determining σ. In the present model, this information is encapsulated in the stratification parameters, e.g., layer thickness or density ratios between the layers.

We find indeed a simple relationship between the standard deviation of the cPDFs σ and the wind forcing τ_0, layer thickness H, and horizontal resolution Δx such that $\sigma = \gamma_{\mathrm{fit}} \frac{\tau_0}{\Delta x H \rho_0}$ where $\gamma_{\mathrm{fit}} = 7.9 \times 10^{-4}$ (non-dimensional) and ρ_0 is the reference density. Fig. 12(a) shows σ as function of the ratio wind/thickness (τ_0/H) for our set of coarse-grained outputs. For each coarse-resolution Δx, the value of σ clearly tends to follow a linear trend, and the slope for each coarse-grained output appears to decrease as the size of the coarse-resolution grid-box increases. Fig. 12(b) shows that such slopes have an inverse dependence on the resolution Δx. In Fig. 13, the standard deviation of the eddy source also scales linearly with the variance of the PV, still showing an inverse dependence on the resolution. In this case, the scaling for the standard deviation is given by $\sigma = \gamma_{\mathrm{fit,q}} \frac{\sigma^2(q)H}{\Delta x}$ where $\gamma_{\mathrm{fit,q}} = 1.7$ such that $\gamma_{\mathrm{fit,q}}$ is order one.[10] For the cases of the periodic channel configuration and the second-order viscosity parameterization, our scaling remains valid – meaning that the standard deviation still increases with the forcing and is inversely proportional to resolution. One main characteristic found is that the standard deviation increases with the Reynolds number and decreases with the coarse-graining size.

Note that the inverse proportionality of the standard deviation to resolution found by Hu and Pierrehumbert (2001) scales as the square root of the resolution length $\left(1/\sqrt{\Delta x}\right)$, whereas we find a directly inverse dependence $(1/\Delta x)$. Two factors could explain this difference. First, the distributions studied by Hu and Pierrehumbert (2001) are not conditional PDFs. Part of the inverse dependence on resolution is expected to be found in the range of variability of the conditional variable itself, leading to a square-root scaling for the non-conditional PDF. Second, Hu and Pierrehumbert (2001) study passive tracers, with particular dependencies on dissipation and dispersion subsumed in the Ching and Kraichnan (1998) formula used to derive their PDFs, and such dependencies are likely not valid for our eddy source term S^*. According to Duan and Nadiga (2007), the inverse proportionality of standard deviation on the resolution Δx implies that the flow of a coarse-resolution model parametrized stochastically using S^* should converge toward a better approximation (in a specific L^2-norm sense) of the high-resolution flow from which S^* is derived as the coarse-resolution size Δx tends to zero.

[10] The standard deviation of enstrophy is also found to be a good predictor of the standard deviation of the eddy source term.

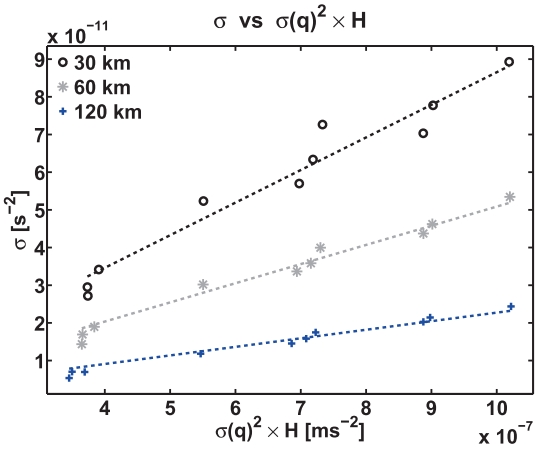

Fig. 13. (a) Linear relationship between the standard deviation σ of the eddy source term and the square of the standard deviation of the PV multiplied by layer thickness H for different resolutions. The fluctuations of the eddy source term are proportional to the fluctuations of PV, as expected, and also inversely proportional to resolution.

5.5. Skewness and kurtosis

Although the standard deviation provides a simple measure of the statistical spread of the conditional dependence of S^* on the forcing, stratification and resolution, it does not provide any information on the shape of this spread. The statistical deviations from the mean of the cPDF are non-Gaussian, as evident from Fig. 10(b), showing specific cPDFs chosen for different values of the conditional variable $\mathcal{R}(\bar{q}) = \overline{\nabla}^2 \frac{\mathrm{D}\bar{q}}{\mathrm{D}t}$ in the upper layer. Besides fat tails (indicative of non-zero kurtosis), we also notice a pronounced asymmetry (non-zero skewness). Such non-Gaussianity in the cPDFs is a general phenomenon often arising from non-linearities, leading to correlations between fluctuations and their non-normal superposition (e.g., Lorenz, 1995; Hu and Pierrehumbert, 2001; Shutts and Palmer, 2007; Franzke et al., 2007). The fat tails can also arise from intermittent transient instabilities (Majda and Kramer, 1999; Bourlioux and Majda, 2002; Majda and Gershgorin, 2013) as the Reynolds (or Peclet) number increases and the flow can transition to turbulent and chaotic regimes with large fluctuations at much greater frequencies than a Gaussian distribution would allow. The skewness of the cPDFs reflects the fact that high fluctuations can be more frequent toward one side than the other with respect to the mean.

Skewness and kurtosis are measured by the third and fourth standardized (non-dimensional) moments μ_3 and μ_4, respectively defined as

$$\mu_n = \sqrt[n]{\frac{\langle (S^* - \mu)^n \rangle}{\sigma^n}}.$$

Both skewness and kurtosis remain remarkably constant with respect to changes in the magnitude of the conditional variable $\nabla^2 D_t \bar{q}$ (not shown). The constant skewness and kurtosis is mostly due to the saturation of the standard deviation as the eddy forcing and conditional variable increase in magnitude. Note that we are always referring to the standardized moments, meaning skewness and kurtosis are scaled by the standard deviation which itself does depend on the wind, stratification, resolution and the value of the conditional variable. Skewness μ_3 and kurtosis μ_4 for all model runs and different coarse-graining size are shown in Fig. 14. Both standardized moments can be shown to be roughly constant and order $O(1)$ independently of model parameters with least square fitting procedure giving values of $\mu_3 \approx 0.61$ and $\mu_4 \approx 1.4$.

The third moment (skewness) shows some spread for a given model resolution but it was not possible to attribute the spread to the forcing, dissipation, stratification or any other coarse-resolution variables or forcing. In addition to the spread, a small dependence of the skewness on the coarse-resolution grid size is found such that the skewness decreases as the resolution increases. The best fit could then be approximated as $\mu_3 \approx \frac{4.5}{\Delta x}$ but the coefficient would take the dimensions of meters. The linear fit being order one captures most of the asymmetry with a fitting coefficient which is dimensionless.

The fourth moment (kurtosis) is independent of any model parameters and forcing. The best fit value of $\mu_4 \approx 1.4$ remains identical for all coarse-graining size values as well such that the extreme events associated with the eddy-mean flow interaction using this conditional variable are well captured by the mean and the standard deviation of the cPDFs leading a constant standardized fourth moment.

The first four moments of $\mathrm{P}(S^*|\overline{\nabla}^2 \mathrm{D}_t \bar{q})$ are determined by three main coarse-resolution parameters: resolution Δx, wind strength τ_0, layer thicknesses (H_1, H_2, H_3), besides the conditional variable itself $\overline{\nabla}^2 \mathrm{D}_t \bar{q}$. Using those parameters, we can therefore reconstruct the four first moments of the cPDFs for any QG model under forcing and dissipation without running a high-resolution model.

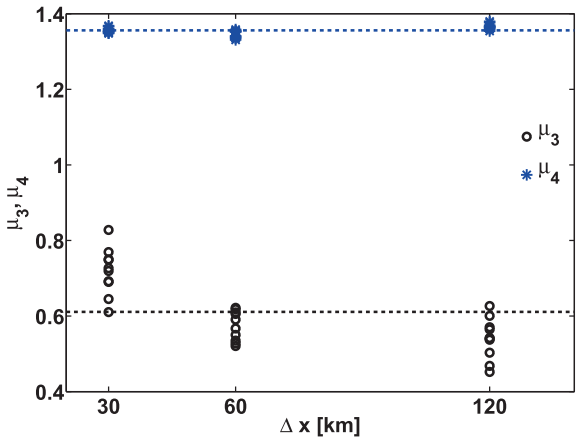

Fig. 14. Third and fourth standardized central moments μ_3, μ_4 plotted against resolution Δx: they have almost constant values $\mu_3 \approx 0.61$ and $\mu_4 \approx 1.4$ with a possible dependence of μ_3 on Δx.

5.6. Reconstruction of the conditional PDFs from the coarse-resolution parameters and basis for a stochastic parametrization

We show that the knowledge of those four moments is enough to reconstruct the shapes of the cPDFs with high accuracy so that knowledge of higher order moments of the cPDFs is unnecessary. The ill-posed inverse problem of reconstructing a probability distribution from some of its moments can be done by several techniques (from quadratures to fitting of pre-established distributions).

In the present work, our goal is to reconstruct the conditional PDFs in such a way as not to decrease the original statistical spread of the distributions, crucial ingredient of a stochastic parametrization. We use the Shannon entropy $H(\mathbf{p}) = -\sum_i p_i \ln p_i$ (Shannon, 1948; Kullback and Leibler, 1951; Bernardo, 1979) as a measure of the spread of a discrete probability distribution $\mathbf{p} = (p_i)$ and the maximum-entropy method (Jaynes, 1982; Sivia, 1990) as our reconstruction technique and described in Appendix D. Using Lagrange multipliers, we estimate the discrete probability distribution functions that satisfy the maximum entropy solution constrained by the mean, standard deviation, skewness and kurtosis of the cPDFs.

The conditional PDFs reconstructed from the four moments via the maximum-entropy method are shown in red in Fig. 15, compared with the original cPDFs diagnosed from the high-resolution model (in black) and with a Gaussian approximation constructed using only mean and standard deviation (in grey). The reconstruction of the cPDFs, which relies simply on knowing the coarse-resolution grid size, forcing and stratification, appears to be an excellent approximation to the originally diagnosed cPDFs and captures more than 95% of the probability range. The Gaussian approximation is visibly poor due to the importance of the skewness and kurtosis of the original cPDFs.

The procedure to reconstruct the cPDFs $\mathrm{P}(S^*|\overline{\nabla}^2 \mathrm{Di}_t \bar{q})$, which represents the subgrid-scale forcing, to be used as forcing/parameterization in a coarse-resolution model can then be summarized as follows:

1. choose the coarse-resolution model parameters: resolution Δx, wind-forcing strength τ_0, and stratification H_i;
2. use Δx to calculate κ, and use Δx, τ_0 and H_i to calculate the standard deviation σ, and the standardized skewness μ_3 and kurtosis μ_4 with the scaling found in the previous section;
3. use the values of the moments in the entropy-maximization procedure (see Appendix D) to find the discretized probabilities $p_i = \mathrm{P}(S_i^*|\overline{\nabla}^2 \mathrm{Di}_t \bar{q})$.

The cPDFs $\mathrm{P}(S^*|\mathcal{R}(\bar{q}))$ could be used as a stochastic parameterization in a coarse-resolution model as follows: (a) at a given time t and grid point $(\bar{x}, \bar{y}, z)$, we calculate the value of the conditional variable $\mathcal{R}(\bar{q}(\bar{x}, \bar{y}, z, t))$; (b) using the PDF $\mathrm{P}(S^*|\mathcal{R}(\bar{q}(\bar{x}, \bar{y}, z, t)))$,

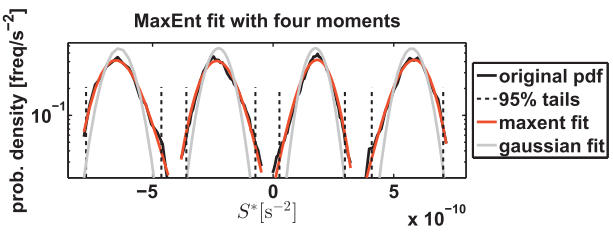

Fig. 15. Sample of diagnosed PDFs and their reconstruction using maximum-entropy (red). A Gaussian approximation, using only the first two moments, is also shown (grey). The dashed black lines show the 2.5% tails. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

conditional on the specific value $\mathcal{R}(\bar{q}(\bar{x},\bar{y},z,t))$, we randomly choose a value S^* of the eddy source term, according to the PDF; (c) we then time-step the PV in Eq. (2) adding to the right-hand side the stochastically chosen value S^* of the source term, and so find the value of PV at the next time step, $\bar{q}(\bar{x},\bar{y},z,t+\Delta t)$; (d) from the updated value of PV, we also update any other diagnostic variables, such as the streamfunction $\bar{\psi}$, at time $t+\Delta t$; (e) we set the new time to $t=t+\Delta t$ and go back to step (a). This procedure could be implemented at each time step and each grid point of the model, or only at particular times using finite decorrelation-time to include information from the higher order moments. The numerical results and stochastic parameterizations and its deterministic counterpart (in which only the mean of the cPDFs κ is chosen to be constant) are left to a follow-up study.

The key points in this procedure are that (1) no run of a high-resolution model is required to diagnose the eddy source term S^* or its distribution (at least in QG), (2) only very simple parameters of the coarse-resolution model are required to construct the cPDFs, which means that in principle no preliminary, exploratory runs of the unparametrized coarse-resolution model are required either.

6. Conclusions

In the panorama of stochastic parameterizations for sub-grid scale processes we can broadly distinguish three approaches: (1) through an explicit spatial and temporal dependence (Buizza et al., 1999; Berloff, 2005b,a); (2) from subgrid activity partially accounted for by an additional prognostic equation (Majdaa and Gershgorin, 2010); (3) from information about the past and present values of resolved variables or past values of the parametrization itself (e.g., Frederiksen, 1999; Khouider et al., 2003; Shutts and Palmer, 2007; Crommelin and Vanden-Eijnden, 2008).

In this study we have attempted to construct a stochastic parameterization of ocean mesoscale eddies similar to the third approach in order to account for the fluctuations in subgrid transport, to represent upscale turbulent cascades, and to account for model error associated with the uncertainty in the parameters and the parametrization. We have coarse-grained the output and equation of a high-resolution quasi-geostrophic (QG) model in a double-gyre configuration, giving rise to an eddy source term S^* which represents the eddy-eddy and eddy-mean flow interactions and the Reynolds stresses.

The complex spatial, temporal and statistical properties of the eddy source term are analyzed as a function of the resolved scales and external parameters. The functional form of the resolved scales, based on a representation of turbulence as a Non-Newtonian viscoelastic medium acting on a mean flow field, is used to describe the eddy source term and its impact on the mean flow. The eddy source term is described using an expression similar to the Rivlin–Ericksen stress for fluids of second-grade (Rivlin and Ericksen, 1955; Holm and Nadiga, 2003), given by $\overline{\nabla}^2 \mathrm{D}_t\bar{q}$ and is related to the rate of strain and shear. The Rivlin–Ericksen stress, which is found to have a high correlation with S^* in most parts of the basin, captures the mean, fluctuations and decorrelation timescale of the eddy source term. We then further explored the impact of the Rivlin–Ericksen stress as a deterministic and stochastic parameterization for transient mesoscale eddies. The Rivlin–Ericksen representation of the eddy source term is shown to introduce the forcing at small scales in order to maintain the jet. The expression for the Rivlin–Ericksen stress used in the present study is mostly understood as a shearing and straining of the flow field but a theoretical derivation remains to be done in order to explain why the functional $\overline{\nabla}^2 \mathrm{D}_t\bar{q}$ is such a good predictor for the eddy source term S^*.

We have used a probability distribution function (PDFs), conditional on the functional $\mathcal{R}(\bar{q})$ of the resolved-scale prognostic variable for the Rivlin-Ericksen stress. The conditional PDFs present high kurtosis (fat tails) and skewness (asymmetry), meaning that extreme fluctuations are not only more frequent than in a Gaussian case but can be more frequent toward one sign than another (with respect to the mean). Such behavior is common in nonlinear turbulent flow (Majda and Kramer, 1999; Bourlioux and Majda, 2002; Majda and Gershgorin, 2013). For this reason, at least four moments of the cPDFs must be preserved. We found that the coarse-grained resolution, wind strength, and stratification are the only parameters necessary to determine the first four moments: the mean μ is proportional to $\overline{\nabla}^2 \mathrm{D}_t\bar{q}$ and to the square of the resolution; the standard deviation σ is proportional to the wind strength, inversely proportional to resolution and layer thickness, and asymptotically independent of the magnitude of the functional $\overline{\nabla}^2 \mathrm{D}_t\bar{q}$; the third and fourth standardized moments are constants of order one. Using these four moments, the cPDFs can be reconstructed with high accuracy via the maximum-entropy method of information theory. Therefore the eddy source term along with its spatial pattern and temporal fluctuations can be reconstructed using the coarse-resolution grid size, the wind forcing and the stratification. In light of these scalings, the implementation of a stochastic closure in a coarse resolution model based on the conditional PDFs requires in principle very little tuning and no preliminary high-resolution (QG) model runs to diagnose the subgrid forcing needed to force the coarse-resolution model. It would be interesting to explore how the relationship holds in primitive equation models and to use observations (Holloway, 1986) to diagnose the eddy forcing as given by the functional $\overline{\nabla}^2 \mathrm{D}_t\bar{q}$.

The implementation of the deterministic and stochastic closures are currently being tested in a variety of numerical models. The underlying functional $\overline{\nabla}^2 \mathrm{D}_t\bar{q}$ shares common features with the deterministic parameterization of Nadiga and Margolin (2001), Holm and Wingate, 2005, and Nadiga and Bouchet, 2011. Therefore there is some confidence that the backscatter and jet rectification can be reproduced with some success. Moreover similarly to Berloff (2005b,a) the eddy source term temporal and spatial correlations can be reproduced by using the cPDFs and their statistical spread therefore taking into account the sub-grid scale fluctuations. However the advantage over the other studies is that we now have a scaling to predict the mean and fluctuations of the eddy source and its dependence on the coarse-resolution model, forcing and stratification, without the need of high-resolution runs (at least in QG).

Yet, the implementation presents some intricacies owing to the presence of the Lagrangian time derivative in the functional $\overline{\nabla}^2 \mathrm{D}_t\bar{q}$ and to a negative proportionality constant κ. We are currently investigating several approaches to evaluate $\overline{\nabla}^2 \mathrm{D}_t\bar{q}$ (e.g., from a previous time step, using implicit method or by Helmholtz inversion). In addition, the stability properties of the parameterization are being investigated in relation to the sign of the κ which represents the relationship between the mean of the cPDFs and the conditional variable. All aspects of the numerical implementation will be addressed in a follow-up paper.

Acknowledgements

We thank James Maddison for helpful advice, and Tomos David for the MaxEnt fit computation. Many thanks to: Dave Munday, Fenwick Cooper, Andrew Majda, Raf Ferrari, Ray Pierrehumbert, Carl Wunsch, and Todd Ringler for very stimulating discussions. We would like to acknowledge Malte Jansen, Pavel Berloff and two anonymous reviewers for their constructive comments on the manuscript. PM thanks Mari, Miri, and Ruxi for warm encouragement. This work was funded by the John Fell Oxford University Press (OUP) Research Fund.

Appendix A. Description of the coarse graining methodology

The coarse-graining methodology used in this study is based on the work by Murdoch and Bedeaux (1994). For an additive quantity (such as mass, momentum or energy) ψ on a domain Ω, we define its coarse-grained value $\overline{\psi}$ as the convolution given by

$$\overline{\psi}(\mathbf{x},t) = \int_\Omega W(\mathbf{x}-\mathbf{y},t)\psi(\mathbf{y},t)\mathrm{d}\mathbf{y}, \tag{A.1}$$

where W is a weight function such that $\int_\Omega W(\mathbf{x},t)\mathrm{d}\mathbf{x} = 1$, and the total ψ is conserved such that $\int_\Omega \overline{\psi}(\mathbf{x},t)\mathrm{d}\mathbf{x} = \int_\Omega \psi(\mathbf{y},t)\mathrm{d}\mathbf{y}$. If the weight function W is positive-definite, Eq. (A.1) is equivalent to a weighted space average of ψ. The weight is typically isotropic, $W(\mathbf{x}) = W(|\,\mathbf{x}\,|)$, and concentrated in a range $\Delta\mathbf{x}$ which defines the coarse-graining scale. However if the convolution (A.1) corresponds to a sharp truncation of ψ in wavenumber space then it would lead to W being non-positive definite. When W is a Gaussian with standard deviation Σ, Eq. (A.1) corresponds to a spatial average at a scale Σ but also to a Fourier averaging to wavenumbers less than $1/\Sigma$.

While it is easy to see that coarse-graining commutes with spatial and temporal derivatives $\partial_t\overline{\psi} = \overline{\partial_t\psi}$, $\nabla\overline{\psi} = \overline{\nabla\psi}$, note that it is not generally true that

$$\overline{\overline{\psi}} = \overline{\psi} \quad \text{exactly}, \tag{A.2}$$

as often assumed in Reynolds averages. Yet, it often correct that $\overline{\overline{\psi}} \approx \overline{\psi}$, approximately (Murdoch and Bedeaux, 1994).

The discrete case is more complex. Suppose that our high-resolution variable ψ_I, where I denotes 2D indexes (I_x, I_y), is defined on a 2D C-grid with cells of areas A_I. The coarse-graining of ψ_I onto a variable $\overline{\psi}_i$ on a low-resolution C-grid with indexes (i_x, i_y) and cells of areas a_i is given by

$$\overline{\psi}_i = \sum_I W_{iI}\psi_I \tag{A.3}$$

with

$$\sum_i W_{iI}a_i = A_I, \quad \sum_I W_{iI} = 1, \tag{A.4}$$

which ensures $\sum_I A_I\psi_I = \sum_i a_i\overline{\psi}_i$ (the time dependence was ignored for simplicity). The indices i and I will have different summation ranges. The discrete spatial-derivative operators do not commute with the discrete coarse-graining such that

$$\overline{\nabla\psi} \neq \overline{\nabla}\overline{\psi}, \tag{A.5}$$

where the finite-difference operators $\overline{\nabla}$ and ∇ operates on the low-resolution and high-resolution grid, respectively.

In the present study we opt for a simple weight function W satisfying (A.4), namely a constant one, having support on the minimal length scale of the coarse-resolution model and isotropic in the interior of the domain. In addition, energy, momentum and vorticity are conserved. For coarse-graining from 7.5 km to 30 km the resulting weight function is shown in Fig. A.16. We want to coarse-grain the streamfunction ψ and PV q from a high-resolution grid to obtain $\overline{\psi}$, $\bar{q}$ onto a low-resolution grid. Eq. (A.3) cannot be applied to both ψ and q simultaneously as it would lead to the following discrepancy:

$$\begin{aligned}\bar{q} &= \overline{\nabla^2\psi} + \beta\bar{y} + \frac{\partial}{\partial z}\left(\frac{f_0^2}{N^2}\frac{\partial\overline{\psi}}{\partial z}\right)\\ &\neq \tilde{q} = \overline{\nabla^2}\overline{\psi} + \beta\bar{y} + \frac{\partial}{\partial z}\left(\frac{f_0^2}{N^2}\frac{\partial\overline{\psi}}{\partial z}\right),\end{aligned} \tag{A.6}$$

with a difference $\bar{q} - \tilde{q} = \overline{\nabla^2\psi} - \overline{\nabla^2}\overline{\psi}$ stemming from the non-commutativity (A.5).

We can choose to coarse-grain ψ and define the low-resolution q in terms of the coarse-grained $\overline{\psi}$ however it would involve a computationally costly Helmholtz inversion. We therefore opt to coarse-grain q and define the low-resolution streamfunction in terms of the coarse-grained $\tilde{q}$ which is a relatively straightforward operation. We define the low-resolution PV as

$$\tilde{q} = \overline{\nabla^2}\overline{\psi} + \beta\bar{y} + \frac{\partial}{\partial z}\left(\frac{f_0^2}{N^2}\frac{\partial\overline{\psi}}{\partial z}\right) \tag{A.7}$$

and the low-resolution velocity $\tilde{\mathbf{u}} = (-\overline{\partial_y\psi}, \overline{\partial_x\psi})$.

The coarse-graining of the prognostic model equation given in Eq. (2) leads to

$$\frac{\partial\bar{q}}{\partial t} + \overline{\nabla\cdot(\mathbf{u}q)} = \nu\overline{\nabla^4\psi} + \overline{F}, \tag{A.8}$$

such that the coarse resolution PV $\tilde{q}$ is given by the following equation

$$\frac{\partial\tilde{q}}{\partial t} + \overline{\nabla}\cdot(\tilde{\mathbf{u}}\tilde{q}) = \tilde{\nu}\overline{\nabla}^4\overline{\psi} + \overline{F} + S^*, \tag{A.9}$$

where

$$\begin{aligned}S^* = {}&\overline{\nabla}\cdot(\tilde{\mathbf{u}}\tilde{q}) - \overline{\nabla\cdot(\mathbf{u}q)} + \nu\overline{\nabla^4\psi} - \tilde{\nu}\overline{\nabla}^4\overline{\psi} + \partial_t(\overline{\nabla}^2\overline{\psi}\\ &- \overline{\nabla^2\psi}).\end{aligned} \tag{A.10}$$

This expression differs from that of Eq. (8) for the small non-commutativity term $\partial_t(\overline{\nabla}^2\overline{\psi} - \overline{\nabla^2\psi})$. In the main text we use $\bar{q}, \bar{\mathbf{u}}$ instead of $\tilde{q}, \tilde{\mathbf{u}}$ for simplicity however, all calculations and figures take into account the non-commutativity (A.5) for completeness.

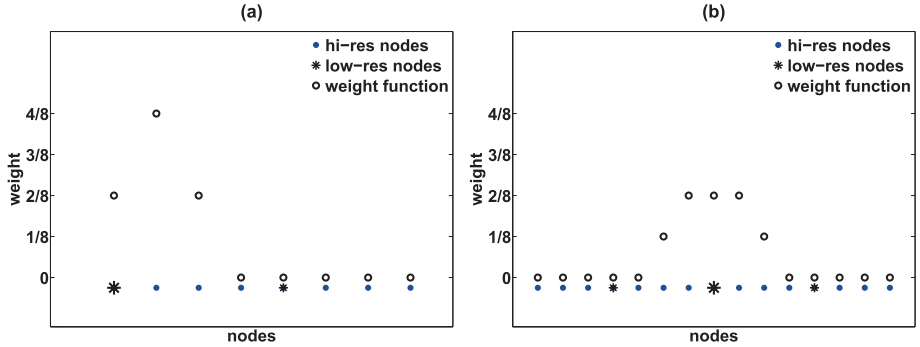

Fig. A.16. One-dimensional weight functions for a coarse-graining from 7.5 km to 30 km at: (a) a boundary node and (b) a generic interior node. The values at the high-resolution grid nodes (blue dots) of the variable to be coarse-grained are summed with the weights shown (black circles) and assigned to the low-resolution grid node (black asterisks). The weights cannot be exactly constant owing to the constraints (A.4). The 2D weight function is the product of two 1D weight functions operating in the meridional and zonal directions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Appendix B. Traditional closures: GM, PV mixing, shear

We examine the correlations of several conditional variables, closely related to the traditional closures of the time-mean effects of eddies onto the mean flow or related to the initiation of instability associated with the eddies. Three main potential "parameterizations" of the eddy source term that could be used as conditional variables are:

1. Gent and McWilliams (1990) mimicking time-mean baroclinic instability, $\mathcal{R}(\bar{q}) = \overline{\nabla} \cdot \frac{\partial}{\partial z}\left(\frac{f_0}{N^2}\overline{\nabla}\bar{b}\right)$;
2. PV homogenization such that the effect of the eddies is to mix PV, $\mathcal{R}(\bar{q}) = \overline{\nabla} \cdot \overline{\nabla}\bar{q}$ (Rhines and Young, 1982b);
3. a shear-based parameterization term, $\mathcal{R}(\bar{q}) = \|\overline{\nabla}_3\bar{\mathbf{u}}\| = \sqrt{\mathrm{Tr}[(\overline{\nabla}_3\bar{\mathbf{u}} + \overline{\nabla}_3\bar{\mathbf{u}}^{\mathrm{T}})^2]}$, where Tr is the trace of the matrix and the superscript T denotes the transpose. The term includes barotropic and baroclinic shears.

Very poor correlations between GM or PV mixing, $\mathcal{R}(\bar{q}) = \overline{\nabla}^2\bar{q}$ and the eddy source terms are found. For example, Fig. B.17a shows the correlation between the PV mixing and the eddy source term (the correlation between the eddy source term and the GM-term is almost identical to that Fig. B.17(a)). The complexity of the eddy source term cannot be approximated by a simple down-gradient closure as the sign of S^* varies significantly along and across the jet, unless a free parameter is allowed to vary with space and time.

Figs. B.17(b) and (c) show the cPDFs calculated for the conditional variables $\mathcal{R}(\bar{q})$ defined as $\overline{\nabla}^2\bar{q}$ and $\|\overline{\nabla}_3\bar{\mathbf{u}}\|$. No correlations between these different conditional variables and the source term are found. The mean of the PDFs is close to zero for most values of the conditional variables and the spread in the PDFs is relatively large, spanning positive and negative values equally for any given value of $\mathcal{R}(\bar{q})$. Any spatial or temporal correlation would have to be brought about by introducing spatially or temporally varying tensorial coefficients using more complex conditional variables of the form $\overline{\nabla} \cdot \kappa_{\mathrm{RY}}\overline{\nabla}\bar{q}$ with $\kappa_{\mathrm{RY}} = \kappa_{\mathrm{RY}}(\bar{x}, \bar{y}, z, t)$. However, spatially varying tensors would ultimately break the frame-invariance of the parametrization. For the shear-based parametrization, the spread in the PDFs increases for large values of the shear leading to possibly extreme values of the eddy source term, in agreement with the expectation that the eddies are derived from strong shears but there is no clear correlation between the mean of the PDF and the conditional variable, which makes the use of such relationship limited. While those parameterizations have been proven successful when implemented in coarse-resolution models, their poor correlation with the eddy source term and its fluctuations will ultimately lead to errors in the representation of the transient sub-grid scale processes and their impact on the mean flow.

Appendix C. Okubo–Weiss decomposition and LANS-α

While the conditional variable is related to the Rivlin–Ericksen tensor it can also be described as a combination of the strain rate and the Lagrangian rate of change of the potential vorticity gradient (Okubo, 1970; Weiss, 1991). Let us consider the potential vorticity q equation given in Eq. (2), $\frac{\mathrm{D}q}{\mathrm{D}t} = \frac{\partial q}{\partial t} + \mathbf{u} \cdot \nabla q = \mathcal{D} + F$. The potential vorticity q is conserved along a Lagrangian trajectory except for the sources and sinks due to dissipation and forcing. Taking the gradient of Eq. (2) and using the relation $\mathrm{D}_t\nabla q = \nabla \mathrm{D}_t q - (\nabla\mathbf{u})^{\mathrm{T}}\nabla q$, the equation for the potential vorticity gradient can be written as

$$\frac{D\nabla q}{Dt} = -(\nabla\mathbf{u})^{\mathrm{T}}\nabla q + \nabla(F + \mathcal{D}), \tag{C.1}$$

where $\nabla\mathbf{u}$ is the velocity gradient tensor. The effect of dissipation is to weaken the gradient of PV while the forcing can intensify and rotate the PV gradient. The stress in our parameterization, $\nabla \mathrm{D}_t q$, can therefore be seen as an additional forcing acting on the local rate of strain $(\nabla\mathbf{u})^{\mathrm{T}}\nabla q$ and the creation of sharper PV gradients in addition on to the Lagrangian rate of change of PV gradient $(\mathrm{D}_t\nabla q)$ giving the orientation of the growth due to the evolution of the water parcels. The Lagrangian rate of change appears to contribute to the orientation of the eddy flow. Fig. C.18(a) shows the mean PV contour lines superimposed on the angle formed by the gradient of $\kappa\overline{\nabla} \cdot \frac{\mathrm{D}\overline{\nabla}\bar{q}}{\mathrm{D}t}$ with the meridional axis. Positive values of the angle translate into a counterclockwise rotation of $\overline{\nabla}\bar{q}$ while negative values into a clockwise rotation. The term related to the Lagrangian rate of change thus acts to rotate the PV gradient. Fig. C.18(b) shows the time mean PV contours in black superimposed on the averaged $\kappa\overline{\nabla} \cdot (\overline{\nabla}\bar{\mathbf{u}})^{\mathrm{T}}\overline{\nabla}\bar{q}$. The term related to the strain rate is seen to coincide with strong PV gradients and therefore believed to be responsible for the sharpening, rectification and intensification of the jet. The PDFs of the eddy source term conditional on $\overline{\nabla} \cdot \frac{\mathrm{D}\overline{\nabla}\bar{q}}{\mathrm{D}t}$ and on $\overline{\nabla} \cdot [(\overline{\nabla}\bar{\mathbf{u}})^{\mathrm{T}}\overline{\nabla}\bar{q}]$ are shown in Figs. C.18(c) and (d). While $\overline{\nabla} \cdot [(\overline{\nabla}\bar{\mathbf{u}})^{\mathrm{T}}\overline{\nabla}\bar{q}]$ appears to enhance the PV gradient and the strength of the jet, it does not produce as good a conditional variable as $\frac{\overline{\nabla}\mathrm{D}\bar{q}}{\mathrm{D}t}$ or $\frac{\mathrm{D}\overline{\nabla}\bar{q}}{\mathrm{D}t}$, mostly due to the poor correlations in certain parts of the basin (not shown).

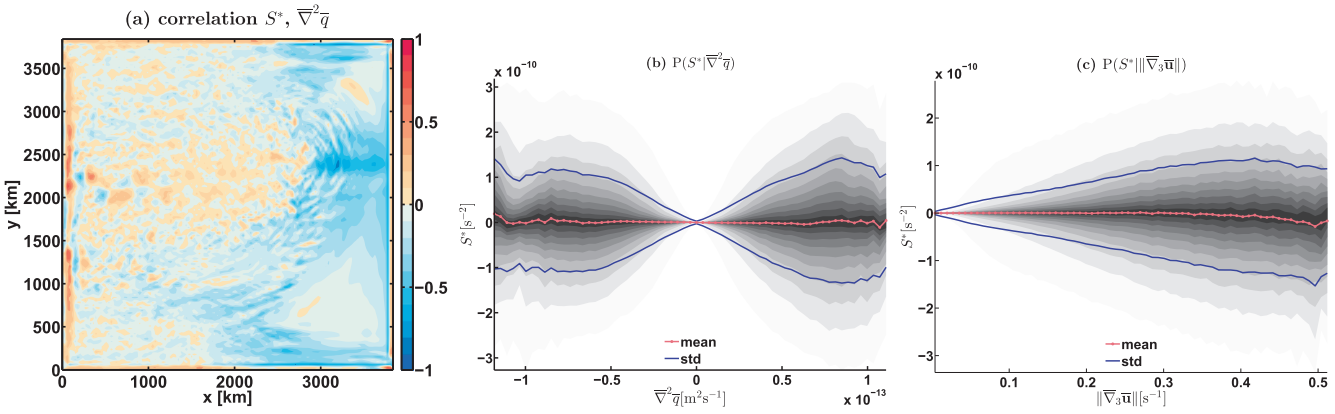

Fig. B.17. (a) Correlations of the eddy source with PV homogeneization/diffusion $\overline{\nabla}^2\bar{q}$. Also shown are cPDFs for the eddy source term S^* conditional on (b) $\overline{\nabla}^2\bar{q}$ and on (c) magnitude of the barotropic and baroclinic shear $\mathcal{R}(\bar{q}) = \|\overline{\nabla}_3\bar{\mathbf{u}}\| = \sqrt{\mathrm{Tr}[(\overline{\nabla}_3\bar{\mathbf{u}} + \overline{\nabla}_3\bar{\mathbf{u}}^{\mathrm{T}})^2]}$. The grey shading shows 5% quantiles increments of the conditional PDFs, the red and blue lines are the mean and standard deviation, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

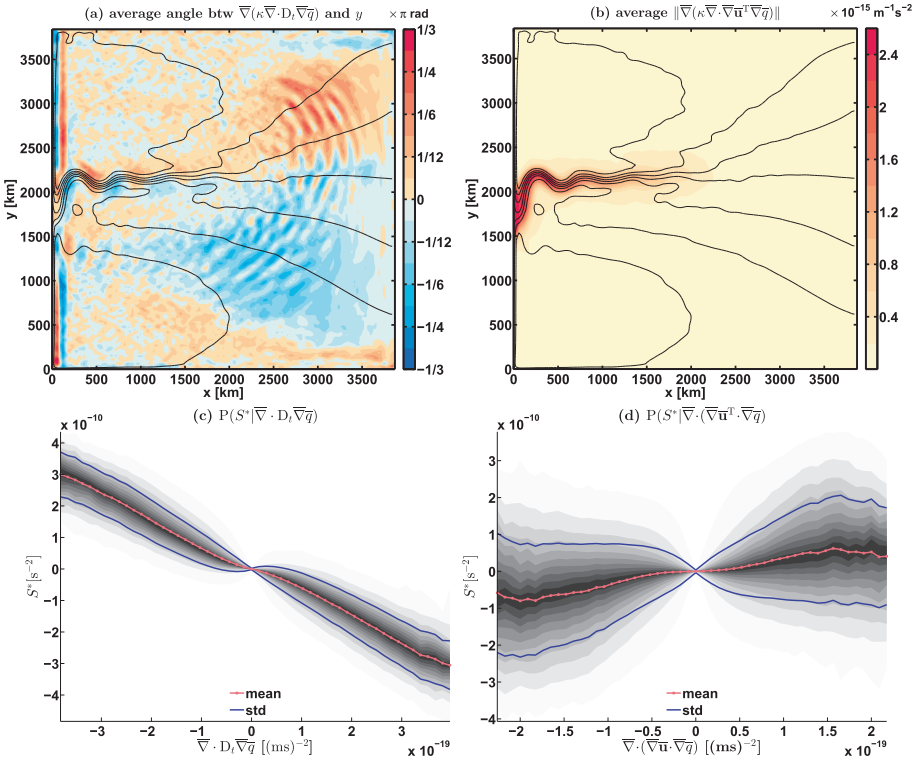

Fig. C.18. Okubo–Weiss Decomposition of the conditional variable: effect of (a) $\kappa\overline{\nabla}\cdot\frac{D\overline{\nabla}\bar{q}}{Dt}$ and (b) $\kappa\overline{\nabla}\cdot[\overline{\nabla}\bar{\mathbf{u}}]^T\overline{\nabla}q$ on potential vorticity. Also shown are PDFs of the eddy source term S^* conditioned on (c) $\overline{\nabla}\cdot\frac{D\overline{\nabla}\bar{q}}{Dt}$ and (d) $\overline{\nabla}\cdot[\overline{\nabla}\bar{\mathbf{u}}]^T\overline{\nabla}\bar{q}$ (30 km resolution).

In addition to the similarities with the Okubo–Weiss straining-shearing of the fluid see also for uniform strain (see also Waugh et al., 2012 for PDF under uniform strain) the RE parameterization shares many common properties with the Lagrange-averaged Navier–Stokes-α (LANS-α) model (e.g., Chen et al., 1998; Nadiga and Margolin, 2001; Holm and Wingate, 2005; Nadiga and Bouchet, 2011) as discussed in Section 4.1. The model is presented as a Lagrangian-averaged filter of the velocities. Nadiga and Margolin (2001) and Holm and Nadiga (2003) derived several baroclinic QG limits of the LANS-α model. One example of such limit can be formulated as a parameterization taking the form: $S_{\mathrm{HN}} = \mathrm{D}_t\overline{\nabla}^2\bar{q} - \overline{\nabla}^2(F+\mathcal{D})$. The cPDFs of the eddy source term conditioned on the Holm and Nadiga (2003) parameterization S_{HN} (not shown) exhibits a good correlation but not as tight as the one formulated in Eq. (9). While both formulations share some common mathematical and physical motivation, the conditional variable $\overline{\nabla}^2\mathrm{D}_t\bar{q}$ has a simpler mathematical form which could potentially lead to an easier numerical implementation of a stochastic parameterization based on the cPDFs. Moreover, the operator defined in Eq. (11) acts upon the Eulerian tendency and the advective term, which is not the case for the LANS-α model. Last, but not least, the parameter κ in our study is negative while the α parameter is positive in the α model.

Appendix D. Reconstruction of the conditional PDFs using Shannon entropy

To reconstruct the conditional PDFs in such a way as not to decrease the original statistical spread of the distributions, consider the Shannon entropy $H(\mathbf{p}) = -\sum_i p_i \ln p_i$ (Shannon, 1948; Bernardo, 1979; Kullback and Leibler, 1951) as a measure of the spread of a discrete probability distribution $\mathbf{p} = (p_i)$. Therefore we can choose the maximum-entropy method (Jaynes, 1982; Sivia, 1990) as our reconstruction technique. The maximum-entropy method chooses, among all distributions that satisfy a set of linear constraints (the four moments in our case), the one which has maximum Shannon entropy. This procedure outlined in Mead and Papanicolaou (1984) proves to be computationally efficient given our set of constraints. For the discrete case, p_i are the probability for the value S_i^* of eddy the source term and the constraints are the values of the four raw moments μ_k (which can easily be found from the standardized moments) such that $\sum_i p_i S_i^{*k} = \mu_k$ with $k = 1,\dots c, 4$. The variational optimization problem is therefore given by

$$\delta\left(H(\mathbf{p}) - \sum_k \lambda_k \sum_i p_i S_i^{*k}\right) = 0, \tag{D.1}$$

where λ_k are Lagrange multipliers, and leads to the probabilities

$$p_i = \exp\left(-\sum_k \lambda_k S_i^{*k}\right) \Big/ Z, \tag{D.2}$$

with $Z(\lambda_k) = \sum_i \exp(-\sum_k \lambda_k S_i^{*k})$. Performing a Legendre transformation of expression (D.1) and substituting (D.2), the sought λ_k are those for which convex function

$$\delta\left(\ln Z(\lambda_k) + \sum_k \lambda_k \mu_k\right) = 0 \tag{D.3}$$

and all other quantities, μ_k and S_i^*, are known. The optimization problem is computationally fast and cheap. Once the four Lagrange multipliers λ_k are found, they can be replaced in the expressions for the probabilities (D.2), so that the discrete conditional PDFs $p_i = p_i(S_i^*, \mu_k)$ are entirely reconstructed.

References

Arakawa, A., 1966. Computational design for long-term numerical integration of the equations of fluid motion: two-dimensional incompressible flow. Part I. J. Comput. Phys. 1 (1), 119–143.

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., Dewar, W., Kravtsov, S., McWilliams, J., 2007a. Ocean eddy dynamics in a coupled ocean-atmosphere model. J. Phys. Oceanogr. 37 (5), 1103–1121.

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

Bernardo, J.-M., 1979. Expected information as expected utility. Ann. Stat. 7 (3), 686–690.

Bourlioux, A., Majda, A.J., 2002. Elementary models with probability distribution function intermittency for passive scalars with a mean gradient. Phys. Fluids 14 (2), 881–897.

Bowler, N.E., Arribas, A., Beare, S.E., Mylne, K.R., Shutts, G.J., 2009. The local ETKF and SKEB: upgrades to the MOGREPS short-range ensemble prediction system. Q.J.R. Meteorol. Soc. 135 (640), 767–776.

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

Buizza, R., Miller, M., Palmer, T.N., 1999. Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. Q.J.R. Meteorol. Soc. 125 (560), 2887–2908.

Burbidge, J.B., Magee, L., Robb, A.L., 1988. Alternative transformations to handle extreme values of the dependent variable. J. Am. Stat. Assoc. 83 (401), 123–127.

Cessi, P., Fantini, M., 2004. The eddy-driven thermocline. J. Phys. Oceanogr. 34 (12), 2642–2658.

Chen, S., Foias, C., Holm, D.D., Olson, E., Titi, E.S., Wynne, S., 1998. Camassa-Holm equations as a closure model for turbulent channel and pipe flow. Phys. Rev. Lett. 81 (24), 5338–5341.

Ching, E.S.C., Kraichnan, R.H., 1998. Exact results for conditional means of a passive scalar in certain statistically homogeneous flows. J. Stat. Phys. 93 (3–4), 787–795.

Cooper, F.C., Haynes, P.H., 2011. Climate sensitivity via a nonparametric fluctuation-dissipation theorem. J. Atmos. Sci. 68 (5), 937–953.

Crommelin, D.T., Vanden-Eijnden, E., 2008. Subgrid-scale parameterization with conditional Markov chains. J. Atmos. Sci. 65 (8), 2661–2675.

Crow, S.C., 1967. Viscoelastic character of finegrained isotropic turbulence. Phys. Fluids 10 (7), 1587–1589.

Danabasoglu, G., McWilliams, J., Gent, P., 1994. The role of mesoscale tracer transports in the global ocean circulation. Science 264 (5162), 1123–1126.

Dorrestijn, J., Crommelin, D.T., Siebesma, A.P., Jonker, H.J.J., 2013. Stochastic parameterization of shallow cumulus convection estimated from high-resolution model data. Theor. Comput. Fluid Dyn. 27 (1–2), 133–148.

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

Dunn, J.E., Fosdick, R.L., 1974. Thermodynamics, stability, and boundedness of fluids of complexity 2 and fluids of second grade. Arch. Ration. Mech. Anal. 56 (3), 191–252.

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.

Eden, C., Greatbatch, R.J., 2008a. Diapycnal mixing by meso-scale eddies. Ocean Modell. 23 (3–4), 113–120.

Eden, C., Greatbatch, R.J., 2008b. Towards a mesoscale eddy closure. Ocean Modell. 20 (3), 223–239.

Farrell, B.F., Ioannou, P.J., 2003. Structural stability of turbulent jets. J. Atmos. Sci. 60 (17), 2101–2118.

Farrell, B.F., Ioannou, P.J., 2007. Structure and spacing of jets in barotropic turbulence. J. Atmos. Sci. 64 (10), 3652–3665.

Foiaş, C., Holm, D.D., Titi, E.S., 2001. The Navier–Stokes-alpha model of fluid turbulence. Physica D 152–153, 505–519.

Franzke, C., Majda, A.J., Branstator, G., 2007. The origin of nonlinear signatures of planetary wave dynamics: mean phase space tendencies and contributions from non-gaussianity. J. Atmos. Sci. 64 (11), 3987–4003.

Frederiksen, J.S., 1999. Subgrid-scale parameterizations of eddy-topographic force, eddy viscosity, and stochastic backscatter for flow over topography. J. Atmos. Sci. 56 (11), 1481–1494.

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.

Frederiksen, J.S., Kepert, S.M., 2006. Dynamical subgrid-scale parameterizations from direct numerical simulations. J. Atmos. Sci. 63 (11), 3006–3019.

Frederiksen, J.S., O'Kane, T.J., Zidikheri, M.J., 2012. Stochastic subgrid parameterizations for atmospheric and oceanic flows. Phys. Scr. 85 (6), 068202.

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

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

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

Henning, C., Vallis, G., 2004. The effects of mesoscale eddies on the main subtropical thermocline. J. Phys. Oceanogr. 34 (11), 2428–2443.

Herring, J.R., Kraichnan, R.H., 1972. Comparison of some approximations for isotropic turbulence. In: Rosenblatt, M., Van Atta, C. (Eds.), Statistical Models and Turbulence, Lecture Notes in Physics, vol. 12. Springer-Verlag, Berlin, pp. 148–194.

Holloway, G., 1986. Estimation of oceanic eddy transports from satellite altimetry. Nature 323 (6085), 243–244.

Holloway, G., 2004. From classical to statistical ocean dynamics. Surv. Geophys. 25 (3–4), 203–219.

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

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

Hu, Y., Pierrehumbert, R.T., 2001. The advection-diffusion problem for stratospheric flow. Part I: concentration probability distribution function. J. Atmos. Sci. 58 (12), 1493–1510.

Hughes, C.W., Thompson, A.F., Wilson, C., 2010. Identification of jets and mixing barriers from sea level and vorticity measurements using simple statistics. Ocean Modell. 32 (1–2, SI), 44–57.

Jaynes, E.T., 1982. On the rationale of maximum-entropy methods. Proc. IEEE 70 (9), 939.

Johnson, N.L., 1949. Systems of frequency curves generated by methods of translation. Biometrika 36 (1/2), 149–176.

Khouider, B., Majda, A., Katsoulakis, M., 2003. Coarse-grained stochastic models for tropical convection and climate. Proc. Natl. Acad. Sci. U.S.A. 100 (21), 11941–11946.

Kirtman, B., Vecchi, G.A. 2011. Why climate modelers should worry about atmospheric and oceanic weather, pp. 511–524.

Kitsios, V., Frederiksen, J.S., Zidikheri, M.J., 2012. Subgrid model with scaling laws for atmospheric simulations. J. Atmos. Sci. 69 (4), 1427–1445.

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

Kitsios, V., Frederiksen, J.S., Zidikheri, M.J., 2013. Scaling laws for parameterisations of subgrid eddy-eddy interactions in simulations of oceanic circulations. Ocean Modell. 68 (8), 88–105.

Kraichnan, R.H., 1959. The structure of isotropic turbulence at very high Reynolds numbers. J. Fluid Mech. 5 (4), 497–543.

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

Kullback, S., Leibler, R.A., 1951. On information and sufficiency. Ann. Math. Stat. 22 (1), 79–86.

LaCasce, J.H. 2008. Lagrangian statistics from oceanic and atmospheric observations. In: Weiss, J.B., Provenzale, A. (Eds.), Transport and Mixing in Geophysical Flows, Lecture Notes in Physics, vol. 744, pp. 165–218.

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

Li, H., von Storch, J.-S., 2013. On the fluctuating buoyancy fluxes simulated in a $\frac{1}{10}^{\circ}$ OGCM. J. Phys. Oceanogr. 43 (7), 1270–1287.

Liepmann, H.W. 1962. Free turbulent flows. In: Favre, A. (Ed.), Mécanique de la Turbulence, Colloques Internationaux du CNRS, CNRS, vol. 108, pp. 211–251.

Lorenz, E.N. 1995. Predictability: a problem partly solved. In: Seminar on Predictability, 4–8 September 1995, Reading. ECMWF.

Lumley, J.L., 1970. Toward a turbulent constitutive relation. J. Fluid Mech. 41 (2), 413–434.

Macvean, M.K., 1983. The effects of horizontal diffusion on baroclinic development in a spectral model. Q.J.R. Meteorol. Soc. 109 (462), 771–783.

Majda, A.J., Gershgorin, B., 2010. Quantifying uncertainty in climate change science through empirical information theory. Proc. Natl. Acad. Sci. U.S.A. 107 (34), 14958–14963.

Majda, A.J., Gershgorin, B., 2013. Elementary models for turbulent diffusion with complex physical features: eddy diffusivity, spectrum and intermittency. Philos. Trans. R. Soc. A 371 (1982), 20120184.

Majda, A.J., Kramer, P.R., 1999. Simplified models for turbulent diffusion: theory, numerical modelling, and physical phenomena. Phys. Rep. 314 (4–5), 237–574.

Marshall, D., 1997. Subduction of water masses in an eddying ocean. J. Mar. Res. 55 (2), 201–222.

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

Mason, P.J., Thomson, D.J., 1992. Stochastic backscatter in large-eddy simulations of boundary layers. J. Fluid Mech. 242 (9), 51–78.

McTigue, D.F., Passman, S.L., Jones, S.J., 1985. Normal stress effects in the creep of ice. J. Glaciol. 31 (108), 120–126.

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

Meredith, M.P., Hogg, A.M., 2006. Circumpolar response of southern ocean eddy activity to a change in the southern annular mode. Geophys. Res. Lett. 33 (16), L16608.

Munday, D.R., Johnson, H.L., Marshall, D.P., 2013. Eddy saturation of equilibrated circumpolar currents. J. Phys. Oceanogr. 43 (3), 507–532.

Murdoch, A.I., Bedeaux, D., 1994. Continuum equations of balance via weighted averages of microscopic quantities. Proc. R. Soc. London A 445 (1923), 157–179.

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

Nadiga, B.T., Margolin, L.G., 2001. Dispersive-dissipative eddy parameterization in a barotropic model. J. Phys. Oceanogr. 31 (8), 2525–2531.

Neelin, J.D., Peters, O., Lin, J.W.-B., Hales, K., Holloway, C.E., 2008. Rethinking convective quasi-equilibrium: observational constraints for stochastic convective schemes in climate models. Philos. Trans. R. Soc. A 366 (1875), 2581–2604.
O'Kane, T.J., Frederiksen, J.S., 2008. Statistical dynamical subgrid-scale parameterizations for geophysical flows. Phys. Scr. 2008, T132.
Okubo, A., 1970. Horizontal dispersion of floatable particles in vicinity of velocity singularities such as convergences. Deep Sea Res. 17 (3), 445–454.
Palmer, T.N., 2001. A nonlocal dynamical perspective on model error: a proposal for nonlocal stochastic-dynamic parametrization in weather and climate prediction models. Q.J.R. Meteorol. Soc. 127, 279–304.
Palmer, T.N., 2012. Towards the probabilistic Earth-system simulator: a vision for the future of climate and weather prediction. Q.J.R. Meteorol. Soc. 138, 841–861.
Penland, C., Sardeshmukh, P.D., 1995. The optimal growth of tropical sea surface temperature anomalies. J. Climate 8 (8), 1999–2024.
Qiu, B., Scott, R.B., Chen, S., 2008. Length scales of eddy generation and nonlinear evolution of the seasonally modulated South Pacific subtropical countercurrent. J. Phys. Oceanogr. 38 (7), 1515–1528.
Rhines, P.B., Young, W.R., 1982a. A theory of wind-driven circulation. I. Mid-ocean gyres. J. Mar. Res. 40 (S), 559–596.
Rhines, P.B., Young, W.R., 1982b. Homogenization of potential vorticity in planetary gyres. J. Fluid Mech. 122, 347–367.
Riesen, P., Hutter, K., Funk, M., 2010. A viscoelastic Rivlin–Ericksen material model applicable to glacier ice. Nonlinear Processes Geophys. 17 (6), 673–684.
Ring, M.J., Plumb, R.A., 2008. The response of a simplified GCM to axisymmetric forcings: applicability of the fluctuation-dissipation theorem. J. Atmos. Sci. 65 (12), 3880–3898.
Rivlin, R.S., 1957. The relation between the flow of non-Newtonian fluids and turbulent Newtonian fluids. Q. Appl. Math. 15 (2), 212–215.
Rivlin, R.S., Ericksen, J., 1955. Stress-deformation relations for isotropic materials. J. Ration. Mech. Anal. 4 (2), 323–425.
Roberts, M.J., Marshall, D.P., 2000. On the validity of downgradient eddy closures in ocean models. J. Geophys. Res. 105 (C12), 28613–28627.
Scott, R.B., Arbic, B.K., 2007. Spectral energy fluxes in geostrophic turbulence: implications for ocean energetics. J. Phys. Oceanogr. 37 (3), 673–688.
Shannon, C.E., 1948. A mathematical theory of communication. Bell Syst. Tech. J. 27 (379–423), 623–656.
Shutts, G.J., 2005. A kinetic energy backscatter algorithm for use in ensemble prediction systems. Q.J.R. Meteorol. Soc. 131 (612), 3079–3102.
Shutts, G.J., Palmer, T.N., 2007. Convective forcing fluctuations in a cloud-resolving model: Relevance to the stochastic parameterization problem. J. Climate 20 (2), 187–202.
Sivia, D.S., 1990. Bayesian inductive inference, maximum entropy & neutron scattering. Los Alamos Sci. 4 (19), 180–207.
Smagorinsky, J., 1963. General circulation experiments with the primitive equations, i, the basic experiment. Mon. Weather Rev. 91, 99–164.
Speziale, C.G., 1987. On nonlinear $k-l$ and $k-\epsilon$ models of turbulence. J. Fluid Mech. 178, 459–475.
Straub, D.N., 1993. On the transport and angular momentum balance of channel models of the antarctic circumpolar current. J. Phys. Oceanogr. 23 (4), 776–782.
Treguier, A., Held, I., Larichev, V., 1997. Parameterization of quasigeostrophic eddies in primitive equation ocean models. J. Phys. Oceanogr. 27 (4), 567–580.
Truesdell III, C.A., Rajagopal, K.R., 2009. An introduction to the mechanics of fluids. In: Modern Birkhäuser Classics, reprint ed. Birkhäuser, Boston, First publ. 1999.
Vallis, G.K., 2006. Atmospheric and Oceanic Fluid Dynamics: Fundamentals and Large-scale Circulation. Cambridge University Press, Cambridge.
Waterman, S., Hoskins, B.J., 2013. Eddy shape, orientation, propagation, and mean flow feedback in western boundary current jets. J. Phys. Oceanogr. 43 (8), 1666–1690.
Waterman, S., Jayne, S.R., 2011. Eddy-mean flow interactions in the along-stream development of a western boundary current jet: an idealized model study. J. Phys. Oceanogr. 41 (4), 682–707.
Waterman, S., Jayne, S.R., 2012. Eddy-driven recirculations from a localized transient forcing. J. Phys. Oceanogr. 42 (3), 430–447.
Waugh, D.W., Keating, S.R., Chen, M.-L., 2012. Diagnosing ocean stirring: comparison of relative dispersion and finite-time Lyapunov exponents. J. Phys. Oceanogr. 42 (7), 1173–1185.
Weiss, J., 1991. The dynamics of enstrophy transfer in 2-dimensional hydrodynamics. Physica D 48 (2–3), 273–294.
Williams, P.D., 2009. A proposed modification to the Robert-Asselin time filter. Mon. Weather Rev. 137 (8), 2538–2546.
Williams, P.D., Haine, T.W.N., Read, P.L., 2004. Stochastic resonance in a nonlinear model of a rotating, stratified shear flow, with a simple stochastic inertia-gravity wave parameterization. Nonlinear Processes Geophys. 11 (1), 127–135.
Wilson, C., Williams, R.G., 2006. When are eddy tracer fluxes directed downgradient? J. Phys. Oceanogr. 36 (2), 189–201.
Zanna, L., 2012. Forecast skill and predictability of observed Atlantic sea surface temperatures. J. Climate 25 (14), 5047–5056.
Zidikheri, M.J., Frederiksen, J.S., 2010. Stochastic modelling of unresolved eddy fluxes. Geophys. Astrophys. Fluid Dyn. 104 (4), 323–348.