

## Tests of Monte Carlo Independent Column Approximation in the ECHAM5 Atmospheric GCM

P. RÄISÄNEN, S. JÄRVENOJA, AND H. JÄRVINEN

*Finnish Meteorological Institute, Helsinki, Finland*

M. GIORGETTA AND E. ROECKNER

*Max Planck Institute for Meteorology, Hamburg, Germany*

K. JYLHÄ AND K. RUOSTEENOJA

*Finnish Meteorological Institute, Helsinki, Finland*

(Manuscript received 21 June 2006, in final form 20 December 2006)

### ABSTRACT

The Monte Carlo Independent Column Approximation (McICA) method for computing domain-average radiative fluxes allows a flexible treatment of unresolved cloud structure, and it is unbiased with respect to the full ICA, but its flux estimates contain conditional random noise. Here, tests of McICA in the ECHAM5 atmospheric GCM are reported. ECHAM5 provides an interesting test bed for McICA because it carries prognostic variables for the subgrid-scale probability distribution of total water content, which allows us to determine subgrid-scale cloud variability directly from the resolved-scale model variables.

Three experiments with differing levels of radiative noise, each consisting of ten 6-yr runs, are performed to estimate the impact of McICA noise on simulated climate. In an experiment that attempted to deliberately maximize McICA noise, a systematic reduction in low cloud fraction occurred. For a more reasonable implementation of McICA, the impact of noise is very small, although statistically discernible.

In terms of the impacts of noise, McICA appears to be a viable approach for use in ECHAM5. However, to improve the simulation of cloud radiative effects, realistic representation of both unresolved and resolved cloud structures is needed, which remains a challenging problem. Comparison of ECHAM5 data with a global cloud system-resolving model dataset and with International Satellite Cloud Climatology Project data suggested two problems related to unresolved cloud structures. First, ECHAM5 appears to underestimate subgrid-scale cloud variability. This problem seems partly related to the use of the beta distribution scheme for total water content in ECHAM5: in its current form, the scheme is unable to generate highly inhomogeneous clouds (relative standard deviation of condensate amount  $>1$ ). Second, it appears that in ECHAM5, overcast cloud layers occur too frequently and partially cloudy layers too rarely. This problem is not unique to the beta distribution scheme; in fact, it is more pronounced when using an alternative, relative humidity-based cloud fraction scheme.

### 1. Introduction

Radiative processes play a central role in determining the climate of the earth, and hence, radiation calculations are among the key physical parameterizations in atmospheric general circulation models (GCMs) used in climate modeling. The radiation schemes in GCMs operate on discrete grid columns and attempt to

provide, for each column, accurate estimates of grid-mean radiative fluxes and heating rates, given the information about atmospheric and surface state provided by the host model. However, the GCMs represent the atmosphere at rather coarse horizontal resolution, a typical gridpoint spacing being 100–500 km. Grid-mean radiative fluxes and heating rates can depend strongly on variations occurring at scales smaller than GCM grid spacing, especially on unresolved cloud structure (e.g., Barker et al. 2003). Unrealistic assumptions such as neglect of subgrid-scale cloud variability can result in large biases in both cloud

---

*Corresponding author address:* Petri Räisänen, Finnish Meteorological Institute, P.O. Box 503, FI-00101 Helsinki, Finland.  
E-mail: petri.raisanen@fmi.fi

shortwave (SW) albedo (e.g., Cahalan et al. 1994) and longwave (LW) emissivity (e.g., Pomroy and Illingworth 2000).

In most current GCM radiation schemes, assumptions about unresolved cloud structure (cloud overlap between different GCM layers and subgrid-scale variability within individual layers) are embedded within the radiative transfer solver. Most often, maximum-random overlap of homogeneous clouds is assumed (cf. Barker et al. 2003). As a radical alternative, Barker et al. (2002) and Pincus et al. (2003) proposed the Monte Carlo Independent Column Approximation (McICA) for estimating domain-average radiative fluxes. McICA separates the description of cloud field structure from the radiative transfer solver by dividing the cloud field into a set of subcolumns. One or more randomly selected subcolumns are then used for each point in the spectral integration. The key features of McICA include the following: (i) it is unbiased with respect to the full ICA, regardless of the assumptions about unresolved optical structure; (ii) it allows us to keep the radiative transfer solver simple and relatively fast, independent of the cloud structure assumed, since each subcolumn is horizontally homogeneous and either cloud free or overcast; and (iii) the radiative fluxes and heating rates produced by McICA contain conditional random errors, or noise.

While the first two points listed above are clear advantages of McICA, the third is a potential problem. Should the unbiased random errors lead, through nonlinear interactions, to significant systematic effects on model climatology, this would call into question the applicability of McICA. This issue needs to be resolved through testing.

The impact of McICA noise has been already addressed in at least three global models. In the studies of Pincus et al. (2003) for the European Centre for Medium Range Weather Forecasts' (ECMWF's) global model and Pincus et al. (2006) for the Geophysical Fluid Dynamics Laboratory's (GFDL's) Atmosphere Model version 2 (AM2) climate model, the impact of noise was found to be negligible. On the contrary, Räisänen et al. (2005, hereafter RBC05) found a slight but distinct signal of McICA noise in tests made with the National Center for Atmospheric Research's (NCAR's) Community Atmosphere Model (CAM): in simulations with a high noise level, low cloud fraction was reduced over the oceans, especially in the Tropics. This feature developed very rapidly, within a couple of days, and probably resulted from short time-scale interaction between low clouds and local radiative heating. It proved feasible to essentially eliminate this feature using the noise reduction techniques devised by

Räisänen and Barker (2004). In general, the above-mentioned studies support the notion that McICA is a viable approach for introducing a flexible description of unresolved cloud structure into GCM radiation calculations. However, it is also evident that the impact of noise is model dependent.

Another important issue related to the use of McICA in GCMs is the specification of unresolved cloud structure. From the technical point of view, the subcolumns needed by McICA can be supplied by the stochastic cloud generator of Räisänen et al. (2004). However, the generator needs as input data not only standard cloud fields provided by GCMs (cloud fraction and grid-mean liquid and ice water contents) but also parameters related to subgrid-scale cloud structure: variability of condensate amount and vertical decorrelation lengths related to overlap rates for cloud fraction and condensate. For an initial implementation of McICA, a simple prescription, such as judiciously selected constant values for these parameters might suffice. However, in the end, the success of McICA in reducing biases in cloud-radiation interaction in GCMs hinges on a realistic description of the parameters. It is therefore desirable to link the unresolved cloud structure to the GCM's prognostic fields.

The current paper reports tests of McICA in the ECHAM5 atmospheric GCM developed by the Max Planck Institute for Meteorology (Roeckner et al. 2003, 2006). ECHAM5 provides an especially interesting test bed for McICA because it carries prognostic variables describing the probability density function (PDF) of total water content in each model grid box (Tompkins 2002). This allows us to derive the PDF of condensate amount needed for the cloud generator without further assumptions. Two basic questions are considered. First, what is the impact of McICA noise in ECHAM5; and second, how realistic are the subgrid-scale cloud structures generated?

Section 2 describes the features of ECHAM5 most relevant for this study. Section 3 outlines a series of ensemble simulations, which are used to evaluate the impact of McICA noise on the climate simulated by ECHAM5 in section 4. The following two sections focus on the properties of subgrid-scale cloud fields in ECHAM5, based on offline radiation calculations (section 5) and comparison of cloud field characteristics with global cloud system-resolving model (CSR) data and satellite data (section 6). The main results are summarized in section 7.

## 2. Model

Version 5.3 of the ECHAM5 atmospheric GCM (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 horizontal resolution of T42 (a grid spacing of  $\sim 2.8^\circ$ ) with 31 layers in the vertical and the model top at 10 hPa. A semi-implicit time integration scheme is used for model dynamics with a time step of 20 min. Model physical parameterizations include turbulent diffusion, cumulus convection, a stratiform cloud scheme, SW and LW radiation schemes, gravity wave drag, a land surface model, and a simple lake model (for details, see Roeckner et al. 2003). The parameterizations are invoked every time step with the exception of radiation, which is computed once in two hours. At intermediate time steps, solar radiative fluxes are adjusted for the change in solar zenith angle, without accounting for changes in the atmospheric state.

The scheme for the representation of stratiform clouds consists of prognostic equations for cloud liquid water and ice, a cloud microphysical scheme based on Lohmann and Roeckner (1996), and, as a unique feature, a statistical cloud cover scheme including prognostic equations for the distribution moments of total water mixing ratio  $q_t$  within GCM grid cells (Tompkins 2002). A beta distribution is assumed for the PDF of  $q_t$ :

$$P(q_t) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{(q_t - a)^{\alpha-1}(b - q_t)^{\beta-1}}{(b - a)^{\alpha+\beta-1}}, \quad (1)$$

$$q_t = q_v + q_l + q_i, \quad (2)$$

where  $\Gamma$  is the gamma function,  $a$  and  $b$  are the lower and upper limits of the distribution,  $\alpha$  and  $\beta$  are shape parameters, and  $q_v$ ,  $q_l$ , and  $q_i$  are mixing ratios of water vapor, cloud liquid water, and cloud ice, respectively. This form of the PDF was chosen by Tompkins as a pragmatic compromise; it is relatively simple and flexible, and provides, in most cases, a reasonable fit to cloud-resolving model simulations of  $q_t$ . Tompkins (2002) mainly considered the use of this approach for the diagnosis of cloud fraction, and it is used solely for this purpose in the standard version of ECHAM5. However, as detailed below, the PDF of  $q_t$  can also be utilized to determine the subgrid-scale distribution of condensate amount.

Four parameters ( $a$ ,  $b$ ,  $\alpha$ , and  $\beta$ ) define the beta distribution. In the current implementation of the Tompkins scheme in ECHAM5, parameter  $\alpha$  is fixed at  $\alpha = 2$  while  $\beta$  is a prognostic variable, constrained to the range  $2 \leq \beta \leq 50$ . This implies that the distribution of total water content is unimodal (“bell shaped”) and either symmetrical or positively skewed. Two conditions are required to determine the bounds  $a$  and  $b$ . In partially cloudy cases, they can be deduced diagnosti-

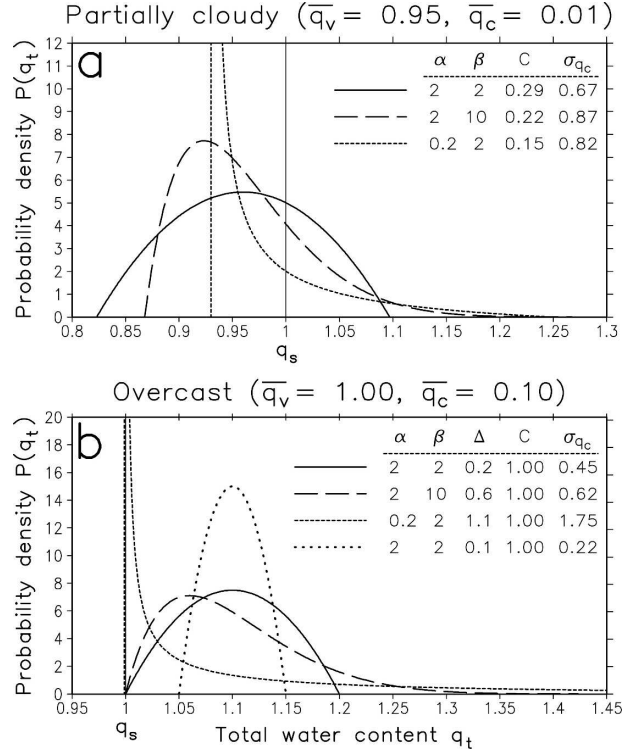


FIG. 1. Examples of  $\beta$  distributions of total water content  $q_t$  and the related cloud parameters. Cloud fraction  $C$  and subgrid-scale cloud variability (characterized here by the relative std dev of condensate amount  $\sigma_{q_c}$ ) are diagnosed using grid-mean specific humidity  $\bar{q}_v$  and cloud condensate content  $\bar{q}_c$  (prognostic variables in ECHAM5), shape parameters  $\alpha$  (prescribed) and  $\beta$  (prognostic), and the width of the distribution  $\Delta = b - a$  (prognostic, but only used in overcast cases). The units are chosen so that the saturated specific humidity  $q_s = 1$ ; hence  $q_t > 1$  indicates cloudy regions. (a) Partially cloudy cases with  $\bar{q}_v = 0.95$  and  $\bar{q}_c = 0.01$ , (b) overcast cases with  $\bar{q}_v = 1.00$  and  $\bar{q}_c = 0.10$  ( $\Delta$  chosen so that the lower limit is  $a = q_s = 1$ , except in the last case). Note that only the cases with  $\alpha = 2$  can actually occur in ECHAM5. See section 6a for more discussion of  $\sigma_{q_c}$ .

cally from the grid-mean values of total water mixing ratio  $q_t$  and condensate mixing ratio  $q_c = q_l + q_i$ , assuming that the saturation mixing ratio  $q_s$  is constant within the grid box. However, to close the system in cloud-free and overcast cases (and only then), a prognostic equation is used to calculate the distribution width  $b - a$  (see Tompkins 2002 for details). Examples of beta distributions of  $q_t$  and the related cloud parameters are displayed in Fig. 1.

ECHAM5 does not explicitly define convective cloud fraction or condensate amount. Rather, the cloud fraction, liquid water, and ice fields in ECHAM5 only represent stratiform clouds, including anvil clouds formed by detrainment of condensate in the upper part of convective updrafts.

The ECHAM5 SW radiation scheme is based on

Fouquart and Bonnel (1980), while the LW calculations use the Rapid Radiative Transfer Model (RRTM; Mlawer et al. 1997). As usual, molecular scattering, gaseous absorption, and effects of clouds and aerosols are taken into account in the calculations. An important detail for the current work is spectral resolution. The SW (LW) region is divided into 4 (16) bands. Furthermore, the LW scheme makes use of the correlated- $k$  distribution method (e.g., Lacis and Oinas 1991), in which the bands are subdivided into cumulative probability space (CPS) intervals according to the strength of gaseous absorption. The total number of CPS intervals in the LW region is 140. There is no such subdivision in the SW calculations, which use the photon path-length distribution method (Bakan and Quenzel 1976; Fouquart and Bonnel 1980) to handle the interaction between gaseous absorption and multiple scattering.

#### *Treatment of subgrid-scale cloud structure*

The McICA experiments employed a version of the radiative transfer code that operates on horizontally homogeneous, overcast or cloud-free subcolumns. The subcolumns were produced using the stochastic cloud generator of Räisänen et al. (2004).

The generator first computes the profile of cloud fraction (0 or 1) for each subcolumn. The state of cloud overlap between two layers  $i$  and  $j$  is described in terms of an overlap parameter  $\alpha_{i,j} \in [0, 1]$ , where  $\alpha_{i,j} = 1$  ( $\alpha_{i,j} = 0$ ) corresponds to maximum (random) overlap. Analyses of observations by Hogan and Illingworth (2000) and Mace and Benson-Troth (2002) suggest that a reasonable formulation of  $\alpha_{i,j}$  is

$$\alpha_{i,j} = \exp \left[ - \int_{z_i}^{z_j} \frac{dz}{L_{cf}(z)} \right], \quad (3)$$

where  $z$  is the altitude. As in RBC05, a globally constant value of 2 km is assumed for the decorrelation depth  $L_{cf}$ .

Second, the subgrid-scale distribution of cloud condensate is generated. The subgrid-scale variability of condensate amount in each layer is derived directly from the beta distribution of total water content. For each cloudy subcell, the generator produces a cumulative probability  $y_c$  defined by

$$y_c = \int_0^{q_c} P_c(q'_c) dq'_c, \quad (4)$$

where  $P_c(q_c)$  is the PDF of cloud condensate amount ( $q_c = q_t + q_i = q_t - q_s$ , when  $q_v = q_s$ ) across the cloudy part of the layer. Since cloudy subcells correspond to the high end of the PDF of total water content,  $y_c$  can be related to the cumulative probability for total water

$$y_t = I(q_t) = \int_a^{q_t} P(q'_t) dq'_t \quad (5)$$

through

$$y_t = (1 - C) + Cy_c, \quad (6)$$

where  $C$  is the cloud fraction in this layer. Hence, the condensate amount in subgrid-scale cloudy cells is obtained as

$$q_c = I^{-1}(y_t) - q_s. \quad (7)$$

The values of  $I^{-1}$  are tabulated as a function of  $y_t$  and the shape parameter  $\beta$ .

The vertical profiles of  $y_c$  depend on the overlap assumption for cloud condensate, which is described here by another decorrelation length  $L_{cw}$ . As in RBC05, a globally constant value  $L_{cw} = 1$  km is used. The use of a smaller decorrelation length for condensate than cloud fraction is supported by the observational analysis of Hogan and Illingworth (2003) and analysis of CSR datasets by Räisänen et al. (2004) and Pincus et al. (2005).

Finally, effective radius of liquid [ $r_{e,l}$  ( $\mu\text{m}$ )] and ice particles [ $r_{e,i}$  ( $\mu\text{m}$ )] is determined for each cloudy subcell following Martin et al. (1994) and Moss et al. (1996):

$$r_{e,l} = 1000 \left( \frac{3\text{LWC}}{4\pi\rho_w k N_d} \right)^{1/3}, \quad (8)$$

$$r_{e,i} = 83.8 \times \text{IWC}^{0.216}, \quad (9)$$

where  $\rho_w = 1000$  kg m<sup>-3</sup> is the density of liquid water,  $k = 0.67$  (0.80) over land (ocean) areas, LWC (g m<sup>-3</sup>) and IWC (g m<sup>-3</sup>) are liquid and ice water content, and  $N_d$  is droplet number concentration, which is prescribed to vary from 50 cm<sup>-3</sup> in the upper troposphere to 220 (80) cm<sup>-3</sup> near the surface over land (ocean). Lower and upper limits of 4 and 24  $\mu\text{m}$  (10 and 150  $\mu\text{m}$ ) are assumed for  $r_{e,l}$  ( $r_{e,i}$ ).

### 3. Ensemble simulations

To test the impact of McICA noise on climate simulated by ECHAM5, three ensemble simulations were performed. All ensembles consisted of 10 members, each of which was initialized with a unique seed number for the random number generator used by the stochastic cloud generator. The simulations were run for six years, using observed distributions of sea surface temperature and sea ice for January 1985–December 1990. The analysis presented below uses data for the last five years (1986–90). While the primary impact of McICA noise appeared within the first few days of the

simulations in the study of RBC05, and indeed also here (see below), relatively long simulations were performed to detect (or rule out) longer time-scale feedbacks related to the land surface and snow cover, and to improve the statistical robustness of the results.

The only difference between the three ensembles lies in the different implementations of the McICA approach, which result in different levels of sampling noise. The whole spectrum of options surveyed by RBC05 is not considered here. Instead, results are presented for an approach with excessive sampling noise (1COL), for one reasonable GCM implementation of McICA (CLDS), and for a reference simulation with low noise level but high computational costs (REF).

#### a. Experiment 1COL

In this experiment, broadband radiative fluxes are computed as

$$\hat{F} = \sum_{k=1}^{\mathcal{K}} \mathcal{F}(s, k), \quad (10)$$

where  $s$  is a *single* randomly generated subcolumn that is used for all  $\mathcal{K}$  spectral intervals. This is by no means an optimal implementation of McICA. The sole motivation for this experiment is to test the impact of large sampling noise on climate simulated by ECHAM5. However, an essentially similar approach has been employed in the Goddard Institute for Space Studies (GISS) GCM (Hansen et al. 1983; Stubenrauch et al. 1997).

#### b. Experiment CLDS

CLDS is an abbreviation for the CLDSAMPL approach of Räisänen and Barker (2004). Broadband fluxes are computed as

$$\begin{aligned} \hat{F} &= (1 - \hat{C}_{\text{tot}}) \sum_{k=1}^{\mathcal{K}} \mathcal{F}(s_{\text{clr}}, k) + \hat{C}_{\text{tot}} \sum_{k=1}^{\mathcal{K}} \mathcal{F}(s_{\text{cl},k}, k) \\ &= (1 - \hat{C}_{\text{tot}}) F_{\text{clr}} + \hat{C}_{\text{tot}} \hat{F}_{\text{cl},d}, \end{aligned} \quad (11)$$

where  $\hat{C}_{\text{tot}}$  is an estimate of total cloud fraction  $C_{\text{tot}}$ ,  $s_{\text{clr}}$  is the clear-sky column,  $s_{\text{cl},k}$  are randomly selected cloudy subcolumns, and  $F_{\text{clr}}$  and  $\hat{F}_{\text{cl},d}$  are clear-sky and cloudy-sky broadband fluxes. Thus random sampling is confined to the cloudy part of the GCM column. Here,  $\hat{C}_{\text{tot}}$  cannot be computed analytically when using the decorrelation length formulation for cloud overlap. Hence,  $\hat{C}_{\text{tot}}$  is estimated by producing first  $j' = 100$  cloud fraction profiles using only the cloud occurrence portion of the generator (see Räisänen and Barker 2004).

There is one difference to the implementation of CLDSAMPL in Räisänen and Barker (2004) and in RBC05. For the ECHAM5 LW scheme, the number of terms (i.e., CPS points for gaseous absorption) on the rhs of (11) is quite high ( $\mathcal{K} = 140$ ). Generating condensate amounts, effective radii, and cloud optical properties separately for each of them would be overkill, since many of the terms  $k$  contribute very little to cloud radiative effects. Therefore, the procedure was streamlined so that a single subcolumn is often used for many terms  $k$  in the LW region. Thus, only  $J_{\text{cl},d} = 50$  cloudy subcolumns are generated instead of  $J_{\text{cl},d} = 140$ .

#### c. Experiment REF

In the reference simulation, broadband fluxes are computed as

$$\begin{aligned} \hat{F} &= (1 - \hat{C}_{\text{tot}}) \sum_{k=1}^{\mathcal{K}} \mathcal{F}(S_{\text{clr}}, k) \\ &\quad + \hat{C}_{\text{tot}} \sum_{k=1}^{\mathcal{K}} \left[ \frac{1}{N_k} \sum_{n=1}^{N_k} \mathcal{F}(s_{\text{cl},n,k}, k) \right] \\ &= (1 - \hat{C}_{\text{tot}}) F_{\text{clr}} + \hat{C}_{\text{tot}} \hat{F}_{\text{cl},d}. \end{aligned} \quad (12)$$

Here,  $N_k \geq 1$  randomly selected cloudy subcolumns  $s_{\text{cl},n,k}$  are used for the  $k$ th interval. Selective spectral sampling is utilized so that  $N_k$  is large for those spectral intervals that contribute strongly to cloud radiative effects. It is demonstrated in section 5d of Räisänen and Barker (2004) that this approach reduces the sampling noise more efficiently than the use of the full ICA. The total number of terms ( $\sum_{k=1}^{\mathcal{K}} N_k$ ) equals 100 (574) for the SW (LW) region, and  $j' = 1000$  subcolumns are used to estimate  $\hat{C}_{\text{tot}}$ .

#### d. Quantification of sampling noise

The sampling noise present in the three McICA ensemble simulations was estimated in offline radiation calculations using the ECHAM5 dataset described in section 5a. Table 1 reports global mean standard deviations for LW, SW, and total (=LW+SW) net fluxes at the surface and for atmospheric radiative heating rates, computed from 30 realizations (i.e., 30 calls of the cloud generator and McICA) for each GCM column. The differences in radiative noise between the three ensembles are large. For the total net flux at the surface and for the total radiative heating rate, the sampling errors for CLDS are about 5 times larger than those for REF but 2–3 times smaller than those for 1COL.

Another point worth noting is that much of the sampling noise is associated with the SW calculations. For example, for CLDS (which is the most relevant of these

TABLE 1. Annual global-mean sampling error statistics for the three McICA ensemble simulations, estimated from offline radiation calculations for the ECHAM5 dataset described in section 5a. The first three rows list mean std dev for LW, SW, and total (TOT = SW + LW) net (down-up) radiative fluxes ( $\text{W m}^{-2}$ ) at the surface. The last three rows give mass-weighted std dev for SW, LW, and total atmospheric heating rates ( $\text{K day}^{-1}$ ).

Quantity	1COL	CLDS	REF
$\sigma(F)_{\text{SW}}$	57.3	25.0	4.8
$\sigma(F)_{\text{LW}}$	13.7	2.2	0.6
$\sigma(F)_{\text{TOT}}$	52.2	25.0	4.8
$\sigma(\partial T/\partial t)_{\text{SW}}$	0.490	0.324	0.059
$\sigma(\partial T/\partial t)_{\text{LW}}$	1.274	0.258	0.064
$\sigma(\partial T/\partial t)_{\text{TOT}}$	1.162	0.401	0.084

experiments for GCM use), the SW radiative flux errors exceed the LW flux errors by an order of magnitude, and even the SW heating rate errors are slightly larger than those in the LW region. The relatively large SW sampling errors stem from the use of the photon pathlength distribution method, in which the spectral bands are not divided into subintervals as in the currently more widely used  $k$  distribution methods. Thus, the number of terms in (11) for the ECHAM5 SW scheme is as low as  $\mathcal{X} = 4$ . In principle, the use of a  $k$  distribution model, such as the RRTM SW scheme (Iacono et al. 2001), would be more advantageous for McICA.

Figure 2 shows that the zonal-mean sampling errors in total net fluxes are largest in the Tropics, and that the errors in total heating rate have a strong maximum in the lowest layer. The latter feature stems at least partly

from high vertical resolution near the surface, the thickness of the lowest layer being  $\sim 65$  m.

On a somewhat philosophical note, it may be argued that rms values for the real world could, in theory, also be included in Table 1 and Fig. 2. Namely, even if we had an optimal parameterization for the decorrelation lengths (instead of our use of constant values) that gives *unbiased* estimates for all combinations of the resolved-scale model variables, such a parameterization would still have *random errors* when compared to the real world. The point is that it is principally impossible to describe the subgrid-scale structure perfectly by using resolved-scale variables only. Consequently, grid-mean radiative fluxes and heating rates should, in principle, not depend deterministically on the resolved-scale variables; rather, they should exhibit some random variations. Of course, it is unlikely that McICA noise would represent the nature of such variations correctly.

#### 4. Impact of McICA noise on simulated climate

Table 2 lists global 5-yr (1986–90) mean values of several variables for the REF simulation, together with mean differences from REF in the CLDS and 1COL simulations. The cloud radiative effect (CRE; also known as cloud radiative forcing) is defined as the difference between all-sky and clear-sky net (down-up) fluxes as

$$\text{CRE} = F_{\text{all-sky}}^{\text{net}} - F_{\text{clear-sky}}^{\text{net}}. \quad (13)$$

The statistical significance of the CLDS – REF and 1COL – REF differences is also indicated in Table 2, based on a two-sided  $t$  test.

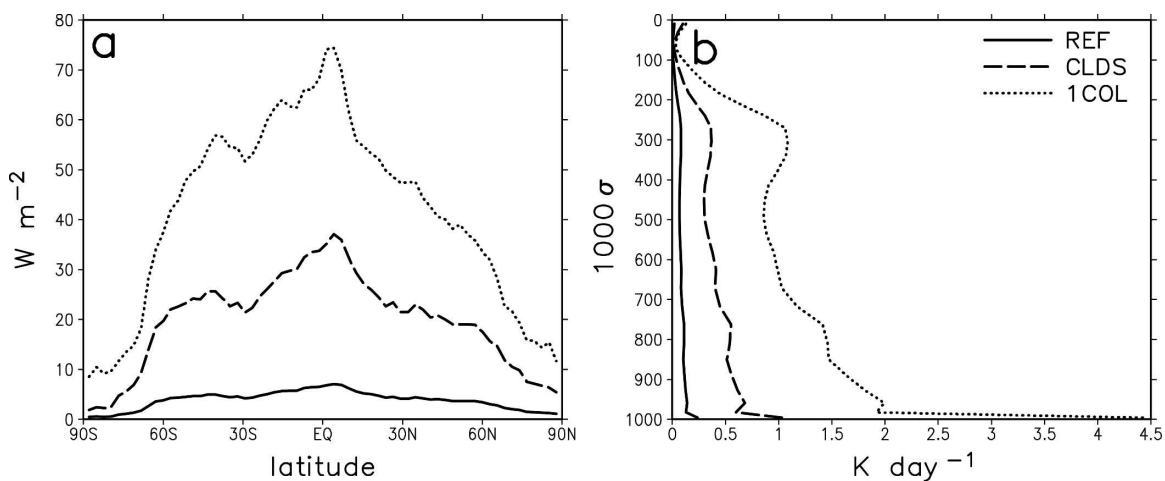


FIG. 2. (a) Zonal-mean std dev of total (SW+LW) net flux at the surface and (b) global-mean vertical profiles of std dev of total radiative heating rate for the REF, CLDS, and 1COL approaches. The results have been computed using the ECHAM5 dataset described in section 5a. They provide a measure for the magnitude of random errors in radiative fluxes and heating rates present in the GCM simulations for individual time steps and GCM columns.

TABLE 2. Summary of selected global-mean climatic parameters for the years 1986–90. The column labeled REF lists values for the reference simulation, and the last two columns give the differences from REF for the CLDS and 1COL simulations. Bold font indicates differences that are statistically significant at the 99.9% level according to a two-sided  $t$  test, while italic font stands for 95% statistical significance. High, middle, and low clouds refer to the areal coverage of clouds above 450 hPa, between 450 and 740 hPa, and below 740 hPa, respectively.

Quantity	REF	$\Delta$ CLDS	$\Delta$ 1COL
SW CRE TOA ( $W m^{-2}$ )	-58.24	<b>0.20</b>	<b>1.56</b>
SW CRE SFC ( $W m^{-2}$ )	-62.42	<b>0.24</b>	<b>1.82</b>
LW CRE TOA ( $W m^{-2}$ )	29.77	-0.03	<b>-0.17</b>
LW CRE SFC ( $W m^{-2}$ )	29.12	<b>-0.10</b>	<b>-0.93</b>
Clouds, high	0.415	-0.000	0.000
Clouds, middle	0.216	-0.000	<b>-0.001</b>
Clouds, low	0.364	<b>-0.001</b>	<b>-0.013</b>
Clouds, total	0.656	<i>-0.001</i>	<b>-0.009</b>
LWP ( $g m^{-2}$ )	62.05	<i>-0.20</i>	<b>-2.27</b>
IWP ( $g m^{-2}$ )	25.65	-0.02	<b>-0.17</b>
$T_2$ m (K)	287.74	0.00	<b>0.05</b>
$T_{300 \text{ hPa}}$ (K)	231.82	0.01	0.01
Precipitation ( $mm day^{-1}$ )	2.934	-0.001	<b>-0.010</b>
PW ( $kg m^{-2}$ )	25.90	<i>0.04</i>	<i>0.04</i>

TOA = top of the atmosphere

SFC = surface

LWP = liquid water path

IWP = ice water path

$T_2$  m = air temperature 2 m above surface

$T_{300 \text{ hPa}}$  = temperature at 300 hPa

Precipitation = convective + stratiform

PW = precipitable water

Clearly, the larger McICA noise in the CLDS and 1COL simulations as compared to REF has some statistically discernible effects on the climate simulated by ECHAM5. In many respects, the results resemble those for NCAR CAM in RBC05. The primary impact of McICA noise is a reduction in low cloud fraction and cloud liquid water path (LWP), and consequently, in SW CRE [both at the surface and at the top of the atmosphere (TOA)] and LW CRE at the surface. Quantitatively, the impact may be considered moderately large for 1COL (e.g., nearly 4% reduction in LWP and over  $1.5 W m^{-2}$  in SW CRE). For the CLDS simulation, the impact of McICA noise is much smaller and, although statistically significant, probably unimportant for practical purposes.

Figure 3 shows that the largest reductions in cloud fraction due to McICA noise occur in the boundary layer ( $\sigma = p/p_s \geq 0.9$ ), with smaller reductions extending higher up, especially at middle and high latitudes. As expected, the differences between 1COL and REF are much larger than those between CLDS and REF. In general, the 1COL – REF differences in low cloud fraction were more pronounced over oceans than over land

(global-mean differences being  $-0.016$  and  $-0.007$ , respectively). The largest absolute differences ( $-0.05$  to  $-0.09$ ) occurred in the stratocumulus areas off the coasts of California, Peru, and Namibia. However, overall, the reduction in cloud fraction due to McICA noise is distributed more uniformly for ECHAM5 than for NCAR CAM, which showed a substantial reduction in the lowest model layer over equatorial oceans and much less elsewhere (see RBC05).

Similar to the results for NCAR CAM, the negative cloud fraction differences between 1COL and REF developed very rapidly, within a couple of days from the start of the simulations (not shown). Moreover, just as for NCAR CAM, the reduction in cloud fraction was accompanied by reduced radiative cooling, increased temperature, and reduced relative humidity (RH) in the lower troposphere. These features were not as striking as those in Fig. 10 of RBC05, but (as noted above for cloud fraction) they were more widespread both horizontally and vertically. Most probably, the physical mechanism behind the reduction in low cloud fraction is basically the same for ECHAM5 as for NCAR CAM. These clouds are partly maintained by local radiative cooling induced by the clouds themselves; McICA noise disrupts the cooling; and a positive feedback loop is induced, in which reductions in low cloud fraction lead to reduced radiative cooling and further reductions in cloud fraction.

Yet this is only a partial explanation. McICA noise can naturally also enhance the local instantaneous radiative cooling, which could, conceivably, lead to increased cloudiness in another positive feedback loop. It remains unclear why the feedback loop leading to reduced low cloud fraction dominates.

Figure 4 displays zonal-mean temperature differences to REF for the CLDS and 1COL simulations. Even the 1COL – REF differences are small ( $\leq 0.35$  K everywhere), partly because the use of prescribed sea surface temperatures also constrains atmospheric temperatures. Nonetheless, some statistically robust features appear. The largest differences (1COL warmer than REF) occur in the lower troposphere at high northern latitudes. In addition, the tropical tropopause region is slightly warmer for 1COL than for REF, while the stratosphere is slightly colder.

The positive temperature differences between 1COL and REF at high northern latitudes resulted mainly from higher summer temperatures in the 1COL simulation over the northern parts of North America and Eurasia (Fig. 5a). This feature was associated with reductions in low cloud fraction and enhanced net SW fluxes at the TOA and at the surface (Figs. 5b–c). The 1COL simulation also features higher zonal land area

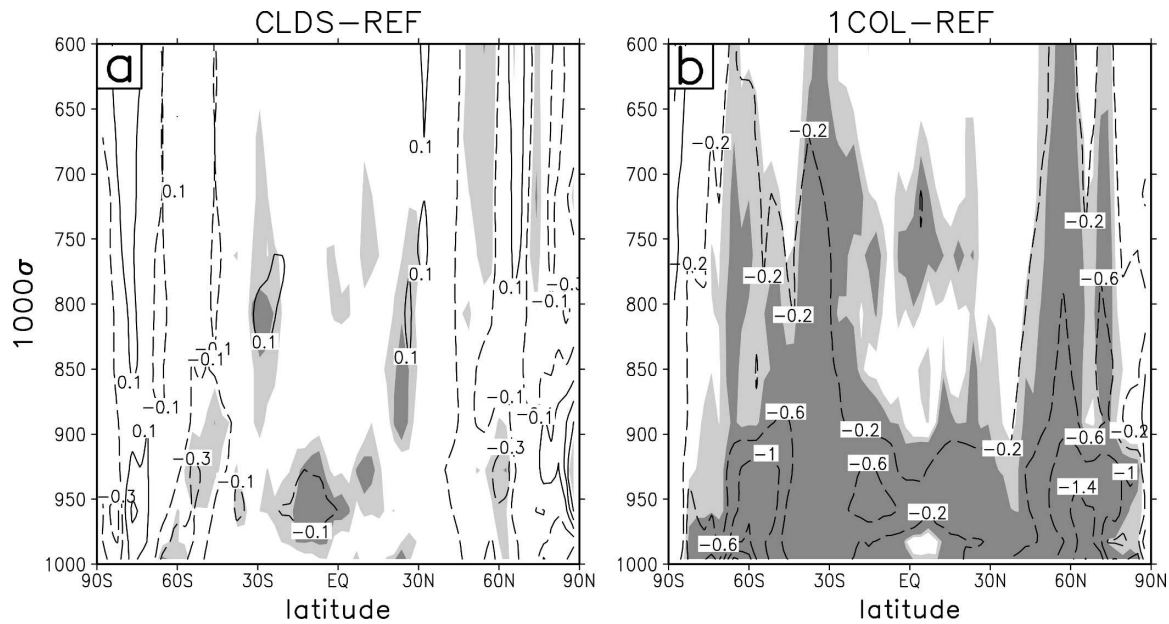


FIG. 3. Differences in zonal-mean cloud fraction (percentage) between (a) the CLDS and REF simulations and (b) the 1COL and REF simulations in the lower and middle troposphere up to  $\sigma = p/p_s = 0.6$ . Light (dark) shading indicates differences that are statistically significant at the 90% (99%) level. Contour intervals are 0.2% in (a) and 0.4% in (b).

mean temperatures than REF during spring, especially between 45° and 60°N. This mainly results from higher temperatures over the Russian territory. There, the springtime warming was associated with a slightly earlier snowmelt, which can be seen as a slight reduction in surface albedo that migrates toward the north when the

spring advances in Fig. 5d. It is not clear whether the higher temperatures are more a consequence of the earlier snowmelt, or the reason for it. An interesting question is whether the large random errors in surface SW fluxes inherent to the 1COL approach contribute to this feature. Since changes in snow cover involve a posi-

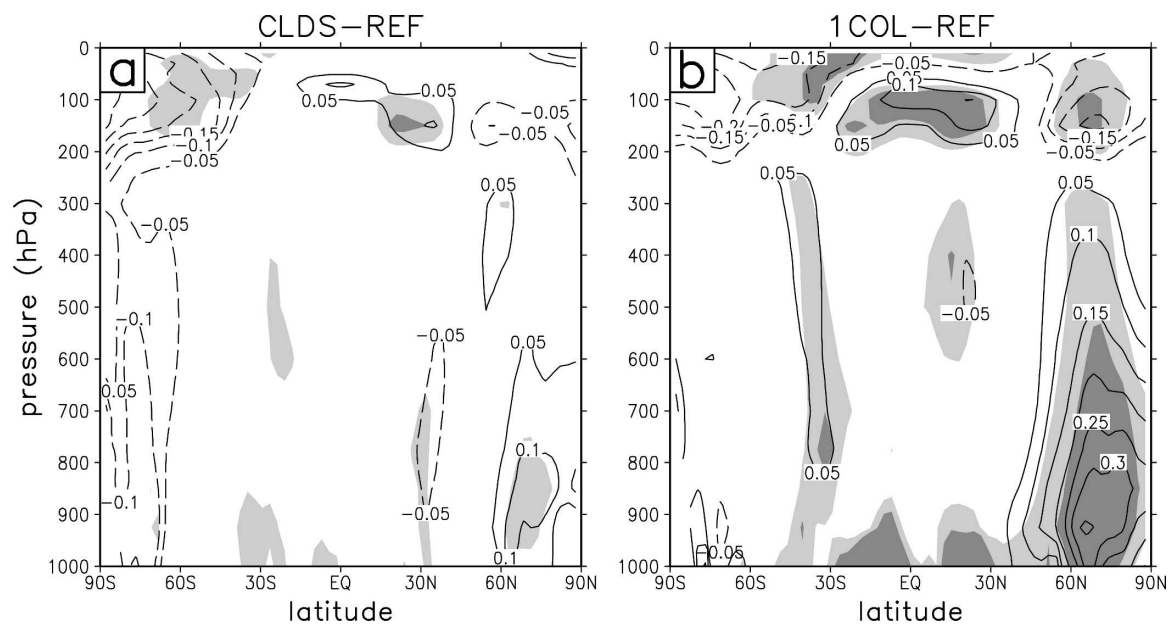


FIG. 4. Zonal-mean temperature differences between (a) the CLDS and REF simulations and (b) the 1COL and REF simulations. Light (dark) shading indicates differences that are statistically significant at the 90% (99%) level. Contour intervals are 0.05 K; zero contours are omitted.



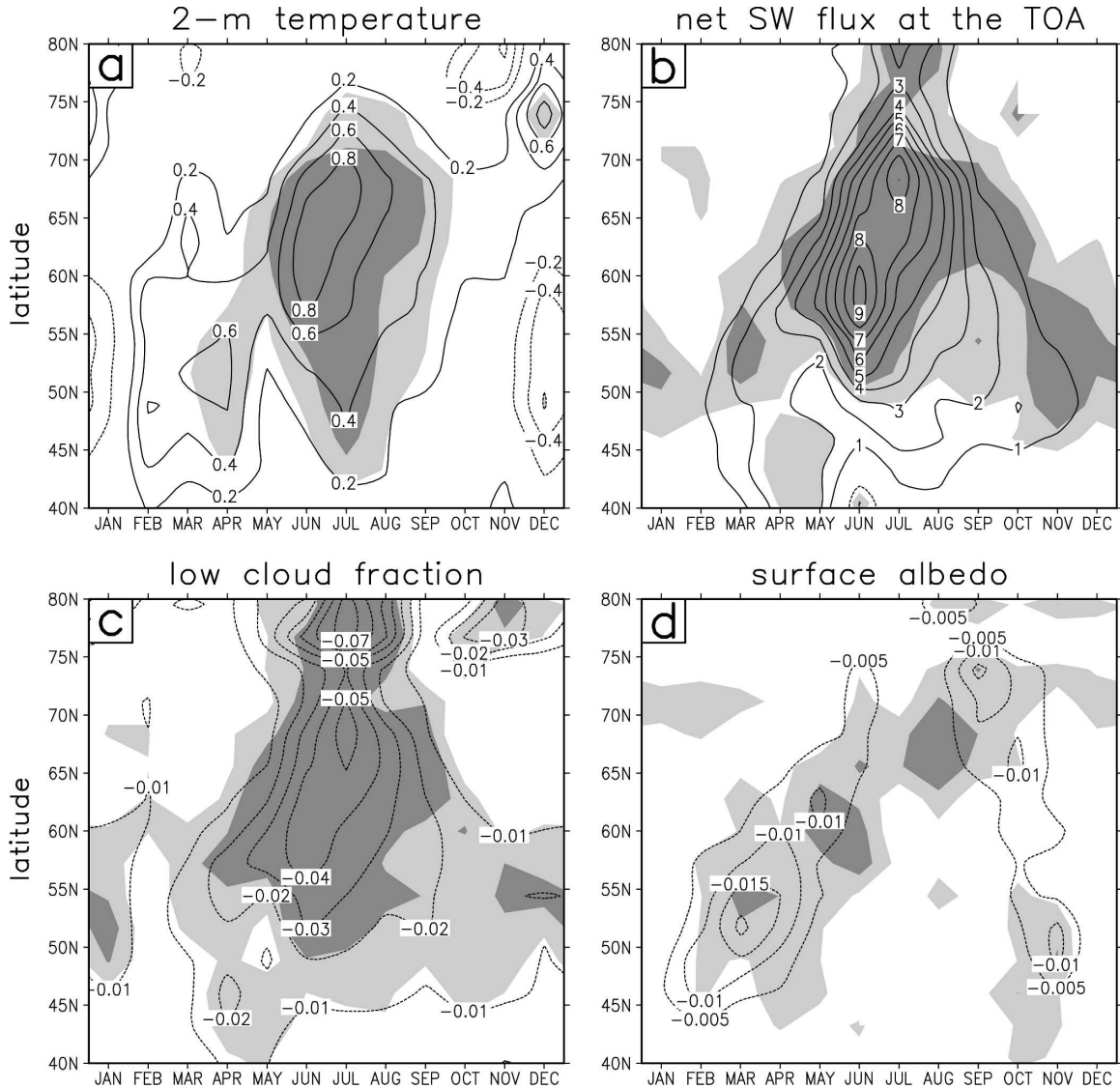


FIG. 5. Mean annual cycle of differences between the 1COL and REF simulations for land grid points between  $40^{\circ}$  and  $80^{\circ}$ N. (a) Two-meter air temperature (K), (b) net SW flux at the top of the atmosphere ( $\text{W m}^{-2}$ ), (c) low cloud fraction, and (d) surface albedo. Light (dark) shading indicates differences that are statistically significant at the 90% (99%) level. Zero contours are omitted.

tive surface albedo feedback, it may be speculated that it would be most likely in this kind of situations, if any, that random errors in surface fluxes could influence the model behavior.

To summarize the key findings of this section, ECHAM5 does show some sensitivity to large random noise in radiation calculations. However, for a reasonable implementation of McICA, the impact of noise seems small enough to allow the use of McICA in ECHAM5.

## 5. Offline radiation calculations

In order for McICA to improve the simulation of cloud radiative effects in a GCM, a reasonable repre-

sentation of unresolved cloud structures is necessary. Thus, we now turn our attention to unresolved cloud features in ECHAM5. It is first investigated, by means of offline calculations, how sensitive the radiative fluxes for ECHAM5 are to assumptions made about unresolved clouds. To better interpret the findings of this exercise, analysis of cloud field characteristics follows in section 6.

### a. Datasets

The primary dataset used here was extracted from one of the CLDS simulations. The input data needed for the radiative transfer scheme and for the stochastic

TABLE 3. Global-mean CREs ( $\text{W m}^{-2}$ ) in offline radiation calculations. SWTOA (SWSFC) and LWTOA (LWSFC) are shortwave and longwave CRE at the top of the atmosphere (at the surface). ECHAM5 refers to results obtained using the ECHAM5 dataset and the McICA version of the ECHAM5 radiation code, while CSRm refers to results for the global cloud system-resolving model dataset.

Experiment	SWTOA	SWSFC	LWTOA	LWSFC
Impact of subgrid-scale cloud variability				
ECHAM5	-58.34	-62.51	29.89	29.24
ECHAM5_PPH	-60.54	-65.01	30.59	29.83
ECHAM5 - ECHAM5_PPH	2.20	2.49	-0.70	-0.60
CSRm - CSRm_PPH	5.48	5.90	-3.64	-1.35
Impact of cloud overlap assumptions				
ECHAM5_MAX	-56.58	-60.51	29.23	28.62
ECHAM5_RAN	-60.48	-64.92	30.45	30.44
ECHAM5_RAN - ECHAM5_MAX	-3.90	-4.41	1.22	1.82
CSRm_RAN - CSRm_MAX	-11.65	-12.95	4.55	3.84

PPH = plane-parallel homogeneous clouds

MAX = maximum overlap

RAN = random overlap

cloud generator were saved for every radiative time step (i.e., at 2-h intervals) for the 15th day of each month for the year 1986. Altogether, data for nearly 1.2 million GCM columns are included.

For comparison, some results are also shown for the global “super-parameterization” dataset used by Räisänen et al. (2004) and Barker and Räisänen (2005). In this dataset, subgrid-scale cloud structure has been simulated explicitly using a 2D cloud system-resolving model (horizontal grid spacing of  $\sim 4$  km) within each GCM column. Therefore, this dataset may provide a more realistic description of subgrid-scale cloud structure than the data extracted from ECHAM5, although it can by no means be regarded as an absolute benchmark. The dataset has the same horizontal resolution as used in the ECHAM5 runs (T42), while the vertical resolution is slightly lower (26 layers), and the data represent a single day only (1 January).

For both datasets (denoted below as the ECHAM5 and CSRm datasets, respectively), ECHAM5 parameterizations for cloud particle size [Eqs. (8) and (9)] and cloud optical properties were adopted in the calculations.

### b. Results

A series of calculations was performed using the ECHAM5 dataset with different assumptions about unresolved cloud structure. Table 3 summarizes global-mean values of CRE for these experiments; in addition, some results based on the CSRm dataset are given.

The first experiment (denoted simply as ECHAM5 in Table 3) used the same parameterizations for cloud overlap [Eq. (3)] and subgrid-scale cloud variability (based on the beta distribution for total water content)

as used in the actual GCM runs. The resulting global-mean CREs are within  $\approx 0.3 \text{ W m}^{-2}$  of the 5-yr ensemble-mean values for the CLDS experiment (in Table 2). In the second experiment (ECHAM5\_PPH), mean values of liquid and ice water content were used throughout the cloudy part of each GCM grid box. The difference between ECHAM5 and ECHAM5\_PPH thus gives the impact of subgrid-scale cloud variability. Compared to the corresponding results for the CSRm dataset (CSRm-CSRm\_PPH), this impact is rather small for ECHAM5. For example, inclusion of subgrid-scale cloud variability weakens the global-mean SW CRE at the TOA only by  $2.2 \text{ W m}^{-2}$  for ECHAM5, as compared to  $5.5 \text{ W m}^{-2}$  for the CSRm dataset.

The lower part of Table 3 addresses the impact of cloud overlap assumptions. ECHAM5\_MAX and ECHAM5\_RAN provide results for the extreme cases of pure maximum overlap (decorrelation lengths  $L_{\text{cf}} = L_{\text{cw}} = \infty$  in the cloud generator) and pure random overlap ( $L_{\text{cf}} = L_{\text{cw}} = 0$ ). As expected, random overlap provides larger values of CRE, for example, by  $3.9 \text{ W m}^{-2}$  for SW CRE at the TOA. However, this is only one-third of the respective difference for the CSRm dataset (CSRm\_RAN-CSRm\_MAX). Similarly, the total cloud fraction  $C_{\text{tot}}$  is quite insensitive to the overlap assumption for the ECHAM5 dataset. Going through the whole range from maximum to random overlap increases  $C_{\text{tot}}$  only by 0.023 (from 0.659 to 0.682), as compared with 0.090 for the CSRm dataset (from 0.596 to 0.686). Recently, Morcrette and Jakob (2000) reported a corresponding difference of 0.056 in  $C_{\text{tot}}$  for the ECMWF model at a horizontal resolution of T42 with 31 layers in the vertical, with slightly larger differences at higher vertical and horizontal resolution.

To the limited extent that the CSRM data can be considered a benchmark, the results suggest that ECHAM5 may underestimate the impact of subgrid-scale cloud variability and cloud overlap assumptions. These issues are discussed further in section 6.

Finally, a short comment on the global-mean values of CRE at the TOA is worth making. While the LW CRE is compatible with the observed value of  $\sim 30 \text{ W m}^{-2}$  (e.g., Kiehl and Trenberth 1997) in all experiments with ECHAM5 data, the global-mean SW CRE ( $-56.6$  to  $-60.5 \text{ W m}^{-2}$ ) is significantly larger than observed ( $\sim -50 \text{ W m}^{-2}$ ). The SW CRE would be overestimated even if the impact of subgrid-scale variability were as large for ECHAM5 as for the CSRM dataset. This serves as a healthy reminder of the fact that biases in model-simulated CRE are not only caused by deficiencies in unresolved cloud structure, but can also depend strongly on errors in grid-mean cloud quantities such as cloud fraction, condensate amount, and effective particle size. Clearly, validation of grid-mean cloud properties in GCMs is a vitally important task, which should proceed hand-in-hand with the improvement and validation of subgrid-scale cloud properties. For a recent comparison of GCM-simulated cloudiness with satellite data, see Zhang et al. (2005).

## 6. Analysis of unresolved cloud structures in ECHAM5

### a. Subgrid-scale cloud variability

Analysis of cloud fields revealed a straightforward reason for the finding that subgrid-scale cloud variability has a substantially smaller radiative impact for ECHAM5 than for the CSRM dataset: the clouds are, in general, less inhomogeneous for ECHAM5. A convenient measure of cloud variability is the relative standard deviation  $\sigma_{q_c}$  (i.e., standard deviation divided by mean) of condensate amount within the cloudy parts of GCM grid boxes. For the ECHAM5 dataset, highest zonal-mean values of  $\sigma_{q_c}$  occur in the tropical lower to middle troposphere ( $\sigma_{q_c} \sim 0.7\text{--}0.8$ ); elsewhere, the typical values are  $\sigma_{q_c} \sim 0.5$ . For the CSRM dataset,  $\sigma_{q_c}$  varies more strongly and is generally higher, in particular, in the tropical middle and upper troposphere where zonal-mean values of  $\sigma_{q_c}$  are  $\sim 1\text{--}1.5$  (Fig. 6b in Barker and Räisänen 2005). There is, however, a caveat in the comparison of the ECHAM5 and CSRM results. As noted in section 2, the condensate fields in ECHAM5 do not include the core parts (i.e., updraft regions) of convective clouds. These regions, which generally have high local condensate amounts, are included in the CSRM fields.

Owing to the above-mentioned discrepancy, it can be

argued that values of  $\sigma_{q_c}$  indeed should be smaller for ECHAM5 than for the CSRM dataset, and that the same applies to the radiative impacts of subgrid-scale cloud variability. At the same time, it is important to note a principal limitation related to the current form of the beta distribution approach used for total water content  $q_t$  in ECHAM5: it is simply impossible to produce cloud layers with as strong subgrid-scale variability as seen in the CSRM data.

It can be shown that when a beta distribution (1) is assumed for  $q_t$ ,  $\sigma_{q_c}$  is determined uniquely by three parameters:  $\alpha$ ,  $\beta$ , and  $(q_s - a)/(b - a)$ . Alternatively, for a given cloud fraction  $C$ ,  $\sigma_{q_c}$  is a unique function of  $\alpha$  and  $\beta$  (except for one special case mentioned in the caption of Fig. 6). Figure 6 shows that  $\sigma_{q_c}$  decreases (increases) with increasing  $\alpha$  (increasing  $\beta$ ), and that the values are larger for the case with low cloud fraction ( $C=0.2$ ) than for the overcast case ( $C=1.0$ ). However, throughout the bell-shaped regime of the beta distribution ( $\alpha > 1$ ;  $\beta > 1$ ) considered in Fig. 6,  $\sigma_{q_c} < 1$ . Values of  $\sigma_{q_c} > 1$  could be attained if  $\alpha < 1$  were used in Eq. (1) (see Fig. 1b for an example). However, it seems unlikely that this would, in general, provide a good approximation for the PDF of total water content [in particular, in this case  $P(q_t) \rightarrow \infty$  when  $q_t \rightarrow a$ ]. This suggests that a more elaborate approach (e.g., a bimodal or multimodal distribution of  $q_t$ , or a separate parameterization for the PDF of condensate amount in the cloudy part of the layer) may be needed for a realistic description of subgrid-scale cloud variability and its radiative effects, especially in convective regions.

Another viewpoint on cloud variability is obtained by considering the parameter  $\epsilon$  (Rossow et al. 2002), which is a measure of subgrid-scale horizontal variations in total column optical thickness  $\tau$ . By definition,

$$\epsilon = 1 - \hat{\tau}/\bar{\tau}, \quad (14)$$

where  $\bar{\tau}$  is the linear-mean value of  $\tau$  for the visible spectral band for cloudy subcolumns within a GCM column,

$$\bar{\tau} = \frac{1}{N} \sum_{i=1}^N \tau_i, \quad (15)$$

and  $\hat{\tau}$  is the corresponding radiative mean value, that is, the optical thickness of a homogeneous cloud that would have the same mean reflectance  $R$  as the actual inhomogeneous cloud field:

$$R(\hat{\tau}) = \frac{1}{N} \sum_{i=1}^N R(\tau_i). \quad (16)$$

Figure 7 compares values of  $\epsilon$  for ECHAM5 data to those derived from International Satellite Cloud Clima-

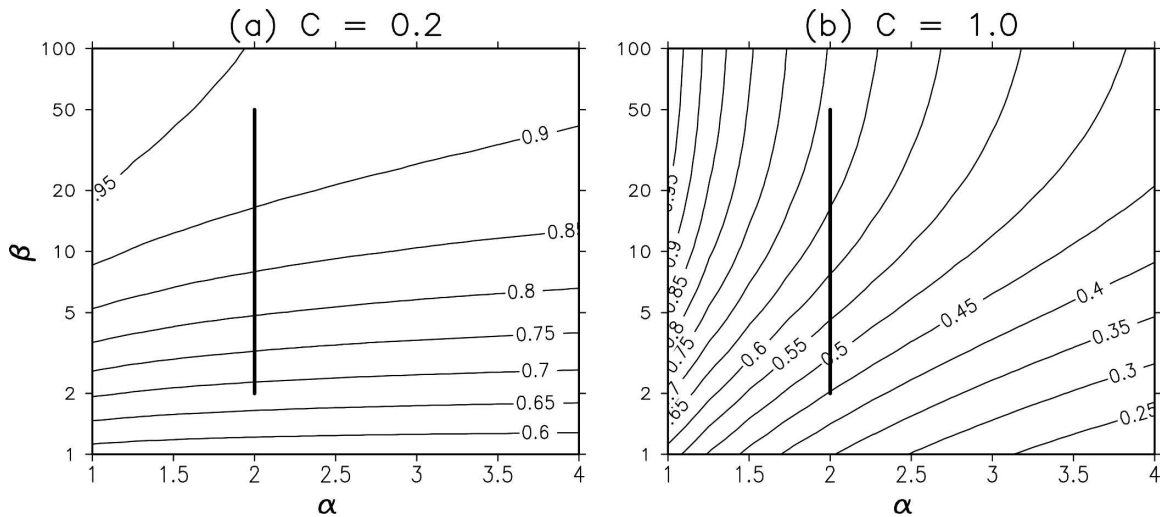


FIG. 6. Relative std dev of cloud condensate amount  $\sigma_{d_c}$  as a function of  $\beta$  distribution shape parameters  $\alpha$  and  $\beta$  for cloud fraction of (a)  $C = 0.2$  and (b)  $C = 1.0$ . The thick vertical lines indicate the parameter range considered in ECHAM5 ( $\alpha = 2$ ,  $\beta = 2 \dots 50$ ). Note that in the overcast case, the results also depend on the lower limit,  $a$ , of the distribution of total water content. Here,  $a = q_s$ , which means that the minimum value of condensate amount is zero. For  $a > q_s$ ,  $\sigma_{d_c}$  would be smaller; an example of this is given in Fig. 1b.

tology Project (ISCCP) D1 data (downloaded from <http://isccp.giss.nasa.gov/products/browsemesos.html>). As discussed by Zhang et al. (2005), the comparison of model results with ISCCP data is influenced by the choice of the lower limit of the optical thickness of model clouds included. Here, a cutoff value of  $\tau_{\min} = 0.3$  was applied to the subcolumns in the ECHAM5 cloud fields, following the current specifications of the ISCCP simulator (Klein and Jakob 1999; Webb et al. 2001). The ISCCP results are 5-yr (1986–90) mean values and represent cloud variations occurring at scales below 280 km, which is comparable with the T42 horizontal resolution of the ECHAM5 and CSRM datasets. Note that inhomogeneity is resolution dependent and generally increases with increasing gridbox size (Rossow et al. 2002; Hogan and Illingworth 2003). Because the ECHAM5 dataset only covers 12 separate days, the results are noisier than those for ISCCP. The ECHAM5 dataset reproduces qualitatively some features seen in ISCCP data, such as relatively high values of  $\epsilon$  along the ITCZ, and relatively low values in the stratocumulus areas west of California, Peru, and Namibia. Quantitatively, however, ECHAM5 underestimates  $\epsilon$  compared to ISCCP data nearly everywhere.

The lower values of  $\epsilon$  for ECHAM5 indicate that *in a vertically integrated sense*, the ECHAM5 cloud fields are less variable than the ISCCP observations suggest. This is consistent with the notion that ECHAM5 may underestimate cloud subgrid-scale variability *within individual GCM layers*, but it provides only supporting,

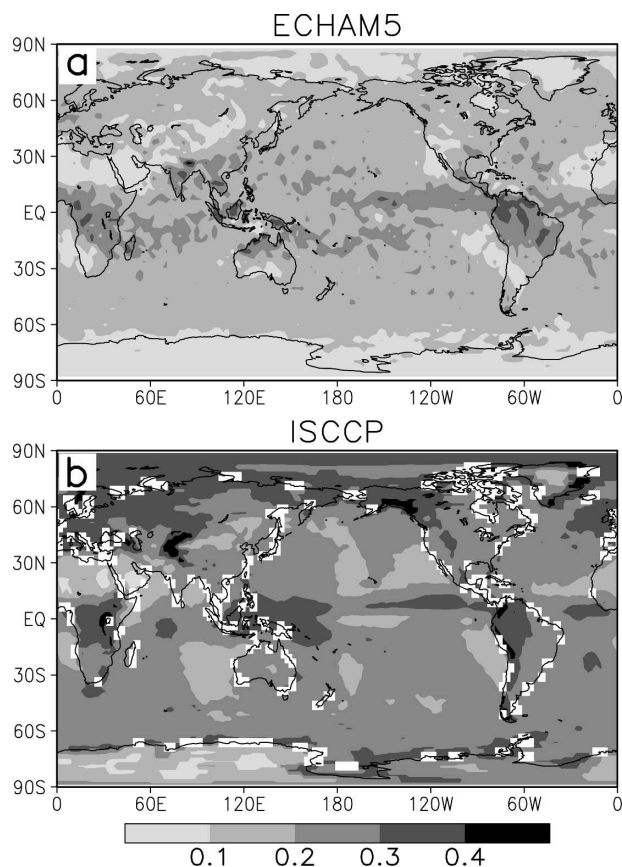


FIG. 7. Cloud variability parameter  $\epsilon$  for (a) the ECHAM5 dataset and (b) ISCCP data. The white spots along coastlines in (b) indicate missing values in the ISCCP data.

not conclusive evidence. Values of  $\epsilon$  can also depend substantially on the vertical structure of clouds.

*b. Why is the impact of cloud overlap assumptions small for ECHAM5?*

This subsection seeks to shed more light on the finding that cloud overlap assumptions have a relatively small impact for ECHAM5. Toward this end, frequency distributions of layer cloud fraction  $C$  and total cloud fraction  $C_{\text{tot}}$  are considered. In addition to the ECHAM5 dataset based on the Tompkins (2002) cloud fraction scheme, it is instructive to consider the results for an alternative dataset denoted below as ECHAM5\_RH. Like the ECHAM5 dataset, this dataset contains data for the 15th day of each month for a single year, but it was extracted from a run using a RH-based cloud fraction scheme available as an option in ECHAM5. Following Sundqvist et al. (1989) and Lohmann and Roeckner (1996), cloud fraction is parameterized as

$$C = 1 - \sqrt{1 - \frac{\text{RH} - \text{RH}_c}{1 - \text{RH}_c}}, \quad (17)$$

where the critical relative humidity  $\text{RH}_c$  is usually defined as a function of the  $\sigma$  coordinate:

$$\text{RH}_c = 0.7 + 0.2 \exp(1 - \sigma^{-4}). \quad (18)$$

As an exception, lower  $\text{RH}_c$  is used in subsidence inversion layers over oceans to improve the simulation of marine stratocumulus clouds.

Figure 8 compares the frequency distribution of  $C$  for the three datasets. For the ECHAM5 dataset, overcast layers are much more frequent than for the CSRSM dataset, while layers with partial cloudiness ( $0 < C < 1$ ) are generally less frequent. For ECHAM5\_RH, these tendencies are even more distinct. In particular, layers with small but nonzero cloud fraction ( $0 < C \leq 0.3$ ) occur rarely in the ECHAM5\_RH dataset. For ECHAM5\_RH, more than half of all cloudy layers are overcast (area-weighted fraction: 54.3%), the corresponding figures being 37.9% for the ECHAM5 dataset and only 15.6% for the CSRSM dataset. These differences become even more striking when considering that the overlap assumption can affect  $C_{\text{tot}}$  only for those GCM columns with (i) at least two partially cloudy layers and (ii) no overcast layers. The fraction of such columns is 51.8% for the CSRSM dataset, as compared to 19.6% for ECHAM5 and only 7.4% for ECHAM5\_RH. Consequently, for ECHAM5\_RH, the cloud overlap assumption matters even less than for ECHAM5, the difference in global-mean  $C_{\text{tot}}$  between

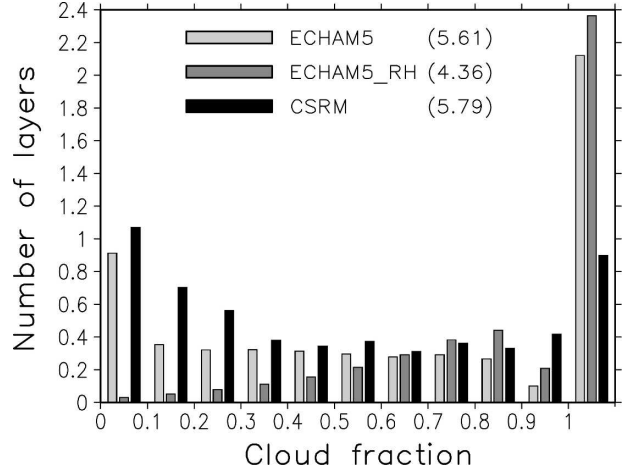


FIG. 8. Global-mean frequency distribution of layer cloud fraction  $C$  for the ECHAM5, ECHAM5\_RH, and CSRSM datasets, plotted as the average number of layers per GCM column with  $C$  in different cloud fraction bins. The first 10 bins give the number of layers with  $0 < C < 0.1$ ,  $0.1 < C < 0.2$ , ...,  $0.9 < C < 1$ , while the last bin represents overcast cloud layers ( $C = 1$ ). Note that cloud-free layers are not included, and that the ECHAM5 and ECHAM5\_RH datasets have a vertical resolution of 31 layers, as compared to 26 layers for the CSRSM dataset. The average number of cloudy layers per GCM column is given in parentheses in the legend. The results have been weighted by GCM grid square area.

random overlap (0.677) and maximum overlap (0.666) being only 0.011.

Figure 9 compares the frequency distribution of  $C_{\text{tot}}$  in the three model datasets to that reported in ISCCP D2 satellite data. Two sets of model results are shown. In Fig. 9a, all model clouds are included (as in Fig. 8), whereas in Fig. 9b, subcolumns with visible optical thickness below  $\tau_{\text{min}} = 0.3$  are omitted, following the specifications of the ISCCP simulator. Screening out the optically thinnest clouds reduces the global-mean  $C_{\text{tot}}$  considerably, by  $\approx 0.10$  for all three datasets. However, the general features of the frequency distributions remain mostly the same. Both for the ECHAM5 and ECHAM5\_RH datasets, cases with  $0 \leq C_{\text{tot}} < 0.1$  and  $C_{\text{tot}} > 0.9$  are more frequent than in the ISCCP data, while cases with  $0.1 < C_{\text{tot}} < 0.9$  are less frequent. For the CSRSM dataset, the global-mean frequency distribution of  $C_{\text{tot}}$  agrees better, although by no means perfectly, with the ISCCP data.

Both Fig. 8 and Fig. 9 suggest that for ECHAM5, overcast cloud layers occur too frequently, while partially cloudy layers occur too rarely. Moreover, this feature is not unique to the Tompkins (2002) scheme. It is, in fact, more pronounced, when the alternative RH-based cloud fraction scheme is employed. Some understanding of this issue can be obtained by considering statistics of cloud fraction as a function of RH.

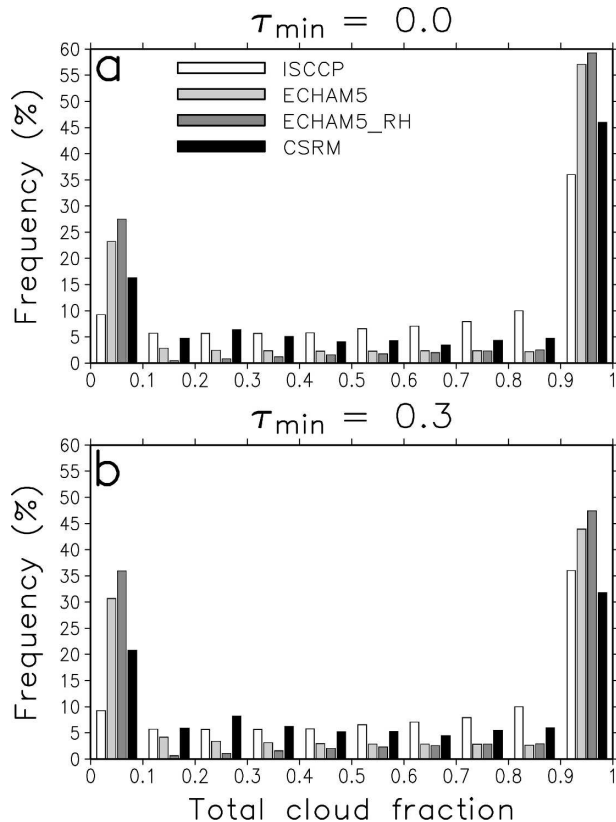


FIG. 9. Global-mean frequency distribution of total cloud fraction  $C_{\text{tot}}$  for ISCCP D2 satellite data, and for the ECHAM5, ECHAM5\_RH, and CSRM model datasets. The bin width is  $\Delta C_{\text{tot}} = 0.1$ . Panel (a) represents the case in which all clouds are included in the model datasets, while in (b), subcolumns with total column optical thickness below  $\tau_{\text{min}} = 0.3$  have been screened out from the model datasets. The results have been weighted by grid square area. Note that unlike Fig. 8, cloud-free cases are included in the lowest bin.

Figure 10 displays, as an example, global-mean values of selected quantities for the ECHAM5 and ECHAM5\_RH datasets at model level 26 (average pressure  $p = 838$  hPa;  $\sigma \approx 0.85$ ), stratified according to RH. At this particular level, the global-mean cloud fraction is slightly larger for ECHAM5\_RH (0.143) than ECHAM5 (0.121), and this holds for almost all values of RH in Fig. 10b. However, Figs. 10c and 10d reveal a distinct difference in the distribution of values of  $C$  for the two datasets. Cases with cloudiness ( $C > 0$ ) are actually more common for the ECHAM5 dataset, but when clouds exist, the mean cloud fraction is substantially larger for the ECHAM5\_RH dataset. Note that even though the cloud fraction scheme used for ECHAM5\_RH [i.e., Eq. (17)] allows  $C > 0$  at this model level for grid-mean RHs down to  $\text{RH}_c \approx 78\%$ ,  $C$  is set to zero whenever the prognostic liquid and ice water contents are zero. For the ECHAM5\_RH

dataset, this is true for most cases with  $\text{RH} \leq 96\%$ , and for the ECHAM5 dataset, for most cases with  $\text{RH} \leq 93\%$ . This difference likely stems from the fact that a change from the Tompkins (2002) scheme to the RH-based cloud fraction scheme [Eq. (17)] also involves changes in the calculation of deposition/sublimation rates of cloud ice and condensation/evaporation rates of liquid water in ECHAM5].

It may also be noted from Fig. 10d that for cases with  $C > 0$  (equivalently, cases with condensate content  $q_c > 0$ ),  $C$  is almost a single-valued function of RH for the ECHAM5\_RH dataset, which, of course, follows from Eqs. (17) and (18). The ECHAM5 dataset based on the Tompkins (2002) scheme shows a wider and presumably more realistic variation of  $C$  for each RH bin.

An important point that Fig. 10 illustrates is that the distribution of layer cloud fraction  $C$  not only depends on the formulation used to diagnose  $C$  (e.g., the explicitly defined or implicitly assumed PDF of total water content), but also on the distribution of simulated grid-mean RH and condensate amounts. For example, the existence of partial cloudiness requires that  $q_c > 0$  and  $\text{RH} < 100\%$ . Thus the relative lack of partially cloudy cases for ECHAM5 and especially ECHAM5\_RH could be related to the inability of the cloud scheme to form and sustain nonzero condensate amounts at grid-mean RHs substantially below 100%. Similarly, the apparent overabundance of overcast cases suggests that grid-mean RH of 100% may occur too often. These statements should be taken as tentative. Rigorous validation of them would be challenging, as it would require reliable observational information on both RH and cloud fraction at GCM grid scales. Neither would it be trivial to pinpoint the specific aspects of model formulation that are responsible for the features seen.

## 7. Summary

This study consisted of two main parts. First, a set of multiyear ensemble simulations was performed with the ECHAM5 atmospheric GCM to evaluate the impact of noise in radiative fluxes and heating rates that is inherent to the Monte Carlo Independent Column Approximation. Second, the radiative impacts of cloud overlap assumptions and subgrid-scale cloud variability were evaluated for a dataset extracted from one of the ECHAM5 simulations, followed by further analysis of unresolved cloud structures to understand the results. The use of a beta distribution scheme for the subgrid-scale variations of total water content in ECHAM5 lends special importance to the latter exercise, since it allows us to derive subgrid-scale cloud variability directly from the resolved-scale model variables.

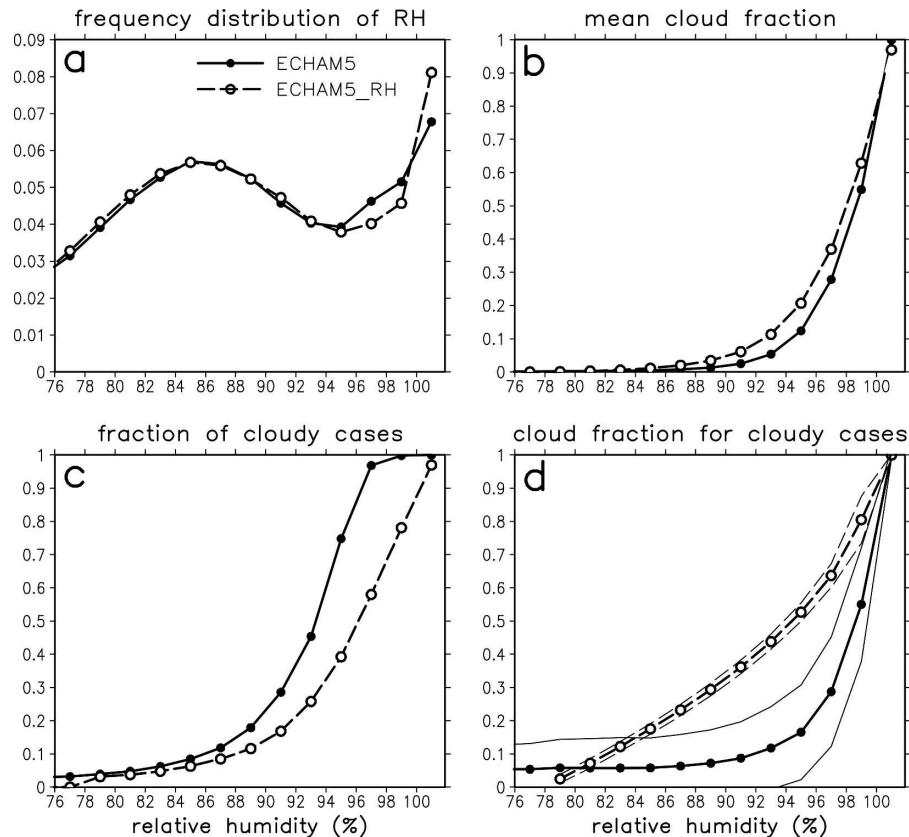


FIG. 10. Global-mean cloud statistics at model level 26 ( $\sigma \approx 0.85$ ) stratified according to grid-mean relative humidity (RH) (bin width = 2% for RH = 76%–100%; the last bin represents saturated cases RH  $\geq 100\%$ ). (a) Frequency distribution of RH for the ECHAM5 and ECHAM5\_RH datasets. (b) Mean cloud fraction (all cases), (c) fraction of cloudy cases (i.e., cloud fraction  $C > 0$ ), and (d) cloud fraction statistics for the cloudy cases only. In (d), the thick lines with symbols give the mean values, while the thin lines represent mean  $\pm 1$  std dev.

The primary impact of McICA noise on climate simulated by ECHAM5 was a slight reduction in low cloud fraction, with associated changes in radiative fluxes. This echoes earlier results obtained with the NCAR CAM, with some differences in details. The underlying physical mechanism could not be resolved completely, but it is most likely related to McICA's local heating rate errors in cloud layers and positive feedbacks that (apparently) preferentially amplify negative rather than positive perturbations in cloud fraction. For a degenerate high-noise version of McICA (the 1COL simulation), the differences from a low-noise reference simulation (REF) were fairly distinct, while they were much smaller, although still statistically robust, for a more reasonable GCM implementation of McICA (the CLDS simulation). For example, the global-mean 1COL – REF difference in the SW cloud radiative effect at the TOA was  $1.6 \text{ W m}^{-2}$ , while the corresponding CLDS – REF difference was only  $0.2 \text{ W m}^{-2}$ . For practical purposes, the main point is that

for a reasonable implementation of McICA, the impacts of noise can be kept very small. It thus appears that McICA provides a viable means to introduce a flexible description of unresolved cloud structure in the radiation calculations in ECHAM5.

A limitation that should be kept in mind is that this study, like all previous tests of McICA, used prescribed distributions of sea surface temperature and sea ice. Tests of McICA including an interactive ocean model are a part of our future plans.

Analysis of unresolved cloud structures and the associated radiative effects revealed two main points. First, ECHAM5 appears to underestimate subgrid-scale cloud variability. The radiative impact of subgrid-scale variability was substantially smaller than that for a global CSR dataset used for comparison, and the subgrid-scale variations in vertically integrated cloud optical thickness were underestimated compared to ISCCP satellite data. Part of the apparent underestimate probably stems from the fact that the core parts of convec-

tive clouds are not included in the cloud fraction and condensate fields in ECHAM5. However, there is also a principal limitation related to the current form of the beta distribution approach used for total water content  $q_t$  in ECHAM5: it is impossible to produce cloud layers with high subgrid-scale variability (relative standard deviation of condensate amount  $\sigma_{q_c} > 1$ ). This suggests that a more elaborate approach (e.g., a bimodal or multimodal distribution of  $q_b$ , or a separate parameterization for the PDF of condensate amount in the cloudy part of the layer) may be needed for a realistic description of subgrid-scale cloud variability.

Second, it was found that the impact of cloud overlap assumptions is quite small for ECHAM5. This was traced back to the frequency distribution of layer cloud fraction: comparisons to CSR data (for layer cloud fraction) and to ISCCP data (for total cloud fraction) both supported the notion that overcast cloud layers occur too frequently in ECHAM5, while layers with partial cloudiness occur too rarely. These features are, however, not necessarily linked to the use of the beta distribution scheme for diagnosis of cloud fraction, but may be related to other aspects of the cloud scheme. In fact, the same tendencies were more pronounced when an alternative relative humidity-based cloud fraction scheme was used.

*Acknowledgments.* Targeted funding for the Finnish Meteorological Institute's climate modeling activities from the Ministry of Transport and Communication Finland (Project Number 36937/2004) is gratefully acknowledged. P. Räisänen also received funding from the Academy of Finland (Project Number 205359). Technical support from the Max Planck Institute for Meteorology, Germany, is likewise acknowledged. Finally, the anonymous reviewers are thanked for their comments, which pushed us to improve the paper. This work is a contribution to the COSMOS Earth System Modeling network (<http://cosmos.enes.org>).

#### REFERENCES

- Bakan, S., and H. Quenzel, 1976: Path length distributions of photons scattered in turbid atmospheres. *Beitr. Phys. Atmos.*, **51**, 28–30.
- Barker, H. W., and P. Räisänen, 2005: Radiative sensitivities for cloud structural properties that are unresolved by conventional GCMs. *Quart. J. Roy. Meteor. Soc.*, **131**, 3103–3122.
- , R. Pincus, and J.-J. Morcrette, 2002: The Monte Carlo Independent Column Approximation: Application within large-scale models. *Extended Abstracts, GCSS-ARM Workshop on the Representation of Cloud Systems in Large-Scale Models*, Kananaskis, AB, Canada, GEWEX, 1–10. [Available online at <http://www.met.utah.edu/skrueger/gcss-2002/Extended-Abstracts.pdf>.]
- , and Coauthors, 2003: Assessing 1D atmospheric solar radiative transfer models: Interpretation and handling of unresolved clouds. *J. Climate*, **16**, 2676–2699.
- Cahalan, R. F., W. Ridgway, W. J. Wiscombe, T. L. Bell, and J. B. Snider, 1994: The albedo of fractal stratocumulus clouds. *J. Atmos. Sci.*, **51**, 2434–2455.
- Fouquart, Y., and B. Bonnel, 1980: Computations of solar heating of the Earth's atmosphere: A new parameterization. *Beitr. Phys. Atmos.*, **53**, 35–62.
- Hansen, J., G. Russell, D. Rind, P. Stone, A. Lacis, S. Lebedeff, R. Ruedy, and L. Travis, 1983: Efficient three-dimensional global models for climate studies: Models I and II. *Mon. Wea. Rev.*, **111**, 609–662.
- Hogan, R. J., and A. J. Illingworth, 2000: Deriving cloud overlap statistics from radar. *Quart. J. Roy. Meteor. Soc.*, **126**, 2903–2909.
- , and —, 2003: Parameterizing ice cloud inhomogeneity and the overlap of inhomogeneities using cloud radar data. *J. Atmos. Sci.*, **60**, 756–767.
- Iacono, M. J., E. J. Mlawer, and S. A. Clough, 2001: Validation of the RRTM shortwave radiation model and comparison to GCM shortwave models. *Proc. 11th Atmospheric Radiation Measurement (ARM) Science Team Meeting*, Atlanta, GA, U.S. Department of Energy. [Available online at [http://www.arm.gov/publications/proceedings/conf11/extended\\_abs/iacono\\_mj.pdf](http://www.arm.gov/publications/proceedings/conf11/extended_abs/iacono_mj.pdf).]
- Kiehl, J. T., and K. E. Trenberth, 1997: Earth's annual global mean energy budget. *Bull. Amer. Meteor. Soc.*, **78**, 197–208.
- Klein, S. A., and C. Jakob, 1999: Validation and sensitivities of frontal clouds simulated by the ECMWF model. *Mon. Wea. Rev.*, **127**, 2514–2531.
- Lacis, A. A., and V. Oinas, 1991: A description of the correlated  $k$  distribution method for modeling nongray gaseous absorption, thermal emission, and multiple scattering in vertically inhomogeneous atmospheres. *J. Geophys. Res.*, **96**, 9027–9063.
- Lohmann, U., and E. Roeckner, 1996: Design and performance of a new cloud micro-physics scheme developed for the ECHAM general circulation model. *Climate Dyn.*, **12**, 557–572.
- Mace, G. G., and S. Benson-Troth, 2002: Cloud-layer overlap characteristics derived from long-term cloud radar data. *J. Climate*, **15**, 2505–2515.
- Martin, G. M., D. W. Johnson, and A. Spice, 1994: The measurement and parameterization of effective radius of droplets in warm stratocumulus clouds. *J. Atmos. Sci.*, **51**, 1823–1842.
- Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated- $k$  model for the longwave. *J. Geophys. Res.*, **102**, 16 663–16 682.
- Morcrette, J.-J., and C. Jakob, 2000: The response of the ECMWF model to changes in the cloud overlap assumption. *Mon. Wea. Rev.*, **128**, 1707–1732.
- Moss, S. J., P. N. Francis, and D. G. Johnson, 1996: Calculation and parameterization of the effective radius of ice particles using aircraft data. *Proc. 12th Int. Conf. on Clouds and Precipitation*, Zürich, Switzerland, International Commission on Clouds and Precipitation and International Association of Meteorology and Atmospheric Science, 1255–1258.
- Pincus, R., H. W. Barker, and J.-J. Morcrette, 2003: A fast, flexible, approximate technique for computing radiative transfer in inhomogeneous cloud fields. *J. Geophys. Res.*, **108**, 4376, doi:10.1029/2002JD003322.



- , C. Hannay, S. A. Klein, K.-M. Xu, and R. Hemler, 2005: Overlap assumptions for assumed probability distribution function cloud schemes in large-scale models. *J. Geophys. Res.*, **110**, D15S09, doi:10.1029/2004JD005100.
- , R. Hemler, and S. A. Klein, 2006: Using stochastically generated subcolumns to represent cloud structure in a large-scale model. *Mon. Wea. Rev.*, **134**, 3644–3656.
- Pomroy, H. R., and A. J. Illingworth, 2000: Ice cloud inhomogeneity: Quantifying bias in emissivity from radar observations. *Geophys. Res. Lett.*, **27**, 2101–2104.
- Räisänen, P., and H. W. Barker, 2004: Evaluation and optimization of sampling errors for the Monte Carlo Independent Column Approximation. *Quart. J. Roy. Meteor. Soc.*, **130**, 2069–2085.
- , —, M. Khairoutdinov, J. Li, and D. A. Randall, 2004: Stochastic generation of subgrid-scale cloudy columns for large-scale models. *Quart. J. Roy. Meteor. Soc.*, **130**, 2047–2067.
- , —, and J. N. S. Cole, 2005: The Monte Carlo Independent Column Approximation's conditional random noise: Impact on simulated climate. *J. Climate*, **18**, 4715–4730.
- Roeckner, E., and Coauthors, 2003: The atmospheric general circulation model ECHAM5. Part I: Model description. Max Planck Institute for Meteorology Rep. 349, 127 pp. [Available from MPI for Meteorology, Bundesstr. 53, 20146 Hamburg, Germany.]
- , and Coauthors, 2006: Sensitivity of simulated climate to horizontal and vertical resolution in the ECHAM5 atmosphere model. *J. Climate*, **19**, 3771–3791.
- Rossow, W. B., C. Delo, and B. Cairns, 2002: Implications of the observed mesoscale variations of clouds for the earth's radiation budget. *J. Climate*, **15**, 557–585.
- Stubenrauch, C. J., A. D. Del Genio, and W. B. Rossow, 1997: Implementation of subgrid cloud vertical structure inside a GCM and its effects on the radiation budget. *J. Climate*, **10**, 273–287.
- Sundqvist, H., E. Berge, and J. E. Kristjansson, 1989: Condensation and cloud parameterization studies with a mesoscale numerical weather prediction model. *Mon. Wea. Rev.*, **117**, 1641–1657.
- Tompkins, A. M., 2002: A prognostic parameterization for the subgrid-scale variability of water vapor and clouds in large-scale models and its use to diagnose cloud cover. *J. Atmos. Sci.*, **59**, 1917–1942.
- Webb, M., C. Senior, S. Bony, and J.-J. Morcrette, 2001: Combining ERBE and ISCCP data to assess clouds in the Hadley Centre, ECMWF and LMD atmospheric climate models. *Climate Dyn.*, **17**, 905–922.
- Zhang, M. H., and Coauthors, 2005: Comparing clouds and their seasonal variations in 10 atmospheric general circulation models with satellite measurements. *J. Geophys. Res.*, **110**, D15S02, doi:10.1029/2004JD005021.

Copyright of *Journal of Climate* is the property of *American Meteorological Society* and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.