The Monte Carlo Independent Column Approximation’s Conditional Random Noise: Impact on Simulated Climate

P. RAISÄNEN

Finnish Meteorological Institute, Helsinki, Finland

H. W. BARKER AND J. N. S. COLE

Meteorological Service of Canada, Downsview, Ontario, Canada

(Manuscript received 29 September 2004, in final form 25 May 2005)

ABSTRACT

The Monte Carlo Independent Column Approximation (McICA) method for computing domain-average radiative fluxes is unbiased with respect to the full ICA, but its flux estimates contain conditional random noise. Results for five experiments are used to assess the impact of McICA-related noise on simulations of global climate made by the NCAR Community Atmosphere Model (CAM). The experiment with the least noise (an order of magnitude below that of basic McICA) is taken as the reference. Two additional experiments help demonstrate how the impact of noise depends on the time interval between calls to the radiation code. Each experiment is an ensemble of seven 15-month simulations.

Experiments with very high noise levels feature significant reductions to cloudiness in the lowermost model layer over tropical oceans as well as changes in highly related quantities. This bias appears immediately, stabilizes after a couple of model days, and appears to stem from nonlinear interactions between clouds and radiative heating. Outside the Tropics, insignificant differences prevail. When McICA sampling is confined to cloudy subcolumns and when, on average, 50% more samples, relative to basic McICA, are drawn for selected spectral intervals, McICA noise is much reduced and the results of the simulation are almost statistically indistinguishable from the reference. This is true both for mean fields and for the nature of fluctuations on scales ranging from 1 day to at least 30 days.

While calling the radiation code once every 3 h instead of every hour allows the CAM additional time to incorporate McICA-related noise, the impact of noise is enhanced only slightly. In contrast, changing the radiative time step by itself produces effects that generally exceed the impact of McICA’s noise.

1. Introduction

Global climate models (GCMs) represent the earth–atmosphere system as a 3D array of cells that have horizontal cross-sectional areas typically well in excess of $10^4$ km$^2$. Conditions in these cells march forward in discrete time steps. At spatial and temporal scales unresolved by GCMs, processes must be parameterized as functions of both the resolved state and ancillary assumptions. Subgrid-scale parameterizations attempt to capture the essence of ubiquitous, generally nonlinear, interactions that GCMs neither portray nor sense directly.

Given their central role in GCMs, subgrid-scale parameterizations must provide unbiased estimates with respect to all independent variables and assumptions. With respect to modeling radiative transfer in GCMs, the only reasonable standard to which GCM-style radiation codes can be expected to aspire to is the Independent Column Approximation (ICA; e.g., Barker et al. 2003). The ICA does not account for true 3D radiative transfer, but any attempt to specify the unresolved 3D structure of a conventional GCM column, given the limited information available, would be overwhelmed by uncertainty. With the ICA, a 1D radiative transfer model is applied to all columns in a specified domain, and flux profiles for individual columns are averaged horizontally to produce domain-average fluxes. To the best of our knowledge, the Monte Carlo Independent Column Approximation (McICA) for computing domain-average radiative fluxes in GCMs is the only ra-
radiative transfer parameterization whose expectation value is identically equal to the full ICA (Barker et al. 2002; Pincus et al. 2003).

The McICA method extricates description of cloud structure from the radiative transfer solver thereby allowing for unlimited, one-point statistical descriptions of unresolved cloud structure. To apply McICA in a GCM, subgrid-scale columns are required. These can be supplied by Räisänen et al.’s (2004) stochastic generator, which is initialized with information from the GCM and assumptions regarding horizontal variance of cloud water and vertical decorrelation lengths that determine overlap rates for cloud fraction and condensate. Both McICA and the generator give rise to conditional random sampling errors (i.e., noise). So, while McICA is unbiased, and thus achieves the primary goal of subgrid-scale parameterizations, it raises questions regarding the impact of random errors on GCM simulations.

The root of our concern over McICA random errors stems from the viewpoint that they represent extraneous information generated within the GCM that could affect the simulation. Presumably, the sensitivity of a nonlinear dynamical system, like a GCM, to these errors depends on the spatial and temporal character of the noise, its magnitude, and the system itself. If the noise is such that it has a statistically significant impact on a simulation, then it can be considered to be above the system’s noise horizon. McICA’s success depends on its conditional sampling errors being beneath the noise horizon of GCMs. Recognizing that this might not be the case, Räisänen and Barker (2004) devised procedures for reducing McICA’s sampling variance.

Pincus et al. (2003) injected a proxy for McICA sampling errors into the European Centre for Medium-Range Weather Forecast’s (ECMWF’s) global model and showed that it has minimal influence on 10-day and 4-month simulations. While their tests are very encouraging, they are not definitive for at least four reasons: (i) the proxy noise lacked the proper conditionality of McICA’s noise; (ii) conditional noise associated with a generator was not included; (iii) McICA noise depends on the radiation codes used, characteristics of the GCM itself, and the nature of clouds realized by the GCM; and (iv) parameterizations of processes used in different GCMs may respond differently to McICA-related noise. The last two points hint at the possibility that the impact of McICA’s random sampling errors might be GCM dependent.

This study takes the first few, of potentially several, steps toward assessing the impact of McICA-related noise on simulations of global climate. As in Pincus et al. (2003), a single atmospheric GCM is employed. Here, however, McICA and a cloud generator are implemented in the GCM thereby eliminating the need for proxy noise. Moreover, the impact of unbiased radiative noise is addressed via several ensemble simulations. These simulations span implementations of McICA ranging from the drawing of a single subgrid-scale column (which maximizes noise) to versions that employ noise reduction techniques (Räisänen and Barker 2004).

Section 2 describes the GCM used here to perform ensemble simulations and summarizes the noise levels associated with various incarnations of McICA. Section 3 discusses results of the simulations, including annual-mean fields, selected aspects of temporal evolution, and the impact on low-level tropical clouds, which are, for this GCM, especially sensitive to noise. Conclusions are given in section 4.

2. Methodology

This section has three subsections. The first describes the models used here. The second introduces the various experimental and reference simulations, and the third quantifies the levels of random noise associated with the simulations.

a. Model

Version 1.8 of the National Center for Atmospheric Research (NCAR) Community Atmosphere Model (CAM) was used as the host GCM. CAM-1.8 is an improved version of the Community Climate Model 3 (Kiehl et al. 1996) that resembles CAM-2 (Collins et al. 2003). The CAM was run at T42 horizontal resolution (a grid spacing of ~2.8°) with 26 layers in the vertical and model top at ~3 hPa. The semi-Lagrangian approach was employed for atmospheric dynamics with a time step of 1 h. Likewise, physical parameterizations, including radiation, were invoked hourly, unless stated otherwise.

The simulations used a prognostic stratiform cloud scheme (Rasch and Kristjansson 1998; Zhang et al. 2003) and Zhang and McFarlane’s (1995) and Hack’s (1994) routines for deep and shallow/midlevel convection, respectively. With two exceptions, other physical parameterizations also followed the standard CAM. First, the CAM radiation scheme was replaced by a correlated k-distribution (CKD) radiation code (Cole 2005). Second, a stochastic cloud generator (Räisänen et al. 2004) was employed to define unresolved cloud structure for each GCM column.

The radiation code used here is characterized in the shortwave (SW) by a multilayer two-stream approxi-
K is the overlap parameter, which is used by the generator. For the SW and Fu and Liou (1992) for the LW. There are $X = 55$ quadrature points in cumulative probability space (CPS) for the SW and $X = 67$ for the LW. Optical properties for cloud droplets and ice crystals are resolved into 15 SW bands and 12 LW bands (Cole 2005). As discussed below, and by virtue of McICA, the radiative transfer algorithms operate on subcolumns whose layers are homogeneous and either cloud free or overcast. Consequently, they act only on profiles of optical properties (e.g., Wiscombe 1977); they do not treat cloud overlap or horizontal inhomogeneity explicitly.

Parameterization of cloud droplet effective radius $r_e$ followed the standard CAM. Over oceans $r_e$ was fixed at 10 $\mu$m, while over land it varied between 7 and 10 $\mu$m depending on layer temperature. Ice crystal $r_e$ varied between 23 and 30 $\mu$m depending on layer pressure.

Subgrid-scale cloud structure is provided by Räisänen et al.’s (2004) stochastic generator. The generator uses, as input, profiles of layer cloud fraction $C$, mean liquid $\bar{w}_l$ and ice $\bar{w}_i$ water contents, probability distributions $p(w)$ of total condensed cloud water $w$ for each layer, and parameters that describe overlap of cloud fraction and cloud condensate. Values of $C$, $\bar{w}_l$, and $\bar{w}_i$ come from the CAM. For the other variables, the following assumptions were made for all simulations. Probability distributions for $w$ were described by

$$p(w) = \frac{1}{\Gamma(\nu)} \left( \frac{w}{\bar{w}} \right)^{\nu-1} e^{-w/\bar{w}},$$

where $\bar{w}$ is mean total cloud water content for the cloudy part of a GCM grid cell, and $\nu$ is a parameter related to the variance of $w$ and was set for each layer as

$$\nu = \begin{cases} 1 & C < 0.9 \\ 2 & 0.9 \leq C < 1 \\ 4 & C = 1. \end{cases}$$

This prescription is based on Barker et al.’s (1996) analysis of satellite imagery. Cloud-system-resolving models (CSRMs) have also been observed to approximate (1) and (2) very roughly (see Barker et al. 2003). By no means is (2) supposed to represent a definitive parameterization; it simply provides a reasonable amount of variability for these sensitivity experiments.

If two layers $k$ and $l$ have cloud fractions $C_k$ and $C_l$, the generator assumes that total vertically projected cloud fraction formed by taking these layers together is

$$C_{k,l} = \alpha_{k,l} C_{k,l}^{\max} + (1 - \alpha_{k,l}) C_{k,l}^{\mathrm{ran}},$$

where

$$C_{k,l}^{\max} = \max(C_k, C_l)$$

$$C_{k,l}^{\mathrm{ran}} = C_k + C_l - C_k C_l,$$

and $\alpha_{k,l}$ is the overlap parameter, which is used by the generator. Analyses of observations by Hogan and Illingworth (2003) and Hogan and Illingworth (2003) showed that overlap of condensate for ice clouds can also be approximated by (5) but that the corresponding decorrelation length $\xi_{sw}$ is generally smaller than $\xi_{lw}$. This was also the case for both liquid and ice clouds within the CSRM datasets analyzed by Räisänen et al. (2004) and Pincus et al. (2005). Thus, in the experiments performed here, $\xi_{sw}$ was set constant over the globe at 1 km.

b. Ensemble simulations

Seven ensembles of simulations were performed. Each ensemble consists of seven members in which each member was initialized with a unique seed for the random number generator used by the subgrid-scale generator. Ensemble size was dictated by computational constraints. Significantly larger ensembles have been used for numerical weather forecasting (e.g., Molteni and Buizza 1999) and other studies (e.g., Pincus et al. 2003). As shown below, however, ensembles of only seven members do not preclude findings that are statistically highly significant.

The simulations were run for 15 months for the period 1 September 1960–30 November 1961 and used prescribed distributions of sea surface temperatures (SSTs) and sea ice. The first 3 months of the simulations were regarded as spinup, and the last 12 months (1 December 1960–30 November 1961) were analyzed.

All ensembles used the same GCM and the same basic radiation codes, made the same assumptions about unresolved cloud structure, and utilized McICA. What distinguished them was the way in which they employed McICA and therefore produced McICA-generator sampling noise. Additionally, two of the ensembles used radiative time steps $\Delta t_{rad}$ of 3 h rather than 1 h, as used by the other five. The purpose of $\Delta t_{rad}$
= 3 h was to explore the impact of allowing random fluctuations more time to influence the CAM.

The ensembles with least sampling noise were considered to be reference simulations and are referred to hereinafter as REF (Δt_rad = 1 h) and REF3 (Δt_rad = 3 h).

For the other ensembles, the impact of sampling noise on simulated climate is defined as the difference between their results and those of the corresponding reference simulation. The seven ensembles are detailed below. They are presented in terms of decreasing sampling noise.

1) EXPERIMENT A: 1COL

The McICA used in this experiment is a reduced version of its most basic form. It resembles the approach employed in the Goddard Institute of Space Studies (GISS) GCM (Hansen et al. 1983; Stubenrauch et al. 1997). Broadband radiative fluxes are computed as

\[ \hat{F} = \sum_{k=1}^{\infty} f(s, k), \]

where \( s \) is a single randomly generated subcolumn that is used for all \( \infty \) spectral intervals. Although to our understanding GISS has only used traditional overlap assumptions in GCM simulations, this approach, like all renditions of McICA, allows for arbitrary flexibility in the definition of unresolved cloud structure. Moreover, it maximizes the magnitude of unbiased radiative noise for the class of McICA models that perform full spectral integrations (cf. Stephens et al. 2004). For this experiment, the radiation codes were called every time step (i.e., Δt_rad = 1 h).

2) EXPERIMENT B: BASIC

The second ensemble uses McICA in its most basic form (cf. Räisänen and Barker 2004). Broadband domain-average fluxes are estimated as

\[ \hat{F} = \sum_{k=1}^{\infty} f(s_k, k), \]

where flux calculations for each CKD quadrature point \( k \) are performed on a unique, stochastically generated (cloudy or clear) subcolumn denoted as \( s_k \). As in 1COL, Δt_rad = 1 h.

3) EXPERIMENT C: CLDS

As discussed by Räisänen and Barker (2004), sampling noise associated with (7) can be reduced by using

\[ \hat{F} = (1 - \hat{C}) \sum_{k=1}^{\infty} f(s_{clr}, k) + \hat{C} \sum_{k=1}^{\infty} f(s_{clr,k}, k) \]

\[ = (1 - \hat{C})F_{clr} + \hat{C}\hat{F}_{clr}, \]

where \( \hat{C} \) is an estimate of total cloud fraction, \( s_{clr} \) is the clear-sky column, \( s_{clr,k} \) are randomly selected cloudy subcolumns, and \( F_{clr} \) and \( \hat{F}_{clr} \) are clear-sky and cloudy broadband fluxes. Values of \( \hat{C} \) equal \( \sum_{k=1}^{\infty} f(s_k, k) \), where \( f = 100 \) is the total number of subcolumns generated per GCM column, and \( \sum_{k=1}^{\infty} \) is the number of subcolumns that contain cloud. Räisänen and Barker (2004) denoted this version of McICA as CLDSAMPL, but it is abbreviated here to CLDS. Since \( F_{clr} \) are computed for diagnostic purposes for all experiments, CLDS takes insignificantly more computer time than both BASIC and 1COL. Again, Δt_rad = 1 h.

4) EXPERIMENT D: SPEC

The fourth experiment reduces McICA variance further by selectively improved spectral sampling: multiple samples are drawn for those, and only those, spectral intervals that contribute much to cloud radiative effects, and consequently, McICA noise (see Räisänen and Barker 2004). In this case, fluxes are computed as

\[ \hat{F} = (1 - \hat{C})F_{clr} + \hat{C} \sum_{k=1}^{\infty} \left( \frac{1}{N_k} \sum_{n=1}^{N_k} f(s_{clr,n,k}, k) \right) \]

\[ = (1 - \hat{C})F_{clr} + \hat{C}\hat{F}_{clr}. \]

where \( N_k \) is the number of samples for the \( k \)th interval, and \( s_{clr,n,k} \) is the \( n \)th cloudy subcolumn used for the \( k \)th interval. Values for \( \hat{C} \) and \( F_{clr} \) are computed as in CLDS. Sets of \( N_k \) were computed a priori based on data from a 1-day simulation. For SW calculations over land, \( [N_k] \) were chosen so as to minimize errors in net surface flux. For SW calculations over ocean and LW calculations over both land and ocean, \( [N_k] \) were based on minimization of sampling errors for atmospheric heating rates (Räisänen and Barker 2004). The total number of samples \( (\sum_{k=1}^{\infty} N_k) \) is fixed at 82 for SW and 100 for LW. This represents ~50% more samples than for CLDS, which used \( N_k = 1 \) for all \( k \) (but it increases CPU time spent for radiation by <25%). As in the previous three experiments, Δt_rad = 1 h.

5) EXPERIMENT E: SPEC3

This experiment is like SPEC in that it uses (9). It differs from SPEC, however, in that Δt_rad = 3 h instead of 1 h. Between radiative time steps, radiative fluxes and heating rates are simply kept constant.
6) **Reference A: REF**

As with SPEC, the reference experiment uses (9) to reduce McICA noise along with $t_{\text{rad}} = 1$ h. However, the number of subcolumns generated per GCM column per time step is $J = 1500$, and $\sum_{k} N_k$ in (9) equals 1100 for SW and 1340 for LW. Hence, sampling errors associated with $\hat{C}$, $\hat{F}_{\text{clt}}$, and ultimately $\hat{F}$ are small relative to the previously mentioned experiments. So, while the radiation calculations for this reference require an order of magnitude more CPU time than the previous experiments, they still contain sampling noise that arises from both the stochastic cloud generator and McICA. The latter source of noise could be eliminated by using the full ICA, but it would then be computationally intractable to use 1500 subcolumns. In fact, ICA with $J = 50$ subcolumns takes as much CPU time as the current reference calculations, but, owing to substantial sampling noise associated with the cloud generator when $J = 50$, sampling errors for radiative fluxes and heating rates are 3–4 times larger.

c. **Quantification of sampling noise**

To evaluate the level of sampling noise associated with the various experiments, diagnostic calculations were performed for one member of each ensemble. On every 45th time step of the CAM, the cloud generator and the radiation codes were invoked 10 times instead of the usual once, and standard deviations were computed for fluxes and heating rates. Table 1 indicates that for the radiation codes used here, sampling errors for net surface fluxes come primarily from SW radiation, while for heating rates they come primarily from LW radiation. Figure 1 shows that surface radiative flux sampling errors are generally largest in the Tropics, while sampling errors for total radiative heating rate maximize near the surface.

The different simulations feature a wide range of radiative noise. By far the smallest sampling errors occur for the REF simulation (nearly 80% smaller than SPEC). Conversely, 1COL features sampling noise that greatly exceeds that for all other experiments, especially for surface fluxes. BASIC, CLDS, and SPEC fall

7) **Reference B: REF3**

This reference is identical to REF except that it uses $t_{\text{rad}} = 3$ h. As with SPEC3, radiative fluxes and heating rates are kept constant between radiative time steps.

![Figure 1](image_url). Zonal-average standard deviations for total (SW + LW) net flux at the surface and global-average profiles of standard deviations for total radiative heating rate. They represent noise levels produced by McICA coupled with the stochastic cloud generator. Results for REF3 and SPEC3 (not shown) are very close to REF and SPEC, respectively.
between these two extremes. The most important difference between these three simulations lies in LW and total heating rate sampling errors, in which values for SPEC are typically ~55% smaller than those for CLDS. Improved spectral sampling has a smaller effect on SW, and thus the total surface fluxes than it did in Räisänen and Barker (2004). This is because the spectral weights of quadrature points in CPS are distributed much more evenly for the current SW code than they were for the code used by Räisänen and Barker (2004). This helps reduce sampling errors in general, but more so for BASIC and CLDS than for SPEC.

For the current tests, noise produced by McICA sampling generally dominates over that produced by the cloud generator. For \( N = 100 \) subcolumns, global rms sampling errors associated with the generator were estimated to be 6.8 W m\(^{-2}\) for total net radiative flux at the surface and 0.088 K day\(^{-1}\) for total radiative heating rate. This implies, for example, that for SPEC (BASIC), 26% (11%) of sampling variance in net surface fluxes, and 28% (3%) of sampling variance in heating rates, comes from the generator, the rest coming directly from McICA.

### 3. Results

#### a. Impact of McICA noise on simulated global climate

The impact of McICA noise on the simulation of mean annual, global climate is assessed by performing Student’s two-sided \( t \) test on several atmospheric variables. Simulations in which radiative transfer calculations were performed every hour are considered first. These are followed by analyses of simulations in which radiative transfer calculations were performed every 3 h.

Table 2 lists global-mean values of several variables for REF. It also lists differences between these values and those simulated by the other experiments and indicates whether they are statistically significant at either the 95% or 99.9% confidence level. It is clear that for CLDS and BASIC, and especially 1COL, McICA-related noise leads to some notable differences relative to REF. The most obvious difference occurs for low cloud fraction, which is underestimated by 0.018 (or 6%) by the 1COL simulation. Consistently, all values of cloud radiative effect (CRE; also known as cloud radiative forcing), which is defined as

\[
CRE = F_{\text{net}}^{\text{all-sky}} - F_{\text{clear-sky}}^{\text{net}},
\]

are significantly smaller than those for REF. Note also that all global-mean values for CRE listed in Table 2 are less than corresponding observed values (e.g., Kiehl and Trenberth 1997). A likely reason for this is that no attempt was made to retune the GCM following the implementation of a new radiative transfer scheme and, in particular, inclusion of subgrid-scale cloud variability. In contrast to the considerable impact of noise manifest in 1COL, the SPEC simulation registered no significant impact at the 95% confidence level.

Taking a slightly different tack, Table 3 lists fractional areas \( f_{\text{Ref}} \) of the globe that exhibit differences relative to
Table 3. Fractional areas of the earth (%) for which annual-mean differences to the reference simulation are statistically significant at the 95% confidence level; $P_i$ is surface pressure and $z_{500 \text{ hPa}}$ is 500-hPa geopotential. Other notations are as in Table 2. SPEC3 refers to the differences between SPEC3 and REF3, while REF3 refers to differences between REF3 and REF.

<table>
<thead>
<tr>
<th>Quantity</th>
<th>1COL</th>
<th>BASIC</th>
<th>CLDS</th>
<th>SPEC</th>
<th>SPEC3</th>
<th>REF3</th>
</tr>
</thead>
<tbody>
<tr>
<td>SW CRE TOA</td>
<td>23.3</td>
<td>14.1</td>
<td>9.6</td>
<td>5.1</td>
<td>6.3</td>
<td>23.0</td>
</tr>
<tr>
<td>SW CRE surface</td>
<td>24.5</td>
<td>15.0</td>
<td>10.1</td>
<td>5.2</td>
<td>6.2</td>
<td>23.1</td>
</tr>
<tr>
<td>LW CRE TOA</td>
<td>8.8</td>
<td>6.6</td>
<td>5.8</td>
<td>3.7</td>
<td>4.3</td>
<td>18.6</td>
</tr>
<tr>
<td>LW CRE surface</td>
<td>33.1</td>
<td>21.0</td>
<td>15.4</td>
<td>4.7</td>
<td>6.1</td>
<td>17.1</td>
</tr>
<tr>
<td>Clouds, high</td>
<td>7.2</td>
<td>6.3</td>
<td>6.5</td>
<td>3.9</td>
<td>4.2</td>
<td>20.8</td>
</tr>
<tr>
<td>Clouds, middle</td>
<td>7.6</td>
<td>6.9</td>
<td>5.0</td>
<td>3.6</td>
<td>5.2</td>
<td>14.3</td>
</tr>
<tr>
<td>Clouds, low</td>
<td>35.6</td>
<td>22.8</td>
<td>15.7</td>
<td>5.2</td>
<td>5.7</td>
<td>18.0</td>
</tr>
<tr>
<td>Clouds, total</td>
<td>16.0</td>
<td>8.3</td>
<td>6.1</td>
<td>3.7</td>
<td>5.3</td>
<td>20.4</td>
</tr>
<tr>
<td>LWP</td>
<td>15.7</td>
<td>10.8</td>
<td>7.0</td>
<td>4.5</td>
<td>5.4</td>
<td>16.7</td>
</tr>
<tr>
<td>IWP</td>
<td>7.4</td>
<td>7.0</td>
<td>6.7</td>
<td>3.4</td>
<td>4.9</td>
<td>18.6</td>
</tr>
<tr>
<td>$T_2 \text{ m}$</td>
<td>25.7</td>
<td>17.8</td>
<td>16.6</td>
<td>3.0</td>
<td>4.9</td>
<td>32.0</td>
</tr>
<tr>
<td>$T_{500 \text{ hPa}}$</td>
<td>32.9</td>
<td>13.6</td>
<td>16.6</td>
<td>3.0</td>
<td>6.4</td>
<td>64.3</td>
</tr>
<tr>
<td>Precipitation</td>
<td>8.0</td>
<td>6.0</td>
<td>5.2</td>
<td>3.2</td>
<td>5.1</td>
<td>18.2</td>
</tr>
<tr>
<td>PW</td>
<td>7.6</td>
<td>6.6</td>
<td>5.1</td>
<td>2.2</td>
<td>4.6</td>
<td>24.3</td>
</tr>
<tr>
<td>$P_i$</td>
<td>6.5</td>
<td>3.7</td>
<td>2.3</td>
<td>1.4</td>
<td>6.2</td>
<td>32.6</td>
</tr>
<tr>
<td>$z_{500 \text{ hPa}}$</td>
<td>18.4</td>
<td>5.7</td>
<td>5.3</td>
<td>0.0</td>
<td>5.4</td>
<td>45.2</td>
</tr>
</tbody>
</table>

REF that are statistically significant at the 95% confidence level. If a variable exhibits adequate horizontal variability (Livezey and Chen 1983), and one draws two random samples from a common population, one can expect $f_{\alpha} \approx (100 - \alpha)\%$. For 1COL this is definitely not the case as $f_{\alpha} > 5\%$ for many variables including the climatologically important air temperature at 2 m above the surface. After restricting McICA samples to cloudy subcolumns and enhancing sampling for important spectral regions, $f_{\alpha}$ for SPEC are typically about 5%. Results for $\alpha = 99\%$ (not listed) behaved similar to $\alpha = 95\%$. This implies that the ensemble produced by SPEC was drawn, statistically speaking, from the same population as REF.

The most dramatic changes in $f_{\alpha}$ as a function of McICA noise level are for low cloudiness and as a consequence SW CRE at the surface and at the top of the atmosphere (TOA), and LW CRE at the surface. There are some interesting features associated with this bias, and they are discussed later in section 3c.

To bolster the claim that $f_{\alpha} \approx 5\%$ implies that simulations are effectively samples drawn from the same population, regions of statistically significant differences should not cluster in coherent spatial patterns. Figure 2 shows regions that exhibit statistically significant differences at the 95% and 99% confidence levels for mean annual SW CRE at TOA for the four experiments relative to REF. For the 1COL, BASIC, and CLDS experiments, reduction of $f_{\alpha}$ with noise is obvious, however, all exhibit strong clustering and confinement of differences to tropical oceans. As seen later, this is due directly to reductions in amounts of low, quite reflective, cloud. For all experiments, differences outside the Tropics display weak coherence.

Consistent with results presented thus far for SPEC, Fig. 2 shows that there is almost no regional pattern to the differences relative to REF. Instead, they appear more like a random scattering of isolated deviations. This suggests that the CAM’s noise horizon, as alluded to in the introduction and associated with McICA noise in particular, lies between the noise produced by SPEC and CLDS. Judging from the noise levels shown in Table 1 and Fig. 1, and the fact that the impact of noise was limited mostly to oceanic areas, it is likely that the much-improved behavior of SPEC relative to CLDS stems from reduction of noise, via additional sampling, in atmospheric LW heating rates.

The impact of sampling noise on zonal-mean temperature distribution is shown in Fig. 3. The most distinct feature is a slight warming of the tropical troposphere for 1COL, BASIC, and CLDS relative to REF. For 1COL, this is accompanied by a cooling near the tropical tropopause and into the adjacent stratosphere. While larger absolute differences to REF occur in polar regions, they are generally statistically insignificant. The tropical tropospheric warming is probably related to reductions in very low cloudiness, which reduces radiative cooling near the surface, destabilizes the lower atmosphere, enhances convection slightly, and thus transports more energy aloft. It should be noted that use of prescribed distributions of SSTs and sea ice may restrict temperature changes substantially, for changes in surface radiation budget have no impact over oceans.

Table 1 shows that noise levels for simulations with a radiative time step of $\Delta t_{\text{rad}} = 3 \text{ h}$ are similar to those using $\Delta t_{\text{rad}} = 1 \text{ h}$. Table 2 shows, however, that the impact of noise going from REF3 to SPEC3 generally exceeds that between SPEC and REF, particularly for
low cloud fraction and LW CRE at the surface where global mean differences are statistically significant at the 95% confidence level. Larger impacts of sampling noise are expected with \( t_{rad} = 3 \) h since lengthening \( t_{rad} \) allows the CAM's dynamics more time to respond to McICA fluctuations. Nevertheless, values of \( f_{SW} \) for SPEC3 are still close to 5% for all quantities (Table 3), and spatial patterns (not shown) resemble those shown for SPEC in Fig. 3.

Tables 2 and 3 also show that the impact of changing \( t_{rad} \) from 1 (REF) to 3 (REF3) h is quite significant. For many of the quantities listed, differences between REF3 and REF are comparable or even larger than those between 1COL and REF. This echoes Morcrette's (2000) finding that large errors can be incurred by increasing \( t_{rad} \). One of the impacts of lengthening \( t_{rad} \) is to shift the diurnal cycle of cloudiness. The current version of CAM features (perhaps erroneously) an early morning maximum and afternoon minimum in total cloud fraction and cloud optical thickness, especially over tropical land areas. When radiation is called once every 3 h, instead of every hour, this cycle is delayed, effectively putting more cloud near noon, and thus enhancing mean SW CRE in REF3 relative to REF. There were also appreciable differences in horizontal distributions of cloud between REF3 and REF, and most of the troposphere was colder for REF3. It is, however, not the purpose of this study to provide a detailed analysis of these non-McICA-induced effects. Rather, these results serve to demonstrate that the impact of changing \( t_{rad} \) in a GCM can easily surpass that of McICA noise.

b. Impact of McICA noise on the evolution of simulated climate

In addition to studying the impact of McICA noise on simulations of mean climate, it is also important to assess whether McICA-related noise alters the evolution of simulated climate. Although an exhaustive treatment of all temporal variations is beyond the scope of the present paper, the impact of McICA noise on the evolution of the system is tested by applying second-order structure function analysis to selected regions and quantities. The second-order structure function is related to the conventional power spectrum but is more intuitive and flexible than Fourier analysis (Marshak et al. 1997). If \( \varphi \) denotes a variable, then the difference between \( \varphi \) at times \( t \) and \( t + \tau \) is

\[
\Delta \varphi(\tau) = \varphi(t + \tau) - \varphi(t). \tag{11}
\]
The $n$th-order structure function over some time interval is denoted as $\langle |\Delta \phi (\tau) |^n \rangle$, where angular brackets signify time average. Only $\langle |\Delta \phi (\tau) |^2 \rangle$ and two climatologically diverse regions are considered: north-central Asia and the tropical west Pacific (TWP). Figure 4 indicates the location of these regions. While the analyses performed here could be done for individual cells, it is more appropriate to analyze current GCMs regionally. Though by no means a proven fact, it seems reasonable to hypothesize that statistical characteristics of extreme values of variables may be overly susceptible to the presence of high-frequency random noise, such as that generated by McICA.

Some characteristics of diurnal minimum air temperature 2 m above the surface $T_{2m,\text{min}}$ averaged over the Asian region are considered first. Most of this region is characterized by continental climatic conditions in which $T_{2m,\text{min}}$ are quite sensitive to changes in cloudiness. Figure 5 (left) shows frequency distributions of regional mean $T_{2m,\text{min}}$ for December–January–February (DJF) for 1COL, SPEC, and REF, based on daily results for seven separate runs per experiment. All three distributions resemble each other closely, especially those for SPEC and REF.

Figure 5 (right) shows the ensemble average and standard deviation of $\langle |\Delta \phi (\tau) |^2 \rangle$ for $T_{2m,\text{min}}$ for 1COL, SPEC, and REF. It also shows the results of $t$ tests at each time lag $\tau$. The quantity shown is $t^*$ and is referred to as the significance. It ranges between zero and one, and is the probability that for random samples drawn from the same underlying distribution, the common $t$ statistic would be as small as or smaller than that computed from the current results. Therefore, if $t^* = 0.95$, this implies that the difference is significant at the 95% confidence level, while $t^* = 0.05$ implies that the difference is significant at only the 5% confidence level. In the latter case, the difference would be deemed by most as insignificant. Thus, these plots indicate that fluctuations in $T_{2m,\text{min}}$ for SPEC and REF for periods ranging from 1 day to 30 days are almost identical, as values of $t^*$ are generally small. Even 1COL exhibits fluctuations that are, for the most part, not even close to being
significantly different from those of REF at the 95% confidence level.

Figure 6 shows 10-day running-mean time series of diurnal mean SW CRE at TOA averaged over the TWP region. Consistent with Fig. 2 and Table 2, one can see clearly a systematic and increasing underestimation in the magnitude of SW CRE as McICA noise increases.

Figure 7 (left) shows frequency distributions of SW CRE at TOA over the TWP region for the September–October–November (SON) period. Clearly, the distribution for 1COL features a systematic shift toward less negative values (i.e., weaker CRE). Figure 7 (right) displays the corresponding structure functional analyses. Here, just as with $T_{2m,min}$ in Asia, errors in fluctuations can be considered to be minor for all experiments, especially those for SPEC. This reinforces the idea that the systematic offset in CRE for 1COL because of reduced low-level cloudiness is the result of a local,

**FIG. 4.** Location of two regions for which results are presented in subsequent figures; A denotes north-central Asia and B denotes TWP.

**FIG. 5.** Analysis of regional mean diurnal minimum air temperature 2 m above the surface for north-central Asia for the period DJF. (left) Frequency distributions for the simulations 1COL, SPEC, and REF (seven separate runs for each ensemble simulation). (right) Ensemble averages and standard deviations of second-order structure functions as functions of time lag $\tau$. In the lower part of the right plot, $t^*$ defines the statistical significance of the differences in structure function between either SPEC or 1COL and REF. For example, $t^* = 0.1$ implies that fluctuations at period $\tau$ as realized by either SPEC or 1COL differ significantly from REF at only the 10% confidence level. Conversely, when $t^*$ exceeds 0.95, the structure functions differ significantly at the 95% confidence level.
short-term feedback and not one associated with larger-scale circulation changes.

Figure 8 (left) shows frequency distributions of LW CRE for the TWP region during SON. It is apparent here that the magnitude of LW CRE for both SPEC and 1COL agrees nicely with that of REF though there is some indication that as McICA noise increases, the tails of the distribution broaden slightly. Note that changes to low-level cloudiness were not expected to influence LW CRE at TOA, and the similarity of these distributions testify to only slight alterations in upper-level clouds due to McICA noise. Again, Fig. 8 (right) shows structure function analyses, and what it indicates is that while differences in fluctuations relative to REF from day to day (i.e., $\tau = 1$) for 1COL are significant at the 90% confidence level, they are negligible for SPEC at all periods less than about 30 days.

c. Low-level cloud over tropical oceans

As noted above several times, the most pronounced impact of radiative noise is a decrease to low cloud fraction over tropical oceans. Furthermore, this decrease is largely confined to the lowermost model layer (i.e., $\leq 130$ m above the surface). Figure 9 shows that even the SPEC simulation’s underestimation of zonal-mean cloud fraction in the lowest layer over equatorial oceans is statistically significant at the 95% confidence level. For the other experiments, especially 1COL, the underestimation is rather severe.

We feel compelled to preface this subsection by noting that cloud fraction in the lowest model layer over tropical oceans for REF is 0.10–0.15, which is suspiciously high compared to ship reports (Warren et al. 1988). Cloud fractions were similar for standard CAM 1.8, so it is not an artifact related to our radiation code. The cloud fractions were slightly smaller for CAM 3, and much smaller for a superparameterized version of CAM 1.8 (Khairoutdinov and Randall 2001). Therefore, the largest impact of McICA noise is on a variable that appears to be simulated poorly by CAM 1.8. Nevertheless, we consider it an interesting impact worth documenting.

To illustrate the nature of the tropical low cloud fraction bias, additional short simulations were made using the REF and 1COL approaches. In one set of simulations, 21 five-day 1COL runs were started from initial states produced by a single REF run (21 initial states during 1–6 September 1960, at 6-h intervals). Figure 10 (upper left) shows the time development of the difference in cloud fraction for the lowest layer between 1COL and REF, averaged over oceans between $15^\circ$S and $15^\circ$N. For a typical 1COL run, a negative bias emerges within 2 h. The bias increases steadily to about $-0.05$ then stabilizes rather abruptly after $\sim 2$ days. This bias was present throughout the 15-month en-
semble simulations (annual-mean bias being $-0.04$) with only modest seasonal variations. The other panels in Fig. 10 show that the negative bias in cloud fraction is accompanied by a positive bias in radiative heating rate—slightly elevated temperature—and a negative bias in relative humidity.

The initial cause of the negative cloud fraction bias is difficult to identify with certainty. However, CAM’s cloud fraction routine is a classic diagnostic scheme in which cloud fraction $C$ depends on relative humidity $RH$, via an assumed quadratic relationship, and hence temperature $T$. Figure 11 shows $C$ as a function of $T$ for a fixed specific humidity $q$. This plot shows several “exact” values of temperature, designated as $T$, with corresponding unbiased distributions that arise from unbiased heating rate fluctuations generated by McICA. Depending on the value of $T$, the associated mean cloud fraction $C$ obtained by integrating over the dis-

![Fig. 8. Same as in Fig. 5, except this is for LW CRE at TOA for the TWP region (B in Fig. 4) for the period SON.](image)

![Fig. 9. Differences relative to the REF simulation in annual- and ensemble-average cloud fractions in the lowermost model layer over land and ocean for the four McICA experiments. Shaded region indicates 95% confidence intervals.](image)
The distribution of $T$ can be equal to, less than, or greater than cloud fraction evaluated at $\bar{T}$. In a GCM, the time-integrated impact of these deviations would also depend on the distribution of $T$ and hence $C$. It would appear that for CAM, cases in which random errors in $T$ act to reduce cloud fraction have a dominant impact. Once a negative cloud fraction bias emerges, a positive feedback loop is induced: reduced cloud fraction $\rightarrow$ reduced radiative cooling $\rightarrow$ increased $T$ and reduced RH $\rightarrow$ further reductions in cloud fraction.

While this seems like a plausible explanation for the reduction in cloud fraction, it is still puzzling as to why it is confined to the layer adjacent to tropical oceans. Equally curious is how and why this bias stabilizes so abruptly after just 2 days. It is not that McICA noise suddenly stops after 2 days, but rather that the system seems to establish some sort of dynamic equilibrium in which local forcings from ongoing noise are balanced by the combined effects of other adjustments.

It is emphasized again that SSTs were prescribed in the simulations. For a model with interactive ocean, the reduction in low cloudiness would lead to increased SSTs, which might feed back on cloudiness, thereby either amplifying or attenuating the original response to McICA noise.

The fact that low-level tropical clouds are maintained, to a large extent, by local radiative cooling is demonstrated in Fig. 12. This plot shows zonal-mean cloud fraction in the lowest layer over oceans for the first week of the ordinary REF and 1COL simulations as well as for seven 1-week experiments (for both REF and 1COL) in which clouds in the lowest layer are neglected in radiation calculations. In the latter case, cloud fraction for REF is reduced by $\sim$0.12, or $\sim$80%, near the equator. An even larger absolute reduction in cloud fraction occurs near the North Pole where very low clouds (or fog) are extremely abundant in the CAM. Consequently, for the case in which clouds in the
lowest layer are radiatively inactive, the difference between 1COL and REF is very small, which further demonstrates the local impact of McICA on radiation fog–like cloud over the tropical ocean. Above the lowest layer, differences in zonal-mean cloud fractions between 1COL and REF are small (i.e., 0.015) regardless of whether the clouds in the lowest layer are included in the radiation calculations or not.

It is possible that the problem discussed in this subsection is specific to this GCM and its overabundance of clouds in the lowest layer. On the other hand, one should expect sensitivity to random errors in radiative heating to depend on model physical parameterizations [cf. Pincus et al. (2003), who observed no significant effects due to random noise]. Nevertheless, the results discussed here point to a mechanism, involving short-term interactions between clouds and radiative heating, through which unbiased radiative noise associated with McICA could lead to biases in simulated climate.

4. Summary and conclusions

The impact of unbiased radiative noise associated with the McICA approach was tested in 15-month ensemble simulations with the NCAR CAM using prescribed distributions of sea surface temperature and sea ice. Contrary to Pincus et al.'s (2003) results for the ECMWF global model, McICA’s sampling noise led to some statistically significant, albeit slight, impacts on the CAM’s simulated climate. The most distinct signal stemming from McICA-related noise was a reduction of cloud fraction in the lowest model layer over tropical oceans. This was accompanied by slight biases in cloud radiative effects and temperature. Additional runs with excessive sampling noise (1COL) indicated that the bias appears during the first few hours of simulation and essentially stabilizes after only a couple of days. This points to the role of short time-scale, local interactions between clouds and radiative heating as the seed of the bias.

Even though the largest effects due to McICA noise observed here are manifest in changes to clouds that are simulated poorly by the CAM, the results suggest that some attention should be paid to the magnitude of conditional random noise for heating rates as generated by McICA. Hence, different ensemble simulations were
performed with varying degrees of McICA sampling noise. It was found that differences relative to the reference simulation (REF) decrease monotonically with improved sampling and decreased noise. The noise reduction techniques devised by Räisänen and Barker (2004) proved quite successful. In particular, it was encouraging that the annual-mean climates of REF and McICA with least sampling noise (which allocated 50% more subcolumns than the base McICA to judiciously selected spectral intervals) were, statistically speaking, virtually indistinguishable.

In addition to assessing the impact of noise on mean climate, second-order structure functional analyses were performed in an attempt to elucidate the impact of noise on the evolution of simulated climate. Regardless of the magnitude of McICA noise, little evidence was found to suggest that the noise had statistically significant impacts, at least at the 90% confidence level, on the nature of fluctuations at periods between 1 and 30 days.

While impacts of sampling noise observed here exceed those in Pincus et al. (2003), the overall impression is that McICA is a viable method that allows for flexible introduction of unresolved cloud structure into GCM radiation calculations. However, the very fact that the impact of noise seems, understandably, to be model dependent implies that this impression cannot be extended safely as yet to all GCMs and simulations of all climatic regimes. It is therefore suggested that simulations addressing the impact of McICA noise be performed with several other GCMs with different physical parameterizations. Moreover, it is recommended highly that noise sensitivity tests be performed eventually with a coupled atmosphere–ocean GCM.

Acknowledgments. This study was supported by grants from the U.S. Department of Energy [Atmospheric Radiation Measurement (ARM) Grant DE-FG02-03ER63521] and the Modelling of Clouds and Climate Proposal funded through the Canadian Foundation for Climate and Atmospheric Sciences, the Meteorological Service of Canada, and the Natural Sciences and Engineering Research Council. P. Räisänen was also financed by the Academy of Finland (project: Treatment of Unresolved Cloud Structures for Earth System Models). We gratefully acknowledge R. Pincus, G. Stephens, and F. Zwiers for discussions.

REFERENCES


characteristics derived from long-term cloud radar data. J. Climate, 15, 2505–2515.


