Towards a turbulence closure based on energy modes*

Jan Viebahn*

CWI, Amsterdam, Netherlands

Daan Crommelin

CWI, and Korteweg-de Vries Institute for Mathematics, University of Amsterdam, Amsterdam, Netherlands

Henk Dijkstra

Institute for Marine and Atmospheric Research Utrecht, Department of Physics, Utrecht University, Utrecht, Netherlands

**Corresponding author address:* CWI, P.O. Box 94079, 1090 GB, Amsterdam, Netherlands.

E-mail: viebahn@cwi.nl

*To appear in the Journal of Physical Oceanography (2019)

ABSTRACT

A new approach to parameterising sub-grid scale processes is proposed: The impact of the unresolved dynamics on the resolved dynamics, i.e. the eddy forcing, is represented by a series expansion in dynamical spatial modes that stem from the energy budget of the resolved dynamics. It is demonstrated that the convergence in these so-called energy modes is faster by orders of magnitude than the convergence in Fourier-type modes. Moreover, a novel way to test parameterisations in models is explored. The resolved dynamics and the corresponding instantaneous eddy forcing are defined via spatial filtering that accounts for the representation error of the equations of motion on the low-resolution model grid. In this way, closures can be tested within the high-resolution model, and the effects of different parameterisations related to different energy pathways can be isolated. In this study, the focus is on parameterisations of the baroclinic energy pathway. The corresponding standard closure in ocean models, i.e. the Gent-McWilliams (GM) parameterisation, is also tested, and it is found that the GM field acts like a stabilising direction in phase space. The GM field does not project well on the eddy forcing and hence fails to excite the model's intrinsic low-frequency variability but it is able to stabilise the model.

1. Introduction

It is crucial that climate models are able to accurately simulate the climate's internal variability, in addition to the climate's mean state and externally forced climate changes (IPCC 2013). For example, a correct representation of internal climate variability is needed in *climate change detection and attribution* studies. Such studies are based on signal-to-noise estimates for which the climate's intrinsic low-frequency variability (LFV) must be estimated, at least in part, from long control integrations of climate models. Also for *climate prediction* studies a correct representation of the intrinsic climate variability is crucial such that internally generated sources of predictability can be exploited. Finally, the ability of models to make quantitative *projections of changes in climate variability*, including the statistics of extreme events under a warming climate, is dependent on an accurate representation of the climate's internal variability.

The climate's intrinsic LFV is typically described by *large-scale modes of climate variability* which are often either statistical eigenmodes (e.g. EOFs) or dynamical eigenmodes (e.g. linear instability modes, see Storch and Zwiers (1999); IPCC (2013); Dijkstra (2016)). The modes of climate variability are characterized as large-scale because they include large spatial structures such as basin-wide coupled modes of ocean-atmosphere variability (e.g. the El Niño-Southern Oscillation), Rossby wave-trains, mid-latitude jets and storm-tracks, etc.

In particular with respect to the ocean, a number of LFV modes (i.e. multi-annual to multi-decadal time scales) have been described (Deser et al. 2010; Dijkstra 2016). It is clear from observations that multidecadal patterns of sea surface temperature variability exist, e.g., the Atlantic Multidecadal Oscillation (Schlesinger and Ramankutty 1994; Kerr 2000) and the Pacific Decadal Oscillation (Mantua et al. 1997; England et al. 2014). Most of these modes have a particular regional or even global manifestation whose amplitude can be larger than that of human-induced climate change. For example, intrinsic multidecadal variability of the ocean heat content has been held responsible for the relatively low recent trend in the global mean surface temperature anomaly, also referred to as the "Global Warming Hiatus" (Meehl et al. 2011, 2013).

However, care is required when interpreting modes of climate variability since (i) their interpretation depends on how one separates modes of variability from forced changes in the time mean, (ii) they may change drastically in space, structure or probability distribution in response to climate change, and (iii) in strongly nonlinear regimes they may be not strictly large scale but the *large-scale structures can be entangled with smaller-scale structures* such that some modes of climate variability may not be entirely representable in climate models with a too coarse spatial resolution.

Moreover, due to the fact that the real ocean dynamics resides in a highly turbulent regime (with a large Reynolds number leading to a high-dimensional unstable manifold on the attractor) it still remains hard to understand the exact physical mechanisms behind the ocean's LFV (Berloff and McWilliams 1999a; Hogg et al. 2005; Dijkstra 2016). Plenty of model studies analyzing eddy-resolving ocean models show that LFV in such models is commonplace (Berloff and McWilliams 1999a; Hogg et al. 2005) and it is now known that the collective action of oceanic mesoscale eddies is one of the main drives of the midlatitude LFV (Kwon 2010). But at the same time the strong eddy field can obscure many features of the circulation, making it difficult to agree upon the mechanisms underpinning the variability (Hogg et al. 2005; Dijkstra 2016).

Central questions still need further clarification: Which part of the ocean's LFV is completely intrinsic to the ocean and which part involves a dynamical coupling to the atmosphere? Which part the ocean's intrinsic LFV can be traced back to stationary modes at high viscosity (i.e. low-order bifurcations) and which part represents a genuinely eddy-driven turbulent phenomenon (i.e.

physical mechanisms solely active at high Reynolds numbers) (Hogg and Blundell 2006; Berloff et al. 2007; Le Bars et al. 2016; Dijkstra 2016)?

Clarification of these questions is hampered by the fact that computational limitations force most studies on climate variability to employ climate models with ocean components that do not resolve the internal Rossby deformation radius (Hallberg 2013). In these coarse-resolution ocean models (typically operating at a horizontal resolution of 1°) usually deterministic eddy parameterizations are applied which are based on diffusive terms that aim to model the potential and kinetic energy transfer from the mean field to the eddy field. These *diffusive* eddy parameterizations achieve a reasonable representation of the time-mean effect of the mesoscale eddy field on the time-mean flow (Bryan et al. 2014; Griffies et al. 2015; Viebahn et al. 2016) but they are not able to excite the ocean's internal LFV observed in eddy-resolving ocean model simulations (Le Bars et al. 2016). Consequently, the estimation of *internal variability uncertainty* (stemming from the chaotic nature of the system) in climate change detection or projections of climate change is still strongly hampered by *model uncertainty* (i.e. limitations of a model's representation of the chaotic nature of the system) in many current climate change studies.

Hence, the search for suitable eddy parameterizations remains a challenging theoretical topic with clear practical dimension. Recently, efforts have been made towards eddy parameterizations that aim to step out of the diffusive parameterization framework and try to represent the eddy effects in terms of *stochastic* eddy forcing (Berloff 2005c; Grooms and Majda 2013; Mana and Zanna 2014; Verheul et al. 2017). Stochastic climate modelling is based on the concept of *scale separation in time* (Franzke et al. 2015), namely, that the state vector of the system can be decomposed into fast modes and slow (low-frequency) modes such that the time scales of these modes strongly differ. The impact of the fast modes on the slow modes appears as eddy forcing in the equations of motion for the slow modes. The development of stochastic climate models then proceeds by accounting for the effects of the unresolved fast modes in a stochastic fashion.

Moreover, for models formulated in physical space (like most ocean models) the essential difference between a high-resolution model and a low-resolution model is the extent of spatial information. The eddy forcing actually represents the impact of the spatially unresolved (or sub-grid scale or small-scale) processes on the spatially resolved (or larger scale) processes. Consequently, for models in physical space time-scale separation should imply *scale separation in space*. That is, the patterns associated with slow variability should exhibit strictly large-scale spatial structures whereas the patterns associated with fast variability should show strictly small-scale spatial structures. Otherwise, the slow modes and the fast modes cannot be disentangled on the low-resolution model grid.

However, scale separation only holds for regimes in which scales are weakly coupled whereas in turbulent regimes different scales are strongly nonlinearly coupled. The lack of time-scale separation introduces *non-Markovian memory effects* and complicates the derivation of systematic parameterizations. The lack of scale separation in space implies that the dynamical modes are *multiscale patterns* both in the horizontal and vertical directions. For example, for the midlatitude ocean gyres it is found that due to the background flow most eigenmodes contain a large variety of scales (Shevchenko et al. 2016). In this case, the LFV is not a single-mode pattern, but it is a coherent pattern phenomenon consisting of a large number of short period phase-related eigenmodes interacting with each other. We note that this can apply to both (high-resolution) statistical eigenmodes like EOFs (Gille and Kelly 1996) and linear eigenmodes on a background flow (Shevchenko et al. 2016). Obviously, the small-scale structures of the dynamical modes are not resolvable on a low-resolution model grid.

In this study, we approach the formulation of eddy parameterisations in the following way: First, we define the resolved dynamics and the corresponding instantaneous eddy forcing via spatial filtering (instead of e.g. temporal averaging) such that we can account for the representation error of the equations of motion on the low-resolution model grid. Second, we represent the impact of the unresolved dynamics on the resolved dynamics (i.e. the eddy forcing) in terms of a series expansion in dynamical spatial modes that stem from the energy budget of the resolved dynamics. These so-called energy modes exhibit strictly large-scale spatial patterns and are equipped with a clear physical interpretation.

In section 2, the resolved dynamics and the related instantaneous eddy forcing are defined: We describe our eddy-resolving ocean model and its LFV in section 2a. Our spatial filtering approach is introduced in section 2b. The corresponding filtered equations of motion and the related eddy forcing terms are presented in section 2c, and in section 2d we analyse the resulting large-scale and small-scale energetics. Subsequently, section 3 deals with developing and testing closures with a focus on the baroclinic energy pathway. We show how we can test parameterisations in a high-resolution model (section 3a), and test the performance of the standard closure of the baroclinic energy pathway in ocean models (i.e. the Gent-McWilliams (GM) parameterisation (Gent and McWilliams 1990)) in section 3b. Finally, in section 3c we define and test the representation of the eddy forcing in energy modes with a focus on representing the baroclinic energy pathway. We end with a summary (section 4) and discussion (section 5).

2. Framework: Eddy forcing of the large-scale flow

Our general starting point is the following (see e.g. Berloff (2005a)): First, an *eddy-resolving (ER) model* is given (section 2a) in order obtain a reference solution, say with state vector ψ, which contains both the large-scale and eddy components. Second, a *non-eddy-resolving (non-*

ER) model is supposed to have the same general set-up as the ER model (e.g. type of governing conservation equations, domain size and boundary conditions, see section 2c), but the former has a *significantly coarser horizontal grid resolution* (by a factor of ten in this study). Consequently, the non-ER model has far fewer degrees of freedom and it can only solve for the large-scale flow evolution. Moreover, the non-ER model may contain *additional dynamical terms* in the governing conservation equations (e.g. the current deterministic eddy parameterisations) which are supposed to parameterise part of the interactions between large-scale components and (sub-)mesoscale eddy components.

Finally, the *eddy forcing (EF)* is a (not necessarily unique) dynamical term that still needs to be added to the governing conservation equations of the non-ER model at hand such that the non-ER solution, say with state vector $\hat{\psi}$, correctly approximates the large-scale structure of the original flow (i.e. of the ER model solution ψ). That is, the EF represents interactions between the large-scale flow and eddy fluctuations that are relevant for the large-scale flow evolution. The precise form of the EF depends on (i) the specific definition of the large-scale structure of the original flow (section 2b), and (ii) the eddy parameterisations already included in the chosen non-ER model equations (section 3).

a. Eddy-resolving ocean model exhibiting low-frequency variability

We consider a standard model of idealised ocean dynamics, namely, quasi-geostrophic (QG), potential-vorticity (PV) equations in a classical double-gyre configuration (see e.g. Vallis (2006)). The fluid-dynamic model describes idealised, wind-driven midlatitude ocean circulation with prescribed density stratification in a flat-bottom square basin with north-south and east-west boundaries. We employ the QG PV conservation equations for two isopycnal layers which represents the simplest description of baroclinically unstable dynamics (Olbers et al. 2012). These are

$$
\begin{aligned}
\partial_t q_1 + J(\psi_1, q_1) &= A_H \nabla^4 \psi_1 - \frac{\partial_y \tau^x}{\rho_0 H_1} \,, &(1)\\
\partial_t q_2 + J(\psi_2, q_2) &= A_H \nabla^4 \psi_2 \,, &(2)
\end{aligned}
$$

where the PV of the two isopycnal layers is given by

$$q_1 = \nabla^2 \psi_1 + \beta y + \frac{f_0}{H_1}\eta \,, \qquad q_2 = \nabla^2 \psi_2 + \beta y - \frac{f_0}{H_2}\eta \,, \tag{3}$$

with the interface displacement $\eta = (f_0/g')(\psi_2 - \psi_1)$, and horizontal velocities given by $\mathbf{u}_i = (u_i, v_i) = \underset{\neg}{\nabla}\psi_i = (-\partial_y \psi_i, \partial_x \psi_i)$.

In our numerical model simulations, the flow is driven at the surface by the asymmetric double-gyre zonal wind stress (as e.g. in Berloff (2005a,c)),

$$\tau^x(y) = \tau_0 \left[\cos\left(\frac{2\pi(y - L/2)}{L} \right) + 2\sin\left(\frac{\pi(y - L/2)}{L} \right) \right] , \qquad \tau^y = 0 \,, \tag{4}$$

where $\tau_0 = 0.04$ Nm^{-2}, and $L = 3500$ km is the size of the square basin with $0 \leq x, y \leq L$. The first internal Rossby radius of deformation, $R = \sqrt{g' H_1 H_2 / (H f_0^2)}$, represents a length scale of baroclinic eddies. It is set to $R = 40$ km, a typical value for the midlatitude ocean circulation. We use mean isopycnal layer thicknesses of $H_1 = 250$ m, and $H_2 = 3750$ m, such that the mean ocean depth is $H = 4000$ m. We also use typical values for the mean density of sea water, $\rho_0 = 1000$ kgm^{-3}, and the reference Coriolis parameter, $f_0 = 8.34\ 10^{-5}$ s^{-1}, such that we have for the meridional variation of the Coriolis parameter, $\beta = 1.87\ 10^{-11}$ m^{-1}s^{-1}, and for the reduced gravity, $g' = g(\rho_2 - \rho_1)/\rho_1 \approx 0.048$ ms^{-2}. Finally, we use an eddy-resolving horizontal resolution of 10 km with a correspondingly small lateral viscosity coefficient, $A_H = 100$ m^{2}s^{-1}, as well as no-slip boundary conditions (similar to Berloff (2005a,c)). The reference simulation is 500 years long and we analyse daily output.

Figure 1 shows a snapshot (Fig. 1c) and a temporal average (Fig. 1d) of the upper layer streamfunction (similar to Fig. 1a,c in Berloff (2005a), and Fig. 1a and Fig. 2a in Berloff (2005c)). The upper-ocean time-mean circulation (Fig. 1d) consists of the southern (subtropical) and northern (subpolar) gyres that fill about 2/3 and 1/3 of the basin, respectively, which is consistent with the wind stress pattern. The time-mean flow is characterised by the Sverdrup balance in most parts of the basin. Only in regions related to the pair of the western boundary currents and their eastward-jet (EJ) extensions non-linear and frictional terms become dominant (Pedlosky 1996). We note that for our specific model setup the boundary currents do not merge with each other but the subpolar gyre enters the subtropical region near the western boundary such that the point of separation from the coast of the subtropical western boundary current is pushed southward relative to the line of zero wind-stress curl (similar to Berloff (2005a,c)). This is a robust regime that appears at large Reynolds number in the stratified and baroclinically unstable double-gyre flow with no-slip boundary conditions (e.g. Haidvogel et al. (1992); Berloff and McWilliams (1999b); Siegel et al. (2001)). In terms of the fluctuations, the basin can be partitioned into the more energetic 'western' part, characterized by strong vortices, and the less energetic 'eastern' part, dominated by the planetary waves (see Berloff et al. (2002) for details).

The corresponding reservoirs of kinetic energy (KE) and available potential energy (PE) are given by

$$\mathrm{KE} = -\frac{\rho_0}{2}\int (H_1\psi_1\nabla^2\psi_1 + H_2\psi_2\nabla^2\psi_2)\,dA\ , \quad \mathrm{PE} = \frac{\rho_0 g'}{2}\int \eta^2\,dA\ . \tag{5}$$

The two reservoirs are governed by the following conservation equations (obtained by multiplying Eq. (1)-(2) with $-\rho_0 H_i\psi_i$ followed by global integration),

$$\frac{d\mathrm{KE}}{dt} = C(\mathrm{PE},\mathrm{KE}) + G(\mathrm{KE}) + D(\mathrm{KE})\ , \tag{6}$$

$$\frac{d\mathrm{PE}}{dt} = -C(\mathrm{PE},\mathrm{KE})\ , \tag{7}$$

where $C(\text{PE},\text{KE})$ represents the conversion between PE and KE, and the generation of KE, $G(\text{KE})$, and the dissipation of KE, $D(\text{KE})$, are given by

$$G(\text{KE}) = \int \psi_1 \partial_y \tau^x \, dA \ , \quad D(\text{KE}) = -A_H \rho_0 \int (H_1 \psi_1 \nabla^4 \psi_1 + H_2 \psi_2 \nabla^4 \psi_2) \, dA \ . \tag{8}$$

Figure 5 shows the temporal evolution of the energetics of the reference simulation. The PE (Fig. 5a) exhibits clear cycles of decadal variability. The about 4 times smaller KE also shows cycles of decadal variability which lags the variability in PE by 1-2 years. The wind energy input $G(\text{KE})$ (Fig. 5b) is balanced by lateral dissipation $D(\text{KE})$ with both showing also significant high-frequency variability on top of low-frequency variability, with higher variance in $D(\text{KE})$ than in $G(\text{KE})$.

Figure 1 also shows upper layer streamfunction anomalies corresponding to a low (Fig. 1e) and a high (Fig. 1f) in PE (see years 52 and 56 in Fig. 5a). These anomaly patterns are similar to those shown in Berloff et al. (2007) (see their Fig. 2 and Fig. 4), and demonstrate that the variability is concentrated around the subtropical EJ. More precisely, the decadal transitions are related to coherent meridional shifts and variations of the intensity of the subtropical EJ, likely governed by the nonlinear adjustment of the combined EJ-eddies system (see Berloff et al. (2007) for details).

b. Flow decomposition into large-scale and eddy components via spatial mode filtering

In this study, the large-scale flow structure is determined by spatial mode filtering. For that the ER model solution ψ is expanded in a set of *spatial filter modes* χ_i,

$$\psi(\mathbf{x},t) = \sum_i^N \Psi_i(t) \chi_i(\mathbf{x}) \ . \tag{9}$$

Note that the spatial filter modes χ_i are time-independent (i.e. non-dynamical). The corresponding large-scale (or filtered) component of ψ is given by a truncated expansion (equivalent to applying a sharp spectral filter)

$$[\psi](\mathbf{x},t) := \sum_{i}^{\hat{N}<N} \Psi_i(t)\chi_i(\mathbf{x}) \ . \tag{10}$$

The cutoff $\hat{N}$ has to be determined such that the retained spatial filter modes $\chi_{i\leq\hat{N}}$ have a consistent representation on the coarse-resolution grid of the non-ER model (see the consistency conditions below). The non-ER model solution $\hat{\psi}$ (denoting spatial fields on the non-ER model grid by a hat) would then optimally be given by[1]

$$\hat{\psi}(\mathbf{x},t) = \sum_{i}^{\hat{N}} \Psi_i(t)\hat{\chi}_i(\mathbf{x}) \ . \tag{11}$$

More precisely, the specification of the spatial filter modes χ_i and the cutoff $\hat{N}$ is guided by the following three consistency conditions:

(SCC) *Scale-content* (or *image*) *consistency*: The scale content of the spatial filter modes $\chi_{i\leq\hat{N}}$ has to be resolvable by the non-ER model grid resolution in order to avoid *aliasing* effects. The scale content of a spatial pattern is typically measured by the familiar Fourier modes (i.e. eigenmodes of the Laplacian). The corresponding cutoff $\hat{N}_{Nyquist}$ is given by the well-known *Nyquist criterion*, which states that the smallest wavelength included in $\chi_{i\leq\hat{N}_{Nyquist}}$ must not be smaller than twice the grid spacing of the non-ER model. In particular, we note that filtering of the ER model reference solution ψ has to be done on the ER model grid (i.e. by using $\chi_{i\leq\hat{N}}$, and not by using $\hat{\chi}_{i\leq\hat{N}}$ and the injection of ψ on the coarse grid) since otherwise aliasing errors occur.

(BCC) *Boundary conditions consistency*: The large-scale flow is supposed to be a solution of the non-ER model equations and, hence, has to satisfy its boundary conditions. In the governing model equations, the differential operator with the highest order derivative (typically related

[1]Identity with respect to time-evolution is meant in a statistical/dynamical sense.

to dissipation) determines the number of boundary conditions that have to be specified. Consequently, the eigenmodes of this differential operator represent a set of spatial modes which are always able to satisfy the boundary conditions (i.e. span the correct function space), and, hence, represent the first choice if the Fourier modes (i.e. eigenfunctions of the Laplacian) cannot satisfy the boundary conditions.

(DC) *Dynamical consistency*: The conservation equations governing the evolution of $[\psi]$ are obtained by filtering the ER model equations (see section 2c). The non-ER model equations are supposed to represent these equations except that the terms including interactions with eddy components are replaced by eddy parameterisations. For that to hold, the spatial derivatives appearing in the governing equations have to be similar for both $\chi_{i\leq\hat{N}}$ and $\hat{\chi}_{i\leq\hat{N}}$, that is, the differences in computing dynamical terms on the different grids must be not be significant. Otherwise, the EF does not solely represent the interactions between the large-scale flow and eddy fluctuations that are relevant for the large-scale flow evolution but it would also have to compensate for differences simply induced by computing the dynamical budget of the large-scale flow on different grids[2]. Consequently, one must generally require $\hat{N} < \hat{N}_{Nyquist}$.

In our ER model no-slip boundary conditions are applied such that Fourier modes cannot be used as the spatial filter modes χ_i (see BCC condition). Consequently, we use as spatial filter modes the eigenmodes of the Bilaplacian,

$$\nabla^4 \chi_i = \lambda_i \chi_i \ , \tag{12}$$

[2]Note that this criterion is expressed here with respect to the equations in physical space. With respect to the equations in modal/wavenumber space it says that the constant interaction coefficients (obtained by computing the amplitude equations of the individual modes) related to the resolved large-scale modes should not significantly change whether computed from the high-resolution or low-resolution representation of the modes.

for which no-slip boundary conditions can be prescribed. Note that λ_i represents the globally integrated lateral dissipation related to the normalised[3] χ_i since $\lambda_i = \int \chi_i \nabla^4 \chi_i \, dxdy$.

Figure 2 shows selected leading eigenmodes (ortho-normalised) of the Bilaplacian with no-slip boundary conditions computed on the high-resolution (i.e. 10 km) grid. The overall structure (i.e. the scale content) of the χ_i is still very similar to the Fourier modes (but note that the quantitative differences are nevertheless global and *not* only localised at the boundary). Computing the eigen-spectrum of ∇^4 on both the ER model grid (i.e. 10 km resolution) and the non-ER model grid (i.e. 100 km resolution) enables us to specify a cutoff $\hat{N}$ in accordance with condition DC. Figure 3 shows the corresponding eigenvalues and their relative difference. As a threshold we choose 10% relative difference in globally integrated lateral dissipation which implies $\hat{N} \approx 54$. The corresponding relative difference in globally integrated kinetic energy (also shown in Fig. 3) is about 5%.

Figure 1 also shows the corresponding snapshot (Fig. 1a) and temporal average (Fig. 1b) of the large-scale (i.e. filtered with $\hat{N} = 54$) upper layer streamfunction. The overall structure of the double-gyre circulation is captured by the large-scale flow in both cases. In particular, the separation point of the subtropical western boundary current is exactly recovered. However, local differences are obvious (also in the time-mean patterns), for example, the locations of the local extremes are shifted. Consequently, spatial filtering and temporal filtering are not equivalent.

c. Conservation equations of the large-scale flow

The conservation equations governing the evolution of the large-scale flow $[\psi]$ are obtained by applying the filtering operation to the QG-PV budget (1)-(2). Filtering and application of the Bilaplacian obviously commute for the eigenmodes of the Bilaplacian. However, for no-slip

[3]Ortho-normalised in the streamfunction norm (equivalent to PE norm), $\int \chi_i \chi_j \, dxdy = \delta_{ij}$.

boundary conditions filtering with the eigenmodes of the Bilaplacian does not commute with both the zonal derivative (i.e. linear beta term) and the Laplacian. Hence, filtering of the governing equations (1)-(2) leads to the following equations governing the filtered flow[4],

$$\partial_t[q_1] + J([\psi_1],[q_1]) = -\mathscr{R}_1 + A_H\nabla^4[\psi_1] - \frac{[\partial_y\tau^x]}{\rho_0 H_1}\,, \tag{13}$$

$$\partial_t[q_2] + J([\psi_2],[q_2]) = -\mathscr{R}_2 + A_H\nabla^4[\psi_2]\,, \tag{14}$$

where the filtered PV reads $[q_i] = \nabla^2[\psi_i] + \beta[y] + \frac{(-1)^{i-1}f_0}{H_i}[\eta]$. The *residual PV fluxes* $\mathscr{R}_i$, representing interactions between the large-scale flow and eddy fluctuations that are relevant for the large-scale flow evolution, are given by

$$\mathscr{R}_i = \mathscr{R}_i^A + \mathscr{R}_i^T\,, \tag{15}$$

with the residual advection of PV, $\mathscr{R}_i^A$, and the residual related to the time tendency of relative PV, $\mathscr{R}_i^T$, given by

$$\mathscr{R}_i^A \equiv [J(\psi_i,q_i)] - J([\psi_i],[q_i])\,, \tag{16}$$

$$\mathscr{R}_i^T \equiv [\partial_t\nabla^2\psi_i] - \partial_t\nabla^2[\psi_i] = [\nabla^2\partial_t\psi_i] - \nabla^2[\partial_t\psi_i]\,. \tag{17}$$

[4]Note that there is a subtlety here: We assume that the terms $\partial_t\nabla^2[\psi_i] = \nabla^2[\partial_t\psi_i]$, $\beta\partial_x[\psi_i]$, $J([\psi_i],[\nabla^2\psi_i])$, and $J([\psi_i],\frac{(-1)^{i-1}f_0}{H_i}[\eta]) = (-1)^{i-1}\frac{f_0^2}{g'H_i}J([\psi_1],[\psi_2])$ do not project on modes which lie outside the subspace defined by the filter cutoff $\hat{N}$. Of course, every model discretised and stepped forward in physical space (and not directly in modal/wavenumber space) suffers from the fact that energy can be transferred to small-scale modes which cannot be adequately represented on the spatial grid (leading e.g. to aliasing). However, since we use the eigenmodes of the frictional term (which typically represents the most small-scale patterns) we expect to essentially remain within the subspace spanned by the large-scale modes (defined via the filter cutoff $\hat{N}$).

The residual advection of PV can be further decomposed into $\mathscr{R}_i^A = \mathscr{R}_i^\beta + \mathscr{R}_i^M + \mathscr{R}_i^B$, with

$$\mathscr{R}_i^\beta \equiv \beta[\partial_x \psi_i] - \beta\partial_x[\psi_i] \;, \tag{18}$$

$$\mathscr{R}_i^M \equiv [J(\psi_i, \nabla^2\psi_i)] - J([\psi_i], \nabla^2[\psi_i]) \;, \tag{19}$$

$$\mathscr{R}_i^B \equiv [J(\psi_i, \frac{(-1)^{i-1} f_0}{H_i}\eta)] - J([\psi_i], \frac{(-1)^{i-1} f_0}{H_i}[\eta]) = \tag{20}$$

$$= (-1)^{i-1}\frac{f_0^2}{g' H_i}\Big([J(\psi_1,\psi_2)] - J([\psi_1],[\psi_2])\Big) \;, \tag{21}$$

which are related to residual planetary vorticity advection, residual nonlinear momentum fluxes, and residual buoyancy fluxes (i.e. interface displacements), respectively.

In the following, we focus on a twofold decomposition of the residual PV fluxes $\mathscr{R}_i$ into

$$\mathscr{R}_i = \mathscr{R}_i^H + \mathscr{R}_i^B \;, \tag{22}$$

where $\mathscr{R}_i^B$ represents the part of $\mathscr{R}_i$ which is related to the vertical density distribution/layer interaction/interface height/APE, whereas $\mathscr{R}_i^H \equiv \mathscr{R}_i^\beta + \mathscr{R}_i^M + \mathscr{R}_i^T$ is related to the horizontal eddy PV fluxes. In this study, the main focus will be on $\mathscr{R}_i^B$ (see section 3).

d. Lorenz energy cycle

The Lorenz energy cycle (LEC) describes the balances of four mechanical energy reservoirs, the large-scale circulation's kinetic energy ([KE]) and available potential energy ([PE]), the eddy kinetic energy (KE$'$) and eddy available potential energy (PE$'$). The four reservoirs are given by

$$[\mathrm{KE}] = -\frac{\rho_0}{2}\int (H_1[\psi_1]\nabla^2[\psi_1] + H_2[\psi_2]\nabla^2[\psi_2])\; dA \;, \quad \mathrm{KE}' = \mathrm{KE} - [\mathrm{KE}] \;, \tag{23}$$

$$[\mathrm{PE}] = \frac{\rho_0 g'}{2}\int [\eta]^2 \; dA \;, \qquad \mathrm{PE}' = \mathrm{PE} - [\mathrm{PE}] \;, \tag{24}$$

and they are governed by the following conservation equations (obtained by multiplying Eq. (13)-(14) with $-\rho_0 H_i[\psi_i]$ and global integration),

$$\frac{d[\mathrm{KE}]}{dt} = C(\mathrm{KE}',[\mathrm{KE}]) + C([\mathrm{PE}],[\mathrm{KE}]) + G([\mathrm{KE}]) + D([\mathrm{KE}]) \;, \tag{25}$$

$$\frac{d\mathrm{KE}'}{dt} = -C(\mathrm{KE}',[\mathrm{KE}]) + C(\mathrm{PE}',\mathrm{KE}') + G(\mathrm{KE}') + D(\mathrm{KE}') \;, \tag{26}$$

$$\frac{d[\mathrm{PE}]}{dt} = C(\mathrm{PE}',[\mathrm{PE}]) - C([\mathrm{PE}],[\mathrm{KE}]) \;, \tag{27}$$

$$\frac{d\mathrm{PE}'}{dt} = -C(\mathrm{PE}',[\mathrm{PE}]) - C(\mathrm{PE}',\mathrm{KE}') \;. \tag{28}$$

The respective generation and dissipation terms are given by

$$G([\mathrm{KE}]) = \int [\psi_1][\partial_y \tau^x] \; dA \;, \tag{29}$$

$$G(\mathrm{KE}') = G(\mathrm{KE}) - G([\mathrm{KE}]) \;, \tag{30}$$

$$D([\mathrm{KE}]) = -A_H \rho_0 \int (H_1[\psi_1]\nabla^4[\psi_1] + H_2[\psi_2]\nabla^4[\psi_2]) \; dA \;, \tag{31}$$

$$D(\mathrm{KE}') = D(\mathrm{KE}) - D([\mathrm{KE}]) \;, \tag{32}$$

and the terms related to energy exchange between the large-scale flow and eddy components read

$$C(\mathrm{KE}',[\mathrm{KE}]) = \rho_0 \int (H_1[\psi_1]\mathscr{R}_1^H + H_2[\psi_2]\mathscr{R}_2^H) \; dA \;, \tag{33}$$

$$C(\mathrm{PE}',[\mathrm{PE}]) = \rho_0 \int (H_1[\psi_1]\mathscr{R}_1^B + H_2[\psi_2]\mathscr{R}_2^B) \; dA \tag{34}$$

$$= -\rho_0 f_0 \int [\eta][J(\psi_1,\psi_2)] \; dA \;. \tag{35}$$

Note that all LEC terms are instantaneously given due to our spatial (instead of temporal) filtering approach.

Figure 4 shows the different terms of the LEC averaged in time (over 500 years of daily output) and summarises the time-mean state and variance of the different energy reservoirs and energy pathways (for the reference simulation described in section 2a). The filtered terms $G([\mathrm{KE}])$ and $[\mathrm{PE}]$ capture 96% and 89% of the full (i.e. unfiltered) $G(\mathrm{KE})$ and PE values, respectively, implying that $G(\mathrm{KE})$ and PE are dominated by large-scale structures. In contrast, $D([\mathrm{KE}])$ is very small

(1.8% of $D(\mathrm{KE})$) implying that $D(\mathrm{KE})$ is dominated by small-scale structures. Consequently, almost all of $G(\mathrm{KE})$ has to be transferred to the eddy field via eddy fluxes. Both conversion terms, $C([\mathrm{PE}],\mathrm{PE}')$ and $C([\mathrm{KE}],\mathrm{KE}')$, have the same order of magnitude but $C([\mathrm{PE}],\mathrm{PE}')$ dominates (almost twice as large both in temporal average and variance). The two eddy energy reservoirs are of similar magnitude with KE' being almost 5 times larger than [KE] (capturing 83% of KE).

The overall picture is similar to the one found in realistic global ocean models (e.g. von Storch et al. (2012)). Common in both the ocean and the atmosphere is that the dominant power pathway is the *baroclinic pathway* $[\mathrm{PE}] \to \mathrm{PE}' \to \mathrm{KE}'$ characterized by a conversion $C([\mathrm{PE}],\mathrm{PE}')$ from the large-scale available potential energy to the eddy available potential energy that has about the same magnitude[5] as the conversion $C(\mathrm{PE}',\mathrm{KE}')$ from the eddy potential energy to the eddy kinetic energy. That is, as in the atmosphere, oceanic mesoscale eddies are, to a large extent, generated by baroclinic instability which is the main mechanism in converting the large-scale available potential energy into the eddy kinetic energy in the ocean. Moreover, and in contrast to the atmosphere, the two conversion terms connected to the large-scale kinetic energy [KE] (i.e., $C([\mathrm{KE}],\mathrm{KE}')$ and $C([\mathrm{KE}],[\mathrm{PE}])$) are directed away from [KE] in the ocean. That is, the two main power pathways in the ocean are $[\mathrm{KE}] \to [\mathrm{PE}] \to \mathrm{PE}' \to \mathrm{KE}'$ and $[\mathrm{KE}] \to \mathrm{KE}'$. The oceanic large-scale circulation, being fuelled by the winds, converts its kinetic energy into the large-scale available potential energy by Ekman pumping. This conversion substantially facilitates density differences and hence the large-scale available potential energy from which the baroclinic pathway originates. The oceanic large-scale circulation converts also its kinetic energy into the eddy kinetic energy.

Figure 5 shows different terms of the LEC evolving in time. The variability in [PE] (Fig. 5a) and $G([\mathrm{KE}])$ (Fig. 5b) is essentially identical to the variability in PE and $G(\mathrm{KE})$, respectively. This

[5] In our model setup the two conversion terms are identical in the time-mean (see Fig. 4) due to the absence of buoyancy sources/sinks.

means that the low-frequency variability in these fields is indeed large-scale, and hence, can in principle be adequately captured by a non-ER model. The converse is true for lateral dissipation $D(\mathrm{KE})$ (Fig. 5b) for which also the variability (next to the time-mean value) of its large-scale component $D([\mathrm{KE}])$ is very small. The KE reservoir (Fig. 5a) represents an intermediate quantity in the sense that its large-scale component [KE] only captures part of the low-frequency variability.

The large-scale wind energy input $G([\mathrm{KE}])$ (Fig. 5b) is balanced by the energy transfer to the eddy components via $C([\mathrm{KE}],\mathrm{KE}')$ (Fig. 5c) and $C([\mathrm{KE}],[\mathrm{PE}]) \rightarrow C([\mathrm{PE}],\mathrm{PE}')$ (Fig. 5d,e). Most importantly, both $C([\mathrm{PE}],\mathrm{PE}')$ and $C([\mathrm{KE}],\mathrm{KE}')$ regularly show backscatter, that is, energy transfer from the eddy components to the large-scale components. Moreover, the variances of $C([\mathrm{PE}],\mathrm{PE}')$ and $d[\mathrm{PE}]/dt$ (both part of Eq. (27)) are significantly larger than the variances of $C([\mathrm{KE}],\mathrm{KE}')$ and $C([\mathrm{KE}],[\mathrm{PE}])$. We note that $C([\mathrm{PE}],\mathrm{PE}')$ and $d[\mathrm{PE}]/dt$ are highly anti-correlated with a correlation coefficient of -0.89 (they tend to be positively correlated when $C([\mathrm{KE}],[\mathrm{PE}]) < C([\mathrm{KE}],\mathrm{KE}')$), whereas the correlation coefficient of $C([\mathrm{PE}],\mathrm{PE}')$ and $C([\mathrm{KE}],[\mathrm{PE}])$ ($d[\mathrm{PE}]/dt$ and $C([\mathrm{KE}],[\mathrm{PE}])$) is 0.42 (0.12).

3. Closures for the baroclinic energy pathway

In stratified flows two distinctively different types of energy conversions between large-scale and eddy components exist. Namely, the energy conversion $C([\mathrm{PE}],\mathrm{PE}')$ involving density perturbations (see Eq. (34)), and the energy conversion $C([\mathrm{KE}],\mathrm{KE}')$ solely related to (horizontal) velocity perturbations (see Eq. (33)). In the temporal average (see Fig. 4), the latter represents a sink of [KE], whereas the former represents a sink of [PE] as part of the baroclinic energy pathway, [PE] $\rightarrow \mathrm{PE}' \rightarrow \mathrm{KE}'$. Instantaneously, both conversion terms can also backscatter, that is, transfer energy from the small-scale components to the large-scale components (Fig. 5c,e). In a non-ER model these two energy transfers have to be adequately modelled. In this study, we focus on closures for

$C([\mathrm{PE}],\mathrm{PE}')$, corresponding to $\mathscr{R}_i^B$ in the large-scale PV budget (see Eq. (34) and Eq. (20)), and leave the development of adequate closures for $C([\mathrm{KE}],\mathrm{KE}')$ (i.e. $\mathscr{R}_i^H$) for future work (note that $C([\mathrm{PE}],\mathrm{PE}')$ generally dominates over $C([\mathrm{KE}],\mathrm{KE}')$, see Fig. 4).

a. Testing closures in an eddy-resolving model

In order to be able to isolate the direct effects of $\mathscr{R}_i^B$ and the performance of corresponding closures we adopt the following approach: For a large-scale flow defined via spatial filtering (section 2b) the corresponding conservation equations (section 2c) can be computed instantaneously from the corresponding eddy-resolving model equations (section 2a). In other words, the non-ER model, Eq. (13)-(14), can be considered as part of the ER model, Eq. (1)-(2). In order to be able to test closures for $\mathscr{R}_i^B$ (Eq. (20)) in an isolated way, that is, without the need to also parameterise $\mathscr{R}_i^H$, we perform simulations with the ER model equations (1)-(2) in which we employ the following decomposition of the Jacobian J at every time step,

$$\begin{aligned} J(\psi_i, \frac{(-1)^{i-1} f_0}{H_i}\eta) \quad &= \quad [J(\psi_i, \frac{(-1)^{i-1} f_0}{H_i}\eta)] + \Delta J = \\ &\overset{Eq.\ (20)}{=} \quad J([\psi_i], \frac{(-1)^{i-1} f_0}{H_i}[\eta]) + \mathscr{R}_i^B + \Delta J \ , \end{aligned} \tag{36}$$

where $\Delta J \equiv J(\psi_i, \frac{(-1)^{i-1} f_0}{H_i}\eta) - [J(\psi_i, \frac{(-1)^{i-1} f_0}{H_i}\eta)]$ corresponds to the small-scale component[6] of $J(\psi_i, \frac{(-1)^{i-1} f_0}{H_i}\eta)$. Note that $J([\psi_i], \frac{(-1)^{i-1} f_0}{H_i}[\eta])$ can be computed form the large-scale fields and only redistributes large-scale energy but does not contribute to large-scale energy dissipation/generation.

Then a parametersation of $\mathscr{R}_i^B$, say $\tilde{\mathscr{R}}_i^B$, can be tested by performing simulations with the ER model equations (1)-(2) and including at every time step the replacement $\mathscr{R}_i^B \rightarrow \tilde{\mathscr{R}}_i^B$ in Eq. (36). That is, the large-scale component of the Jacobian, $\mathscr{R}_i^B$ (which is needed in the non-ER model), is parameterised whereas the small-scale component, ΔJ, remains explicitly computed. We empha-

[6] Note that $[\Delta J] = 0$ since we use a sharp filter.

sise that the 'true' $\mathscr{R}_i^B$ is always available since we solely perform simulations with the ER model. Hence, quantities like the relative error $||\mathscr{R}_i^B - \tilde{\mathscr{R}}_i^B||/||\mathscr{R}_i^B||$ can be computed at every time step. As demonstrated in the following (sections 3b and 3c), it is by no means a trivial task to parameterise $\mathscr{R}_i^B$ in such a way that the energy level and low-frequency variability of the large-scale flow are captured.

b. Standard GM parameterisation

In general, the GM parameterisation is interpreted as the standard down-gradient parameterisation for the horizontal component of the isopycnal eddy flux (Vallis 2006; Olbers et al. 2012). In a layer model, this corresponds to down-gradient diffusion of interface displacement η. More precisely, the isopycnal interface PV flux is given by $\mathbf{u}_i\eta$. Assuming a Reynolds decomposition into mean (denoted by an overbar) and eddy (denoted by a prime) components (e.g. via temporal averaging, see e.g. Pope (2000)), the isopycnal interface eddy PV flux is given by $\overline{\mathbf{u}_i'\eta'}$, and the GM parameterisation reads

$$\overline{\mathbf{u}_i'\eta'} = -K_{GM}\nabla\overline{\eta} \ , \tag{37}$$

where K_{GM} is an interfacial diffusivity typically $O(1000\ m^2s^{-1})$. Finally, the divergence of the interface eddy PV flux (which actually appears in the mean PV budget) becomes

$$\overline{J(\psi_i', \eta')} = \nabla \cdot \overline{\mathbf{u}_i'\eta'} = -\nabla \cdot K_{GM}\nabla\overline{\eta} \ . \tag{38}$$

The equivalent to $\overline{J(\psi_i', \eta')}$ in case the eddy components are defined via spatial filtering is $\mathscr{R}_i^B$ (Eq. (20)) . That is, in our case the GM parameterisation reads

$$\tilde{\mathscr{R}}_i^B = -\frac{(-1)^{i-1}f_0}{H_i}\nabla \cdot K_{GM}\nabla[\eta] \ , \tag{39}$$

such that the unresolved buoyancy fluctuations are represented as local interfacial diffusion. Inserting Eq. (39) into Eq. (34) and assuming a spatially constant $K_{GM} > 0$ we get

$$\tilde{C}(\mathrm{PE}', [\mathrm{PE}]) = \rho_0 g' K_{GM} \int [\eta] \nabla^2 [\eta] \; dA \leq 0 \; . \tag{40}$$

Consequently, the GM parameterisation represents a sink of [PE] at every instant of time and, hence, excludes any backscatter (Eq. (40) actually corresponds to the kinetic energy of the large-scale baroclinic mode).

1) Constant GM diffusivity

A constant K_{GM} can be directly estimated from the energetics of the reference simulation by combining the temporal average (denoted by an overbar) of Eq. (34), shown in Fig. 4, and the temporal average of Eq. (40), such that[7]

$$K_{GM} = \frac{\overline{C(\mathrm{PE}', [\mathrm{PE}])}}{\rho_0 g' \int \overline{[\eta] \nabla^2 [\eta]} \; dA} \; . \tag{41}$$

This way the GM parameterisation accounts exactly for the time-mean [PE] dissipation, given the reference large-scale flow. For our model results we get a typical value of $K_{GM} \approx 1067$ m^2/s.

Figure 6a shows time series of PE resulting from simulations in which the GM parameterisation with a constant K_{GM} is employed (blue and green lines). In order to assure numerical stability $K_{GM} \geq 1500$ m^2/s is necessary in our model[8]. The GM parameterisation does its job by extracting [PE] from the large-scale flow such that a statistical equilibrium results. However, the low-frequency variability exhibited by the reference simulation (black line) is absent. The dynamics exclusively reside below the PE-level of the low-PE regime of the reference simulation. That is, the low-frequency transitions in phase space to the high-PE regime are suppressed in case the

---

[7]Note that this estimation is not affected by rotational eddy fluxes since it is not computed on the level of fluxes (like Eq. (37)) but on the level of dynamical terms appearing in the PV budget.

[8]Note that the model blows up if $\tilde{\mathscr{R}}_i^B$ is simply set to zero, consistent with the fact that $\mathscr{R}_i^B$ acts as a sink of time-mean [PE] (see Fig. 4).

GM parameterisation with a constant K_{GM} is used. Presumably, backscatter is necessary for the dynamics in order to be able to reach high-PE states.

2) Time-dependent GM diffusivity

For a time-dependent K_{GM} the GM parameterisation (Eq. (39)) reads

$$\tilde{\mathscr{R}}_i^B = -\frac{(-1)^{i-1} f_0}{H_i} K_{GM}(t) \nabla^2[\eta] \, . \tag{42}$$

We diagnose K_{GM} from the model results via projection on $\nabla^2[\eta]$ (equivalent to a least-squares estimation), that is,

$$K_{GM}(t) = -\frac{H_i}{(-1)^{i-1} f_0} \frac{\int \mathscr{R}_i^B \nabla^2[\eta] \, dxdy}{\int (\nabla^2[\eta])^2 \, dxdy} \, , \tag{43}$$

where $\mathscr{R}_i^B$ represents the explicitly computed residual PV flux. That is, K_{GM} represents the expansion coefficient of $\mathscr{R}_i^B$ in $\nabla^2[\eta]$ (see also section 3c, Eq. (44)).

Figure 6a shows the time series of PE resulting from a simulation in which the GM parameterisation is employed with K_{GM} obtained from projection (i.e. Eq. (43)) at every time step (red line). Now a form of low-frequency variability is indeed excited but the corresponding high-PE regime resides at and below the low-PE regime of the reference simulation (black line), and the PE variability has smaller variance (see also the second row of Tab. 1). The low-frequency variability actually oscillates around the PE level of the simulation in which the GM parameterisation with $K_{GM} = 1500$ m^2/s is employed (blue line). This is consistent with the fact that the time-mean of K_{GM} obtained from projection is given by $\overline{K_{GM}} \approx 1585$ m^2/s (see the second row of Tab. 1 and also the next paragraph). Consequently, also in case of the GM parameterisation with a time-dependent K_{GM} obtained from Eq. (43) the transitions in phase space to the high-PE regime of the reference simulation are not captured.

Figure 6b shows the estimated pdf of K_{GM} computed from Eq. (43) and either employed in the GM parameterisation (red) or just diagnosed from the reference simulation (black). Both K_{GM}-distributions are unimodal and slightly positively skewed. Most strikingly, however, is that in both cases K_{GM} captures a significant amount of negative values. Negative K_{GM}-values are not consistent with a diffusion model. Hence, the low-frequency variability (red line in Fig. 6a) presumably emerges from the wrong reason, namely, backscatter due to a negative diffusivity. We also note that the temporal average and standard deviation of K_{GM} is significantly smaller when only diagnosed from the reference simulation (448 ± 697 m^2/s) than when the GM parameterisation is actually applied (1585 ± 1374 m^2/s, see also the second row of Tab. 1). We discuss this difference in detail in the next section (subsections c3,4).

Finally, we emphasise that the relative error $||\mathscr{R}_i^B - \tilde{\mathscr{R}}_i^B||/||\mathscr{R}_i^B||$ of the GM parameterisation is about 97% and hence extremely high (see the second row of Tab. 1). This holds for when the GM parameterisation is employed as well as for when the GM parameterisation is just diagnosed from the reference simulation (see also Fig. 7a discussed below in subsections c3,4).

c. Dynamical spatial mode representation of the eddy forcing based on energetics

It is well-known that the diffusive closure approach is limited since eddies also act up-gradient in geophysical turbulence (Starr 1968; Berloff 2005a), implying energy transfer from the eddy components to the large scale (i.e. backscatter, see Fig. 5c,e). Consequently, instead of aiming for an improved turbulent diffusion closure (e.g. via a spatially/temporally/stochastically varying eddy diffusivity tensor) we seek for additional dynamical large-scale spatial fields (next to the large-scale isopycnal gradient) to represent the eddy forcing more adequately. That is, in order to extend or replace the GM parameterisation we think in terms of a dynamical[9] spatial mode expansion of the eddy forcing,

$$\tilde{\mathscr{R}}_i^B(\mathbf{x},t) = \sum_{k=1}^{l} \xi_k(t)\varphi_k(\mathbf{x},t) \ , \tag{44}$$

with time-dependent spatial modes $\varphi_k(\mathbf{x},t)$, and evolution coefficients $\xi_k(t)$. The GM parameterisation (42) represents a special case with $l = 1$, $\xi_1 = -\frac{(-1)^{i-1} f_0}{H_i} K_{GM}(t)$, and $\varphi_1 = \nabla^2[\eta]$.

Optimally, the spatial modes $\varphi_k(\mathbf{x},t)$ can be efficiently obtained from terms of the large-scale flow equations, and the evolution coefficients $\xi_k(t)$ have clear dynamical or statistical properties such that they may be modelled deterministically or stochastically, $\xi_k(t) \to \xi_k(t;\omega)$. Also a small set of modes should be sufficient in order to assure feasibility. However, dynamical modes are typically constructed via generalised eigenproblems (e.g. linear instability modes (Dijkstra 2005; Berloff 2005b; Shevchenko et al. 2016)) or optimisation problems (e.g. Lyapunov vectors, CNOPs (Dijkstra 2013; Dijkstra and Viebahn 2015)) and, hence, are generally expensive to compute, if at all.

1) Specification of spatial energy modes

In this study, we explore whether spatial fields that stem from the large-scale energetics can suit as *dynamical spatial modes* $\varphi_k(\mathbf{x},t)$ (as in Eq. (44)) to parameterise the eddy forcing. More precisely, we focus on large-scale available potential energy budget, Eq. (27), since the eddy forcing related to the baroclinic route, $\mathscr{R}_i^B$, directly appears therein. The capital letters in Eq. (27) denote globally integrated LEC-terms. In order to express the respective LEC-terms as spatially extended PV fields (i.e. as integral kernels of the globally integrated energetics) we use lower-case letters.

[9]Dynamical modes are time-dependent and budget-based in the sense that their computation explicitly involves the governing conservation equations (see e.g. Dijkstra (2016)). In contrast, for example, statistical modes (e.g. EOFs) are data-based and not budget-based.

We then have

$$\partial_t[\mathrm{pe}] = c(\mathrm{pe}', [\mathrm{pe}]) - c([\mathrm{pe}], [\mathrm{ke}]) \quad \Leftrightarrow \quad c(\mathrm{pe}', [\mathrm{pe}]) = \partial_t[\mathrm{pe}] + c([\mathrm{pe}], [\mathrm{ke}]) \;, \tag{45}$$

with

$$\begin{aligned}
\partial_t[\mathrm{pe}] &= -\frac{1}{R^2}\partial_t[\eta] \;, && (46)\\
c(\mathrm{pe}', [\mathrm{pe}]) &= \frac{f_0}{g'}\left(\mathscr{R}_1^B - \mathscr{R}_2^B + \frac{1}{R^2}J([\psi_1], [\psi_2])\right) = && (47)\\
&= \frac{f_0}{g'R^2}[J(\psi_1, \psi_2)] \;, && (48)\\
-c([\mathrm{pe}], [\mathrm{ke}]) &= \frac{f_0}{g'}\left(\mathscr{R}_1^H - \mathscr{R}_2^H + J([\psi_1], \nabla^2[\psi_1])] - J([\psi_2], \nabla^2[\psi_2]))\right) - \\
&\quad -\beta\partial_x[\eta] - \partial_t\nabla^2[\eta] + A_H\nabla^4[\eta] + \frac{f_0[\partial_y\tau^x]}{\rho_0 H_1 g'} = && (49)\\
&= \frac{f_0}{g'}\left([J(\psi_1, \nabla^2\psi_1)] - [J(\psi_2, \nabla^2\psi_2)]\right) - \beta[\partial_x\eta] - [\partial_t\nabla^2\eta] + \\
&\quad +A_H\nabla^4[\eta] + \frac{f_0[\partial_y\tau^x]}{\rho_0 H_1 g'} \;. && (50)
\end{aligned}$$

Note that multiplication of Eq. (45) with $-\rho_0 g' R^2[\eta]$ and global integration gives Eq. (27). For the 2-layer model considered in this study, Eq. (45) corresponds to the PV evolution equation of the large-scale interface displacement $[\eta]$ (i.e., first baroclinic mode) which is obtained by subtracting Eq. (13) from Eq. (14). Combining Eq. (20) with Eq. (48) gives

$$\begin{aligned}
\mathscr{R}_i^B &= \frac{(-1)^{i-1}}{H_i}\left(f_0R^2c(\mathrm{pe}', [\mathrm{pe}]) - \frac{f_0^2}{g'}J([\psi_1], [\psi_2])\right) = && (51)\\
&\overset{Eq.\ (45)}{=} \frac{(-1)^{i-1}}{H_i}\left(f_0R^2(\partial_t[\mathrm{pe}] + c([\mathrm{pe}], [\mathrm{ke}])) - \frac{f_0^2}{g'}J([\psi_1], [\psi_2])\right) \;, && (52)
\end{aligned}$$

such that $\mathscr{R}_i^B$ is solely expressed in terms of large-scale (i.e. filtered) quantities.

This motivates us to consider the following two types of *dynamical spatial energy modes*

$$\varphi_\tau^r(\mathbf{x}, t) := \partial_t[\mathrm{pe}](\mathbf{x}, t-\tau) \quad \text{and} \quad \varphi_\tau^c(\mathbf{x}, t) := c([\mathrm{pe}], [\mathrm{ke}])(\mathbf{x}, t-\tau) \;, \tag{53}$$

where φ_τ^r is related to the temporal change of the APE reservoir at previous time $t - \tau < t$, and φ_τ^c is related to the conversion between large-scale APE and KE at previous time $t - \tau < t$. Here τ

represents the lag relative to the current time t. Note that $\partial_t[\mathrm{pe}]$ and $c([\mathrm{pe}],[\mathrm{ke}])$ are not available in a numerical model at time t but are given only after the equations of motion are solved. Additionally, the temporal derivatives of the spatial energy modes φ^r_τ and φ^c_τ can be considered (since e.g. these may improve the convergence behaviour of Eq. (44) analogous to a Taylor expansion).

Hence, in terms of a numerical model with a discrete time step Δt the overall set of spatial energy modes reads

$$\Phi := \bigcup_{k,l=1}^{\infty} \{\partial_t^{l-1}\varphi^r_{k\Delta t}, \partial_t^{l-1}\varphi^c_{k\Delta t}\} \,. \tag{54}$$

The set Φ is obviously infinite. Moreover, the energy modes are generally non-orthogonal. Note that the eddy forcing of the previous time step, i.e. $\mathscr{R}^B_i(\mathbf{x}, t-\Delta t)$, is exactly given via the energy fields $\varphi^r_{\Delta t}$ and $\varphi^c_{\Delta t}$ (see Eq. (52)). Consequently, it is essentially the increment of the eddy forcing, $\mathscr{R}^B_i(\mathbf{x},t) - \mathscr{R}^B_i(\mathbf{x}, t-\Delta t)$, that has to be modelled by Eq. (44) with energy modes.

2) Selection of finite subset of spatial energy modes

In order to compute a dynamical spatial mode expansion of the eddy forcing as in Eq. (44) in a numerical model one has to select a finite subset of energy modes out of Φ. A detailed analysis of (finding) the optimal subset of energy fields is a topic for future research (see discussion section 5). In this study we investigate the following subsets of Φ,

$$\Phi^n_{\Delta\tau} := \{\varphi^r_{\Delta t}, \varphi^c_{\Delta t}\} \bigcup_{k=1}^{n} \{\partial_t \varphi^r_{(\Delta t+(k-1)\Delta\tau)}, \partial_t \varphi^c_{(\Delta t+(k-1)\Delta\tau)}\} \,, \tag{55}$$

where $\Delta\tau$ represents the lag step size, and n determines the cardinality of $\Phi^n_{\Delta\tau}$, given by $|\Phi^n_{\Delta\tau}| = 2+2n$. Note that for each $\Phi^n_{\Delta\tau}$ the contained energy modes vary in time but $\Delta\tau$ and n are fixed.

In other words, the subspace spanned by $\Phi^n_{\Delta\tau}$ is enlarged by increasing n which corresponds to additionally including realisations of the fields $\partial_t\varphi^r, \partial_t\varphi^c$ further in the past. Enlarging the subspace used to approximate the eddy forcing by field realisations further in the past is a form of delay embedding (Takens 1981). Moreover, it is motivated by the Mori-Zwanzig formalism which demonstrates that the representation of unresolved physics includes (the estimation of) a memory term that involves the past history of the resolved physics (Wouters and Lucarini 2013; Gottwald et al. 2016). The possible relevance of the flow history for ocean eddy parameterizations has also been pointed out recently by Bachman et al. (2018) in the context of a non-Newtonian fluid mechanics approach to eddy parameterization.

More precisely, in the following sections we investigate the convergence behaviour of the following dynamical spatial mode expansion of the eddy forcing,

$$\begin{aligned}\tilde{\mathscr{R}}_i^{\Phi^n_{\Delta\tau}} \quad := \quad & \xi_0^J(t)J([\psi_i],\frac{(-1)^{i-1}f_0}{H_i}[\eta]) + \xi_0^r(t)\varphi^r_{\Delta t} + \xi_0^c(t)\varphi^c_{\Delta t} + \\ & + \sum_{k=1}^{n}\left[\xi_k^r(t)\partial_t\varphi^r_{(\Delta t+(k-1)\Delta\tau)} + \xi_k^c(t)\partial_t\varphi^c_{(\Delta t+(k-1)\Delta\tau)}\right]. \end{aligned} \tag{56}$$

In this study we consider $0 \leq n \leq 16$ and $\Delta\tau \in \{3\text{hrs}, 6\text{hrs}, 12\text{hrs}\}$. For completeness we also include the large-scale Jacobian, $J([\psi_i], \frac{(-1)^{i-1}f_0}{H_i}[\eta])$, in the expansion (see Eq. (52)). For each choice of n and $\Delta\tau$ the expansion (56) represents a parameterisation of the eddy forcing $\mathscr{R}_i^B$.

The expansion coefficients in Eq. (56) are computed at each model time step by using ordinary least squares with respect to $\mathscr{R}_i^B$. Analysing the dynamical and statistical behaviour of the expansion coefficients as well as proposing a (possibly stochastic) model for the expansion coefficients in order to build a fully self-consistent closure is a topic for future research (see discussion section 5). Here the aim is to investigate how well the expansion (56) approximates (converges to) $\mathscr{R}_i^B$.

Finally, we contrast the convergence behaviour of Eq. (56) in two ways. We first consider the similar dynamical spatial mode expansion,

$$\begin{aligned}\tilde{\mathscr{R}}_i^{\Phi^n_{\Delta\tau},GM} \quad := \quad & \xi_0^J(t)J([\psi_i],\frac{(-1)^{i-1}f_0}{H_i}[\eta]) + \xi_0^{GM}(t)\nabla^2[\eta] + \xi_0^r(t)\varphi^r_{\Delta t} + \xi_0^c(t)\varphi^c_{\Delta t} + \\ & + \sum_{k=1}^{n}\left[\xi_k^r(t)\partial_t\varphi^r_{(\Delta t+(k-1)\Delta\tau)} + \xi_k^c(t)\partial_t\varphi^c_{(\Delta t+(k-1)\Delta\tau)}\right], \end{aligned} \tag{57}$$

where the GM term (see Eq.(42)) is additionally included in the expansion. In this way we investigate the impact of the GM field on the convergence behaviour.

In addition, we also analyse the convergence behaviour of the spatial filter modes χ_i as given[10] in section 2b. That is, we consider the same expansions as in (56) and (57) but instead of using the energy modes $\Phi^n_{\Delta\tau}$ we use the filter modes χ_i (ordered by decreasing eigenvalue/wavenumber). The expansions read

$$\tilde{\mathscr{R}}_i^{\mathscr{F}^n} := \xi_0^J(t)J([\psi_i], \frac{(-1)^{i-1}f_0}{H_i}[\eta]) + \sum_{k=1}^{n} \xi_k^f(t)\chi_k \,, \tag{58}$$

$$\tilde{\mathscr{R}}_i^{\mathscr{F}^n,GM} := \xi_0^J(t)J([\psi_i], \frac{(-1)^{i-1}f_0}{H_i}[\eta]) + \xi_0^{GM}(t)\nabla^2[\eta] + \sum_{k=1}^{n} \xi_k^f(t)\chi_k \,. \tag{59}$$

Again, the expansion coefficients are computed at each model time step by using ordinary least squares with respect to $\mathscr{R}_i^B$.

3) Approximation of the eddy forcing on the reference attractor

In the following we analyse the approximation of the eddy forcing $\mathscr{R}_i^B$ by the different series expansions defined in the previous section (see Eq. (56)-(59)). In this section the terms in Eq. (56)-(59) are diagnosed from the reference simulation (described in section 2a). That is, the replacement $\mathscr{R}_i^B \to \tilde{\mathscr{R}}_i^B$ (see section 3a) is not applied in the simulation and, hence, the state vector is always on the attractor of the reference ER model.

(i) Relative error of eddy forcing. Figure 7a shows the time-mean relative error of the eddy forcing for the filter mode expansion either with GM term (black, Eq. (59)) or without GM term (red, Eq. (58)). The two curves are almost identical which demonstrates that the GM term is not able to significantly reduce the relative error of the eddy forcing. In other words, the GM field (as a direction in phase space) is largely orthogonal to the eddy forcing field. For both

[10] Loosely speaking, these are Fourier-type modes. More precisely, the spatial filter modes χ_i are eigenmodes of the Bilaplacian in this study.

curves the decrease in relative error (i.e. the slope of the curve) is minimal at the beginning and monotonically increasing with increasing number of filter modes. The value of the relative error for the filter modes is in the order of 10^{-1} and only reaches a very small value (i.e. $O(10^{-13})$) when all filter modes are used. That is, the convergence of Eq. (58)-(59) is slow.

Figure 7b shows the time-mean relative error of the eddy forcing for the energy mode expansion Eq. (57) with $\Delta\tau = 3$hrs (blue), $\Delta\tau = 6$hrs (black), $\Delta\tau = 12$hrs (magenta). Note the logarithmic scale on the ordinate. The effect of the GM term on the relative error is again very small such that the curves related to Eq. (56) are indistinguishable from the shown curves. In contrast to the filter modes, the decrease in relative error (i.e. the slope of the curve) is maximal at the beginning and monotonically decreasing with increasing number of energy modes (for comparison the curve of the filter modes is shown by the blue dashed line). With only 4 energy modes used in Eq. (56) the relative error drops to $O(10^{-5})$ and for $\Delta\tau = 3$hrs a relative error of $O(10^{-12})$ is reached with 30 energy modes. That is, the convergence of Eq. (56)-(57) is very fast since adding energy modes reduces the order of magnitude of the relative error. Finally, it holds for the reference simulation that the smaller $\Delta\tau$ the smaller the relative error.

(ii) GM diffusivity. Figure 8a shows the time-mean GM diffusivity K_{GM} for the filter mode expansion with GM term (blue, Eq. (59)). The GM diffusivity K_{GM} decreases nearly linearly due to the subsequent inclusion of more and more filter modes. However, the value of K_{GM} remains in the order of 100 m^2/s. Only when almost all filter modes are included the value of K_{GM} becomes small and, hence, the impact of the GM term is insignificant.

Figure 8b shows the time-mean GM diffusivity K_{GM} for the energy mode expansion with GM term (Eq. (57)) with $\Delta\tau = 3$hrs (blue), $\Delta\tau = 6$hrs (black), $\Delta\tau = 12$hrs (magenta). The behaviour of K_{GM} resembles the behaviour of the relative error (Fig. 7b). Note again the logarithmic scale on the ordinate. Including energy modes drastically reduces the value K_{GM}, that is, by orders of magnitude. With only 4 energy modes used in Eq. (57) the value of K_{GM} drops to $O(10^{-3}m^2/s)$ indicating that the GM term is essentially without impact. Finally, it holds for the reference simulation that the smaller $\Delta\tau$ the smaller K_{GM}.

4) Approximation of the eddy forcing in the presence of error perturbations

In this section we analyse the approximation of the eddy forcing $\mathscr{R}_i^B$ by the different series expansions defined by Eq. (56)-(59). At each time step the corresponding replacement $\mathscr{R}_i^B \to \tilde{\mathscr{R}}_i^B$ is performed (see section 3a) resulting into a different simulation for each parameterisation (e.g. series expansion). Consequently, error perturbations due to the approximate representation of the eddy forcing $\mathscr{R}_i^B$ are introduced in each simulation and, hence, the state vector can be pushed away from the attractor of the reference simulation. If the parameterisation of the eddy forcing is accurate enough it can compensate for the error perturbations and can keep the system within or near the attractor of the reference simulation. On the other hand, if the parameterisation of the eddy forcing is not accurate enough then the respective model will exhibit a different attractor.

(i) Relative error of eddy forcing. Figure 7a shows the time-mean relative error of the eddy forcing for the filter mode expansion either with GM term (blue, Eq. (59), see also Tab. 1) or without GM term (magenta, Eq. (58)). The relative error for the simulations with error perturbations is slightly smaller than for the reference simulation (black and red curves). Nevertheless, the overall behaviour is very similar to the reference simulation: The blue and magenta curves are nearly identical which indicates that the GM term is not able to significantly reduce the relative error of the eddy forcing. For both curves the decrease in relative error (i.e. the slope of the curve) is minimal at the beginning and monotonically increasing with increasing number of filter modes.

The convergence of Eq. (58)-(59) is slow since the value of the relative error is in the order of 10^{-1} and only reaches a very small value (i.e. $O(10^{-13})$) when all filter modes are used.

A crucial point in Fig. 7a is that the filter mode expansion without GM term (magenta, Eq. (58)) leads to a model blow-up if less than 10 filter modes are used (the magenta curve only starts at $\#modes = 10$). On the other hand, the filter mode expansion with GM term (blue, Eq. (59)) leads to stable model simulations for any number of filter modes (see also Tab. 1). Hence, the effect of the GM term becomes clearer: the GM term cannot not significantly reduce the relative error of the eddy forcing but it can stabilise the model. In dynamical systems terms the GM term acts as a stablising direction in phase space. That is, the GM term cannot direct the system's state along the attractor (it cannot excite the intrinsic low-frequency variability transitions in phase space as done by unstable directions) but it mainly keeps the system from diverging.

Figure 7b shows the time-mean relative error of the eddy forcing for the energy mode expansion Eq. (57) with $\Delta\tau = 3$hrs (red, see also Tab. 1), $\Delta\tau = 6$hrs (green), $\Delta\tau = 12$hrs (cyan). Note the logarithmic scale on the ordinate. The effect of the GM term on the relative error is again very small such that the curves related to Eq. (56) are indistinguishable from the shown curves. On the other hand, the stabilising effect of the GM term also appears for the energy modes: for the application of Eq. (56) (i.e. energy mode expansion without GM term) with only 2 energy modes the model blows up whereas for the application of Eq. (57) (i.e. energy mode expansion with GM term) the model is stable.

The overall behaviour of the relative error for the simulations with error perturbations (red, green, cyan) is similar to the results of the reference simulation (blue, black, magenta). That is, the relative error decreases much faster (adding energy modes reduces the order of magnitude of the relative error) than for the filter modes (shown for comparison by the blue dashed line). However, due to the induced error perturbations the decrease in relative error is weaker than for the reference simulation. For example, for 30 energy modes and $\Delta\tau = 3$hrs the relative error is $O(10^{-5})$ instead of $O(10^{-12})$ for the reference simulation. Moreover, the impact of $\Delta\tau$ is more complicated than for the reference simulation. Roughly speaking, if less than 20 energy modes are used in Eq. (56) or Eq. (57) then the relative error is slightly smaller for larger $\Delta\tau$ whereas if more than 20 energy modes are used then the situation of the reference situation is reencountered (i.e. the smaller $\Delta\tau$ the smaller the relative error).

(ii) GM diffusivity. Figure 8a shows the time-mean GM diffusivity K_{GM} for the filter mode expansion with GM term (red, Eq. (59), see also Tab. 1). The behaviour is largely similar to the results of the reference simulation (blue), namely, the GM diffusivity K_{GM} decreases nearly linearly due to the subsequent inclusion of more and more filter modes. However, the value of K_{GM} is significantly larger (about one order of magnitude) than when diagnosed from the reference simulation. This in accordance with the interpretation of the GM term as a stabilising direction in phase space because in the presence of error perturbations (driving the system away from the attractor) the eddy forcing will project more on stable directions (driving the system back to the attractor). In other words, in the presence of error perturbations the GM term has work to do.

Figure 8b shows the time-mean GM diffusivity K_{GM} for the energy mode expansion with GM term (Eq. (57)) with $\Delta\tau = 3$hrs (red, see also Tab. 1), $\Delta\tau = 6$hrs (green), $\Delta\tau = 12$hrs (cyan). The behaviour is largely similar to the results of the reference simulation (blue, black, magenta), namely, including energy modes drastically reduces the value K_{GM}, that is, by orders of magnitude. It also largely holds that the smaller $\Delta\tau$ the smaller K_{GM}. On the other hand, the value of K_{GM} is larger than when diagnosed from the reference simulation (i.e. the stabilising direction projects on the error perturbations). Nevertheless, the value of K_{GM} is still significantly smaller ($O(1)$ for only 4 energy modes) compared to the values typically used in ocean models ($O(1000)$).

(iii) Time series of potential energy. Figure 9 shows time series of PE related to simulations employing the filter mode expansion with GM term (red, Eq. (59), see also Tab. 1) and the energy mode expansion with GM term and $\Delta\tau = 3$hrs (blue, Eq. (57), see also Tab. 1). For comparison the time series of the PE of the reference simulation is shown in black.

The energy mode expansion exhibits monotonic and fast convergence behaviour in terms of PE (i.e. low-frequency variability). If only 2 energy modes are used (panel d) the PE variability is still significantly different from the reference PE. Intense low-frequency variability is present but it is situated between the low-PE regime of the reference simulation and another very-low-PE regime. Already with 4 energy modes in the expansion the high-PE regime of the reference simulation is regularly reached (not shown). But the low-PE regime is still bit lower than for the reference case. For 6 $\geq$ energy modes (panels f, h, j) the PE variability of the reference simulation appears to be essentially recovered.

As expected, the situation is different for the filter mode expansions. The convergence behaviour is non-monotonic. Even for 20 filter modes (panel i) the PE variability is significantly different from the reference simulation. When using filter modes it appears to be difficult to reach the high-PE regime of the reference simulation. Either the PE variance is significantly smaller than for the reference simulation (panels c, g) or the low-PE regime is lower than for the reference simulation (panels e, i). This is also visible in Tab. 1.

4. Summary

The three key points of this study can be summarized as follows: First, we propose a new approach to parameterising sub-grid scale processes. In this approach the impact of the unresolved dynamics on the resolved dynamics, that is the eddy forcing, is represented by a series expansion in dynamical spatial modes that stem from the energy budget of the resolved dynamics. More precisely, the so-called energy modes are directly obtained from the equations of motion of the resolved flow by identifying the integral kernels that lead to the different reservoir, generation, dissipation, and conversion terms in the large-scale energy budget. Hence, the energy modes exhibit strictly large-scale patterns and they are equipped with a clear physical interpretation in terms of energetics. Convergence towards the eddy forcing is accomplished via delay embedding by including additional realisations of these fields further in the past. We also note the relation to the Mori-Zwanzig formalism which indicates that the representation of unresolved physics needs to include a memory term that involves the past history of the resolved physics. For the 2-layer QG ocean model considered in this study, we demonstrate that the convergence of a series expansion in the energy modes is by orders of magnitude faster than the convergence of a series expansion in Fourier-type modes. That is, the eddy forcing can be accurately approximated with a very limited number of energy modes which enables a feasible parameterisation.

Second, we explore a novel way to test parameterisations in models. The resolved dynamics and the corresponding instantaneous eddy forcing are defined via spatial filtering which accounts for the representation error of the equations of motion on the low-resolution model grid. In this way closures can be tested within the high-resolution model. Whereas in low-resolution models all energy pathways between large-scale and eddy components must be parameterised simultaneously, testing parameterisations in the high-resolution model offers the possibility to isolate the effects of a single parameterisation (related to a single energy pathway) while the other large-scale eddy energy conversions are correctly computed. For the 2-layer QG ocean model considered in this study, we focus on parameterisations of the baroclinic energy pathway while the barotropic energy pathway is correctly computed by the high-resolution model.

Third, we test the standard closure of the baroclinic energy pathway in the ocean components of state-of-the-art climate models, i.e. the Gent-McWilliams (GM) parameterisation with a scalar diffusivity, in the high-resolution QG ocean model considered in this study. It turns that the GM field steers trajectories along a stabilising direction in phase space. That is, the GM field does not project well on the eddy forcing (it exhibits a very high relative error) and fails to excite the model's intrinsic low-frequency variability (i.e. it is not able to propagate the model's state along the correct attractor e.g. along an unstable direction). The GM field mainly stabilises the model. That is, if the representation of the eddy forcing is very inaccurate (e.g. small number of modes used in expansion) the GM term performs the necessary dissipation of available potential energy such that the model does not diverge.

5. Discussion

Finally, we elaborate on open issues of this study and related future research directions:

Self-consistent closure of the baroclinic energy pathway. A closure of the baroclinic energy pathway is self-consistent if it does not involve the actual ("true") baroclinic eddy forcing. However, in this study we still use $\mathscr{R}_i^B$ for the computation of expansion coefficients (i.e. the coefficients that appear in a spatial mode expansion) via ordinary least squares. Determining a self-consistent closure of the baroclinic energy pathway is related to three intricate and intimately related issues: (i) determining the optimal subset of energy fields (see Eq. (54)), (ii) diagnosing the corresponding expansion coefficients, and (iii) proposing a (possibly stochastic) self-consistent model for the expansion coefficients. The choices made with respect to $(i)-(iii)$ can have an effect on the accuracy of the approximation (as indicated in this study by the different choices for $\Delta\tau$), the computational cost and complexity of the model, the regularity of the expansion coefficients, and the uniqueness and hence physical interpretation of the series expansion in energy modes.

For example, a problem related to these issues and well-known in statistics and machine learning is the issue of overfitting-versus-underfitting or the bias-variance tradeoff. Low-bias approaches can usually give accurate representation of the data but produce large variances. In contrast, models with higher bias produce lower variances but less accurate representations. Regularization methods introduce bias into the regression solution that can reduce variance considerably. In this way the behaviour of the expansion coefficients becomes simpler and easier to model but the approximation becomes less accurate.

Self-consistent closure of the barotropic energy pathway. In order to make the equations for the large-scale flow (Eq. (13)-(14)) completely self-consistent one also has to specify a self-consistent closure of the barotropic energy pathway (i.e. $\mathscr{R}_i^H$). The standard closure of the barotropic energy pathway is lateral viscous dissipation with an enhanced 'eddy' viscosity coefficient. Similar to the GM parameterisation the lateral viscosity parameterisation suffers from the lack of backscatter (see Fig. 5). But an adequate closure of the energy exchange between large-scale and eddy components is necessary in order to be able to perform low-resolution model simulations exhibiting eddy-driven low-frequency variability. One option is to proceed in a way similar to this study: explore whether spatial fields that stem from the large-scale kinetic energy budget can suit as dynamical modes to parameterise the eddy forcing $\mathscr{R}_i^H$.

Dynamical systems analysis of the large-scale flow in the turbulent regime. As soon as adequate closures for both the baroclinic and the barotropic energy pathways are available it is in principle possible (i.e. feasible due to low model resolution) to analyse the dynamics of the large-scale flow in the turbulent regime in a systematic way. In case of deterministic closures this is related to the existence of multiple equilibria, stability properties, bifurcations and chaotic attractors (Dijkstra 2005). In case of stochastic closures the investigation will be from the perspective of random dynamical systems which is related to stochastic bifurcations (i.e. changes in the probability density function), pullback attractors and invariant measures (Dijkstra 2013). We note that in case of low model resolutions a whole set of numerical techniques to investigate transitions in stochastic dynamical systems becomes feasible (Dijkstra et al. 2016). For example, it becomes possible to numerically solve the SPDEs via dynamical mode expansions (Sapsis and Lermusiaux 2009) and to investigate the interaction of external noise forcing with internal nonlinear variability in the turbulent regime (Sapsis and Dijkstra 2013).

Comparison with other approaches to eddy parameterization. In this study, we compared turbulence closures based on energy modes with the GM eddy parameterisation approach. We focussed on a positive and spatially constant K_{GM} because it is straightforward to diagnose (e.g. not entering issues around rotational eddy fluxes), and, more importantly, because a spatially homogenous K_{GM} is still regularly applied in state-of-the-art realistic ocean models. On the other hand, the estimation and performance of a spatially inhomogeneous (and possibly tensor-valued) K_{GM} remains a crucial topic (Eden et al. 2007, 2009; Viebahn and Eden 2010). The relation between energy modes and the GM parameterisation, as well as other approaches to eddy parameterisation (Mana and Zanna 2014; Jansen and Held 2014; Bachman et al. 2018), will hopefully be further elucidated in future studies.

More realistic ocean model configurations. The ocean model considered in this study is situated at the more idealised end in the hierarchy of ocean models. Several features and processes must be included in order to make the details more realistic. These include higher vertical resolution, diabatic terms like buoyancy forcing and buoyancy sinks, and realistic topography and coastlines. We are currently extending our results to a 3-layer model including realistic topographic interactions. Eventually, one also has to consider the primitive equations in order to be able to investigate global realistic ocean models. The corresponding energy budgets are more complicated but detailed analyses are becoming available nowadays (von Storch et al. 2012; Wu et al. 2017; Jüling et al. 2018).

Climate model simulations subject to intrinsic (eddy-driven) low-frequency variability. Finally, when adequate closures for the energy pathways in realistic ocean models are available then long-period low-resolution climate model simulations exhibiting eddy-driven low-frequency variability become possible. This is crucial since then issues related to anthropogenic climate change (forced variability) versus intrinsic low-frequency variability (internal variability) can be addressed in a statistically significant manner.

Acknowledgments. The authors would like to thank Inti Pelupessy, Alexis Tantet, and Fred Wubs for helpful discussions. This research is funded by the Netherlands Organization for Scientific Research (NWO) through the Vidi project "Stochastic models for unresolved scales in geophysical flows".

References

Bachman, S. C., J. A. Anstey, and L. Zanna, 2018: The relationship between a deformation-based eddy parameterization and the LANS-α turbulence model. *Ocean Modell.*, **126**, 56–62.

Berloff, P., 2005a: On dynamically consistent eddy fluxes. *Dyn. Atmos. Oceans*, **38**, 123146.

Berloff, P., 2005b: On rectification of randomly forced flows. *J. Mar. Res.*, **63**, 497–527.

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

Berloff, P., A. Hogg, and W. Dewar, 2007: The turbulent oscillator: A mechanism of low-frequency variability of the wind-driven ocean gyres. *J. Phys. Oceanogr.*, **37**, 2363–2386.

Berloff, P., and J. C. McWilliams, 1999a: Large-scale, low-frequency variability in wind-driven ocean gyres. *J. Phys. Oceanogr.*, **29**, 1925–1949.

Berloff, P., and J. C. McWilliams, 1999b: Quasigeostrophic dynamics of the western boundary current. *J. Phys. Oceanogr.*, **29**, 26072634.

Berloff, P., J. C. McWilliams, and A. Bracco, 2002: Material transport in oceanic gyres. Part I: Phenomenology. *J. Phys. Oceanogr.*, **32**, 764–796.

Bryan, F. O., P. R. Gent, and R. Tomas, 2014: Can Southern Ocean eddy effects be parameterized in climate models. *J. Climate*, **27**, 411–425.

Deser, C., M. A. Alexander, S.-P. Xie, and A. S. Phillips, 2010: Sea surface temperature variability: Patterns and mechanisms. *Annu. Rev. Mar. Sci.*, **2**, 11543.

Dijkstra, H. A., 2005: *Nonlinear Physical Oceanography*. Springer, 532 pp.

Dijkstra, H. A., 2013: *Nonlinear Climate Dynamics*. Cambridge University Press, 367 pp.

Dijkstra, H. A., 2016: A normal mode perspective of intrinsic ocean-climate variability. *Annu. Rev. Fluid Mech.*, doi:10.1146/annurev-fluid-122414-034506.

Dijkstra, H. A., and J. P. Viebahn, 2015: Sensitivity and resilience of the climate system: A conditional nonlinear optimization approach. *Communications in Nonlinear Science and Numerical Simulation*, **22**, 13–22.

Dijkstra, H. A., and Coauthors, 2016: A numerical framework to understand transitions in high-dimensional stochastic dynamical systems. *Dynamics and Statistics of the Climate System*, doi: 10.1093/climsys/dzw003.

Eden, C., R. J. Greatbatch, and J. Willebrand, 2007: A diagnosis of thickness fluxes in an eddy-resolving model. *J. Phys. Oceanogr.*, **37**, 727–742.

Eden, C., M. Jochum, and G. Danabasoglu, 2009: Effects of different closures for thickness diffusivity. *Ocean Modell.*, **26**, 47–59.

England, M. H., and Coauthors, 2014: Recent intensification of wind-driven circulation in the Pacific and the ongoing warming hiatus. *Nature Climate Change*, **4**, 222–227.

Franzke, C., T. O'Kane, J. Berner, P. Williams, and V. Lucarini, 2015: Stochastic climate theory and modeling. *WIREs Clim Change*, **6**, 63–78.

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

Gille, S. T., and K. A. Kelly, 1996: Scales of spatial and temporal variability in the Southern Ocean. *J. Geophys. Res.*, **101**, 8759–8773.

Gottwald, G., D. Crommelin, and C. Franzke, 2016: Stochastic climate theory. *Nonlinear and Stochastic Climate Dynamics*.

Griffies, S. M., and Coauthors, 2015: Impacts on ocean heat from transient mesoscale eddies in a hierarchy of climate models. *J. Climate*, **28**, 952–977.

Grooms, I., and A. Majda, 2013: Efficient stochastic superparameterization for geophysical turbulence. *Proc. Natl. Acad. Sci. (USA)*, **110**, doi:10.1073/pnas.1302548110.

Haidvogel, D., J. C. McWilliams, and P. R. Gent, 1992: Boundary current separation in a quasi-geostrophic, eddy-resolving ocean circulation model. *J. Phys. Oceanogr.*, **22**, 882902.

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

Hogg, A., and J. Blundell, 2006: Interdecadal variability of the Southern Ocean. *J. Phys. Oceanogr.*, **36**, 1626–1645.

Hogg, A., P. D. Killworth, J. Blundell, and W. K. Dewar, 2005: Mechanisms of decadal variability of the wind-driven ocean circulation. *J. Phys. Oceanogr.*, **35**, 512–531.

IPCC, 2013: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press.

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

Jüling, A., J. Viebahn, S. Drijfhout, and H. Dijkstra, 2018: Energetics of the southern ocean mode. *Journal of Geophysical Research Oceans*, submitted.

Kerr, R. A., 2000: A North Atlantic climate pacemaker for the centuries. *Science*, **288**, 1984–1985.

Kwon, Y.-O., 2010: Role of the gulf stream and kuroshio-oyashio systems in large-scale atmosphere-ocean interaction: A review. *Journal of Climate*, **23**, 32493281.

Le Bars, D., J. P. Viebahn, and H. A. Dijkstra, 2016: A Southern Ocean mode of multidecadal variability. *Geophys. Res. Lett.*, doi::10.1002/2016GL068177.

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

Mantua, N. J., S. R. Hare, Y. Zhang, J. M. Wallace, and R. Francis, 1997: A Pacific interdecadal climate oscillation with impacts on salmon production. *Bull. Amer. Meteor. Soc.*, **78**, 1069–1079.

Meehl, G. A., J. M. Arblaster, J. T. Fasullo, A. Hu, and K. E. Trenberth, 2011: Model-based evidence of deep-ocean heat uptake during surface-temperature hiatus periods. *Nature Climate Change*, **1**, 360–364.

Meehl, G. A., A. Hu, J. Arblaster, J. Fasullo, and K. E. Trenberth, 2013: Externally forced and internally generated decadal climate variability associated with the interdecadal Pacific oscillation. *J. Climate*, **26**, 7298–7310.

Olbers, D., J. Willebrand, and C. Eden, 2012: *Ocean Dynamics*. Springer, 704 pp.

Pedlosky, J., 1996: *Ocean Circulation Theory*. Springer, 455 pp.

Pope, S., 2000: *Turbulent Flows*. Cambridge University Press.

Sapsis, T., and H. Dijkstra, 2013: Interaction of additive noise and nonlinear dynamics in the double-gyre wind-driven ocean circulation. *Journal of Physical Oceanography*, doi:10.1175/JPO-D-12-047.1.

Sapsis, T., and P. Lermusiaux, 2009: Dynamically orthogonal field equations for continuous stochastic dynamical systems. *Physica D*, **238**, 2347–2360.

Schlesinger, M. E., and N. Ramankutty, 1994: An oscillation in the global climate system of period 6570 years. *Nature*, **367**, 723–726.

Shevchenko, I. V., P. S. Berloff, D. Guerrero-Lopez, and J. E. Roman, 2016: On low-frequency variability of the midlatitude ocean gyres. *J. Fluid Mech.*, **795**, 423–442.

Siegel, A., J. B. Weiss, J. Toomre, J. C. McWilliams, P. S. Berloff, and I. Yavneh, 2001: Eddies and vortices in ocean basin dynamics. *Geophys. Res. Lett.*, **28**, 31833187.

Starr, V., 1968: *Physics of negative viscosity phenomena*. McGraw-Hill, New York, 256 pp.

Storch, H., and F. W. Zwiers, 1999: *Statistical Analysis in Climate Research*. Cambridge University Press.

Takens, F., 1981: Detecting strange attractors in turbulence. *Dynamical Systems and Turbulence, Lecture Notes in Mathematics*, **898**.

Vallis, G. K., 2006: *Atmospheric and Oceanic Fluid Dynamics*. Cambridge University Press, 745 pp.

Verheul, N., J. Viebahn, and D. Crommelin, 2017: Covariate-based stochastic parameterization of baroclinic ocean eddies. *Mathematics of Climate and Weather Forecasting*, doi: 10.1515/mcwf-2017-0005.

Viebahn, J., and C. Eden, 2010: Towards the impact of eddies on the response of the Southern Ocean to climate change. *Ocean Modell.*, **34**, 150–165.

Viebahn, J., A. von der Heydt, D. Le Bars, and H. Dijkstra, 2016: Effects of Drake Passage on a strongly eddying global ocean. *Paleoceanography*, doi:10.1002/2015PA002888.

von Storch, J.-S., C. Eden, I. Fast, H. Haak, D. Hernandez-Deckers, E. Maier-Reimer, J. Marotzke, and D. Stammer, 2012: An estimate of the Lorenz Energy Cycle for the world ocean based on the 1/10 STORM/NCEP simulation. *J. Phys. Oceanogr.*, **42**, 2185–2205.

Wouters, J., and V. Lucarini, 2013: Multi-level dynamical systems: Connecting the Ruelle response theory and the Mori-Zwanzig approach. *Journal of Statistical Physics*, **151**, 850–860.

Wu, Y., Z. Wang, and C. Liu, 2017: On the response of the Lorenz energy cycle for the Southern Oocean to intensified westerlies. *Journal of Geophysical Research*, doi:10.1002/2016JC012539.

LIST OF TABLES

TABLE 1. Temporal average and standard deviation of PE, the GM diffusivity K_{GM}, and the relative error of the eddy forcing for different model setups with $\tilde{\mathscr{R}}_i^B$ expanded in either filter modes (indicated by F(#$filter\ modes$) and based on Eq. (59)) or energy modes (indicated by E(#$energy\ modes$) and based on Eq. (57) with $\Delta\tau = 3$ hours). The values are based on daily output of about 200 years. Note that the values are still subject to small trends since for perfect convergence simulation lengths of $O(1000\ years)$ would be necessary (see also Fig. 9).

setup	PE $[PJ]$	K_{GM} $[m^2/s]$	$\|\|\mathscr{R}_i^B - \tilde{\mathscr{R}}_i^B\|\|/\|\|\mathscr{R}_i^B\|\|$
reference	967 ± 116	-	-
GM	685 ± 58.3	1585 ± 1374	0.97 ± 0.03
J-GM	743 ± 74.9	1126 ± 1140	0.80 ± 0.11
J-GM-F(2)	855 ± 56.3	1713 ± 1477	0.78 ± 0.11
J-GM-F(4)	843 ± 78.1	1931 ± 1607	0.77 ± 0.11
J-GM-F(6)	770 ± 88.9	1720 ± 1446	0.77 ± 0.11
J-GM-F(8)	803 ± 96.7	1750 ± 1511	0.77 ± 0.11
J-GM-F(10)	868 ± 74.8	1726 ± 1508	0.77 ± 0.11
J-GM-F(20)	778 ± 107	1315 ± 1223	0.73 ± 0.11
J-GM-F(30)	952 ± 104	1063 ± 1054	0.67 ± 0.12
J-GM-E(2)	782 ± 103	8.0 ± 28	0.67 ± 0.11
J-GM-E(4)	893 ± 126	2.2 ± 33	0.51 ± 0.10
J-GM-E(6)	940 ± 99.6	1.1 ± 24	0.37 ± 0.08
J-GM-E(8)	946 ± 99.9	0.7 ± 18	0.28 ± 0.07
J-GM-E(10)	959 ± 113	0.6 ± 14	0.21 ± 0.06
J-GM-E(20)	971 ± 123	0.02 ± 1	0.01 ± 10^{-3}
J-GM-E(30)	973 ± 117	$10^{-3} \pm 0.02$	$10^{-5} \pm 10^{-5}$

LIST OF FIGURES

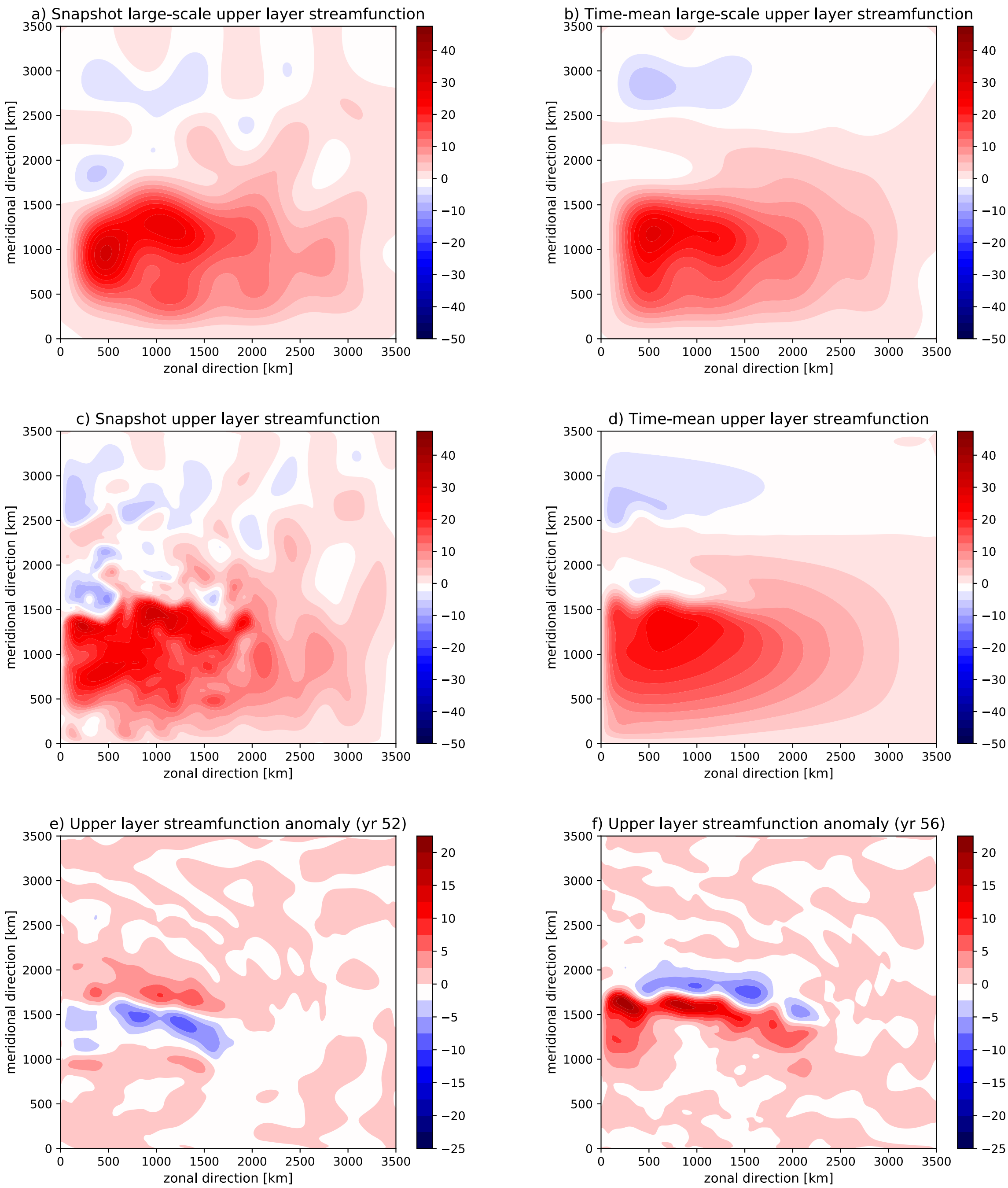


FIG. 1. Large-scale component of an upper layer streamfunction (a) snapshot and (b) time-mean. (c) and (d) show the corresponding reference (i.e. unfiltered) upper layer streamfunctions. Anomalies of the reference (i.e. unfiltered) upper layer streamfunction (with respect to the time-mean shown in d) corresponding to (e) a low and (f) a high in the low-frequency variability of PE (see Fig. 5a). The contour interval in all panels is 2.5 Sv.

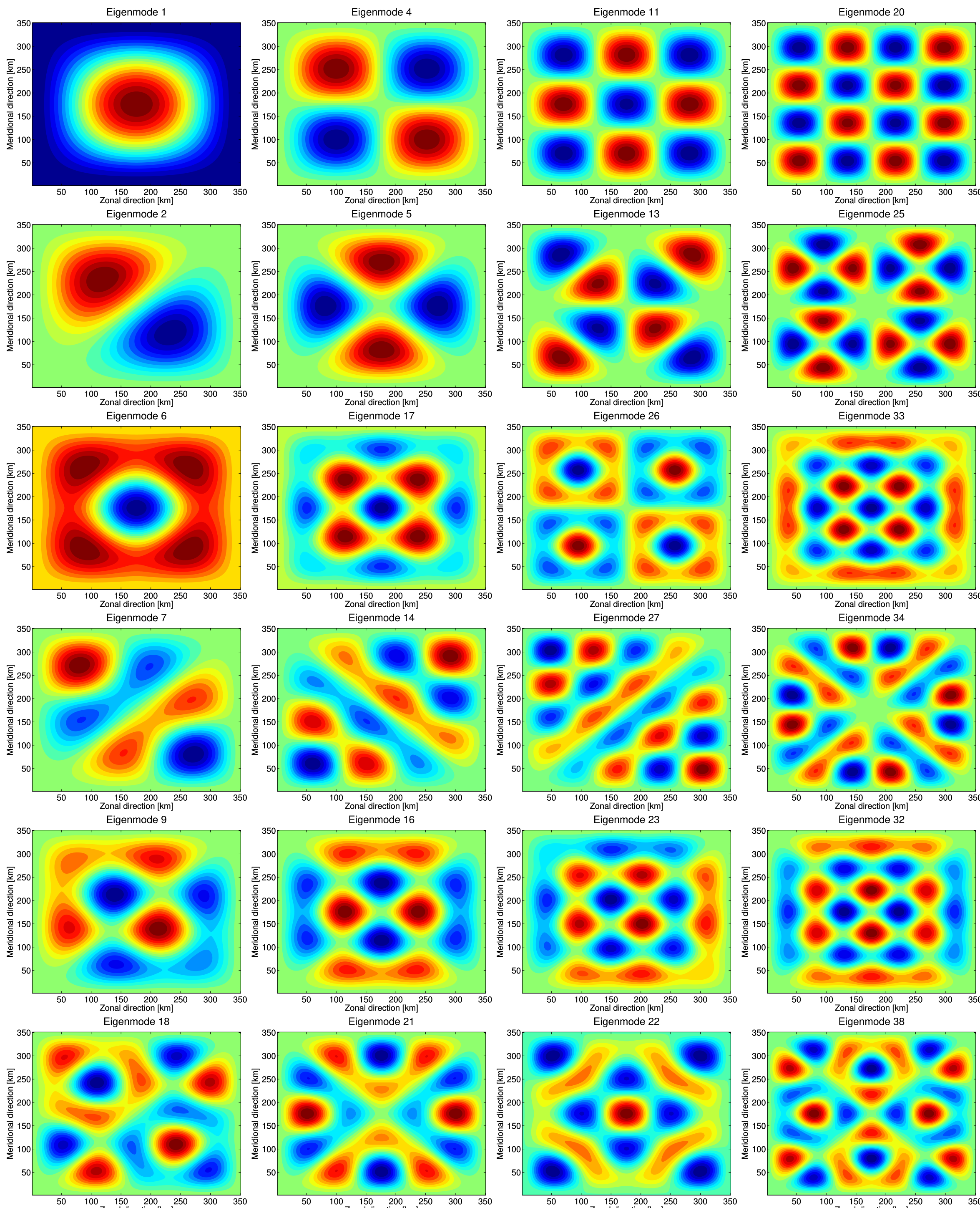


FIG. 2. Selected leading eigenmodes (ortho-normalised) of the Bilaplacian with no-slip boundary conditions computed on the high-resolution grid (i.e 349^2 grid points).

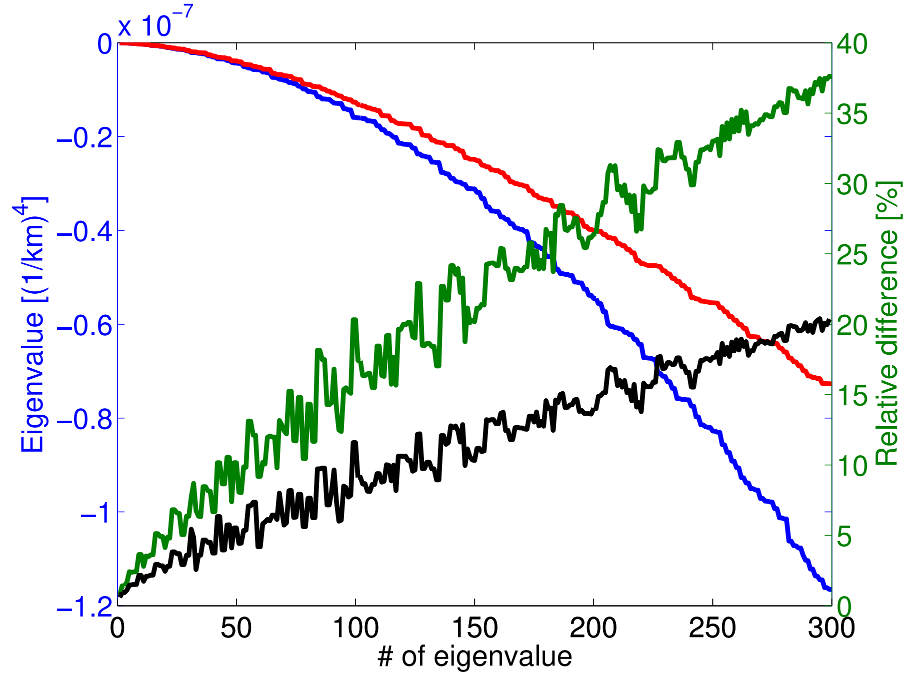


FIG. 3. First 300 eigenvalues of the Bilaplacian (left axis) for the ER model (blue) and the non-ER model (red). The relative difference (right axis), i.e. $|\left(\hat{\lambda}_i/\lambda_i - 1\right) * 100|$, is also shown (green). Moreover, the relative difference in globally integrated kinetic energy, i.e. $|\left((\int \hat{\chi}_i \nabla^2 \hat{\chi}_i \; dxdy)/(\int \chi_i \nabla^2 \chi_i \; dxdy) - 1\right) * 100|$, is shown (black). Note that the number of eigenmodes for the ER (non-ER) model is 349^2 (34^2). The Nyquist cutoff for the non-ER model is $\hat{N}_{Nyquist} = 17^2 = 289$.

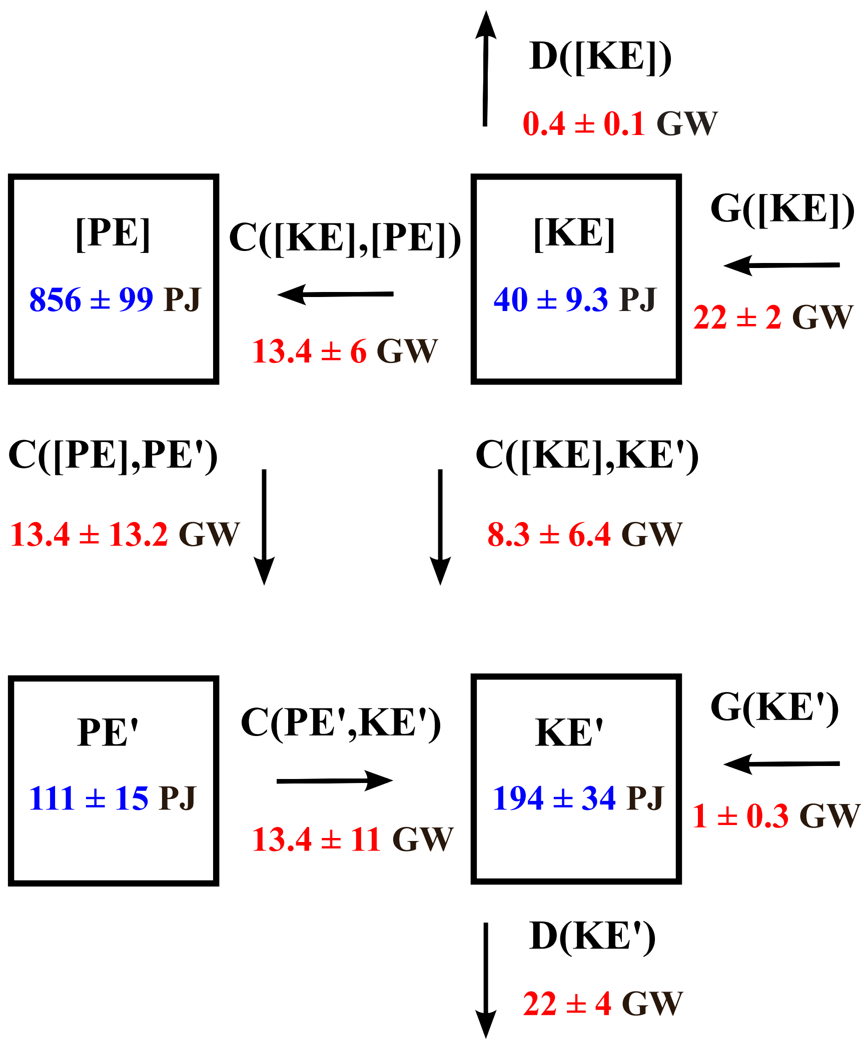
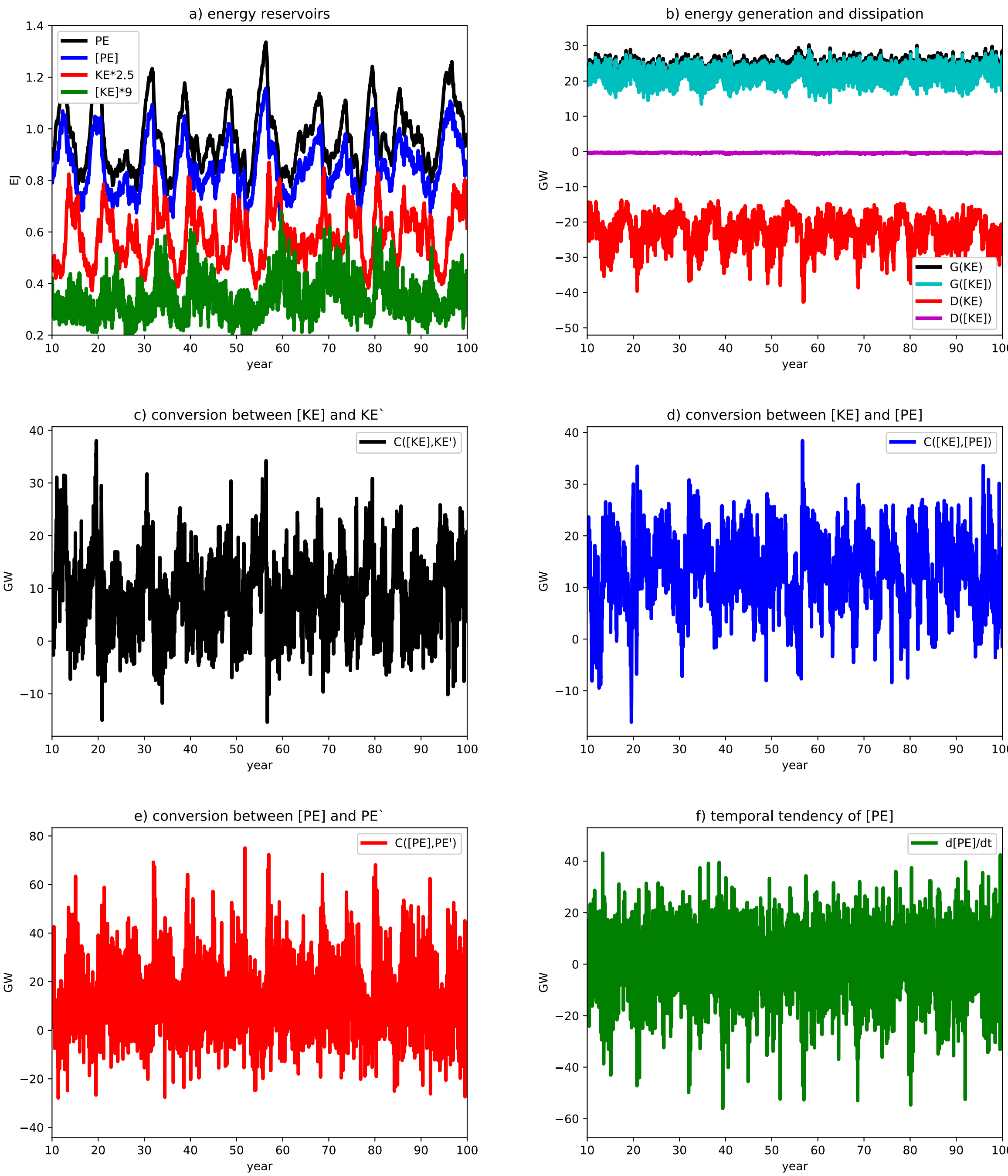


FIG. 4. Lorenz energy cycle (temporal average and standard deviation) of the 2-layer QG model based on spatial filtering and 500 years of daily output. Shown are the reservoirs of large-scale available potential energy ([PE]), large-scale kinetic energy ([KE]), eddy available potential energy (PE'), and eddy kinetic energy (KE'), as well as the corresponding energy generation (G), dissipation (D), and conversion (C) terms.



1006 FIG. 5. Time series of (a) energy reservoirs, (b) energy generation and dissipation, (c) conversion between
1007 large-scale and small-scale kinetic energy, (d) conversion between large-scale kinetic energy and large-scale
1008 available potential energy, (e) conversion between large-scale available potential energy and small-scale avail-
1009 able potential energy, and (f) temporal tendency of the large-scale available potential energy reservoir. Note that
1010 the last three terms constitute the large-scale available potential energy budget (Eq. (27)).

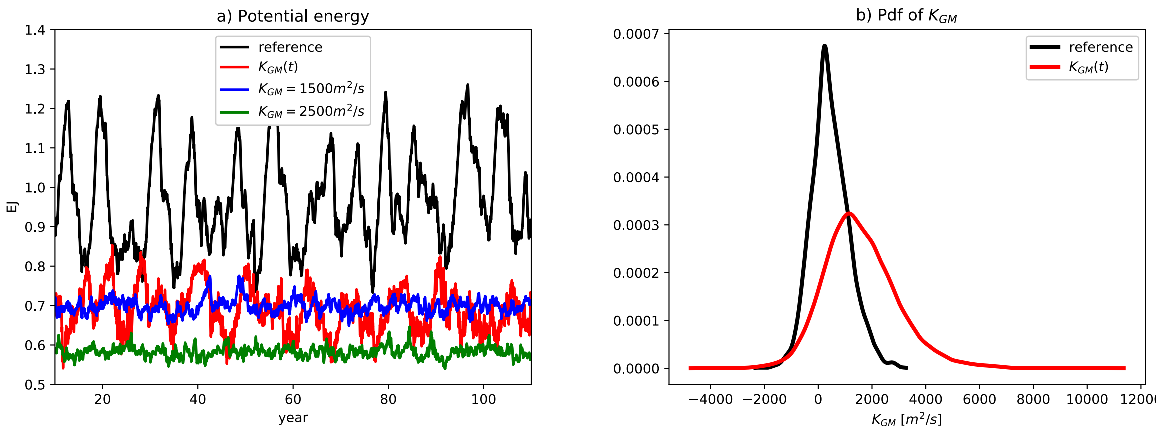


FIG. 6. a) Potential energy corresponding to the reference simulation (black, same as in Fig. 5a), and for simulations in which the GM parameterisation (Eq. (39)) is employed with K_{GM} either a constant (blue, green) or given via Eq. (43) (red). b) Estimated probability density function of K_{GM} (computed via Eq. (43)) for a simulation in which the GM parameterisation is employed (red) and for the reference simulation (black, the GM parameterisation is not employed in the model but K_{GM} is just diagnosed). The average and standard deviation are $1585 \pm 1373\ m^2/s$ (red) and $448 \pm 697\ m^2/s$ (black).

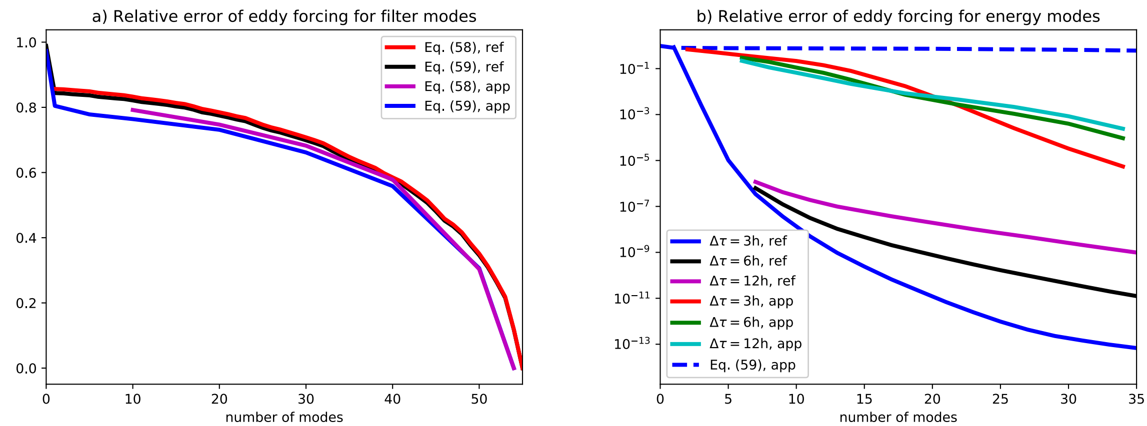


FIG. 7. Time-mean relative error of eddy forcing, $||\mathscr{R}_i^B - \tilde{\mathscr{R}}_i^B||/||\mathscr{R}_i^B||$, with $\tilde{\mathscr{R}}_i^B$ given by (a) the series expansions (58) or (59) related to the filter modes and (b) the series expansion (57) related to the energy modes (the results for the series expansion (56) are virtually identical). Here 'ref' refers to the reference simulation ($\tilde{\mathscr{R}}_i^B$ is only diagnosed) and 'app' refers to simulations in which $\tilde{\mathscr{R}}_i^B$ is applied. The lag step size $\Delta\tau$ is given in hours.

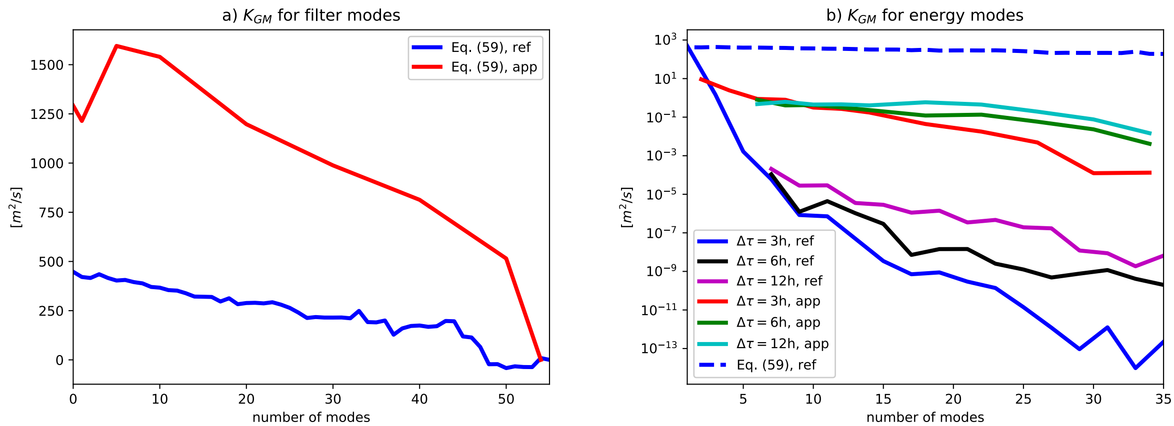


FIG. 8. Time-mean of the GM diffusivity K_{GM} for (a) the series expansion (59) and (b) the series expansion (57). Here 'ref' refers to the reference simulation ($\tilde{\mathscr{R}}_i^B$ is only diagnosed) and 'app' refers to simulations in which $\tilde{\mathscr{R}}_i^B$ is applied. The lag step size $\Delta\tau$ is given in hours.

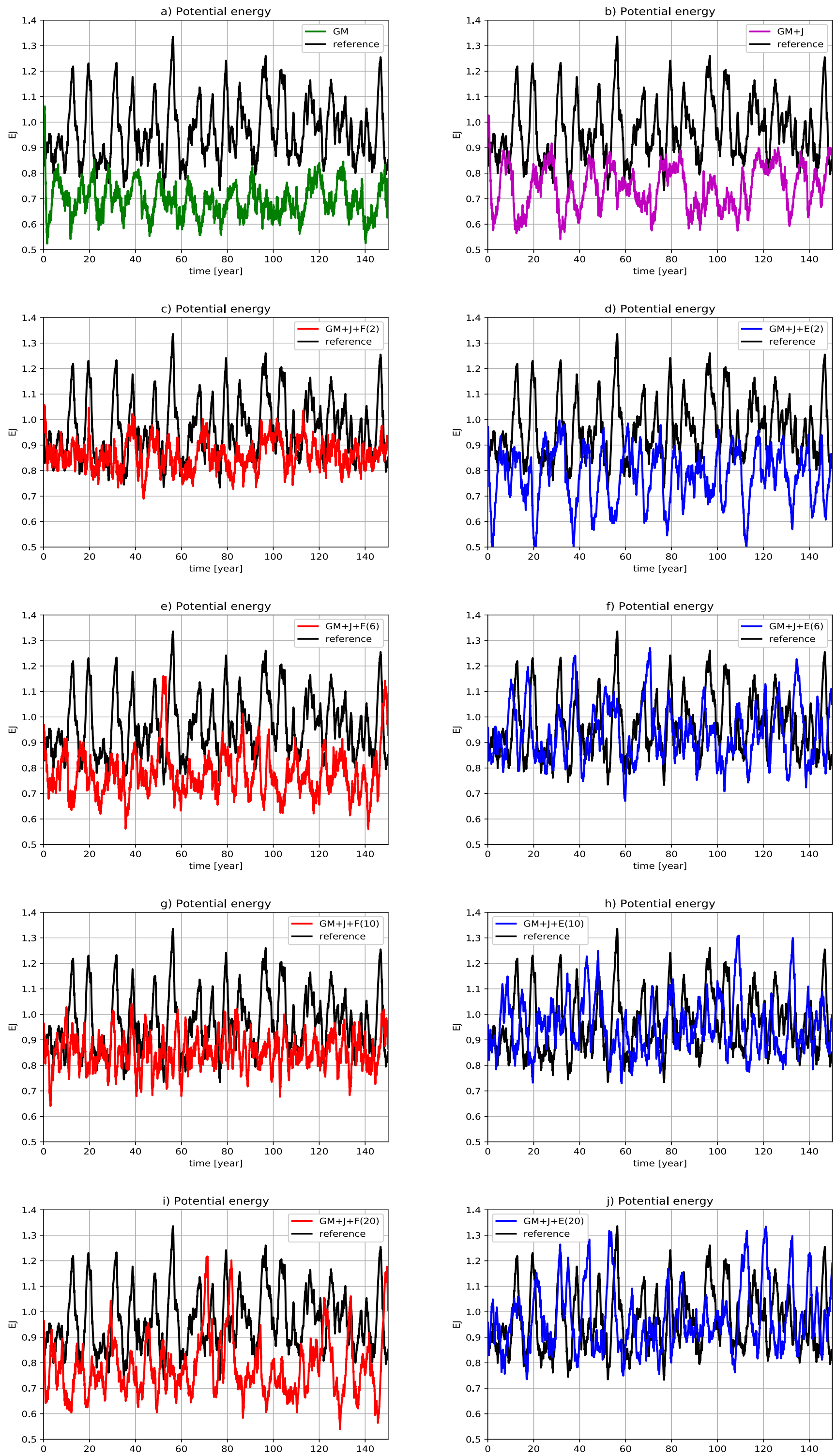


FIG. 9. Time series of PE for the reference simulation (black), and for simulations with $\tilde{\mathscr{R}}_i^B$ expanded in either filter modes (red, with F(#$filter\ modes$) and based on Eq. (59)) or energy modes (blue, with E(#*energy modes*) and based on Eq. (57) with $\Delta\tau = 3$ hours).