Clim Dyn
DOI 10.1007/s00382-007-0325-y

Optimisation of simplified GCMs using circulation indices and maximum entropy production

Torben Kunz · Klaus Fraedrich · Edilbert Kirk

Received: 2 August 2006 / Accepted: 12 September 2007

Abstract Two kinds of objective functions for parameter optimisation in simplified general circulation models (SGCMs) are introduced and tested with an SGCM employing linear parameterisations for diabatic heating, surface friction and horizontal diffusion. (a) A set of circulation indices is introduced to characterise the zonal mean primary and secondary circulation and the global energetics. The objective function is then given by the distance between the modelled and a reference (e.g. observed) circulation in a state space spanned by these indices. (b) The global and time mean entropy production and kinetic energy dissipation are introduced as additional objective functions, following the maximum entropy production principle. It is found that both methods lead to optimal parameter values close to the standard configuration of the model, though the method of the second kind is restricted to those model parameters associated with internal processes such as heat and momentum fluxes.

1 Introduction

Atmospheric circulation models are represented by a set of nonlinear differential equations, describing conservation of energy, mass and momentum. Depending on the scale of the phenomena to be modelled and the aim of the application, adequate approximations are made. Solving these equations necessitates discretisation and numerical solution which limits the spatial resolution. Processes below this spatial scale and those excluded by the approximations made cannot be simulated explicitly by the model and are taken into account by proper parameterisations. The formulation of these parameterisations include parameters which, in general, need to be tuned to obtain reasonable solutions.

There is a number of studies on the optimal choice of parameters of different types of general circulation models (GCMs). For example, Severijns and Hazeleger (2005) optimise parameters of the parameterisations of radiation, clouds and convection in an intermediate-complexity atmospheric general circulation model (GCM) by defining several multi-dimensional fields of model variables as an optimisation target and applying a downhill simplex method to minimise an objective function, which is based on errors between model and target fields. Jones et al. (2005) optimise model parameters of a low resolution version of a complex atmospheric GCM by defining the climatology of the high resolution version as the target and choosing as the objective function to be maximised the Arcsin–Mielke score, also taking into account several multi-dimensional fields of different model variables. Lunkeit et al. (1998) and Blessing et al. (2004) optimise the simplified general circulation model PUMA (Portable University Model of the Atmosphere; see Fraedrich et al. 1998, 2005b), which employs simple linear parameterisations for diabatic heating, surface friction and horizontal diffusion, and is a further development of the multi-layer spectral model described in Hoskins and Simmons (1975), also comparable with the Held and Suarez (1994) type dynamical core. In these two cases the involved parameters are tuned with respect to the simulated zonal and time mean temperature, and objective functions are defined by an Euclidean distance between the modelled and target zonal mean temperature. Lunkeit et al. (1998) use a

T. Kunz (✉) · K. Fraedrich · E. Kirk
Meteorologisches Institut, Universität Hamburg, Bundesstraße 55, 20146 Hamburg, Germany
e-mail: torben.kunz@zmaw.de

nudging method, where the forcing field is corrected at every time step by a damping term dependent on the difference between the actual zonal mean model temperature and the target. The forcing field then quickly converges towards a reasonable solution. Blessing et al. (2004) use an adjoint version of the model to obtain information about the gradient of the objective function with respect to the relaxation temperature field.

This study proposes parameter optimisation methods for simplified general circulation models (SGCMs), ranging from dynamical cores to intermediate-complexity models. SGCMs are often used for basic process studies, focussing on certain key aspects of the atmospheric general circulation, and are generally characterised by low computational costs. This type of model is also suitable for testing different basic parametrisation approaches, for example, for investigation of the effect of stochastic forcing on the large scale circulation in the context of stochastic parameterisation (e.g. Seiffert et al. 2006). The optimisation methods proposed here support the modeller in SGCM development or setup of new model experiments. They are tested with the aforementioned SGCM PUMA, and are of the following two different kinds. (a) An objective function is defined based on a small number of circulation indices, which characterise the general circulation and can be calculated for the circulation state at a given parameter configuration or, similarly, the circulation state of the real atmosphere, or another model. By introducing a metric on the state space, spanned by these indices, the objective function is given by the distance between the model circulation state and a reference state (e.g. real atmosphere), which is to be minimised. Characterising the circulation by a set of indices allows focusing on certain features of the circulation, according to the respective application of the model. Furthermore, it reflects the purpose of SGCMs to simulate certain characteristic processes of the general circulation by reducing the complexity of the system. More complex GCMs, on the other hand, simulating a more realistic climate, may be better optimised by directly comparing the (in general zonally varying) fields of model variables with corresponding fields from observations or higher resolved model simulations, as suggested by Severijns and Hazeleger (2005) and Jones et al. (2005). For the test case shown in this study, a set of indices is chosen to describe the mean primary and secondary circulation and the global energetics in terms of the Lorenz energy cycle (see Sect. 2). (b) Alternatively to this objective function, which is dependent on the choice of indices and relates the model to an externally prescribed reference state, the global mean entropy production and the kinetic energy dissipation are calculated, motivated by the selection principle of maximum entropy production of non-equilibrium stationary states of open systems (e.g. Dewar 2003). That is, with this method parameters are optimised by a thermodynamic principle inherent to the system under consideration. It turns out that both quantities are maximised near the optimal values of the subset of those model parameters, which can be related to internal processes (internal parameters).

Given the aim of the study to find optimal parameters for SGCMs, the outline is as follows: in Sect. 2 the parameters of the PUMA model are introduced and the different objective functions are defined. The results from applications to the PUMA model are presented in Sect. 3 and the limitations of the methods are discussed. A summary and conclusions follow in Sect. 4.

2 Model and Methods

For this study the SGCM PUMA is used, which is documented in detail in Fraedrich et al. (1998, 2005b). The dry hydrostatic model solves the primitive equations on the sphere and utilises the spectral transform method (Eliasen et al. 1970; Orszag 1970) and a semi-implicit time differencing scheme. The model integrations presented in the next section are performed with a triangular truncation at wavenumber $n_T = 21$ and with five equally spaced σ-levels in the vertical (T21L5). The model equations further include the following linear parameterisations.

Diabatic heating is parameterised by Newtonian cooling, where the model temperature T is relaxed towards a constant three-dimensional relaxation temperature field T_R within the timescale τ_H:

$$\left(\frac{\partial T}{\partial t}\right)_{\text{Heating}} = \frac{T_R - T}{\tau_H}. \tag{1}$$

The relaxation temperature field T_R provides the thermal forcing, which drives the general circulation, and is essentially determined by the equator-to-pole difference $(\Delta T_R)_{\text{EP}}$ at the surface and its global mean vertical lapse rate L. The values of the timescale τ_H at the five model levels are specified as follows (from top to bottom):

$$\begin{aligned} &\text{levels 1 to 3}: \quad \tau_H = \tau_H{}^*, \\ &\text{level 4}: \quad \tau_H = (0.8\alpha + 0.2)\tau_H{}^*, \\ &\text{level 5}: \quad \tau_H = \alpha\tau_H{}^*. \end{aligned} \tag{2}$$

The parameter τ_H^* sets the heating timescale in the free atmosphere and, by the factor for boundary layer diabatic heating α $(0 < \alpha \leq 1)$, the timescale at the lower model levels can be reduced accounting for the faster heating timescale there due to turbulent vertical heat fluxes from the surface.

Surface friction is applied to relative vorticity ζ and divergence D by a Rayleigh friction term at the lowermost model level, with a timescale τ_F:

Journal : Large 382 | Dispatch : 8-10-2007 | Pages : 11
Article No. : 325 | LE | TYPESET
MS Code : 325 | CP | DISK

$$\left(\frac{\partial(\zeta, D)}{\partial t}\right)_{\text{Friction}} = -\frac{(\zeta, D)}{\tau_F}. \tag{3}$$

Horizontal diffusion is represented by a scale selective 8th order hyperdiffusion term, which parameterises subgrid scale mixing by non-resolved eddies and their effect on the energy and enstrophy cascade, with vertically independent timescale τ_D:

$$\left(\frac{\partial(T, \zeta, D)}{\partial t}\right)_{\text{Diffusion}} = -k\frac{\nabla^8(T, \zeta, D)}{\tau_D},$$

$$\text{where} \quad k = \left(\frac{a^2}{n_T(n_T+1)}\right)^4, \tag{4}$$

and a is the Earth's radius.

Thus, with this formulation the model contains six model parameters. $(\Delta T_R)_{\text{EP}}$ and L, which set up the external forcing of the general circulation, are referred to as external parameters, whereas τ_H^*, τ_F, τ_D and α, which are related to internal processes like diabatic heating, dissipation and mixing, are termed internal parameters. The standard values for these parameters are listed in Table 2, second column, as used for various experiments with PUMA (e.g. Frisius et al. 1998).

In the following three different objective functions for optimising parameters of SGCMs are defined, which are then tested with the PUMA model. The index based objective function (a) differs from those based on entropy maximisation (b) in the following points: the advantage appears to be the fact that the characteristics of the model circulation to be tuned can be determined externally in this method, since the choice of indices is, in general, arbitrary and the target of the optimisation can be set manually (to, for example, the circulation of the real atmosphere or that of the same model with higher resolution) by specifying corresponding reference index values (see below). On the other hand, the entropy maximisation based methods are independent of any external information and, therefore, objectivity is fully ensured. Thus, while method (a) may be used by the modeller to directly tune SGCM parameters with respect to the desired aspects of the general circulation, method (b) allows to check consistency of any parameter configuration (for example, that found by the former method) with the entropy maximisation principle, though it may also be used as an objective function itself.

2.1 Index based objective function

A common way of interpreting the atmospheric general circulation is to analyse the dynamics of the zonally averaged circulation and the global energetics in terms of the Lorenz energy cycle. This is particularly useful for investigating the circulation modelled by SGCMs with zonally symmetric forcing and boundary conditions, where zonal asymmetries are restricted to transient eddies and the time mean climate is independent of longitude. Additionally, the zonal mean circulation is usually separated into the zonal component (primary circulation) and the ageostrophic components in the vertical-meridional plane (secondary circulation). The basic features of the primary circulation are the westerlies in mid and high latitudes with maxima in the upper tropospheric jet streams and easterlies in low latitudes, whereas the secondary circulation is dominated by the meridional overturning in the Hadley and Ferrel cells and can be expressed by the meridional mass streamfunction ψ, defined in the "Appendix". The Lorenz energy cycle quantifies the conversions between available potential energy A and kinetic energy K and additionally those between the zonal mean, AZ and KZ, and the eddy parts, AE and KE, respectively. It thereby gives information in a concise way about the globally integrated interactions between the different parts of the circulation and about processes involved, like for example baroclinic activity and dissipation of kinetic energy. The various energy conversions are also given in the "Appendix".

To construct an objective function for the parameter optimisation problem, first, a set of circulation indices I_i is to be defined for this method. For the test case with the PUMA model presented in this study, the following choice of indices is made to cover the above mentioned features of the general circulation resolved by this SGCM (see Table 1 for detailed definitions): The zonal mean circulation is characterised by the strength of the tropospheric circumpolar vortex at 500 hPa, also known as the zonal index measuring strong/weak mid-latitude westerlies (I_1), the intensity of the tropospheric jet stream, resulting from non-linear eddy-zonal flow feedbacks (I_2) (primary circulation) and by the width (I_3) and intensity (I_4) of the Hadley cell (secondary circulation). The global mean energetics in terms of the Lorenz energy cycle are characterised by the baroclinic energy conversions (I_5 and I_6) and the barotropic energy conversion (I_7). No additional index is used for characterisation of the Ferrel cell, since they are driven by mid-latitude baroclinic wave activity, already measured by the energy conversions. These indices are then combined to the *state vector* $\vec{I} := (I_1, I_2, \ldots, I_N) \in Z \subset \mathbb{R}^N$, where Z is the *state space*, and, for this particular choice of indices, it is $N = 7$. This state vector can be calculated from model output as well as for any observed circulation state. In this context it represents a long term mean over at least 10 years. For another than this test case different indices may be chosen, depending on the respective application of the method. Again, this appears to be an advantage of this method. If, for example, the model is extended to the stratosphere, indices for the winter stratospheric polar vortex and for the flux of wave activity across the tropopause might be reasonable, or

Journal : Large 382 | Dispatch : 8-10-2007 | Pages : 11
Article No. : 325 | ☐ LE | ☐ TYPESET
MS Code : 325 | ☑ CP | ☑ DISK

Table 1 Circulation indices $I_1,\ldots,I_7$: z_{500} is the 500 hPa geopotential height, u_{300} the zonal wind at 300 hPa, ϕ the latitude, and ϕ_1, ϕ_2 are the first and second zero crossings of the time, zonal and vertical mean mass streamfunction ψ (see "Appendix") counted from the equator, A is the available potential energy and K the kinetic energy, each separated into zonal mean (Z) and eddies (E). Also specified are the reference state {\vec{I}}_R=((I_R)_1,\ldots,(I_R)_7) taken from observations and the weighting vector {\vec{w}}=(w_1,\ldots,w_7) used in Sect. 3

Index	Definition	Units	$(I_R)_i$	w_i
Zonal index	$I_1 := \overline{[z_{500}(30°) - z_{500}(60°)]}$	gpm	520	15
Jet intensity	$I_2 := \max(\overline{[u_{300}]}(\phi))$	m s^{-1}	24	1
Hadley cell width	$I_3 := \Delta\phi = \phi_2 - \phi_1,\ \widehat{\overline{[\psi]}}(\phi_{1,2}) = 0$	° lat	28	5
Hadley cell intensity	$I_4 := \max(\overline{[\psi]}(\phi, p))$	10^{10}kg s^{-1}	6.5	2.0
Baroclinic conversion I	$I_5 := \langle AZ \rightarrow AE \rangle$	W m^{-2}	1.25	0.5
Baroclinic conversion II	$I_6 := \langle AE \rightarrow KE \rangle$	W m^{-2}	2.00	0.5
Barotropic conversion	$I_7 := \langle KE \rightarrow KZ \rangle$	W m^{-2}	0.35	0.5

in the case of applying a stochastic forcing to the model dynamics (e.g. Perez-Munuzuri et al. 2005; Seiffert et al. 2006), an index for the dominant mid-latitude zonal wave-number. Also the number of indices may vary for different applications of this method.

Next, for the construction of the objective function a metric d is introduced on the state space Z as *distance* between two states $\vec{I}_a$ and $\vec{I}_b$:

$$d : Z \times Z \rightarrow \mathbb{R},\quad d(\vec{I}_a, \vec{I}_b) := \frac{\sqrt{\sum_{i=1}^{N} \left[w_i^{-1} \left((I_a)_i - (I_b)_i \right) \right]^2}}{\sqrt{N}}, \tag{5}$$

with $w_i > 0$. The w_i are called *metric weights* and $\vec{w} := (w_1, \ldots, w_N)$ the *weighting vector*. Note that $d(\vec{I}_a, \vec{I}_b) = 1$, if $|(I_a)_i - (I_b)_i| = w_i$ for all i.

Finally, with the *model* $M : P \rightarrow Z$, $M(\vec{P}) = \vec{I}_M$ symbolising an integration of the circulation model with a parameter configuration represented by the *parameter vector* $\vec{P} \in P$, with the m-dimensional *parameter space* P, that results in the model circulation state $\vec{I}_M$, the *objective function* F can be defined as

$$F : P \times Z \rightarrow \mathbb{R},\quad F(\vec{P}, \vec{I}_R) = d(M(\vec{P}), \vec{I}_R) = d(\vec{I}_M, \vec{I}_R), \tag{6}$$

with reference state $\vec{I}_R$ (e.g. state of observed circulation). In the context of the parameter optimisation problem this objective function F is to be minimised by variation of the parameter vector $\vec{P}$ at fixed $\vec{I}_R$. The required termination condition for the minimisation of F is specified as $F(\vec{P}, \vec{I}_R) \leq 1$.

2.2 Entropy production and kinetic energy dissipation

Alternatively to the objective function of the previous subsection, which is dependent on the choice of indices, the weighting vector and the reference state, two further objective functions are introduced here, which are based on maximisation principles. Open systems with many degrees of freedom in quasi stationary states far from equilibrium appear to always approach such states associated with the maximum possible entropy production (MEP) under the given boundary conditions. Several investigations, starting from Paltridge (1975), appear to support this principle and show that the configuration of the long term mean large scale horizontal atmospheric and oceanic heat fluxes corresponds to a state of maximum entropy production (Dewar 2003). Ozawa and Ohmura (1997) also show this for the atmospheric vertical convective heat fluxes. This is also associated with a state of maximum vertical heat fluxes and maximum kinetic energy dissipation (MKD; see below). In this context long term mean states of the circulation are interpreted as quasi stationary states, which are subject to internally generated as well as to externally forced (e.g. annual cycle) fluctuations allowing the system to transit between different states and, therefore, to select the state of MEP or MKD, respectively (Paltridge 1979).

In a circulation model some internal processes, like diabatic heating, surface friction and horizontal diffusion in the case of the PUMA model used in this study, are parameterised and, therefore, cannot adjust to and select the state of MEP/MKD. Instead that state can be found by tuning the corresponding model parameters. Kleidon et al. (2003) vary the timescale of surface friction in the same model and find a state of MEP at a reasonable value for this parameter. Here, this behaviour of entropy production and additionally that of kinetic energy dissipation is used to define these quantities as alternative objective functions for optimisation of model parameters, in comparison to the results of the index based objective function. The change of entropy dS at heat supply dQ and at temperature T is given by

$$\mathrm{d}S = \frac{\mathrm{d}Q}{T}. \tag{7}$$

The global and time mean of entropy production in the atmosphere η is then

$$\eta = \frac{c_p}{g}\left\{\overline{\int_0^{p_s} \frac{1}{T}\frac{\partial T}{\partial t}\mathrm{d}p}\right\}, \tag{8}$$

in units of $[\mathrm{Wm}^{-2}K^{-1}]$, $\{X\}$ and $\overline{X}$ denote a global horizontal and a time mean, respectively. The global and time mean dissipation of kinetic energy D_{kin} must be equal to its production, given by the conversion of available potential energy A into kinetic energy K in terms of the Lorenz energy cycle [see Eqs. (13) and (15) in the "Appendix"]:

$$D_{kin} = \langle AE \rightarrow KE\rangle + \langle AZ \rightarrow KZ\rangle. \tag{9}$$

From Eqs. (13) and (15) it is clear that this sum represents vertical atmospheric heat fluxes. Thus, a state of MKD is associated with a state of maximum vertical heat fluxes. Since η is related to fluxes of heat energy and D_{kin} is related to fluxes of available potential energy and kinetic energy, both objective functions part of global mean energetics.

However, in spite of the large evidence it is important to note that the MEP principle is still controversial, and its applicability to the climate has not been convincingly demonstrated (Whitfield 2005). Therefore, the provisional status of MEP for climate should be considered when using the MEP based objective functions (8) and (9).

3 Results

Employing PUMA a test is performed of the sensitivity of the index based objective function and that of the entropy production and dissipation of kinetic energy to changes of the model parameters. The following six parameters (introduced in Sect. 2) are varied through a range, where the model can be stably integrated: The equator-to-pole difference $(\Delta T_R)_{\mathrm{EP}}$ and vertical lapse rate L of the relaxation temperature field, the timescales τ_H^*, τ_F and τ_D for the diabatic heating, surface friction and horizontal hyperdiffusion, respectively, and the factor for the boundary layer diabatic heating α (see Eqs. 1–4). Thus, each parameter configuration can be represented by a parameter vector $\vec{P} \in P$ in the six-dimensional parameter space P. For each parameter configuration a 13.5 years model integration with perpetual equinox conditions was performed and the first 18 months were discarded to avoid effects of the model's spinup phase.

3.1 Index based objective function

To calculate the index based objective function F (Sect. 2), it is necessary to specify a reference circulation state $\vec{I}_R$ and a weighting vector $\vec{w}$. Here, $\vec{I}_R$ is chosen to represent the long term annual mean state of the observed circulation. The primary circulation indices $(I_R)_1$ and $(I_R)_2$ are calculated from the NCEP/NCAR reanalysis dataset (Kalnay et al. 1996), the secondary circulation indices $(I_R)_3$ and $(I_R)_4$ are estimates taken from Peixoto and Oort (1992) and the indices for the energy conversions $(I_R)_5$, $(I_R)_6$ and $(I_R)_7$ are estimates from values given in Oort and Peixoto (1974), Oort and Peixoto (1983) and Arpe et al. (1986). The metric weights w_i reflect the estimated uncertainties in determining the indices by comparing different climatologies. The values of the $(I_R)_i$ and w_i are given in Table 1.

First, each of the six model parameters is varied separately (with other parameters held fixed at their standard values). The response of the index based objective function F to changes of the model parameters, together with its inter-annual variability σ_d (standard deviation of annual means), are shown in Fig. 1. Except for the boundary layer diabatic heating α, F exhibits clear minima at certain values for all parameters. The minimum for α is only weakly pronounced. The minimum values $P_{\min}$ (after a second iteration with smaller parameter increments around the minima of the first iteration) are almost identical to the original standard parameter configuration $\vec{P}_{\mathrm{standard}}$ except for the timescale of the hyperdiffusion (τ_D = 8 days) and are summarised in Table 2, together with the corresponding minima of $F(P_{\min}, \vec{I}_R)$ and its variability σ_d. Note that neither the value of F of its secondary minimum at τ_D = 1.5 h, nor that at τ_D = 6 h (i.e. $\vec{P}_{\mathrm{standard}}$) differs significantly in terms of the inter-annual variability from the absolute minimum at τ_D = 8 days. The minima of $F(P_{\min}, \vec{I}_R)$ are all less than the value for the standard parameter configuration, $F(\vec{P}_{\mathrm{standard}}, \vec{I}_R) = 1.43$, thus indicating a slight improvement of the parameter tuning already in the case of single parameter variation.

Next, as an example for two parameter variation the timescale of the diabatic heating τ_H^* and that of the boundary layer heating α are varied. The objective function F as a function of these two parameters is shown in Fig. 2. The topography of F exhibits a valley including the corresponding standard parameter configuration $\vec{P}_{\mathrm{standard}}$ (with τ_H^* = 30 days, α = 0.17) for small α and the absolute minimum of F at large α (τ_H^* = 15 days, α = 1). It is interesting that the latter one, with uniform diabatic heating timescale, indeed corresponds to another standard parameter configuration of the model, with F = 1.25, which is only, but not significantly, outperformed by the minimum found in the single parameter variation of τ_D, with F = 1.20 (see Table 2). These alternative standard values

Journal : Large 382 Dispatch : 8-10-2007 Pages : 11
Article No. : 325 □ LE □ TYPESET
MS Code : 325 ✔ CP ✔ DISK

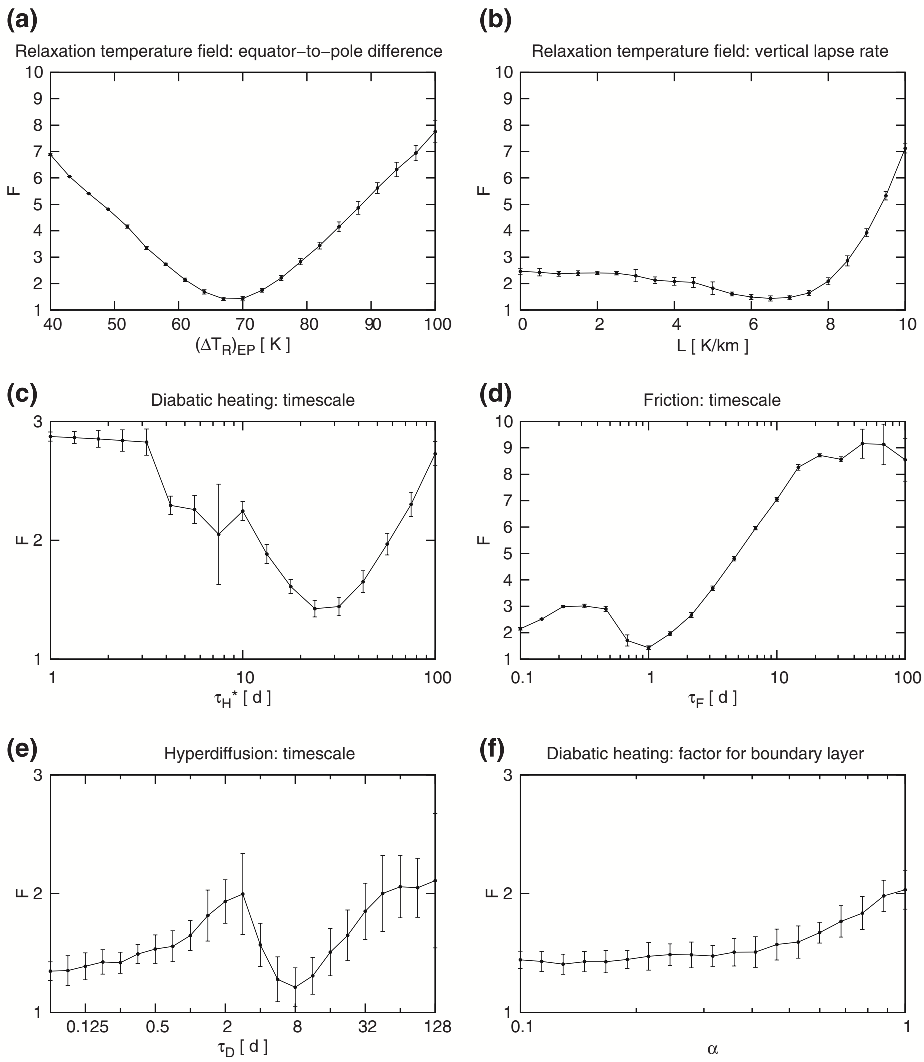

Fig. 1 Single parameter optimisation: index based objective function F as function of the parameters $(\Delta T_R)_{\mathrm{EP}}$, L, τ_H^*, τ_F, τ_D and α (see text for details). Vertical bars represent inter-annual variability ($\pm\sigma_d$, standard deviation of annual means)

Table 2 Single parameter optimisation: optimal model parameter values $P_{\min}$ and corresponding minima of index based objective function F with inter-annual variability σ_d (standard deviation of annual means), and optimal model parameter values $P_{\max}$ with respect to global entropy production $[\eta]$ and kinetic energy dissipation $[D_{\mathrm{kin}}]$

Parameter	P_{standard}	$P_{\min}$ $[F]$	$F(P_{\min}, \vec{I}_R) \pm \sigma_d$	$P_{\max}$ $[\eta]$	$P_{\max}$ $[D_{\mathrm{kin}}]$
$(\Delta T_R)_{\mathrm{EP}}$	70 K	69.0 K	1.39 ± 0.06	–	–
L	6.5 K km^{-1}	6.5 K km^{-1}	1.42 ± 0.08	–	–
τ_H^*	30 days	29 days	1.41 ± 0.08	24 days	18 days
τ_F	1 days	0.85 days	1.35 ± 0.12	2.2 days	2.2 days
τ_D	6 h	8 days	1.20 ± 0.13	1.4 days	–
α	0.17	0.13	1.40 ± 0.09	0.22	0.28

For comparison the standard parameter configuration P_{standard} is also given (with $F(\vec{P}_{\mathrm{standard}}, \vec{I}_R) = 1.43$)

for the diabatic heating, $\tau_H^* = 30$ days and $\alpha = 0.17$, were used, for example, by [1,000] to enhance the effect of zonally asymmetric low level thermal forcing of localised storm tracks.

When the two parameter optimisation is extended to three parameters by adding the timescale of the hyperdiffusion τ_D two not significantly different minima (in terms of σ_d) are found, which are close to the two different

Journal : Large 382 | Dispatch : 8-10-2007 | Pages : 11
Article No. : 325 | ☐ LE | ☐ TYPESET
MS Code : 325 | ☑ CP | ☑ DISK

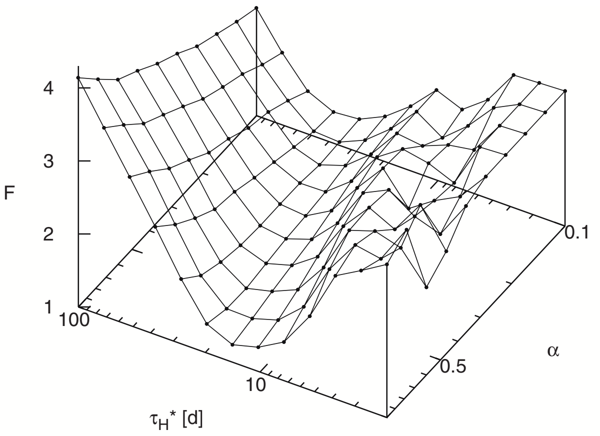

Fig. 2 Two parameter optimisation: Index based objective function F as function of the timescale of diabatic heating τ_H^* and the factor for boundary layer diabatic heating α

standard parameter configurations with respect to τ_H^* and α. In particular, the minima are at $\tau_H^* = 40$ days, $\alpha = 0.15$, $\tau_D = 4$ days with $F = 1.09$ ($\sigma_d = 0.17$) and at $\tau_H^* = 14$ days, $\alpha = 1$, $\tau_D = 1.5$ h with $F = 1.11$ ($\sigma_d = 0.08$). At these minima F is close to unity and, therefore, the termination condition (see Sect. 2) is approximately fulfilled. They further correspond to the primary and secondary minimum of F of the single parameter optimisation of τ_D. However, here the values of τ_D are less separated, suggesting that optimisation of this parameter is strongly dependent on the other model parameter. This indicates the need of multiple parameter optimisation. The results shown here from this three parameter as well as those of the single parameter optimisation by the index based objective function F, together with the choice of indices made for this test application, yield hyperdiffusion timescales not significantly different from the standard value of 6 h. Also, the minima found are insignificant over a wide parameter range. This might be overcome by an additional index measuring the slope of a log–log kinetic energy spectrum with respect to total wavenumber, since the hyperdiffusion essentially parameterises the energy and enstrophy cascades, in particular the interaction of nonresolved subgrid scale dynamics with the explicitly simulated large scale dynamics.

3.2 Entropy production and kinetic energy dissipation

To evaluate the suitability of the entropy production η and the kinetic energy dissipation D_{kin} as alternative objective functions for model parameter tuning independent of external information (choice of indices, metric weights, reference state), the response of these quantities to changes of the six model parameters is also investigated. The results are shown in Fig. 3 and some details are listed in Table 2. The response to those parameters setting up the geometrical form of the forcing field, $(\Delta T_R)_{\text{EP}}$ and L, is characteristically different to the response to those associated with the timescales τ_H^*, τ_F, τ_D and α. Whereas variation of $(\Delta T_R)_{\text{EP}}$ and L, related to the external forcing of the general circulation (external parameters), leads to monotonic behaviour of η and D_{kin}, by variation of τ_H^*, τ_F, τ_D and α, related to internal processes like heat and momentum fluxes (internal parameters), maxima of η and D_{kin} are attained at certain parameter values $P_{\max}$, except for the kinetic energy dissipation as a function of the timescale of the hyperdiffusion τ_D.

The response of η and D_{kin} to changes of external and internal parameters can be understood as follows. The external parameters set the forcing that drives the whole system, that is, larger values of $(\Delta T_R)_{\text{EP}}$ and L lead to greater horizontal and vertical temperature gradients, respectively, which, in turn, lead to stronger equator-to-pole and vertical heat fluxes by more effective mid-latitude baroclinic eddies and zonal mean meridional circulation. This is directly associated with increased entropy production and kinetic energy dissipation. Variation of internal parameters, on the other hand, leads to changes of the efficiency of the atmospheric heat engine. Strong surface friction (small τ_F) largely suppresses any circulation, in particular, mid-latitude baroclinic systems, reducing atmospheric heat transport, entropy production and kinetic energy dissipation. Weak surface friction makes the circulation less baroclinic, which is then dominated by the barotropic governer and heat transport is again little effective (Kleidon et al. 2003). Thus, in between these extremes a maximum of η, and also D_{kin}, must exist. Similar arguments hold for the diabatic heating. Strong heating (small τ_H^* or α) suppresses zonal eddy structures and vertical motion, reducing heat transport and dissipation. Weak heating leads to weak temperature gradients, and thus, heat transport is ineffective and little kinetic energy is dissipated. A maximum is attained at intermediate heating rates.

The hyperdiffusion timescale, however, is the only internal parameter with a characteristically different behaviour of η and D_{kin}, respectively. Variation of the other model parameters leads to qualitatively similar behaviour of these two quantities, since the corresponding changes of entropy production are mainly due to changes of the large scale heat transport by the explicitly simulated circulation, which is dominated by the mid-latitude baroclinic eddies, and therefore also induce changes of the kinetic energy dissipation in terms of the Lorenz energy cycle (9). Strongly increasing the horizontal diffusivity (decreasing τ_D) reduces the amplitude and thereby the efficiency of the large scale eddies, and thus the associated heat transport

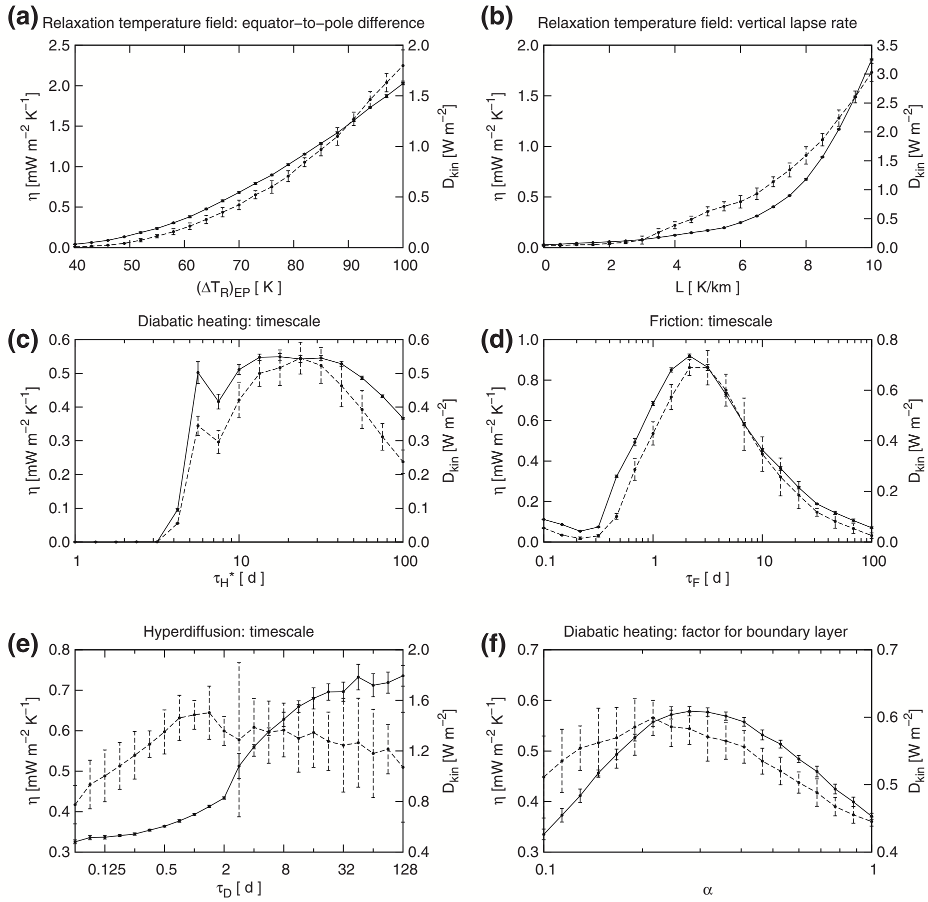

Fig. 3 Global and time mean entropy production η (*dashed*) and dissipation of kinetic energy D_{kin} (*solid*) as function of the parameters $(\Delta T_R)_{\mathrm{EP}}$ and L (external parameters) and of τ_H^*, τ_F, τ_D and α (internal parameters). Vertical bars represent inter-annual variability (±SD, standard deviation of annual means)

and entropy production as well as the kinetic energy dissipation. Decreasing the horizontal diffusivity (increasing τ_D), on the other hand, essentially reduces the entropy production by subgrid scale mixing parameterised by the hyperdiffuion, but does not affect the large scale circulation pattern above some value of τ_D, leading to a maximum of η at intermediate values of horizontal diffusivity, but monotonically increasing and saturating D_{kin} for increasing τ_D.

In Table 2 the P_{max} are compared to the P_{min} where F attains its minima, and to P_{standard}. Most of the standard parameter values and those found by minimisation of F are not significantly different from the P_{max}, in the sense that they fall into the parameter range given by η and D_{kin}, respectively, together with the corresponding year-to-year standard deviation (see Fig. 3). These parameters found by the index based objective function can, therefore, be interpreted as being consistent with the entropy maximisation principle. Only the frictional timescale τ_F maximises the entropy production and kinetic energy dissipation at values significantly different from P_{standard} and P_{min} (at $P_{\mathrm{max}} = 2.2$ days). This result suggests that the standard value of $\tau_F = 1$ days should be slightly increased, at least in the context of this single parameter variation, and, moreover, it demonstrates that any model tuning to externally prescribed circulation features (e.g. minimisation of F) can lead to parameter configurations, which are inconsistent with the underlying physics, here, with the entropy maximisation principle. For other model setups τ_F may maximise η and D_{kin} at values closer to its standard value. For comparison, in the Held and Suarez (1994) scheme for the frictional timescale, τ_F is set to 1.5 days at $\sigma = 0.9$, used for dynamical core models similar to PUMA. Finally, it should be noted that in the PUMA model the frictional dissipation of kinetic energy is not taken into account by the thermodynamic equation, which is, however, strongly related to the entropy budget.

In the context of η and D_{kin} as stand alone objective functions, reasonable parameter values are found for the frictional and heating timescales, though the entropy production maximum with respect to τ_D is only weakly pronounced. This suggests that η and D_{kin} are also valuable as objective functions for the optimisation of internal model parameters. Nevertheless, a detailed inspection of the results is an important step when using this method, as it is the case for the method based on minimisation of F. It

Journal : Large 382 | Dispatch : 8-10-2007 | Pages : 11
Article No. : 325 | ☐ LE | ☐ TYPESET
MS Code : 325 | ☑ CP | ☑ DISK

is noteworthy the smaller standard deviation and, hence, the higher significance of the maxima of D_{kin} compared to η, making D_{kin} the more favourable objective function.

4 Summary and conclusions

Two kinds of objective functions for parameter optimisation in SGCMs are introduced and successfully tested with the primitive equation model PUMA. First, an objective function based on a small set of circulation indices is defined. For the test case presented in this study these indices characterise the zonal mean primary and secondary circulation and the global energetics in terms of the Lorenz energy cycle. On the state space spanned by these indices, a metric is defined and thus the parameter optimisation reduces to a minimisation of the distance between the modelled and a reference state, which is chosen as that of the observed general circulation. The examples of single, two and three parameter optimisation yield optimum parameter values close to that of the model's standard configurations. Also the value of the timescale of the hyperdiffusion does not significantly differ from its standard value. Nevertheless, for this parameter multiple not significantly different minima are obtained, at values being shorter as well as longer than the standard parameter value, and over a relatively wide parameter range, which, however, appears to reduce for multiple parameter optimisation. An improvement of a single parameter optimisation of this parameter might be achieved by including an additional index in the set of circulation indices, which characterises the slope of the kinetic energy spectrum at large total wavenumbers. This clearly demonstrates the limitations of this method and the importance of a suitable choice of indices, according to the nature of parameters to be tuned.

Second, the global and time mean entropy production and kinetic energy dissipation as additional objective functions are calculated, motivated by the selection principle of maximum entropy production of non-equilibrium stationary states of open systems (e.g. Dewar 2003). Since by this principle the state of maximum entropy production is selected under the given boundary conditions, variation of external parameters, which set up the external forcing of the model and, therefore, the boundary conditions, leads to a monotonic increase of the these two quantities and an optimisation is not possible. Variation of internal parameters associated with internal processes such as heat and momentum fluxes, on the other hand, leads to maxima at certain parameter values, which appear to be close to the optimum values obtained from the index based objective function and to the standard values. Therefore, the optimal parameter values found by the index based method are consistent with the entropy maximisation principle. Only for the single parameter optimisation of the frictional timescale the entropy maximisation method suggests a slightly greater value compared to its standard value as well as that obtained by the minimisation method. This demonstrates one advantage of the combined use of both methods, since the parameters found by the index based method, that is, the tuning towards externally prescribed circulation features, may lead to a circulation state inconsistent with the underlying physics in terms of the entropy maximisation principle. More generally, with the index based method the model can be tuned towards any circulation state that is impossible to realise in the particular model, leading to unplausible parameter configurations. Finally, optimisation of the timescale of the hyperdiffusion by the entropy method does not give significant results since the entropy production maximum is not significant and the kinetic energy dissipation behaves characteristically different in that it does not maximise at an intermediate parameter value. It is also important to note that the numerical scheme used in the PUMA model as well as in many other state-of-the-art SGCMs introduces errors to the entropy budget of the atmosphere. Specifically, conservation of entropy following the adiabatic part of the dynamics is not strictly fulfilled by the numerics, as discussed by Woollings and Thuburn (2006), who investigate the various sources of entropy in an SGCM very similar to PUMA.

The advantages of the index based objective function can be summarised as follows. First, application of this method is not restricted to internal parameters; second, the characteristics of the circulation to be tuned can be determined externally by an appropriate choice of indices; and third, the aim of the optimisation can be set manually by specification of a suitable reference state of circulation. At the same time, this external information represents a limitation of this method, since the parameter configuration found does not need to result in any physically meaningful circulation state. The entropy production and kinetic energy dissipation as alternative objective functions, on the other hand, are independent of any external information and, therefore, are more objective for tuning the model towards a realistic circulation, using a thermodynamic principle inherent to the system (presumed the applicability of MEP to the climate). However, these maximisation methods are restricted to internal model parameters. Thus, combined use of both methods is advantageous, in the sense that the entropy maximisation based methods are used to check consistency of any parameter configuration with the entropy maximisation principle, and the index based method is used, in particular, to find new parameter configurations including external parameters.

There are many possible further applications of the methods introduced here. These include optimisation of additional external model parameters, for example, when extending the model to the stratosphere to investigate basic mechanisms of stratosphere troposphere dynamical coupling, as it is actually done with the PUMA model. This is associated with parameters setting up the strength and size of the stratospheric winter time polar vortex, the upper boundary conditions and others. Furthermore, SGCMs are useful tools for testing different approaches of stochastic parameterisation, as, for example, the study by Seiffert et al. (2006), also using PUMA. In this context, noise parameters such as its amplitude and spatial and temporal correlations enter the model as new parameters, which need to be tuned. In these cases the proposed optimisation methods can be used to look at the new model parameters from different perspectives. But also parameter tuning in intermediate-complexity SGCMs is a further possible application of these optimisation methods. The Planet Simulator (Fraedrich et al. 2005a), which is based on the dynamical core of the PUMA but includes many additional parameterisations of, e.g. radiation, clouds, convection, precipitation, vegetation, sea-ice, etc. is frequently used for different process studies of the climate system and is still under development. To this end the proposed methods support the modeller in further development. Finally we note that with respect to the optimisation technique, different methods may be applied such as, for example, the downhill simplex method or genetic algorithms.

Acknowledgments We thank Frank Lunkeit for helpful comments, Axel Kleidon for discussion of the MEP principle and also three anonymous reviewers for their valuable comments on the first draft of the manuscript.

5 Appendix

The zonal mean meridional mass streamfunction ψ is defined by (Peixoto and Oort 1992):

$$[\bar{v}] = \frac{g}{2\pi a \cos\phi}\frac{\partial\psi}{\partial p} \qquad (10)$$

$$[\bar{\omega}] = -\frac{g}{2\pi a^2 \cos\phi}\frac{\partial\psi}{\partial\phi}, \qquad (11)$$

with $\psi = 0$ at $p = 0$, and where $[\bar{v}]$ and $[\bar{\omega}]$ denote the time and zonal mean meridional and vertical velocity components, respectively, a the Earth's radius, g the gravitational acceleration, ϕ is latitude and p pressure.

The energy conversions of the Lorenz energy cycle as outlined by Ulbrich and Speth (1991) and Arpe et al. (1986) are given by:

$$\langle AZ \to AE\rangle = -\frac{\gamma}{g}\left([v^*T^*]\frac{1}{a}\frac{\partial[T]}{\partial\phi} + [\omega^*T^*] \times\left(\frac{\partial}{\partial p}([T]-\{T\}) - \frac{R}{c_p p}([T]-\{T\})\right)\right) \qquad (12)$$

$$\langle AE \to KE\rangle = -[\omega^*T_v^*]\frac{R}{gp} \qquad (13)$$

$$\langle KE \to KZ\rangle = \frac{1}{g}\left([u^*v^*]\frac{1}{a}\frac{\partial[u]}{\partial\phi} + [u^*v^*][u]\frac{\tan\phi}{a} + [v^*v^*]\frac{1}{a}\frac{\partial[v]}{\partial\phi} - [u^*u^*][v]\frac{\tan\phi}{a} + [\omega^*u^*]\frac{\partial[u]}{\partial p} + [\omega^*v^*]\frac{\partial[v]}{\partial p}\right) \qquad (14)$$

$$\langle AZ \to KZ\rangle = -([\omega]-\{\omega\})([T_v]-\{T_v\})\frac{R}{gp} \qquad (15)$$

with virtual temperature T_v, specific heat capacity at constant pressure c_p, gas constant of dry air R, zonal wind u, zonal mean [], deviation from zonal mean * (eddy part), global horizontal mean { } and the stability parameter

$$\gamma = -\frac{R}{p}\left(\frac{\partial[T]}{\partial p} - \frac{R}{c_p}\frac{[T]}{p}\right)^{-1}. \qquad (16)$$

References

Arpe K, Brankovic C, Oriol E, Speth P (1986) Variability in time and space of energetics from a long series of atmospheric data produced by ECMWF. Contrib Atmos Phys 59:321–355

Blessing S, Fraedrich K, Lunkeit F (2004) Climate diagnostics by adjoint modelling: a feasibility study. In: Fischer H, Kumke T, Lohmann G, Flöser G, Miller H, von Storch H, Negendank JFW (eds) The KIHZ project: towards a synthesis of Holocene proxy data and climate models. Springer, Berlin, pp 383–396

Dewar R (2003) Information theory explanation of the fluctuation theorem, maximum entropy production and self-organized criticality in non-equilibrium stationary states. J Phys A 36:631–641

Eliasen E, Machenhauer B, Rasmussen E (1970) On a numerical method for integration of the hydrodynamical equations with a spectral representation of the horizontal fields. Inst of Theor Met Univ Copenhagen

Fraedrich K, Kirk E, Lunkeit F (1998) Portable University Model of the Atmosphere. Tech. Rep. 16 DKRZ

Fraedrich K, Jansen H, Kirk E, Luksch U, Lunkeit F (2005a) The planet simulator: Towards a user friendly model. Meteorol Z 14:299–304

Fraedrich K, Kirk E, Luksch U, Lunkeit F (2005b) The portable university model of the atmosphere (PUMA): storm track dynamics and low-frequency variability. Meteorol Z 14:735–745

Frisius T, Lunkeit F, Fraedrich K, James IA (1998) Storm-track organization and variability in a simplified atmospheric global circulation model (SGCM). Q J R Meteorol Soc 124:1019–1043

Held IM, Suarez MJ (1994) A proposal for the intercomparison of the dynamical cores of atmospheric general circulation models. Bull Am Meteorol Soc 75:1825–1830

Hoskins BJ, Simmons AJ (1975) A multi-layer spectral model and the semi-implicit method. Q J R Meteorol Soc 101:637–655

Jones C, Gregory J, Thorpe R, Cox P, Murphy J, Sexton D, Valdes P (2005) Systematic optimisation and climate simulation of FAMOUS, a fast version of HadCM3. Clim Dyn 25:189–204

Kalnay E, Konamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu Y, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo KC, Ropelewski C, Wang J, Leetmaa A, Reynolds R, Denne R, Joseph D (1996) The NCEP/NCAR 40-year reanalysis project. Bull Am Meteorol Soc 77:437–471

Kleidon A, Fraedrich K, Kunz T, Lunkeit F (2003) The atmospheric circulation and states of maximum entropy production. Geophys Res Lett 30:2223. doi:10.1029/2003GL018363

Lunkeit F, Fraedrich K, Bauer SE (1998) Storm tracks in a warmer climate: sensitivity studies with a simplified global circulation model. Clim Dyn 14:813–826

Oort AH, Peixoto JP (1974) Annual cycle of energetics of atmosphere on a planetary scale. J Geophys Res 79:2705–2719

Oort AH, Peixoto JP (1983) Global angular-momentum and energy-balance requirements from observations. Adv Geophys 25:355–490

Orszag SA (1970) Transform method for the calculation of vector-coupled sums: application to the spectral form of the vorticity equation. J Atmos Sci 27:890–895

Ozawa H, Ohmura A (1997) Thermodynamics of a global-mean state of the atmosphere—a state of maximum entropy iIncrease. J Clim 10:441–445

Paltridge GW (1975) Global dynamics and climate—a system of minimum entropy exchange. Q J R Meteorol Soc 101:475–484

Paltridge GW (1979) Climate and thermodynamic systems of maximum dissipation. Nature 279:630–631

Peixoto JP, Oort AH (1992) Physics of Climate. American Physical Society, New York, p 520

Perez-Munuzuri V, Deza R, Fraedrich K, Kunz T, Lunkeit F (2005) Coherence resonance in an atmospheric global circulation model. Phys Rev E 71. doi:10.1103/PhysRevE.71.065602

Seiffert R, Blender R, Fraedrich K (2006) Subscale forcing in a global atmospheric circulation model and stochastic parameterisation. Q J R Meteorol Soc 132:1627–1643

Severijns CA, Hazeleger W (2005) Optimizing parameters in an atmospheric general circulation model. J Clim 18:3527–3535

Ulbrich U, Speth P (1991) The global energy cycle of stationary and transient atmospheric waves—results from ECMWF analyses. Meteorol Atmos Phys 45:125–138

Whitfield J (2005) Order out of chaos. Nature 436:905–907

Woollings T, Thuburn J (2006) Entropy sources in a dynamical core atmosphere model. Q J R Meteor Soc 132:43–59

Journal : Large 382 | Dispatch : 8-10-2007 | Pages : 11
Article No. : 325 | ☐ LE | ☐ TYPESET
MS Code : 325 | ☑ CP | ☑ DISK