**Atmospheric
Chemistry
and Physics**

Estimation of ECHAM5 climate model closure parameters with adaptive MCMC

H. Järvinen[1], P. Räisänen[1], M. Laine[1], J. Tamminen[1], A. Ilin[2], E. Oja[2], A. Solonen[3], and H. Haario[3]

[1]Finnish Meteorological Institute, Helsinki, Finland
[2]Aalto University School of Science and Technology, Espoo, Finland
[3]Lappeenranta University of Technology, Lappeenranta, Finland

Abstract. Climate models contain closure parameters to which the model climate is sensitive. These parameters appear in physical parameterization schemes where some unresolved variables are expressed by predefined parameters rather than being explicitly modeled. Currently, best expert knowledge is used to define the optimal closure parameter values, based on observations, process studies, large eddy simulations, etc. Here, parameter estimation, based on the adaptive Markov chain Monte Carlo (MCMC) method, is applied for estimation of joint posterior probability density of a small number ($n = 4$) of closure parameters appearing in the ECHAM5 climate model. The parameters considered are related to clouds and precipitation and they are sampled by an adaptive random walk process of the MCMC. The parameter probability densities are estimated simultaneously for all parameters, subject to an objective function. Five alternative formulations of the objective function are tested, all related to the net radiative flux at the top of the atmosphere. Conclusions of the closure parameter estimation tests with a low-resolution ECHAM5 climate model indicate that (i) adaptive MCMC is a viable option for parameter estimation in large-scale computational models, and (ii) choice of the objective function is crucial for the identifiability of the parameter distributions.

1 Introduction

Atmospheric general circulation models (GCMs) consist of dynamical laws of atmospheric motions and physical parameterizations of sub-grid scale processes, such as cloud formation and boundary layer turbulence. Specified parameters appear in physical parameterization schemes where some un-

resolved variables are expressed by predefined parameters rather than being explicitly modeled. These are called closure parameters. A simple example of such a parameter is provided by turbulent transfer in the atmosphere. In a first order closure, the transfer of a quantity q is assumed to be proportional to the gradient of q multiplied by a fixed diffusion coefficient – note that a whole hierarchy of closures of different orders exists, each with different closure parameters (Mellor and Yamada, 1974). Another example is cloud shortwave optical properties which depend on cloud optical thickness. This can be related to resolved cloud liquid water amount via the mean effective radius of cloud water droplets. If the cloud micro-physics is not resolved, the mean effective radius has to be prescribed (Martin et al., 1994). The modelled shortwave radiation flux is sensitive to the specified value of this parameter, and it can act as an effective "tuning handle" of the simulated climate.

An underlying principle in climate model development is to aim at few rather than many closure parameters. In the model development process, best expert knowledge is used to define the optimal parameter values. They can be constrained to some degree based on observations, process studies, large eddy simulations, etc. but they do not necessarily represent any directly observable quantity. Additionally, parameter values can depend on the discretization details, such as grid interval or choices made regarding modeling of other physical processes. This is a dilemma since observations do not provide guidance towards resolution or modeling environment dependent parameter values. In summary, the closure parameters are determined such that (i) they are consistent with prior knowledge, and (ii) simulations prove to be realistic in posterior validation. In fact, both can be used in an iterative manner to optimize model performance.

The closure parameters of atmospheric general circulation models are, by definition, constant during the model run. Therefore they should perform well independent of particular weather situations, both locally and in a global sense.

Various approaches are available for solving the closure parameter estimation problem. First, the review paper of Navon (1993) concentrates on adjoint techniques (e.g., Rinne and Järvinen, 1993) and stresses the questions of parameter identifiability and stability. This implies that both the estimation method and the parameters to be estimated need to be selected carefully. Sequential state estimation in numerical weather prediction aims at fitting the initial condition and model parameters to prior information and to observations (e.g., Dee, 2005). Only the maximum-likelihood fit and a Gaussian error covariance are obtained from solving the tangent-linear analysis equation. If closure parameters are estimated in this framework, their values partly reflect the latest observations – this is in fact in slight contradiction to the notion that the closure parameter distributions should be stationary.

Annan and Hargreaves (2007) provide a review of the available parameter estimation methods in climate modelling. They also discuss the Markov chain Monte Carlo (MCMC) method and consider it too computationally expensive for estimating climate model closure parameters. Their treatment of MCMC is, however, somewhat restricted to the Metropolis algorithm (Metropolis et al., 1953), and recent advances in adaptive methods are not fully covered.

Jackson et al. (2008) used a stochastic optimization method, Multiple Very Fast Simulated Annealing (MVFSA), for the search of optimal closure parameters of the CAM3.1 climate model, with respect to a cost function based on multi-source observations. The search paths of the stochastic optimizer were used to infer about the parametric uncertainty. Villagran et al. (2008) further evaluated the performance of MVFSA and compared it against different MCMC methods with a surrogate climate model. These methods included Adaptive Metropolis (AM), the Single Component Adaptive Metropolis (SCAM) and the Delayed Rejection Adaptive Metropolis (DRAM), see (Haario et al., 2001, 2004, 2005, 2006; Andrieu and Moulines, 2006). Villagran et al. (2008) note that MVFSA may be efficient for optimization, but statistical inference about the parameter distribution tends to be biased. The results were in favor of DRAM and SCAM, especially in test cases with short sampling chains.

In this article, we demonstrate the use of MCMC in the context of the atmospheric general circulation model ECHAM5. A central outcome is that it is, indeed, viable to use MCMC for parameter estimation for a climate model, at least in the context of a coarse-resolution atmospheric GCM. Research methods are presented in Sect. 2, experimental setup and results in Sects. 3 and 4, and discussion and conclusions in Sects. 5 and 6.

2 Materials and methods

2.1 The adaptive MCMC approach

Markov chain Monte Carlo (MCMC) methods are widely used in parameter estimation and computational inverse problems. A mathematically solid way of describing the estimation problem is to use Bayesian approach where the measurements and unknown parameters are considered as random variables and the solution is described as a combination of prior information and the evidence that comes from the measurements via the objective function (i.e., the likelihood). The solution, i.e., the estimated distribution of the retrieved parameters, is known as the posterior distribution. Instead of just finding the "best estimate", the MCMC technique simulates the full distribution of the solution in the n dimensional model parameter space, where n equals the number of parameters to be estimated.

The classical Metropolis algorithm (Metropolis et al., 1953) proceeds in two steps. In the proposal step, a candidate value is sampled using a "proposal distribution". In the acceptance step, the candidate value is either accepted or rejected. The Metropolis acceptance probability depends on the values of the objective function at the candidate value and the present value. If the value is accepted, it becomes the new value in the chain and if it is rejected, the chain just repeats the present value. More probable values are always accepted but there is a positive probability to accept less probable values, too. In this way it is assured that the whole target distribution is explored. The exact formula for the acceptance probability is selected such that the distribution of the simulated values converges to the target probability.

The original Metropolis algorithm is simple and straightforward. In practice, however, an effective performance, i.e., convergence towards the correct target distribution with reasonably few model evaluations, may require preliminary test runs or laborious hand tuning of the proposal distribution. Such methods, however, are not easily available in case of a computationally demanding climate model like ECHAM5. Recent developments speed up the search of the proposal by using adaptive techniques that 'learn' during the sampling process. In this article, we have applied the Delayed Rejection Adaptive Metropolis (DRAM) algorithm. Textbook treatment of MCMC methods can be found, e.g., in Robert and Casella (2005).

2.2 ECHAM5 model and the closure parameters

Version 5.4 of the ECHAM5 atmospheric general circulation model (Roeckner et al., 2003, 2006) was used. The dynamical part of ECHAM5 is formulated in spherical harmonics, while physical parameterizations are computed in grid point space. The simulations reported here used a coarse horizontal resolution of T21, i.e., triangular truncation at wave number 21, corresponding to a grid-spacing of 5.625 deg. The

Table 1. The considered sub-set of ECHAM5 closure parameters.

Parameter	Description
CAULOC	A parameter influencing the accretion of cloud droplets by precipitation (rain formation in stratiform clouds)
CMFCTOP	Relative cloud mass flux at the level above non-buoyancy (in cumulus mass flux scheme)
CPRCON	A coefficient for determining conversion from cloud water to rain (in convective clouds)
ENTRSCV	Entrainment rate for shallow convection

model vertical grid had 19 layers with model top at 10 hPa. A semi-implicit time integration scheme is used for model dynamics with a time step of 40 min. Model physical parameterizations (see Roeckner et al., 2006) are invoked every time step with the exception of radiation, which is computed once in two hours.

Four ECHAM5 closure parameters were considered (Table 1). These parameters are related to physical parameterizations of clouds and precipitation. The choice of these parameters is motivated by their substantial influence on model cloud fields and therefore the radiative fluxes at the top of the atmosphere (TOA). It is thus plausible that they can be constrained by a suitable formulation of the objective function.

2.3 Observational data sets

In this initial study, the definition of the objective function is based solely on the net (longwave + shortwave) radiative flux at the TOA. The observational estimates are taken from the Clouds and the Earth's Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) dataset (Loeb et al., 2009). The CERES EBAF dataset is, however, only used for the mean values. Since this dataset contains data for five years only, we chose to derive the interannual standard deviations from the the European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis data (ERA-40; Uppala et al., 2005), which provides a much longer (44-year) timeseries.

2.4 The objective function

The parameter posterior probability distribution is conditional to the choice of the objective function. In case of ECHAM5, it is a measure of the accuracy of the climate simulation – a trained human eye would be very efficient in selecting "good" and "bad" simulations and the aim here is to construct an objective function which would replace this human element. On one hand, the objective function should be physically justified, i.e., being capable of distinguishing accurate climate simulations from inaccurate ones. On the other hand, it should be constructed such that the parameter

distributions are identifiable with respect to the chosen objective function. If this is the case, the parameter posterior probability distribution should be compact and limited. If not, either the objective function does not provide the desired guidance for the parameters, or they are simply not relevant in tuning the model with respect to the objective function.

Five alternative formulations of the objective function are tested, all of which are related to the net radiative flux at the TOA in the ECHAM5 model (F) and in CERES EBAF data (F^o). Annual and monthly mean fluxes are denoted by $\overline{F}$ and $\overline{\overline{F}}$, and global and zonal means by $\langle F \rangle$ and $[F]$, respectively. Subscripts x and y refer to geographical location in zonal and meridional direction, and t refers to time (in months). The first of the five alternative formulations of the objective function is denoted by $J^G(\theta)$, and it uses only the global-annual mean value of F:

$$J^G(\theta) = \frac{\left(\langle \overline{F} \rangle - \langle \overline{F^o} \rangle\right)^2}{(\sigma^o_{\langle \overline{F} \rangle})^2}, \tag{1}$$

where θ is the vector of four closure parameters. It penalizes climate simulations which deviate from the global annual-mean net radiative flux in CERES EBAF data ($0.9\,\mathrm{Wm^{-2}}$). The squared net flux difference is normalized by the standard deviation $\sigma^o_{\langle \overline{F} \rangle}$ of the inter-annual variability of the global annual mean net flux, which is estimated from ERA-40 data ($0.53\,\mathrm{Wm^{-2}}$).

The second formulation is denoted by $J^{XY}(\theta)$

$$J^{XY}(\theta) = \frac{1}{12}\sum_{t=1}^{12}\sum_x\sum_y w_{x,y} \frac{\left(\overline{\overline{F}}_{x,y,t} - \overline{\overline{F^o}}_{x,y,t}\right)^2}{(\sigma^o_{\overline{\overline{F}}})^2} \tag{2}$$

It accounts for local differences in monthly mean net fluxes. The weights $w_{x,y}$ represent grid point area fractions. The squared net flux difference is normalized by the standard deviation of the inter-annual variability of the local monthly mean net fluxes, based on ERA-40 data. The third formulation, denoted by $J^{\mathrm{ZONAL}}(\theta)$, uses zonal mean values of monthly mean net fluxes:

$$J^{\mathrm{ZONAL}}(\theta) = \frac{1}{12}\sum_{t=1}^{12}\sum_y w_y \frac{\left([\overline{\overline{F}}_y] - [\overline{F^o}_y]\right)^2}{(\sigma^o_{[\overline{F}]})^2} \tag{3}$$

Here, the weights w_y represent area fractions for the zonal bands, and the normalizing factor is the standard deviation of the inter-annual variability in monthly and zonal mean net fluxes.

The last two formulations

$$J^{G+XY}(\theta) = J^G(\theta) + J^{XY}(\theta) \tag{4}$$

$$J^{G+\mathrm{ZONAL}}(\theta) = J^G(\theta) + J^{\mathrm{ZONAL}}(\theta) \tag{5}$$

are combinations of the objective function Eq. (1) with Eqs. (2) and (3), respectively. Equations (4) and (5) attempt to emphasize the weight of the global annual mean net flux in addition to the regional details in net radiative fluxes.

Table 2. Parameter values applied in the MCMC tests. The first column gives the default values for resolution T21L19, the second column the initial estimate of one-sigma uncertainty used to initialize the MCMC chain, the third column minimum and maximum parameter values allowed, and the fourth column the range of parameter values applied in standard ECHAM5. The upper limit of parameter CAULOC was 100 for the first four experiments, but it was reduced to 30 for the $J^{G+\text{ZONAL}}$ experiment, when it was realized that all values of CAULOC higher than about 30 produce *identical* results.

Parameter	Default value	Initial std.dev.	Range in MCMC tests	Range in ECHAM5 (other model resolutions etc.)
CAULOC	1	1	0–30/100	1–5
CMFCTOP	0.10	0.08	0–1	0.10–0.35
CPRCON	8×10^{-4}	4×10^{-4}	0–1.5×10^{-2}	1×10^{-4}–10^{-3}
ENTRSCV	3×10^{-4}	3×10^{-4}	0–5×10^{-3}	3×10^{-4}–10^{-3}

3 Experimental setup

Five separate experiments were performed, one for each of the five objective functions listed above (Eqs. 1–5). An MCMC chain consisting of 1000 model runs was applied, with one exception: a longer chain of 4500 runs was carried out for the experiment using the $J^{G+\text{ZONAL}}$ objective function (Eq. 5)[1]. Each model evaluation represents a one-year climate simulation with the low-resolution ECHAM5 model. Prescribed distributions of sea surface temperature and sea ice for year 1990 were used (AMIP Project Office, 1996), and the model initial condition was 1 January 1990. One simulation step took about 17 min using 30 CPUs on a Cray XT5m computer.

Default parameter values and prior distributions (or ranges) applied in the experiments are provided in Table 2. The MCMC algorithm was broadly as follows:

Step 0: Initialize the four closure parameters to their default values; Initialize proposal distribution to reflect the a priori knowledge about parameter uncertainty; Run the model for one year; Post-process the model data and evaluate the objective function.

Step 1: Draw new parameters from the proposal distribution centered at the current parameter values; Run the model with new parameter values and evaluate the objective function.

Step 2: Accept or reject new parameter values based on the difference of objective functions at current vs. previous step; Update the proposal distribution according to the adaptive MCMC algorithm.

Step 3: Return to Step 1 if the chain has not yet been completed.

---

[1]The actual length of the MCMC chain is slightly longer than the number of ECHAM5 runs, because the parameter values proposed by MCMC may fall outside the given parameter bounds, in which case they are rejected immediately, without running ECHAM5.

Note that the difficulty in providing a correct initial proposal covariance in Step 0 makes the adaptation method applied in Step 2 crucial for the sampling to be efficient.

4 Results

The MCMC tests with the low-resolution ECHAM5 climate model are discussed in the next four subsections, with emphasis on general aspects of the results.

4.1 Parameter chains

The random walk process is started in each experiment from the default parameter values (Table 2). The parameter values for the subsequent runs depend on the definition of the objective function. We illustrate this by showing the MCMC chains for the five different objective functions and for two parameters with contrasting behaviour: CMFCTOP (Fig. 1) and CPRCON (Fig. 2). In Figs. 1 and 2, the dots represent parameter values included in the MCMC chain. If a parameter combination is rejected by the MCMC algorithm, the previous accepted parameter values are repeated. The horizontal grey line represents the default parameter value, 0.1 for CMFCTOP and 8×10^{-4} for CPRCON, respectively. Note that in Fig. 1, the scale is different in different panels.

For CMFCTOP (Fig. 1), the parameter values are generally well-bounded from above. Only for J^{XY} the constraint on CMFCTOP is somewhat weak, the largest accepted parameter values approaching the upper limit of physically meaningful values (CMFCTOP=1). For J^G, the accepted parameter values vary on both sides of the default value, while for the three remaining cost functions J^{ZONAL}, J^{G+XY} and $J^{G+\text{ZONAL}}$, there is a clear tendency for parameter values smaller than the default. Overall, CMFCTOP is an example of a parameter which behaves quite in an expected way.

The MCMC chains for the parameter CPRCON (Fig. 2) behave rather differently from those for CMFCTOP (Fig. 1). Generally, the values of CPRCON are weakly bounded from above for all formulations of the objective function – sooner (J^{XY}) or later (J^G) the upper limit of the prior range of
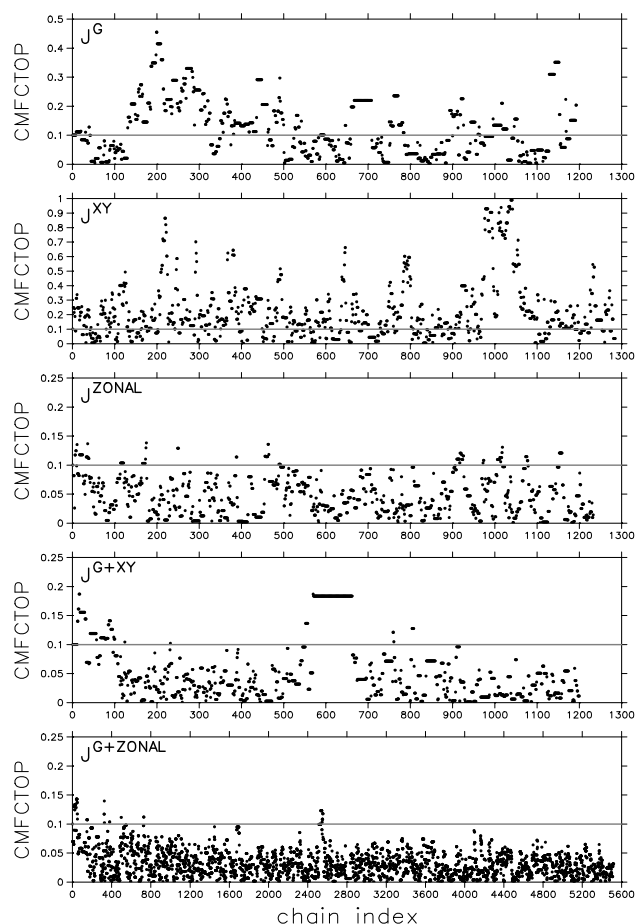
Fig. 1. The MCMC chain for parameter CMFCTOP. The horizontal grey line at parameter value 0.1 is the default value. Note the different scaling in different panels. Also note that the length of the MCMC chains is slightly larger than the number of ECHAM5 runs performed (1000 for the first four chains, 4500 for the $J^{G+\text{ZONAL}}$ chain).

Fig. 2. Same as Fig. 1 but for parameter CPRCON. The horizontal grey line at parameter value 8×10^{-4} is the default value.

parameter values is met. There seems to be a tendency towards parameter values larger than the default. Figure 2 is an example of a parameter which is weakly constrained by any of the objective functions, and the overall behaviour is not very desirable. A possible explanation is that for changes in CPRCON, the corresponding changes in longwave and shortwave fluxes at the TOA tend to cancel each other, leading to smaller changes in the TOA net flux.

In general, Figs. 1 and 2 point to the importance of the choice of the objective function. This is seen both in the fact that some objective functions constrain CMFCTOP much more tightly than others (e.g., compare $J^{G+\text{ZONAL}}$ with J^{XY}) and in the failure of *all* objective functions to constrain CPRCON properly. The latter point calls for an improved definition of the objective function in future work. Specifically for CPRCON, compensation between longwave and shortwave fluxes suggests that an objective function that
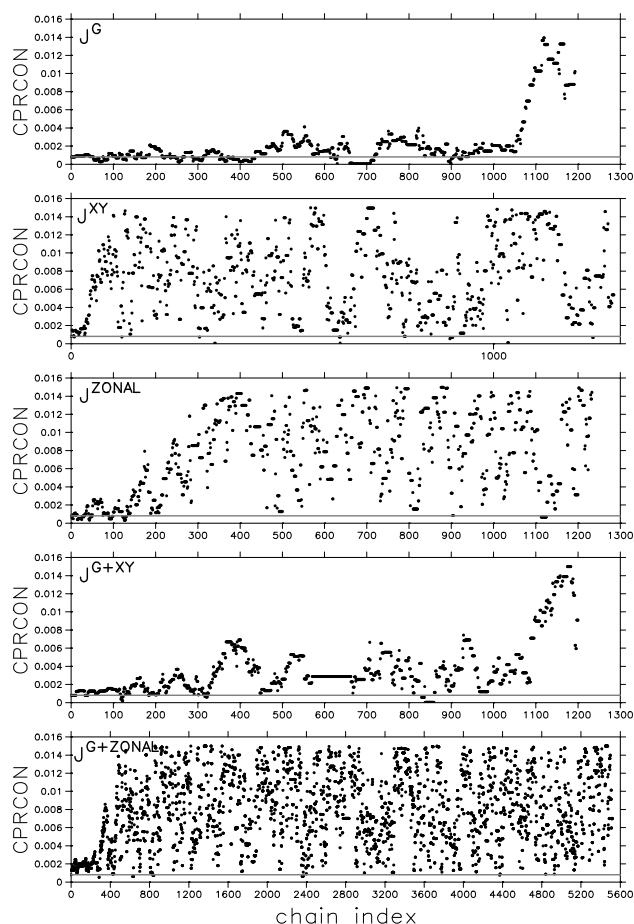
utilizes these fluxes separately, rather than only the net flux, might better constrain this parameter. Other model fields sensitive to CPRCON include precipitation rate, and middle and high cloud fractions.

Another general point evident in particular in Fig. 2 concerns the convergence of the MCMC chains. For example, when using the objective function J^G, the values of CPRCON remain relatively close to the default value nearly until the end of the chain, but then rapidly drift to a much higher level. Had we stopped the chain after (e.g.) 800 runs, we could have falsely concluded that J^G constraints CPRCON rather well. This suggests that a MCMC chain length of 1000 runs is not long enough for analyzing the posterior parameter distributions properly. The issue of chain convergence is addressed more comprehensively in the following section, based on the $J^{G+\text{ZONAL}}$ experiment.

4.2 Analysis of the $\mathbf{J}^{G+\text{ZONAL}}$ experiment

To visualize how the statistical characteristics of the MCMC parameter chains evolve in the $J^{G+\text{ZONAL}}$ experiment, we
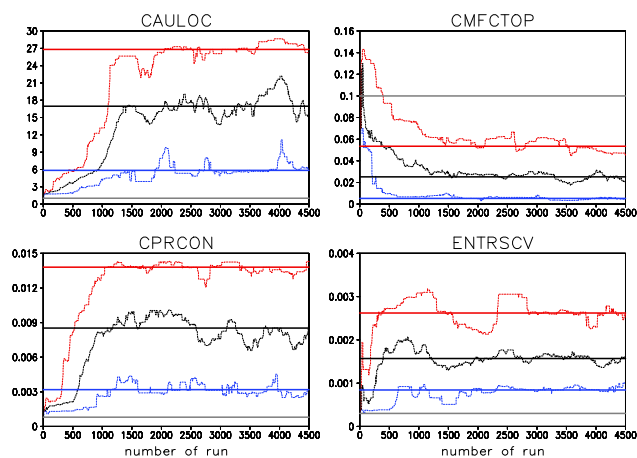
Fig. 3. A demonstration of how the MCMC posterior distribution converges with increasing chain length for the $J^{\text{G+ZONAL}}$ objective function. The dotted blue, black and red lines show the 10th, 50th and 90th percentiles of the posterior distribution for the previous 500 runs (or for all runs performed so far, for runs 1–499). The horizontal blue, black and red lines indicate the 10th, 50th and 90th percentile of the posterior distribution computed using the last 3500 runs (i.e, runs 1001–4500) of the chain. The horizontal grey lines indicate the default values of the parameters.

evaluated the posterior distributions after each run, always *using the last 500 runs only*. The blue, black and red dotted lines in Fig. 3 show the 10th, 50th and 90th percentage points computed from these distributions for each of the four parameters. The solid horizontal lines show the respective percentage points computed from the last 3500 runs (i.e., runs 1001 through 4500).

While the parameters start from their default values, all of them experience a drift in the early part of the chain, so that the distributions of CAULOC, CPRCON and ENTRSCV drift towards higher values and that of CMFCTOP towards lower values. By visual inspection of Fig. 3, the drift lasts for roughly 500–1000 runs, CPRCON and ENTRSCV stabilizing slightly earlier than CAULOC and CMFCTOP. For all parameters, percentage points computed from runs 1001–1500 are close to those computed from runs 1001–4500. This suggests that for the analysis of the parameter posterior distributions, it is sufficient to omit the first 1000 runs.

Figure 4 displays posterior distributions of the four parameters computed from runs 1001 through 4500. As shown before (Fig. 1), for CMFCTOP there is a strong preference for values smaller than default (0.1). Also, as suggested by Fig. 2, CPRCON is poorly constrained from above. In broad terms, all values are deemed equally likely by the MCMC algorithm, except that the very lowest values are ruled out. CAULOC behaves quite similarly to CPRCON. Finally, the posterior distribution for ENTRSCV features a broad maximum centred around ENTRSCV$\approx 1.5 \times 10^{-3}$. Both the lowest values and very high values of ENTRSCV are deemed

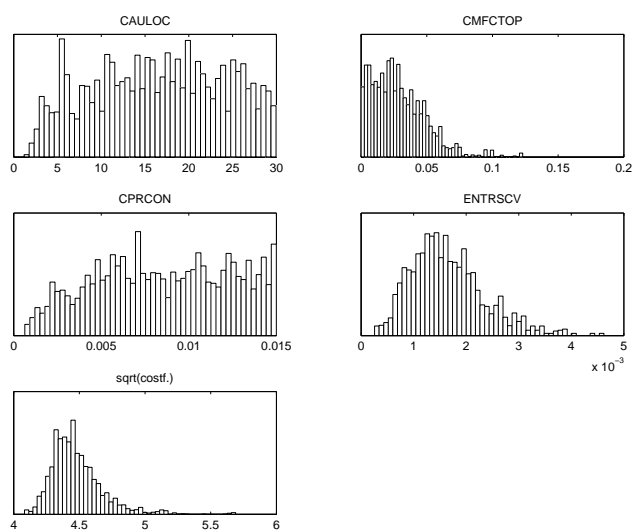**Fig. 4.** Estimated posterior distributions as chain histograms for the $J^{\text{G+ZONAL}}$ cost function. Histogram is calculated from the chain with the first 1000 values removed. The lowest panel shows the histogram of square root of values of the cost function.

unlikely. The 2-dimensional marginal posteriors (Fig. 5) do not display significantly large correlations between the parameters and further reveal the rather poor identifiability of CAULOC and CPRCON. In summary, when using the objective function $J^{\text{G+ZONAL}}$, the MCMC algorithm is able to provide *some* constraints on all four parameters considered, although for CAULOC and CPRCON, the constraints are rather weak.

4.3 Objective function versus radiative fluxes

Trivially, parameter retuning by the MCMC process can improve (i.e., decrease) the value of the objective function compared to its value for default parameter settings. A crucial question is, however, whether the MCMC process helps to reduce errors in those quantities not explicitly included in the objective function. A simple test illustrated in Fig. 6 indicates that this is, again, dependent on the choice of the objective function.

Figure 6 displays the five different objective functions versus global annual mean net, longwave (LW) and shortwave (SW) radiative fluxes at the TOA – recall that only the net flux, rather than LW and SW fluxes separately, is used in the objective functions (1)–(5). The vertical grey line represents the observed global annual mean fluxes from CERES EBAF data, and the grey dot corresponds to the default parameter values. For J^G (Fig. 6, panels a–c), the cloud of points of the MCMC chain is exactly parabolic for net radiation, as J^G penalizes of squared differences in global annual mean net radiation. The default parameter values correspond quite closely to the objective function minimum. Obviously, this has been used as a criterion in the ECHAM5 model tuning.
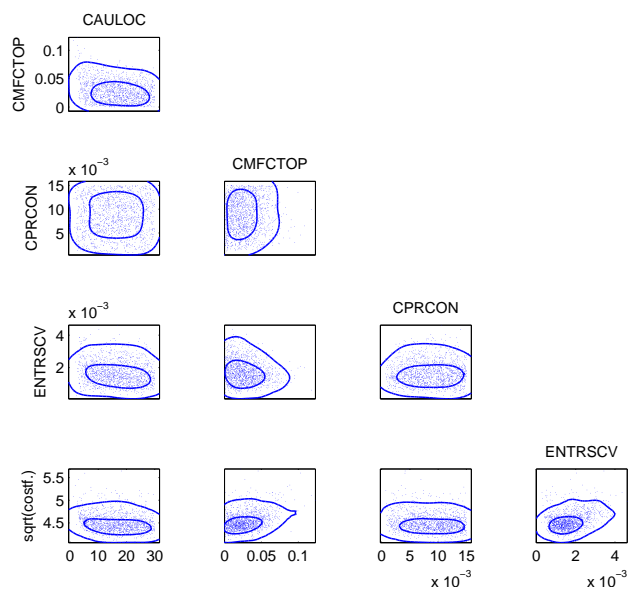
Fig. 5. Pairwise MCMC chain plots for the $J^{G+\text{ZONAL}}$ cost function. The first 1000 points are removed as burn-in time, as in the histograms of Fig. 4. The contours estimate 50% and 95% posterior probability levels of the marginal 2 dimensional distributions. Note that in addition to the four model parameters, the square root of value of the cost function is included as the fifth chain member.

For J^G, the default parameter values correspond to LW and SW biases of 7–8 Wm^{-2}. It is possible to select parameter values for an unbiased model in net radiation which correspond to LW and SW biases in the interval of about 3 to 20 Wm^{-2}, but not smaller. In particular, an overestimate of the (down-up) LW radiation at the TOA compared to CERES EBAF data seems to be an inherent bias of ECHAM5 at T21 resolution.

For J^{XY} (Fig. 6d–f), the cloud of points of the MCMC chain is diffuse and weakly parabolic for net radiation, and J^{XY} varies rather little from one MCMC step to another. There is a strong tendency for a positive net flux bias. Thus, minimization of errors in the geographical distribution of the monthly net flux is not a sufficiently strong constraint for obtaining correct global annual net flux. Note, however, that J^{XY} tends to decrease when the LW and SW biases decrease, which is a very desirable property of J^{XY}.

For J^{ZONAL} (Fig. 6g–i), the main cloud of points has a weak tendency for a positive bias in the global annual net flux, implying that J^{ZONAL} constrains somewhat better the global annual mean flux than J^{XY}. There is a very clear tendency for J^{ZONAL} to decrease when the LW and SW biases decrease. The default model is somewhat outlying in the LW/SW fluxes compared to the main cloud of points.

Next, the formulations J^{G+XY} and $J^{G+\text{ZONAL}}$, which utilize both the global annual net flux and the geographical distribution on monthly basis, are examined. The behaviour of J^{G+XY} versus net radiation is largely dominated by the
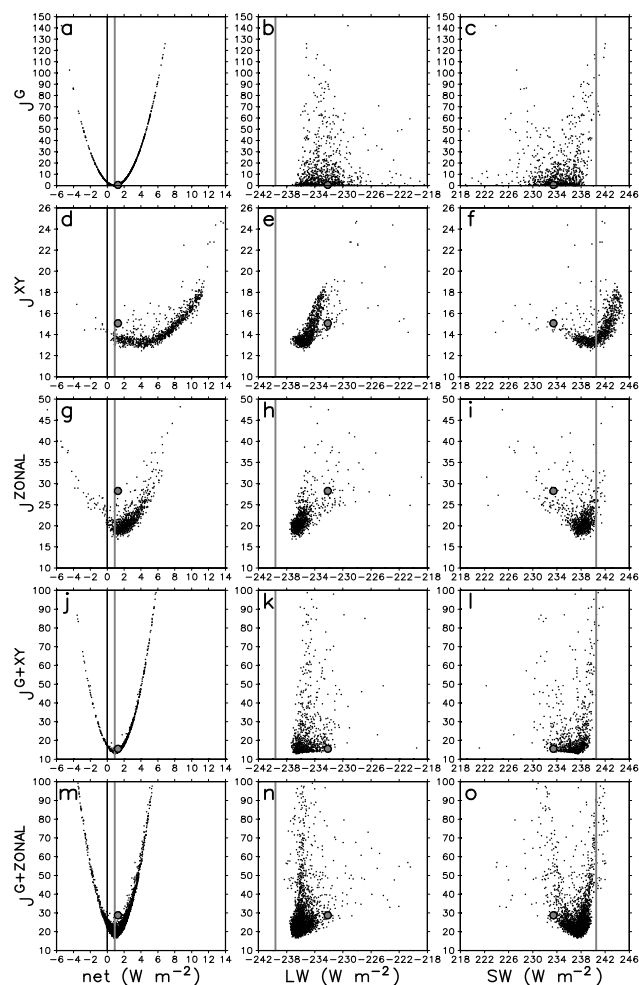

Fig. 6. The objective function versus global annual mean radiative fluxes (net, LW, and SW). The vertical grey line represents the observed global annual mean value in CERES data (0.9 Wm^{-2}, -239.6 Wm^{-2} and 240.5 Wm^{-2} for net, LW, and SW fluxes, respectively), and the grey dot corresponds to the default parameter values.

global annual mean term (Fig. 6, j–l). This is mainly because the normalizing factor σ is much smaller in Eq. (1) than in Eq. (2) (i.e., the global annual mean flux varies much less than local monthly mean values, and therefore provides a stricter constraint on the parameters). However, J^{G+XY} constrains the LW and SW parts somewhat better than J^G alone (Fig. 6a–c). Finally, the behaviour of $J^{G+\text{ZONAL}}$ versus net radiation is to some extent dominated by the global annual mean term (Fig. 6m–o), but the zonal net flux distribution makes a significant contribution. The LW and SW parts are nicely constrained such that their biases decrease as $J^{G+\text{ZONAL}}$ decreases. Overall, the behaviour of $J^{G+\text{ZONAL}}$ is probably the most attractive of the five tested objective functions. In conclusion, addition of the global annual net flux term in J^{G+XY} and $J^{G+\text{ZONAL}}$ (Fig. 6, last two rows) has the desired effect that the results are unbiased with
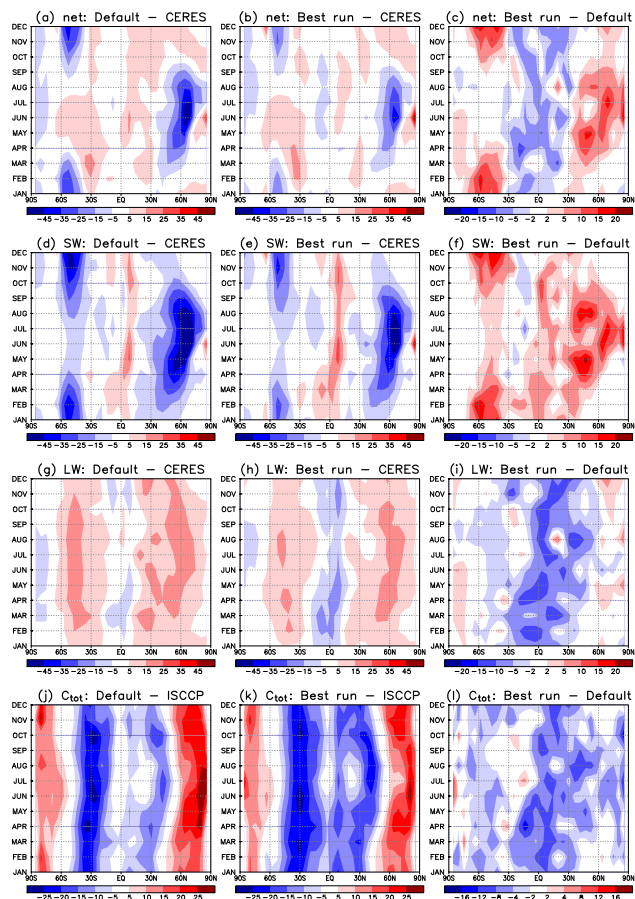
Fig. 7. Time-latitude cross section of TOA net flux difference between the default ECHAM5 and CERES observations (panel **a**), the corresponding difference for the ECHAM5 run with the smallest value of $J^{G+\text{ZONAL}}$ (panel **b**), and the difference between these two ECHAM5 runs (panel **c**; note the different scale for shading). **(d)**–**(f)**: Same as (a)–(c) but for the (down-up) shortwave flux at the TOA. **(g)**–**(i)**: Same as (a)–(c) but for the (down-up) longwave flux at the TOA. **(j)**–**(l)**: Same as (a)–(c) but for total cloud fraction (C_{tot}) compared with ISCCP satellite observations. The parameter values corresponding to default ECHAM5 are CAULOC= 1, CMFCTOP=0.1, CPRCON= 8×10^{-4}, and ENTRSCV= 3×10^{-4}; while those for the best run are CAULOC=29.83, CMFCTOP= 0.0333, CPRCON= 1.05×10^{-2}, and ENTRSCV= 1.35×10^{-3}. The corresponding values of $J^{G+\text{ZONAL}}$ are 28.7 and 16.7.

respect to the net flux and the geographical distributions are respected to some extent.

Based on Fig. 6 (and also Figs. 1 and 2), what can we conclude regarding the choice of the objective function? First, given that a reasonable simulation of the global annual-mean net flux at the TOA is necessary to avoid climate drift in coupled atmosphere-ocean GCMs, it seems prudent to include this term explicitly in the objective function (cf. Jackson et al., 2008). Second, use of the global-mean TOA radiation balance alone does not work well. It is possible to get a single number "right" with very different model climates, as

demonstrated by the wide range of global mean LW and SW radiation corresponding to low values of J^G in Fig. 6b and c. Thus, terms addressing modeled spatio-temporal structures should also be included in the objective function. In this respect, use of zonal and monthly-mean values (J^{ZONAL} and $J^{G+\text{ZONAL}}$) appears a better choice than the use of local monthly-mean values (J^G and J^{G+XY}). The reason for this is that zonal values exhibit smaller interannual "random" variations than the local values (i.e., values of $\sigma^o_{\left[\overline{\overline{F}}\right]}$ in Eq. (3) are substantially smaller than those of $\sigma^o_{\overline{\overline{F}}}$ in Eq. 2). In order for MCMC to be able to efficiently distinguish "good" from "bad" parameter combinations, the systematic impact of parameter changes needs to be relatively large compared to the random variations.

Lastly, it is by no means our purpose to imply that $J^{G+\text{ZONAL}}$ is the ultimate solution to the problem of defining the objective function. Other model fields beyond the net flux could/should be included. Also, it should be stressed that while the use of zonal means is simple, it is hardly the optimal choice for the description of spatial structures. For example, the use of empirical orthogonal functions, as in Jackson et al. (2008), is certainly an option worth considering.

4.4 Illustration of the simulation errors

Figure 7 illustrates the impact that a parameter retuning through the MCMC process has on the climate simulated by ECHAM5. Two model runs are considered: the "default run" using the default parameter setting, and the "best run" corresponding to the smallest value of the objective function $J^{G+\text{ZONAL}}$. The corresponding values of $J^{G+\text{ZONAL}}$ are 28.7 and 16.7, respectively.

For the default run, the largest net flux errors appear at high latitudes ($\sim$55° S and $\sim$60° N) during local summer, with differences of about $-40\,\text{Wm}^{-2}$ from CERES EBAF data (Fig. 7a). At lower latitudes, smaller and predominantly positive biases prevail. For the optimized closure parameters (Fig. 7b), the maximum monthly mean errors are reduced by about $10\,\text{Wm}^{-2}$. The pattern of differences between the two runs (Fig. 7c) is, for the most part, opposite to that of the original biases.

The SW and LW fluxes at the TOA are considered in Figs. 7d–f and g–i, respectively. It can be seen that the improved simulation of the net flux in the "best run" mainly results from reduced biases in SW radiation, especially at high latitudes. The impact of parameter tuning on zonal-mean longwave fluxes is, overall, neutral: the positive bias at mid- to high-latitudes is reduced, but at the same time, the negative bias in the tropics is increased. For precipitation (not shown), the impact of tuning also appeared neutral: the "default run" and the "best run" both featured essentially similar biases when compared to Climate Prediction Center's (CPC) Merged Analysis of Precipitation (CMAP) data (Xie and Arkin, 1997).

Finally, total cloud fraction is considered in Figs. 7j–l. Compared to the International Satellite Cloud Climatology Project (ISCCP) D2 data (Rossow et al., 1996; Rossow and Dueñas, 2004), the default run features too much cloudiness at high latitudes, and too little cloudiness at lower latitudes, with largest negative biases around 30°S (Fig. 7j). In the "best run", cloudiness is, overall, reduced (Fig. 7l). While this alleviates the positive bias compared to ISCCP data at high latitudes, the negative bias prevailing over most of the globe is increased (Fig. 7k). In summary, some of the quantities not included in the cost function are improved, while others are deteriorated.

5   Discussion

The MCMC approach requires long chains of model runs and is therefore best applicable to models that can be run relatively fast. In the present work, we have demonstrated (as far as we know, for the first time) that it is viable to apply MCMC to parameter estimation in an atmospheric general circulation model (GCM) used for climate simulations. This is based on three facts: the low spatial resolution of the model, application of the adaptive MCMC algorithm (DRAM), and the relatively fast response of atmospheric processes to "external forcing" (in our case, changes in parameter values). As to the limits of the approach, we note that in, e.g., ocean GCMs, the response time scales are much longer and MCMC would be computationally more demanding. Also, additional care is needed in selecting the cost function if we model systems which include important reservoirs associated with long time scales, such as carbon pools. This is the case with comprehensive Earth system models with sub-models for terrestrial biosphere and ocean biogeochemistry. One can of course estimate parameters off-line for terrestrial biosphere models (Tuomi et al., 2009), for instance, but interactions and feedbacks with the rest of the modelling systems are omitted in this procedure.

Traditional model parameter sensitivity analysis applies perturbations on model parameters, and draws conclusions about the sensitivity of model simulations on parameter values. This is typically done separately for different model parameters. This study illustrates that the range of parameter values that can produce good simulations in terms of an objective function can be much wider when more than one parameter is considered simultaneously. This is because the combined effect of two or more parameters can keep the model simulation in an acceptable region. Traditional sensitivity analysis thus makes the parameter space to appear more limited than it really is. Also, it is extremely hard to find these combined effects with traditional methods.

One issue of concern with the MCMC approach is related to error compensation. The optimal values of the closure parameters may depend on processes that these parameters do not directly influence. For example, all-sky net radiation at the TOA is a sum of clear-sky net radiation and cloud radiative forcing. Any bias in clear-sky radiative transfer calculations could influence the posterior distribution of closure parameters that affect cloudiness. The problem of error compensation is, however, not inherent to MCMC but applies to model retuning in general. Presumably the best way to mitigate this problem in the framework of MCMC is to carefully select an objective function that accounts for multiple aspects of climate. Various other fields beyond the TOA net radiation could be used, as in Jackson et al. (2008).

In this article, the definition of the objective function has been based on very simple statistics such as global and zonal mean values. More sophisticated formulations would account for observed climate phenomena, especially those associated with three-dimensional distributions and possibly including also their temporal evolution. The spatial characteristics can be captured using standard statistical techniques, such as empirical orthogonal functions. Their extensions (e.g., Ilin et al., 2006) can account for more distinctive features of the observed climate variability. Formulation of such an advanced cost function is one of the future directions of our research. Other questions that have to be addressed in the cost function formulation are, e.g., how to combine several similarity criteria in one objective function, and what is the length of climate simulation required to alleviate the effects of purely random variations in the objective function.

6   Conclusions

All general circulation models of the atmosphere or ocean – including climate models – contain closure parameters to which the model simulations are sensitive. These parameters appear in physical parameterization schemes where some unresolved variables are expressed by predefined parameters. In climate modeling, typically, best available expertise is used to define the optimal closure parameter values, based on observations, process studies, large eddy simulations, etc. This procedure has the drawback that little is learned about the parameter posterior distributions: is the optimum local or global, are parameters correlated, etc. Here, parameter estimation, based on the adaptive Markov chain Monte Carlo (MCMC) method, was applied for estimation of joint posterior probability distribution of closure parameters in the ECHAM5 climate model run at a coarse horizontal resolution T21. The four selected parameters related to clouds and precipitation were sampled by an adaptive random walk process, subject to an objective function. Five alternative formulations of the objective function were tested, all of which were related to the net radiative flux at the top of the atmosphere.

We have demonstrated the viability of MCMC methods, especially adaptive MCMC, for the objective estimation of the uncertainties related to closure parameters in climate models. For the four closure parameters, we found an

MCMC chain consisting of 4500 one-year ECHAM5 runs sufficient for a full exploration of the posterior distribution. Chains of 1000 runs were somewhat too short due to an initial drift of the parameter distributions. Even with the most promising cost function, only two parameters (CMFCTOP and ENTRSCV) were found to identify and produce reasonable tight posterior uncertainties within the predefined parameter limits (Fig. 4). However, it is a strength of the MCMC methodology that it allows one to study the identifiabilities and the correlation structures of the parameters (Fig. 5) even when the parameters do not identify. Further work is ongoing in the construction of more refined objective functions for the estimation of the closure parameters and their uncertainties.

Edited by: P. Jöckel

References

AMIP Project Office: AMIP II guidelines, AMIP Newsletter, 8, http://www-pcmdi.llnl.gov/projects/amip/NEWS/amipnl8.php, 1996.

Andrieu, C. and Moulines, E.: On the ergodicity properties of some adaptive MCMC algorithms, Ann. Appl. Prob., 16, 1462–1505, doi:10.1214/105051606000000286, 2006

Annan, J. D. and Hargreaves, J. C.: Efficient estimation and ensemble generation in climate modeling, Phil. Trans. R. Soc. A, 365, 2077–2088, doi:10.1098/rsta.2007.2067, 2007.

Dee, D. P.: Bias and data assimilation, Q. J. R. Meteorol. Soc., 131, 3323–3343, doi:10.1256/qj.05.137, 2005.

Haario, H., Saksman, E., and Tamminen, J.: An adaptive Metropolis algorithm, Bernoulli, 7, 223–242, 2001.

Haario, H., Laine, M., Lehtinen, M., Saksman, E., and Tamminen, J.: MCMC methods for high dimensional inversion in remote sensing (with discussion), J. Roy. Stat. Soc. B, 66, 591–607, 2004.

Haario, H., Saksman, E., and Tamminen, J.: Componentwise adaptation for high dimensional MCMC, Comput. Stat., 20, 265–273, doi:10.1007/BF02789703, 2005.

Haario, H., Laine, M., Mira, A., and Saksman, E.: DRAM: Efficient adaptive MCMC, Stat. Comput., 16, 339–354, doi:10.1007/s11222-006-9438-0, 2006.

Ilin, A., Valpola, H., and Oja, E.: Exploratory Analysis of Climate Data Using Source Separation Methods, Neural Networks, 19, 155–167, 2006.

Jackson, C. S., Sen, M. K., Huerta, G., Deng, Y., and Bowman, K. P.: Error Reduction and Convergence in Climate Prediction, J. Climate, 21, 6698–6709, doi:10.1175/2008JCLI2112.1, 2008.

Loeb, N. G., Wielicki, B. A., Doelling, D. R., Smith, G. L., Keyes, D. F., Kato, S., Manalo-Smith, N., and Wong, T.: Toward optimal closure of the Earth's top-of-atmosphere radiation budget, J. Climate, 22, 748–766, 2009.

Martin, G. M., Johnson, D. W., and Spice, A.: The measurement and parameterization of effective radius of droplets in warm stratocumulus, J. Atmos. Sci., 51, 1823–1842, 1994.

Mellor, G. L. and Yamada, A.: A hierarchy of turbulence closure models for planetary boundary layers, J. Atmos. Sci., 31, 1791–1806, 1974.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E.: Equation of State Calculations by Fast Computing Machines, J. Chem. Phys., 21, 1087–1092, 1953.

Navon, I. M.: Practical and theoretical aspects of adjoint parameter estimation and identifiability in meteorology and oceanography, Dynam. Atmos. Ocean., 27, 55–79, 1993.

Rinne, J. and Järvinen, H.: Estimation of the Cressman term for a barotropic model through optimization with use of the adjoint model, Mon. Weather Rev., 121, 825–833, 1993.

Robert, C. P. and Casella, G.: Monte Carlo Statistical Methods, Springer, New York, USA, second edn., 2005.

Roeckner, E., Bäuml, G., Bonaventura, L., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kirchner, I., Kornblueh, L., Manzini, E., Rhodin, A., Schlese, U., Schulzweida, U., and Tompkins, A.: The atmospheric general circulation model ECHAM5, Part I Model Description, Tech. Rep. No. 349, Max-Planck-Institut fr Meteorologie, 2003.

Roeckner, E., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kornblueh, L., Manzini, E., Schlese, U., and Schulzweida, U.: Sensitivity of simulated climate to horizontal and vertical resolution in the ECHAM5 atmosphere model, J. Climate, 19, 3771–3791, 2006.

Rossow, W. and Dueñas, E. N.: The International Satellite Cloud Climatology Project (ISCCP) web site: An online resource for research., Bull. Am. Meteorol. Soc., 85, 167–172, 2004.

Rossow, W. B., Walker, A., and Roiter, M.: International Satellite Cloud Climatology Project (ISCCP) documentation of new cloud datasets, Tech. Rep. WMO/TD-737 349, World Meteorological Organization, 1996.

Tuomi, M., Thum, T., Järvinen, H., Fronzek, S., Berg, B., Harmon, M., Trofymow, J. A., Sevanto, S., and Liski, J.: Leaf litter decomposition – Estimates of global variability based on Yasso07 model, Ecological Modelling, 220, 3362–3371, doi:10.1016/j.ecolmodel.2009.05.016, 2009.

Uppala, S. M., Kållberg, P. W., Simmons, A. J., Andrae, U., Bechtold, V. D. C., Fiorino, M., Gibson, J. K., Haseler, J., Hernandez, A., Kelly, G. A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R. P., Andersson, E., Arpe, K., Balmaseda, M. A., Beljaars, A. C. M., Berg, L. V. D., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Hólm, E., Hoskins, B. J., Isaksen, L., Janssen, P. A. E. M., Jenne, R., McNally, A. P., Mahfouf, J., Morcrette, J., Rayner, N. A., Saunders, R. W., Simon, P., Sterl, A., Trenberth, K. E., Untch, A., Vasiljevic, D., Viterbo, P., and Woollen, J.: The ERA-40 re-analysis, Q. J. Roy. Meteorol. Soc., 131, 2961–3012, doi:10.1256/qj.04.176, 2005.

Villagran, A., Huerta, G., Jackson, C. S., and Sen, M. K.: Computational Methods for Parameter Estimation in Climate Models, Bay. Anal., 3, 1–27, 2008.

Xie, P. and Arkin, P.: Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs, Bull. Am. Meteorol. Soc., 78, 2539–2558, 1997.