Climate-Dependence in Empirical Parameters of Subgrid-Scale Parameterizations using the Fluctuation-Dissipation Theorem

Martin Pieroth*, Stamen I. Dolaptchiev, and Matthias Zacharuk

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

Tobias Heppelmann

Deutscher Wetterdienst, Offenbach am Main, Germany

Andrey Gritsun

Institute of Numerical Mathematics, Russian Academy of Sciences, Moscow, Russia

Ulrich Achatz

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

Corresponding author address: Martin Pieroth, Institut für Atmosphäre und Umwelt, Goethe-Universität Frankfurt, Altenhöferallee 1, D-60438 Frankfurt am Main, Germany.

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

ABSTRACT

Many subgrid-scale (SGS) parameterizations in climate models contain empirical parameters and are thus data dependent. In particular, it is not guaranteed that the SGS parameterization still helps the model to produce the correct climate projection in the presence of an external perturbation (e.g., due to climate change). Therefore, a climate dependence of tuning parameters is proposed, using the fluctuation-dissipation theorem (FDT). The FDT provides an estimation of the changes in the statistics of a system, caused by a small external forcing. These estimations are then used to update the SGS parameterization. This procedure is tested for a toy atmosphere given by a quasi-geostrophic three-layer model (QG3LM). We construct a low-order climate model for this toy atmosphere, based on a reduced number of its empirical orthogonal functions (EOFs), equipped with either an empirical deterministic or an empirical stochastic SGS parameterization. External forcings are considered that are either a local anomalous heat source in the extratropics or a global dynamical forcing represented by individual EOF patterns. A quasi-Gaussian variant of the FDT is able to successfully update the SGS parameterization leading to an improvement in both amplitude and correlation between the low-order climate model and the QG3LM, in case of a perturbed system. The stochastic closure exhibits nearly no improvement compared to the deterministic parameterization. The application of a more sophisticated non-Gaussian FDT algorithm (i.e., the blended Short-time/quasi-Gaussian FDT) yields only marginal improvement over the simple quasi-Gaussian FDT.

1. Introduction

One key problem in climate modeling is the response of the climate system to an external forcing. In particular, the anthropogenic influence on the climate is of interest. This includes the global mean temperature, the sea level, precipitation patterns, or atmospheric phenomena such as ENSO. The current procedure to estimate such a response is by means of sensitivity studies using general circulation models (GCMs). However, due to the complexity of GCMs, those sensitivity studies are quite expensive.

It might be tempting to use alternatively so-called low-order climate models, based, for example, on optimal basis patterns (Achatz et al. 1995; Kwasniok 1996; Selten 1997; Achatz and Schmitz 1997; Achatz and Opsteegh 2003a,b; Kwasniok 2004, 2007). The result is a reduction of dimension while the overall quality is supposed to stay comparable to that of a regular GCM. Due to their lower state-space dimension such models are potentially less demanding on computing power than the latter. The dimension reduction also entails, however, as in any multi-scale system, the use of additional subgrid-scale (SGS) parameterizations that take the effect of neglected basis patterns into account.

Such SGS parameterizations are often data-driven (e.g., Achatz and Branstator 1999; Achatz and Opsteegh 2003a,b; Kravtsov et al. 2009). That is, a suitable function is tuned against simulated data (or real observations) to minimize the residual error between the low-order climate model and the data. However, such a data-driven SGS parameterization is inherently dependent on the data against which it is tuned. This was shown by Achatz and Branstator (1999) for a low-order model consisting of a filtered two-layer model projected onto empirical orthogonal functions (EOFs) of a GCM and a simple linear SGS parameterization. The reduced model was successful in capturing the characteristics of the GCM dynamics, such as the average transient eddy fluxes and the

statistics of the first and second moment. However, the low-order model was not able to capture the dynamical changes due to the presence of a local anomalous heating in the tropics. A similar result was found by Achatz and Opsteegh (2003b) for a reduced EOF model based on the primitive equations.

This problem is not limited to low-order models. Even state-of-the-art GCMs contain empirical parameters. Those models also employ SGS parameterizations, partially based on physical considerations, but nonetheless are usually highly tuned to the present-day climate. Consequently, they suffer from data dependence as well. One could argue that the effect of data dependence in a physical SGS parameterization is negligible. However, the fact that the tuning of GCMs and climate models is a common and even necessary step underlines the impact of the empirical parameters. One example for the dependence on the empirical parameters is the study of Rockel and Geyer (2008) which shows that regional climate models cannot be transferred to a different climatic zone without retuning.

There are approaches to derive SGS parameterizations directly from model equations. Verkley and Severijns (2014) introduced a SGS parameterization based on maximum entropy principle (Verkley et al. 2016) which does not require any fitting. In the stochastic mode reduction (Majda et al. 2001, 2003) the nonlinear self-interaction of the SGS processes is described by an empirical Ornstein-Uhlenbeck process. This Ornstein-Uhlenbeck process is then used to derive a stochastic SGS parameterization explicitly from the model equations. Franzke et al. (2005) introduced a seamless variant of the stochastic mode reduction which drops the assumption of an empirical Ornstein-Uhlenbeck process. Consequently, the required lag correlations for the stochastic SGS closure cannot be calculated analytically from the Ornstein-Uhlenbeck process but are obtained directly from data. Either way, while there is still some tuning involved, this explicitly derived SGS closure should be more robust to climate change than traditional parameterizations. Nevertheless,

the stochastic mode reduction requires a strict time-scale separation which is not necessarily given in (atmospheric) GCMs (Franzke and Majda 2006). An alternative method based on response theory, which instead of a time-scale separation relies on weak coupling, was proposed by Wouters and Lucarini (2012) and used to derive parameterizations (Wouters et al. 2016; Demaeyer and Vannitsem 2017). Recently, Lucarini and Wouters (2017) derived explicit expressions on how this parameterization changes when the background state of the system is perturbed. This response-theory ansatz is conceptually close to the approach we take in this study.

In particular, in this paper we follow the route of Achatz et al. (2013). For cases where the tuning parameters of the SGS parameterization are based on the statistics of the system in question, the authors propose to use the fluctuation-dissipation theorem (FDT, see Marconi et al. 2008, for a recent review) to estimate the change of the statistics due to an external forcing. In general, the FDT connects a suitable covariance function of a dynamical system at statistical equilibrium with the response of the system, caused by a small external forcing or a sufficiently small change in model parameters (Risken 1984). Since the FDT only depends on a covariance function no detailed knowledge of the governing equations of the system is required. However, for the FDT to work several constraints have to be fulfilled. In general, the system in question must be in statistical equilibrium and have a time-invariant probability density function (PDF) which must be differentiable. Although extensions to nonequilibrium systems (Lucarini and Sarno 2011; Ragone et al. 2016) and to time periodic cases are possible (Majda and Wang 2010; Gritsun 2010), the condition of differentiability is generally violated in systems which exhibit deterministic chaos. Usually, the system attractor of such systems is fractal. The differentiablity can be ensured if some suitable random noise is added to the system to smooth the PDF (Zeeman 1988).

Even though the atmosphere and the climate system are not strictly consistent with the conditions above, Leith (1975) suggested the use of the FDT for an approximation of the climate

response. Since then numerous studies have applied the so-called quasi-Gaussian FDT (qG-FDT, Majda et al. 2005, where the equilibrium PDF is assumed to be close to Gaussian) to data from various idealized climate models (e.g, Bell 1980; North et al. 1993; Gritsoun and Dymnikov 1999; Gershgorin and Majda 2010; Achatz et al. 2013; Fuchs et al. 2015; Lutsko et al. 2015), simple GCMs (Gritsun and Branstator 2007; Gritsun et al. 2008; Ring and Plumb 2008), and even coupled atmosphere-ocean GCMs (Gritsun and Branstator 2016). However, all studies showed mixed results regarding the FDT performance. It remains unclear why this is the case. A recent study of Hassanzadeh and Kuang (2016) indicated that this could be due to nonnormal response operators. This nonnormality might lead to a strong interaction between resolved and unresolved EOFs which cannot be captured by an FDT-response operator in an EOF subspace. Another possibility is that the forcing might project onto a stable direction of the attractor, resulting in a response that is not covered by the unperturbed model PDF (Gritsun and Lucarini 2017). Furthermore, it is possible that the Gaussian assumption may contribute to the error more than expected. Therefore, Cooper and Haynes (2011) suggest to relax the Gaussian approximation of the equilibrium PDF by the use of a nonparametric kernel method. Alternatively, Abramov and Majda (2007, 2008, 2009) introduced a blended Short-time/quasi-Gaussian (ST/qG-FDT) algorithm which avoids the differentiation of the PDF by use of a tangent linear model. Extensions of non-Gaussian response theory to higher orders in the perturbation amplitude are possible as well (Lucarini and Wouters 2017). While those methods generally yield superior results, compared to the simple qG-FDT, they are significantly more expensive.

Therefore, Achatz et al. (2013) decided to use the qG-FDT. The authors considered a toy atmosphere based on the barotropic vorticity equation on the sphere and a corresponding reduced model based on EOFs. The linear SGS parameterization of the latter model is determined completely by first and second moments of the toy atmosphere. In addition, they used a local anomalous heating

located in the midlatitudes to investigate the robustness of the closure due to a changing climate. Their qG-FDT estimations of the first moments were reasonably good. However, the second moments exhibited a considerably higher error. Consequently, the resulting changes in the tuning parameters of the closure were not usable. Therefore, Achatz et al. (2013) introduced a reduced form of qG-FDT, basically discarding the estimations for the second moments, resulting in reasonably good closure updates. Furthermore, they showed that the use of the reduced-qG-FDT (rqG-FDT) modified closure yielded a better agreement with the reference model than the unmodified reduced model, in the case of a perturbed climate. Moreover, for sufficiently high EOF dimension the reduced model with modified SGS parameterization even outperformed the direct qG-FDT estimation. Nevertheless, due to the negligence of the qG-FDT estimation of the second moments, only a part of the SGS parameterization could be updated. The failure of the qG-FDT could be associated with the constraints of the theorem. The nonlinear dynamics of the smallest-scale processes might not act as a stochastic forcing and thus a sufficiently smooth PDF may not be given, since the toy atmosphere has no clear time-scale separation. Furthermore, only barotropic Rossby waves and no fast synoptic-scale processes (e.g., baroclinic instability) or gravity waves are present in this model.

Consequently, the present study pursues the approach of Achatz et al. (2013) further, however, now applies it to a baroclinic toy atmosphere, given by a quasi-geostrophic three-layer model (QG3LM). As it turns out, the increased complexity of the model and the presence of baroclinic instability leads to an overall better performance of the qG-FDT, allowing a successful update of the closure parameters.

This paper is organized as follows. In section 2 the QG3LM and the low-order climate model, based on the leading EOFs of the QG3LM, are introduced. Furthermore, data-driven SGS parameterizations are constructed. The models are perturbed by both local anomalous forcings and global

anomalous forcings presented in section 3. Moreover, we recapitulate the qG-FDT and explain the climate-dependence of the SGS parameterization in more detail. Afterwards we present the results of the experiments in section 4. Section 5 summarizes and discusses the findings and gives some conclusions.

2. Models

a. The quasi-geostrophic three-layer model

In this paper the QG3LM of Marshall and Molteni (1993) is considered as a toy atmosphere. The QG3LM is a model of medium complexity containing baroclinic Rossby waves. The model is governed by the quasi-geostrophic potential vorticity equation on the sphere

$$\frac{\partial q_i}{\partial t} = -J(\Psi_i, q_i) + D_i(\Psi_{i-1}, \Psi_i, \Psi_{i+1}) + S_i \tag{1}$$

$$q_i = \nabla^2 \Psi_i + \eta_i + f\left(1 + \delta_{i,3}\frac{h}{H}\right), \tag{2}$$

at the pressure levels 200 hPa, 500 hPa, and 800 hPa, denoted by $i = 1,2,3$, respectively. Here q is the quasi-geostrophic potential vorticity; Ψ is the streamfunction; $J(\cdot,\cdot)$ denotes the standard Jacobian operator; D is a function containing temperature relaxation, Ekman friction, and hyper-diffusion. Furthermore, η denotes the stretching vorticity, f is the Coriolis parameter, and h/H is the normalized orography (cf., Marshall and Molteni 1993). The constant vorticity forcing S is tuned against 10 winter seasons of reanalysis data from ECMWF (Liu and Opsteegh 1995) which enables the model to simulate a realistic northern hemispheric winter.

The QG3LM has a spectral discretization with a triangular truncation of T21. This leads to $N =$ 1449 degrees of freedom. The model is integrated forward in time using the leapfrog scheme with a time step of $\Delta t = 1/36$ days. For all following results a daily model output is used. Furthermore, for all integrations the first 10,000 days have been discarded to eliminate spinup effects.

Equation (1) can be rewritten by the introduction of a state vector $\mathbf{x} \in \mathbb{R}^N$, i.e.

$$\frac{d\mathbf{x}}{dt} = \mathbf{G}(\mathbf{x}), \tag{3}$$

where the function $\mathbf{G}(\cdot)$ represents the right-hand side of (1).

b. The low-order climate model

For the construction of the low-order climate model the QG3LM is projected onto EOF space. In particular, the deviation $\mathbf{x}'$ of the state vector from its time-mean $\langle \mathbf{x} \rangle$ (i.e., $\mathbf{x}' = \mathbf{x} - \langle \mathbf{x} \rangle$, where $\langle \cdot \rangle$ denotes a time average) is projected onto the leading M ($M < N$) EOFs. Therefore, the state vector is expressed as

$$\begin{aligned} \mathbf{x} &= \langle \mathbf{x} \rangle + \mathbf{x}' && (4) \\ &= \langle \mathbf{x} \rangle + \sum_{k=1}^{M} a_k(t)\mathbf{e}_k + \boldsymbol{\epsilon}(t) && (5) \\ &= \langle \mathbf{x} \rangle + \mathbf{P}\mathbf{a}(t) + \boldsymbol{\epsilon}(t), && (6) \end{aligned}$$

where $\mathbf{a} \in \mathbb{R}^M$ is the so-called principal component vector, $\mathbf{P} \in \mathbb{R}^{N \times M}$ is a matrix containing the leading M EOFs $\mathbf{e}$ as columns, and $\boldsymbol{\epsilon} \in \mathbb{R}^N$ is the time-dependent truncation error.

The EOFs are defined by the eigenvalue problem

$$\left\langle \mathbf{x}'\mathbf{x}'^{\mathrm{T}} \right\rangle \mathbf{M}\mathbf{e}_k = \lambda_k \mathbf{e}_k, \tag{7}$$

where λ_k denotes the explained energy variance of the QG3LM by the kth EOF and $\mathbf{M}$ is the total energy metric of Ehrendorfer (2000). The exact formula of $\mathbf{M}$ is given in appendix A. The EOFs have been calculated from a 100,000 day integration of the QG3LM. It turns out that 540 EOFs suffice to explain 90% of the variance of the analyzed data. Spectra with this kind of flatness are typical when an energy metric is applied to data which is not filtered in time (e.g., Achatz and Branstator 1999; Achatz and Opsteegh 2003a).

Using the decomposition of (6) in the QG3LM (3) yields

$$\frac{d\mathbf{a}}{dt} = \mathbf{P}^{\mathrm{T}}\mathbf{M}\mathbf{G}(\langle\mathbf{x}\rangle + \mathbf{P}\mathbf{a}) + \mathbf{s}(\mathbf{a},\mathbf{x})\,, \tag{8}$$

where $\mathbf{s}$ is the SGS tendency error resulting from the truncation error within $\mathbf{G}(\cdot)$.

A suitable parameterization $p(\mathbf{a})$ is required to address the SGS error. This results in the low-order model equation

$$\frac{d\mathbf{a}}{dt} = \tilde{\mathbf{G}}(\mathbf{a}) + p(\mathbf{a}) + \varepsilon_p(\mathbf{a},\mathbf{x})\,, \tag{9}$$

where $\tilde{\mathbf{G}}(\mathbf{a}) = \mathbf{P}^{\mathrm{T}}\mathbf{M}\mathbf{G}(\langle\mathbf{x}\rangle + \mathbf{P}\mathbf{a})$ and $\varepsilon_p(\mathbf{a},\mathbf{x})$ denotes the parameterization error (i.e., $\varepsilon_p(\mathbf{a},\mathbf{x}) = \mathbf{s}(\mathbf{a},\mathbf{x}) - p(\mathbf{a})$). To address the SGS error, we choose a linear deterministic parameterization (Achatz and Branstator 1999; Achatz and Opsteegh 2003a), i.e.

$$p_{\mathrm{det}}(\mathbf{a}) = \mathbf{r} + \mathbf{L}\mathbf{a}\,, \tag{10}$$

where $\mathbf{r} \in \mathbb{R}^M$ and $\mathbf{L} \in \mathbb{R}^{M\times M}$ are a constant vector and matrix, respectively. Minimizing the norm of the parameterization error (i.e., $\langle \| \varepsilon_p \|^2 \rangle$) yields as optimal closure parameters

$$\begin{aligned} \mathbf{L} &= \langle \mathbf{s}'\mathbf{a}'^{\mathrm{T}} \rangle \langle \mathbf{a}'\mathbf{a}'^{\mathrm{T}} \rangle^{-1} \\ \mathbf{r} &= \langle \mathbf{s} \rangle - \mathbf{L}\langle \mathbf{a} \rangle\,, \end{aligned} \tag{11}$$

where $\mathbf{s}(t) = [\mathbf{a}(t+\Delta t) - \mathbf{a}(t-\Delta t)]/(2\Delta t) - \tilde{\mathbf{G}}(\mathbf{a}(t))$ denotes the SGS tendency error approximated by centered differences in time. Thus, the closure is completely determined by first and second statistical moments of the QG3LM. For the calculation of the closure parameters $\mathbf{r}$ and $\mathbf{L}$, a $6\cdot 10^6$ day integration of the QG3LM is used to safely exclude sampling errors. This huge amount of data is not necessary for a reliable application of the FDT, but needed for the convergence of the response due to the external forcing (see section 5). Hence only for consistency we use a time series of the same length here.

Furthermore, we consider a stochastic parameterization given by an Ornstein-Uhlenbeck process

$$p_{\text{stoch}}(\mathbf{a}) = \mathbf{r} + \mathbf{L}\mathbf{a} + \boldsymbol{\Sigma}\dot{\mathbf{W}}\,, \tag{12}$$

where $\boldsymbol{\Sigma} \in \mathbb{R}^{M\times M}$ is a constant diagonal matrix and $\dot{\mathbf{W}}dt = d\mathbf{W} \in \mathbb{R}^M$ denotes an increment of a Wiener process. The noise in (12) (and in all following equations) is interpreted in the sense of Stratonovich (i.e., as a limit of physical noise with vanishing memory). Using a maximum likelihood approach (Honerkamp 1994) to estimate the closure parameters, we obtain for the deterministic part of (12) the same equations as for $p_{\text{det}}(\mathbf{a})$ (11), whereas the optimal noise amplitude is given by

$$\boldsymbol{\Sigma}_{ii} \;=\; \sqrt{2\Delta t \left\langle [s_i - (r_i + \sum_k \mathbf{L}_{ik} a_k)]^2 \right\rangle}\,. \tag{13}$$

In conclusion, the low-order climate model with parameterization reads

$$\frac{d\mathbf{a}}{dt} = \tilde{\mathbf{G}}(\mathbf{a}) + \mathbf{r} + \mathbf{L}\mathbf{a} + \boldsymbol{\Sigma}\dot{\mathbf{W}}\,, \tag{14}$$

where $\boldsymbol{\Sigma} = 0$ in the case of a deterministic closure p_{det} and $\boldsymbol{\Sigma}$ given by (13) in the case of a stochastic parameterization p_{stoch}. Thus, for the remainder of this paper we write the reduced model in this general form.

Generally, such semi-empirical low-order models tend to overestimate the variance (Achatz and Branstator 1999; Achatz and Opsteegh 2003a). To counter this behavior, we tune the hyperdiffusion of the reduced model by adjusting the diffusion time scale τ_H (see Marshall and Molteni 1993) to minimize the relative error (25) of the variance. The numerical values for both the deterministic and stochastic parameterization are given in Table 1.

Figure 1 displays the performance of the reduced model with deterministic parameterization in the 500 EOF case, in terms of both mean (top) and variance (bottom) of streamfunction. In particular, Figure 1 a) and c) show the result of the QG3LM at 200 hPa (where the response is

strongest among all three layers) projected onto the 500 EOF space, whereas Figure 1 b) and d) show the equivalent result of the reduced model (14). In general, the parameterization works quite well: The mean is virtually identical, whereas in the variance only minor differences are visible. For smaller EOF truncations of the low-order models as well as for the models with stochastic parameterization the results are qualitatively similar (not shown).

3. Climate-dependent parameterization

a. External anomalous forcing

The low-order model with parameterization reproduces both qualitatively and quantitatively the QG3LM. However, it remains to be seen how well the reduced model responds to an external forcing. For this purpose we use a time-independent local anomalous heating, representing the effect of a sea surface anomaly due to a change in the ocean circulation (Branstator and Haupt 1998; Achatz and Branstator 1999; Achatz and Opsteegh 2003b; Achatz et al. 2013). As a quasi-geostrophic potential vorticity forcing $\delta\mathbf{S}$, the discretized heating can be written as

$$\delta S_i = -\frac{AR\Delta p}{f_0}\left[(1-\delta_{i,3})\frac{\sin\left(\frac{\pi p_{i+\frac{1}{2}}}{p_s}\right)}{r^2_{i+\frac{1}{2}}p_{i+\frac{1}{2}}} - (1-\delta_{i,1})\frac{\sin\left(\frac{\pi p_{i-\frac{1}{2}}}{p_s}\right)}{r^2_{i-\frac{1}{2}}p_{i-\frac{1}{2}}}\right]\cos^2\left(\frac{\lambda-\lambda_c}{\Delta\lambda}\right)\cos^2\left(\frac{\phi-\phi_c}{\Delta\phi}\right), \quad (15)$$

where $A = 1.25$ K day^{-1} is the amplitude of the heating, R is the universal gas constant, $\Delta p = 300$ hPa denotes the pressure difference of the respective layers, $p_s = 1000$ hPa is the surface pressure, f_0 is the Coriolis parameter at 45N, λ_c, ϕ_c denote the position at which the forcing is centered, and $\Delta\lambda = \Delta\phi = 20^\circ$ is the horizontal extension of the forcing. For $\lambda \notin [\lambda_c - \Delta\lambda/2, \lambda_c + \Delta\lambda/2]$ and $\phi \notin [\phi_c - \Delta\phi/2, \phi_c + \Delta\phi/2]$ the anomalous forcing is set to zero. The pressure in between the layers are given by $p_{1\frac{1}{2}} = 350$ hPa and $p_{2\frac{1}{2}} = 650$ hPa, respectively. Furthermore r denotes the Rossby deformation radius (see Marshall and Molteni 1993, for the exact values).

In contrast to Branstator and Haupt (1998), the anomalous heating is placed in the midlatitudes since the QG3LM exhibits relatively small variances at the equator, resulting in a poor EOF representation of an anomalous forcing placed there. Therefore, for all following experiments the latitude position is fixed at $\phi_c = 45^\circ$, whereas the longitude position varies at $\lambda_c = \{0^\circ, 30^\circ, \ldots, 330^\circ\}$. The most compact model to be investigated is the 20-EOF model. Hence, the anomalous forcing has always been projected onto the first 20 EOFs. Projection onto more EOFs also increases the response amplitude, resulting in nonlinear behavior of the response, especially in the second moments. Perturbing the system with the same external forcing but with opposite signs (i.e., $\pm\delta\mathbf{S}$) should result in a pattern correlation of -1 if the response is purely linear. However, even with restriction onto the first 20 EOFs, this test shows an average pattern correlation of only about -0.75 for the second moments. On the other hand, the first moments exhibit on average a correlation stronger than -0.95.

The left column of Figure 2 shows the equilibrium response of the QG3LM at 200 hPa due to a local anomalous heating at $\lambda_c = 60^\circ$, calculated from a $6 \cdot 10^6$ day time series. Panel a) displays the response in mean streamfunction, whereas panel c) shows the response in mean zonal wind. The maximum values in wind speed are about 2 m s^{-1} which is comparable to the expected change in mean zonal wind speed due to anthropogenic climate change (Lorenz and DeWeaver 2007).

It might be a strong simplification to model a climate change by a single local anomalous forcing placed in the extratropics. One would rather expect that climate change manifests on a global scale. Therefore, in order to drop the locality in the external forcing and thus be more realistic, we are also considering global anomalous forcings given by individual EOFs multiplied by a small constant, i.e. (14) is supplemented by

$$\delta\mathbf{f} = \varepsilon_k \mathbf{e}_k \,, \tag{16}$$

where $\varepsilon_k > 0$ and $k = 1, 2, \ldots, 5$. The ε_k's are chosen in such a way that the turbulent energy $\mathbf{x}'^{\mathrm{T}}\mathbf{M}\mathbf{x}'$ in the response of the global anomalous forcing is comparable to that of the local anomalous forcing ($E_{\mathrm{turb}} \approx 470$ TJ kg^{-1}). The exact numerical values are given in Table 2.

The right column of Figure 2 shows, as a representative example, the response to the global anomalous forcing of EOF 1. The response is similar to the response of the local anomalous forcing. Overall, the amplitude is slightly weaker with maximum zonal wind speeds of about 1.5 m s^{-1}.

b. The perturbed low-order climate model

The simplest ansatz to implement an anomalous forcing in the reduced model is the à-priori (i.e., the naïve) low-order model

$$\frac{d\mathbf{a}}{dt} = \tilde{\mathbf{G}}(\mathbf{a}) + \mathbf{r} + \mathbf{L}\mathbf{a} + \boldsymbol{\Sigma}\dot{\mathbf{W}} + \delta\mathbf{f}, \tag{17}$$

where $\delta\mathbf{f}$ is either given by (16) or $\delta\mathbf{f} = \mathbf{P}_{20}^{\mathrm{T}}\mathbf{M}\delta\mathbf{S}$ (where $\mathbf{P}_{20}$ denotes projection onto the first 20 EOFs) in the case of a global anomalous forcing or local anomalous forcing, respectively.

The performance of the à-priori model is illustrated in Figure 4. Shown is the response in variance of streamfunction for a local anomalous forcing placed at $\lambda_c = 270^\circ$. Panel a) shows the result of the QG3LM projected onto the first 500 EOFs; panel b) shows the result of the 500 EOF à-priori low-order model (17). In the response of the à-priori low-order model the minima over the Pacific ocean and Atlantic ocean are barely visible. Furthermore, the amplitude of the maxima over the Pacific ocean are too strong. On the other hand, the minimum over Greenland is too pronounced. Still, the pattern correlation is reproduced well. Qualitatively similar results are found for the global anomalous forcings (not shown).

The reason for the incorrect response of the à-priori low-order model is the dependence of the closure parameters on the training data-set, which is from the unperturbed model (Achatz et al. 2013; Lucarini and Wouters 2017). The actual reduced climate model is given by

$$\frac{d\mathbf{a}}{dt} = \tilde{\mathbf{G}}(\mathbf{a}) + \mathbf{r} + \delta\mathbf{r} + (\mathsf{L} + \delta\mathsf{L})\mathbf{a} + (\Sigma + \delta\Sigma)\dot{\mathbf{W}} + \delta\mathbf{f}, \tag{18}$$

where $\delta\mathbf{r}$ and $\delta\mathsf{L}$ are the changes in the SGS parameterization due to the altered climate of the model. In particular, the updated closure parameters read

$$\mathsf{L} + \delta\mathsf{L} = \left(\langle \mathbf{s}'\mathbf{a}'^{\mathrm{T}}\rangle + \delta\langle \mathbf{s}'\mathbf{a}'^{\mathrm{T}}\rangle\right)\left(\langle \mathbf{a}'\mathbf{a}'^{\mathrm{T}}\rangle + \delta\langle \mathbf{a}'\mathbf{a}'^{\mathrm{T}}\rangle\right)^{-1} \tag{19}$$

$$\mathbf{r} + \delta\mathbf{r} = \langle \mathbf{s}\rangle + \delta\langle \mathbf{s}\rangle - (\mathsf{L} + \delta\mathsf{L})\left(\langle \mathbf{a}\rangle + \delta\langle \mathbf{a}\rangle\right) \tag{20}$$

$$\begin{aligned}\Sigma_{ii} + \delta\Sigma_{ii} = \Bigg\{ & 2\Delta t\Big\{ \left(\langle s_i^2\rangle + \delta\langle s_i^2\rangle\right) \\ & -2\left[(r_i + \delta r_i)\left(\langle s_i\rangle + \delta\langle s_i\rangle\right) + \sum_k (\mathsf{L}_{ik} + \delta\mathsf{L}_{ik})\left(\langle s_i a_k\rangle + \delta\langle s_i a_k\rangle\right)\right] \\ & + (r_i + \delta r_i)^2 + 2\sum_k (\mathsf{L}_{ik} + \delta\mathsf{L}_{ik})(r_i + \delta r_i)\left(\langle a_k\rangle + \delta\langle a_k\rangle\right) \\ & + \sum_{k,j} (\mathsf{L}_{ik} + \delta\mathsf{L}_{ik})\left(\mathsf{L}_{ij} + \delta\mathsf{L}_{ij}\right)\left(\langle a_k a_j\rangle + \delta\langle a_k a_j\rangle\right)\Big\}\Bigg\}^{\frac{1}{2}},\end{aligned} \tag{21}$$

where $\delta\langle \mathbf{s}'\mathbf{a}'^{\mathrm{T}}\rangle = \delta\langle \mathbf{s}\mathbf{a}^{\mathrm{T}}\rangle - \delta\langle \mathbf{s}\rangle\delta\langle \mathbf{a}\rangle^{\mathrm{T}}$ and $\delta\langle \mathbf{a}'\mathbf{a}'^{\mathrm{T}}\rangle = \delta\langle \mathbf{a}\mathbf{a}^{\mathrm{T}}\rangle - \delta\langle \mathbf{a}\rangle\delta\langle \mathbf{a}\rangle^{\mathrm{T}}$. Consequently, if the changes in the statistical moments (i.e., $\delta\langle \mathbf{a}\rangle, \delta\langle \mathbf{s}\rangle, \delta\langle s_i^2\rangle, \delta\langle \mathbf{a}\mathbf{a}^{\mathrm{T}}\rangle$, and $\delta\langle \mathbf{s}\mathbf{a}^{\mathrm{T}}\rangle$) are known, it will be easy to update the SGS parameterization.

One way to obtain the changes in the statistical moments is to directly analyze data from the perturbed QG3LM. Using the resulting à-posteriori parameterization clearly fixes the issues of the à-priori model. This can be seen in panel c) of Figure 4. Similar results as for the response in the variance are found for the response in the mean streamfunction (not shown), although, in this case, already the à-priori low-order model reproduces the response quite well.

Nevertheless, a recalculation of the QG3LM with the anomalous forcing contradicts the very idea behind efficient low-order climate models. In addition, the problem outlined in Figure 4 (panels a) and b)) is not only limited to low-order climate models. All models containing empirical tuning parameters suffer from data dependence, including state-of-the-art GCMs, paleoclimate models, and many SGS parameterizations. However, for climate projections it is impossible to retune the empirical parameters against a new data-set. Therefore, other approaches are necessary.

c. Fluctuation-Dissipation Theorem

The climate-dependence of the SGS parameterization (18) is introduced by the FDT, as suggested by Achatz et al. (2013). In this study we focus mainly on the conventional qG-FDT. However, to evaluate the approximation of Gaussianity we additionally investigate selected cases with the more sophisticated ST/qG-FDT.

The qG-FDT states for the steady state response of an arbitrary function of the state $\mathbf{h}(\mathbf{a})$ in the reduced EOF space (Gritsun et al. 2008)

$$\delta\langle\mathbf{h}(\mathbf{a})\rangle = \int_0^\infty \langle\mathbf{h}[\mathbf{a}(\tau)]\mathbf{a}'^{\mathrm{T}}(0)\rangle d\tau \langle\mathbf{a}'(0)\mathbf{a}'^{\mathrm{T}}(0)\rangle^{-1}\delta\mathbf{f}, \tag{22}$$

where $\delta\mathbf{f}$ is a constant forcing. A detailed derivation of expression (22) can be found, for example, in Risken (1984). For practical purposes we integrate (22) up to 100 days. The calculation of the lag integral is done via the efficient "Cooper-Haynes" algorithm presented in the appendix of Lutsko et al. (2015), however, here applied to the Simpsons rule.

In addition to the qG-FDT, we also apply the blended ST/qG-FDT algorithm of Abramov and Majda (2008) to selected cases. In contrast to the qG-FDT, the ST-FDT makes no assumption on the shape of the PDF of the system. Thus, the resulting response is generally more accurate than (22). The price for the higher accuracy, however, is the need of the tangent linear model. However,

since the tangent linear model is inherently unstable the ST-FDT response is only useful for short times (up to τ_*). Therefore, Abramov and Majda (2008) propose to combine the ST-FDT and qG-FDT by replacing the qG-FDT response operator (22) with the ST-FDT response operator at short times, while keeping the qG-FDT operator for longer times. The resulting blended ST/qG-FDT reads

$$\delta\langle \mathbf{h}(\mathbf{a})\rangle = \left[\int_0^{\tau_*} \langle \nabla \mathbf{h}[\mathbf{a}(\tau)]\mathbf{T}^{\tau}_{\mathbf{a}(t)}\rangle d\tau + \int_{\tau_*}^{\infty} \langle \mathbf{h}[\mathbf{a}(\tau)]\mathbf{a}'^{\mathrm{T}}(0)\rangle d\tau \langle \mathbf{a}'(0)\mathbf{a}'^{\mathrm{T}}(0)\rangle^{-1}\right]\delta\mathbf{f}, \tag{23}$$

where the tangent linear model $\mathbf{T}^{\tau}_{\mathbf{a}(t)} \in \mathbb{R}^{N\times M}$, initialized with $\mathbf{a}(t)$ and integrated to time τ, is calculated solving

$$\frac{d\mathbf{T}^{\tau}_{\mathbf{a}(t)}}{dt} = \nabla \mathbf{G}[\mathbf{a}(t+\tau)]\mathbf{T}^{\tau}_{\mathbf{a}(t)}, \tag{24}$$

and $\nabla\mathbf{G}[\mathbf{a}(\tau)] \in \mathbb{R}^{N\times N}$ is the Jacobian of the right-hand side of the QG3LM (3) in the full EOF space. Furthermore, in (23) $\nabla\mathbf{h} \in \mathbb{R}^{\bullet\times N}$, where $\bullet$ depends on the dimension of $\mathbf{h}$. Both the instability of the tangent linear model and the ambiguous choice of the parameter τ_*, however, limit the robustness of the blended ST/qG-FDT algorithm. Here $\tau_* = 2$ days, determined by trial and error (i.e., by comparing the response of the qG-FDT and ST-FDT for various τ_*).

4. Experiments

a. Setup

For the test of the FDT-adjusted closure we consider 12 different cases, corresponding to the different positions of the local anomalous forcing (see section 3 a). In the case of the global anomalous forcing we use the first five EOFs as stated in (16). For the computation of the FDT response operators the same $6\cdot 10^6$ day time series is used as for the calculation of the closure. Furthermore, we run the QG3LM with the anomalous forcing for each case. With the resulting

perturbed time series the actual response of the respective moments are calculated to quantify the performance of the FDT algorithms. We restrict our investigation to five low-order models with truncations of 20 (30% explained variance), 50 (46%), 100 (60%), 200 (73%), and 500 (89%) EOFs, respectively.

For the evaluation of the FDT-adjusted closure we consider four different low-order models

- à-priori: no climate-dependence ($\delta\mathbf{r} = \delta\mathbf{L} = \delta\boldsymbol{\Sigma} = 0$) as in (17).

- qG-FDT: closure corrections calculated by qG-FDT.

- à-posteriori: the true closure corrections, obtained from a run of the QG3LM with anomalous forcing.

- rqG-FDT: as the qG-FDT but with $\delta\mathbf{L} = 0$ (see Achatz et al. 2013).

As it turns out, no useful closure updates can be obtained for $\boldsymbol{\Sigma}$ for the local anomalous forcing. In this case, we could keep $\boldsymbol{\Sigma}$ constant and set $\delta\boldsymbol{\Sigma} = 0$. However, this might suppress the effect of the update of the remaining part of the closure with the FDT. Therefore, for the local anomalous forcing we consider only reduced models with a deterministic parameterization (i.e., $\boldsymbol{\Sigma} = \delta\boldsymbol{\Sigma} = 0$). To test the effect of the stochasticity we focus on the cases with the global anomalous forcing.

For the quantification of the results we calculate for $\mathbf{a}, \mathbf{b} \in \mathbb{R}^M$ and $\mathsf{A}, \mathsf{B} \in \mathbb{R}^{M\times M}$ a relative error given by

$$\varepsilon(\mathbf{a},\mathbf{b}) = \frac{\|\mathbf{a}-\mathbf{b}\|^2}{\|\mathbf{a}\|\|\mathbf{b}\|} \qquad\qquad \varepsilon(\mathsf{A},\mathsf{B}) = \frac{\|\mathsf{A}-\mathsf{B}\|^2}{\|\mathsf{A}\|\|\mathsf{B}\|}\,, \tag{25}$$

where the norm $\|\cdot\|$ is either the 2-norm (if the argument is a vector) or the Frobenius norm (if the argument is a matrix).

b. Results

1) LOCAL ANOMALOUS FORCING

Figure 3 (top) shows a boxplot of relative error of the various statistical moments entering (20), summarizing the qG-FDT predictions for all local anomalous forcing cases and each EOF truncation. Clearly the qG-FDT is unable to provide a correct estimate of $\delta\langle s_i^2\rangle$ for any EOF truncation. Apart from that, the remaining first moments (i.e., $\delta\langle\mathbf{a}\rangle$ and $\delta\langle\mathbf{s}\rangle$) are predicted remarkably well with an average median of around 0.1 for all truncations. The median of relative error of the second moments (i.e., $\delta\langle\mathbf{a}\mathbf{a}^{\mathrm{T}}\rangle$ and $\delta\langle\mathbf{s}\mathbf{a}^{\mathrm{T}}\rangle$) are on average with 0.3 to 0.4 considerably higher. In general, the performance of the qG-FDT for $\delta\langle\mathbf{a}\rangle$ and $\delta\langle\mathbf{a}\mathbf{a}^{\mathrm{T}}\rangle$ is independent of the EOF truncation. In fact, for higher EOF truncations the spread of relative error decreases for both moments (although, for $\delta\langle\mathbf{a}\rangle$ this is not visible). However, all moments containing the discretized SGS error $\mathbf{s}$ (i.e., $\delta\langle\mathbf{s}\rangle$, $\delta\langle s_i^2\rangle$, and $\delta\langle\mathbf{s}\mathbf{a}^{\mathrm{T}}\rangle$) exhibit an increase in both median and spread with rising EOF truncation.

Figure 3 (bottom) shows the equivalent evaluation of the closure corrections (20) resulting from the qG-FDT estimation of the statistical moments. The constant term $\delta\mathbf{r}$ has the best result with an average median of 0.15 for all EOF truncations. In contrast, the linear term $\delta\mathbf{L}$ has a considerably higher error in the median which even reaches $\varepsilon \approx 1$ for the 500 EOF low-order model. Generally, the quality of the qG-FDT predicted closure corrections decrease with increasing EOF dimension. This is due to the increase of the error of $\delta\langle\mathbf{s}\mathbf{a}^{\mathrm{T}}\rangle$ (see Figure 3, top) which directly influences $\delta\mathbf{L}$. Consequently, this trend translates to $\delta\mathbf{r}$ as well, since it is dependent on $\delta\mathbf{L}$ (20). The same holds for $\delta\boldsymbol{\Sigma}$, although the high error of $\delta\langle s_i^2\rangle$ renders the closure update of the noise amplitude useless for all EOF truncations.

Figure 3 is not sufficient for a final evaluation of the qG-FDT-predicted closure corrections. The utility of the method is eventually decided by the performance of the reduced models including the

adjusted closures. However, before showing summarizing plots, we first focus on a representative example, given by the 500 EOF low-order model with a local anomalous forcing located at $\lambda_c = 270°$.

Figure 4 shows the response in variance of streamfunction of a) the QG3LM projected onto 500 EOFs, b) - d) three low-order models with different deterministic SGS parameterizations, and e) the direct qG-FDT estimation. The QG3LM shows mainly a response in the northern (winter) hemisphere. In particular, the response of the QG3LM exhibits a band of maxima at roughly 60°N, spanning nearly the whole globe. Furthermore, multiple minima are located over the Pacific ocean, the Atlantic ocean, and Greenland. Overall, the pattern correlation of all reduced models and the direct qG-FDT estimation are, with about 0.90, quite high. Yet, prominent differences are visible in the amplitude. The à-priori low-order model (Figure 4 b)) has a distinct maximum over the Pacific ocean. Furthermore, the minimum over Greenland is too pronounced, whereas the remaining minima are too weak in amplitude. Consequently, the relative error reads $\varepsilon = 0.24$. The à-posteriori low-order model (Figure 4 c)) fairs better with $\varepsilon = 0.17$, due to stronger minima and a slightly reduced maximum over the Pacific ocean. The response of the low-order model with the qG-FDT-adjusted SGS parameterization (Figure 4 d)) is virtually identical to the à-posteriori low-order model which can also be seen from the relative error of $\varepsilon = 0.18$. In contrast, the direct qG-FDT estimation (Figure 4 e)) overestimates the amplitude of the response by a factor of two resulting in a relative error of $\varepsilon = 0.40$.

Similar results are found for the meridional momentum flux $\langle \mathbf{u}'\mathbf{v}' \rangle$ (Figure 5) and the meridional entropy flux (not shown), although for these fluxes the estimation using the qG-FDT directly also provides useful results.

Finally, we summarize the results of the reduced models with adjusted closures in Figure 6. Figure 6 has a similar structure as Figure 3, however, it shows the performance of the response

of the first moment ($\delta\langle\mathbf{a}\rangle$, left panel) and covariance ($\delta\langle\mathbf{a}'\mathbf{a}'^{\mathrm{T}}\rangle$, right panel) of streamfunction of the reduced models with the various adjusted deterministic closures and the direct qG-FDT estimation, compared to the QG3LM. In general, for the response in the first moment the à-priori low-order model with the unmodified closure performs worst for all EOF truncations, whereas the à-posteriori model yields generally the best result of all low-order models. Furthermore, we see a systematic improvement for the low-order model with qG-FDT-adjusted SGS parameterization compared to the à-priori low-order model and the model with the rqG-FDT-adjusted closure. Still, for EOF truncations smaller than 500 EOFs, the direct qG-FDT yields by far the best correspondence with the actual response. However, the higher the EOF truncation, the better fair the reduced models in general. In fact, the qG-FDT-adjusted low-order model with 500 EOFs yields a comparable result to the direct qG-FDT estimation.

The findings of the response of the covariance (right panel) is qualitatively similar, independently of the EOF truncation. However, only for low-order models with more than 200 EOFs the results are useful (i.e., $\varepsilon < 1$). Additionally, for 500 EOFs we find that the reduced model with adjusted closure begins to outperform the direct qG-FDT estimation of the response in covariance. Although, in those cases even the à-priori low-order model yields a better result than the qG-FDT.

Even though the FDT estimations of the change of the statistics are linear, we can significantly improve the description of the perturbed dynamics by the low-order model compared to the à-priori case. In particular, we do not use the FDT estimations directly but apply them on the empirical parts of the SGS closure. This allows us to account for the nonlinearity explicitly when running the low-order model.

2) Global Anomalous Forcing

Figure 7 shows the qG-FDT estimation of the changes in the statistical moments and the skill of the resulting closure updates, in the case of the global anomalous forcing. Qualitatively the findings in the case of the local anomalous forcing are reproduced. However, the overall performance of the qG-FDT estimations is increased. In general, the spread of estimates is significantly reduced for all statistical moments. In the second moments, the average median of the relative error of $\delta\langle \mathbf{a}\mathbf{a}^{\mathrm{T}}\rangle$ is below 0.2, whereas the relative error of $\delta\langle \mathbf{s}\mathbf{a}^{\mathrm{T}}\rangle$ is basically unchanged in comparison to the local anomalous forcing case. In addition, the estimation of $\delta\langle \mathbf{s}_i^2\rangle$ is with an average median of 0.15 comparable to the second moments. Furthermore, the remaining first moments exhibit on average an relative error below 0.1. Interestingly, the quality of the resulting closure corrections are (apart from $\delta\boldsymbol{\Sigma}$) nearly identical compared to the local anomalous forcing case. This indicates that the considered SGS parameterization is highly sensitive to $\delta\langle \mathbf{s}\mathbf{a}^{\mathrm{T}}\rangle$. Nevertheless, the good estimation of $\delta\langle s_i^2\rangle$ allows a useful update of $\delta\boldsymbol{\Sigma}$ which was not possible for the local anomalous forcing case. Consequently, the stochastic parameterization (12) can be studied under the perturbation with a global forcing.

For direct comparison we first show in Figure 8 the results of the low-order models with adjusted deterministic closures in the case of the global anomalous forcing. In general, the results of the local anomalous forcing are reproduced, however, the performance of all low-order models and the direct qG-FDT is systematically better. In particular, the spread is significantly reduced. Figure 9 shows the reduced models with adjusted stochastic parameterization in the case of the global anomalous forcing. Overall, the performance is qualitatively similar to the deterministic case. A slight improvement in terms of spread is visible for all reduced models and all EOF truncations. Nevertheless, the net effect by the introduction of the noise term is quite small. The results of the

low-order models with à-posteriori and rqG-FDT closure are not shown, since their behavior is qualitatively the same as in Figure 6.

3) ST/QG-FDT

In order to estimate the error of the Gaussian approximation in the qG-FDT, we predict the changes in the statistical moments of the 20 EOF and 200 EOF case using the ST/qG-FDT by Abramov and Majda (2007). The result is given in Figure 10 (top). Generally, the performance of the ST/qG-FDT is qualitatively similar to that of the qG-FDT (Figure 7). For both FDT algorithms the first moments are significantly better estimated compared to the second moments. Furthermore, the quality of the response of the first moments are equivalent. Nevertheless, we see an improvement in the ST/qG-FDT estimations in terms of spread for the second moments, in particular for $\delta\langle \mathbf{a}\mathbf{a}^{\mathrm{T}}\rangle$. However, running the reduced models with ST/qG-FDT-adjusted closure yield similar results compared to the models with qG-FDT-updated closure in Figure 6 (not shown). This might indicate an inherent issue of the ST/qG-FDT. However, the median of $\delta\langle \mathbf{s}\mathbf{a}^{\mathrm{T}}\rangle$ estimated by ST/qG-FDT (Figure 10) stays roughly the same compared to the qG-FDT estimation (Figure 7). It turns out that for our simulations the deterministic part of the SGS parameterization is highly sensitive towards this specific moment. Thus, this might explain the lack of improvement in the low-order models with ST/qG-FDT adjusted closures.

5. Summary and Discussion

In this study we addressed the problem of data dependence of empirical tuning parameters and the resulting inability of models containing such parameters to respond correctly to external perturbations. We considered the QG3LM of Marshall and Molteni (1993) as a toy atmosphere. In addition, we constructed a low-order climate model based on EOFs and equipped with a purely

data-driven deterministic or stochastic SGS parameterization. While the low-order model with both closures reproduced the statistics (in terms of both amplitude and correlation) of the QG3LM correctly, it generated an incorrect response if the system was perturbed by an anomalous forcing. This is a known issue of (semi) empirical low-order models (Achatz and Branstator 1999; Achatz and Opsteegh 2003b) which is due to the fact that the parameterization is tuned to the unperturbed statistics. Consequently, such low-order models are unsuitable for climate projections. However, we want to emphasize that this is a fundamental problem in climate modeling which is not only limited to low-order climate models. Even physical parameterization schemes such as radiation or convection contain to some degree empirical components (e.g., aerosol distribution, vegetation, or entrainment of convective cells). One could hope that the physical basis should make those closures more robust to climate change than their purely data-driven counterparts. However, there are indications that physical SGS closures suffer from data-dependence as well. Rockel and Geyer (2008), e.g., show that a regional climate model, tuned to simulate regional climate over middle Europe well, yields incorrect results if applied over Southern America without retuning. Schirber et al. (2015) show that several gravity-wave parameterization schemes can be tuned successfully to help a GCM simulate a realistic quasi-biennial oscillation. Yet, in a climate-change experiment the response of the quasi-biennial oscillation depends strongly on the chosen scheme.

The present study follows the ansatz of Achatz et al. (2013) by tackling the problem by means of the FDT which is related to a corresponding response-theory approach of Lucarini and Wouters (2017). If the empirical parameters are based on a minimization of an objectively formulated error (i.e., it depends on the statistics of the system), à-priori estimations of the changes in the corresponding statistics could help reducing the data dependence. The study of Achatz et al. (2013), however, potentially suffered from the barotropic toy atmosphere used there, with relatively weak high-frequency variability, and hence limited applicability of stochastic approaches. The hypothe-

sis is that a setting with baroclinic instability and corresponding synoptic-scale activity should be better suited to an FDT ansatz. For a corresponding test, this study focusses on the most simple form of FDT: the qG-FDT. In addition, we considered two types of perturbations, namely local anomalous forcings (localized heat sources) and global anomalous forcings, represented up to a factor by individual EOFs.

For both anomalous forcing types the estimations of the response of the first moments were remarkably accurate. The estimations of the second moments exhibited a systematically higher error. This is consistent with Gritsun et al. (2008) who also show a generally higher error for second moments. In contrast to the literature (e.g., Gritsun and Branstator 2007; Gritsun et al. 2008), the problem in this study is even more challenging, since the reference used for the evaluation of the FDT is the true à-posteriori response of the perturbed QG3LM, including all nonlinearities. The large amount of data used in this study allows us to exclude a convergence problem. Moreover, we analyzed the response in the second moments with respect to its linearity. There we found no link between the (non)linearity of the response and the quality of the FDT estimations. At least for a transient response Majda et al. (2005, page 68) proved that the qG-FDT estimation of the first moment ($\delta\langle\mathbf{a}\rangle$) is more accurate than that of the second moment ($\delta\langle\mathbf{a}\mathbf{a}^{\mathrm{T}}\rangle$).

Furthermore, we found that the statistical moments containing the discretized SGS error $\mathbf{s}$ (i.e., $\delta\langle\mathbf{s}\rangle$, $\delta\langle s_i^2\rangle$ and $\delta\langle\mathbf{s}\mathbf{a}^{\mathrm{T}}\rangle$) experienced a trend, that is the error of the qG-FDT estimation increased with higher EOF truncations. This could be due to a larger contribution of the stable manifold (Gritsun and Lucarini 2017) at higher EOF truncations. On the other hand, it can also be understood by observing that the size of the SGS error gets smaller as the EOF truncation is relaxed. This results in a small signal-to-noise ratio for moments containing the discretized SGS error. Repeating the experiments with only half of the time series shows an increase in this trend.

The qG-FDT estimations of the statistics were used to update the closures. Here we distinguished four cases: the à-priori low-order model (no change in the parameterization), the à-posteriori low-order model (retuned parameterization), the qG-FDT low-order model (closure updated with the qG-FDT estimations), and the rqG-FDT low-order model (closure updated with qG-FDT estimations of first moments only, see Achatz et al. 2013, for details).

In general, for the response in mean and covariance of streamfunction the update of the deterministic closure by the qG-FDT yielded a better agreement with the QG3LM compared to the à-priori low-order model. However, for the response in covariance of streamfunction the result of the low-order models with EOF truncations below 200 EOFs were not useful (i.e., $\varepsilon > 1$). In those cases, the simple linear closure does not suffice to describe the SGS processes. For the 500 EOF low-order models we found that the qG-FDT-updated closure starts to outperform the direct qG-FDT estimation of the response in both mean and covariance. Nevertheless, even a comparable quality to the direct qG-FDT result is already a success, since the low-order model provides a time series which can be further analyzed, whereas the FDT only provides a response.

Overall, the qG-FDT estimations of the changes of the statistics worked better than in Achatz et al. (2013). This was to be expected since the QG3LM with its relatively fast baroclinic instability better fulfills the constrains of the qG-FDT. Consequently, our simulations show that discarding the FDT estimations of the second moments (i.e., using the rqG-FDT low-order model) yields inferior results compared to the qG-FDT low-order model, for all considered cases.

If the forcing projects only on one of the first five EOFs (i.e., global anomalous forcing) instead of a combination of the first 20 EOFs (i.e., local anomalous forcing), the FDT performs significantly better. Lutsko et al. (2015) argue that taking more EOFs into account results in higher uncertainty of the response operator, since higher EOFs are inherently more noisy. However, we consider quite long time series, so one would expect that the first 20 EOFs are sufficiently well

sampled. Furthermore, we are confident that enough data was used to construct the FDT response operators. In fact, using only $3 \cdot 10^6$ (or even less) days to generate the qG-FDT operators and comparing them to the $6 \cdot 10^6$-days-à-posteriori response yielded nearly the same results as presented in this paper. Moreover, if we split up the $6 \cdot 10^6$ days time series into six parts and thus generate an ensemble of qG-FDT operators we can show that the mean operators are quite similar to the operators used in this study, while the ensemble spread is quite small. Thus, it seems that the worse performance of the FDT in case of the local anomalous forcing is due to its "hot-spot" nature which is in agreement with findings of Fuchs et al. (2015).

Due to the overall worse performance of the FDT in the case of the local anomalous forcing, we were unable to obtain useful updates for the noise amplitude in the stochastic closure. Therefore, only for the global anomalous forcing a stochastic parameterization could be considered. Qualitatively, the reduced models with stochastic closure produce similar results compared to the deterministic case described above. However, quantitatively no significant improvement is evident.

The superior ST/qG-FDT blended algorithm, which was exemplary applied on the 20 EOF and 200 EOF case, yielded an improvement, especially for $\delta\langle \mathbf{a}\mathbf{a}^{\mathrm{T}} \rangle$. However, the SGS parameterization used in this study is highly sensitive to $\delta\langle \mathbf{s}\mathbf{a}^{\mathrm{T}} \rangle$. Therefore, the resulting closure correction was only marginally better compared to the qG-FDT result. Apparently, the derivations from Gaussianity are negligible in our study. Yet, because of the calculation of the tangent linear model, the computational cost of the ST/qG-FDT algorithm is significantly higher than that for the simple qG-FDT. Thus, we see our comparatively simple approach (i.e., using the qG-FDT) supplementary to the more sophisticated ST/qG-FDT, at least if one is interested in the equilibrium response of the model.

All in all, the FDT provides useful estimations of the changes of the statistics in a perturbed climate which allows for a successful update of the empirical parameters in an atmospheric model. However, since the FDT is a linear response theory, this method can only be applied to small forcings. Furthermore, it is restricted to forcings which project into unstable directions of the attractor (Gritsun and Lucarini 2017). Still, FDT offers a first step towards climate-dependent tuning parameters in data-driven SGS parameterizations. One possible route to reduce further the climate dependence of parameterizations would be to apply the present approach within the stochastic mode reduction procedure (Majda et al. 2001, 2003). If instead of the seamless approach (Franzke et al. 2005; Franzke and Majda 2006) the coefficients of the Ornstein-Uhlenbeck process are estimated directly (Dolaptchiev et al. 2013a,b; Zacharuk et al. 2018), the latter can be updated using the FDT. For future studies more complex models (e.g., GCMs), containing unbalanced motion (e.g., gravity waves), or moisture should be considered. In addition, it might be interesting to study whether the FDT is able to estimate the changes due to a realistic external forcing. A related open issue is how the FDT (or other response-theory approaches) might be used to predict the sensitivity of tuning parameters of standard physical SGS parameterizations (e.g., convection, turbulence, gravity waves, clouds, etc.) to external conditions. Since climate change might operate to a considerable degree via such SGS processes, this could eventually be quite relevant for sensitivity studies of the climate system as a whole.

Acknowledgments. The authors thank Valerio Lucarini and the two anonymous reviewer for their helpful comments which led to an improvement of this manuscript. MP thanks Erwan Brisson for the useful discussion. MP and UA thank German Research Foundation (DFG) for partial support through grant AC 71/7-1. SD and MZ thank the DFG for partial support through grant DO 1819/1-1. AG received support from the Russian Foundation for Basic Research (grant 16-55-12015).

APPENDIX A

Total Energy Metric

The state vector is defined as

$$\mathbf{x} = \begin{pmatrix} {}_{\nu}x^{m}_{n,1} \\ {}_{\nu}x^{m}_{n,2} \\ {}_{\nu}x^{m}_{n,3} \end{pmatrix} \equiv \begin{pmatrix} {}_{\nu}\psi^{m}_{n,i=1} \\ {}_{\nu}\psi^{m}_{n,i=2} \\ {}_{\nu}\psi^{m}_{n,i=3} \end{pmatrix}, \tag{A1}$$

where ${}_{\nu}\psi^{m}_{n,i}$ is the spectral expansion coefficient of the streamfunction Ψ, m is the zonal wavenumber, n is the total wavenumber, and $\nu = 1$ ($\nu = 2$) denotes the real (imaginary) part of the coefficient, respectively. Furthermore, the boundary conditions of the QG3LM leads to trivial components for $m = n = 0$ and zero imaginary components for $m = 0$ and $n \geq 1$ (Ehrendorfer 2000). Consequently, the state vector can be reduced to

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} \equiv \begin{pmatrix} {}_{1}\psi^{0}_{1,1} \\ {}_{1}\psi^{0}_{1,2} \\ \vdots \\ {}_{2}\psi^{21}_{21,3} \end{pmatrix}, \tag{A2}$$

where the maximum wavenumbers of 21 are given by the spectral truncation of T21. The total energy metric by Ehrendorfer (2000) is defined as

$$\mathbf{M} = \begin{pmatrix} \widehat{\mathbf{M}}_{0,1} & & & 0 \\ & \widehat{\mathbf{M}}_{0,2} & & \\ & & \ddots & \\ 0 & & & \widehat{\mathbf{M}}_{21,21} \end{pmatrix}, \tag{A3}$$

where the matrices on the diagonal are given by

$$\widehat{\mathbf{M}}_{m,n} = \frac{\delta_m}{3}\begin{pmatrix} n(n+1)+a^2R_1^{-2} & -a^2R_1^{-2} & 0 \\ -a^2R_1^{-2} & n(n+1)+a^2R_1^{-2}+a^2R_2^{-2} & -a^2R_2^{-2} \\ 0 & -a^2R_2^{-2} & n(n+1)+a^2R_2^{-2} \end{pmatrix}. \tag{A4}$$

Here $\delta_m = 0.5$ for $m = 0$ and equal to one otherwise, $a = 6370$ km is the earth's radius, and $R_1^{-2} = 700$ km and $R_2^{-2} = 450$ km are the Rossby radii of deformation of the 200-500 hPa and 500-800 hPa layer, respectively. In combination with the state vector (A2) this energy metric yields

$$E = \mathbf{x}^{\mathrm{T}}\mathbf{M}\mathbf{x}, \tag{A5}$$

i.e., the total energy of the flow.

References

Abramov, R. V., and A. J. Majda, 2007: Blended response algorithms for linear fluctuation-dissipation for complex nonlinear dynamical systems. *Nonlinearity*, **20 (12)**, 2793–2821, doi: 10.1088/0951-7715/20/12/004.

Abramov, R. V., and A. J. Majda, 2008: New Approximations and Tests of Linear Fluctuation-Response for Chaotic Nonlinear Forced-Dissipative Dynamical Systems. *Journal of Nonlinear Science*, **18 (3)**, 303–341, doi:10.1007/s00332-007-9011-9.

Abramov, R. V., and A. J. Majda, 2009: A New Algorithm for Low-Frequency Climate Response. *J. Atmos. Sci.*, **66 (2)**, 286–309, doi:10.1175/2008JAS2813.1.

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

Achatz, U., U. Löbl, S. Dolaptchiev, and A. Gritsun, 2013: Fluctuation-Dissipation Supplemented by Nonlinearity: A Climate-Dependent Subgrid-Scale Parameterization in Low-Order Climate Models. *J. Atmos. Sci.*, **70 (6)**, 1833–1846, doi:10.1175/JAS-D-12-0229.1.

Achatz, U., and J. D. Opsteegh, 2003a: Primitive-Equation-Based Low-Order Models with Seasonal Cycle. Part I: Model Construction. *J. Atmos. Sci.*, **60 (3)**, 465–477, doi:10.1175/1520-0469(2003)060⟨0465:PEBLOM⟩2.0.CO;2.

Achatz, U., and J. D. Opsteegh, 2003b: Primitive-Equation-Based Low-Order Models with Seasonal Cycle. Part II: Application to Complexity and Nonlinearity of Large-Scale Atmosphere Dynamics. *J. Atmos. Sci.*, **60 (3)**, 478–490, doi:10.1175/1520-0469(2003)060⟨0478:PEBLOM⟩2.0.CO;2.

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. *J. Atmos. Sci.*, **54 (20)**, 2452–2474, doi:10.1175/1520-0469(1997)054⟨2452:OTCPIT⟩2.0.CO;2.

Achatz, U., G. Schmitz, and K.-M. Greisiger, 1995: Principal Interaction Patterns in Baroclinic Wave Life Cycles. *J. Atmos. Sci.*, **52 (18)**, 3201–3213, doi:10.1175/1520-0469(1995)052⟨3201:PIPIBW⟩2.0.CO;2.

Bell, T. L., 1980: Climate Sensitivity from Fluctuation Dissipation: Some Simple Model Tests. *J. Atmos. Sci.*, **37 (8)**, 1700–1707, doi:10.1175/1520-0469(1980)037⟨1700:CSFFDS⟩2.0.CO;2.

Branstator, G., and S. E. Haupt, 1998: An Empirical Model of Barotropic Atmospheric Dynamics and Its Response to Tropical Forcing. *J. Climate*, **11 (10)**, 2645–2667, doi:10.1175/1520-0442(1998)011⟨2645:AEMOBA⟩2.0.CO;2.

Cooper, F. C., and P. H. Haynes, 2011: Climate Sensitivity via a Nonparametric Fluctuation-Dissipation Theorem. *J. Atmos. Sci.*, **68 (5)**, 937–953, doi:10.1175/2010JAS3633.1.

Demaeyer, J., and S. Vannitsem, 2017: Stochastic parametrization of subgrid-scale processes in coupled ocean-atmosphere systems: benefits and limitations of response theory. *Quart. J. Roy. Meteor. Soc.*, **143 (703)**, 881–896, doi:10.1002/qj.2973.

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.

Ehrendorfer, M., 2000: The Total Energy Norm in a Quasigeostrophic Model. *J. Atmos. Sci.*, **57 (20)**, 3443–3451, doi:10.1175/1520-0469(2000)057⟨3443:NACTEN⟩2.0.CO;2.

Franzke, C., and A. J. Majda, 2006: Low-Order Stochastic Mode Reduction for a Prototype Atmospheric GCM. *J. Atmos. Sci.*, **63 (2)**, 457–479, doi: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. *J. Atmos. Sci.*, **62 (6)**, 1722–1745, doi:10.1175/JAS3438.1.

Fuchs, D., S. Sherwood, and D. Hernandez, 2015: An Exploration of Multivariate Fluctuation Dissipation Operators and Their Response to Sea Surface Temperature Perturbations. *J. Atmos. Sci.*, **72 (1)**, 472–486, doi:10.1175/JAS-D-14-0077.1.

Gershgorin, B., and A. J. Majda, 2010: A test model for fluctuation-dissipation theorems with time-periodic statistics. *Physica D*, **239 (17)**, 1741–1757, doi:10.1016/j.physd.2010.05.009.

Gritsoun, A. S., and V. P. Dymnikov, 1999: Barotropic atmosphere response to small external action: theory and numerical experiments. *Izvestiya, Atmospheric and Oceanic Physics*, **35 (5)**, 511–525.

Gritsun, A., 2010: Construction of the response operators onto small external forcings for the general circulation atmospheric models with time-periodic right hand sides. *Izvestiya, Atmospheric and Oceanic Physics*, **46 (6)**, 748–756, doi:10.1134/S000143381006006X.

Gritsun, A., and G. Branstator, 2007: Climate Response Using a Three-Dimensional Operator Based on the Fluctuation-Dissipation Theorem. *J. Atmos. Sci.*, **64 (7)**, 2558–2575, doi:10.1175/JAS3943.1.

Gritsun, A., and G. Branstator, 2016: Numerical aspects of applying the fluctuation dissipation theorem to study climate system sensitivity to external forcings. *Russian Journal of Numerical Analysis and Mathematical Modelling*, **31 (6)**, 339–354, doi:10.1515/rnam-2016-0032.

Gritsun, A., G. Branstator, and A. Majda, 2008: Climate Response of Linear and Quadratic Functionals Using the Fluctuation-Dissipation Theorem. *J. Atmos. Sci.*, **65 (9)**, 2824–2841, doi:10.1175/2007JAS2496.1.

Gritsun, A., and V. Lucarini, 2017: Fluctuations, Response, and Resonances in a Simple Atmospheric Model. *Physica D*, **349**, 62–76, doi:10.1016/j.physd.2017.02.015.

Hassanzadeh, P., and Z. Kuang, 2016: The Linear Response Function of an Idealized Atmosphere. Part II: Implications for the Practical Use of the Fluctuation-Dissipation Theorem and the Role of Operator's Nonnormality. *J. Atmos. Sci.*, **73 (9)**, 3441–3452, doi:10.1175/JAS-D-16-0099.1.

Honerkamp, J., 1994: *Stochastic Dynamical Systems: Concepts, Numerical Methods, Data Analysis*. Wiley-VCH, 535 pp.

Kravtsov, S., D. Kondrashov, and M. Ghil, 2009: Empirical model reduction and the modeling hierarchy in climate dynamics and the geosciences. *Stochastic Physics and Climate Modeling*, T. Palmer, and P. Williams, Eds., Cambridge University Press, 35–72.

Kwasniok, F., 1996: The reduction of complex dynamical systems using principal interaction patterns. *Physica D*, **92 (1-2)**, 28–60, doi:10.1016/0167-2789(95)00280-4.

Kwasniok, F., 2004: Empirical Low-Order Models of Barotropic Flow. *J. Atmos. Sci.*, **61 (2)**, 235–245, doi:10.1175/1520-0469(2004)061⟨0235:ELMOBF⟩2.0.CO;2.

Kwasniok, F., 2007: Reduced Atmospheric Models Using Dynamically Motivated Basis Functions. *J. Atmos. Sci.*, **64 (10)**, 3452–3474, doi:10.1175/JAS4022.1.

Leith, C. E., 1975: Climate Response and Fluctuation Dissipation. *J. Atmos. Sci.*, **32 (10)**, 2022–2026, doi:10.1175/1520-0469(1975)032⟨2022:CRAFD⟩2.0.CO;2.

Liu, Q., and T. Opsteegh, 1995: Interannual and decadal variations of blocking activity in a quasi-geostrophic model. *Tellus A*, **47 (5)**, 941–954, doi:10.1034/j.1600-0870.1995.00202.x.

Lorenz, D. J., and E. T. DeWeaver, 2007: Tropopause height and zonal wind response to global warming in the IPCC scenario integrations. *Journal of Geophysical Research Atmospheres*, **112 (10)**, 1–11, doi:10.1029/2006JD008087.

Lucarini, V., and S. Sarno, 2011: A statistical mechanical approach for the computation of the climatic response to general forcings. *Nonlinear Processes in Geophysics*, **18 (1)**, 7–28, doi:10.5194/npg-18-7-2011.

Lucarini, V., and J. Wouters, 2017: Response formulae for n -point correlations in statistical mechanical systems and application to a problem of coarse graining. *Journal of Physics A: Mathematical and Theoretical*, **50 (35)**, 355 003, doi:10.1088/1751-8121/aa812c.

Lutsko, N. J., I. M. Held, and P. Zurita-Gotor, 2015: Applying the Fluctuation-Dissipation Theorem to a Two-Layer Model of Quasigeostrophic Turbulence. *J. Atmos. Sci.*, **72 (8)**, 3161–3177, doi:10.1175/JAS-D-14-0356.1.

Majda, A. J., R. V. Abramov, and M. J. Grote, 2005: *Information Theory and Stochastics for Multiscale Nonlinear Systems*, CRM Monograph Series, Vol. 25. American Mathematical Society, 133 pp.

Majda, A. J., I. Timofeyev, and E. Vanden-Eijnden, 2001: A Mathematical Framework for Stochastic Climate Models. *Communications on Pure and Applied Mathematics*, **54 (8)**, 891–974, doi:10.1002/cpa.1014.

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

Majda, A. J., and X. Wang, 2010: Linear response theory for statistical ensembles in complex systems with time-periodic forcing. *Communications in Mathematical Sciences*, **8 (1)**, 145–172, doi:10.4310/CMS.2010.v8.n1.a8.

Marconi, U. M. B., A. Puglisi, L. Rondoni, and A. Vulpiani, 2008: Fluctuation-dissipation: Response theory in statistical physics. *Physics Reports*, **461 (4-6)**, 111–195, doi:10.1016/j.physrep.2008.02.002.

Marshall, J., and F. Molteni, 1993: Toward a Dynamical Understanding of Planetary-Scale Flow Regimes. *J. Atmos. Sci.*, **50 (12)**, 1792–1818, doi:10.1175/1520-0469(1993)050⟨1792:TADUOP⟩2.0.CO;2.

North, G. R., R. E. Bell, and J. W. Hardin, 1993: Fluctuation dissipation in a general circulation model. *Climate Dyn.*, **8 (6)**, 259–264, doi:10.1007/BF00209665.

Ragone, F., V. Lucarini, and F. Lunkeit, 2016: A new framework for climate sensitivity and prediction: a modelling perspective. *Climate Dyn.*, **46 (5-6)**, 1459–1471, doi:10.1007/s00382-015-2657-3.

Ring, M. J., and R. A. Plumb, 2008: The Response of a Simplified GCM to Axisymmetric Forcings: Applicability of the Fluctuation-Dissipation Theorem. *J. Atmos. Sci.*, **65 (12)**, 3880–3898, doi:10.1175/2008JAS2773.1.

Risken, H., 1984: *The Fokker-Planck Equation: Methods of Solution and Applications*. Springer-Verlag, 454 pp.

Rockel, B., and B. Geyer, 2008: The performance of the regional climate model CLM in different climate regions, based on the example of precipitation. *Meteor. Z.*, **17 (4)**, 487–498, doi:10.1127/0941-2948/2008/0297.

Schirber, S., E. Manzini, T. Krismer, and M. Giorgetta, 2015: The quasi-biennial oscillation in a warmer climate: sensitivity to different gravity wave parameterizations. *Climate Dyn.*, **45 (3-4)**, 825–836, doi:10.1007/s00382-014-2314-2.

Selten, F. M., 1997: Baroclinic Empirical Orthogonal Functions as Basis Functions in an Atmospheric Model. *J. Atmos. Sci.*, **54 (16)**, 2099–2114, doi:10.1175/1520-0469(1997)054⟨2099:BEOFAB⟩2.0.CO;2.

Verkley, W. T. M., P. C. Kalverla, and C. A. Severijns, 2016: A maximum entropy approach to the parametrization of subgrid processes in two-dimensional flow. *Quart. J. Roy. Meteor. Soc.*, **142 (699)**, 2273–2283, doi:10.1002/qj.2817.

Verkley, W. T. M., and C. A. Severijns, 2014: The maximum entropy principle applied to a dynamical system proposed by Lorenz. *The European Physical Journal B*, **87 (1)**, 7, doi:10.1140/epjb/e2013-40681-2.

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

Wouters, J., and V. Lucarini, 2012: Disentangling multi-level systems: averaging, correlations and memory. *Journal of Statistical Mechanics: Theory and Experiment*, **2012 (03)**, P03 003, doi:10.1088/1742-5468/2012/03/P03003.

Zacharuk, M., S. I. Dolaptchiev, I. Timofeyev, and U. Achatz, 2018: Effective subgrid-scale parameterization for one-dimensional shallow water dynamics using stochastic mode reduction. *Quart. J. Roy. Meteor. Soc., under revision*.

Zeeman, E. C., 1988: Stability of dynamical systems. *Nonlinearity*, **1 (1)**, 115–155, doi:10.1088/0951-7715/1/1/005.

LIST OF TABLES

TABLE 1. Adjusted diffusion time scale [days] used in the low-order models (original value 2 days).

number of EOFs	deterministic τ_H	stochastic τ_H
20	1.0	0.4
50	1.3	0.6
100	1.2	0.6
200	1.9	1.1
500	2.6	2.2

TABLE 2. Constant factors used for the calculation of the global anomalous forcing in units of $4\Omega^2$ where $\Omega = 2\pi$ day^{-1}.

ε_1	ε_2	ε_3	ε_4	ε_5
$\frac{1}{\sqrt{45}} \cdot 10^{-5}$	$\frac{1}{\sqrt{10}} \cdot 10^{-5}$	$\frac{1}{\sqrt{3}} \cdot 10^{-5}$	$\frac{1}{2} \cdot 10^{-5}$	$\frac{1}{\sqrt{2}} \cdot 10^{-5}$

LIST OF FIGURES

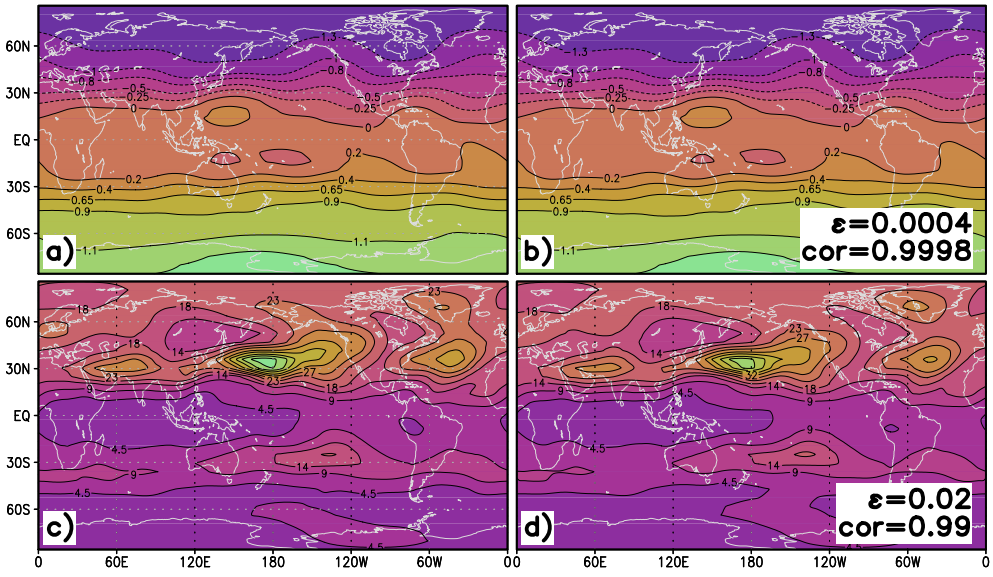


FIG. 1. Mean streamfunction (upper row, 10^8 [m^2 s^{-1}]) and variance of streamfunction (lower row, 10^{13} [m^4 s^{-2}]) at 200 hPa. Panels a) and c) show the result of the QG3LM projected onto the first 500 EOFs; panels b) and d) show the result of the 500 EOF low-order climate model with deterministic parameterization. ε denotes the relative error (25) and *cor* is the pattern correlation calculated between the low-order model and the QG3LM, respectively. Note that the color shading is adjusted and the same units are used for each row.

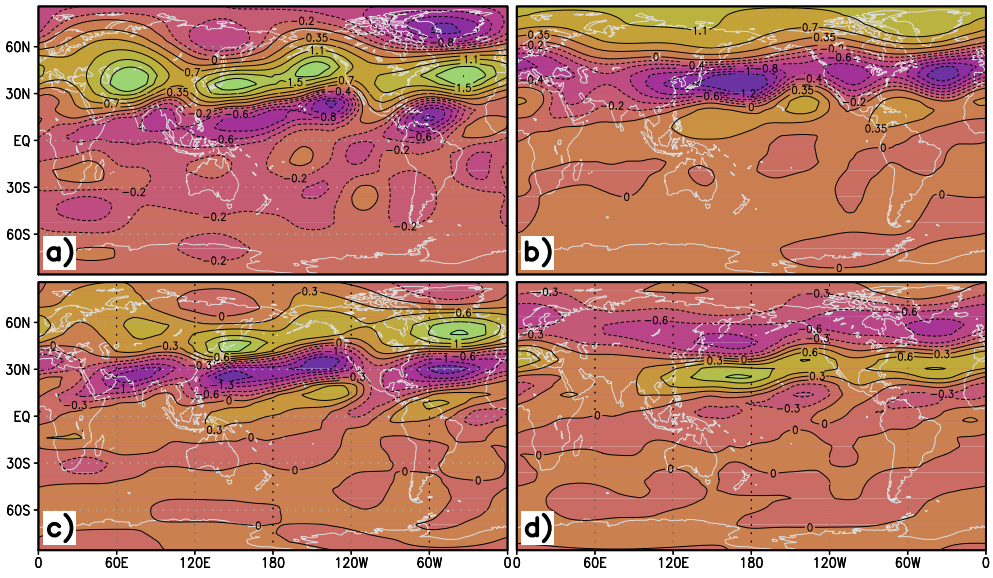
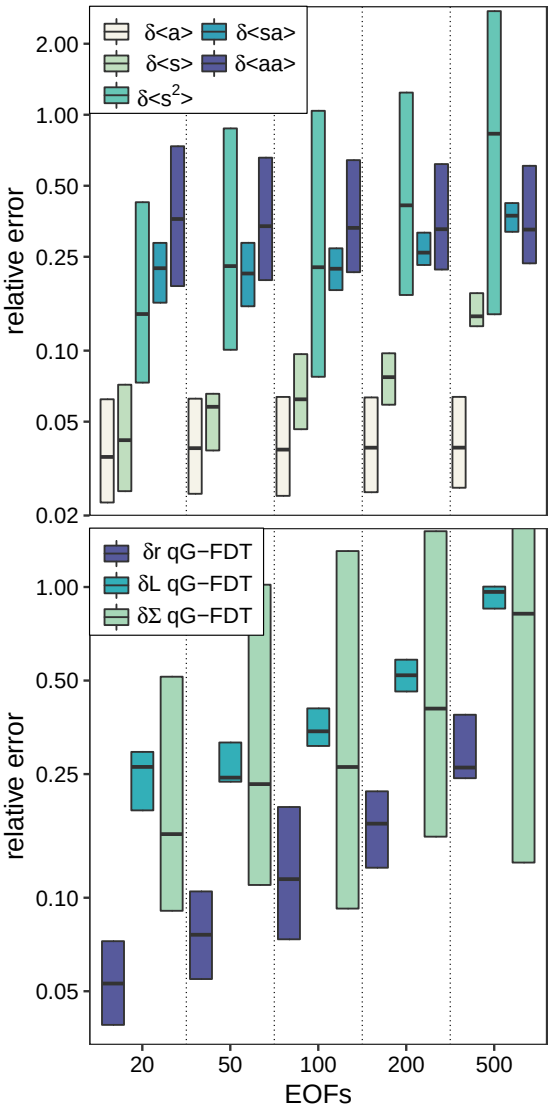


FIG. 2. Response of the QG3LM at 200 hPa due to the local anomalous forcing at $\lambda_c = 60^\circ$ in a) mean streamfunction 10^6 [m^2 s^{-1}] and c) zonal wind [m s^{-1}]. Panels b) and d) show the response due to the global anomalous forcing using the first EOF. The color shading is adjusted and the same units are used for each row.



816 FIG. 3. Boxplot of the relative error between the qG-FDT and the true response of the moments (top) and
817 à-posteriori closure corrections (bottom) to the local anomalous forcing, against the number of EOFs. The
818 statistics for the boxplot come from the different forcing positions $\lambda_c \in \{0°, 30°, \ldots, 330°\}$. The median is given
819 by the horizontal black line and the colored area denotes the interquartile range (25% to 75% percentile).

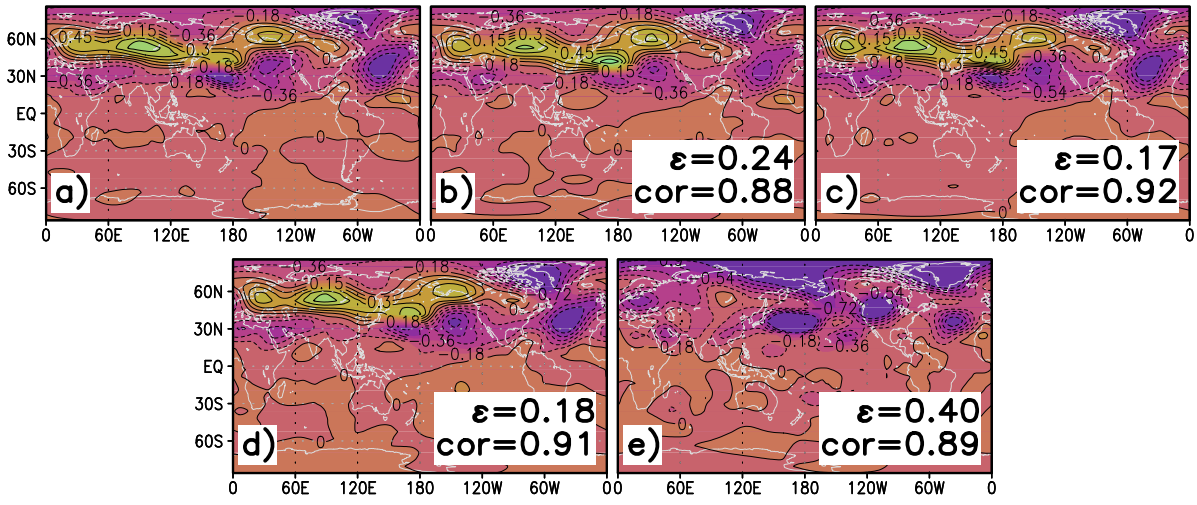


FIG. 4. Response in variance of streamfunction 10^{13} [m^4 s^{-2}] at 200 hPa due to a local anomalous forcing located at $\lambda_c = 270°$. a) shows the response of the QG3LM projected onto 500 EOFs. The response of the 500 EOF low-order model with b) the à-priori SGS parameterization; c) the à-posteriori closure; d) with qG-FDT predicted closure corrections. For comparison the direct qG-FDT estimation of the response in variance of streamfunction is given in e). ε denotes the relative error (25) and *cor* is the pattern correlation calculated between the low-order model and the QG3LM, respectively.

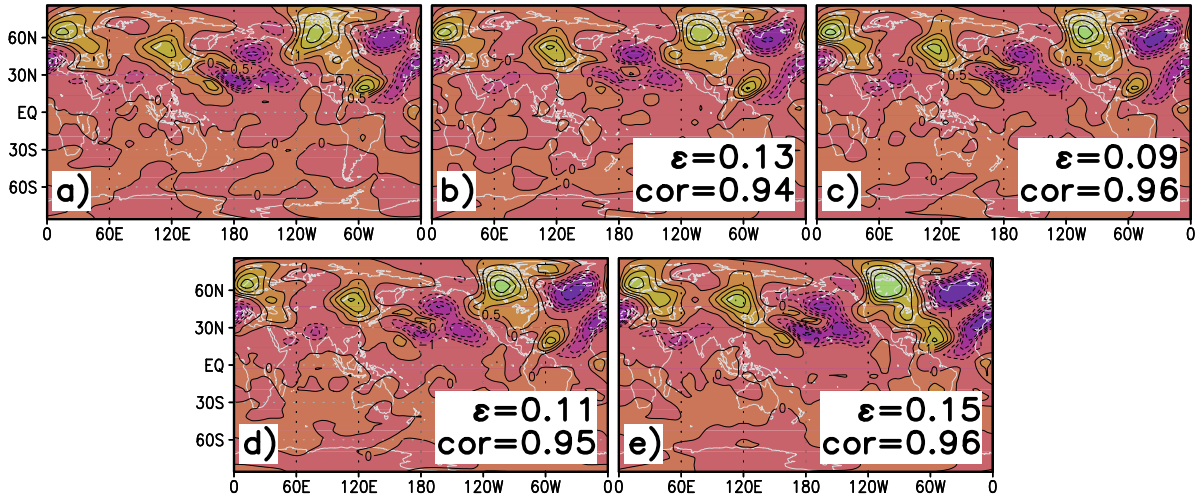
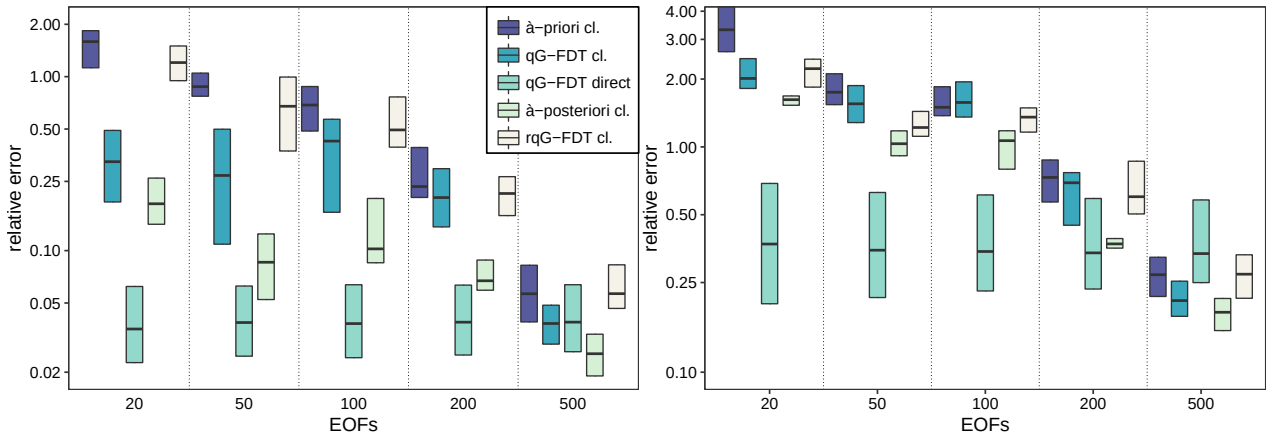


FIG. 5. As Figure 4 but for response in meridional momentum flux [$m^2\ s^{-2}$].



826 FIG. 6. Boxplot of the relative error of the response of the first moment (left) and covariance of streamfunction
827 (right) between the low-order climate model with adjusted closures and the QG3LM with local anomalous forc-
828 ing against the number of EOFs. For comparison also the direct qG-FDT estimation of the respective moment
829 is plotted. The statistics for the boxplot come from the different forcing positions $\lambda_c \in \{0°, 30°, \ldots, 330°\}$.

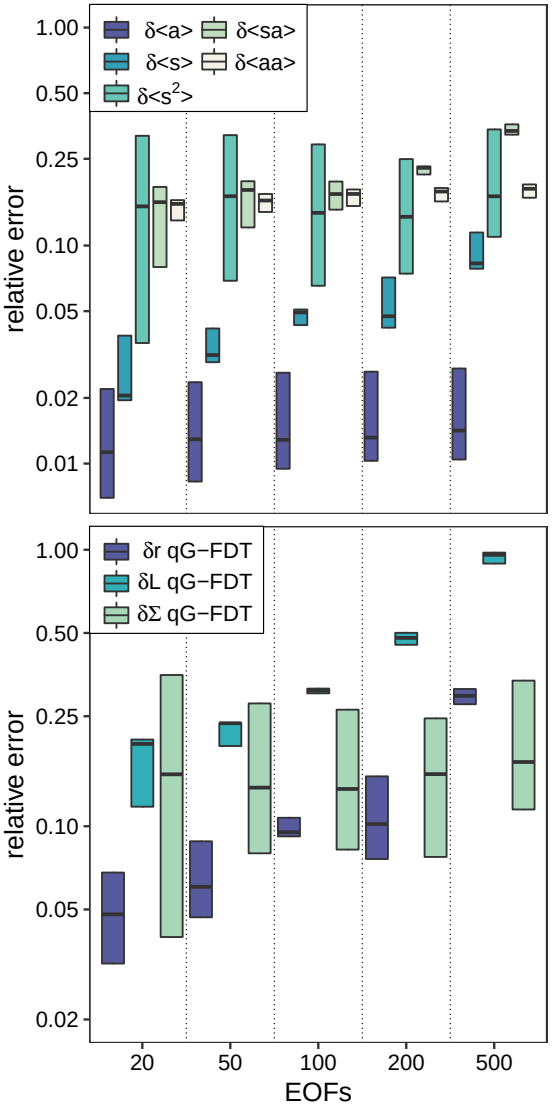
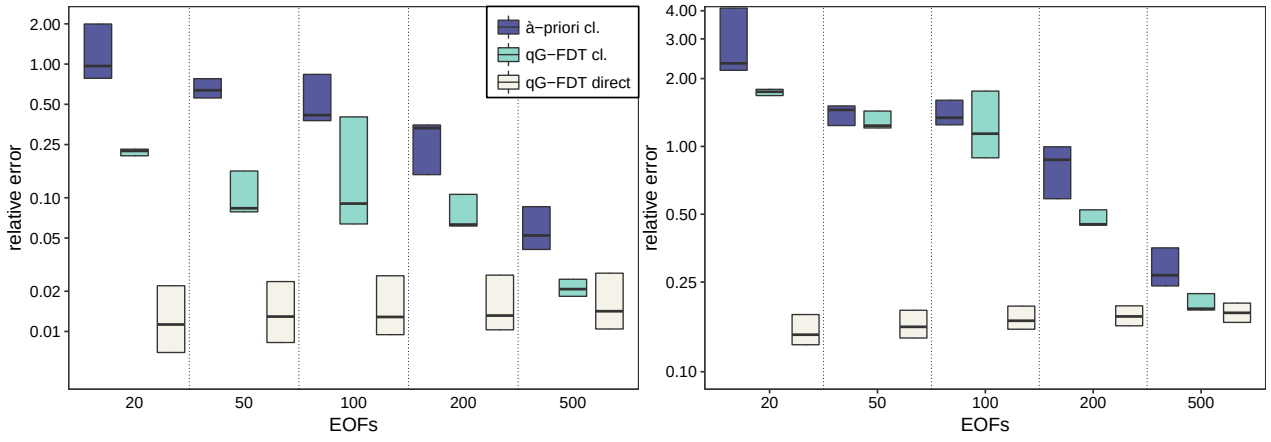


FIG. 7. Same as Figure 3 but for the global anomalous forcings represented by EOFs $\mathbf{e}_k$ with $k \in \{1,2,\ldots,5\}$. For comparison with the local anomalous forcing case the results are shown as a boxplot.



832 FIG. 8. As Figure 6 but for the global anomalous forcings represented by EOFs $\mathbf{e}_k$ with $k \in \{1,2,\ldots,5\}$. For
833 comparison with the local anomalous forcing case the results are shown as a boxplot.

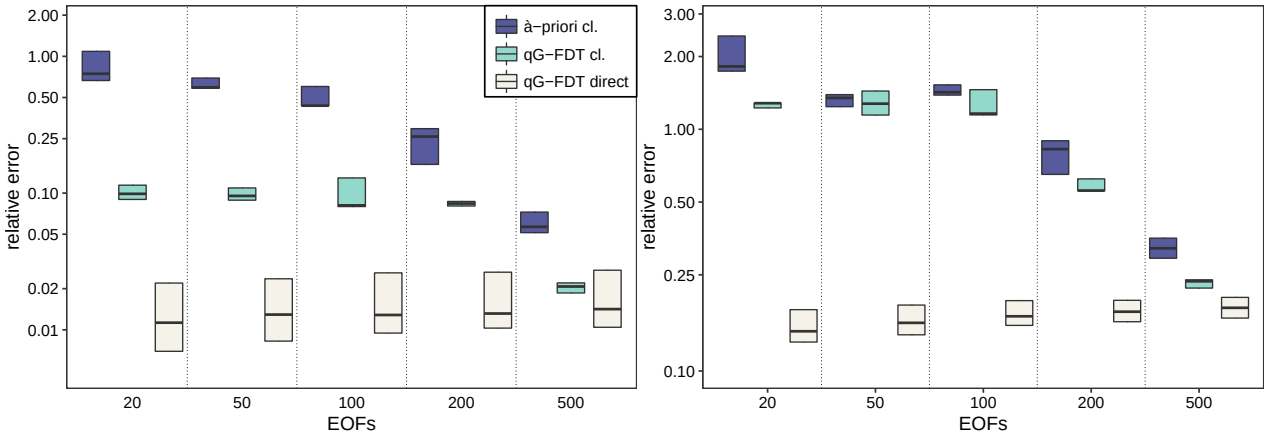


FIG. 9. As Figure 8 but for the stochastic parameterization.

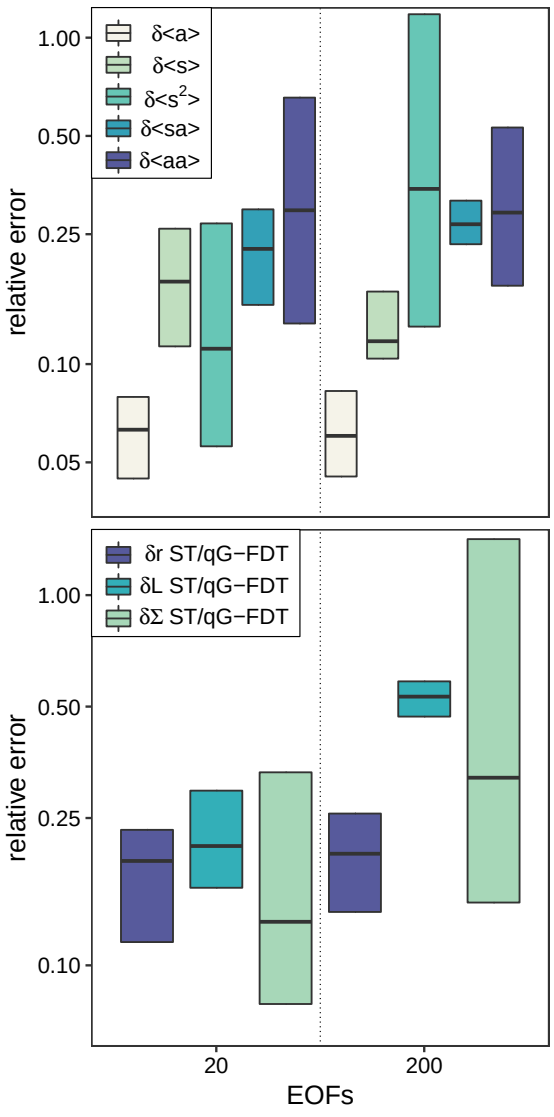


FIG. 10. Same as Figure 3, however, showing the result of the ST/qG-FDT algorithm.