Stochastic subgrid-scale parameterization for one-dimensional shallow water dynamics using stochastic mode reduction

Matthias Zacharuk, Stamen I. Dolaptchiev*, Ulrich Achatz

Institut für Atmosphäre und Umwelt, Goethe-Universität Frankfurt, Frankfurt am Main, Germany

Ilya Timofeyev

University of Houston, Houston, Texas, USA

Corresponding author address: Institut für Atmosphäre und Umwelt, Fachbereich Geowissenschaften/Geographie, Goethe-Universität, Altenhöferallee 1, 60438 Frankfurt/Main, Germany

E-mail: dolaptchiev@iau.uni-frankfurt.de

ABSTRACT

We address the question of parameterizing the subgrid scales in simulations of geophysical flows by applying stochastic mode reduction to the one-dimensional stochastically forced shallow water equations. The problem is formulated in physical space by defining resolved variables as local spatial averages over finite-volume cells and unresolved variables as corresponding residuals. Based on the assumption of a time-scale separation between the slow spatial averages and the fast residuals, the stochastic mode reduction procedure is used to obtain a low-resolution model for the spatial averages alone with local stochastic subgrid-scale parameterization coupling each resolved variable only to a few neighboring cells. The closure improves the results of the low-resolution model and outperforms two purely empirical stochastic parameterizations. It is shown that the largest benefit is in the representation of the energy spectrum. By adjusting only a single coefficient (the strength of the noise) we observe that there is a potential for improving the performance of the parameterization, if additional tuning of the coefficients is performed. In addition, the scale-awareness of the parameterizations is studied.

1. Introduction

Atmospheric processes encompass a large spectrum of spatial and temporal scales. These range from several millimeters and seconds for boundary layer turbulence up to 10^7 meters and several weeks (and even longer) for planetary wave dynamics. Due to limited computer resources numerical atmospheric models cannot describe all these processes on all scales simultaneously. However, the different scales are interacting in a complex manner and this leads to the challenging problem of parameterizing the effect of the unresolved subgrid-scale (SGS) processes onto the resolved ones. Examples include the parameterization of synoptic and mesoscale eddies in planetary scale atmospheric models (e.g. Petoukhov et al. 2000; Weaver et al. 2001), momentum and temperature fluxes in the atmospheric boundary layer (Stull 1988) or SGS Reynold stresses in large eddy simulations (e.g. Pope 2000).

In this context, stochastic elements have become increasingly popular. Stochastic parameterizations can reduce a systematic model error, represent uncertainty in predictions, or trigger regime transitions (e.g., Palmer 2001; Berner et al. 2017). Typically some *ad-hoc* SGS model is assumed and the corresponding coefficients are optimized (tuned) so as to obtain the best possible agreement, in some sense, with observations or high-resolution simulations. Examples in comprehensive climate and weather models are stochastically perturbed parameterization tendencies (Buizza et al. 1999; Palmer et al. 2009) or stochastic kinetic energy backscatter (Shutts 2005; Berner et al. 2009). Empirical Ornstein-Uhlenbeck (OU) processes have been used in some studies of low-frequency and large-scale atmospheric variability (e.g. Winkler et al. 2001; Newman et al. 2003; Pegion and Sardeshmukh 2011), which can be extended to include quadratic nonlinearities as well as a time correlated stochastic forcing (Kravtsov et al. 2005; Kondrashov et al. 2005).

With regard to SGS parameterizations in climate models issues can arise from the fact that they are typically tuned to optimally represent the statistics of the present-day climate. If climate changes due to some external forcing, it is not guaranteed that the tuned parameters are still optimal. The fluctuation-dissipation theorem might be able to provide corrections (Achatz et al. 2013; Pieroth et al. 2018) in some cases, but such an approach relies on the perturbations being sufficiently weak. Moreover, there is a need for scale-aware parameterizations in atmosphere modeling, as model resolution increases continuously and mesh refinement techniques become widely used. In addition, the consistency between particular SGS parameterizations and the numerical discretization becomes important.

These considerations motivate the development of other approaches where the SGS parameterization is derived from first principles, if possible without any empirical parameter optimization. The direct interaction approximation (DIA) introduced by Kraichnan (1959) allowed to successfully apply statistical dynamical closure theory in relevant geophysical flows (Frederiksen and Davies 1997; Frederiksen 1999; Frederiksen et al. 2003). In the presence of time scale separation, the asymptotic method of averaging has been applied (Hasselmann 1976; Imkeller and von Storch 2001; Arnold et al. 2003; Monahan and Culina 2011). This method requires an estimation of the invariant measure of the fast scales conditioned on the slow scales, which might limit its applicability when going to high-dimensional systems. Another promising approach, without any empirical component, is based on the maximum entropy principle (Verkley 2011; Verkley and Severijns 2014; Verkley et al. 2016). Recently, Wouters and Lucarini (2012, 2013) have introduced a new method originating from response theory. This method relies on a weak coupling between resolved and unresolved scales and it has been applied to simple and more complex settings (Wouters et al. 2016; Demaeyer and Vannitsem 2017; Vissio and Lucarini 2017).

The DIA parameterization has been successfully applied to barotropic (Frederiksen and Davies 1997) and primitive equations model (Frederiksen et al. 2003), it has been extended to include the effects of mean flow and topography (Frederiksen 1999). The DIA closure is derived in spectral space by considering the evolution of second-order cumulant and response function. Next, the nonlinear damping rate and nonlinear noise are introduced. This results in a globally coupled SGS model in spectral space. However, techniques have been proposed to simplify the equations and obtain locally coupled models reproducing the spectra from direct numerical simulations (Frederiksen and Davies 1997; Frederiksen et al. 2003).

Another nearly self-consistent possibility that exploits a separation of time scales between resolved and unresolved scales is the stochastic mode reduction (SMR) procedure proposed by Majda et al. (2001, 2002, 2003). The SMR is a homogenization technique for multiscale systems (Khasminsky (1966b,a); Kurtz (1973); Papanicolaou (1976) and a recent overview Pavliotis and Stuart (2008) and references therein) and it is supplemented by an empirical step, where the fast SGS self-interactions in the evolution equation for the unresolved modes are replaced by an OU process. Following this step, an analytical derivation of a stochastic parameterization for the fast modes is possible, rigorously valid in the limit of infinite time scale separation.

So far the SMR procedure has already been applied to balanced models, such as barotropic (Franzke et al. 2005) and quasi-geostrophic (Franzke and Majda 2006) dynamics. The separation between resolved and unresolved scales has been performed by using empirical orthogonal functions (EOFs). However, EOFs are sometimes not able to guarantee a sufficient separation of the underlying time scales (e.g., see Figure 3 in Franzke and Majda 2006). The SMR carried out in spectral space (Franzke et al. 2005; Franzke and Majda 2006) is quite similar to the DIA closure approach. In particular, the main goal of both techniques is to represent the subgrid pro-

cesses by a nonlinear damping and a state-dependent noise and both techniques have been utilized successfully in geophysical flows.

In applications of the SMR to spectral space the resulting reduced model is globally coupled with linear, quadratic and even cubic terms. This hampers the applicability of the technique when high-dimensional systems with large number of resolved modes are considered. However, the latter problem can be avoided by applying the SMR in physical space to a finite-volume discretization of the equations. Such discretization does not per se include global coupling as in spectral discretizations, since grid cells interact directly only with a small number of neighbors. Finite-volume schemes are traditionally applied in ocean models (e.g. Haidvogel and Beckmann 1999), regional atmospheric modeling (e.g. Skamarock and Coauthors 2008) and recently even for global atmospheric models as well (Rípodas et al. 2009; Majewski et al. 2002; Satoh et al. 2008). In the examples above complex boundaries, such as continental boundaries in ocean modeling or non-periodic lateral boundaries in regional area atmospheric modeling, necessitate the use of discretizations and SGS parameterizations formulated in physical space. This motivated Dolaptchiev et al. (2013a,b) to consider a local approach where the resolved variables are defined by local spatial averages and the SGS flow by deviations from these averages, a configuration typically encountered in large-eddy turbulence parameterization (e.g. Pope 2000). The local definition leads to a local SGS parameterization, coupling only near neighbors, as shown for the Burgers equation (Dolaptchiev et al. 2013a,b). The efficient local stochastic SGS parameterization allows to consider large numbers of resolved scales. In addition, the clear gap of spatial scales between the resolved and unresolved variables enables a more pronounced time-scale separation.

Obviously Burgers equation represents a highly idealized prototype model for testing various statistical and closure methods and it is necessary to verify the applicability of the SMR for local spatial averages for more realistic fluid-dynamical models. One step in this direction is performed

in this work by applying the approach to a stochastically forced one-dimensional shallow water layer (1DSW). It incorporates at least two issues of relevance in the general context. First, in contrast to the Burgers equation the 1DSW allows for gravity waves. Secondly, if formulated in flux form, the shallow-water flow dynamical equations entail non-polynomial nonlinearities. This problem is of broader relevance, since such highly nonlinear terms appear in the general compressible fluid flow equations as well, in the pressure-gradient acceleration.

The work presented here can be summarized as follows. Based on a high-resolution finite-volume discretization of the shallow-water equations we use in Sec. 2 local spatial averages to define coarse and slow (resolved) variables and, via corresponding residuals, fine and fast (unresolved) variables. The assumed time-scale separation is verified numerically. The SMR theory for obtaining an SGS parameterization of the unresolved modes is then introduced and applied to the specific problem. In Sec. 3 we discuss the practical implementation, and also introduce, for comparison, two purely empirical approaches. Results from model simulations with the various SGS parameterizations are then compared in Sec. 4. Here we also investigate the scale awareness of the approaches, i.e. their ability to be applied at different resolutions without any re-tuning. Conclusions are finally drawn in Sec. 5.

2. Method

a. Shallow water model

We consider a stochastically forced one dimensional shallow water layer with periodic boundaries, using as variables the height of the fluid h and the momentum hu, where u is the velocity. The governing equations with plane topography (e.g. Vallis 2006) read in flux form

$$\partial_t \begin{pmatrix} h \\ hu \end{pmatrix} = -\partial_x \begin{pmatrix} hu - \nu \partial_x h \\ \frac{(hu)^2}{h} + \frac{g}{2}h^2 - \nu \partial_x hu \end{pmatrix} + \boldsymbol{\rho}, \tag{1}$$

with a large-scale stochastic forcing $\boldsymbol{\rho}$ (see Sec. 3) and a mass weighted diffusion with the constant parameter ν.

For a high-resolution spatial discretization the domain of length L is divided into N fine intervals $\Delta x = L/N$, labelled by a small index $i \in \{0, 1, ..., N-1\}$. With this the equations in (1) can be discretized by a symmetric finite-volume scheme

$$\frac{d}{dt} \begin{pmatrix} h_i \\ hu_i \end{pmatrix} = -\frac{1}{\Delta x} \left(\mathbf{F}_{i+\frac{1}{2}} - \mathbf{F}_{i-\frac{1}{2}} \right) + \boldsymbol{\rho}_i \,, \tag{2}$$

with the discrete forcing $\boldsymbol{\rho}_i$ and the flux at the boundary given by

$$\mathbf{F}_{i+\frac{1}{2}} = \frac{1}{2} \begin{pmatrix} (hu)_{i+1} + (hu)_i - 2\nu \dfrac{h_{i+1} - h_i}{\Delta x} \\ \dfrac{(hu)_{i+1}^2}{h_{i+1}} + \dfrac{(hu)_i^2}{h_i} + \dfrac{g}{2}h_{i+1}^2 + \dfrac{g}{2}h_i^2 - 2\nu \dfrac{(hu)_{i+1} - (hu)_i}{\Delta x} \end{pmatrix} . \tag{3}$$

The discrete flux form (2) conserves total mass $\frac{1}{N}\sum_{k=0}^{N-1} h_k$ and total momentum $\frac{1}{N}\sum_{k=0}^{N-1} hu_k$ in the absence of forcing. Given our choice of forcing and dissipation, the number of fine cells is chosen large enough so as to resolve all processes occurring. Hence in the following simulations using (2) , with N large enough, will be called direct numerical simulation (DNS).

b. Local averages

As a representation of the typical situation of atmospheric models with insufficient resolution, we introduce a second discretization, with $N_c = N/n$ coarse cells, each consisting of n fine cells of the initial discretization, and labelled by the capital index $I \in \{0, 1, ..., N_c - 1\}$. Associated with

the coarse grid coarse variables H and HU, also called resolved variables, are defined by local spatial averages inside a coarse box

$$\begin{pmatrix} H_I \\ HU_I \end{pmatrix} = \frac{1}{n} \sum_{k=nI}^{n(I+1)-1} \begin{pmatrix} h_k \\ hu_k \end{pmatrix} . \tag{4}$$

Further, fine variables h' and hu', referred to as unresolved or SGS variables, are defined using the deviations of the initial variables from the corresponding coarse variables

$$\begin{pmatrix} h'_i \\ hu'_i \end{pmatrix} = \begin{pmatrix} h_i \\ hu_i \end{pmatrix} - \begin{pmatrix} H_{I[i]} \\ HU_{I[i]} \end{pmatrix} . \tag{5}$$

$I[i]$ denotes here the index of the coarse cell with the i-th fine cell placed inside. The coarse and fine variables can be used to express (2) as

$$\frac{d}{dt} \begin{pmatrix} H_I \\ HU_I \end{pmatrix} = - \frac{\mathbf{F}_{(I+1)n-\frac{1}{2}} - \mathbf{F}_{nI-\frac{1}{2}}}{n\Delta x} + \boldsymbol{\rho}_I \, , \tag{6}$$

$$\frac{d}{dt} \begin{pmatrix} h'_i \\ hu'_i \end{pmatrix} = - \frac{\mathbf{F}_{i+\frac{1}{2}} - \mathbf{F}_{i-\frac{1}{2}}}{\Delta x} + \frac{\mathbf{F}_{(I[i]+1)n-\frac{1}{2}} - \mathbf{F}_{I[i]n-\frac{1}{2}}}{n\Delta x} \, , \tag{7}$$

where we assume that the forcing $\boldsymbol{\rho}_I$ acts only onto the coarse variables. By collecting all resolved variables H_I, HU_I in one vector $\mathbf{x} \in \mathbb{R}^{M_c}$ and all SGS variables h'_i, hu'_i in another vector $\mathbf{z} \in \mathbb{R}^M$ with $M_c = 2N_c$ and $M = 2N$, (6) and (7) can be rewritten as

$$\dot{x}_i = \rho_i^x + a_i^x(\mathbf{x}) + b_i^{xz}(\mathbf{x}, \mathbf{z}) \, , \tag{8}$$

$$\dot{z}_i = b_i^z(\mathbf{x}, \mathbf{z}) + c_i^z(\mathbf{z}) \, . \tag{9}$$

Here ρ_i^x results from the forcing and terms have been regrouped so as to identify the coarse variable self-interactions $a_i^x(\mathbf{x})$, the coupling terms $b_i^{xz}(\mathbf{x},\mathbf{z})$, $b_i^z(\mathbf{x},\mathbf{z})$ and the fine variable self-interactions $c_i^z(\mathbf{z})$. Complete neglect of the SGS variables yields the bare-truncation model

$$\dot{x}_i = \rho_i^x + a_i^x(\mathbf{x}) \; . \tag{10}$$

This low resolution model is defined on the coarse grid with N_c grid cells and it lacks an SGS parameterization.

c. Stochastic Mode Reduction

1) QUADRATIC APPROXIMATION

A difficulty in the application of SMR is caused by the terms involving $1/h$ in the flux (3) , as they represent a nonlinearity of an arbitrary order. So far SMR has only been applied to systems with quadratic nonlinearities. It can, however, handle nonlinearities of arbitrary polynomial form. Hence a solution could be expanding everywhere but in $a_I^x(\mathbf{x})$ terms with $1/h$ in a finite Taylor series around the mean fluid height $\mathscr{H}$. It turns out sufficient, however, to simply replace $1/h \approx 1/\mathscr{H}$ in order to reproduce the statistics of the DNS. Thus, the bare truncation part of the model is computed exactly but in all other terms involving SGS modes this approximation is used, leading to an approximation of (8) and (9), where SGS-variable nonlinearities take a quadratic form

$$\dot{x}_i = \rho_i^x + a_i^x(\mathbf{x}) + \left(L_{ij}^{xz} z_j + B_{ijk}^{xxz} x_j z_k + B_{ijk}^{xzz} z_j z_k \right) \; , \tag{11}$$

$$\dot{z}_i = \left(L_{ij}^{zx} x_j + B_{ijk}^{zxx} x_j x_k + B_{ijk}^{zxz} x_j z_k \right) + \left(L_{ij}^{zz} z_j + B_{ijk}^{zzz} z_j z_k \right) \; . \tag{12}$$

Here and in the following we make use of Einstein's summation convention and the summation index is running up to either M_c or M depending on if $\mathbf{x}$ or $\mathbf{z}$ is involved. The linear and quadratic interaction coefficients L_{ij}^{xz}, L_{ij}^{zx}, L_{ij}^{zz}, B_{ijk}^{xxz}, B_{ijk}^{xzz}, B_{ijk}^{zxx}, B_{ijk}^{zxz}, B_{ijk}^{zzz} are given in Appendix A.

2) Empirical OU process

The SGS variables z_i are not independent, since the corresponding local spatial average over a coarse cell vanishes by definition. Thus, one degree of freedom for each coarse cell has to be eliminated. This is achieved by Fourier-transforming h' and hu' locally inside each coarse cell. The Fourier amplitude of the zero-wavenumber is equal to the vanishing local average inside the coarse cell. Hence by discarding this wavenumber component one degree of freedom can be eliminated. This defines the new independent SGS variables θ_i

$$\theta_i = \hat{T}_{ij} z_j \tag{13}$$

$$z_i = \hat{R}_{ij} \theta_j \,, \tag{14}$$

where the matrices $\hat{R} \in \mathbb{R}^{M \times M_f}$ and $\hat{T} \in \mathbb{R}^{M_f \times M}$ are constructed from the inverse and forward Fourier transformation and $M_f = M - M_c$ is the number of independent SGS variables. With this preparation one can move to the next step of SMR, i.e. replacing the SGS self-interactions $L_{ij}^{zz} z_j + B_{ijk}^{zzz} z_j z_k$ by an empirical OU process. The SGS equation (12) becomes

$$d\theta_i = \hat{T}_{ik} \left(L_{kj}^{zx} x_j + B_{kjl}^{zxx} x_j x_l + B_{kjl}^{zxz} x_j \hat{R}_{lm} \theta_m \right) dt + \Gamma_{ij} \theta_j dt + \sigma_i dW_i \,, \tag{15}$$

where Γ_{ij} denote the coefficients of the negative-definite OU drift, $\sigma_{ij} = \sigma_i \delta_{ij}$ those of a diagonal diffusion tensor, and dW_i Wiener increments. Note that no sum over i is taken in the Wiener term in (15). We assume that Γ couples only SGS modes corresponding to the same coarse cell. Under

this assumption Γ has a block-diagonal form and the resulting SGS closure is local, coupling only neighbors and next-neighbors of a coarse cell. Since SMR assumes further that the OU process is the dominant term in the SGS equation (see below), the OU drift Γ can be estimated from the lagged covariance of $\boldsymbol{\theta}$ (Gardiner 2009)

$$\overline{\boldsymbol{\theta}(t)\boldsymbol{\theta}^T(t+\tau)} = C(\tau) = C(0)e^{\Gamma^T\tau}\,, \tag{16}$$

where $\overline{(\cdot)}$ denotes a time average. By integrating over time one can solve for Γ

$$\left(\Gamma^T\right)^{-1} = -C(0)^{-1}\int_0^\infty C(\tau)d\tau\,. \tag{17}$$

The computation of the time integral in (17) is performed using the numerically efficient Cooper-Haynes algorithm (Lutsko et al. 2015). Note that despite the block-diagonal form, Γ still allows for a coupling between both SGS variables h_i' and hu_i' inside each coarse cell. Because of the spatial homogeneity of the considered shallow-water model the coefficients of Γ are the same for each matrix block and can be obtained by averaging over the estimates from the different coarse cells. Using the steady Lyapunov equation $\Gamma C(0) + C(0)\Gamma^T = -\sigma\sigma^T$, the diagonal diffusion coefficients are found from

$$\sigma_i = \sqrt{-2\Gamma_{ik}C(0)_{ki}}\,. \tag{18}$$

Moreover, it has turned out to be useful to observe that in typical applications Γ is diagonalizable and has distinct eigenvalues (e.g., Delsole 2004). We hence introduce new variables y_i

$$y_i = U_{ij}^{-1}\theta_j\,, \tag{19}$$

where the real invertible matrix U is from the real Jordan canonical form decomposition of Γ

$$\Gamma = U\Lambda U^{-1} \,. \tag{20}$$

By applying the transformation matrices $R = \hat{R}U$ and $T = U^{-1}\hat{T}$, the model equations (11) and (15) can be written in terms of the new variables as

$$\dot{x}_i = \rho_i^x + a_i^x(\mathbf{x}) + b_i^x(\mathbf{x},\mathbf{y}) \,, \tag{21}$$

$$\dot{y}_i = b_i^y(\mathbf{x},\mathbf{y}) + \Lambda_{ij} y_j + \Sigma_i \dot{W}_i \,. \tag{22}$$

Here we use the notation

$$\begin{aligned} b_i^x(\mathbf{x},\mathbf{y}) =& L_{ij}^{xz} R_{jk} y_k + B_{ijk}^{xxz} R_{kl} x_j y_l + B_{ijk}^{xzz} R_{jl} R_{km} y_l y_m \\ =& L_{ij}^{xy} y_j + B_{ijk}^{xxy} x_j y_k + B_{ijk}^{xyy} y_j y_k \,, \end{aligned} \tag{23}$$

$$\begin{aligned} b_i^y(\mathbf{x},\mathbf{y}) =& T_{ij} \left(L_{jk}^{zx} x_k + B_{jkl}^{zxx} x_k x_l + B_{jkl}^{zxz} R_{lm} x_k y_m \right) \\ =& L_{ij}^{yx} x_j + B_{ijk}^{yxx} x_j x_k + B_{ijk}^{yxy} x_j y_k \,, \end{aligned} \tag{24}$$

and we have also introduced effective drift coefficients $\Sigma_i = \sqrt{U_{ij}^{-1} U_{ij}^{-1} \sigma_j^2}$ with pairwise identical noise parameters for pairs of complex eigenvalues, see Appendix C in Dolaptchiev et al. (2013a) for the details.

3) HOMOGENIZATION

The remaining step is the derivation of an effective equation for the coarse variable $\mathbf{x}$ alone, using the homogenization technique (Majda et al. 2001; Pavliotis and Stuart 2008), with terms taking SGS effects into account. The main assumption of the SMR is the presence of distinct

time scales in the considered variables. So far the model is spatially separated into coarse and fine variables. This does not necessarily imply a separation of the underlying time scales. However, as it will be shown later, for the considered regime the separation in space also induces a separation in time, with the resolved variable $\mathbf{x}$ acting on a slower time scale than the SGS variable $\mathbf{y}$. Hence a time-scale separation factor $\varepsilon \ll 1$ is introduced to characterize the different time scales associated with the different terms on the right hand side of (21) and (22). We replace $b^x \to b^x/\varepsilon$, $b^y \to b^y/\varepsilon$, and $\Lambda_{ij}y_j + \Sigma_i \dot{W}_i \to \Lambda_{ij}y_j/\varepsilon^2 + \Sigma_i \dot{W}_i/\varepsilon$, where then b^x, b^y, $\Lambda_{ij}y_j$, and $\Sigma_i \dot{W}_i$ are all $O(1)$, obtaining

$$\dot{x}_i = \rho_i^x + a_i^x(\mathbf{x}) + \frac{1}{\varepsilon} b_i^x(\mathbf{x},\mathbf{y}) \,, \tag{25}$$

$$\dot{y}_i = \frac{1}{\varepsilon} b_i^y(\mathbf{x},\mathbf{y}) + \frac{1}{\varepsilon^2}\Lambda_{ij}y_j + \frac{1}{\varepsilon}\Sigma_i \dot{W}_i \,. \tag{26}$$

The above scaling implies that the bare truncation part $\rho^x + a^x(\mathbf{x})$ acts on the slowest, the coupling terms b^x and b^y on a faster and the SGS self-interactions on the fastest time scale. In the following the corresponding backward Fokker-Planck equation (FPE)

$$\partial_t p = L_3 p + \frac{1}{\varepsilon} L_2 p + \frac{1}{\varepsilon^2} L_1 p \,, \tag{27}$$

for the probability density function (PDF) p is considered, in the limit of an infinite time scale separation $\varepsilon \to 0$. The operators on the right hand side are defined as

$$L_3 = -\left(\rho_i^x + a_i^x(\mathbf{x})\right)\partial_{x_i} \,, \tag{28}$$

$$L_2 = -\, b_i^x(\mathbf{x},\mathbf{y})\partial_{x_i} - b_i^y(\mathbf{x},\mathbf{y})\partial_{y_i} \,, \tag{29}$$

$$L_1 = -\,\Lambda_{ij}y_j\partial_{y_i} - \frac{\Sigma_i^2}{2}\partial_{y_i}^2 \,. \tag{30}$$

Next, the PDF in (27) is expanded in terms of ε: $p = p^0 + \varepsilon p^1 + \varepsilon^2 p^2 + ...$, which leads to the following set of equations

$$O(\varepsilon^{-2}): \quad 0 = L_1 p^0 \,, \tag{31}$$

$$O(\varepsilon^{-1}): \quad 0 = L_2 p^0 + L_1 p^1 \,, \tag{32}$$

$$O(\varepsilon^{0}): \quad p^0 = L_3 p^0 + L_2 p^1 + L_1 p^2 \,. \tag{33}$$

The leading order $O(\varepsilon^{-2})$-equation shows that p^0 is in the null space of L_1 and therefore it does not depend on the fast variables: $p^0 = p^0(\mathbf{x})$. The $O(\varepsilon^{-1})$-equation can be solved for p^1: $p^1 = -L_1^{-1} L_2 p^0$ if the solvability condition

$$P L_2 p^0 = 0 \tag{34}$$

is satisfied, where the projection operator P, projecting onto the null space of L_1 is utilized. With this result the last $O(\varepsilon^0)$-equation can be written as an effective FPE for p^0 only

$$\partial_t p^0 = L_3 p^0 - P L_2 L_1^{-1} L_2 p^0 \,. \tag{35}$$

This is the backward FPE of a low-resolution model, for the coarse variables alone, which consists of the bare truncation part $L_3 p^0$ and an SGS parameterization $P L_2 L_1^{-1} L_2 p^0$. The null-space projection P and inverse of the OU operator L_1 are detailed in Appendix B. Using these, one finds that the stochastic differential equation corresponding to the effective FPE (35) can be written as

$$dx_i = [\rho_i^x + a_i^x(\mathbf{x}) + \beta_i(\mathbf{x})]\, dt + d\xi_i(\mathbf{x}) \,. \tag{36}$$

Here β_i represents the deterministic part and $d\xi_i$ the stochastic part of the SGS parameterization, containing both additive and multiplicative noise terms. One finds that the deterministic closure β_i is

$$\beta_i = \int_0^\infty d\tau \left\langle b_j^x(\mathbf{x},\mathbf{y}) \frac{\partial b_i^x(\mathbf{x},\mathbf{y}(\tau))}{\partial x_j} \right\rangle + \langle \mathbf{y}\mathbf{y}^T \rangle_{jm}^{-1} \int_0^\infty d\tau \left\langle y_m b_j^y(\mathbf{x},\mathbf{y}) b_i^x(\mathbf{x},\mathbf{y}(\tau)) \right\rangle - \int_0^\infty d\tau \left\langle \frac{\partial b_j^y(\mathbf{x},\mathbf{y})}{\partial y_j} b_i^x(\mathbf{x},\mathbf{y}(\tau)) \right\rangle , \tag{37}$$

where the expectations $\langle\cdot\rangle$ are taken over the OU statistics, and $\mathbf{y}$ represents an OU trajectory with initial condition $\mathbf{y} = \mathbf{y}(0)$. The stochastic closure $d\xi_i$ takes the form

$$d\xi_i = \sqrt{2} B_{ij} dW_j , \tag{38}$$

where the matrix elements B_{ij} are obtained from the decomposition

$$B_{ik}B_{jk} = \int_0^\infty d\tau \left\langle b_i^x(\mathbf{x},\mathbf{y}(0)) b_j^x(\mathbf{x},\mathbf{y}(\tau)) \right\rangle . \tag{39}$$

With these results one can also show that back-transforming $b^x/\varepsilon \to b^x$, $b^y/\varepsilon \to b^y$, and $\Lambda_{ij}y_j/\varepsilon^2 + \Sigma_i \dot{W}_i/\varepsilon \to \Lambda_{ij}y_j + \Sigma_i \dot{W}_i$ leaves β and $d\xi$ unchanged so that (36) is the desired low-resolution model.

Finally getting back to the specific case, (21) – (24) , Appendix C shows that the solvability condition (34) is satisfied. Moreover, inserting (23) and (24) for b^x and b^y yields

$$\begin{aligned}\beta_i = &\left(L^{xy}_{mj} + B^{xxy}_{mlj} x_l\right) B^{xxy}_{imk} (C_S)_{jk} \\ &+ \left(L^{yx}_{mo} x_o + B^{yxx}_{mop} x_o x_p\right) \left(L^{xy}_{ik} + B^{xxy}_{ijk} x_j\right) \langle \mathbf{y}\mathbf{y}^T \rangle^{-1}_{mn} (C_S)_{nk} \\ &+ B^{xyy}_{ijk} B^{yxy}_{mpo} x_p \langle \mathbf{y}\mathbf{y}^T \rangle^{-1}_{mn} (C_T)_{onjk} \,, \end{aligned} \tag{40}$$

where the tensors C_S and C_T are given by

$$(C_S)_{jk} = \int_0^\infty d\tau \langle y_j(0) y_k(\tau) \rangle \,, \tag{41}$$

$$(C_T)_{onjk} = \int_0^\infty d\tau \left(\langle y_o(0) y_j(\tau) \rangle \langle y_k(\tau) y_n(0) \rangle + \langle y_o(0) y_k(\tau) \rangle \langle y_j(\tau) y_n(0) \rangle \right) \,. \tag{42}$$

With this the decomposition (39) becomes

$$B_{ik} B_{jk} = B^{xyy}_{imn} (C_T)_{mnkl} B^{xyy}_{jkl} + \left(L^{xy}_{in} + B^{xxy}_{ikn} x_k\right) (C_S)_{nm} \left(L^{xy}_{jm} + B^{xxy}_{jlm} x_l\right) \,. \tag{43}$$

The prescription would be to perform this decomposition every time step. However, this would be very expensive. Therefore we neglect cross-correlations between the $d\xi_i$ in different cells and approximate them by

$$\begin{aligned} d\xi_i \approx &\sqrt{2 B_{ij} B_{ij}} dW_i \,, \\ = &\sqrt{2 B^{xyy}_{ijn} (C_T)_{jnkl} B^{xyy}_{ikl}} dW^1_i + \sqrt{2 \left(L^{xy}_{in} + B^{xxy}_{ijn} x_j\right) (C_S)_{nk} \left(L^{xy}_{ik} + B^{xxy}_{ilk} x_l\right)} dW^2_i \,, \end{aligned} \tag{44}$$

which we call effective stochastic forcing. In each coarse cell the approximated stochastic term has thus the same variance as its exact counterpart. The stochastic term (44) consists of an additive part, which acts on both variables, and a multiplicative part, that acts only on HU. An important feature of the SGS parameterization, with deterministic part (40) and stochastic component (44),

is that it couples a volume cell only to its neighbors and next neighbors. This allows application of the approach to large systems.

3. Test case and model suite

For the validation of our approach we consider a stochastically forced periodic shallow-water layer of horizontal extent $L = 10^4$ km and mean height $\mathscr{H} = 10$ km. The diffusion constant is $\nu = 10^5$ km^2 day^{-1} . A large-scale stochastic forcing (Chekhlov and Yakhot 1995) is applied to the momentum equation

$$\boldsymbol{\rho}_I = \begin{pmatrix} 0 \\ \sum_{k=1}^{3} \frac{\mu\alpha_k}{\sqrt{k\Delta t}} \cos\left(2\pi\left(\frac{kIn\Delta x}{L_x} + \psi_k\right)\right) \end{pmatrix} . \tag{45}$$

Normally distributed random numbers α_k and ψ_k are used, the amplitude parameter μ is 10^5 km^2/day$^{\frac{3}{2}}$ and the forcing acts onto the leading Fourier modes $1 \leq k \leq 3$. Various model set-ups have been chosen as follows, a summary is given in Table 2. In all cases the integrations have been done over 10^4days with 10^3 outputs per day.

a. High-resolution simulations

1) DNS

Reference is provided by direct numerical simulations, integrating (2) with $N = 512$ volume cells. A 4th-order Runge-Kutta-scheme is used, with a time step $\Delta t = 10^{-4}$ days.

2) OU-DNS

In two intermediate steps in the application of the SMR, first the nonlinearities affected by SGS dynamics have been kept quadratic by replacing $1/h \to 1/\mathscr{H}$, and then the SGS nonlinear self-interactions have been replaced by an empirical OU process, leading to the system (21) and

(22). Direct integration of these equations, henceforth termed OU-DNS, thus appears as a useful check of the validity of the SMR approach. However, directly using the OU parameters estimated from (17) and (18) turned out not to be stable enough. Therefore following Achatz and Branstator (1999) an additional scale-selective damping has been supplemented to the OU drift in each coarse cell. This has been done in the spectral representation of the latter, see (15), by replacing

$$\Gamma \to \Gamma + \begin{pmatrix} \gamma & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \gamma \end{pmatrix}, \qquad \text{with } \gamma = -\alpha \begin{pmatrix} 1^2 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & (n-1)^2 \end{pmatrix} \in \mathbb{R}^{(n-1)\times(n-1)} .$$

The diagonal matrix γ represents damping of the Fourier modes inside each coarse cell with an amplitude proportional to the squared wave number, with here $\alpha = 90$ day^{-1}. As also in all other stochastic integrations outlined below, the time integration has been done by a split-step method with a 4th-order Runge-Kutta step for the deterministic part and an Euler-Mayurama step for the stochastic part. The time step is $\Delta t = 10^{-4}$day, as in the deterministic DNS.

b. Low-resolution simulations

With the high-resolution simulations as reference we can validate the SMR approach for providing an SGS parameterization for low-resolution models that only use the coarse cells. We compare the performance of this approach also to that of more simple purely empirical parameterizations. Considered approaches are as follows where, unless otherwise stated, the number of coarse cells employed was always $N_c = 64$ with an averaging interval of $n = 8$.

1) Low-resolution simulations without SGS parameterization

Two slightly different approaches have been chosen to obtain low-resolution models. The first is defined by the original discretized equations (2), but with a lower spatial resolution of $N_c = N/n$. This is henceforth referred to as *low-resolution model (LRM)*. The second variant is the *bare-truncation model (BRT)* defined in (10), with a resolution of N_c as well. The difference between BRT and LRM is the diffusion in the models. In the BRT it is proportional to $\nu/(n\Delta x^2)$, and in the LRM to $\nu/(n\Delta x)^2$, implying that the BRT has an effective diffusion by a factor n stronger, when compared to the LRM.

2) Low-resolution simulations with SGS parameterization

Three types of low-resolution simulations with stochastic SGS parameterization have been tested.

(i) SMR parameterization. The low-resolution model (36) to be validated is the BRT supplemented by the SMR SGS parameterization consisting of the deterministic and stochastic components (40) and (44), it is referred to as BRT-SMR. For stability reasons the BRT-SMR diffusivity had to be increased in corresponding simulations to $\nu = 2 \cdot 10^5$ km^{2}day^{-1}. The time step employed was $\Delta t = 2 \cdot 10^{-5}$day.

(ii) Empirical OU parameterizations for BRT and LRM. As a quality measure for the SMR approach we also consider low-resolution simulations with an empirical OU SGS parameterization, denoted by BRT-OU or LRM-OU, depending on the low-resolution dynamical core used together with the empirical OU SGS parameterization. As in the SMR parameterization only coupling to neighbors and next neighbors is taken into account. The BRT-OU, for example, can be written as

$$dx_i = \left(\rho_i^x + a_i^x(\mathbf{x}) + \tilde{\Gamma}_{ij}\hat{x}_j^I\right) dt + \tilde{\sigma}_i dW_i \,. \tag{46}$$

where the vector $\hat{\mathbf{x}}^I$ has 10 components and encompasses the values of H and HU in the five coarse cell from $I-2$ up to $I+2$, where I is the cell index corresponding to the variable x_i. The OU parameters $\tilde{\Gamma}$ and $\tilde{\sigma}$ have been estimated using a standard maximum likelihood approach (Honerkamp 1994), yielding

$$\tilde{\Gamma}_{ij} = \overline{b_i^x \hat{x}_k} \left(\overline{\hat{\mathbf{x}}\hat{\mathbf{x}}^T}^{-1}\right)_{kj}, \tag{47}$$

$$\tilde{\sigma}_i^2 = \Delta t \overline{\left[b_i^x - \tilde{\Gamma}_{ij}\hat{x}_j^I\right]^2}, \tag{48}$$

with the superscript of $\hat{\mathbf{x}}^I$ suppressed in (47). For the LRM-OU the corresponding parameters have been determined in the same manner. In contrast to the estimation of the OU processes in the SMR by (17) and (18), here the integrated lagged covariance function could not be used, because the SGS effects $b_i^x(\mathbf{x},\mathbf{y})$ as such do not satisfy a prognostic equation dominated by an OU process. Replacing (17) and (18) by maximum-likelihood estimates would have been an option as well. Corresponding tests have shown a slightly deteriorated performance, however.

4. Results

In the following we show step by step the essential results from our various simulation experiments. Autocorrelations and spectra turned out to be qualitatively similar for momentum and surface-height. We therefore focus below on the latter.

a. DNS of the shallow-water layer

Fig. 1 (left) displays the time dependence of the autocorrelation of the resolved variable H and of the SGS variable h' in the DNS. A slowly decaying oscillation is visible in the autocorrelation of H. The period of this oscillation is nearly equal to the time $\tau = L/\sqrt{g\mathscr{H}} \approx 0.37$ day, required for gravity waves to pass once through the domain. This shows that the model has some intrinsic dynamics and is not dominated by forcing and diffusion. The autocorrelation of the SGS variable h' decays much faster to zero than that of H. The large difference in the correlation time between SGS and resolved variables indicates that the assumption of time-scale separation between $\mathbf{x}$ and $\mathbf{y}$ is met to a good agreement.

The spatial distribution of the variance of h' is displayed in Fig. 1 (right). The variance is lowest in the middle of a coarse cell, and gradually increases towards the cell boundaries. This spatial shape is explained in Appendix E as being due to a spatially decreasing autocorrelation of h.

The potential-energy spectrum from the DNS is displayed in the left panel of Fig. 2. With the considered forcing and diffusion parameters one obtains an inertial range with spectral index 2 up to around wavenumber $kL/2\pi = 64$. There is a small kink in the spectrum after wavenumber 3, due to the forcing acting only onto the first three modes.

The deviations from a Gaussian in the fourth order moments of H and HU are less than 4% and 2%, respectively. In addition we find nearly vanishing odd moments (not shown). We conclude that the statistics of the resolved variables are close to Gaussian.

b. OU-DNS

As described above, the replacement $1/h \to 1/\mathscr{H}$ in the SGS nonlinearities leads to the system (11) and (12) with strictly quadratic nonlinearities, as required for the application of the SMR method. Simulations with this model reproduce the DNS data nearly perfectly (not shown).

Replacing the SGS self-interactions by an empirical OU process leads to the system (21) and (22). The corresponding OU-DNS reproduces the correlations of the DNS with minor differences (not shown). The energy spectrum from the OU-DNS, projected onto the coarse grid is displayed in Fig. 2 (right). It follows the DNS spectrum for the first 7 wavenumbers and then drops below it. This indicates a too strong damping at high wavenumbers, which seems to be due to the introduction of the deterministic part in the OU-process. The spatial variance of h' from the OU-DNS model is presented in Fig. 1 (right). It follows with small deviations the structure from the DNS model.

c. Low-resolution simulations without SGS parameterization

Before considering results from low-resolution simulations with the various SGS parameterizations, we first address low-resolution simulations without any parameterizations, to provide a useful reference. The time dependence of the autocorrelation function of H from these simulations is shown in Fig. 1 (left). The amplitude of the oscillation of the auto-correlation from the LRM simulations is significantly weaker than from the DNS and has a relative error of 6.3%, computed for time lags between 0 and 1 day. The corresponding oscillation from the BRT simulation is slightly stronger correlated with that from the DNS, with a relative error of 2.3%. The corresponding period matches that from the DNS whereas that from the LRM simulation is shorter. Comparing the corresponding energy spectra in Fig. 2 (right) one can see that the energy from the BRT simulation is overall less than from the DNS, and that the spectrum is steeper. In contrast to this, LRM simulations yield too much energy between wavenumbers 4 and 15, and too little at smaller scales. At all scales LRM simulations yield more energy than the BRT simulations.

The relative errors in Table 1 show that the LRM simulation significantly overestimates the all statistical moments, the fourth moment in HU even by 111.9%. The BRT simulation yields moments that are too small, in the case of the fourth moment by 29%.

We also verified numerically that in order to reproduce the spectra of the first 32 wavenumbers in the DNS with $N = 512$ grid points, it is possible to perform low-resolution DNS with at least 256 spatial points. However, further reducing the number of points in the low-resolution DNS significantly corrupts the spectra of the first 32 wavenumbers. Thus, this demonstrates the need for SGS parameterizations if one wants to reduce the spatial resolution beyond $N = 256$.

d. Low-resolution simulations with SGS parameterization

Energy spectrum. The potential-energy spectra obtained from low-resolution simulations with the different parameterizations are shown in Fig. 3 (right). The overall qualitative behavior of the shallow-water layer can be reproduced with the SMR parameterization but there is too much energy in scales up to around wavenumber 12 and too little energy in higher wavenumbers. On the other hand, the bare truncation model with empirical stochastic corrections (BRT-OU) and the low-resolution model (LRM-OU) do not reproduce all the details of the spectra sufficiently accurately. In particular, the LRM-OU spectrum contains significantly too little energy in all wave numbers. The BRT-OU simulation can reproduce the true spectrum well in the first 15 wavenumbers but fails completely at higher wavenumbers.

Moments. The statistical moments from the various simulations are summarized in Table 1. In general all closures show high relative errors, the empirical OU parameterizations underestimate and the SMR parameterization overestimates the moments. Errors in the HU-moments are smallest for the low-resolution simulation using the SMR parameterization. In the H-moments BRT-SMR has lower errors than LRM-OU but larger than BRT-OU. The BRT-SMR model over-

estimates the fluctuations in the coarse H-variable, which is consistent with the result for potential energy spectra discussed above. This implies that the SGS stochastic forcing representing the energy backscatter is too strong.

Improvement of Energy Balance. To further improve the performance of our SGS parametrization using the stochastic mode reduction, we consider BRT-SMR model with reduced SGS stochastic forcing. This is motivated by the fact that there are several assumptions in the SMR approach (e.g. time-scale separation, representing the fast variables to the leading order by the OU process, polynomial form approximation of the interaction terms in the equation for momentum). Thus, we consider BRT-SMR models where the stochastic part in the SMR parameterization is reduced by 40% or completely neglected. Stochastic terms in the SMR parametrization represent the energy backscatter of small scales onto resolved large scales. It has been recognized that proper modeling of this phenomena is particularly important in the context of geophysical turbulence (see e.g. Palmer (2001); Palmer et al. (2009); Berner et al. (2009)). Therefore, we study how well the SMR parametrization reproduces this process and whether various approximations introduced in the context of applying the SMR to the shallow water equation impose additional sensitivity of the BRT-SMR model to the stochasticity of the closure.

The reduction of the stochastic part by 40% (defining BRT-SMR-0.6) significantly reduces the error in the variance of H to 2.9% and of HU to -5.2%, see Table 1. To avoid extensive tuning of the BRT-SMR model, we consider the uniform SGS noise reduction of both variables. Alternatively, SGS noises on H and HU can be reduced by a different percentage, thus further optimizing the performance of the BRT-SMR model. With the choice of the 40% reduction of SGS noise, the performance in the energy spectrum can be improved for the first 8 wave numbers, but for higher wave numbers the energy content drops, see Fig. 4. However, we wold like to emphasize that the first 8 wavenumbers contain approximately 97% of the potential energy. From Fig. 4 it is also

visible that already the deterministic part of the closure can significantly improve the spectrum as compared to BRT.

Correlation. The time autocorrelations from low-resolution simulations (BRT or LRM) with either the empirical OU or the SMR parameterization are depicted in Fig. 3 (left). One can see that application of the SMR parameterizations leads reproducing the autocorrelation from the DNS with small differences in amplitude. The relative error of the correlation is 3.4%. Application of the OU parameterization in the LRM leads to simulations with an oscillation in the autocorrelation that is too weak in amplitude and exhibits a small phase shift, whereas use of the OU approach in the BRT leads to simulations with an autocorrelation similar to that obtained with the SMR parameterization. The relative error of the correlation is 10.5% for the LRM-OU simulation and 6.6% for the BRT-OU simulation. The SGS noise reduction in the BRT-SMR-0.6 model does not significantly affect the correlation function (not depicted for this model). The correlation function for the BRT-SMR-0.6 overlaps with the correlation function computed using the BRT-SMR. This can be intuitively understood since the correlation function in many stochastic models is determined primarily by the strength of the deterministic terms.

e. Scale adaptivity

The advantage of the SMR parameterization is that it can be adapted easily to changes in the model setup, and in many situations it does not have to be recalculated. This has been investigated by considering larger averaging intervals of $n = 16,\ 32$, resulting in different spatial resolutions $N_c = N/n = 16,\ 32$. To adjust the SMR closure to the changed resolution, we use (D2) and (D4) from the Appendix D. Note that no re-determination of the model is necessary. In contrast to this, no modification rule exists for the empirical closures in the BRT-OU and the LRM-OU model. We

keep those parameterizations unchanged for the considered cases. Whereas the LRM-OU remains stable, the BRT-OU is unstable in both cases.

The potential energy spectra from integrations of the resulting stable models are displayed in Fig. 5. For comparison the corresponding DNS projection is shown as well. In both cases integration of the low-resolution models with SGS parameterization yield less energy in the resolved flow than the DNS. However, application of the SMR SGS parameterization leads to better agreement with the DNS, especially for $n = 16$. Both low-resolution simulations can capture the time correlation well (not shown).

5. Conclusion

The applicability of subgrid-scale (SGS) parameterizations to a wide range of parameters of a dynamical system such as the atmosphere, and their ability to be easily used at different model configurations, requires that they are based on first principles as much as possible. Stochastic mode reduction (SMR) as suggested by Majda et al. (2001, 2002, 2003), i.e. homogenization applied to a system with its nonlinear fast-variable self-interactions replaced by an empirical Ornstein-Uhlenbeck (OU) process, is a promising option in this direction. Geophysical applications of the SMR so far were performed always in spectral space (Franzke et al. 2005; Franzke and Majda 2006). However, in many applications, such as ocean modeling or regional climate modeling, SGS parameterizations in physical space are required. In order to construct such parameterization we use the local approach suggested by Dolaptchiev et al. (2013a,b), and tested within the framework of the Burgers equation. A central aspect of this approach is the discrimination, within a finite-volume formulation of the high-resolution dynamics, between slowly varying spatial averages, that are resolved explicitly, and more rapidly varying deviations from those, that are to be parameterized.

As a next step towards the application of this technique to real atmospheric flows, our work validates the applicability of the SMR in the context of one-dimensional shallow-water (1DSW) flow. This introduces two general features. (1) Gravity waves are included as well as (2) high-order non-polynomial nonlinearities that generally affect compressible flows. After the validation of the required time-scale separation between local averages and small-scale flow, the latter issue has been handled by replacing, in all dynamical terms affecting or affected by the fast small-scale flow, the inverse of the water-column height by the inverse of its global equilibrium value. This limits the corresponding nonlinearities to quadratic. Further replacing all small-scale self-interactions by an empirical OU process yields a representation of the dynamics that allows model simulations in rather good agreement with simulations of the unmodified 1DSW equations. We could hence proceed and apply the homogenization technique to obtain an explicit low-resolution model for the local averages, with an SMR SGS parameterization of the small-scale flow coupling only a small number of neighboring cells.

This model has been validated against data from high-resolution simulations of 1DSW flow. It is shown that the SMR SGS parameterization improves the energy spectrum at the smaller resolved scales in comparison with both simulations without SGS parameterizations and simulations using an empirical OU SGS parameterization. In the error of some statistical moments no clear benefit of the SMR SGS parameterization is present. However, we demonstrated that the error can be considerably lowered by diminishing the stochasticity in the SMR closure. In particular, the variance error of the SMR SGS model can be reduced to 2.9% in H. We also found that this comes along with less energy at high wavenumbers. We conjecture that the performance of the BRT-SMR model can be improved further by empirically adjusting the coefficients of the SMR SGS parametrization. Finally, we also show that the closure can easily be adapted to changes in the model parameters. This enables a scale awareness, which allows to utilize the SMR SGS pa-

rameterization for different spatial resolutions and leads to improvements compared to empirical SGS schemes.

In a related study within the framework of the Lorenz 96 model, Vissio and Lucarini (2017) have recently demonstrated parameter-awareness of a parameterization derived using response theory (Wouters and Lucarini 2012) by changing the time-scale separation between resolved and unresolved scales. How far this extends to our setting, where scale-adaptivity is considered with regard to the number of resolved modes, remains to be investigated. Moreover, as shown by Wouters et al. (2016) for comparatively simple models, the SGS scheme of (Wouters and Lucarini 2012) does outperform the SMR parameterization at smaller time lags, but on longer time scales it converges to the SMR result in the limit of infinite time scale separation. Still, it would be interesting to extend the work of Demaeyer and Vannitsem (2018) by comparing both approaches in more complex applications.

Motivated by the DIA closure of Frederiksen and Davies (1997), Frederiksen and Kepert (2006) applied a stochastic modeling approach accounting for memory effects of the turbulent eddies and constructed SGS parameterization from a high-resolution simulation. The resulting SGS model is local in spectral space and includes linear eddy drain viscosity and stochastic backscatter viscosity. The same approach was successfully used by Kitsios et al. (2012, 2013) to construct scale-aware SGS parameterizations, which reproduce the spectra exactly. Interestingly, the SMR provides additional nonlinear deterministic correction terms and multiplicative noise terms. Such terms might become important in situations where effects due to topography, intermittency or large-scale flow are relevant. In addition, the scaling laws found by Kitsios et al. (2012, 2013) for the eddy viscosities suggest that similar scaling laws might be valid for the parameters of the OU process in Fourier space, used in the SMR approach. This might improve further the results on scale-adaptivity presented here.

Potentially an issue is that we had to increase diffusivity in order to stabilize the low-resolution model with SMR SGS parameterization. This is a well known issue with purely empirical SGS parameterizations (e.g. Achatz and Schmitz 1997). In the present semi-analytical approach it might be overcome by using the energy conserving discretization of Fjordholm et al. (2011). As pointed out by Majda et al. (2009) there is a connection between energy conservation by the discretized nonlinearities and the cubic damping term in the SMR SGS parameterization. Indeed, in the studies of Franzke et al. (2005); Franzke and Majda (2006); Dolaptchiev et al. (2013a) the discrete treatment of the nonlinear terms conserves energy and the resulting SMR SGS schemes are stable.

Even from the present results, however, we conclude that it appears worthwhile further moving towards the application of the SMR SGS parameterizations to low-resolution simulations in general compressible flows. Next step would be to increase the complexity by considering two-dimensional shallow-water flow and by including rotational effects. Such system contains dispersive inertial gravity waves as well as geostrophic balanced flow. One interesting question in this regard is if the effect of high-frequency, small-scale gravity waves on the large-scale gravity waves and geostrophic flow can be parameterized using the present local SMR approach.

Acknowledgments. We want to thank the reviewers for their comments and suggestions which helped to improve the draft version of the manuscript. The code for the SGS parameterizations is available upon request. MZ, SD and IT thank the German Research Foundation (DFG) for partial support through grant DO 1819/1-1. IT thanks for partial support by the grant ONR N00014-17-1-2845. UA thanks DFG for partial support through grant AC 71/7-1.

APPENDIX A

Interaction coefficients

To define the interaction coefficients in (11), (12), the following notation is used

$$\left(L^{xz}_{lm}z_m, B^{xxz}_{lmn}x_m z_n, B^{xzz}_{lmn}z_m z_n\right) = \begin{cases} \left(L^{Hz}, B^{Hxz}, B^{Hzz}\right) & \text{if } x_l \text{ denotes } H_I \\ \left(L^{HUz}, B^{HUxz}, B^{HUzz}\right) & \text{if } x_l \text{ denotes } HU_I \end{cases}, \tag{A1}$$

$$\left(L^{zz}_{lm}z_m, L^{zx}_{lm}x_m\right) = \begin{cases} \left(L^{h'z}, L^{h'x}\right) & \text{if } z_l \text{ denotes } h'_i \\ \left(L^{hu'z}, L^{hu'x}\right) & \text{if } z_l \text{ denotes } hu'_i \end{cases}, \tag{A2}$$

$$\left(B^{zxx}_{lmn}x_m x_n, B^{zxz}_{lmn}x_m z_n, B^{zzz}_{lmn}z_m z_n\right) = \begin{cases} \left(B^{h'xx}, B^{h'xz}, B^{h'zz}\right) & \text{if } z_l \text{ denotes } h'_i \\ \left(B^{hu'xx}, B^{hu'xz}, B^{hu'zz}\right) & \text{if } z_l \text{ denotes } hu'_i \end{cases}, \tag{A3}$$

with $I \in \{0,1,...,N_c-1\}$ and $i \in \{0,1,...,N-1\}$. The linear interaction coefficients read

$$\begin{aligned} L^{Hz} =& -\frac{1}{2n\Delta x}\left(hu'_{n(I+1)} + hu'_{n(I+1)-1} - hu'_{nI} - hu'_{nI-1}\right) && \text{(A4)}\\ &+\frac{\nu}{n\Delta x^2}\left(h'_{n(I+1)} - h'_{n(I+1)-1} - h'_{nI} + h'_{nI-1}\right) \\ L^{HUz} =& \frac{\nu}{n\Delta x^2}\left(hu'_{n(I+1)} - hu'_{n(I+1)-1} - hu'_{nI} + hu'_{nI-1}\right) && \text{(A5)}\\ L^{h'x} =& -\frac{1}{2\Delta x}\left(HU_{I[i+1]} - HU_{I[i-1]}\right) + \frac{\nu}{\Delta x^2}\left(H_{I[i-1]} - 2H_{I[i]} + H_{I[i+1]}\right) && \text{(A6)}\\ &+\frac{1}{2n\Delta x}\left(HU_{I[i]+1} - HU_{I[i]-1}\right) - \frac{\nu}{n\Delta x^2}\left(H_{I[i]+1} - 2H_{I[i]} + H_{I[i]-1}\right) && \text{(A7)}\\ L^{hu'x} =& \frac{\nu}{\Delta x^2}\left(HU_{I[i-1]} - 2HU_{I[i]} + HU_{I[i+1]}\right) && \text{(A8)}\\ &-\frac{\nu}{n\Delta x^2}\left(HU_{I[i]+1} - 2HU_{I[i]} + HU_{I[i]-1}\right) && \text{(A9)}\\ L^{h'z} =& -\frac{1}{2\Delta x}\left(hu'_{i+1} - hu'_{i-1}\right) + \frac{\nu}{\Delta x^2}\left(h'_{i-1} - 2h'_i + h'_{i+1}\right) - L^{Hz} && \text{(A10)}\\ L^{hu'z} =& \frac{\nu}{\Delta x^2}\left(hu'_{i-1} - 2hu'_i + hu'_{i+1}\right) - L^{HUz}, && \text{(A11)} \end{aligned}$$

where in (A10), (A11) the terms L^{Hz} and L^{HUz} are given in (A4), (A5), but with the index I replaced by $I[i]$. The nonlinear interaction coefficients read

$$B^{Hxz} = B^{Hzz} = B^{h'xx} = B^{h'xz} = B^{h'zz} = 0 \tag{A12}$$

$$\begin{aligned} B^{HUxz} = -\frac{1}{n\Delta x}\Bigg[& \frac{1}{\mathscr{H}}\Big(HU_{I+1} hu'_{n(I+1)} + HU_I hu'_{n(I+1)-1} \\ & - HU_I hu'_{nI} - HU_{I-1} hu'_{nI-1} \Big) \\ & + \frac{g}{2}\left(H_{I+1} h'_{n(I+1)} + H_I h'_{n(I+1)-1} - H_I h'_{nI} - H_{I-1} h'_{nI-1} \right) \Bigg] \end{aligned} \tag{A13}$$

$$\begin{aligned} B^{HUzz} = -\frac{1}{2n\Delta x}\Bigg[& \frac{1}{\mathscr{H}}\left((hu'_{n(I+1)})^2 + (hu'_{n(I+1)-1})^2 - (hu'_{nI})^2 - (hu'_{nI-1})^2 \right) \\ & + \frac{g}{2}\left((h'_{n(I+1)})^2 + (h'_{n(I+1)-1})^2 - (h'_{nI})^2 - (h'_{nI-1})^2 \right) \Bigg] \end{aligned} \tag{A14}$$

$$B^{hu'xx} = -\frac{1}{2\Delta x}\left(\frac{1}{\mathscr{H}}\left(HU^2_{I[i+1]} - HU^2_{I[i-1]} \right) + \frac{g}{2}\left(H^2_{I[i+1]} - H^2_{I[i-1]} \right) \right) \tag{A15}$$

$$+ \frac{1}{2n\Delta x}\left[\frac{1}{\mathscr{H}}\left(HU^2_{I[i]+1} - HU^2_{I[i]-1} \right) + \frac{g}{2}\left(H^2_{I[i]+1} - H^2_{I[i]-1} \right) \right] \tag{A16}$$

$$\begin{aligned} B^{hu'xz} = & -\frac{1}{\Delta x}\left[\frac{1}{\mathscr{H}}\left(HU_{I[i+1]} hu'_{i+1} - HU_{I[i-1]} hu'_{i-1} \right) + \frac{g}{2}\left(H_{I[i+1]} h'_{i+1} - H_{I[i-1]} h'_{i-1} \right) \right] \\ & - B^{HUxz} \end{aligned} \tag{A17}$$

$$B^{hu'zz} = -\frac{1}{2\Delta x}\left(\frac{1}{\mathscr{H}}\left((hu'_{i+1})^2 - (hu'_{i-1})^2 \right) + \frac{g}{2}\left((h'_{i+1})^2 - (hu'_{i-1})^2 \right) \right) - B^{HUzz}, \tag{A18}$$

where in (A17), (A18) the terms B^{HUxz} and B^{HUzz} are given in (A13), (A14), but with the index I replaced by $I[i]$.

APPENDIX B

Null-space projection and inverse of the OU backward Fokker-Planck operator

To determine the projection operator P and the inverse operator L_1^{-1} one can consider an auxiliary process described by the backward FPE $\partial_t \chi = L_1 \chi$ with the conditional PDF $\chi(\tilde{\mathbf{y}}, 0 \mid \mathbf{y}, \tau)$. The invariant measure of this process $p_s(\tilde{\mathbf{y}}) = \lim_{\tau \to -\infty} \chi(\tilde{\mathbf{y}}, 0 \mid \mathbf{y}, \tau)$ defines the projection operator

$$(Pg)(\mathbf{x}) = \int d\mathbf{y}\; g(\mathbf{x},\mathbf{y}) p_s(\mathbf{y}) = \langle g(\mathbf{x},\mathbf{y})\rangle \;, \tag{B1}$$

with the expectation $\langle\cdot\rangle$ of $g(\mathbf{x},\mathbf{y})$ with respect to the invariant measure p_s. The inverse operator L_1^{-1} applied onto a function $f(\mathbf{x},\mathbf{y})$ is found to be given by

$$\left(L_1^{-1} f\right)(\mathbf{x},\mathbf{y}) = \int_0^\infty d\tau \int d\mathbf{y}\; f(\mathbf{x},\tilde{\mathbf{y}}) \chi(\tilde{\mathbf{y}},\tau \mid \mathbf{y},0) \;. \tag{B2}$$

Both operators applied consecutively yield

$$\begin{aligned}
\left(PgL_1^{-1} f\right)(\mathbf{x}) &= \int d\mathbf{y}\; g(\mathbf{x},\mathbf{y}) p_s(\mathbf{y}) \int_0^\infty d\tau \int d\tilde{\mathbf{y}}\; f(\mathbf{x},\tilde{\mathbf{y}}) \chi(\tilde{\mathbf{y}},\tau \mid \mathbf{y},0) \\
&= \int_0^\infty d\tau \int d\mathbf{y}\; g(\mathbf{x},\mathbf{y}) p_s(\mathbf{y}) \int d\tilde{\mathbf{y}}\; f(\mathbf{x},\tilde{\mathbf{y}}) \chi(\tilde{\mathbf{y}},\tau \mid \mathbf{y},0) \\
&= \int_0^\infty d\tau\; \langle g(\mathbf{x},\mathbf{y}) f(\mathbf{x},\tilde{\mathbf{y}}(\tau))\rangle \;,
\end{aligned} \tag{B3}$$

where in the lagged covariance in the last line $\tilde{\mathbf{y}}(\tau)$ is understood to be an OU trajectory with initial condition $\tilde{\mathbf{y}}(0) = \mathbf{y}$.

APPENDIX C

Solvability condition

The solvability condition (34) can be rewritten to $Pb_i^x \partial_{x_i} p^0 = 0$ since $p^0 = p^0(\mathbf{x})$. It is fulfilled if the even stronger condition $Pb_i^x = 0$ holds. This is the case here. Using (23), defining

$$\kappa_{jk} = R_{jl} R_{km} \langle y_l y_m \rangle \tag{C1}$$

and inserting the transformation $\mathbf{z} = R\mathbf{y}$, one obtains with $\langle y_i \rangle = 0$

$$Pb_i^x = B_{ijk}^{xyy} \langle y_j y_k \rangle = B_{ijk}^{xzz} \kappa_{jk} = B_{ijj}^{xzz} \kappa_{jj} \;. \tag{C2}$$

We have used in the last step that $B^{xzz}_{ijk} = B^{xzz}_{ijk}\delta_{jk}$, as can be seen from Appendix A. Since $B^{xzz}_{ijj}\kappa_{jj}$ is a difference between fluxes at the right and at the left of a cell, the total sum over i vanishes

$$\sum_i B^{xzz}_{ijj}\kappa_{jj} = 0\,. \tag{C3}$$

The homogeneity of $\langle y_l y_m\rangle$ implies that each element in the sum is identical and thus must vanish. This proofs the solvability condition.

APPENDIX D

The SMR SGS parameterization for changed resolution

The SMR SGS parameterization can be adapted to different coarse grid resolutions without any recalculation. This can be achieved by collecting interaction coefficients proportional to different powers of n. For example the linear coefficients L^{yx} can be written as $L^{yx} = \tilde{L}^{yx} + \frac{1}{n}\hat{L}^{yx}$, where $\hat{L}^{yx}$ results from all terms in L^{yx} multiplied by $\frac{1}{n}$ and $\tilde{L}^{yx}$ from terms independent of n. Similarly all other coefficients can be split in this way, in particular we have $L^{xy} = \frac{1}{n}\hat{L}^{xy}$, $B^{xxy} = \frac{1}{n}\hat{B}^{xxy}$, etc.. This separation leads to the deterministic closure

$$\begin{aligned}
\beta_i = &\left(\frac{1}{n}\hat{L}^{xy}_{mj} + \frac{1}{n}\hat{B}^{xxy}_{mlj}x_l\right)\frac{1}{n}\hat{B}^{xxy}_{imk}(C_S)_{jk} \\
&+ \left(\left[\tilde{L}^{yx}_{mo} + \frac{1}{n}\hat{L}^{yx}_{mo}\right)\right]x_o + \left[\tilde{B}^{yxx}_{mop} + \frac{1}{n}\hat{B}^{yxx}_{mop}\right]x_o x_p\Bigg)\frac{1}{n}\left(\hat{L}^{xy}_{ik} + \hat{B}^{xxy}_{ijk}x_j\right)\langle \mathbf{y}\mathbf{y}^T\rangle^{-1}_{mn}(C_S)_{nk} \\
&+ \frac{1}{n}\hat{B}^{xyy}_{ijk}\left[\tilde{B}^{yxy}_{mpo} + \frac{1}{n}\hat{B}^{yxy}_{mpo}\right]x_p\langle \mathbf{y}\mathbf{y}^T\rangle^{-1}_{mn}(C_T)_{onjk} \qquad \text{(D1)} \\
= &\frac{1}{n}\tilde{\beta}_i + \frac{1}{n^2}\hat{\beta}_i \qquad \text{(D2)}
\end{aligned}$$

where in the last steps the results are summarized with respect to the power of n. The effective stochastic closure analogously yields

$$d\xi_i \approx \frac{1}{n}\sqrt{2\hat{B}^{xyy}_{ijn}(C_T)_{jnkl}\hat{B}^{xyy}_{ikl}}dW_i^1 + \frac{1}{n}\sqrt{2\left(\hat{L}^{xy}_{in}+\hat{B}^{xxy}_{ijn}x_j\right)(C_S)_{nk}\left(\hat{L}^{xy}_{ik}+\hat{B}^{xxy}_{ilk}x_l\right)}dW_i^2 \tag{D3}$$

$$= \frac{1}{n}d\hat{\xi}_i\,. \tag{D4}$$

APPENDIX E

Spatial shape of the SGS variance

In continuous space the surface-height mean in the first coarse cell, with length L_c, and the corresponding SGS deviations can be written as

$$H = \frac{1}{L_c}\int_0^{L_c} h(x)\,dx\,, \qquad\qquad h'(x) = h(x) - H\,. \tag{E1}$$

This leads to spatial dependence in the variance of h'

$$\overline{h'(x)^2} - \overline{h'}^2 = \overline{h(x)^2} - 2\overline{h(x)H} + \overline{H^2} - \overline{h'}^2\,, \tag{E2}$$

Due to spatial homogeneity of $h(x)$ and $\overline{h'} = 0$, only the second term can be assumed to be spatially dependent. Now by assuming an exponentially decaying spatial correlation $\overline{h(x)h(x')} \sim \exp(-\alpha|x-x'|)$, with a decay rate $\alpha > 0$, the middle term becomes

$$
\begin{aligned}
\overline{h(x)H} &= \frac{1}{L_c}\int_0^{L_c} \overline{h(x)h(x')}\, dx' \\
&\sim \frac{1}{L_c}\left\{\int_0^{x} \exp\left[\alpha(x'-x)\right]\, dx' + \int_x^{L_c} \exp\left[\alpha(x-x')\right]\, dx'\right\} \\
&= \frac{1}{\alpha L_c}\left\{2 - \mathrm{e}^{-\alpha x} - \mathrm{e}^{\alpha(x-L_c)}\right\} \\
&= \frac{2}{\alpha L_c}\left\{1 - \mathrm{e}^{-\frac{\alpha L_c}{2}} \cosh\left[\alpha\left(x - \frac{L_c}{2}\right)\right]\right\}\,,
\end{aligned}
\tag{E3}
$$

describing the characteristic U-shape of the fine variable variance in the first coarse cell $x \in [0, L_c]$. The same considerations hold for all other coarse cells.

References

Achatz, U., and G. Branstator, 1999: A Two-Layer Model with Empirical Linear Corrections and Reduced Order for Studies of Internal Climate Variability. *Journal of the Atmospheric Sciences*, **56 (17)**, 3140–3160, doi:10.1175/1520-0469(1999)056⟨3140:ATLMWE⟩2.0.CO;2.

Achatz, U., U. Löbl, S. I. Dolaptchiev, and A. Gritsun, 2013: FluctuationDissipation Supplemented by Nonlinearity: A Climate-Dependent Subgrid-Scale Parameterization in Low-Order Climate Models. *Journal of the Atmospheric Sciences*, **70 (6)**, 1833–1846, doi:10.1175/JAS-D-12-0229.1, URL http://journals.ametsoc.org/doi/abs/10.1175/JAS-D-12-0229.1.

Achatz, U., and G. Schmitz, 1997: On the Closure Problem in the Reduction of Complex Atmospheric Models by PIPs and EOFs: A Comparison for the Case of a Two-Layer Model with Zonally Symmetric Forcing. *Journal of the Atmospheric Sciences*, **54 (20)**, 2452–2474, doi:10.1175/1520-0469(1997)054⟨2452:OTCPIT⟩2.0.CO;2, URL http://journals.ametsoc.org/doi/abs/10.1175/1520-0469(1997)054{\%}3C2452:OTCPIT{\%

}3E2.0.CO;2{\%}5Cnhttp://journals.ametsoc.org/doi/abs/10.1175/1520-0469{\%}281997{\%}29054{\%}3C2452{\%}3AOTCPIT{\%}3E2.0.CO{\%}3B2.

Arnold, L., P. Imkeller, and Y. Wu, 2003: *Reduction of deterministic coupled atmosphere-ocean models to stochastic ocean models: A numerical case study of the Lorenz-Maas system*, Vol. 18. 295–350 pp., doi:10.1080/14689360310001607979.

Berner, J., G. J. Shutts, M. Leutbecher, and T. N. Palmer, 2009: A Spectral Stochastic Kinetic Energy Backscatter Scheme and Its Impact on Flow-Dependent Predictability in the ECMWF Ensemble Prediction System. *Journal of the Atmospheric Sciences*, **66 (3)**, 603–626, doi:10.1175/2008JAS2677.1, URL http://journals.ametsoc.org/doi/abs/10.1175/2008JAS2677.1.

Berner, J., and Coauthors, 2017: Stochastic parameterization toward a new view of weather and climate models. *Bulletin of the American Meteorological Society*, **98 (3)**, 565–587, doi:10.1175/BAMS-D-15-00268.1, 1510.08682.

Buizza, R., M. Miller, and T. Palmer, 1999: Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System. *Quarterly Journal of the Royal Meteorological Society*, **125 (560)**, 2887–2908, doi:10.1256/smsqj.56005, URL http://ci.nii.ac.jp/naid/40020798400/http://www.ingentaselect.com/rpsv/cgi-bin/cgi?ini=xref{\&}body=linker{\&}reqdoi=10.1256/smsqj.56005.

Chekhlov, A., and V. Yakhot, 1995: Kolmogorov turbulence in a random-force-driven Burgers equation: Anomalous scaling and probability density functions. *Physical Review E*, **52 (5)**, 5681–5684, doi:10.1103/PhysRevE.52.5681, arXiv:1011.1669v3.

Delsole, T., 2004: Stochastic models of quasigeostrophic turbulence. *Surveys in Geophysics*, **25 (2)**, 107–149, doi:10.1023/B:GEOP.0000028164.58516.b2.

Demaeyer, J., and S. Vannitsem, 2017: Stochastic parametrization of subgrid-scale processes in coupled ocean-atmosphere systems: benefits and limitations of response theory. *Quarterly Journal of the Royal Meteorological Society*, **143 (703)**, 881–896, doi:10.1002/qj.2973, URL http://arxiv.org/abs/1605.00461http://dx.doi.org/10.1002/qj.2973http://doi.wiley.com/10.1002/qj.2973, 1605.00461.

Demaeyer, J., and S. Vannitsem, 2018: Comparison of stochastic parameterizations in the framework of a coupled ocean-atmosphere model.

Dolaptchiev, S. I., U. Achatz, and I. Timofeyev, 2013a: Stochastic closure for local averages in the finite-difference discretization of the forced Burgers equation. *Theoretical and Computational Fluid Dynamics*, **27 (3-4)**, 297–317, doi:10.1007/s00162-012-0270-1.

Dolaptchiev, S. I., I. Timofeyev, and U. Achatz, 2013b: Subgrid-scale closure for the inviscid Burgers-Hopf equation. *Communications in Mathematical Sciences*, **11 (3)**, 757–777, doi:10.4310/CMS.2013.v11.n3.a5.

Fjordholm, U. S., S. Mishra, and E. Tadmor, 2011: Well-balanced and energy stable schemes for the shallow water equations with discontinuous topography. *Journal of Computational Physics*, **230 (14)**, 5587–5609, doi:10.1016/j.jcp.2011.03.042, URL http://dx.doi.org/10.1016/j.jcp.2011.03.042.

Franzke, C., and A. J. Majda, 2006: Low-Order Stochastic Mode Reduction for a Prototype Atmospheric GCM. *Journal of the Atmospheric Sciences*, **63 (2)**, 457–479, doi:10.1175/JAS3633.1, URL papers2://publication/uuid/A13C96BD-C761-469A-A3F1-BD7E98DA3808{\%}5Cnhttp://journals.ametsoc.org/doi/abs/10.1175/JAS3633.1.

Franzke, C., A. J. Majda, and E. Vanden-Eijnden, 2005: Low-Order Stochastic Mode Reduction for a Realistic Barotropic Model Climate. *Journal of the Atmospheric Sciences*, **62 (6)**, 1722–1745, doi:10.1175/JAS3438.1, URL papers2://publication/uuid/A13C96BD-C761-469A-A3F1-BD7E98DA3808{\%}5Cnhttp://journals.ametsoc.org/doi/abs/10.1175/JAS3633.1http://journals.ametsoc.org/doi/abs/10.1175/JAS3438.1.

Frederiksen, J., 1999: Subgrid-scale parameterizations of eddy-topographic force, eddy viscosity, and stochastic backscatter for flow over topography. *JOURNAL OF THE ATMOSPHERIC SCIENCES*, **56 (11)**, 1481–1494, doi:{10.1175/1520-0469(1999)056⟨1481:SSPOET⟩2.0.CO;2}.

Frederiksen, J. S., and A. G. Davies, 1997: Eddy Viscosity and Stochastic Backscatter Parameterizations on the Sphere for Atmospheric Circulation Models. *Journal of Atmospheric Sciences*, **54**, 2475–2492, doi:10.1175/1520-0469(1997)054⟨2475:EVASBP⟩2.0.CO;2.

Frederiksen, J. S., M. R. Dix, and A. G. Davies, 2003: The effects of closure-based eddy diffusion on the climate and spectra of a gcm. *Tellus A*, **55 (1)**, 31–44, doi:10.1034/j.1600-0870.2003.201329.x, URL https://onlinelibrary.wiley.com/doi/abs/10.1034/j.1600-0870.2003.201329.x.

Frederiksen, J. S., and S. M. Kepert, 2006: Dynamical Subgrid-Scale Parameterizations from Direct Numerical Simulations. *Journal of Atmospheric Sciences*, **63**, 3006–3019, doi:10.1175/JAS3795.1.

Gardiner, C., 2009: *Stochastic Methods*. 4th ed., Springer-Verlag Berlin Heidelberg, 447 pp.

Haidvogel, D. B., and A. Beckmann, 1999: *Numerical Ocean Circulation Modeling*. Imperial College Press, London.

Hasselmann, K., 1976: Stochastic climate models Part I. Theory. *Tellus*, **28 (6)**, 473–485, doi:10.3402/tellusa.v28i6.11316, URL http://link.springer.com/10.1007/978-3-0348-8287-3http:

//tellusa.net/index.php/tellusa/article/view/11316https://www.tandfonline.com/doi/full/10.3402/tellusa.v28i6.11316.

Honerkamp, J., 1994: *Stochastic Dynamical Systems: Concepts, Numerical Methods, Data Analysis*. Wiley-VCH.

Imkeller, P., and J.-S. von Storch, Eds., 2001: *Stochastic Climate Models*. Birkhäuser Basel, Basel, doi:10.1007/978-3-0348-8287-3, URL http://link.springer.com/10.1007/978-3-0348-8287-3.

Khasminsky, R. Z., 1966a: A limit theorem for the solutions of differential equations with random right-hand sides. *Theory Prob. Applications*, **11**, 390–406.

Khasminsky, R. Z., 1966b: On stochastic processes defined by differential equations with a small parameter. *Theory Prob. Applications*, **11**, 211–228.

Kitsios, V., J. S. Frederiksen, and M. J. Zidikheri, 2012: Subgrid Model with Scaling Laws for Atmospheric Simulations. *Journal of Atmospheric Sciences*, **69**, 1427–1445, doi:10.1175/JAS-D-11-0163.1.

Kitsios, V., J. S. Frederiksen, and M. J. Zidikheri, 2013: Scaling laws for parameterisations of subgrid eddy-eddy interactions in simulations of oceanic circulations. *Ocean Modelling*, **68**, 88–105, doi:10.1016/j.ocemod.2013.05.001.

Kondrashov, D., S. Kravtsov, A. W. Robertson, and M. Ghil, 2005: A hierarchy of data-based ENSO models. *Journal of Climate*, **18 (21)**, 4425–4444, doi:10.1175/JCLI3567.1.

Kraichnan, R. H., 1959: The structure of isotropic turbulence at very high reynolds numbers. *Journal of Fluid Mechanics*, **5 (4)**, 497–543, doi:10.1017/S0022112059000362.

Kravtsov, S., D. Kondrashov, and M. Ghil, 2005: Multilevel Regression Modeling of Nonlinear Processes: Derivation and Applications to Climatic Variability. *Journal of Climate*,

687 **18 (21)**, 4404–4424, doi:10.1175/JCLI3544.1, URL http://journals.ametsoc.org/doi/abs/10.
688 1175/JCLI3544.1.

689 Kurtz, T. G., 1973: A limit theorem for perturbed operator semigroups with applications to random
690 evolution. *J. Funct. Anal.*, **12**, 55–67.

691 Lutsko, N. J., I. M. Held, and P. Zurita-Gotor, 2015: Applying the FluctuationDissipation The-
692 orem to a Two-Layer Model of Quasigeostrophic Turbulence. *Journal of the Atmospheric Sci-*
693 *ences*, **72 (8)**, 3161–3177, doi:10.1175/JAS-D-14-0356.1, URL http://journals.ametsoc.org/doi/
694 abs/10.1175/JAS-D-14-0356.1.

695 Majda, A., I. Timofeyev, and E. Vanden-Eijnden, 2002: A priori tests of a stochastic mode
696 reduction strategy. *Physica D: Nonlinear Phenomena*, **170 (3-4)**, 206–252, doi:10.1016/
697 S0167-2789(02)00578-X.

698 Majda, A. J., C. Franzke, and D. Crommelin, 2009: Normal forms for reduced stochastic climate
699 models. *Proceedings of the National Academy of Sciences of the United States of America*,
700 **106 (10)**, 3649–53, doi:10.1073/pnas.0900173106, URL http://www.ncbi.nlm.nih.gov/pubmed/
701 19228943{\%}5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2645348.

702 Majda, A. J., I. Timofeyev, and E. V. Eijnden, 2001: A mathematical framework for stochastic
703 climate models. *Communications on Pure and Applied Mathematics*, **54 (8)**, 891–974, doi:
704 10.1002/cpa.1014.

705 Majda, A. J., I. Timofeyev, and E. Vanden-Eijnden, 2003: Systematic Strategies for Stochastic
706 Mode Reduction in Climate. *Journal of the Atmospheric Sciences*, **60 (14)**, 1705–1722, doi:
707 10.1175/1520-0469(2003)060⟨1705:SSFSMR⟩2.0.CO;2.

Majewski, D., and Coauthors, 2002: The Operational Global IcosahedralHexagonal Gridpoint Model GME: Description and High-Resolution Tests. *Monthly Weather Review*, **130 (2)**, 319–338, doi:10.1175/1520-0493(2002)130⟨0319:TOGIHG⟩2.0.CO;2, URL http://journals.ametsoc.org/doi/abs/10.1175/1520-0493(2002)130{\%}3C0319:TOGIHG{\%}3E2.0.CO;2.

Monahan, A. H., and J. Culina, 2011: Stochastic averaging of idealized climate models. *Journal of Climate*, **24 (12)**, 3068–3088, doi:10.1175/2011JCLI3641.1.

Newman, M., P. D. Sardeshmukh, C. R. Winkler, and J. S. Whitaker, 2003: A Study of Subseasonal Predictability. *Monthly Weather Review*, **131 (8)**, 1715–1732, doi:10.1175//2558.1, URL http://journals.ametsoc.org/doi/abs/10.1175//2558.1.

Palmer, T. N., 2001: A nonlinear dynamical perspective on model error: A proposal for non-local stochastic-dynamic parametrization in weather and climate prediction models. *Quarterly Journal of the Royal Meteorological Society*, **127 (572)**, 279–304, doi:10.1002/qj.49712757202, URL http://doi.wiley.com/10.1002/qj.49712757202.

Palmer, T. N., R. Buizza, F. Doblas-Reyes, T. Jung, M. Leutbecher, G. J. Shutts, M. Steinheimer, and A. Weisheimer, 2009: Stochastic Parametrization and Model Uncertainty. *ECMWF Technical Memoranda*, **598**, 1–42, URL http://www.ecmwf.int/publications/.

Papanicolaou, G., 1976: Some probabilistic problems and methods in singular perturbations. *Rocky Mountain J. Math*, **6**, 653–673.

Pavliotis, G. A., and A. M. Stuart, 2008: *Multiscale Methods*, Texts Applied in Mathematics, Vol. 53. Springer New York, New York, NY, doi:10.1007/978-0-387-73829-1, URL http://link.springer.com/10.1007/978-0-387-73829-1.

729 Pegion, K., and P. D. Sardeshmukh, 2011: Prospects for Improving Subseasonal Predictions.
730 *Monthly Weather Review*, **139 (11)**, 3648–3666, doi:10.1175/MWR-D-11-00004.1, URL http:
731 //journals.ametsoc.org/doi/abs/10.1175/MWR-D-11-00004.1.

732 Petoukhov, V., A. Ganopolski, V. Brovkin, M. Claussen, A. Eliseev, C. Kubatzki, and S. Rahm-
733 storf, 2000: CLIMBER-2: A Climate System Model of Intermediate Complexity. Part I: Model
734 Description and Performance for Present Climate. *Climate Dynamics*, **16**, 1–17.

735 Pieroth, M., S. I. Dolaptchiev, M. Zacharuk, A. Gritsun, and U. Achatz, 2018: Climate-
736 Dependence in Empirical Parameters of Subgrid-Scale Parameterizations using the Fluctuation-
737 Dissipation Theorem. *Journal of the Atmospheric Sciences*, submitted.

738 Pope, S. B., 2000: *Turbulent Flows*. Cambridge University Press, Cambridge, doi:10.1017/
739 CBO9780511840531, URL http://ebooks.cambridge.org/ref/id/CBO9780511840531.

740 Rípodas, P., and Coauthors, 2009: Icosahedral Shallow Water Model (ICOSWM): results of shal-
741 low water test cases and sensitivity to model parameters. *Geoscientific Model Development*
742 *Discussions*, **2 (1)**, 581–638, doi:10.5194/gmd-2-231-2009.

743 Satoh, M., T. Matsuno, H. Tomita, H. Miura, T. Nasuno, and S. Iga, 2008: Nonhydrostatic icosa-
744 hedral atmospheric model (NICAM) for global cloud resolving simulations. *Journal of Compu-*
745 *tational Physics*, **227 (7)**, 3486–3514, doi:10.1016/j.jcp.2007.02.006.

746 Shutts, G., 2005: A kinetic energy backscatter algorithm for use in ensemble prediction systems.
747 *Quarterly Journal of the Royal Meteorological Society*, **131 (612)**, 3079–3102, doi:10.1256/qj.
748 04.106.

749 Skamarock, W., and Coauthors, 2008: A Description of the Advanced Research WRF Version 3.
750 Tech. Rep. NCAR/TN-475+STR, NCAR.

751 Stull, R. B., Ed., 1988: *An Introduction to Boundary Layer Meteorology*. Springer Nether-
752 lands, Dordrecht, doi:10.1007/978-94-009-3027-8, URL http://link.springer.com/10.1007/
753 978-94-009-3027-8.

754 Vallis, G. K., 2006: *Atmospheric and Oceanic Fluid Dynamics: Fundamentals and Large-Scale*
755 *Circu- lation*. 747 pp.

756 Verkley, W. T. M., 2011: A maximum entropy approach to the problem of parametrization. *Quar-*
757 *terly Journal of the Royal Meteorological Society*, **137 (660)**, 1872–1886, doi:10.1002/qj.860.

758 Verkley, W. T. M., P. C. Kalverla, and C. A. Severijns, 2016: A maximum entropy approach to the
759 parametrization of subgrid processes in two-dimensional flow. *Quarterly Journal of the Royal*
760 *Meteorological Society*, **142 (699)**, 2273–2283, doi:10.1002/qj.2817.

761 Verkley, W. T. M., and C. A. Severijns, 2014: The maximum entropy principle applied to a dy-
762 namical system proposed by Lorenz. *European Physical Journal B*, **87 (1)**, doi:10.1140/epjb/
763 e2013-40681-2.

764 Vissio, G., and V. Lucarini, 2017: A proof of concept for scale-adaptive parameterizations: the
765 case of the Lorenz '96 model. *Quarterly Journal of the Royal Meteorological Society*, doi:
766 10.1002/qj.3184.

767 Weaver, A. J., and Coauthors, 2001: The UVic earth system climate model: Model description,
768 climatology, and applications to past, present and future climates. *Atmosphere-Ocean*, **39 (4)**,
769 361–428, doi:10.1080/07055900.2001.9649686, URL http://www.tandfonline.com/doi/abs/10.
770 1080/07055900.2001.9649686.

771 Winkler, C. R., M. Newman, and P. D. Sardeshmukh, 2001: A linear model of wintertime low-
772 frequency variability. Part I: Formulation and forecast skill. *Journal of Climate*, **14 (24)**, 4474–
773 4494, doi:10.1175/1520-0442(2001)014⟨4474:ALMOWL⟩2.0.CO;2.

774 Wouters, J., S. Iankov Dolaptchiev, V. Lucarini, and U. Achatz, 2016: Parameterization of
775 stochastic multiscale triads. *Nonlinear Processes in Geophysics*, **23 (6)**, 435–445, doi:10.5194/
776 npg-23-435-2016.

777 Wouters, J., and V. Lucarini, 2012: Disentangling multi-level systems: averaging, correlations and
778 memory. *Journal of Statistical Mechanics: Theory and Experiment*, **2012 (03)**, P03 003, doi:10.
779 1088/1742-5468/2012/03/P03003, URL http://stacks.iop.org/1742-5468/2012/i=03/a=P03003,
780 arXiv:1110.6113v1.

781 Wouters, J., and V. Lucarini, 2013: Multi-level Dynamical Systems: Connecting the Ruelle Re-
782 sponse Theory and the Mori-Zwanzig Approach. *Journal of Statistical Physics*, **151 (5)**, 850–
783 860, doi:10.1007/s10955-013-0726-8, 1208.3080.

LIST OF TABLES

TABLE 1. First column in both tables shows the spatially averaged second ($k = 2$) and fourth ($k = 4$) order centered statistical moments values of DNS. The other columns contain the relative errors of different low-resolution simulations. Upper table for H and lower table for HU.

H	DNS	LRM	BRT	LRM-OU	BRT-OU	BRT-SMR	BRT-SMR-0.6	BRT-SMR-0.0
$k = 2$	2.846 $(\mathrm{km})^2$	0.138	-0.167	-0.589	-0.125	0.195	0.029	-0.049
$k = 4$	23.56 $(\mathrm{km})^4$	0.542	-0.290	-0.821	-0.207	0.455	0.080	-0.079

HU	DNS	LRM	BRT	LRM-OU	BRT-OU	BRT-SMR	BRT-SMR-0.6	BRT-SMR-0.0
$k = 2$	2082 $(10^3(\mathrm{km})^2/\mathrm{d})^2$	0.255	-0.159	-0.597	-0.141	0.099	-0.052	-0.124
$k = 4$	1.277E+07 $(10^3(\mathrm{km})^2/\mathrm{d})^4$	1.119	-0.274	-0.834	-0.251	0.185	-0.119	-0.247

TABLE 2. Summary of the different models used

DNS	Direct Numerical Simulation with spatial resolution of N cells $\dot{x}_i = \rho_i^x + a_i^x(\mathbf{x}) + b_i^{xz}(\mathbf{x},\mathbf{z})$; $\dot{z}_i = b_i^z(\mathbf{x},\mathbf{z}) + c_i^z(\mathbf{z})$
BRT	Bare-Truncation Model with spatial resolution of $N_c = N/n$ $\dot{x}_i = \rho_i^x + a_i^x(\mathbf{x})$
LRM	Low-Resolution Model DNS with spatial resolution of $Nc = N/n$.
OU-DNS	DNS with $\frac{1}{h}$-approximation and SGS self-interactions replaced by an OU process $\dot{x}_i = \rho_i^x + a_i^x(\mathbf{x}) + b_i^x(\mathbf{x},\mathbf{y})$, $\dot{y}_i = b_i^y(\mathbf{x},\mathbf{y}) + \Lambda_{ij} y_j + \Sigma_i \dot{W}_i$.
BRT-OU	BRT with empirical Ornstein-Uhlenbeck parameterization $dx_i = \left(\rho_i^x + a_i^x(\mathbf{x}) + \tilde{\Gamma}_{ij}\hat{x}_j^I\right) dt + \tilde{\sigma}_i dW_i$.
LRM-OU	LRM with empirical Ornstein-Uhlenbeck parameterization
BRT-SMR	BRT with Stochastic Mode Reduction parameterization $dx_i = [\rho_i^x + a_i^x(\mathbf{x}) + \beta_i(\mathbf{x})]\, dt + d\xi_i(\mathbf{x})$
BRT-SMR-0.6	BRT-SMR with stochastic forcing reduced to 60% $dx_i = [\rho_i^x + a_i^x(\mathbf{x}) + \beta_i(\mathbf{x})]\, dt + 0.6 d\xi_i(\mathbf{x})$.
BRT-SMR-0.0	BRT-SMR without stochastic forcing $dx_i = [\rho_i^x + a_i^x(\mathbf{x}) + \beta_i(\mathbf{x})]\, dt$

LIST OF FIGURES

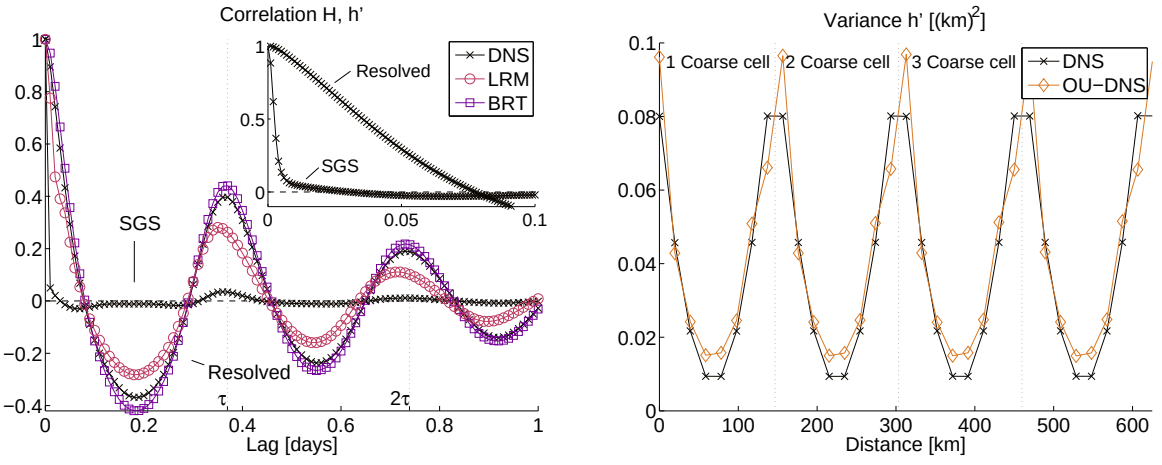


FIG. 1. Left: spatially averaged time autocorrelation of the resolved variable H and SGS variable h'. Results from the DNS and two low-resolution simulations without SGS parameterizations (LRM and BRT). On the time axis the characteristic gravity wave time τ is marked, see Sec. 4 a. In the upper right corner a shorter time interval is presented, which resolves the decay of h'. Right: the variance of the SGS variable h' for 32 fine grid points and an averaging interval $n = 8$. Results from the DNS and OU-DNS.

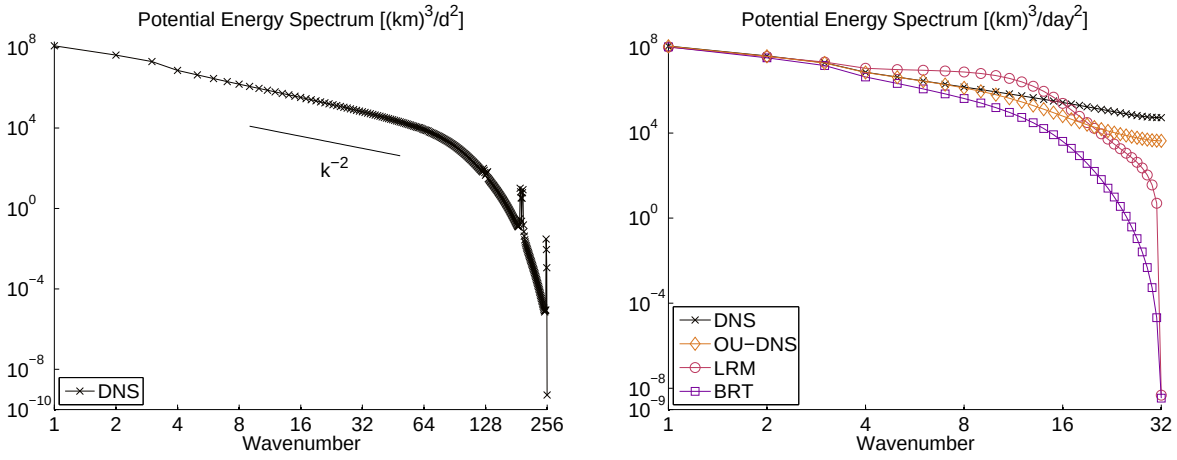
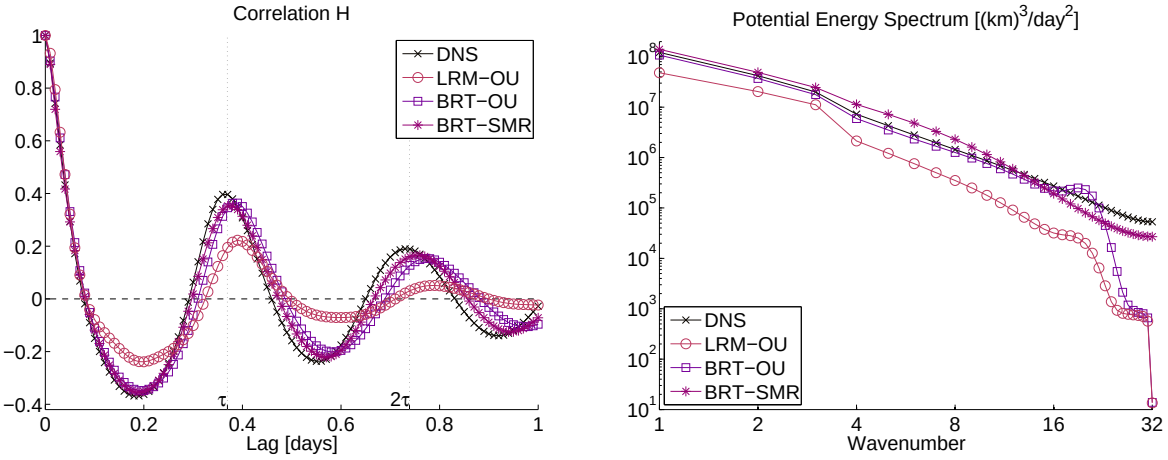
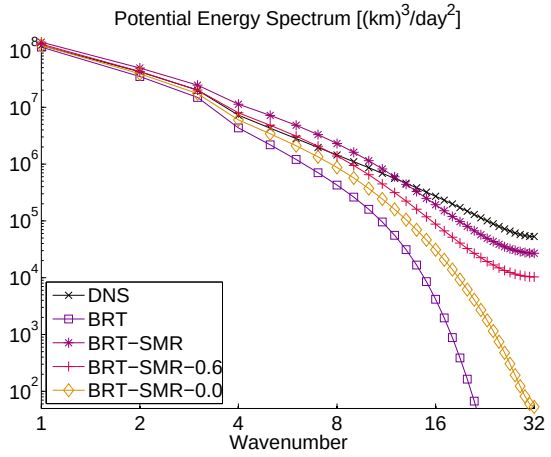


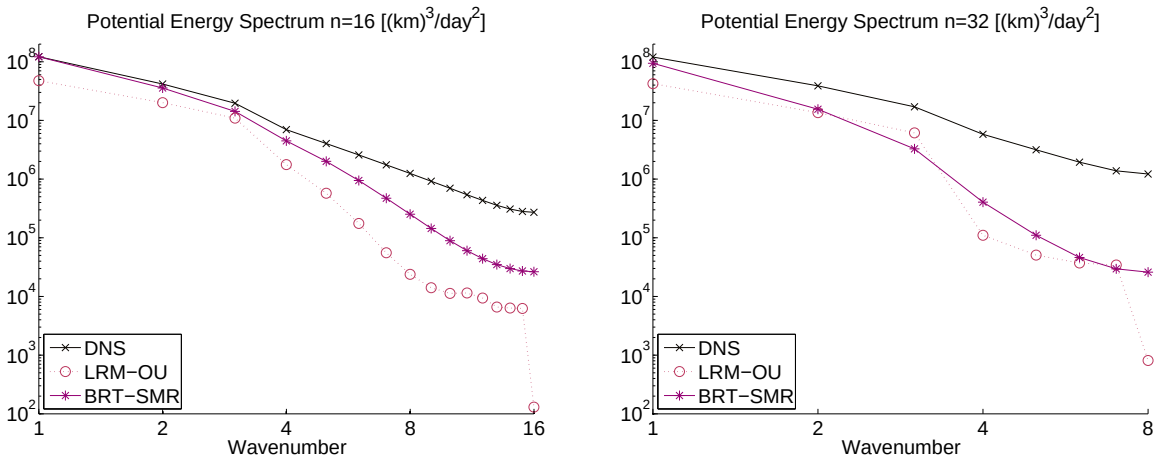
FIG. 2. Left: the potential energy spectrum computed from h in the DNS. Right: the potential energy spectrum computed from H in the DNS, OU-DNS and two low-resolution simulations without SGS parameterizations (LRM and BRT).



820 FIG. 3. The spatially averaged time autocorrelation (left) and the potential energy spectrum (right) from DNS
821 and three low-resolution simulations with SGS parameterizations (LRM-OU, BRT-OU, BRT-SMR).



822 FIG. 4. The potential energy spectrum in DNS, BRT, BRT-SMR, BRT-SMR with a damped stochastic forcing
823 $d\xi \to 0.6d\xi$ (BRT-SMR-0.6) and BRT-SMR with neglected stochastic forcing $d\xi \to 0$ (BRT-SMR-0.0).



824 FIG. 5. The Spectrum for an averaging interval of $n = 16$ (left) and $n = 32$ (right) in DNS, LRM-OU and
825 BRT-SMR. The simulations with LRM-OU and BRT-SMR have a resolution of $N_c = 32$ (left) and $N_c = 16$
826 (right).