Ocean kinetic energy backscatter parametrizations on unstructured grids: Impact on mesoscale turbulence in a channel

Stephan Juricke[a,b,*], Sergey Danilov[a,b], Anton Kutsenko[b], Marcel Oliver[b]

[a] *Alfred Wegener Institute for Polar and Marine Research, Am Handelshafen 12, 27570 Bremerhaven, Germany*
[b] *Jacobs University, Campus Ring 1, 28759 Bremen, Germany*

Abstract

We present a new energy backscatter parametrization for primitive equation ocean models at eddy-permitting resolution, specifically for unstructured grids. Traditional eddy parametrizations in terms of viscosity closures lead to excessive dissipation of kinetic energy when used with eddy-permitting meshes. Implemented into the FESOM2 ocean model, the backscatter parametrization leads to a more realistic total dissipation of kinetic energy. It maintains a reservoir of dissipated energy and reinjects this subgrid energy at larger scales at a controlled rate. The separation between dissipation and backscatter scales is achieved by using different-order differential operators and/or spatial smoothing. This ensures numerical model stability.

We perform sensitivity studies with different choices of parameter settings and viscosity schemes in a configuration with a baroclinically unstable flow in a zonally reentrant channel with a horizontally uniform mesh. The best backscatter setup substantially improves eddy-permitting simulations at 1/4° and 1/6° resolution, bringing them close to a 1/12° eddy-resolving reference. Improvements are largest for levels of kinetic energy and variability in temperature and vertical velocity. A selected optimal default scheme is then tested in a mixed resolution setup – a channel with narrow transitions between an eddy-permitting and an eddy-resolving subdomain. The

*Corresponding author

Email addresses: s.juricke@jacobs-university.de (Stephan Juricke), sergey.danilov@awi.de (Sergey Danilov), a.kutsenko@jacobs-university.de (Anton Kutsenko), m.oliver@jacobs-university.de (Marcel Oliver)

backscatter scheme is able to adapt dynamically to the different resolutions and moves the diagnostics closer to the high resolution reference throughout the domain.

Our study is a first step toward using backscatter in global variable-mesh ocean models and suggests potential for substantial improvements of ocean mean state and variability at reduced computational cost.

Keywords: Kinetic energy backscatter, subgrid eddy parametrization, inverse energy cascade, viscosity closure, eddy-permitting resolution

---

1. Introduction

Mesoscale eddies contribute to the transport of heat and tracers in many regions of the ocean (e.g. Jayne and Marotzke, 2002; Griffies et al., 2015). They affect the shape and temporal behavior of the ocean currents (e.g. Kirtman et al., 2012) and modify the atmosphere-ocean exchange (e.g. Bryan et al., 2010; Frenger et al., 2016; Roberts et al., 2016). They further play a crucial role in the response to changes in atmospheric forcing in a warming climate, and can, for example, counteract the effect of increasing Westerlies on the Southern Ocean Circumpolar Current (Munday et al., 2013).

The resolution of any ocean or atmosphere model defines, to a large degree, which processes can be explicitly resolved and which need to be parametrized, i.e., represented through theoretical or heuristic equations that describe the mean impact of unresolved subgrid processes on the resolved mean flow. The traditional approach to simulating subgrid mesoscale turbulence is via the Gent–McWilliams parametrization (Gent and McWilliams, 1990) in the tracer equations and an eddy-viscosity closure in the momentum equation. When mesoscale eddies are resolved, only the latter is used. It simulates the effects of baroclinic instability due to subgrid structures and maintains model stability by dissipating enstrophy at the grid scale (see, e.g., Danilov et al., 2019). However, if resolution is insufficient, horizontal viscosity closures – while dissipating enstrophy at small scales – are also excessively dissipating energy. They are thereby degrading the effective resolution of the ocean model even further by damping eddy structures and reducing mean and eddy kinetic energy.

At 1° resolution – still in use today in many climate models (Taylor et al., 2012) – the Gent–McWilliams parametrization is widely used to simulate the mean impact of unresolved eddies on potential energy release. Present-day

global ocean models used for climate research generally move toward an ocean grid resolution of 1/4° (Haarsma et al., 2016) which is eddy-permitting over broad ocean domains. Such models start to resolve mesoscale turbulence. The size of mesoscale eddies, however, depends on the first internal Rossby radius which varies from a few kilometers on shelves in the high latitudes to around 100 km at around 10° N/10° S (Hallberg, 2013; Sein et al., 2017). According to Hallberg (2013), the local Rossby radius should be resolved with at least two mesh points to represent eddies, and even finer resolution may be needed to reduce over-dissipation on scales where physical instabilities occur (Soufflet et al., 2016). Thus, only the largest eddies are fully resolved on 1/4° meshes and most of the mesoscale eddy activity remains part of the unresolved subgrid.

An eddy-permitting model is able to represent some of the inverse energy cascade whereby energy is transferred from smaller to larger scales, a process predicted in theory (e.g. as summarized in Danilov et al., 2019) and observed in the ocean (e.g. Scott and Wang, 2005; Wang et al., 2015). Although the arguments for the underlying physical mechanisms driving an inverse cascade may vary depending on the analyzed flow regime, its existence has been observed in many studies (see, e.g., Wang et al., 2015, and references therein). The challenge of eddy-permitting meshes comes from combining resolved eddies in low latitudes with unresolved, i.e., parametrized eddies in high latitudes through wide areas where eddies could be simulated but suffer from excessive damping by the viscosity closure.

In the real ocean, mesoscale eddy kinetic energy is removed from the system by many processes, such as bottom friction and dissipation in the mixed layer, with the horizontal viscosity contribution being small. To improve eddy-permitting ocean models, the ideal concept is to keep the level and distribution of kinetic energy dissipation close to what is observed in the real ocean, where horizontal viscous energy dissipation is much smaller than in the models. Conceptually, this means that energy dissipated due to horizontal viscosity needs to be reintroduced into the system without violating model stability, one of the main constraints for any numerical simulation. As the inverse energy cascade is at least partly resolved at eddy-permitting resolution, this gives an opportunity to inject energy back into the system at scales larger than those of dissipation, and make use of this inverse energy cascade in the sense of energy backscatter, i.e. energy transferred from smaller back to larger scales, energizing the entire spectrum of wavenumbers.

The concept of energy backscatter parametrizations has a long history

(see, e.g., Frederiksen and Davies, 1997; Kitsios et al., 2013; Jansen and Held, 2014; Jansen et al., 2015; Danilov et al., 2019, and references therein), predominantly in the context of idealized setups. Energy backscatter parametrizations have also been used in the atmospheric sciences and in global atmospheric models (Shutts, 2005; Berner et al., 2009; Leutbecher et al., 2017). In this context, backscatter parametrizations are, to a large degree, used as a tool for estimating the uncertainty due to the lack of information about the unresolved scales. In particular, stochastic backscatter can be used to inflate the dynamical growth of model ensemble spread (Shutts, 2005).

Recently, in the context of increased ocean model resolution, the idea of backscatter as an energetically consistent closure has gathered new momentum and generated promising results in idealized ocean model simulations (Jansen et al., 2015; Klöwer et al., 2018). The approaches in these studies are making efficient use of the special situation of model resolution we currently encounter in ocean models: Energy backscatter can be employed to reinject energy and improve both the mean and variability of connected model quantities.

The next step is to investigate the impact of ocean kinetic energy backscatter in full primitive equation global ocean models. A wide variety of ocean models are currently used in coupled climate models, applying different parametrizations, discretizations, resolutions, and dynamical cores. The finite element sea ice-ocean model FESOM (Danilov et al., 2004; Wang et al., 2014) was the first multiresolution global ocean model coupled to an atmospheric component in a climate model setup (Sidorenko et al., 2015; Rackow et al., 2018). It allows to locally vary grid resolutions by designing telescoped (i.e. locally increased resolution) triangular meshes. The advantage is that mesh refinement can be made where it is necessary, for example, to locally resolve the mesoscale eddy field (Sein et al., 2016, 2017). FESOM has moved to a finite volume dynamical core in FESOM2. In variable-grid simulations, the effect of a coarse domain region upstream or nearby may still be visible in the high resolution areas (see Danilov and Wang, 2015; Sein et al., 2016). Backscatter parametrizations may be used to partly compensate for these effects, making FESOM and FESOM2 an ideal testbed to implement, study, and improve backscatter parametrizations, not only on classical equidistant grids, but also on more challenging grids with varying resolution. On such grids, backscatter is expected to be most active in the low resolution parts, helping to energize the flow there and consequently also in the downstream higher resolution regions.

In this paper, we provide a systematic evaluation of different backscatter implementations in the context of FESOM2 in a 3D zonally reentrant channel setup. Our backscatter model, as in Jansen et al. (2015), applies an unresolved kinetic energy (UKE) equation to the scalar UKE field which keeps track of the energy dissipated by the viscous operator, making it available for reinjection into the momentum equation. Different from Jansen et al. (2015), we use a three-dimensional rather than vertically averaged two dimensional UKE field. We test the backscatter setup on horizontally uniform and variable meshes, and for different viscosity parametrizations with temporally and spatially varying viscosity coefficients. We compare the eddy-permitting simulations with and without backscatter with a high resolution, eddy-resolving reference to investigate the impact of the backscatter on mean and eddy kinetic energy, but also on flow variability in quantities such as temperature and vertical velocity.

The paper is structured as follows. In Section 2, we introduce the model and the experimental setup. We also introduce the viscosity parametrization options and present the new backscatter parametrizations including relevant parameters. Section 3 discusses simulations with backscatter on two uniform grids of $1/4^\circ$ and $1/6^\circ$ resolution and compares them to an eddy-resolving reference of $1/12^\circ$ resolution. We further explore the sensitivity of the backscatter parametrization to different parameter settings and identify an optimal "default" scheme. In Section 4, we test the default scheme in a mixed-resolution setup. We compare simulations with and without backscatter, focusing on the transitions between high and low resolution mesh domains. Section 5 discusses the results and their potential with regard to global ocean simulations.

2. Model and backscatter parametrization

2.1. Model setup

The model used for this study is the current version of FESOM2 (Danilov et al., 2017). It solves the primitive equations using triangular surface meshes and a cell-vertex (quasi-B-grid) discretization, with tracers, sea surface height, and vertical velocities on mesh vertices, and horizontal velocities on triangle centroids. Details to the model setup, temporal discretization and the advection schemes can be found in Danilov et al. (2017). Here we use a simplified setup that does not include the sea ice model and uses a linear equation of state with temperature being the only advected scalar field.

The classical ocean momentum equation is given by

$$\partial_t \boldsymbol{u} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{u} + w\, \partial_z \boldsymbol{u} + f\, \boldsymbol{u}^{\perp} + \frac{1}{\rho_0} \boldsymbol{\nabla} P = \boldsymbol{V}(\boldsymbol{u}) + \partial_z (A_v\, \partial_z \boldsymbol{u}) \qquad (1)$$

with $\boldsymbol{u} = (u, v)$ the horizontal velocity field, w the vertical velocity, time t, vertical coordinate z, $\boldsymbol{\nabla} = (\partial_x, \partial_y)$ the horizontal gradient operator, Coriolis parameter f, $\boldsymbol{u}^{\perp} = (-v, u)$, reference water density ρ_0, pressure P, gravitational acceleration g, vertical viscosity coefficient A_v which is specified by the vertical mixing parametrization in FESOM2, and horizontal subgrid viscosity parametrization $\boldsymbol{V}$ to be discussed further. We use the usual quadratic frictional stress term as the bottom boundary condition where $A_v\, \partial_z \boldsymbol{u} = C_d\, |\boldsymbol{u}| \boldsymbol{u}$ with $C_d = 0.001$. No wind stress is applied at the surface. With regards to viscous stresses at vertical walls, free-slip boundary conditions were applied. Impermeability at the boundaries is always implied. Furthermore, for the biharmonic operator there is one more boundary condition: we set to zero contributions from boundary edges in Eq. B.6, which is equivalent to the assumption that the velocity derivative in the direction of the boundary $\partial \boldsymbol{u} / \partial \boldsymbol{n}$ is zero.

The flow is maintained by temperature relaxation near the boundaries via the temperature equation

$$\partial_t T + \boldsymbol{\nabla} \cdot (\boldsymbol{u} T) + \partial_z (w T) = \partial_z (K_v\, \partial_z T) + \gamma\, (T_{\mathrm{clim}} - T)\,. \qquad (2)$$

The density ρ is given by the linear equation of state

$$\rho = \rho_0\, (1 - \alpha\, (T - T_0))\,. \qquad (3)$$

Here, T denotes the temperature, T_0 its reference value, K_v the vertical diffusivity (provided by the vertical mixing parametrization), T_{clim} the given (climatological) temperature distribution, γ the relaxation parameter, different from zero only in narrow zones (see Section 2.3), and α is the thermal expansion coefficient in the linear equation of state. Horizontal diffusivity is provided by the transport scheme which combines 3rd order upwind (with weight 0.25) and 4th order centered (with weight 0.75) fluxes. Heat fluxes through all surfaces are assumed to be zero. The setup of the numerical experiments will be explained further in Section 2.3, together with the selection of γ and T_{clim}.

2.2. Backscatter scheme

The focus of this study is the subgrid closure for the momentum equation of the ocean model. Eq. 1 is augmented to include a negative viscosity term that acts as kinetic energy backscatter. The new momentum equation reads:

$$\partial_t \boldsymbol{u} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{u} + w\, \partial_z \boldsymbol{u} + f\, \boldsymbol{u}^{\perp} + \frac{1}{\rho_0} \boldsymbol{\nabla} P = \boldsymbol{V}(\boldsymbol{u}) + \boldsymbol{B}(\boldsymbol{u}) + \partial_z (A_v\, \partial_z \boldsymbol{u}) \quad (4)$$

with $\boldsymbol{B}(\boldsymbol{u})$ as the backscatter term. It should be emphasized here that the backscatter term $\boldsymbol{B}(\boldsymbol{u})$ is not designed to represent the physical inverse energy transfer from unresolved to resolved scales in the sense of an inverse cascade near the model grid's truncation scale. Instead, it is a numerically stable method to compensate for the unphysical forward energy transfer caused by the viscous closure $\boldsymbol{V}(\boldsymbol{u})$. This is achieved by reinjecting excessively dissipated kinetic energy on larger scales and making use of the already resolved inverse energy cascade to allow the resolved flow dynamics to simulate well-developed mesoscale turbulence (see also the discussion in Jansen et al. 2015). In the following, we will discuss the design of both the viscosity term $\boldsymbol{V}(\boldsymbol{u})$ and the backscatter term $\boldsymbol{B}(\boldsymbol{u})$.

2.2.1. Viscosity parametrization $\boldsymbol{V}(\boldsymbol{u})$

On triangular meshes, triangles with a common edge differ in the orientation relative to one another. For example, on regular triangular grids one can imagine that a central triangle is pointing northwards with one of its tips, while all three neighbouring triangles are pointing southwards with one of their tips. The neighbouring triangles are effectively rotated by 180° relative to the central triangle. This may lead to a grid mode (grid imprinting) for discrete quantities located on triangles (see also Danilov et al., 2017). Since the horizontal velocities of FESOM2 are defined as volume averages over triangular prisms, they may suffer from such a mode. In order to keep it under control, a special form of viscous dissipation operator is needed, capable of dissipating a grid pattern in velocity. FESOM2 is therefore run with approximate harmonic and biharmonic operators, detailed in Appendix B, which act strongly dissipatively on the grid mode and are thus capable of effectively suppressing it.

For simulations of turbulent flows we seek such a viscous operator that would provide minimum dissipation consistent with numerical stability. In this study we test two different vicosity schemes. In the following we will

only discuss the Reynolds type viscosity scheme that showed the best results as discussed in Section 3. The alternative Leith viscosity parametrization is detailed in Appendix A.

The Reynolds type viscosity uses the discrete biharmonic operator given by Eq. B.6 with a coefficient for each grid cell c

$$\nu_c^b = S_c\,\nu_c^R \quad \text{with} \quad \nu_c^R = \max\{u_0^R, |\boldsymbol{u}_c|\}\,\sqrt{S_c}/30\,. \tag{5}$$

Here, S_c is the area of the top of the prism c, $u_0^R = 0.2\,\mathrm{m/s}$ is a velocity scale selected experimentally, as well as the factor $1/30$, and $\boldsymbol{u}_c$ is the cell horizontal velocity. For large velocities, this form of dissipation is similar to that of the third-order upwind scheme, but with a reduced coefficient. We note that ν_c^R has the dimension of a harmonic viscosity; it keeps the cell Reynolds number constant throughout the simulation, hence the superscript R.

2.2.2. Energetic considerations of $\boldsymbol{V}(\boldsymbol{u})$

To know how much UKE is available for reinjection into the resolved flow, we need to know how much energy has been dissipated by the viscous operator. The energy dissipation rate due to horizontal viscosity at a given location in space $\boldsymbol{x}$ and time t in the continuum is given by

$$\mathcal{V} = \boldsymbol{u}\cdot\boldsymbol{V}(\boldsymbol{u})\,. \tag{6}$$

This expression is not sign definite and it only leads to consistent energy dissipation when averaged over the entire domain. The subgrid viscous force $\boldsymbol{V}$ is the divergence of the subgrid viscous stress tensor $\boldsymbol{\sigma}$, i.e. $\boldsymbol{V} = \boldsymbol{\nabla}\cdot\boldsymbol{\sigma}$. In the case of the classical harmonic viscosity, $\boldsymbol{\sigma} = 2\nu(\dot{\boldsymbol{\varepsilon}} - \operatorname{tr}\dot{\boldsymbol{\varepsilon}})$ with $\dot{\boldsymbol{\varepsilon}} = \frac{1}{2}(\boldsymbol{\nabla}\boldsymbol{u} + (\boldsymbol{\nabla}\boldsymbol{u})^T)$, so that the local rate of energy dissipation $\mathcal{V}$ takes the form

$$\mathcal{V} = \boldsymbol{u}\cdot(\boldsymbol{\nabla}\cdot\boldsymbol{\sigma}) = \boldsymbol{\nabla}\cdot(\boldsymbol{\sigma}\boldsymbol{u}) - 2\nu\,|\dot{\boldsymbol{\varepsilon}} - \operatorname{tr}\dot{\boldsymbol{\varepsilon}}|^2 = \boldsymbol{\nabla}\cdot(\boldsymbol{\sigma}\boldsymbol{u}) + \overline{\mathcal{V}}, \tag{7}$$

where $\overline{\mathcal{V}}$ is nonpositive. A similar relation holds for the biharmonic viscosity. $\mathcal{V}$ and $\overline{\mathcal{V}}$ will become equal if averaged over some area such that the flux contribution $\boldsymbol{\nabla}\cdot(\boldsymbol{\sigma}\boldsymbol{u})$ becomes small. Both forms of the dissipation rate, i.e. the full form $\mathcal{V}$ or the nonpositive form $\overline{\mathcal{V}}$, can be used in calculations of UKE available for backscatter, and their discrete analogs are given in Appendix B. The advantage of the original form $\mathcal{V}$ is that it contains the full viscous energy contribution. However, it contains grid-scale flux contributions and has to

be smoothed before it is applied to computations of UKE. Spatial smoothing is done by applying the filters defined in Appendix C.

Because of the properties of the discretization and the selection of viscous operators, dissipation related to mesoscale turbulence is accompanied by dissipation of the grid mode of the cell-vertex discretization. This is common for staggered discretizations on unstructured meshes and may lead to increased dissipation on the grid scale.

2.2.3. Backscatter parametrization $\boldsymbol{B}(\boldsymbol{u})$

Conceptually, the term $\boldsymbol{B}(\boldsymbol{u})$ in Eq. 4 reinjects the energy that is excessively dissipated by the viscous parametrization, but on a larger spatial scale and at a controlled rate over time. In fully-developed quasi-geostrophic turbulence, viscosity is thought to dissipate enstrophy near the truncation scale while energy dissipation is small. Hence, by reinjecting energy at larger scales, we shift the dynamics toward such a well-developed turbulent state. In this new equilibrium, excessive energy dissipation is substantially reduced. The system still dissipates a sufficient amount of enstrophy to prevent model instabilities.

Our backscatter operator takes the form of the approximate harmonic operator given by Eq. B.2, except that its coefficient has the opposite sign to the coefficient of a viscosity parametrization and that its magnitude is related to the amount of available UKE. In our model, the scalar UKE field $e = e(\boldsymbol{x}, z, t)$, defined on cells c, follows the prognostic equation

$$\partial_t e = -c_{\mathrm{dis}}\, \dot{E}_{\mathrm{dis}} - \dot{E}_{\mathrm{back}} - \boldsymbol{\nabla} \cdot (\nu^C\, \boldsymbol{\nabla} e)\,. \tag{8}$$

The first term on the right, $c_{\mathrm{dis}}\, \dot{E}_{\mathrm{dis}}$, is the resolved kinetic energy dissipation rate, the second describes the rate at which UKE is returned back to the resolved flow through backscatter, and the last is some UKE harmonic diffusion with coefficient $\nu_c^C = (S_c/S_r)^{1/2}\, 600\,\mathrm{m}^2/\mathrm{s}$ at a specific cell c which follows Jansen et al. (2015) but is scaled due to changes in grid resolution, with $S_r = 5.8 \cdot 10^9\,\mathrm{m}^2$ the area of a reference triangle. Since we apply smoothing to the first two terms, additional UKE diffusion is small and will not be discussed in detail.

Note that our UKE field varies in z as well, in contrast to Jansen et al. (2015) who applied vertical averaging to the UKE contributions. The same is true for the spatially varying viscosity coefficient used in this study. In our setup, this choice is physically more realistic as the observed rates of

dissipation and backscatter are about 1/3 smaller in the bottom layer than at the surface; see Section 3.3.2 below.

UKE advection is neglected since it is not clear whether UKE should simply be advected by the mean flow. Jansen et al. (2015) have shown that even without UKE advection the backscatter parametrization can achieve good results. Adding an advective term will increase computational costs, but more importantly will affect model stability. Dissipation of resolved kinetic energy by the spatial and temporal discretization schemes is also neglected here and may be considered in future studies. Furthermore, we do not dissipate UKE through an additional term in Eq. 8, in line with Jansen et al. (2015). Adding such a dissipative term could be seen as an alternative approach which we did not follow in this study. Instead, the fact that some energy dissipates is modeled by the coefficient c_{dis} which takes values between 0 and 1 and determines which fraction of the energy that is removed from the resolved flow enters the UKE budget. The coefficient c_{dis} varies in time and space and follows the design by Klöwer et al. (2018), where

$$c_{\mathrm{dis}}(\boldsymbol{u}) = \left(1 + \frac{R(\boldsymbol{u})}{R_{\mathrm{dis}}}\right)^{-1} \tag{9}$$

with $R(\boldsymbol{u}) = |D_r(\boldsymbol{u})|/f$ the local Rossby number, calculated for each cell by a discrete estimate of the local horizontal deformation rate

$$|D_r(\boldsymbol{u})| = \sqrt{(\partial_x u - \partial_y v)^2 + (\partial_y u + \partial_x v)^2}\,, \tag{10}$$

where x and y are local zonal and meridional directions, and $\boldsymbol{u} = (u, v)$. R_{dis} is a tunable scaling parameter that defines for which flow regime – distinguished by the local Rossby number – dissipated energy will enter the UKE. Large values, i.e. values greater than one, correspond to a more effective recycling of dissipated kinetic energy; see Klöwer et al. (2018) for a detailed discussion.

The first two terms in Eq. 8 can be calculated in two different ways. The first is to use the respective tendency contributions from viscosity and backscatter directly and take the dot product with the local value of $\boldsymbol{u}$. This yields the sign non-definite form, corresponding to the $\mathcal{V}$-form of Eq. 7.

However, for reasons of numerical stability, it is desirable to keep these contributions sign-definite, i.e., to have local UKE injection by the viscosity term, and local UKE removal by the backscatter term. This is achieved

by applying spatial filtering as described in Appendix C. $\mathcal{V}$ is not necessarily sign-definite, only its area mean is non-positive. Likewise, the local values of harmonic backscatter are not sign-definite; injection of energy is only achieved in an area-mean sense.

The second alternative is to use a sign definite form, i.e. the $\overline{\mathcal{V}}$-form, of the respective terms to ensure positivity of UKE as well as numerical stability. Sign-definite forms of their respective operators were also used by Jansen et al. (2015) and Klöwer et al. (2018). The corresponding sign-definite discrete forms for our scheme are detailed in Appendix B.1 and Appendix B.2.

In both cases, spatial filtering is used for the following reasons:

- The viscous dissipation rate may include contributions from a grid mode (cf. the discussion at the end of Section 2.2.1), and so does the backscatter term as it is based on the same differencing scheme. We do not want grid scales to be emphasized by the backscatter and carried into the subgrid energy equation. Spatial filtering of the tendencies will smooth it.

- Spatial filtering can be used for scale separation (see also Berloff, 2018). Smoothing of the backscatter term and the UKE will ensure that the backscatter scheme is active at different, larger scales than the corresponding (unfiltered) viscous operator. Some scale separation can also be achieved by using the less scale selective harmonic operator for the backscatter and the more scale selective biharmonic operator for the viscous parametrization (see Jansen et al., 2015).

- For the sign-indefinite form of the energy contributions, spatial filtering will reduce the relative amplitude of the flux term in Eq. 7. If filtering is large enough, the contributions from the flux term become small, making the remaining energy tendency almost sign definite, hence bringing $\mathcal{V}$ closer to $\overline{\mathcal{V}}$.

The spatial filter applied here is the one described in Eq. C.3. This filter has a spatial scale that depends on the grid, as it incorporates only information from neighboring grid points. Its advantage is a minimal halo size for exchange of information across processor boundaries, which is crucial for massively parallel implementations.

Thus, the first two terms in the UKE model, Eq. 8, for cell c take the

form

$$(c_{\mathrm{dis}}\,\dot{E}_{\mathrm{dis}})_c = (\mathsf{F}^{n_1})_{cc'}\,(c_{\mathrm{dis}}\,\mathcal{V})_{c'} \quad \text{with} \quad \mathcal{V}_c = (\boldsymbol{u}\cdot\boldsymbol{V}(\boldsymbol{u}))_c = (\boldsymbol{u}\cdot\mathsf{V}^b\boldsymbol{u})_c \tag{11a}$$

for the UKE source, and

$$(\dot{E}_{\mathrm{back}})_c = (\mathsf{F}^{n_2})_{cc'}\,\mathcal{B}_{c'} \quad \text{with} \quad \mathcal{B}_c = (\boldsymbol{u}\cdot\boldsymbol{B}(\boldsymbol{u}))_c = \boldsymbol{u}_c\cdot(\mathsf{F}^{n_3})_{cc'}\,(\mathsf{V}^B\boldsymbol{u})_{c'} \tag{11b}$$

for the UKE sink. F^n denotes n cycles of the smoothing filter Eq. C.3; n_1, n_2, and n_3 are the experimentally selected numbers of smoothing cycles. Summation over the repeated index c' is implied. These expressions are written out for the Reynolds viscosity scheme where V^b denotes the discrete biharmonic operator Eq. B.6 with coefficient given by Eq. 5 and V^B denotes the discrete harmonic backscatter operator Eq. B.2 with the coefficient scaled by the local UKE amplitude e_c

$$\nu_c^B = -c_0\,\sqrt{S_c}\,\sqrt{\max(2e_c, 0)} < 0\,. \tag{12}$$

The constant c_0 determines how quickly the dissipated energy is reinjected into the resolved flow, i.e. it controls the temporal response of the parametrization. The maximum function ensures that the backscatter is only active when there is a local reservoir of UKE. Physically, negative values of e_c correspond to a local depletion of UKE and can potentially occur because both the backscatter and the viscosity operators act on different scales. For the Leith viscosity scheme, V^b in Eq. 11a must be replaced by the sum of the harmonic operator V (Eq. B.2) with coefficient given by Eq. A.1 and the biharmonic (background) operator V^b (Eq. B.6) with coefficient given by Eq. A.2 (see Appendix A).

Finally, the actual discrete backscatter contribution to the momentum equation (Eq. 4) takes the form

$$\boldsymbol{B}_c(\boldsymbol{u}) = (\mathsf{F}^{n_3})_{cc'}\,(\mathsf{V}^B\boldsymbol{u})_{c'}\,. \tag{13}$$

Note that the smoothing filter applied in the computation of $\mathcal{B}_c$ in Eq. 11b must also be used in the backscatter tendency contribution $\boldsymbol{B}_c(\boldsymbol{u})$ to remove the grid mode from the operator.

Alternatively, we may replace $\mathcal{V}_c$ and $\mathcal{B}_c$ in Eq. 11 by their sign-definite alternatives, Eq. B.4 and B.7. We retain the smoothing filters to potentially

increase the scales on which the backscatter acts. The backscatter term in the momentum equation also retains its smoothing filter to prevent backscatter into the grid mode. The sign-definite formulation typically requires fewer smoothing cycles, i.e. smaller values of n_1, n_2, and n_3.

As a consequence of these different choices, we end up with four separate backscatter implementations, two with sign definite energy tendencies ($\overline{\mathcal{V}}$) and two with non-definite ones ($\mathcal{V}$), one of each for the Leith and one for the Reynolds viscosity schemes (Appendix A and Section 2.2.1, respectively). We will further discuss the effect of different filter cycle choices, the coefficients R_{dis} and c_0, and the choice of the backscatter implementation in Section 3.2.6 below.

2.3. Experimental setup

We simulate a zonally reentrant, baroclinically unstable flow in a channel with walls at 30°N and 45°N and spherical geometry. As for global model simulations, the grid is rotated with the rotated pole over Greenland, located at 40°W/75°N. The channel has 24 unevenly spaced vertical layers, with thicknesses ranging from 10 m at the surface to 100 m for the last few layers, going down to the constant depth of 1600 m. The flow is analyzed in two distinct sets of experimental setups that differ in their meshes and zonal extent, described below. For both sets of experiments, when varying the horizontal resolution, the vertical resolution is kept constant to enable a better comparison of the effect of increased horizontal resolution vs. the backscatter parametrization.

Both sets of experiments use the same temperature relaxation to force the mean flow. This takes place in 1.5° strips at the northern and the southern boundaries. The relaxation parameter γ in Eq. 2 takes the value 1/(3 days) directly at the boundary and is linearly reduced to zero at 43.5° and 31.5°, respectively. The reference temperature climatology T_{clim} to which the temperature near the boundaries is relaxed is the same as the temperature initial condition, which is linear with vertical gradient of $8 \times 10^{-3}\,\text{K/m}$ and meridional gradient of $-0.5 \times 10^{-5}\,\text{K/m}$. The surface temperature at the southern boundary is 25°C. Initial velocities are zero. A small temperature perturbation – the structure of which is not relevant for the discussion of the results – is added to the initial state to create small instabilities that quickly grow on a mean eastward flow. All runs reach a statistically steady state of kinetic energy within the first year.

The first set of experiments compares three equidistant meshes with resolutions of 1/4°, 1/6° and 1/12°, corresponding to triangle areas of about 270 km^2, 120 km^2 and 30 km^2. The zonal extent of the channel is 20°. The lower resolution runs were performed with and without backscatter, the high resolution run is used as a reference without backscatter. The Rossby radius is about 25 km, so that the 1/4° and 1/6° grids are eddy-permitting while the 1/12° grid is eddy-resolving. For the 1/4° setup different viscosity and backscatter schemes were tested in short 5-year integrations. The default backscatter and viscosity configurations – selected after carrying out the sensitivity studies of the 1/4° setup (see Section 3) – are extended to a total of ten years each to improve statistics, which are well equilibrated over this period. Timestep lengths for the 1/4°, 1/6°, and 1/12° grid configurations are 15 min, 10 min, and 1 min respectively. These choices are relatively conservative (i.e. small), especially for the 1/12° grid configuration; a moderate increase of timestep did not significantly affect the results (not shown).

The second set of experiments has a zonal extent of 40° to accommodate subdomains of different resolution. The horizontal grid is unstructured and has a high and a low resolution part with two transition regions at 7.5° and 32.5° longitude (see Fig. 1 of Danilov and Wang (2015), and later in this paper Fig. 8). The resolution varies from 1/12° in the high resolution domain to around 1/3° in the low resolution part, incorporating both eddy-permitting and eddy-resolving grids. Two simulations are carried out, with and without the default backscatter parametrization discussed in Section 3 for the equidistant grid. The general setup of the experimental domain corresponds to "Setup A" of Danilov and Wang (2015). Simulation length is 10 years with the timestep set to 1 min. As a high resolution reference for these simulations we use the same 1/12° equidistant mesh setup with zonal extent of 20° as in the first set of experiments.

Further details on the parameter choices of all experimental setups of this study – especially for the viscosity and backscatter parametrizations – can be found in the next Sections 3 and 4, and are summarized in Table 1.

2.4. Diagnostics

Diagnostics are based on daily mean data if not stated otherwise, which provides sufficient temporal resolution. The first year of each simulation is excluded as spinup unless stated otherwise. We focus on zonally averaged surface mean and temporal standard deviation fields, vertical, layer averaged mean and temporal standard deviation fields, and total integrated fields over

Label	$\Delta x\,[^\circ]$	$\Delta t\,[\mathrm{min}]$	**Viscosity**	$\nu_0\,[\frac{\mathrm{m}^2}{\mathrm{s}}]$	**B/S**	**S/D**	R_{dis}	c_0	**Filter**
$\mathbf{REF}_{1/12}$	$\frac{1}{12}$	1	Reynolds	–	No	–	–	–	–
$\mathbf{REF}_{1/6}$	$\frac{1}{6}$	10	Reynolds	–	No	–	–	–	–
$\mathbf{REF}_{1/4}$	$\frac{1}{4}$	15	Reynolds	–	No	–	–	–	–
$\mathbf{REF}_{1/4}^{L500}$	$\frac{1}{4}$	15	Leith	500	No	–	–	–	–
$\mathbf{REF}_{1/4}^{L2000}$	$\frac{1}{4}$	15	Leith	2000	No	–	–	–	–
$\mathbf{BACK}_{1/6}$	$\frac{1}{6}$	15	Reynolds	–	Yes	No	1	0.1	(2,2,4)
$\mathbf{BACK}_{1/4}$	$\frac{1}{4}$	15	Reynolds	–	Yes	No	1	0.1	(2,2,4)
$\mathbf{BACK}_{1/4}^{L2000}$	$\frac{1}{4}$	15	Leith	2000	Yes	No	1	0.1	(2,2,4)
$\mathbf{BACK}_{1/4}^{A}$	$\frac{1}{4}$	15	Reynolds	–	Yes	No	8	0.1	(2,2,4)
$\mathbf{BACK}_{1/4}^{B}$	$\frac{1}{4}$	15	Reynolds	–	Yes	No	100	0.1	(2,2,4)
$\mathbf{BACK}_{1/4}^{C}$	$\frac{1}{4}$	15	Reynolds	–	Yes	No	0.1	0.1	(2,2,4)
$\mathbf{BACK}_{1/4}^{D}$	$\frac{1}{4}$	15	Reynolds	–	Yes	No	1	0.4	(2,2,4)
$\mathbf{BACK}_{1/4}^{E}$	$\frac{1}{4}$	15	Reynolds	–	Yes	No	1	0.01	(2,2,4)
$\mathbf{BACK}_{1/4}^{F}$	$\frac{1}{4}$	15	Reynolds	–	Yes	No	1	0.1	(2,2,2)
$\mathbf{BACK}_{1/4}^{G}$	$\frac{1}{4}$	15	Reynolds	–	Yes	No	1	0.1	(1,1,4)
$\mathbf{BACK}_{1/4}^{H}$	$\frac{1}{4}$	15	Reynolds	–	Yes	Yes	1	0.1	(2,2,2)
$\mathbf{BACK}_{1/4}^{I}$	$\frac{1}{4}$	15	Reynolds	–	Yes	Yes	1	0.1	(1,1,2)
$\mathbf{BACK}_{1/4}^{J}$	$\frac{1}{4}$	15	Leith	2000	Yes	Yes	1	0.1	(2,2,2)
$\mathbf{REF}_{m}$	$\frac{1}{3}$ to $\frac{1}{12}$	1	Reynolds	–	No	–	–	–	–
$\mathbf{BACK}_{m}$	$\frac{1}{3}$ to $\frac{1}{12}$	1	Reynolds	–	Yes	No	1	0.1	(2,2,4)

Table 1: Table of simulation labels and configurations. Column Δx refers to resolution, Δt to the timestep, **Viscosity** to the viscosity scheme used, ν_0 to the coefficient of minimum viscosity when the Leith scheme is used (see Appendix A), **B/S** states whether backscatter is active or not, **S/D** states whether sign-definite operators are used for the UKE tendencies, R_{dis} and c_0 are parameters defined in Eq. 9 and 12, respectively, and **Filter** specifies the number of smoothing cycles (n_1, n_2, n_3) in Eq. 11 and Eq. 13. See text for details.

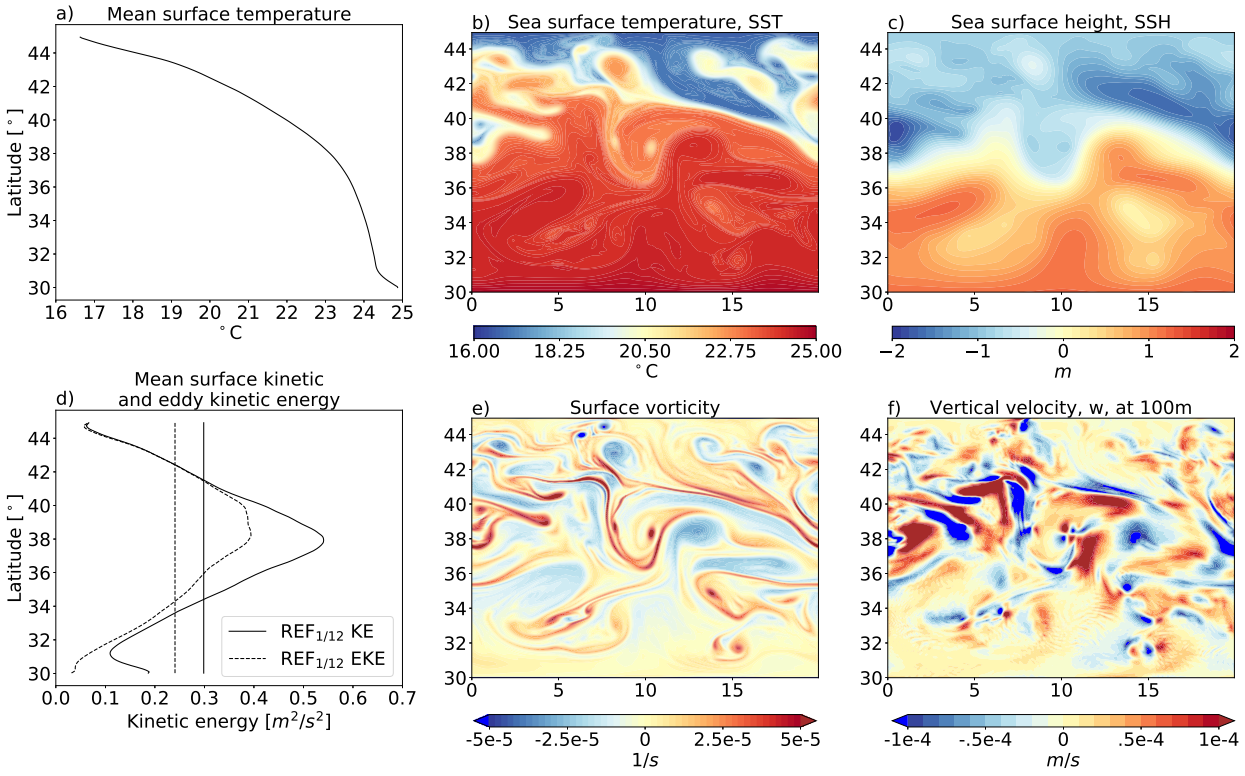

Figure 1: Reference simulation $\mathbf{REF}_{1/12}$. Left column: 9-year zonal averages; the vertical lines in (d) show total means. Color panels: daily mean snapshots, all taken at the same instance in time.

different vertical levels. Eddy kinetic energy (EKE) is here defined as the kinetic energy (KE) of the velocity fluctuations after removing the temporal model mean state.

3. Equidistant grid simulations

3.1. Reference simulations without backscatter

For the reference simulations, we use the discrete form of the Reynolds viscosity operator (Eq. B.2) with coefficient given by Eq. 5. In our experiments, this choice resulted in the least amount of spurious dissipation and the least amount of grid scale noise even on high-resolution unstructured grids. We will discuss the effect of the Leith compared to the Reynolds parametrization in Section 3.2.2 below.

The circulation at 1/12° resolution is characterized by a mean eastward flow with a mean KE peak between about 38°N and a trough at about 31°N and right at the northern boundary (Fig. 1d), where the troughs can be seen as artificially forced by the temperature boundary layer relaxation. The temperature relaxation forcing is also visible as a steep gradient in zonally averaged temperature close to the southern boundary (Fig. 1a). Snapshots of sea surface temperature SST, sea surface height SSH, surface vorticity, and vertical velocity w at $100\,m$ depth reveal that turbulence is strongest in the middle of the channel (Fig. 1, right two columns). Eddy size varies from larger to smaller going northward, in accordance with the increase of the Coriolis parameter. Vertical velocity (Fig. 1f) is the most sensitive field when it comes to grid scale noise. Some small wave-like structures are visible in w, but not in the other fields. Increased viscosity would damp these waves, but also reduce the simulated variability and provide a less competitive reference for the backscatter.

Comparing resolutions of 1/4°, 1/6°, and 1/12° with focus on zonal-temporal mean KE and EKE reveals that there is not only a strong resolution dependence of the amplitude of EKE, but also of its meridional location and of the feedback of eddies on the mean KE (Fig. 2, left two columns). While layer averaged KE and especially EKE amplitudes generally increase with increasing resolution, the resolution change also impacts the general flow structure: At 1/12°, the flow has only one major meridional peak, seen both in mean EKE and KE. At 1/4° and 1/6°, the flow has two distinct peaks at slightly different locations (around 32° and 40°N at 1/4° resolution and around 32°–34° and 40°–42°N at 1/6° resolution). Increasing resolution from 1/4° to 1/6° spreads out the two peaks, but the additional eddy forcing at 1/6° is still not strong enough to break the bimodal flow structure and create the unimodal flow seen at 1/12°.

In the vertical, an increase in horizontal resolution from 1/4° to 1/6° and finally to 1/12° leads to an intensification of KE (EKE) by a factor of around 1.5 (2) between the coarsest and the finest resolution, in all layers (Fig. 3). Similar differences are seen in the layer-averaged temporal standard deviations of temperature and w (Fig. 3, right two columns). We note that KE and EKE increase monotonically from bottom to surface. The standard deviation of temperature has peaks near bottom and surface, and a minimum at a depth of about 600–800 m. The opposite is true for variations in w, which are most intense at mid-depth and weakest near the surface and the bottom.

Total surface KE varies substantially between resolutions (Fig. 3, left

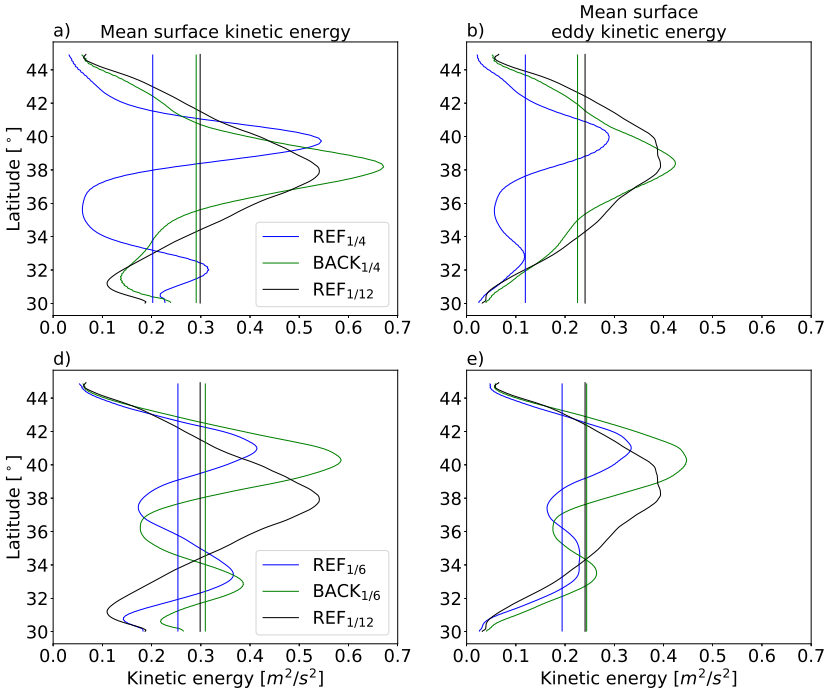

Figure 2: Surface KE diagnostics comparing low-resolution (blue), low-resolution with backscatter (green), and high-resolution reference (black). Top row: 1/4° vs. 1/12° resolution. Bottom row: 1/6° vs. 1/12° resolution. Shown are 9-year zonal means and surface layer averages (vertical lines).

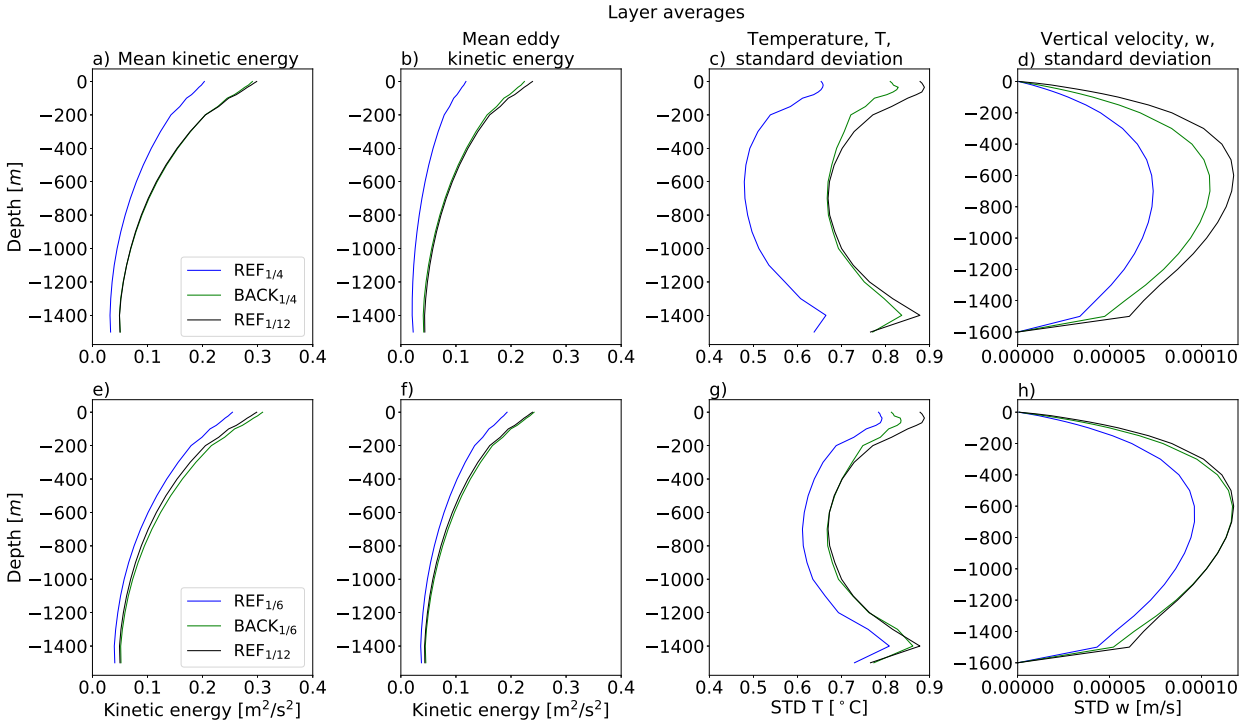

Figure 3: Vertical profile diagnostics comparing low-resolution (blue), low-resolution with backscatter (green), and high-resolution reference (black). Top row: 1/4° vs. 1/12° resolution. Bottom row: 1/6° vs. 1/12° resolution. Shown are 9-year layerwise averages.

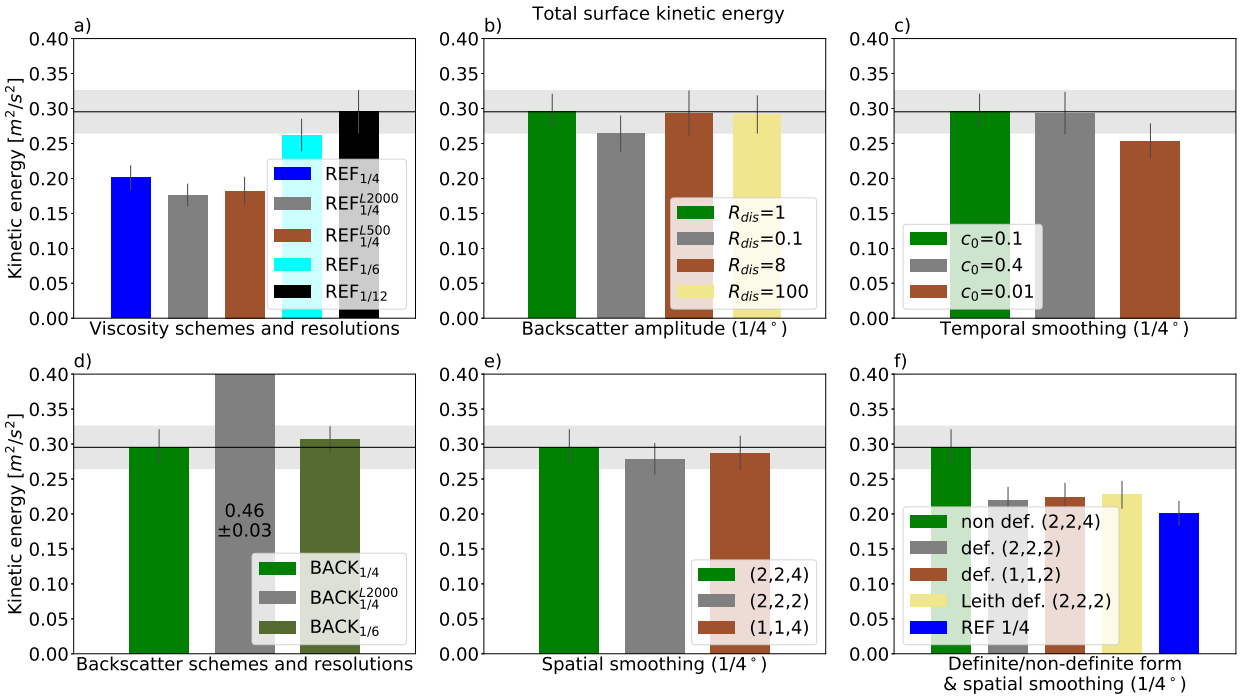

Figure 4: Summary of the sensitivity studies in terms of surface KE statistics. Bar heights are 9-year averages of total surface KE, vertical lines indicate a spread of one temporal standard deviation. The horizontal line marks the value for the reference **REF**$_{1/12}$, the shaded region the corresponding spread of one temporal standard deviation. Panel (a): different viscosity schemes and resolutions without backscatter. Panels (b)–(f): Default backscatter scheme (always in green) vs. variations in the indicated parameter. See Section 3.2 for details.

column, Fig. 4a): at 1/12° resolution, it is about 1.5 times larger than at 1/4° and about 1.2 times larger than at 1/6°. This change is accompanied by an increase of the temporal variability in the total KE (vertical lines in Fig. 4a).

3.2. Impact of backscatter

3.2.1. Sensitivity studies

To test the backscatter scheme, several parameter choices have to be made which we will briefly discuss. We study the effect of these choices in structured sensitivity studies. A summary of the simulations carried out for the sensitivity studies is given in Table 1.

3.2.2. Viscosity schemes

To provide baseline benchmarks, we carried out a set of simulations without backscatter, but with different viscosity schemes and different resolutions. The biharmonic Reynolds viscosity scheme is taken as described in Section 2.2.1. The harmonic Leith scheme with additional biharmonic background viscosity (see Appendix A) has a a tunable coefficient of minimum viscosity ν_0; we tested different values.

At 1/4° resolution, the Reynolds viscosity scheme shows the least depletion of total KE; the two Leith viscosity settings – with minimum viscosities of $\nu_0 = 500$ and $2000\,\mathrm{m^2\,s^{-1}}$ – perform slightly worse with respect to minimizing spurious dissipation while remaining numerically stable (Fig. 4a). Increased minimum viscosity tends to stabilize the Leith viscosity scheme, but also reduces total KE slightly (Fig. 4a). Consequently, we select the Reynolds viscosity as our default viscosity scheme.

3.2.3. Viscosity-backscatter combinations

Related to the choice of the viscosity is the backscatter implementation, as it uses the viscous energy dissipation to estimate the amount of UKE available for the backscatter. Even though the Leith scheme without backscatter tends to be more dissipative than the Reynolds scheme, in combination with backscatter the Leith scheme over-energizes the flow and produces far too much surface KE in the 1/4° setup when compared to the high resolution reference $\mathbf{REF}_{1/12}$. Consequently, the Reynolds scheme performs generally better with backscatter than the Leith scheme (Fig. 4d).

Still, the Leith scheme also equilibrates within the five-year run and does not spin up infinitely (Fig. 4d and 5c). Looking at the detailed daily averaged

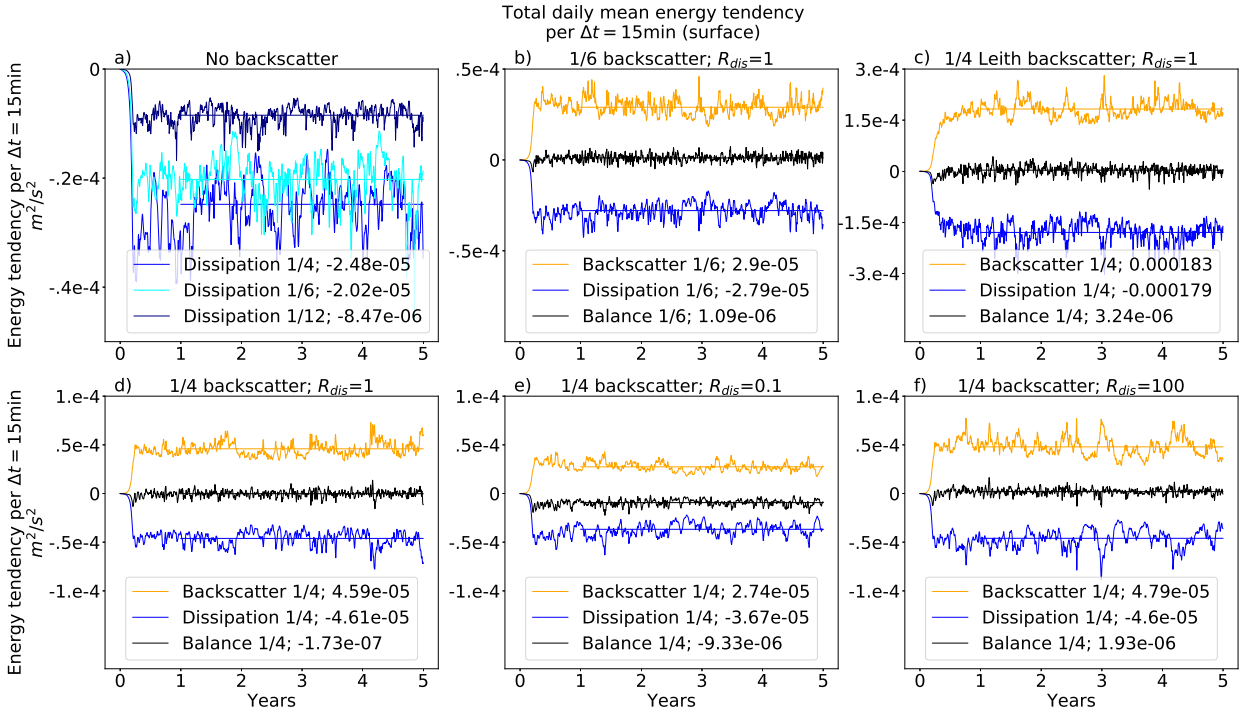

Figure 5: Time evolution of surface layer total dissipation and backscatter energy tendencies. Top left: different resolutions without backscatter. All others: different backscatter schemes; the black curve is the difference between backscatter tendency (orange) and viscous dissipation (blue). Horizontal lines denote the mean over years 2–5; numbers in the legends are 4-year mean values. Tendencies are the integrated contributions to the momentum equation for a time interval of 15 min (i.e. the timestep size of the low resolution setup). Notice the different y-axis limits.

energy tendency contributions, we see that in the Leith case backscatter tends to inject more energy than viscosity removes (Fig. 5c). Even though the excessive injection of energy is small at the surface, it is – relative to the rate of viscous dissipation – much larger in the subsurface and near the bottom.

While the backscatter equation (Eq. 8 and 12) is designed to balance energy removal and injection, e varies in time and space, and the balance is neither local nor instantaneous. The scheme may potentially inject more (or alternatively considerably less) energy than is globally removed at a particular instance of time. Furthermore, the inclusion of backscatter increases the spatially and temporally varying viscosity coefficient as well, as it energizes the flow which the viscosity scheme tries to counteract (compare amplitudes in Fig. 5a with those in Fig. 5b–f). Therefore, a new balance between viscosity and backscatter is created which varies in both components and for both time and space. This balance can potentially either lean towards too much dissipation or too much backscatter. For the Leith scheme, the latter is the case.

The backscatter is only directly controlling the injection of KE into the flow, but not the conversion of potential to kinetic energy. As the flow is fed by a temperature relaxation at the boundaries (Eq. 2), an excessively turbulent flow may modify eddy heat fluxes from an effectively infinite heat reservoir at the walls, increasing the release of potential energy. In the Leith backscatter setup, the standard deviations in w and temperature are much larger than in the $1/12^\circ$ reference (not shown). This might be a result of the spatially strongly varying Leith viscosity which also imprints onto the backscatter. As a result, the Reynolds scheme is chosen as the default backscatter scheme.

3.2.4. Backscatter parameter R_{dis} – backscatter amplitude

Changing R_{dis} has only minor effects on the total KE of the simulation (Fig. 4b). Low values of R_{dis} correspond to a reduced backscatter amplitude and show some effect by reducing total KE (Fig. 4b), although sensitivity to this parameter is considerably smaller than what is reported by Klöwer et al. (2018). One of the reasons is that our setup does not have a pronounced turbulent boundary layer as can be found near complex bottom topography, coast lines, or in the case of western boundary currents. We expect that this parameter will have more impact once we test the backscatter in global ocean simulations. Aside from the surface KE, another indicator for the

effect of this parameter is the balance of backscatter vs. viscosity energy tendencies (Fig. 5d–f). Decreasing this parameter leads to a balance that is more dissipative, i.e. favors the viscosity energy tendencies. Increasing this parameter leads to a balance that injects more energy. The equilibrium levels of viscous and backscatter energy tendencies change, with lower values of R_{dis} injecting less energy and consequently leading to reduced total viscous dissipation as the flow is slightly less turbulent (Fig. 5e).

3.2.5. Backscatter parameter c_0 – temporal smoothing

The temporal UKE smoothing controls the timescale of the energy release from the UKE reservoir and is governed by c_0. It shows an impact on KE only when the parameter is reduced substantially, i.e. when temporal smoothing is high (Fig. 4c). However, it has a considerable impact on the UKE equation (not shown), as it can prevent the occurrence of negative UKE values when c_0 is decreased. We found a good balance between a smooth, mostly positive UKE and high values of total KE with a parameter choice of 0.1, which is smaller than reported by Jansen et al. (2015), probably due to the fact that our implementation needs a stronger spatial and temporal smoothing for the UKE equation.

3.2.6. Backscatter parameters (n_1, n_2, n_3) – spatial smoothing

Spatial smoothing is applied to each of the following three terms: (i) The dissipated energy term which forces the UKE equation, (ii) the backscattered energy term which dissipates energy in the UKE equation, and (iii) the actual backscatter tendency contribution to the velocity field $\boldsymbol{u}$. The filter settings are described by a triple of integers (n_1, n_2, n_3) which specify the number of iterated filter cycles in (i), (ii), and (iii), respectively. Spatial smoothing does not affect the total surface KE much (Fig. 4e). It is, however, crucial for model stability and the prevention of negative UKE. Also, the backscatter tendency in the momentum equation needs sufficient smoothing to reinject energy at scales larger than the dominant scales of viscous dissipation. Especially higher resolution setups are sensitive to the number of smoothing cycles (not shown).

We found that $(n_1, n_2, n_3) = (2, 2, 4)$ is a good choice with regards to model stability and smallest amount of total energy dissipation for all tested resolutions. This suggests a sufficient scale separation of the backscatter and viscosity operators even for high-wavenumber-resolving simulations where the backscatter is small.

3.2.7. Sign definite vs. sign-indefinite formulation

We analyzed the effect of the sign-indefinite ($\mathcal{V}$) compared to the sign-definite ($\overline{\mathcal{V}}$) formulation of the energy tendency operators in the UKE equation (Fig. 4f). The sign-definite formulation remains numerically stable and is more energetic than the reference simulation without backscatter, $\mathbf{REF}_{1/4}$. It also does not need as many spatial smoothing cycles or potentially as low a value for c_0 to keep the model numerically stable compared to the sign-indefinite version. Furthermore, the Leith sign-definite version performs just as well as the Reynolds sign-definite version, without the excessive injection of energy seen for the sign-indefinite Leith backscatter scheme. This points to the flux components of the viscous dissipation rate $\mathcal{V}$ (i.e. the first term on the right of Eq. 7) as the main cause of excessive energy injection for the sign-indefinite Leith scheme.

However, all sign-definite versions are clearly outperformed by the sign-indefinite Reynolds scheme in terms of improved surface KE. This observation indicates that the flux components in Eq. 7 are relevant and necessary contributions for a good performance of the backscatter scheme in this setup.

3.2.8. Decision on the default scheme

Based on the sensitivity studies, we choose the following *default scheme*: Viscosity is implemented by the biharmonic Reynolds scheme, the energy source and sink terms to the UKE equation are chosen in their sign-indefinite form with parameter $R_{\text{dis}} = 1$, the temporal smoothing parameter is taken to be $c_0 = 0.1$, and the number of filtering cycles is $(2, 2, 4)$ for smoothing the source term in the UKE model, the sink term in the UKE model, and the backscatter forcing in the momentum equation, respectively. The two default backscatter setups for the two lower resolutions are $\mathbf{BACK}_{1/4}$ and $\mathbf{BACK}_{1/6}$ (see Table 1). Both setups reach similar levels of surface KE as the reference $\mathbf{REF}_{1/12}$ (Fig. 4d).

3.3. Impact of the default backscatter scheme

3.3.1. Simulated flow

Hereafter, we give a detailed assessment of the flow which is simulated with the default backscatter setup. Total KE in the lower-resolution runs with backscatter is restored to a level comparable to the level at much higher, eddy-resolving resolution. We note that backscatter actually increases the rate of viscous dissipation. However, this loss is balanced by the backscatter energy injection (Fig. 5b and 5d) up to a level where the balance between

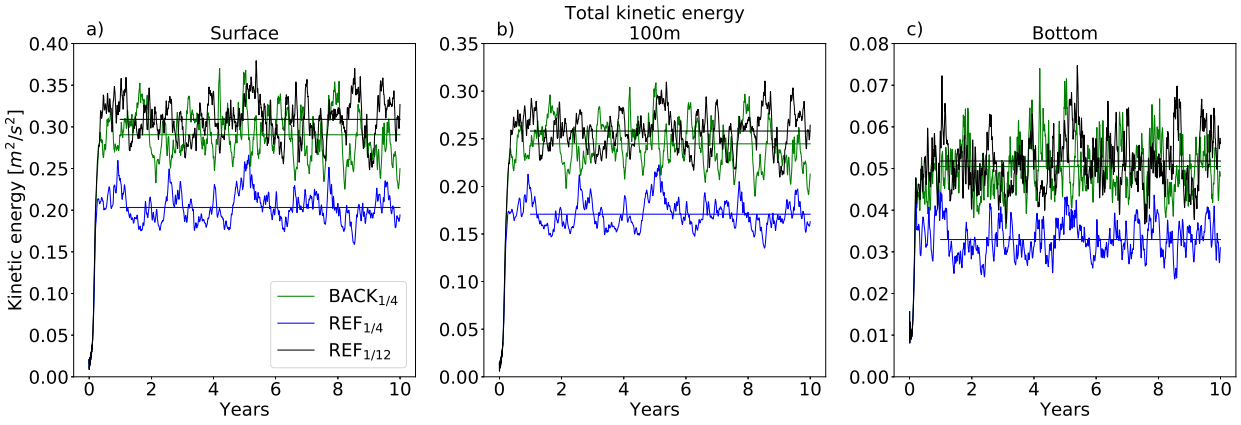

Figure 6: Comparison of low-resolution (blue), low-resolution with backscatter (green), and high-resolution (black) reference runs: 10 year time series of total surface KE at different depths, for $\mathbf{REF}_{1/12}$, $\mathbf{REF}_{1/4}$, and $\mathbf{BACK}_{1/4}$. Horizontal lines denote the mean over the last 9 years of the respective simulation.

the two is closer to total KE conservation than the rate of viscous energy loss in the high resolution 1/12° reference (compare to Fig. 5a). The improved energy balance is also visible in the total KE of the default scheme at 1/4° resolution at different model layer depths (Fig. 6a, 6b, and 6c). KE is increased by about 50% in all layers to a level comparable to $\mathbf{REF}_{1/12}$.

In the surface layer energy spectra, backscatter leads to an increased energy across the entire spectrum near the northern and southern boundaries and in the center of the channel, both for the 1/4° and 1/6° simulations (not shown).However, the increase in spectral power through backscatter does not necessarily rectify the structure of the mean flow at lower resolution to the structure observed in $\mathbf{REF}_{1/12}$. The lower resolution runs without backscatter, $\mathbf{REF}_{1/4}$ and $\mathbf{REF}_{1/6}$, have a bimodal zonal mean flow, but $\mathbf{REF}_{1/12}$ is unimodal. Even though the bimodal structure collapses to a unimodal one in $\mathbf{BACK}_{1/4}$, the location and width of the mean flow peak is not quite the same as in $\mathbf{REF}_{1/12}$ (Fig. 2a and b). In $\mathbf{BACK}_{1/6}$, the flow stays bimodal, although the jets are getting closer (Fig. 2d and e). As a consequence, line spectra at different locations do not always match the high resolution reference (not shown). Total mean KE and EKE in the different layers, on the other hand, are similar between the backscatter simulations and the high resolution reference (Fig. 2 and Fig. 4d). The question to which degree backscatter schemes are capable of rectifying the spatial structure and

strength of the mean flow will need to be addressed in follow-up studies.

Aside from KE and EKE, backscatter also considerably improves other layer mean quantities such as the standard deviation of w and of T (Fig. 3). While the improvement is not quite strong enough to totally alleviate errors in these fields in $\mathbf{BACK}_{1/4}$, the results in $\mathbf{BACK}_{1/6}$ show a remarkably good match with $\mathbf{REF}_{1/12}$. This is especially noteworthy since the backscatter scheme injects KE only into the horizontal momentum equation (Eq. 4); it does not directly affect the variability of T or w. Nonetheless, the improved representation of eddies in the system leads to a corresponding improvement in the variability of other fields as well. On the other hand, the horizontal mean flow biases which are still present in the backscatter simulations are also visible in these other flow variables as they are affected by the position of the mean current (not shown). However, many flow properties can be represented on a $1/4°$ mesh with backscatter nearly as well as on a $1/12°$ mesh without backscatter. The additional computational costs for backscatter are around 50% in this setup with only one tracer equation and including all the high frequency output. This is a very small amount of additional costs compared to those resulting from a resolution increase from $1/4°$ to $1/12°$. The $1/12°$ is at least 30 times more expensive in theory, and is often even more expensive in practice.

3.3.2. Unresolved kinetic energy details

The UKE of the backscatter scheme nicely scales with resolution as expected (Fig. 7b). Although the general horizontal structure of the subgrid energy is quite similar between the resolutions, the difference between the northern and southern part of the domain is more pronounced in $\mathbf{BACK}_{1/4}$ than in $\mathbf{BACK}_{1/6}$. This might be one of the reasons why in $\mathbf{BACK}_{1/4}$ the mean flow representation looks better than in $\mathbf{BACK}_{1/6}$. Another noteworthy aspect is that UKE drops to about one third its maximum value from surface to bottom (Fig. 7a), justifying the use of a vertically variable UKE.

As designed, the backscatter parametrization tends to act on larger scales than the viscosity parametrization (compare Fig. 7h and Fig. 7i). The smoothing of the two energy tendency contributions to the UKE equation (Fig. 7e,f) ensures that the UKE field is sufficiently smooth and mostly positive at each timestep (Fig. 7d) and positive when layer averaged (Fig. 7a). Both contributions are large at similar locations, following the flow structure but smoothed in time and space. The energy contribution from the backscatter term (Eq. 13) to the momentum equation has a very similar

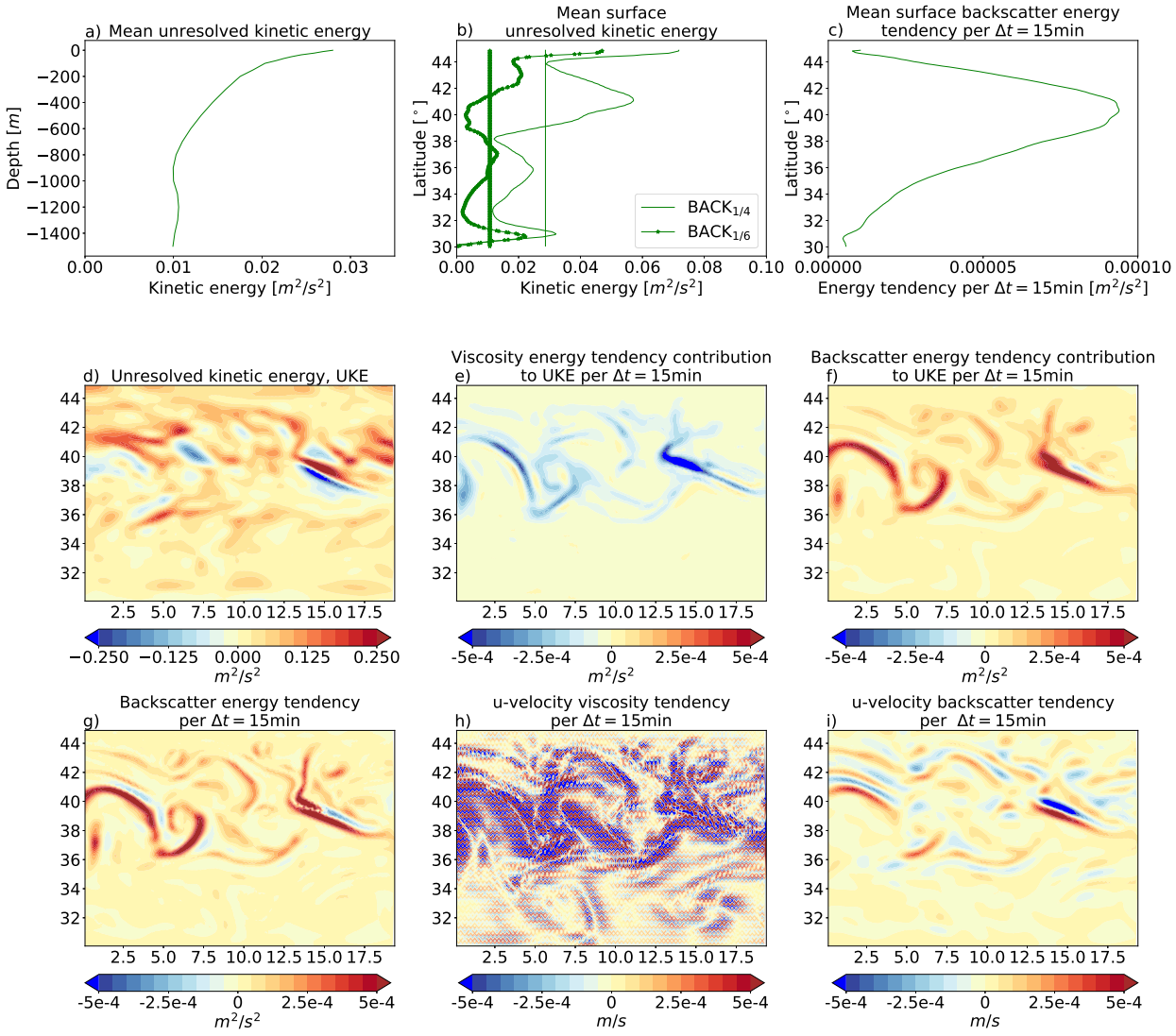

Figure 7: Unresolved kinetic energy and its tendency contributions for $\mathbf{BACK}_{1/4}$. Top row: 9-year (a) layer averaged UKE, (b) zonally averaged surface UKE and (c) backscatter energy tendency. (b) also includes UKE for $\mathbf{BACK}_{1/6}$. Bottom two rows: Daily mean surface snapshots taken at the same instance in time. Tendencies are the integrated contributions to (e and f) the UKE equation or (h and i) the momentum equation, and (c and g) the discrete backscatter energy tendencies, for a time interval of 15 min (i.e. the timestep size of the low resolution setup). Note that the grid mode is visible in the viscosity tendency (h).

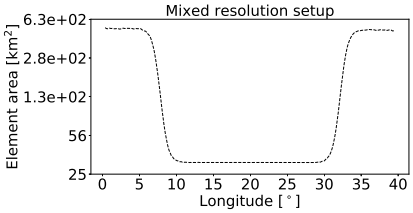

Figure 8: (Meridionally averaged) Resolution of the mixed mesh simulation, given in element area [km^2].

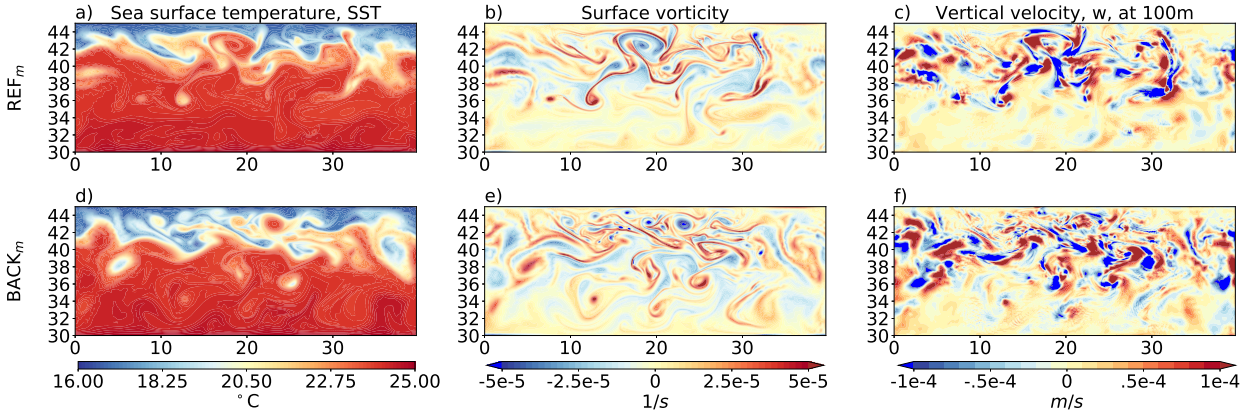

Figure 9: Daily mean snapshots of (top row) $\mathbf{REF}_m$ and (bottom row) $\mathbf{BACK}_m$ for (left column) surface temperature, (middle column) surface vorticity and (right column) vertical velocity at 100 m depth. All snapshots are taken at the same instance in time.

form as the sink term in the UKE equation (Eq. 11b), with the exception of being spatially slightly more distinct in structure (compare Fig. 7f to 7g).

The actual energy injection by the backscatter has a very similar unimodal structure at the surface compared to the surface total KE and EKE of $\mathbf{REF}_{1/12}$ (Fig. 7c). Only the peak of the energy injection is slightly more northward (around 40°N–41°N) compared to the reference simulation KE peak (around 38°N) and the EKE peak (around 38°N–40°N). The zonally averaged backscatter energy tendency of the momentum equation also looks much smoother in space than the actual UKE (compare Fig. 7b to 7c).

4. Variable grid simulations

We test the effect of backscatter on a grid with zonally varying resolution. The zonal extent is increased to 40° where about half of the domain is

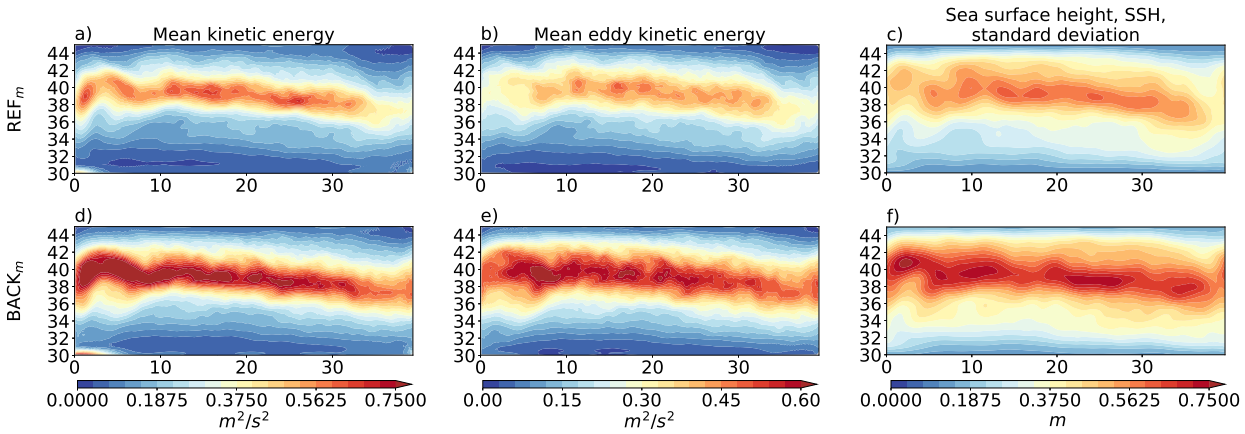

Figure 10: 9 year averaged mean state of (top row) $\mathbf{REF}_m$ and (bottom row) $\mathbf{BACK}_m$ for (left column) mean surface KE, (middle column) mean surface EKE, and (right column) standard deviation of sea surface height.

eddy-permitting while the other half is eddy-resolving, with transition zones between the two grid resolutions (Fig. 8). These transition zones are relatively sharp in this setup and therefore present a somewhat extreme testcase. As a high resolution reference for the mixed resolution simulations we keep the equidistant grid $\mathbf{REF}_{1/12}$ simulation with zonal extent of 20° used in the previous section.

4.1. Reference simulation without backscatter

In the reference simulation without backscatter, $\mathbf{REF}_m$, an imprint of the local resolution is clearly visible in the daily means of all fields (Fig. 9a,b,c). The high resolution part of the domain is also visible in the temporal means of, for example, surface EKE (Fig. 10b) and in the standard deviation of sea surface height (Fig. 10c). Further imprints of the resolution change can also be found in the vertical structure of mean EKE (Fig. 11a), and standard deviation of temperature (Fig. 11b) and w (Fig. 11c). Once the flow enters the high resolution part of the domain, all these fields become more intense and reach deeper down. Once they leave the high resolution part, they return back to much smaller values. For a detailed discussion of these observations, see Danilov and Wang (2015).

The mean flow shows an artifact of the resolution change, visible for example in the mean KE (Fig. 10a) in the low resolution region just upstream

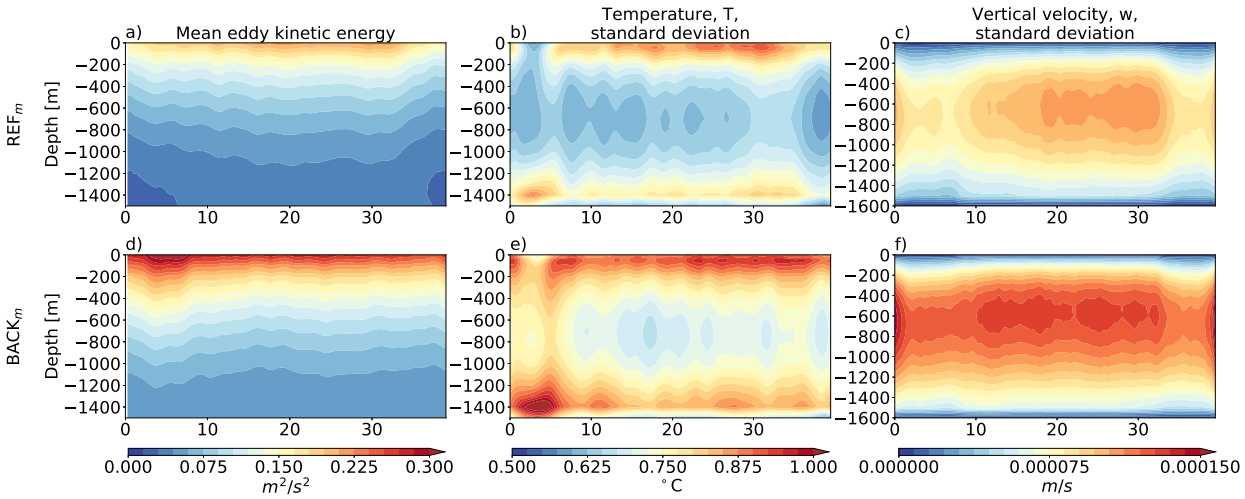

Figure 11: 9 year averaged mean state of (top row) $\mathbf{REF}_m$ and (bottom row) $\mathbf{BACK}_m$ for layer averaged (left column) mean KE and standard deviation of (middle column) temperature and (right column) vertical velocity.

of the high resolution part (around 0°–5° longitude): The flow shows a wave-like excursion as if confronted with an obstacle. Further, the length of the high resolution domain in this grid configuration does not seem sufficient to saturate the vertical structure of, for example, EKE, since there is still a zonal slope in this field at the downstream end of the high resolution domain (Fig. 11a). This suggests that equilibrium amplitudes, i.e. those corresponding to a purely high resolution simulation, of this and other fields are not reached.

Compared to $\mathbf{REF}_{1/12}$, $\mathbf{REF}_m$ underestimates the amplitudes of the flow fields when looking at zonal averages (Fig. 12a,b,c,d) and the vertical structure of layer averages (Fig. 12e,f,g,h). Furthermore, even though $\mathbf{REF}_m$ does reproduce the unimodal structure of the mean flow, the peak in the mean flow is slightly too far north. This is caused by the northward excursion of the spurious wave pattern described above, from which the flow cannot fully recover even in the high resolution domain.

4.2. Impact of default backscatter setup

4.2.1. Simulated flow

The effect of backscatter in the mixed resolution setup, $\mathbf{BACK}_m$, is very similar to the effect in the equidistant mesh configurations. Generally, KE

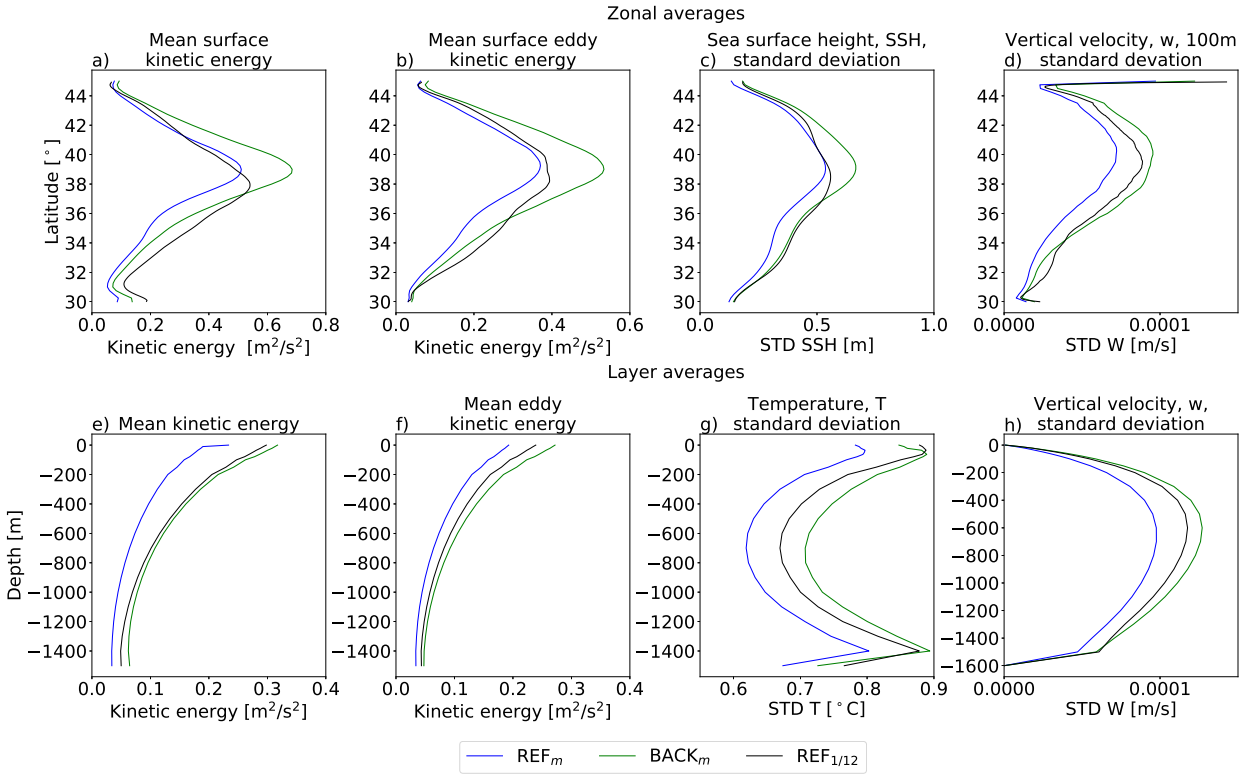

Figure 12: 9 year (top row) zonally averaged surface mean states and (bottom row) layer averaged mean states of the mixed resolution simulations $\mathbf{REF}_m$ (blue) and $\mathbf{BACK}_m$ (green), and the 1/12° reference 20° periodicity simulation $\mathbf{REF}_{1/12}$ (black), for (left column) mean KE and (second column from left) EKE, and standard deviation (c) of sea surface height, (d) of vertical velocity at 100 m depth, (g) of temperature, and (h) of vertical velocity.

and EKE are considerably increased, especially in the low resolution part of the channel (Fig. 10d,e). The pattern of KE, and also of T and SSH variability is much more zonally uniform for $\mathbf{BACK}_m$ (Fig. 10d,e,f), reducing the visible imprint of the resolution change on these fields. This is also true for the vertical structure of the flow, i.e. meridionally averaged KE and EKE, and meridionally averaged standard deviation of T and w (Fig. 11d,e,f). Daily means of w at 100 m depth, SST and surface vorticity are also much more uniform (Fig. 9d,e,f) to the point where the visible imprint of the change of resolution is substantially reduced.

Backscatter does, however, not reduces the wave-like disturbance upstream of the high resolution subdomain (Fig 10a) – if anything, it reenergizes this flow anomaly. As the structure is already present in $\mathbf{REF}_m$, the inclusion of backscatter just strengthens it rather than removing it from the flow. Backscatter-enhanced imprints of the anomaly are also visible in the vertical flow structure for layer averaged EKE and for the variability of T (Fig. 11d,e).

$\mathbf{BACK}_m$ is slightly more energetic in terms of total KE and slightly more variable in fields such as T, SSH and w when compared to $\mathbf{REF}_{1/12}$ (see the zonal surface layer averages and the vertical structure of total layer averages in Fig. 12). The reason is that $\mathbf{BACK}_m$ also allows for backscatter in the part of the domain which is resolved at $1/12°$, while $\mathbf{REF}_{1/12}$ does not include any backscatter. In general, backscatter reenergizes the flow not only at the surface, but at all vertical layers. The typical increase of layer KE is about 30% between $\mathbf{REF}_m$ and $\mathbf{BACK}_m$, and brings the backscatter simulation close to $\mathbf{REF}_{1/12}$ (Fig. 12e).

4.2.2. Unresolved kinetic energy details

Backscatter is especially active in regions with low resolution, hence strong dissipation (Fig. 13d). In the high resolution part of the domain, backscatter is more active in the northern central part of the domain, where the mean flow is strong and eddy scales are small, therefore even pushing the limits of an eddy-resolving simulation at $1/12°$ resolution (Fig. 13d). The same pattern is visible in daily mean snapshots of UKE (Fig. 13c). Interestingly, time-averaged UKE actually exhibits slightly negative values near the surface in the middle of the channel in the high resolution part of the domain (dark blue areas in Fig. 13a,b). While backscatter is sometimes active to energize the flow in this region (Fig. 13c,d), there are substantial periods where the backscatter scheme is switched off due to local depletion

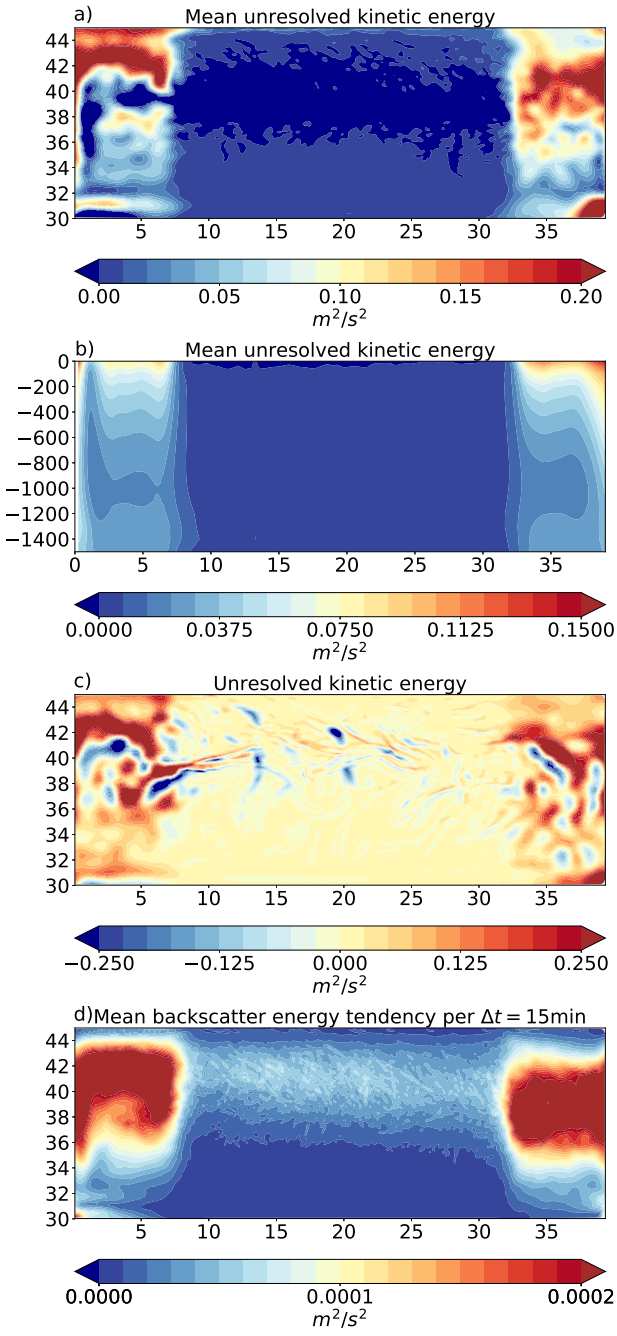

Figure 13: Backscatter fields for $\mathbf{BACK}_m$, with 9 year averaged mean state of the (a) surface UKE, (b) vertical structure of meridionally averaged UKE, and (d) surface backscatter energy tendency, as well as (c) a daily mean snapshot of surface UKE. Tendencies are the integrated contributions for a time interval of 15 min (to make it comparable to Fig. 7). The snapshot is taken at the same instance in time compared to Fig. 9.

of UKE. The 1/12° resolution is already so high that backscatter is inactive more often than not. Backscatter is also active in the high resolution part of the domain because (a) the mean flow transports under-resolved eddies from the low-resolution part into the high resolution part, and (b) the mean flow distortion is rectified when the flow enters the high-resolution region.

The vertical UKE structure shows a slight increase of UKE when moving from mid-depth of about 1000 m further down towards the bottom. As discussed above, the flow reaches higher energy levels and flow variability once it enters the high resolution region, even though it does not saturate before it returns to the low resolution region. For mean EKE this is especially pronounced at depth (see Fig. 11a for $\mathbf{REF}_m$). This in turn affects local dissipation. Consequently, UKE is affected in $\mathbf{BACK}_m$ as well. Such visible variations in UKE with depth highlight the necessity to maintain a three-dimensional UKE field to capture the varying dissipation rates between vertical layers.

5. Discussion and outlook

We developed an energy backscatter parametrization based on previous studies by Jansen et al. (2015) and Klöwer et al. (2018). The parametrization is energetically more consistent and less dissipative than classical eddy-viscosity parametrizations. At the same time, it maintains sufficient dissipation of enstrophy to preserve numerical stability. Our sensitivity studies showed that energy backscatter is able to balance the viscosity energy tendencies and that it reinjects the right amount of energy back into the flow, on larger scales, without violating model stability. The nonlinear upscale energy transfer of the resolved flow then distributes the reinjected energy in a physically realistic way.

An important aspect both for the theoretical design of backscatter and for the observed numerical stability of the scheme is the necessity of temporal and spatial smoothing. Temporal smoothing simulates the delay of UKE release back into the mean flow; spatial smoothing is necessary to release the energy on larger scales than where it is dissipated. Certainly, the existence of a grid mode in our finite volume discretization adds to the amount of smoothing cycles that are needed to keep the model numerically stable and the unresolved energy smooth and mostly positive.

In the reentrant channel setup of this study the low resolution simulations without backscatter were far too dissipative, thus deficient in the level of to-

tal kinetic energy and in the level of flow variability when compared to an eddy-resolving 1/12° simulation. Without backscatter, the 1/4° simulation had only about 60–70% of the reference kinetic energy, the 1/6° simulation had an energy defect of around 20%. With backscatter, these energy deficits are entirely rectified; moreover, the variability of, for example, temperature and of vertical velocity improved to the point where they are nearly indistinguishable from the reference. While characteristics of the mean flow are also improved, the backscatter parametrization cannot entirely rectify a mean flow bias.

The next, more challenging step following this idealized channel setup will be the implementation of backscatter in a global ocean configuration. We expect that some adjustments will be necessary. First, the current test case has no complex bottom topography. Along complex ocean floor structures, physical dissipation is a relevant KE sink, so that backscatter would need to be reduced or switched off entirely to prevent reinjection of the dissipated energy. Backscatter might also lead to numerical instabilities if it is too strong. Similar considerations apply to complex boundaries, e.g., along to the western flanks of the strong western boundary currents. The parameter R_{dis} might take care of such backscatter damping in a physically consistent way as reported by Klöwer et al. (2018). But it might become necessary to adapt R_{dis} to the basin geometry or to otherwise redesign the part of the parametrization that defines how much dissipated energy is actually entering the UKE regime and is available for backscatter.

Second, changes in resolution on unstructured grids might still become a problem. Although we showed that the scheme works well for rapidly changing grid resolution, a more complex structure of the transition zones might still lead to issues that are not covered by our current study. While UKE should scale with resolution, issues might still arise for very high resolution parts of a global mesh configuration, where backscatter is no longer desired at all.

Future studies will have to address issues regarding model stability brought about by the coupling with the UKE equation. While the highly spatially and temporally varying viscosity coefficient employed in this study might be a better subgrid eddy closure than the one varying only according to the grid resolution – which has previously been applied in the context of kinetic energy backscatter – it also is more difficult to combine with the backscatter scheme. This is evidenced by the overenergizing of the mean flow when combining backscatter with the harmonic Leith and biharmonic background

viscosity. Adding vertical layers to the UKE field requires extra effort, but was found necessary for reasons of stability and model performance. Furthermore, there is a strong interplay between the details of the numerical scheme and the details of the backscatter closure. Different schemes will raise additional questions that need to be solved more or less individually. Smoothing in this context is a powerful tool to distinguish between the large scale flow and eddy and backscatter scales, as also discussed by Berloff (2018).

One aspect that is certainly worth further investigation is the question with what kind of spatial structure the energy should be backscattered. While we need to have a scale separation between dissipation and backscatter, it is not clear what the exact structure of the respective operators or the related smoothing should be. It seems sensible that energy should be backscattered close to where it is dissipated, unless UKE advection is significant, which is currently not clear. Generally, energy should mostly be injected where the model is dynamically active or should be active compared to higher resolutions; the latter is not necessarily predictable.

Alternative approaches to ocean backscatter also include stochastic implementations (Jansen and Held, 2014; Mana and Zanna, 2014; Cooper and Zanna, 2015; Cooper, 2017) where the unresolved eddy effects are simulated by stochastic patterns. A combination of a stochastic and the here presented deterministic approach are feasible, where the deterministic part tracks the backscatter amplitude and approximate location, while the stochastic patterns distribute the energy locally in a random fashion. Furthermore, certain parameters as part of the backscatter scheme are related to high uncertainties that could be tackled by stochastic components, following implementations by, for example, Juricke et al. (2017, 2018). An example in this context would be the parameter R_{dis} which – when chosen in a specific regime – does not seem to affect model stability very much but would affect the total amount of backscattered energy.

Additionally, it is of interest to study how backscatter affects higher order eddy statistics such as eddy size and life time. These properties are not directly included in the backscatter scheme, but are certainly affected by the backscatter forcing of the flow. Furthermore, the impact of backscatter on the mean flow deserves much closer attention.

Finally, there are other aspects of the energy cycle that are not addressed by our backscatter parametrization. One example is numerical energy dissipation through time stepping and horizontal discretization. This loss of kinetic energy is not captured by the UKE equation. Other examples are the

generally underestimated growth rates of baroclinic instabilities (Barham et al., 2018) and the conversion of potential to kinetic energy, which are also affected by the resolution and the capability of the model to fully resolve mesoscale and even submesoscale structures. With backscatter, the conversion might be improved by increased EKE, but we will have to investigate the full energy cycle including potential energy conversion in future studies with more realistic forcing.

The current results suggest that there is potential for savings in terms of computing cost of more than one order of magnitude without degrading the level of mean kinetic and eddy kinetic energy on a lower resolution grid.

Acknowledgments

This paper is a contribution to the projects M3 and S2 of the Collaborative Research Centre TRR 181 "Energy Transfer in Atmosphere and Ocean" funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under project number 274762653. The computational resources were supplied by the supercomputing facilities at the Alfred Wegener Institute in Bremerhaven.

References

Barham, W., Bachman, S., and Grooms, I. (2018). Some effects of horizontal discretization on linear baroclinic and symmetric instabilities. *Ocean Modelling*, 125:106 – 116.

Berloff, P. (2018). Dynamically consistent parameterization of mesoscale eddies. Part III: Deterministic approach. *Ocean Model.*, 127:1–15.

Berner, J., Shutts, G., Leutbecher, M., and Palmer, T. (2009). A spectral stochastic kinetic energy backscatter scheme and its impact on flow-dependent predictability in the ECMWF ensemble prediction system. *J. Atmos. Sci.*, 66:603–626.

Bryan, F. O., Tomas, R., Dennis, J. M., Chelton, D. B., Loeb, N. G., and McClean, J. L. (2010). Frontal scale air–sea interaction in high-resolution coupled climate models. *J. Climate*, 23(23):6277–6291.

Cooper, F. C. (2017). Optimisation of an idealised primitive equation ocean model using stochastic parameterization. *Ocean Model.*, 113:187–200.

Cooper, F. C. and Zanna, L. (2015). Optimisation of an idealised ocean model, stochastic parameterisation of sub-grid eddies. *Ocean Model.*, 88:38–53.

Danilov, S., Juricke, S., Kutsenko, A., and Oliver, M. (2019). Toward consistent subgrid momentum closures in ocean models. In Eden, C. and Iske, A., editors, *Energy Transfers in Atmosphere and Ocean*, pages 145–192. Springer-Verlag.

Danilov, S., Kivman, G., and Schröter, J. (2004). A finite-element ocean model: principles and evaluation. *Ocean Model.*, 6:125–150.

Danilov, S., Sidorenko, D., Wang, Q., and Jung, T. (2017). The Finite-volumE Sea ice–Ocean Model (FESOM2). *Geosci. Model Dev.*, 10(2):765–789.

Danilov, S. and Wang, Q. (2015). Resolving eddies by local mesh refinement. *Ocean Model.*, 93:75–83.

Fox-Kemper, B. and Menemenlis, D. (2008). Can large eddy simulation techniques improve mesoscale rich ocean models? In Hecht, M. W. and Hasumi, H., editors, *Ocean Modeling in an Eddying Regime*, pages 319–337. American Geophysical Union.

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

Frenger, I., Gruber, N., Knutti, R., and Münnich, M. (2016). Imprint of Southern Ocean eddies on winds, clouds and rainfall. *Nat. Geosci.*, 6:608–612.

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

Griffies, S. M., Winton, M., Anderson, W. G., Benson, R., Delworth, T. L., Dufour, C. O., Dunne, J. P., Goddard, P., Morrison, A. K., Rosati, A., Wittenberg, A. T., Yin, J., and Zhang, R. (2015). Impacts on ocean heat from transient mesoscale eddies in a hierarchy of climate models. *J. Climate*, 28(3):952–977.

Haarsma, R. J., Roberts, M. J., Vidale, P. L., Senior, C. A., Bellucci, A., Bao, Q., Chang, P., Corti, S., Fučkar, N. S., Guemas, V., von Hardenberg, J., Hazeleger, W., Kodama, C., Koenigk, T., Leung, L. R., Lu, J., Luo, J.-J., Mao, J., Mizielinski, M. S., Mizuta, R., Nobre, P., Satoh, M., Scoccimarro, E., Semmler, T., Small, J., and von Storch, J.-S. (2016). High resolution model intercomparison project (HighResMIP v1.0) for CMIP6. *Geosci. Model Dev.*, 9(11):4185–4208.

Hallberg, R. (2013). Using a resolution function to regulate parameterizations of oceanic mesoscale eddy effects. *Ocean Model.*, 72:92–103.

Jansen, M. F. and Held, I. M. (2014). Parameterizing subgrid-scale eddy effects using energetically consistent backscatter. *Ocean Model.*, 80:36–48.

Jansen, M. F., Held, I. M., Adcroft, A. J., and Hallberg, R. (2015). Energy budget-based backscatter in an eddy permitting primitive equation model. *Ocean Model.*, 94:15–26.

Jayne, S. R. and Marotzke, J. (2002). The oceanic eddy heat transport. *J. Phys. Oceanogr.*, 32(12):3328–3345.

Juricke, S., MacLeod, D., Weisheimer, A., Zanna, L., and Palmer, T. N. (2018). Seasonal to annual ocean forecasting skill and the role of model and observational uncertainty. *Quart. J. R. Meteorol. Soc.*, 144(715):1947–1964.

Juricke, S., Palmer, T. N., and Zanna, L. (2017). Stochastic sub-grid scale ocean mixing: Impacts on low frequency variability. *J. Climate*, 30(30):4997–5019.

Kirtman, B. P., Bitz, C., Bryan, F., Collins, W., Dennis, J., Hearn, N., Kinter, J. L., Loft, R., Rousset, C., Siqueira, L., Stan, C., Tomas, R., and Vertenstein, M. (2012). Impact of ocean model resolution on CCSM climate simulations. *Clim. Dynam.*, 39(6):1303–1328.

Kitsios, V., Frederiksen, J., and Zidikheri, M. (2013). Scaling laws for parameterisations of subgrid eddy–eddy interactions in simulations of oceanic circulations. *Ocean Model.*, 68:88–105.

Klöwer, M., Jansen, M. F., Claus, M., Greatbatch, R. J., and Thomsen, S. (2018). Energy budget-based backscatter in a shallow water model of a double gyre basin. *Ocean Model.*, 132:1–11.

Leutbecher, M., Lock, S.-J., Ollinaho, P., Lang, S. T. K., Balsamo, G., Bechtold, P., Bonavita, M., Christensen, H. M., Diamantakis, M., Dutra, E., English, S., Fisher, M., Forbes, R. M., Goddard, J., Haiden, T., Hogan, R. J., Juricke, S., Lawrence, H., MacLeod, D., Magnusson, L., Malardel, S., Massart, S., Sandu, I., Smolarkiewicz, P. K., Subramanian, A., Vitart, F., Wedi, N., and Weisheimer, A. (2017). Stochastic representations of model uncertainties at ECMWF: state of the art and future vision. *Quart. J. Roy. Meteorol. Soc.*, 143(707):2315–2339.

Mana, P. P. and Zanna, L. (2014). Toward a stochastic parameterization of ocean mesoscale eddies. *Ocean Model.*, 79:1–20.

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

Pearson, B., Fox-Kemper, B., Bachman, S., and Bryan, F. (2017). Evaluation of scale-aware subgrid mesoscale eddy models in a global eddy-rich model. *Ocean Model.*, 115:42–58.

Rackow, T., Goessling, H. F., Jung, T., Sidorenko, D., Semmler, T., Barbi, D., and Handorf, D. (2018). Towards multi-resolution global climate modeling with ECHAM6–FESOM. Part II: climate variability. *Clim. Dynam.*, 50(7):2369–2394.

Roberts, M. J., Hewitt, H. T., Hyder, P., Ferreira, D., Josey, S. A., Mizielinski, M., and Shelly, A. (2016). Impact of ocean resolution on coupled air-sea fluxes and large-scale climate. *Geophys. Res. Lett.*, 43(19):10430–10438.

Roux, D. Y. L. (2012). Spurious inertial oscillations in shallow-water models. *Journal of Computational Physics*, 231(24):7959 – 7987.

Scott, R. B. and Wang, F. (2005). Direct evidence of an oceanic inverse kinetic energy cascade from satellite altimetry. *J. Phys. Oceanogr.*, 35(9):1650–1666.

Sein, D. V., Danilov, S., Biastoch, A., Durgadoo, J. V., Sidorenko, D., Harig, S., and Wang, Q. (2016). Designing variable ocean model resolution based on the observed ocean variability. *J. Adv. Model. Earth Syst.*, 8(2):904–916.

Sein, D. V., Koldunov, N. V., Danilov, S., Wang, Q., Sidorenko, D., Fast, I., Rackow, T., Cabos, W., and Jung, T. (2017). Ocean modeling on a mesh with resolution following the local Rossby radius. *J. Adv. Model. Earth Syst.*, 9(7):2601–2614.

Shutts, G. (2005). A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Quart. J. Roy. Meteorol. Soc.*, 131:3079–3102.

Sidorenko, D., Rackow, T., Jung, T., Semmler, T., Barbi, D., Danilov, S., Dethloff, K., Dorn, W., Fieg, K., Goessling, H. F., Handorf, D., Harig, S., Hiller, W., Juricke, S., Losch, M., Schröter, J., Sein, D. V., and Wang, Q. (2015). Towards multi-resolution global climate modeling with ECHAM6–FESOM. Part I: model formulation and mean climate. *Clim. Dynam.*, 44(3):757–780.

Soufflet, Y., Marchesiello, P., Lemarié, F., Jouanno, J., Capet, X., Debreu, L., and Benshila, R. (2016). On effective resolution in ocean models. *Ocean Model.*, 98:36–50.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bull. Amer. Meteorol. Soc.*, 93(4):485–498.

Wang, Q., Danilov, S., Sidorenko, D., Timmermann, R., Wekerle, C., Wang, X., Jung, T., and Schröter, J. (2014). The finite element sea ice–ocean model (FESOM) v.1.4: formulation of an ocean general circulation model. *Geosci. Model Dev.*, 7(2):663–693.

Wang, S., Liu, Z., and Pang, C. (2015). Geographical distribution and anisotropy of the inverse kinetic energy cascade, and its role in the eddy equilibrium processes. *J. Geophys. Res. Oceans*, 120(7):4891–4906.

Appendix A. Leith viscosity scheme

An alternative to the biharmonic Reynolds viscosity with coefficient following Eq. 5 relies on the Leith and modified Leith viscosities (Fox-Kemper

and Menemenlis, 2008)

$$\nu_c^{L0} = 0.2 \, \min\Big\{ S_c \sqrt{S_c \left(c_D \, |\boldsymbol{\nabla\nabla} \cdot \boldsymbol{u}|_c^2 + c_L \, |\boldsymbol{\nabla} q|_c^2 \right)} + \nu_0 \sqrt{S_c / S_r}, S_c / \Delta t \Big\}, \tag{A.1}$$

with the minimum function serving to taper the viscosity to ensure that the CFL condition is not violated. Here, $q = \boldsymbol{k} \cdot (\boldsymbol{\nabla} \times \boldsymbol{u})$ denotes the vertical vorticity and c_D, c_L are constants on the order of one. The term $\nu_0 \sqrt{S_c / S_r}$ provides a resolution-scaled minimum viscosity where ν_0 is the constant minimum viscosity coefficient and $S_r = 5.8 \cdot 10^9 \, \mathrm{m}^2$ as before. The Leith viscosity can be associated with the scaling laws for 2D turbulence in the enstrophy range (Fox-Kemper and Menemenlis, 2008). Since it is already scale selective, it is used with the harmonic operator. A recent study by Pearson et al. (2017) shows that this form of viscosity generally leads to smaller total dissipation than the biharmonic one with quasi-uniform coefficient. However, it is nonuniform in space, and may be insufficient to damp a grid mode where ν_c^{L0} is small. It is therefore run in combination with small background biharmonic viscosity with coefficient

$$\nu_c^{bb} = \max(0, u_0^{bb} \, S_c^{1/2} - \nu_c^{L0}) \tag{A.2}$$

where $u_0^{bb} = 0.01 \, \mathrm{ms}^{-1}$ so that the background viscosity is active only when the harmonic Leith viscosity is small.

The numerical factors in the formulas above are obtained experimentally as to prevent buildup of grid-scale noise in the vertical velocity field. They differ from those in Pearson et al. (2017) because of the differences in mesh geometry and discretization. The Leith viscosity coefficient is rather patchy in our simulations and is smoothed with a spatial filter F defined in Eq. C.3, so that $\nu_c^L = (\mathsf{F}^2)_{cc'} \, \nu_{c'}^{L0}$. Here, F^2 denotes the composition of F with itself, and Einstein summation over the repeated index c', running over neighbors of c, is implied.

Appendix B. Approximate viscosity operators

Appendix B.1. Approximate harmonic viscosity operator

To compute the viscous force in the discrete case, one needs to estimate the components of the stress tensor $\boldsymbol{\sigma}$ at the edges of the triangular cell. The components can be easily computed either at cell centroids or at vertices, and then averaged to edges. As it turns out, the resulting divergence of $\boldsymbol{\sigma}$ on the

triangle c involves the velocities at neighboring triangles – i.e. those sharing the same edges – only with small weights, which reduce to zero on equilateral meshes. As a consequence, the difference between the velocities on triangle c and neighboring triangles c' only weakly contributes to the viscous force. It will therefore fail to damp a grid mode (for a discussion on spurious modes on triangular grids see, e.g., Roux, 2012).

For this reason, we define an approximate viscous operator V which acts dissipatively on the grid mode via

$$(\mathsf{V}\boldsymbol{u})_c = \frac{1}{S_c} \sum_{c' \in \mathcal{N}(c)} (\boldsymbol{u}_{c'} - \boldsymbol{u}_c) \, \frac{\ell_{c'c}}{|\boldsymbol{r}_{c'c}|} \, \frac{\nu_{c'} + \nu_c}{2} \tag{B.1}$$

where $\mathcal{N}(c)$ is the set of triangles having a common edge with c, $\ell_{c'c}$ is the length of the edge between cells c' and c, $\boldsymbol{r}_{c'c}$ is the vector pointing from the centroid of cell c' to the centroid of cell c, and S_c as before.

When the mesh cells are equilateral triangles, $(\boldsymbol{u}_{c'} - \boldsymbol{u}_c)/|\boldsymbol{r}_{c'c}|$ is the velocity gradient in the direction of the outward normal to the edge between c and c', so that the right of Eq. B.1 is the sum of viscous fluxes leaving cell c, thus defining a consistent discretization of the harmonic viscosity operator $\boldsymbol{V}(\boldsymbol{u}) = \boldsymbol{\nabla} \cdot (\nu \boldsymbol{\nabla} \boldsymbol{u})$. Although this interpretation is only approximate on general meshes, we use the formula above as a substitute for the harmonic viscosity operator (see also the discussion of different viscosity operator options for FESOM2 in Danilov et al., 2017).

Since the geometric quantities in Eq. B.1 and the mean viscosity are defined on edges, they can be incorporated into a generalized edge viscosity $\nu_{c'c}$, which is symmetric between c and c', so that Eq. B.1 takes the form

$$(\mathsf{V}\boldsymbol{u})_c = \frac{1}{S_c} \sum_{c' \in \mathcal{N}(c)} (\boldsymbol{u}_{c'} - \boldsymbol{u}_c) \, \nu_{c'c} \, . \tag{B.2}$$

Summing over all cells c, $\sum_c S_c \, (\mathsf{V}\boldsymbol{u})_c = 0$ because the difference between $\boldsymbol{u}_{c'}$ and $\boldsymbol{u}_c$ appears with opposite signs in the expressions for cell c' and c. Hence, momentum is globally preserved. Furthermore, taking the total viscous dissipation $\sum_c S_c \mathcal{V} = \sum_c S_c \, \boldsymbol{u}_c \cdot (\mathsf{V}\boldsymbol{u})_c$, we observe that the contribution from c' and c appears twice, one time as $\nu_{cc'} \, \boldsymbol{u}_c \cdot (\boldsymbol{u}_{c'} - \boldsymbol{u}_c)$ and the other time as $\nu_{c'c} \, \boldsymbol{u}_{c'} \cdot (\boldsymbol{u}_c - \boldsymbol{u}_{c'})$, which sum to $-\nu_{c'c} \, |\boldsymbol{u}_{c'} - \boldsymbol{u}_c|^2$. (Note that this rearrangement is equivalent to operations in Eq. 7, with the flux contribution disappearing because of the area sum.) This proves that area mean kinetic

energy dissipation is non-positive and introduces a discrete equivalent of the sign-definite dissipation rate $\overline{\mathcal{V}}$ (see Eq. 7), namely

$$\overline{\mathcal{V}}_{c'c} = -\frac{\nu_{c'c}}{S_{c'c}} \left| \boldsymbol{u}_{c'} - \boldsymbol{u}_c \right|^2 , \tag{B.3}$$

which is associated with the area $S_{c'c} = (S_{c'} + S_c)/3$ around the edge $c'c$. It is conservatively remapped to a cell dissipation rate

$$\overline{\mathcal{V}}_c = \frac{1}{3} \sum_{c' \in \mathcal{N}(c)} \overline{\mathcal{V}}_{c'c} \, . \tag{B.4}$$

If layer thicknesses vary, the coefficients must vary with volume rather than area.

Appendix B.2. Approximate biharmonic viscosity operator

A discrete biharmonic viscosity operator is obtained by applying an expression of the form of Eq. B.1 twice.

To avoid splitting the viscosity coefficient between the two harmonic operators, we define

$$(\mathsf{L}\boldsymbol{u})_c = \sum_{c' \in \mathcal{N}(c)} (\boldsymbol{u}_{c'} - \boldsymbol{u}_c) \, , \tag{B.5}$$

which we can represent by a symmetric matrix $\mathsf{L}_{c'c}$, where $\mathsf{L}_{c'c} = 1$ if $c' \neq c$ have a common edge and zero otherwise, and $\mathsf{L}_{cc} = -\sum_{c' \in \mathcal{N}(c)} \mathsf{L}_{c'c}$. Then $(\mathsf{L}\boldsymbol{u})_c = \mathsf{L}_{cc'} \, \boldsymbol{u}_{c'}$ where summation over repeated indices is implied. We further define an operator N whose matrix representation is diagonal with $\mathsf{N}_{cc} = \nu^b_c$. Then the discrete biharmonic operator defined via

$$(\mathsf{V}^b \boldsymbol{u})_c = -\frac{1}{S_c} (\mathsf{LNL}\boldsymbol{u})_c = -\frac{1}{S_c} \mathsf{L}_{cn} \, \mathsf{N}_{nm} \, \mathsf{L}_{mj} \, \boldsymbol{u}_j \tag{B.6}$$

is symmetric and negative, as can be shown by taking the dot product with $\boldsymbol{u}_c$ and summing over all cells. As for the case of harmonic viscosity, this suggests a sign-definite expression for the energy dissipation rate, defined on cells,

$$\overline{\mathcal{V}}^b_c = -\frac{\nu^b_c}{S_c} \left| (\mathsf{L}\boldsymbol{u})_c \right|^2 . \tag{B.7}$$

It is convenient to specify the coefficient of biharmonic viscosity in terms of its harmonic equivalent by setting $\nu^b_c = S_c \, \nu_c$ (no summation implied), which is assumed in our implementation.

Appendix C. Smoothing filter

As discussed in the main text, we require smoothing filters in several places of our implementation of backscatter: to smooth the Leith coefficient of viscosity, to smooth the diagnosed dissipation rate $\mathcal{V}$, and, most importantly, to ensure that the energy returned back to the flow is free of small scales.

Our filters are powers of a single filtering operator: the composition of averaging from cell centroids to common vertices, denoted X, with averaging from vertices back to centroids, denoted C. Let a_c be any quantity defined at cells. Its vertex values $\bar{a}_v$ are obtained as

$$\bar{a}_v = (\mathsf{X}a)_v = \sum_{c \in \mathcal{C}(v)} a_c \, S_c \Big/ \sum_{c \in \mathcal{C}(v)} S_c \,, \tag{C.1}$$

where $\mathcal{C}(v)$ denotes the set of cells containing vertex v. Letting X_{vc} denote the associated matrix symbol, we can also write $\bar{a}_v = \mathsf{X}_{vc} \, a_c$.

Similarly, for a quantity b_v defined on vertices, we define its cell values $\bar{b}_c$ via

$$\bar{b}_c = (\mathsf{C}b)_c = \frac{1}{3} \sum_{v \in \mathcal{V}(c)} b_v \,, \tag{C.2}$$

where $\mathcal{V}(c)$ is the set of vertices of cell c.

Finally, we set

$$\mathsf{F} = \mathsf{C}\mathsf{X} \,. \tag{C.3}$$

The effective spatial smoothing length of Eq. C.3 varies with grid resolution. However, the filter can be applied several times to increase the smoothing length. For example, the Leith viscosity is filtered twice, which can be written as F^2.