Q. J. R. Meteorol. Soc. (1998), **124**, pp. 949–981

A cumulus parametrization with a prognostic closure

By DZONG-MING PAN* and DAVID. A. RANDALL
Colorado State University, USA

(Received 5 September 1996; revised 5 June 1997)

SUMMARY

The paper describes the introduction of a prognostic cumulus kinetic energy (CKE) as a replacement for the quasi-equilibrium closure hypothesis of Arakawa and Schubert (AS).

In the original version of the AS parametrization, the cloud work function, a measure of the convective available potential energy, is assumed to be maintained at 'small' values through a quasi-equilibrium between the cumulus convection and the 'large-scale forcing'. It is argued here, however, that the distinction between the convective and large-scale processes is ambiguous and subjective. It is demonstrated that the need for such a distinction can be avoided by relaxing the quasi-equilibrium assumption, through the introduction of a prognostic CKE; referred to as prognostic closure. A dimensional parameter, α, is introduced to relate the CKE to the square of the cloud-base convective mass flux. It is shown that 'adjustment time' defined by AS is related to α, so that when the adjustment time approaches zero the prognostic closure reduces to quasi-equilibrium closure. A second dimensional parameter, τ_D, is used to determine the rate at which the CKE is dissipated. In the limit of small α and τ_D, the convective mass flux is formally independent of both α and τ_D if the environmental sounding is assumed to be given, but in reality the results of a prognostic model do depend on these two parameters because they affect the time-dependent sounding.

For simplicity, a single constant value of α is used for all cloud types in tests with a general-circulation model, and this gives reasonably good results. Larger values of α lead to more frequent shallow cumulus convection and a cooler and more humid troposphere, in which stratiform condensation is more active and more large-scale precipitation can reach the surface. A longer dissipation time-scale leads to a warmer tropical troposphere. The interactions between stratiform cloudiness and convection prove to be quite important, leading to the conclusion that the convection parametrization really cannot be evaluated independently of the stratiform cloud parametrization with which it interacts.

KEYWORDS: Adjustment time Convections Cumulus kinetic energy Planetary boundary-layer

1. INTRODUCTION

Cumulus convection plays an essential role in the atmospheric general circulation. Large-scale numerical models, such as general-circulation models (GCMs), have grid cells with horizontal dimensions on the order of a hundred kilometres or more, and so cannot resolve the cumuli, which have diameters on the order of 1–10 km. Representation of the cumulus effects in terms of grid-scale variables is called cumulus parametrization. The simulated atmospheric circulation produced by a GCM is extremely sensitive to the formulation of the cumulus parametrization.

Variants of the cumulus parametrization proposed by Arakawa and Schubert (1974; hereafter AS) are being used in many GCMs today. Key elements of the AS parametrization are as follows:

• *The use of a convective mass flux to parametrize the vertical transports by the convective updraughts.* This idea was first proposed by Arakawa (1969). It has now been almost universally adopted, especially considering that moist convective adjustment can be formulated in terms of a mass flux (see Suarez *et al.* 1983).

• *The introduction of a simple but explicit model of a cumulus cloud as a conceptual component of the cumulus parametrization.* The particular cloud model used by AS to represent the life-cycle averaged properties of each cloud type consists of an entraining plume with a constant fractional entrainment rate, and detrainment only at the cloud-top level. Although this cloud model has been criticized by many authors (e.g. Warner 1970; Raymond and Blyth 1986), Lin (1994) has recently shown, using a cloud-resolving

* Corresponding author: Department of Atmospheric Science, Colorado State University, Ft. Collins, CO 80523, USA.

model, that the entraining-plume model works reasonably well for the purpose of cumulus parametrization. A discussion of the advantages and disadvantages of this type of cloud model is beyond the scope of this paper.

• *The introduction of a spectrum of cloud 'types', distinguished by their fractional entrainment rates.* In effect these cloud types differ according to their cloud-top heights, for a given sounding. Arakawa (1969) introduced a spectrum of three cloud types for use in a three-layer GCM; two of these cloud types started in the lowest model layer and topped-out in the middle and upper layer, respectively, while the third cloud type started in the middle layer and ended in the top layer. When the AS parametrization is applied in numerical models, cloud types are often characterized by their cloud-top pressures rather than their fractional entrainment rates (e.g. Lord *et al.* 1982).

• *A closure assumption based on the hypothesis that the convective clouds quickly act to convert whatever moist convective available potential energy is present in convectively active atmospheric columns into convective kinetic energy.* This idea is implicit in the 'moist convective adjustment' parametrization of Manabe *et al.* (1965), and was also used by Arakawa (1969). AS referred to this closure, as applied to a spectrum of cloud types sharing a common large-scale environment, as 'quasi-equilibrium', hereafter referred to as QE. As formulated by AS, QE closure is based on a distinction between convective and large-scale or non-convective processes.

• *An explicit coupling between the convective clouds and the planetary boundary-layer (PBL).* AS assumed that all convective clouds start from the PBL, and so drain mass out of the PBL, thus tending to reduce the PBL's depth. The cumulus updraughts were assumed to remove from the PBL air with the vertically averaged PBL values, so that the convection did not alter those average values.

• *Neglect of the effects of convective downdraughts.* Following the paper of AS, the importance of convective downdraughts, which AS had neglected, was quickly pointed out by Johnson (1978) and others, and several variants of the AS parametrization now include them.

• *Neglect of convection originating above the PBL top.* The AS parametrization did not include clouds whose bases occur above the top of the PBL; in some models that use the AS parametrization such upper-level cumuli are represented through the supplemental use of moist convective adjustment (e.g. Randall *et al.* 1989).

This paper reports on a prognostic closure which we have formulated as an alternative to the QE closure. We introduce a prognostic cumulus kinetic energy (CKE) for each cloud type, thus endowing the cumulus convection itself with a memory, over and above the memory represented by the prognostic variables for the large-scale atmospheric state. We show that the prognostic closure has both conceptual and practical advantages over the QE closure.

The paper is organized as follows. Following a brief review of the AS parametrization in section 2, section 3 explains the concepts underlying the prognostic convective closure. We then present the results of some preliminary tests in a GCM. A brief description of the model itself is given in section 4. The tests, which are discussed in section 5, show that the radiatively active clouds of the model are quite sensitive to the convective closure, and because of this we find it necessary, as discussed in section 6, to change the way in which radiatively active cloud formation is parametrized. After making this change, we performed additional simulations using the prognostic closure, including some sensitivity tests. These further calculations are discussed in sections 7 and 8. The paper closes with a summary and conclusions.

2. A BRIEF REVIEW OF THE AS PARAMETRIZATION

In a large-scale model, such as a GCM, many cumuli can co-exist within a single grid cell. A cumulus parametrization deals with only the collective effects of such a cumulus ensemble. The space- and time-scale separations between the cumuli and the resolvable large-scale motions make parametrization possible. Let an overbar denote a suitable large-scale average, and let a prime denote a departure from this average. AS wrote the grid-cell average budget equations for energy and moisture as:

$$\rho\frac{\partial}{\partial t}\bar{s} = -\overline{\nabla\cdot(\rho s\mathbf{V})} - \frac{\partial}{\partial z}(\rho\overline{sw}) + Q_1, \tag{1}$$

$$\rho\frac{\partial}{\partial t}\bar{q} = -\overline{\nabla\cdot(\rho q\mathbf{V})} - \frac{\partial}{\partial z}(\rho\overline{qw}) - \frac{Q_2}{L}. \tag{2}$$

Here ρ is the density of the air; $s \equiv c_pT + gz$ is the dry static energy, where T is temperature, g is gravity and c_p is the specific heat at constant pressure; q is the water vapour mixing ratio; w is the vertical velocity; $\mathbf{V}$ is the horizontal velocity; L is the latent heat of evaporation; and Q_1 and Q_2 are the 'apparent heat source' and 'apparent moisture sink' defined by Yanai *et al.* (1973):

$$Q_1 \equiv Q_R + LC - \frac{\partial}{\partial z}(\rho\overline{w's'}), \tag{3}$$

$$Q_2 \equiv LC + L\frac{\partial}{\partial z}(\rho\overline{w'q'}) \tag{4}$$

where Q_R is the radiative heating rate and C is the net condensation. We assume that the vertical eddy transports above the boundary layer are mainly due to the effects of cumulus convection. As discussed by Arakawa and Chen (1987), the key problem of cumulus parametrization is to provide closure for Eqs. (1) and (2), which include the four unknowns $(\partial/\partial t)\bar{s}$, $(\partial/\partial t)\bar{q}$, $(Q_1 - Q_R)$, and Q_2.

Consider a fractional area, σ, of a GCM grid cell that is covered by convective updraughts of an ensemble of cumulus clouds. The convective flux of dry static energy by the cumulus ensemble can be written as:

$$\rho\overline{w's'} = M_c(s_c - \bar{s}) \tag{5}$$

where

$$M_c \equiv \rho\sigma w_c \tag{6}$$

is the convective mass flux. Here we consider only a single cloud type, for simplicity. The approximations in (5) and (6) are valid provided that $\sigma \ll 1$. Similarly, the convective moisture flux is given by

$$\rho\overline{w'q'} = M_c(q_c - \bar{q}). \tag{7}$$

By using (5), (7), as well as the budget equations for the clouds themselves, assuming that the horizontal eddy transport is negligible, and neglecting storage of dry air, moisture, and static energy in the cloud ensemble, AS rewrote (1) and (2) as:

$$\rho\frac{\partial}{\partial t}\bar{s} = D_s + (M_c - \rho\overline{w})\frac{\partial}{\partial z}\bar{s} - \rho\overline{\mathbf{V}}\cdot\nabla\bar{s} + \overline{Q_R}, \tag{8}$$

and

$$\rho\frac{\partial}{\partial t}\bar{q} = D_q + (M_c - \rho\overline{w})\frac{\partial}{\partial z}\bar{q} - \rho\overline{\mathbf{V}}\cdot\nabla\bar{q}. \tag{9}$$

Equations (8) and (9) show explicitly how the cumulus clouds modify their environment. The first terms on the right-hand sides of these equations, D_s and D_q (functions of detrained mass flux and cloud properties), represent the effects of detrainment on $\bar{s}$ and $\bar{q}$. The second terms represent the effects of cumulus-induced environmental subsidence.

AS introduced a simple cumulus-cloud model, essentially describing the life-cycle-averaged properties of a cumulus updraught. The cumulus ensemble was assumed to consist of a collection of subensembles, each composed of convective updraughts characterized by different fractional rates of entrainment, λ, all originating near the top of the PBL, and all sharing a common large-scale environment. Given the cloud-base conditions and the environmental sounding, the fractional entrainment rate determines the in-cloud moist static energy and mixing ratio. From this in-cloud sounding we can determine the neutral-buoyancy level, at which the clouds are assumed to detrain, as well as the properties of the detrained air. Larger entrainment rates lead to more rapid dilution and so lower cloud tops, for a given sounding. Using such a cloud model we can obtain the in-cloud sounding for each subensemble. The remaining problem is to determine the convective mass flux, M_c.

Cumulus convection occurs as a result of moist convective instability, in which the potential energy of the mean state is converted into the kinetic energy of cumulus convection. AS defined the 'cloud work function', A, for a cumulus subensemble, as a vertical integral of the buoyancy of the cloud air with respect to the large-scale environment:

$$A(\lambda) = \int_{z_B}^{z_D(\lambda)} \frac{g}{c_p \bar{T}(z)} \eta(z, \lambda)\{s_{vc}(z, \lambda) - \bar{s}_v(z)\}\, dz. \tag{10}$$

Here z_D is the height of the detrainment level; z_B is the cloud base, which is assumed to be at the top of the PBL for all types of clouds; η is the normalized cloud mass flux, satisfying $M_c = \eta M_B$; M_B is the cloud-base mass flux; $\bar{s}_v = \bar{s} + c_p \bar{T}(0.608\bar{q} - \bar{l})$ denotes the virtual static energy; $\bar{l}$ is the mixing ratio of liquid water; and subscript c denotes the in-cloud sounding, taking into account dilution by the entrainment of environmental air at each level. From (10) we see that the function $A(\lambda)$ is a property of the large-scale environment. A positive value of $A(\lambda)$ means that a cloud with fractional entrainment rate λ can convert the potential energy of the mean state into convective kinetic energy. For $\lambda = 0$, $A(\lambda)$ is essentially identical to the convective available potential energy (CAPE), as conventionally defined.

Numerical models use (8) and (9) to prognosticate $\bar{T}(z)$ and $\bar{q}(z)$, from which $A(\lambda)$ can be determined; therefore, these models indirectly forecast $A(\lambda)$. By taking the time derivative of (10), and using (8) and (9), AS showed that

$$\frac{d}{dt} A(\lambda) = J M_B(\lambda) + F(\lambda), \tag{11}$$

where M_B is the non-negative cloud-base mass flux; F is the 'large-scale forcing', which represents the rate of increase of cloud work function due to non-convective processes; and $J M_B$ includes all of the terms involving M_B. Equation (11) is written in simplified form. The $J M_B$ term actually represents an integral over all cloud types, and is written as a product here merely to simplify the discussion. The quantity J symbolically represents the kernel of the integral, which is a property of the large-scale sounding (see AS for details). The $J M_B$ term of (11) tends to reduce $A(\lambda)$, because cumulus convection stabilizes the environment, so that J is usually negative. Keep in mind that (11) holds for each cumulus subensemble. As discussed later, (11) depends on a subtle assumption which, we argue, represents a conceptual weakness of the QE closure.

AS assumed QE of the cloud work function, i.e.

$$\frac{\mathrm{d}}{\mathrm{d}t}A(\lambda) = JM_{\mathrm{B}}(\lambda) + F(\lambda) \cong 0. \tag{12}$$

This QE closure assumption means that the moist convective instability generated by the large-scale forcing, $F(\lambda)$, is very rapidly consumed by cumulus convection, i.e. the two terms on the right-hand side of (11) approximately balance each other. The QE hypothesis implies that cumulus clouds lack any memory of their past history. Instead, they closely follow the large-scale forcing, quickly reacting to changes in the large-scale forcing in a very short 'adjustment time', like a basketball player in a man-to-man defence. The QE hypothesis is valid if the *time-scale for changes* in $F(\lambda)$ is much longer than the adjustment time, so that the convection can keep up with the changes in $F(\lambda)$. AS estimated that the adjustment time is on the order of 10^3 to 10^4 s, so that use of (12) would be justified for the simulation of fluctuations whose time-scales are on the order of one day or longer. The hypothesis of cloud work function QE has been observationally tested by AS, Lord and Arakawa (1980), Lord (1982), Kao and Ogura (1987), and Grell *et al.* (1991), among others, and has held up well.

The significance of (12) is that it can be solved as a Fredholm integral equation for the convective mass flux as a function of cloud type, λ. After discretization this leads to a system of linear equations (Lord *et al.* 1982). Although the system is linear, the mass flux distribution function, $M_{\mathrm{B}}(\lambda)$, is required to be non-negative for all λ. This cannot be guaranteed without making additional assumptions (e.g. Hack *et al.* 1984).

A key assumption in writing (11), and thus in formulating the QE closure, is that atmospheric processes can be divided into the large-scale forcing and the convective response. Can such a distinction really be cleanly drawn? Stratiform clouds usually cover a large area and last for hours or even days, so it might seem reasonable to consider the heating and drying associated with them as part of the large-scale processes; this is what AS did. The distinction is not so clear, however, when we consider the interactions between stratiform clouds and cumulus convection. Stratiform clouds are often directly produced by cumulus detrainment (e.g. Rutledge and Houze 1987), and this convective generation of stratiform clouds can occur on time-scales of a few hours. From this point of view it seems reasonable to consider the heating and drying associated with the stratiform clouds to be part of the 'convective processes'. Similar ambiguities arise in connection with the surface sensible- and latent-heat fluxes, which certainly can occur in the absence of cumulus convection, but which can be enhanced by the effects of cumulus convection, on short time-scales (e.g. Jabouille *et al.* 1996). Should these enhanced surface fluxes be considered part of the large-scale forcing, or part of the convective response?

The point is that the QE closure, (11), can only be applied after the convective processes have been conceptually separated from the non-convective processes. The subjective nature of such a distinction, as pointed out earlier, is therefore a conceptual weakness of the QE closure. To avoid this problem we have to avoid defining the large-scale forcing and convective response, while at the same time finding a useful alternative closure assumption for the convection parametrization. One of the main purposes of this paper is to show how this can be done. A second purpose of this paper is to show how the AS parametrization can be significantly simplified without losing its basic idea.

3. Prognostic closure

The starting point for our prognostic closure is the formulation of a set of prognostic equations governing the vertically integrated CKE per unit area, for each cloud type.

Similar equations were discussed by Lord and Arakawa (1980). We begin with the eddy kinetic-energy equation (e.g. Stull 1988) which can be written, using Cartesian tensor notation, as

$$\frac{\partial}{\partial t}(\rho_0\overline{e}) + \frac{\partial}{\partial x_j}(\rho_0\overline{U_j}\overline{e} + \rho_0\overline{u_j u_i^2} + \overline{u_i p'}) = \delta_{i3}\frac{g}{\overline{\theta_v}}\rho_0\overline{u_i\theta_v} - \rho_0\overline{u_i u_j}\frac{\partial}{\partial x_j}\overline{U_i} + \varepsilon_V. \quad (13)$$

Here $\overline{U_j}$ and u_j are the mean and eddy velocities, respectively; $e \equiv \frac{1}{2}u_i^2$ is the eddy kinetic energy per unit mass; ρ_0 is the basic-state density; p is pressure; θ_v is the virtual potential temperature; δ_{i3} is the Kronecker delta; and ε_V is the viscous dissipation rate. We assume here that the perturbations governed by (13) are predominantly due to cumulus convection, although of course both boundary-layer turbulence and gravity waves may co-exist with the convective circulations, at some levels. The left-hand side of (13) includes the local time-rate-of-change, as well as transport terms due to the mean flow, the triple-correlations of the velocity components, and pressure–velocity correlations. The first term on the right-hand side of (13) is the buoyancy generation term, which is expected to be positive when cumulus convection is active. The second term is the shear generation term, which can be either positive or negative.

We define the 'bulk' CKE as the vertical integrated eddy kinetic energy per unit area:

$$K \equiv \frac{1}{2}\int_{z_S}^{z_T} \rho_0\overline{u_i^2}\,\mathrm{d}z, \quad (14)$$

where z_T and z_S are the heights of the highest cloud top and the earth's surface, respectively. Note that, with this definition, the bulk CKE includes all of the kinetic energy associated with all branches of the convective circulations, including the updraughts, the downdraughts, and the horizontal currents that connect them together. As discussed later, numerical simulations by Xu and Arakawa (1992) suggest, perhaps surprisingly, that the horizontal branches of the convective circulations actually contribute a large fraction of the CKE.

A prognostic equation for the bulk CKE can be derived by vertically integrating (13), which gives

$$\frac{\partial K}{\partial t} = B + S - D, \quad (15)$$

where

$$B \equiv \int_{z_S}^{z_T} \delta_{i3}\frac{g}{\overline{\theta_v}}\rho_0\overline{u_i\theta_v}\,\mathrm{d}z$$

is the buoyant production term,

$$S \equiv -\int_{z_S}^{z_T} \rho_0\overline{u_i u_j}\frac{\partial}{\partial x_j}\overline{U_i}\,\mathrm{d}z$$

is the shear production term, and $D \equiv \int_{z_S}^{z_T} \varepsilon_V\,\mathrm{d}z$ is the vertically integrated dissipation rate. In deriving (15) we have assumed that the various transport terms on the left-hand side of (13) vanish when integrated through the depth of the convective layer; these neglected terms could represent, for example, a sink of convective kinetic energy due to the production of a gravity-wave energy flux by the convection. We also define a dissipation time-scale by

$$\tau_D \equiv \frac{K}{D}. \quad (16)$$

Lord and Arakawa (1980) pointed out that τ_D cannot be much longer than a cloud lifetime, i.e. about 10^3 s.

The importance of vertical wind shear for convection has been reported in many studies (e.g. Asai 1964; Seitter and Kuo 1983; Lilly and Jewett 1990; Xu, 1991). Shear can potentially influence the convection in a variety of ways. One of the simplest is through the shear term of (15). Wu and Moncrieff (1996) showed that shear generation of CKE can be comparable with buoyancy generation, when the latter is computed by taking into account the loading due to condensed water. Nevertheless, in the present study we neglect the shear-production term of (15), as well as the other potential effects of shear on the convection. Parametrizing the effects of shear on moist convection is an interesting and challenging area for future research.

With these simplifying assumptions, (15) reduces to

$$\frac{\partial K}{\partial t} = B - \frac{K}{\tau_D}. \tag{17}$$

We now assume that an equation similar to (17) applies for each cumulus subensemble, i.e.

$$\frac{\mathrm{d}}{\mathrm{d}t}K(\lambda) = B(\lambda) - \frac{K(\lambda)}{\tau_D(\lambda)}. \tag{18}$$

This assumption can be broken down into two sub-assumptions, as follows. First, we assume that perturbation velocity and virtual potential-temperature distributions can be defined for each subensemble. Each convective cloud is considered to be associated with a cellular circulation, so that the complicated convective motions filling the large-scale space represent the superposition of many individual convective cells. In effect this was already assumed by AS. Second, we assume that direct dynamical interactions among the subensembles are negligible in the sense that the kinetic-energy budget of each subensemble can be formulated without making reference to the other subensembles, i.e. without 'cross-talk'. This would be the case, for example, if the convective circulations 'tiled the plane', as Rayleigh–Benard convection does (e.g. Chandrasekhar 1961). Obviously this assumption can only be true in an approximate sense.

Finally, we rewrite the buoyancy term of (18) in terms of the cloud work function and the cloud-base mass flux, to obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}K(\lambda) = M_B(\lambda)A(\lambda) - \frac{K(\lambda)}{\tau_D(\lambda)}. \tag{19}$$

Here $M_B(\lambda)A(\lambda) = B(\lambda)$, the rate of CKE generation by buoyancy, for cloud type λ, as discussed by AS. Equation (19) predicts the CKE using the cloud work function and the cloud-base mass flux; while Eq. (11) effectively determines the time-rate-of-change of the cloud work function, for a given cloud-base mass flux. Equations (19) and (11) together contain three unknowns, namely $(\mathrm{d}A/\mathrm{d}t)$, K, and M_B. We need one more equation to close this problem.

To derive a relation between M_B and K we consider a unit horizontal area of which a small fraction σ is covered by cumulus updraughts. The area outside the cumulus cloud undergoes compensating subsidence. Let subscripts u and d represent upward and compensating downward motions, respectively. We have

$$\overline{w} = \sigma w_u + (1-\sigma)w_d, \tag{20}$$

and therefore

$$\overline{w'^2} = \sigma(w_u - \overline{w})^2 + (1-\sigma)(w_d - \overline{w})^2 \cong \sigma(w_u)^2, \tag{21}$$

provided that $\sigma \ll 1$. Using (21) and (6), we find that

$$\frac{1}{2}\rho\overline{w'^2} \cong \frac{M_B^2\eta^2}{2\rho\sigma}. \tag{22}$$

Here we have used $M_c = \eta M_B$. Following Arakawa and Xu (1990) we can define a parameter, α, with dimensions of (length)^4(mass)$^{-1}$, by

$$K = \alpha M_B^2. \tag{23}$$

We can interpret α as a 'conversion factor' relating the cumulus mass flux to the CKE. From (22) and (23), we see that

$$\alpha \equiv \frac{1}{2\varepsilon}\int_{z_S}^{z_T} \frac{\eta^2}{\rho\sigma}\,dz, \tag{24}$$

where

$$\varepsilon \equiv \frac{1}{K}\int_{z_S}^{z_T} \frac{1}{2}\rho\overline{w'^2}\,dz \tag{25}$$

is the fraction of the total kinetic energy that comes from the vertical component of the velocity. Equation (24) shows that α depends on ε, σ, and the depth of the clouds.

The parameters η and z_T can be obtained using the cloud model. The simplest way to close Eqs. (23), (19), and (11) is to assume α = constant. The alternative is to formulate ε and σ in terms of K, M_B, A, and presumably other large-scale variables, in which case (24) could be used to determine α.

To estimate the value of α we use Eq. (24) with cumulus ensemble model (CEM) output of Xu (1991). The CEM was developed by Krueger (1988); it has a domain size comparable with a grid box of a GCM, but resolves individual clouds explicitly, in a two-dimensional framework. Xu (1991) used a prescribed periodic large-scale advective cooling and moistening rate as large-scale forcing for the CEM. Details of his experiment are described by Xu *et al.* (1992). The eddy kinetic energy diagnosed from the CEM output includes all eddies over the whole domain, since it is difficult to distinguish individual cloud types in the CEM output. As a result our computation yields a 'bulk' value of α. The computed time series of ε is shown in Fig. 1, from which we see that ε is on the order of 10^{-2} or larger. The slow oscillation apparent in the figure (approximately six cycles during the 7 days) is due to the periodic large-scale forcing prescribed in Xu's experiment. Smaller values of ε occur when the convection is stronger. If we choose $\sigma \sim 0.01$, $\eta \sim 1$ (for non-entraining clouds), and $\rho \sim 1$ kg m^{-3} or smaller, and integrate over a depth of 10^4 m, Eq. (24) gives $\alpha \sim 10^8$ m^{4}kg^{-1} or larger.

Equations (19), (23), and (11) have the steady-state solution:

$$M_B = -\frac{F}{J}, \tag{26}$$

$$A = -\frac{\alpha}{\tau_D}\frac{F}{J}, \tag{27}$$

$$K = \alpha\left(\frac{F}{J}\right)^2. \tag{28}$$

Note that in (26), (27), and (28), F and J appear only in the ratio (F/J). Because (26) can be obtained directly from (11) with the QE closure, we conclude that the QE closure

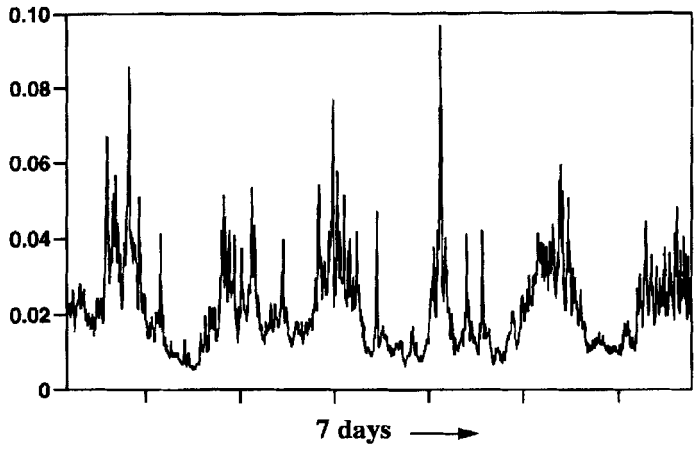

Figure 1. Time series of ε (see text) from a cumulus-ensemble-model run.

can be interpreted as a 'steady-state' limiting case of our prognostic closure. Because (26) does not involve α or τ_D, it *appears* that the convective mass flux should be independent of these two parameters. That would be important if true, because we currently lack a theory to determine α and τ_D. Note, however, that F and/or J could depend implicitly on α or τ_D; later we present evidence that this is in fact the case. The steady-state cloud work function and CKE, given by (27) and (28), do depend explicitly on the values of α and τ_D. Because we have no theory to determine α and τ_D, we do not claim that we can actually predict the correct values of the cloud work function or CKE. We do claim, however, that we can predict physically meaningful values of M_B. In preliminary tests we found that our results depend only on the ratio α/τ_D. Since 10^3 s seems to be a reasonable value for τ_D, we will focus only on the sensitivity of our results to α. Later we present numerical experiments to show the effects of varying α.

Xu and Arakawa (1992) found that shear acts to increase the fraction of the CKE in the horizontal components of the motion, and to decrease the fraction in the vertical component, so that shear reduces ε and increases α. For a given value of the CKE, Eq. (23) gives a smaller M_B when α is larger. Since M_B is the agency through which the convection modifies its environment, shear inhibits the feedback of the convection on the mean flow. In particular, shear inhibits convective stabilization of the environment, allowing larger values of the cloud work function (more CAPE) to build up in response to the large-scale forcing.

To examine further the difference between the prognostic closure and the QE closure, we can estimate the adjustment time in the framework of the prognostic closure. Combining (19), (23), and (11), and assuming that J and α are independent of time, we get

$$\frac{d^2M_B}{dt^2} + \frac{1}{2\tau_D}\frac{dM_B}{dt} - \frac{J}{2\alpha}M_B = \frac{F}{2\alpha}, \tag{29}$$

which describes a damped oscillation. We can derive a similar equation for A:

$$\frac{d^2A}{dt^2} + \frac{1}{2\tau_D}\frac{dA}{dt} - \frac{J}{2\alpha}A = \frac{F}{2\tau_D} + \frac{dF}{dt}. \tag{30}$$

The second term on the left-hand side of (29) is the damping term, with a time-scale of

$2\tau_D$. In the limit $\tau_D \to \infty$, with $F = 0$, (29) reduces to

$$\frac{d^2 M_B}{dt^2} - \frac{J}{2\alpha} M_B = 0, \tag{31}$$

which describes free oscillations about $M_B = 0$, because $(J/\alpha) < 0$. The condition that $M_B \geq 0$ implies that these oscillations will halt, however, as soon as M_B has decreased to zero. The time-scale for this to occur is proportional to $(\alpha/|J|)^{1/2}$. We can thus interpret $(\alpha/|J|)^{1/2}$ as τ_{adj}, the 'adjustment time', defined by AS as the time required for convective processes to reduce A to zero in the absence of large-scale forcing. This shows that τ_{adj} is closely related to α. Larger values of α correspond to longer adjustment times. So long as τ_D and τ_{adj} are small enough, QE of the cloud work function is a good approximation. More discussion is given by Randall and Pan (1993).

The AS parametrization and our prognostic closure are summarized using a flow chart in Fig. 2, which illustrates the interactions between cumulus convection and other processes. The box in the upper half of the figure includes all the non-convective (large-scale) processes, such as PBL turbulence, radiation (Q_R), advection, and the effects of stratiform clouds. Given a sounding of temperature and mixing ratio ($\overline{T}_o, \overline{q}_o$), without convection, the large-scale processes cause the sounding to change to ($\overline{T}_f, \overline{q}_f$). The corresponding change of the cloud work function is from A_o to A_f, and the difference $(A_o - A_f)$ is the so-called 'large-scale forcing' F. The kernel, J, can be obtained from the initial sounding. With F and J known, M_B and hence the cumulus heating, moistening, and precipitation rate ($\delta\overline{T}_{cu}$, $\delta\overline{q}_{cu}$, and *precip.*) can be obtained using the QE assumption. This is indicated in the shaded area. Note that here we classify the stratiform clouds with the large-scale processes, as did AS. With the prognostic closure (indicated by the thick solid-arrow), however, all the computations in the shaded area (including those of the large-scale forcing and the kernel) can be dropped. The prognostic closure eliminates any need to define or calculate F and J, and so sidesteps the ambiguity involved in the separation between the cumulus response and the large-scale forcing.

Before going on, we summarize briefly the practical application of the prognostic closure, and contrast it with the practical application of QE. The 'cloud model' and 'feedback' components of the parametrization are essentially identical between the two approaches; only the method used to determine the cloud mass flux distribution function, $M_B(\lambda)$, is different. With QE closure, $M_B(\lambda)$ is determined by solving a discretized form of the Fredholm integral equation discussed by AS, subject to the condition that $M_B(\lambda) \geq 0$ for all λ. In order to solve the integral equation it is necessary to evaluate the kernel matrix and the large-scale forcing. With our prognostic closure the integral equation does not have to be solved, and so the kernel and the forcing do not have to be evaluated (or even defined). With the prognostic closure we give initial conditions for $K(\lambda)$, and then compute $M_B(\lambda)$ using (23). The in-cloud sounding and cloud work function can then be evaluated using the cloud model, and the feedback on the large-scale state can then be determined using the methods of AS. Finally, new values of $K(\lambda)$ are predicted using (19).

4. Model description and experiment design

Pan (1995) conducted tests of the prognostic closure in a one-dimensional model forced with GATE* data; a discussion of these one-dimensional tests is omitted here for brevity. We now proceed directly to a discussion of the results of tests of the prognostic

* GARP (Global Atmospheric Research Program) Atlantic Tropical Experiment, a joint World Meteorological Organization/International Council of Scientific Unions project.

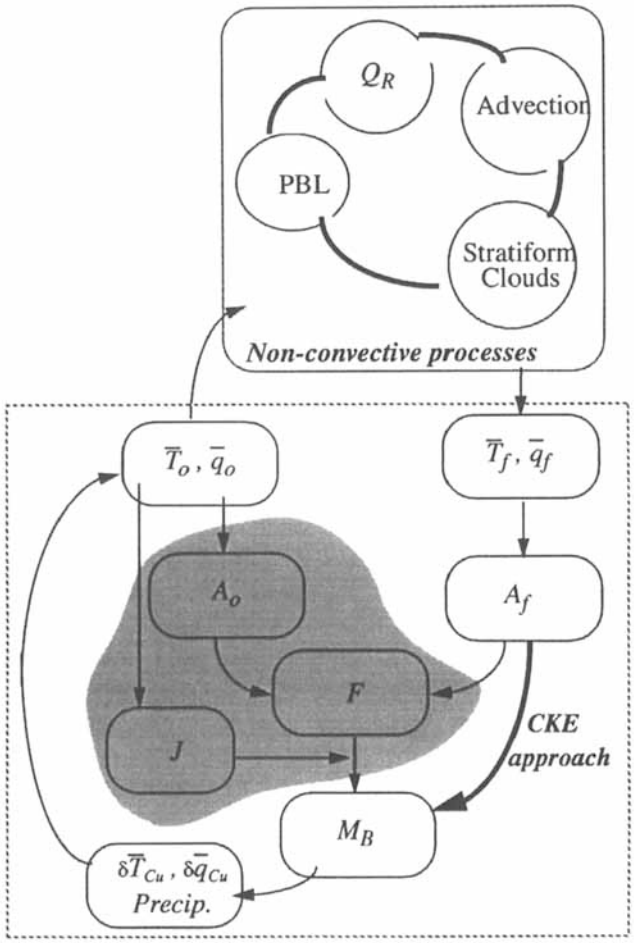

Figure 2. Flow chart of the Arakawa–Schubert cumulus parametrization. The upper box represents the large-scale (non-convective) processes, including stratiform clouds, advection, radiation (Q_R), and planetary boundary-layer (PBL) processes. $\overline{T}_o$, $\overline{q}_o$ and $\overline{T}_f$, $\overline{q}_f$ are the temperature and mixing ratio, respectively, before and after the large-scale processes. A_o and A_f are the cloud work functions, and the difference ($A_o - A_f$) is the 'large-scale forcing' F. J is the 'kernel', M_B is the cloud-base mass flux, $\delta\overline{T}_{cu}$, $\delta\overline{q}_{cu}$, and *Precip.* denote the cumulus heating, moistening, and precipitation rate, respectively, and CKE is the cumulus kinetic energy. See text for explanation of shaded area.

closure in the three-dimensional Colorado State University (CSU) GCM. Except as otherwise described below, the version of the CSU GCM used here is essentially the same as that used by Randall *et al.* (1991). It is an offshoot of the 1982 version of the University of California at Los Angeles (UCLA) GCM. The governing equations are finite-differenced, using highly conservative schemes (Arakawa and Lamb 1977, 1981; Arakawa and Suarez 1983). The simulations discussed in this paper were made with a grid spacing of 4° of latitude by 5° of longitude, with 17 layers. The prognostic variables of the GCM are potential temperature, water vapour mixing ratio, horizontal velocity components, surface pressure, depth of the PBL and the PBL's vertically averaged turbulence kinetic energy, ground temperature and snow depth at land points, and the ice temperature at land-ice and sea-ice points. The top of the PBL is an internal coordinate surface of the model, marking the top of the model's lowest layer (Suarez *et al.* 1983; Randall *et al.* 1985). Stratocumulus clouds are assumed to be present in the PBL whenever the temperature and mixing ratio at the PBL

TABLE 1. SUMMARY OF THE SIMULATIONS

Run	Closure	Anvils
BQ	Quasi-equilibrium ($\alpha \to 0$)	Binary
BK	$\alpha = 10^8$ m^{4}kg^{-1}, $\tau_D = 600$ s	Binary
FK8	$\alpha = 10^8$ m^{4}kg^{-1}, $\tau_D = 600$ s	Fractional
FK9	$\alpha = 10^9$ m^{4}kg^{-1}, $\tau_D = 600$ s	Fractional
FK7	$\alpha = 10^7$ m^{4}kg^{-1}, $\tau_D = 600$ s	Fractional
FKT	$\alpha = 10^8$ m^{4}kg^{-1}, $\tau_D = 1200$ s	Fractional

For further explanation see text.

top (as determined by a mixed-layer assumption) correspond to supersaturation, provided that cloud-top entrainment instability does not occur. The prescribed boundary conditions of the GCM include realistic topography, and the observed climatological seasonally varying global distributions of sea surface temperature and sea-ice thickness. We also prescribe the soil properties and the seasonally varying morphological and physiological parameters characterizing the land-surface vegetation.

The parametrized PBL stratocumulus clouds interact with the solar and terrestrial radiation (Harshvardhan *et al.* 1989). In addition, the GCM includes parametrizations of the hydrological and radiative effects of stratiform cloudiness in the free atmosphere. One of the themes of this paper is that the cumulus convection parametrization interacts very strongly with the parametrized stratiform cloudiness in the free atmosphere. We used several parametrizations of free-atmospheric stratiform cloudiness in the course of this study, and they are discussed at appropriate places in the remainder of this paper.

In the 'control' version of the GCM used in this study, radiatively active cloudiness can be associated with large-scale saturation, with PBL stratocumulus clouds, and/or with optically thick upper tropospheric clouds associated with cumulus detrainment, hereafter referred to as 'anvils', following the approach of Harshvardhan *et al.* (1989). When and where cloudiness occurs, the clouds are assumed to fill the entire grid box. As our study progressed, the response of our model to changes in the convection parametrization forced us to make improvements to the parametrization of the anvil clouds. The various anvil parametrizations used are discussed later.

The other physical parametrizations of the model include moist convective adjustment to represent cumulus convection originating above the PBL, and dry convective adjustment.

A summary of the simulations discussed in this paper is given in Table 1. The control run, hereafter referred to as BQ, used the original AS parametrization with the cloud-work-function QE closure (as implemented by Lord *et al.* (1982)). The anvils were assumed to cover the entire grid box, so that their cloud fractions are either zero or one; for this reason we refer to them as 'binary' anvils. All of the other runs used the prognostic closure. As discussed later, the values of α varied among the various runs with the prognostic closure. With one exception, the runs with the prognostic closure used $\tau_D = 600$ s as the dissipation time-scale for the cumulus kinetic energy; the run using $\tau_D = 1200$ s we call FKT.

All simulations are for January conditions. The first month (December) is treated as a spin-up, and only the results of the second month are used. To obtain statistically significant results for monthly means, we performed an 'ensemble' of five realizations, each based on a 2-month run. These realizations are identical except for small perturbations of the initial temperature field (of order 0.5 K) at each grid point. An ensemble average is obtained by averaging the five runs together. In the following, when we discuss the results of a particular run, e.g. the FK8 run, we are actually referring to the corresponding ensemble of five runs.

5. Initial tests of the prognostic closure

The BQ run used a version of the AS cumulus parametrization that employs the QE closure of AS, as implemented by Lord *et al.* (1982). The BK run is identical to the BQ run except that it uses the prognostic closure described earlier in this paper.

In both the BQ and BK runs a conventional 'large-scale precipitation' (LSP) parametrization was used, in which condensation is assumed to occur when the relative humidity exceeds 100%, and falling rain is allowed to evaporate into unsaturated layers on the way down. Radiatively active anvils were assumed to occur whenever convection penetrated above the 500 mb level. The anvils have their bases at 500 mb and their tops at the highest model level reached by convection. The short-wave optical thickness of the anvils depends only on the physical depth of the convection above this level, as discussed by Harshvardhan *et al.* (1989). The long-wave optical depth of the anvil clouds is assumed to be proportional to the short-wave optical thickness, and is related to the long-wave emittance, which is given by a simple exponential formula (Harshvardhan *et al.* 1989). The binary anvil parametrization and LSP were also used, along with the QE closure, in the simulations described by Randall *et al.* (1989), and in the 1991 AMIP (Atmospheric Model Intercomparison Project; Gates 1992) simulation with the CSU model. In fact, the BQ version of the GCM as discussed in this paper is very similar to the CSU 91 AMIP version.

Randall *et al.* (1989) conducted GCM experiments to examine effects of anvils, with the QE closure. If the long-wave effect of the anvils dominates, the effect is to warm the atmospheric column, which favours large-scale rising motion, leading to more convection. If the short-wave effect of the anvils dominates, increased anvil coverage reduces the surface evaporation rate over moist land, and so tends to inhibit convection. The results obtained by Randall *et al.* (1989) depend on the anvil formation parametrization and also on the closure assumption which is used to determine the convective intensity. In their QE closure, AS included the radiative effects of anvil clouds as part of the large-scale forcing, but as discussed earlier this choice is subjective. With our prognostic closure the large-scale forcing does not have to be defined at all, so the point becomes moot. For this reason we were particularly interested to see how the interactions between convective and stratiform clouds would change when the QE closure was replaced by the prognostic closure.

Figure 3 shows the January total precipitation rates from the BQ and BK runs, the observations of Legates and Willmott (1990), and 9-year average data from the Global Precipitation Climatology Project (GPCP; Xie and Arkin 1996). The general pattern is similar for both runs and the two sets of observations. Substantial differences between the two runs do occur in the global means, however. The prognostic closure produces a global-mean January precipitation rate of 3.33 mm day^{-1}, which is a 7% decrease relative to the control run (3.58 mm day^{-1}). The observed global-mean January precipitation rate is 3.63 mm day^{-1} according to Legates and Willmott (1990) and only 2.58 mm day^{-1} from the GPCP. All four maps show intertropical convergence zone (ITCZ) precipitation maxima, and subtropical precipitation minima. In middle latitudes, precipitation maxima occur in the storm tracks off the east coasts of Asia and North America. The observed maxima over the Amazon basin and South Africa are not produced by either of the simulations; these observed maxima are more realistically captured in other simulations discussed later.

Compared with the QE closure, the prognostic closure produces generally weaker precipitation in the tropics, except for the tropical western Pacific. Increases of the simulated precipitation rate occur in some mid-latitude locations. From BQ to BK, the zonal-mean precipitation decreases in the tropics but increases in the mid-latitudes. The zonal-mean January precipitation rates from the two simulations and the observations are shown in Fig. 4, along with the differences between the two runs in terms of the cumulus precipitation

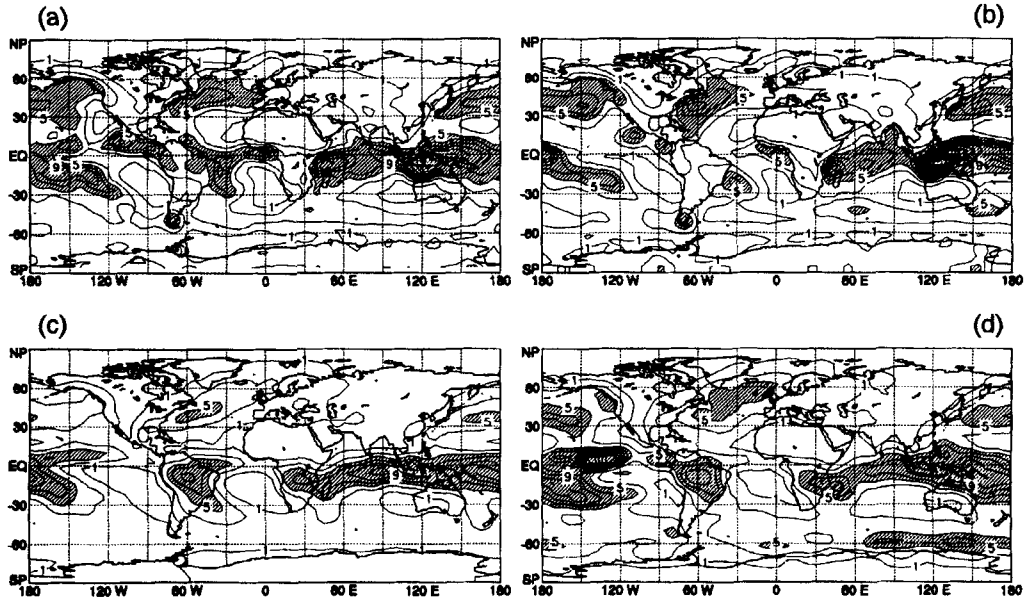

Figure 3. January monthly-mean total precipitation rates. (a) The control run, (b) the BK run (see text), (c) GPCP data (Xie and Arkin 1996), and (d) the observations of Legates and Willmott (1990). The contour interval is 2 mm day^{-1}. Areas with precipitation rates larger than 5 mm day^{-1} are shaded.

(CUP) rate and the large-scale precipitation (LSP) rate. For both runs the ITCZ precipitation maxima occur between the equator and 10°S, as would be expected for January conditions. The prognostic closure reduces the precipitation rate everywhere between 30°N and 30°S. Figure 4 shows that the CUP rate decreases at almost all latitudes (mostly in the tropics) while the LSP rate changes very little in the tropics, but mostly increases in middle latitudes. The global mean CUP rate is reduced by 22% (from 2.1 mm day^{-1} to 1.6 mm day^{-1}), while the global mean LSP rate increases by 13% (from 1.5 to 1.7 mm day^{-1}).

Cumulus convection occurs much more frequently in the BK run than in the BQ run, as shown in Fig. 5. The figure also shows the cumulus anvil incidence, which is defined here as the fraction of all time steps for which cumulus clouds penetrate upward through the 500 mb level. As a global mean in the control run, cumulus convection covers 11% of the earth, while anvil clouds cover 8%. In the BK run the cumulus incidence has increased dramatically to 25% and the anvil incidence has increased to 18%. The larger cumulus and anvil incidences of the BK run indicate the more persistent cumulus convection, as might be expected with the prognostic closure, which endows the parametrized cumulus convection with a memory from one time step to the next. We therefore expect cumulus convection to be more persistent with the prognostic closure than with the QE closure.

The substantially larger anvil incidence in the BK run drastically changes the simulation of the earth's radiation budget, making it much less realistic in the BK run. Figure 6 shows the January zonal-mean absorbed solar radiation at the top of the atmosphere from the two runs and observations. The observations are taken from the 1987 data of the Earth Radiation Budget Experiment (ERBE; Ramanathan *et al.* 1989). Major differences between the two simulations occur between 40°S and 20°N. The prognostic closure reduces the absorbed solar radiation throughout the tropics, and by up to 78 W m^{-2} close to the equator.

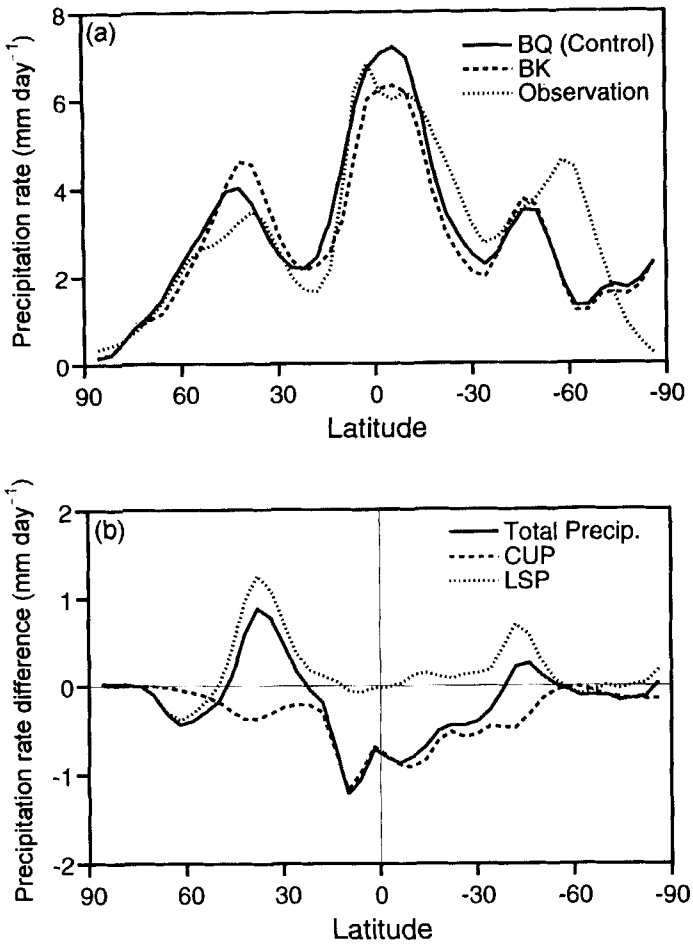

Figure 4. (a) Zonal mean of the January monthly-mean precipitation rate and (b) precipitation rate difference between run BK and BQ (see text) in components: cumulus precipitation (dashed), large-scale precipitation (dotted) and total precipitation. Three curves in (a) are observations (Legates and Willmott, dotted) and results from the BQ run (solid) and the BK run (dashed) (see text).

Figure 7 shows the changes of the zonal-mean energy budgets from run BQ to run BK. With the prognostic closure, both the outgoing long-wave radiation (OLR) and the absorbed solar radiation at the top of the atmosphere are reduced. This reduces the net gain of radiative energy at the top of the atmosphere. The flow of radiative energy into the earth's surface is reduced by an even larger amount. As a net result the prognostic closure results in a gain of radiative energy by the atmosphere. This is consistent with the reduced surface latent-heat flux, as shown in the same figure, and the reduced surface latent-heat flux is consistent with the reduced precipitation rate.

These results make it clear that a better method is needed to represent the radiative effects of the anvil clouds. We return to this issue in the next section.

Figure 8 shows the zonally averaged cumulus heating, LSP heating, total latent heating, and the total radiative cooling, as obtained in run BQ. Evaporation of detrained condensates causes the cumulus cooling above 200 mb, where LSP then recondenses and causes warming (see Randall *et al.* 1989). Meanwhile, the cumulus-induced subsidence causes warming throughout the convective layer below 200 mb. Evaporation of the falling

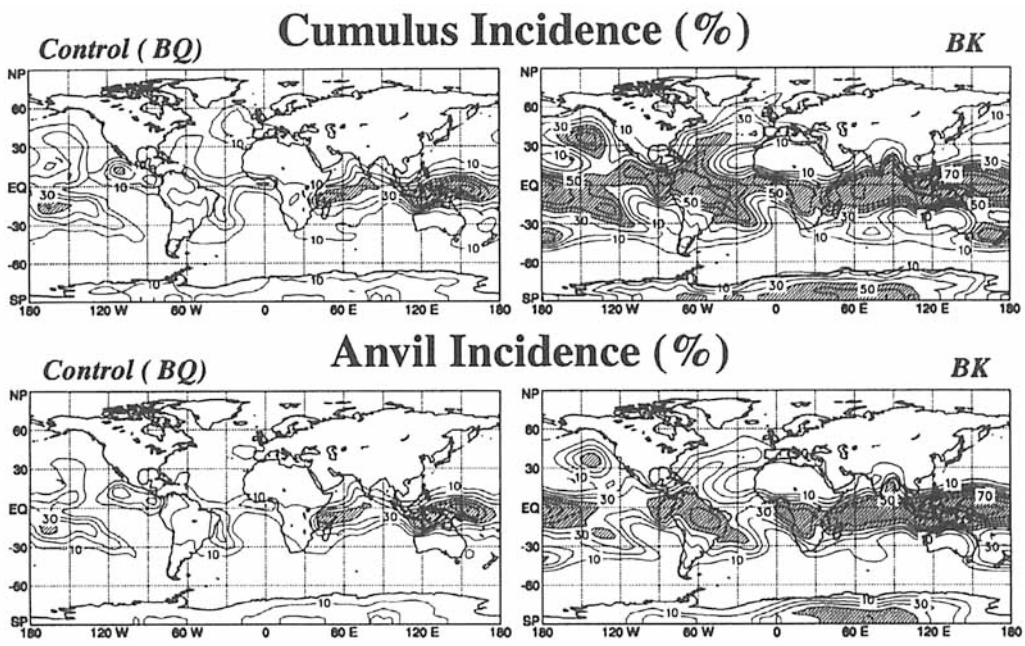

Figure 5. Monthly mean cumulus incidence (upper panels) and anvil incidence (lower panels) from the control run (left panels) and the BK run (right panels) (see text). The contour interval is 10%. Shading is used where the incidence is larger than 40%.

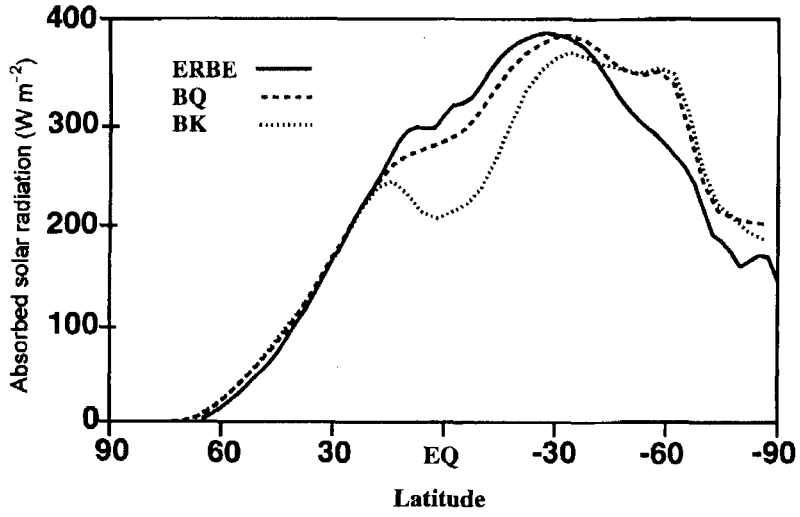

Figure 6. January zonal mean of the absorbed solar radiation at the top of the atmosphere. The three curves shown are from the BK and BQ runs (see text) and observations from the Earth Radiation Budget Experiment.

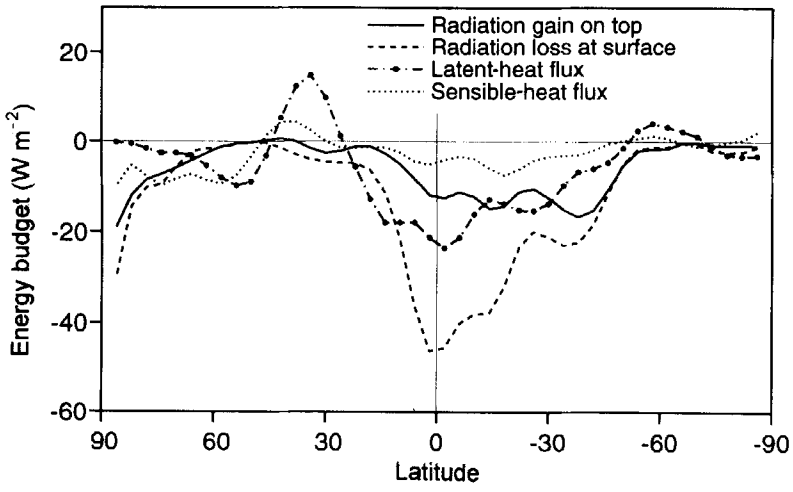

Figure 7. Differences (run BK − run BQ) in radiation gain on the top of the atmosphere, radiation loss into the surface, surface latent-heat and sensible-heat fluxes.

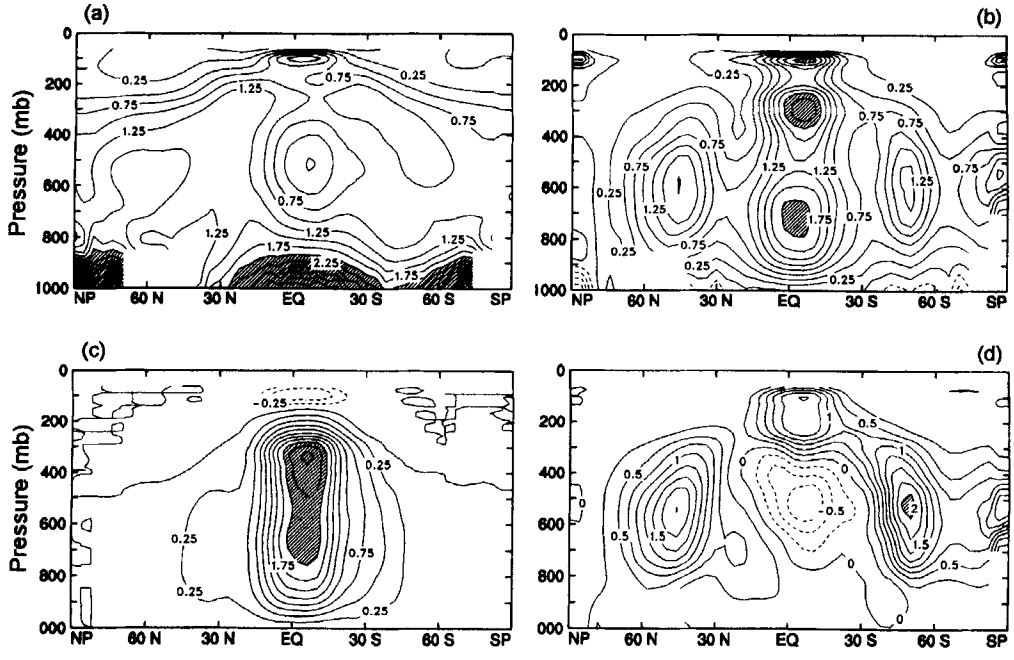

Figure 8. Zonal-mean January (a) total radiative cooling, (b) total latent heating, (c) cumulus heating, and (d) large-scale precipitation heating in the BQ run (see text). The contour interval is 0.25 K day^{-1} with negative values shown dashed. Areas larger than 2 K day^{-1} and smaller than −2 K day^{-1} are shaded.

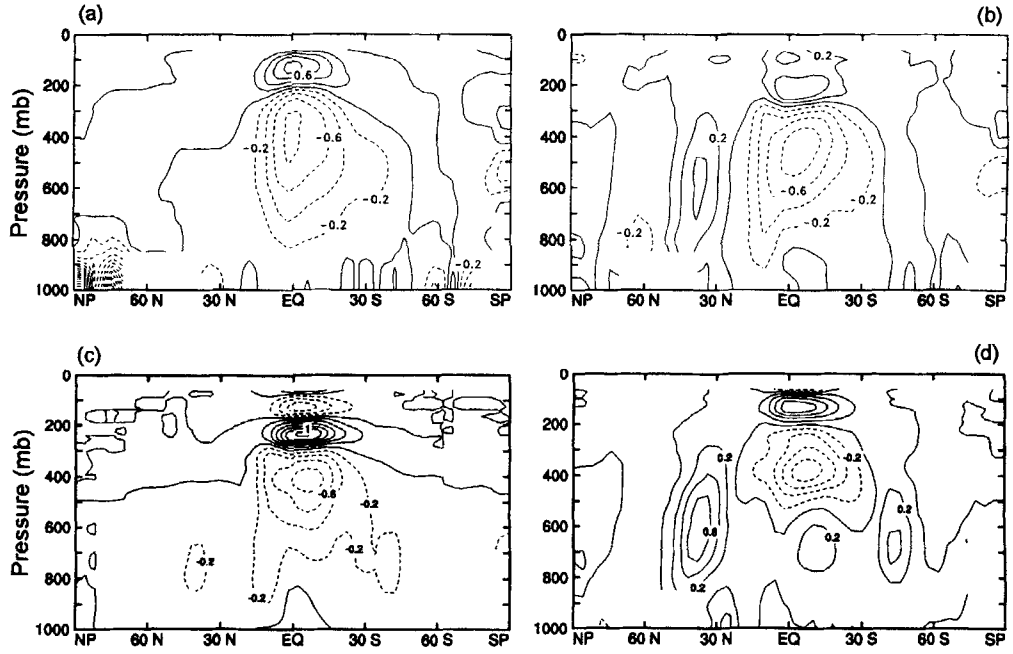

Figure 9. Diabatic heating difference (K) between the BK and BQ runs (see text). (a) Total radiative cooling, (b) total latent heating, (c) cumulus heating, and (d) large-scale precipitation heating. Negative values shown dashed.

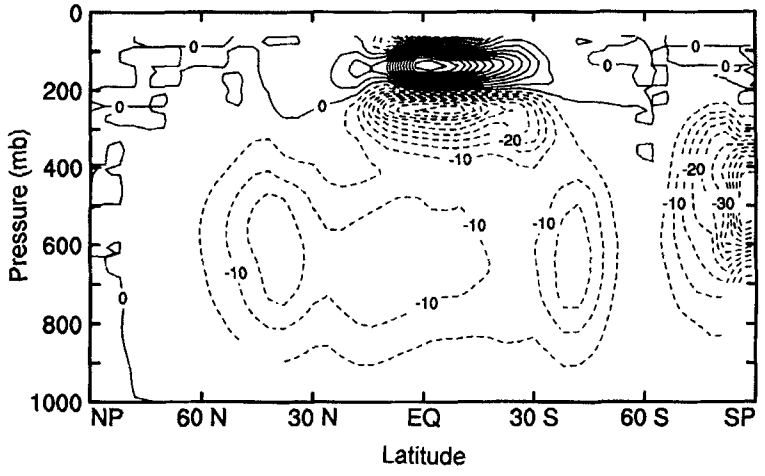

Figure 10. Difference of the cumulus detrainment mass flux between runs BK and BQ (see text). Contour interval is 0.005 h^{-1} with negative values shown dashed.

LSP leads to cooling between 400 mb and 700 mb, near the tropical anvil cloud base. LSP accounts for the two mid-latitude maxima of the latent heating, and exhibits a maximal evaporative cooling. Strong radiative cooling due to the water vapour continuum appears near the surface in the summer hemisphere. The total radiative cooling has a minimum (that is, relative heating) at 500 mb in the tropics, which corresponds to the anvil cloud base. This is due to the radiative effects of the anvil clouds, which are assumed to have their bases at 500 mb in all cases. The anvil-induced radiative heating at 500 mb represents a forcing

(destabilization) for cumuli that penetrate higher than 500 mb, and it also increases the probability of clouds detraining at that level and hence the incidence of relatively shallow cumuli.

Figure 9 shows the differences between the BK and BQ runs. The results of the BK run are qualitatively similar to those of the BQ run, but the intensities are different. The minima of radiative cooling and cumulus heating at 500 mb in the tropics both intensify. The radiative cooling minimum at 500 mb in the BQ run actually becomes a radiative warming in the BK run. We can see in Fig. 9 that both the total latent heating and the total radiative cooling generally decrease from runs BQ and BK. The decreased radiative cooling (actually increased warming) at the anvil cloud base in the BK run is associated with the increased anvil incidence, mentioned earlier. This further suppresses shallow convection. The increased cumulus incidence from run BQ and run BK is thus mostly associated with deep convection. The difference in the detrainment mass flux between runs BK and BQ is shown as a function of latitude and pressure in Fig. 10. In the BK run the cumulus clouds detrain less mass almost everywhere except near the tropical tropopause above 200 mb. The increased detrainment mass flux near the tropical tropopause in the BK run corresponds to the stronger detrainment cooling there, which in turn leads to stronger large-scale condensation, and to stronger evaporative cooling between 300 and 500 mb. In summary, as we pass from run BQ to run BK, the mass flux distribution function shifts so that we obtain more deep convection and less shallow convection. The total cumulus mass flux at the PBL top is reduced, but most of the mass convectively drawn from the PBL into the free atmosphere travels to almost the tropopause.

Figure 11 shows differences between runs BK and BQ (i.e. BK minus BQ) for monthly mean and zonally averaged temperature, relative humidity, and water vapour mixing ratio. The prognostic closure results in a warming of up to 3.5 K at 500 mb and a slight cooling close to the tropopause. This coincides with the smaller radiative cooling there even though the latent heating is also weaker. As discussed later, these results are sensitive to the value of α used. Because of the warming at 500 mb the static stability above this level is weaker, and so cumulus-induced subsidence causes less warming. We can say that the anvil-induced 500 mb warming of the BK run is balanced by weaker 500 mb cumulus warming. The warmer temperatures are associated with decreased relative humidities, but the change of the zonal-mean mixing ratio is relatively small, and so is the change of the cumulus drying rate (not shown).

The computer time used by the cumulus parametrization in the BK run is cut in half, relative to run BQ. Most of the savings is due to the elimination of the kernel calculation. The most expensive remaining portion of the parametrization is the determination of the fractional entrainment rate. This cost can be drastically cut by using a linear mass flux profile, as discussed by Ding and Randall (personal communication).

6. Fractional anvils

In order to improve the simulation of the earth's radiation budget with the more frequently occurring deep convection as simulated with the prognostic closure, we abandoned the binary anvil parametrization used in the BQ and BK runs. Here we introduce a simple (and temporary) 'fractional anvil' parametrization. In this parametrization when convection penetrates above the 500 mb level the fractional area covered by the anvils is assumed to be proportional to the total convective cloud-base mass flux, i.e.

$$f = \mathrm{Min}\{aM_{\mathrm{B}}, 1\}, \tag{32}$$

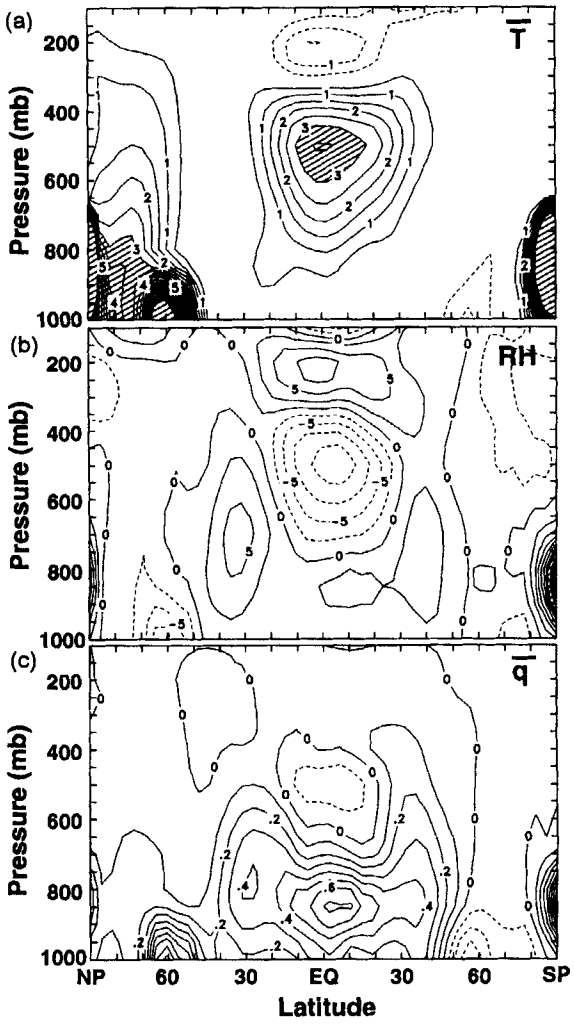

Figure 11. Difference between runs BK and BQ (see text) for monthly mean and zonally averaged (a) temperature, (b) relative humidity, and (c) mixing ratio. The contour intervals are 0.5 K, 2.5%, and 0.1 g kg^{-1}, respectively. Hatched areas are where temperature difference is more than 3 K. Negative values are shown dashed.

where a is a constant. The short-wave optical depth was then calculated using

$$\tau_{sw} = f\tau_{anv} + (1 - f)\tau_{cs}, \tag{33}$$

where τ_{anv} and τ_{cs} are the optical depths with and without anvils, respectively. The effective area-averaged long-wave emittance of the anvils was assumed to be given by

$$\tilde{\varepsilon} = f\tilde{\varepsilon}_0. \tag{34}$$

We chose $a = 75$ m^{2}s kg^{-1}, which makes the global-mean planetary albedo close to that observed.

The FK8 run (actually an ensemble of five two-month December–January runs) described below is identical to the BK run, except that the fractional anvil parametrization was used. Here the '8' denotes $\alpha = 10^8$ m^{4}kg^{-1}; other values of α are considered in the next section.

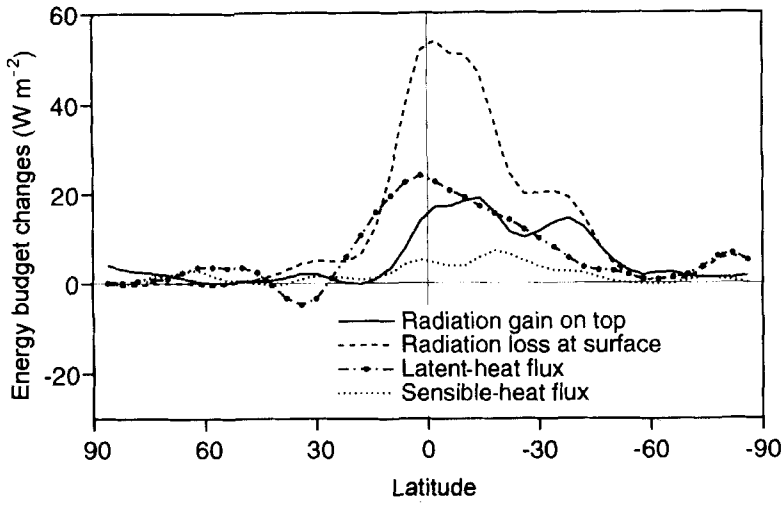

Figure 12. The zonal-mean energy-budget changes caused by the introduction of fractional anvils. The solid line shows the change in the net radiation at the top of the atmosphere, positive into the atmosphere. The dashed line shows the change in the net radiation at the surface, positive into the surface. The dashed line with dots along it shows the change in the surface latent-heat flux, positive upward, and the dotted line shows the change in the surface sensible-heat flux, again positive upward.

Figure 12 shows the zonal-mean energy-budget changes caused by the introduction of fractional anvils (i.e. as we pass from run BK to run FK8). In most of the tropics much more solar radiation (up to 60 W m^{-2}) is absorbed by the earth–atmosphere system. The OLR increases where the absorbed solar radiation increases, but not as much. As a result there is a net gain of radiative energy at the top of the atmosphere, from 10°N to 60°S. Also shown is the difference in the net radiation flux at the surface. As might be expected, most of the increased solar radiation absorbed by the earth–atmosphere system is absorbed at the surface. The balance of the land-surface energy budget is preserved mainly by an increase in the rate of surface evaporation, especially in the Amazon basin and South Africa (not shown). As discussed later (see Fig. 15), this leads to particularly strong increases of the precipitation rate in those locations, and represents an improvement in the realism of the simulated precipitation distribution.

Figure 13 shows the temperature and mixing ratio differences between the FK8 and BK runs. The fractional anvils eliminate most of the warming around the level of the anvil cloud bases, and reduce the cooling above the 400 mb level in the tropics, but have little effect on the zonally averaged mixing ratio.

7. Sensitivity to α and τ_D

In this section we present and compare January simulations obtained with the prognostic closure and fractional anvils, and with α equal to 10^8 m^{4}kg^{-1}, 10^9 m^{4}kg^{-1} and 10^7 m^{4}kg^{-1}, in runs FK8, FK9, and FK7, respectively. (Each of these cases is actually an ensemble of five two-month December–January runs.) All of these runs use $\tau_D = 600$ s. We also present the results of a test with $\tau_D = 1200$ s.

Figure 14 shows maps of the simulated January-mean precipitation rates obtained with the three different α's. The precipitation distributions are generally very similar between runs FK7 and FK8. In contrast, run FK9 shows much weaker precipitation in the tropical central and western Pacific, and over the Indian Ocean. The maximum over the Amazon

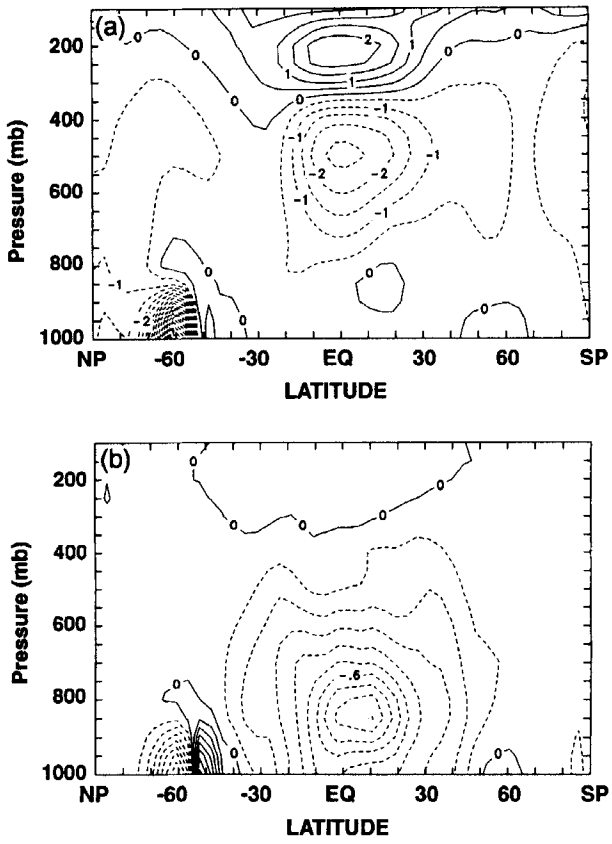

Figure 13. Differences of (a) temperature (K) and (b) mixing ratio (g kg^{-1}) between runs FK8 and BK (see text). The contour intervals are the same as in Fig. 11.

basin is most prominent in run FK7, is present in run FK8, and is almost absent in run FK9. Figure 15 shows the CUP. There is a general and almost uniform decrease in the CUP as α is increased from 10^7 m^4kg^{-1} to 10^9 m^4kg^{-1}. Associated with this is a nearly compensating increase in LSP (not shown).

Figure 16 shows the January zonal-mean differences between runs FK9 and FK8, for the precipitation rate, cumulus and anvil incidences, surface sensible-heat flux and evaporation rate. As α increases, the precipitation rate decreases in the tropics but increases in mid-latitudes. The decrease of the cumulus precipitation accounts for the total precipitation decrease in the tropics, while the mid-latitude precipitation increase comes from LSP. These changes are similar to those from the BQ run to the BK run (refer back to Fig. 4). Surface evaporation also decreases, slightly, at almost all latitudes. Although the CUP decreases, cumulus convection actually occurs much more often, up to about 20% more often in the tropics. The change of anvil incidence is much less than the change in cumulus incidence. This means that most of the additional convection is shallow. Shallow cumulus clouds tend to increase the precipitable water by detraining and moistening the environment, without precipitating much. The more humid environment allows falling LSP to reach the surface more easily. On the other hand, the moistening and cooling due to the evaporation of falling stratiform precipitation promotes shallow convection, so there is a synergism at work here.

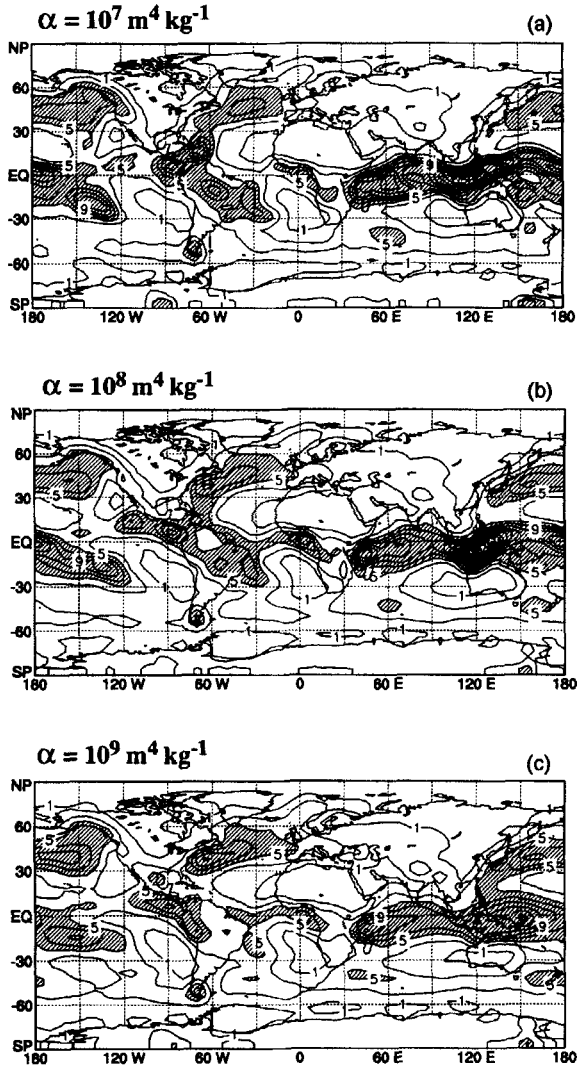

Figure 14. Simulated January-mean precipitation rate (mm day^{-1}) with runs (a) FK7, (b) FK8, and (c) FK9 (see text). Contour interval and shading are the same as in Fig. 3.

Figure 17 shows the stream function of the mean meridional circulation (MMC) for runs BQ, BK, FK7, FK8 and FK9, and the European Centre for Medium-Range Weather Forecasts (ECMWF) analyses. A negative stream function means that a cell is running 'clockwise', i.e. rising in the north and sinking in the south. The Hadley cells obtained in runs BK and FK8 are stronger than the ECMWF analysis and stronger than in the control run. The similar strength of the MMC in the BK and FK8 runs suggests that the intensification of the MMC cannot be directly explained by the use of the fractional anvil or the prognostic closure. Run FK9 reduces the intensity of the Hadley cell by about 10%, and brings it into closer agreement with the ECMWF analysis. It should also be kept in mind that observational estimates of the MMC can vary by a factor of two or more. The MMCs obtained with both the BK and FK8 runs are about as intense as those shown in the 1988–92 ECMWF analyses.

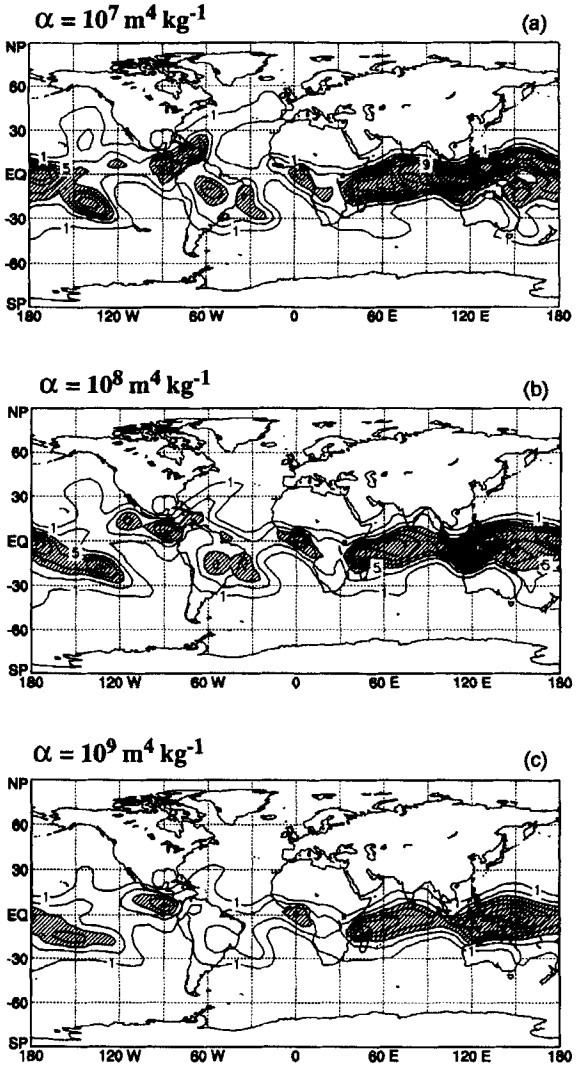

Figure 15. As Fig. 14 but for the simulated January cumulus precipitation rate.

Figure 18 shows the zonally averaged temperature and relative humidity differences between each of the simulations BQ, FK8, and FK9, and the 1985–93 mean ECMWF analyses. The current model, with its top at 50 mb and only two layers above 100 mb, cannot simulate physical processes of the middle-to-high-latitude stratosphere, so the large temperature differences at those upper levels should be ignored. The ECMWF-analysed relative humidity above 400 mb is unreliable. Run BQ produces a warmer-than-observed tropical troposphere. The zonal mean temperature with run FK8 is actually less realistic. Run FK9 produces better results. The simulated tropical troposphere is generally drier than the ECMWF analysis, with not much difference between the simulations (the mixing ratios are similar and are not shown here). With the larger α, the tropical troposphere is (realistically) cooler. In this respect, $\alpha = 10^9$ m^{4}kg^{-1} is a better choice than $\alpha = 10^8$ m^{4}kg^{-1}. Run FK7 produces an even warmer tropical troposphere (not shown).

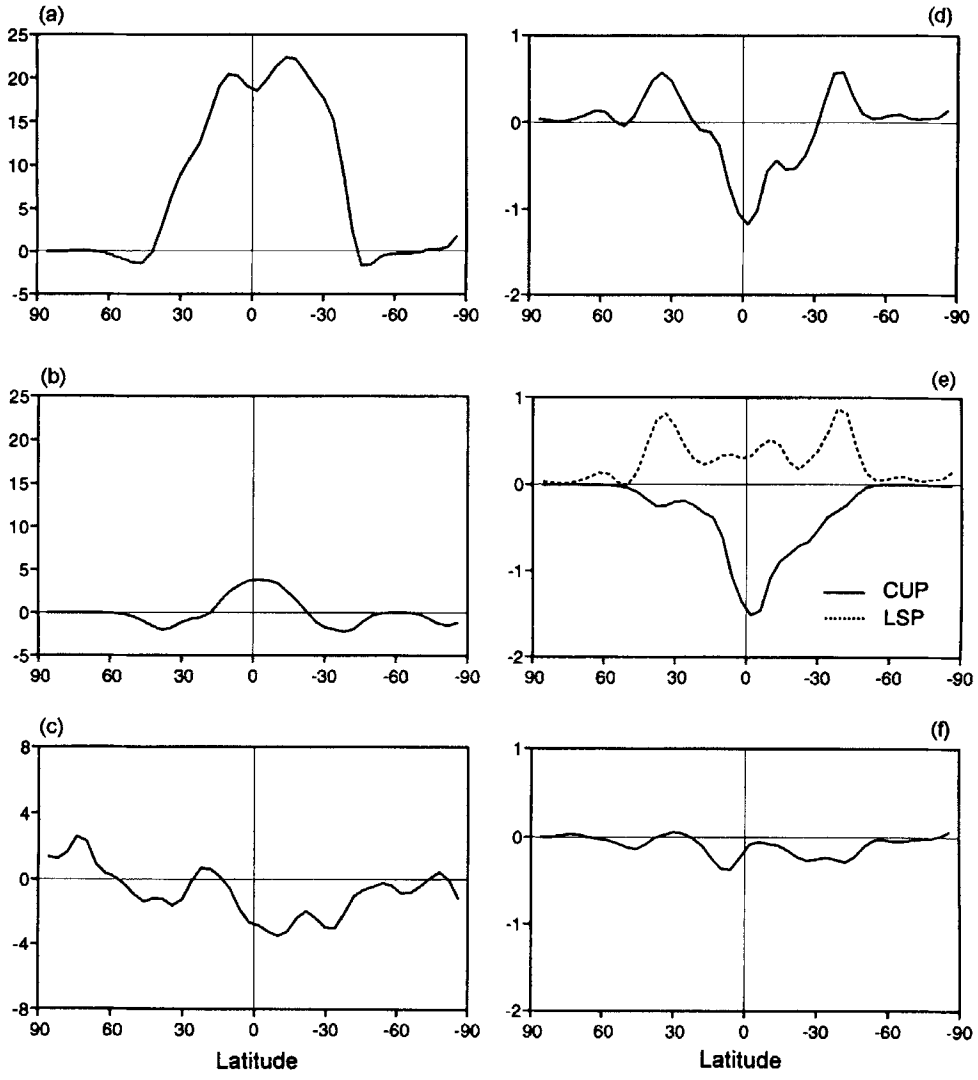

Figure 16. January zonal-mean differences between runs FK9 and FK8 (see text) for: (a) cumulus incidence (%), (b) anvil incidence (%), (c) surface sensible-heat flux (W m^{-2}), (d) total precipitation rate (mm day^{-1}), (e) cumulus (CUP), and large-scale precipitation (LSP) rates (mm day^{-1}) and (f) surface evaporation rate (mm day^{-1}).

We close this section by briefly describing the results of a test in which $\alpha = 10^8\ m^4 kg^{-1}$ and τ_D is increased to 1200 s; this run, FKT, can be compared with run FK8. As shown in Fig. 19, the effect of increasing τ_D is to warm the tropical troposphere and cool the middle latitudes. The results with $\tau_D = 600$ s are more realistic than those with $\tau_D = 1200$ s.

8. A FURTHER SENSITIVITY TEST

Cheng and Arakawa (1990) discussed several alternative definitions of the cloud work function. The first, which they called A_1, is the same as the cloud work function introduced by AS, and is the one used in the present study. It includes the thermal effects on buoyancy in updraughts, and the water-vapour correction associated with the virtual-temperature effect. A refined definition of the cloud work function can, however, include, for example,

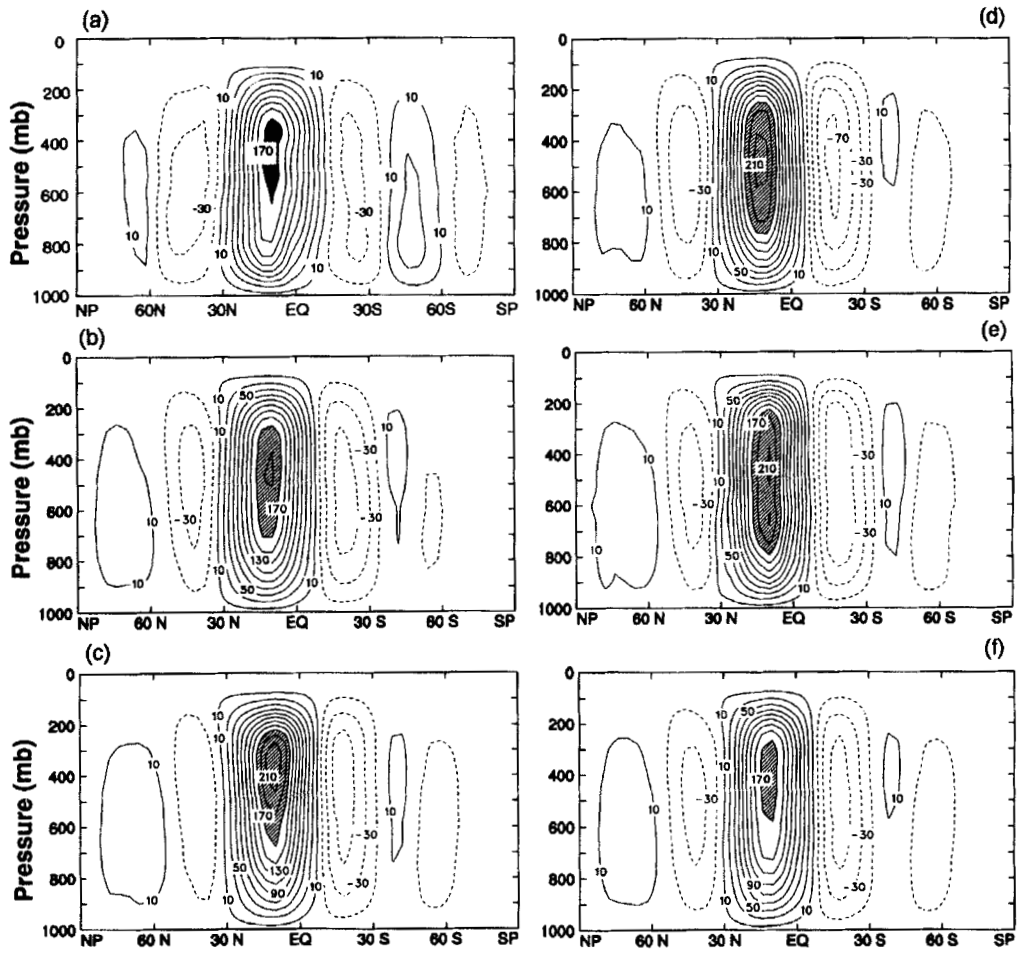

Figure 17. Stream function of the mean meridional circulation from (a) the ECMWF observations, and simulations (b) BQ, (c) BK, (d) FK7, (e) FK8, and (f) FK9 (see text). The contour interval is 20×10^9 kg s^{-1} with negative values shown dashed. Regions with values larger than 170×10^9 kg s^{-1} are hatched.

the effects of liquid water and ice loading in the updraughts. These effects are included in A_2. Finally, A_3 includes kinetic energy generation by convective downdraughts. Cheng and Arakawa showed that these three versions of the cloud work function vary with time in proportion to each other, at least for the cases that they examined. They concluded that it may not be important to distinguish between them.

Here we focus on the difference between A_1 and A_2. Liquid-water and ice-water loading reduce the net buoyancy, so that $A_2 < A_1$ for a given sounding. We therefore propose the following hypothesis: If we run the GCM using A_1, and then suddenly replace A_1 by A_2, the instantaneous effect is to decrease the buoyancy available to convective updraughts. As a result, convective activity will slow down. This convective 'time out' will allow the atmosphere to cool (due to the reduced rate of latent-heat release) and to moisten (due to the reduced precipitation rate). Such a cooling and moistening tends to increase the cloud work function, so that after some time convection will become re-established, but with a cooler and more humid mean state.

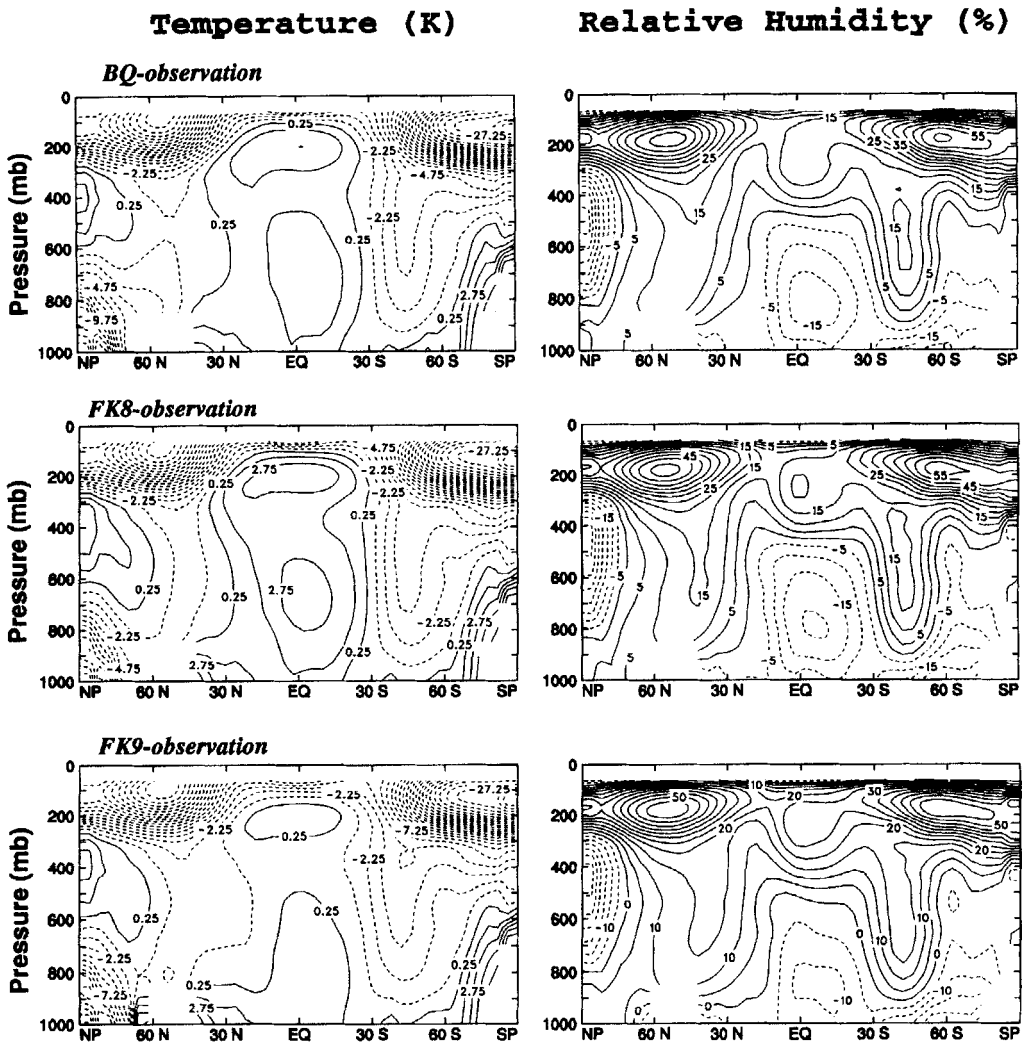

Figure 18. Difference plots showing (top) BQ − observations; (centre) FK8 − observations, and (bottom) FK9 − observations (see text) for the zonally-averaged monthly-mean temperature (left) and relative humidity (right). The contour intervals are 1.25 K for temperature and 5% for relative humidity with negative values shown dashed.

To explore this hypothesis in a simplified way, we performed the following idealized numerical experiment. We performed a run just like FK8, except that we arbitrarily replaced (19) by

$$\frac{\mathrm{d}}{\mathrm{d}t}K(\lambda) = M_{\mathrm{B}}(\lambda)\frac{A(\lambda)}{10} - \frac{K(\lambda)}{\tau_{\mathrm{D}}(\lambda)}, \tag{35}$$

i.e. we arbitrarily divided the cloud work function by 10 before using it to evaluate the buoyancy term of the CKE equation. Such an arbitrary reduction of A can be considered as a crude and exaggerated analogue of the effects of introducing liquid-water and ice-water loading.

This seemingly drastic alteration of our model has only rather minor effects on the distribution of precipitation, as shown in Figs. 20 and 21. There is a slight reduction of the precipitation rate, on the order of 10%, in response to the ten-fold reduction in the

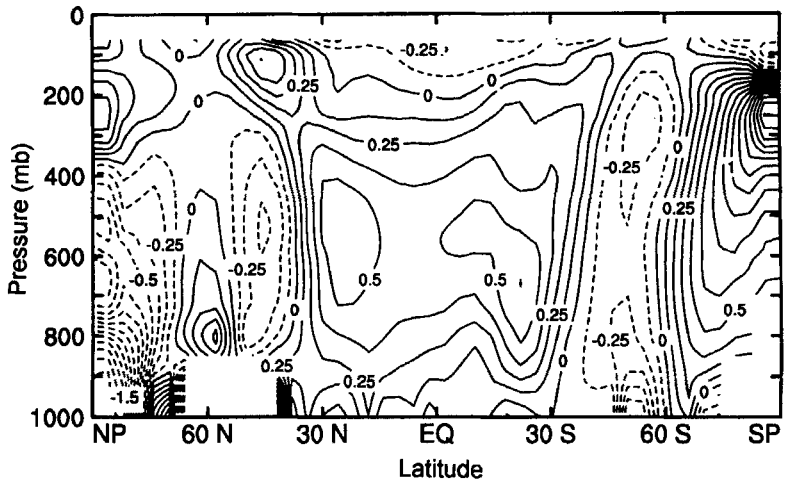

Figure 19. The zonally averaged temperature difference (K) between runs FKT and FK8 (see text). Negative values are shown dashed.

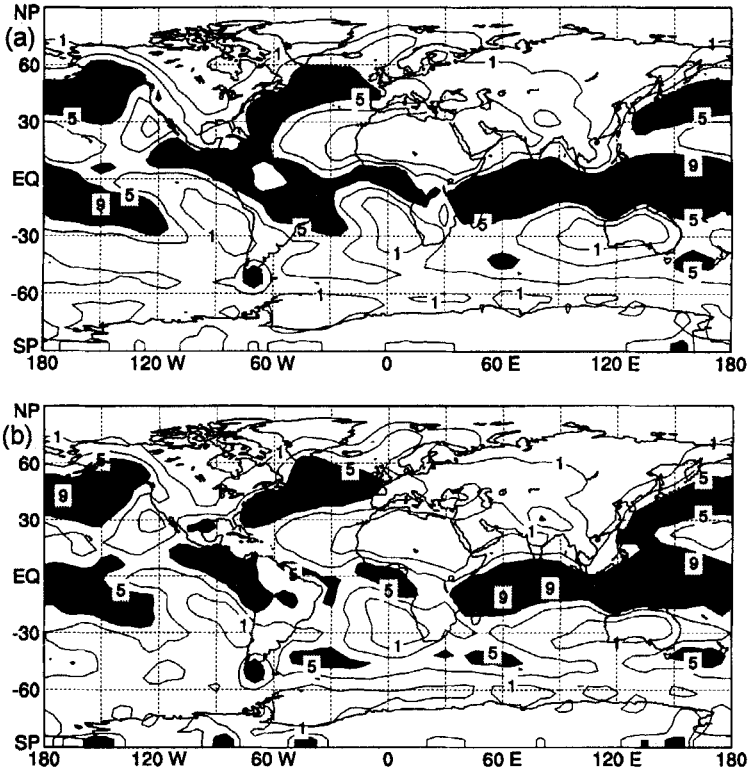

Figure 20. January total precipitation (mm day^{-1}) for (a) run FK8 and (b) the experiment in which (35) is used in place of (19). Shading above 5 mm day^{-1}.

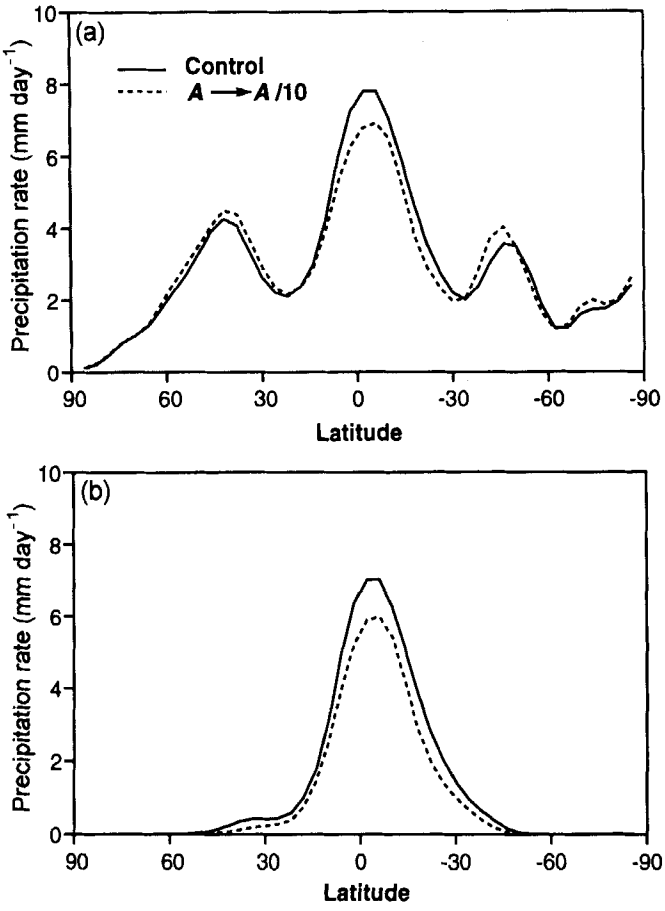

Figure 21. Zonal means of the results shown in Fig. 20; (a) total precipitation rate and (b) cumulus precipitation rate.

effective value of the cloud work function, and the geographical pattern of the precipitation changes very little. In contrast, however, there are large changes in the mean distributions of temperature and moisture, as shown in Fig. 22. In line with the earlier discussion, the temperature has cooled, by as much as 3.5 K in the upper troposphere, and the water vapour mixing ratio has increased, by up to 2 g kg^{-1}; the combined effects of these two changes imply a very large increase of the relative humidity throughout most of the troposphere, as shown in the bottom panel of Fig. 22.

We draw three conclusions from this simple experiment. First, the pattern of precipitation is not sensitive to the actual value of the CAPE, as measured by the cloud work function. According to our interpretation, this is consistent with the QE hypothesis of AS. Second, the structure of the mean sounding is in fact sensitive to the CAPE, a point perhaps overlooked and certainly not stressed by AS. Finally, our results provide motivation for further research to formulate and test a more refined definition of the cloud work function.

9. Summary and discussion

Arakawa and Schubert (1974) assumed a QE between cumulus convection and the 'large-scale forcing'. It is not clear, however, that the distinction between the large-scale

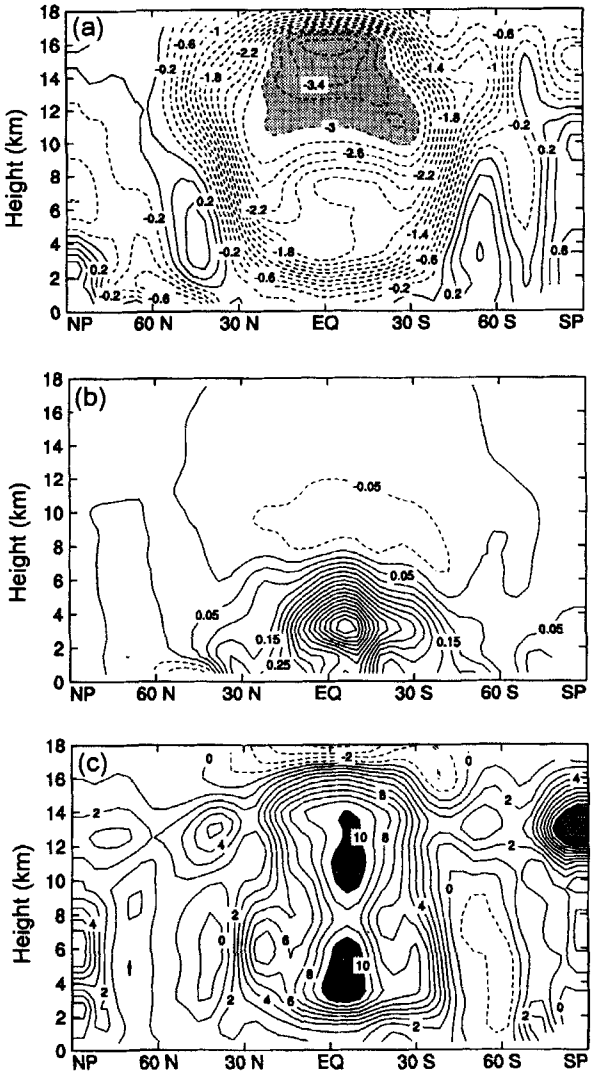

Figure 22. Differences in (a) temperature (K), (b) water vapour mixing ratio (g kg^{-1}), and (c) relative humidity (%) between the experiment in which (35) was used and the control run (FK8) in which (19) was used instead. Negative values are shown dashed. The experiment is cooler and more humid than the control.

forcing and the convective response can be uniquely drawn. Partly for this reason we have relaxed the QE closure by explicitly predicting the cumulus kinetic energy for each cumulus subensemble. The prognostic closure, which bypasses the explicit definitions of the kernel and the large-scale forcing, sidesteps the ambiguity involved in the distinction between the cumulus response and the large-scale forcing. This represents a step towards improving the coupling among the model's physical parametrizations.

The prognostic closure involves a parameter, α, which relates the cumulus kinetic energy to the cloud-base mass flux. We showed that the 'adjustment time' is related to α. When α approaches zero, the prognostic closure gives QE. According to the definitions of CKE and cloud-base mass flux, α should be a function of both cloud depth and ambient wind shear. As a preliminary test of the prognostic closure we have used a single, constant

value of α for all cloud types in all of our simulations. The results appear to be acceptable. Since we have not proposed a theory to determine α, we do not claim that we can actually predict the observed values of the CKE or the cloud work function. We do claim, however, that we can simulate physically meaningful values of M_B.

The use of a crude 'binary anvil' parametrization with the prognostic closure gives unrealistic radiation budgets. This forced us to reconsider the representation of anvil clouds. A revisited 'fractional anvil' parametrization gives both an improved radiation budget and a more realistic precipitation distribution. This illuminates the importance of the coupling between different components of the model physics.

Our sensitivity tests show that a larger α leads to more shallow cumulus convection and less cumulus precipitation. The more frequent shallow convection moistens the middle troposphere and enables more large-scale precipitation to reach the surface. Larger values of α give more realistic temperatures in the tropical troposphere. Our results should also depend on the way we parametrize the large-scale condensation process, however, and so we conclude that the convection parametrization really cannot be evaluated independently of the stratiform cloud parametrization with which it interacts.

The simplicity of the prognostic closure makes including shear effects in the CKE equation relatively straightforward. Clearly, the effects of shear should be incorporated into future cumulus parametrizations. In the context of the present parametrization, this will entail (at least) two steps: including the effects of vertical wind shear in the CKE equation, and allowing α to depend on the shear. A convective Richardson number that measures the relative size of buoyancy and shear effects may be a useful parameter for this purpose. This is an interesting project for the future.

ACKNOWLEDGEMENTS

D. Dazlich of CSU provided valuable assistance with the GCM. We are grateful to Dr K. Xu for permitting us to use his CEM results. This research has been supported by the National Science Foundation under grant ATM-9121629, by the Department of Energy under grant DE-FG-03-94ER61929, and by subcontract No. 0965-G-5B387 from UCLA. Computing resources have been provided by the Scientific Computing Division at the National Center for Atmospheric Research, and by the National Energy Research Computer Center at the Lawrence Livermore National Laboratory.

REFERENCES

Arakawa, A. 1969 'Parameterization of cumulus convection'. Proceedings of the WMO/IUGG Symposium on numerical weather prediction, Tokyo, 26 November–4 December, 1968. Japan Meteorol. Agency, IV, 8, 1–6

Arakawa, A. and Chen, J.-M. 1987 'Closure assumptions in the cumulus parameterization problem'. Pp. 107–130 in WMO/IUGG Symposium on short- and medium-range numerical weather prediction, Tokyo

Arakawa, A. and Lamb, V. R. 1977 Computational design of the basic dynamical processes of the UCLA general circulation model. *Methods Comput. Phys.*, **17,** 173–265 Academic Press, New York

1981 A potential enstrophy and energy conserving scheme for the shallow-water equations. *Mon. Weather Rev.*, **109,** 18–36

Arakawa, A. and Schubert, W. H. 1974 Interaction of a cumulus cloud ensemble with the large-scale environment, Part I. *J. Atmos. Sci.*, **31,** 674–701

Arakawa, A. and Suarez, M. J. 1983 Vertical differencing of the primitive equations in sigma coordinates. *Mon. Weather Rev.*, **111,** 34–45

Arakawa, A. and Xu, K.-M. 1990 'The macroscopic behavior of simulated cumulus convection and semiprognostic tests of the Arakawa–Schubert cumulus parameterization'. Proceedings of the Indo-US seminar on parameterization of sub-grid scale processes in dynamical models of medium-range prediction and global climate. Pune, India

Asai, T. 1964 Cumulus convection in the atmosphere with vertical wind shear: numerical experiment. *J. Meteorol. Soc. Japan*, **42**, 245–259

Chandrasekhar, S. 1961 *Hydrodynamic and hydromagnetic stability*. Clarendon Press, Oxford, UK

Cheng, M.-D. and Arakawa, A. 1990 'Inclusion of convective downdrafts in the Arakawa–Schubert cumulus parameterization'. Tech. Rep., Dept. of Atmospheric Sciences, UCLA, USA

Gates, W. L. 1992 AMIP: The Atmospheric Model Intercomparison Project. *Bull. Am. Meteorol. Soc.*, **73**, 1962–1970

Grell, G. A., Kuo, Y.-H. and Pasch, R. J. 1991 Semiprognostic tests of cumulus parameterization schemes in the middle latitudes. *J. Atmos. Sci.*, **119**, 5–31

Hack, J. J., Schubert, W. H. and Silva Dias, P. L. 1984 A spectral cumulus parameterization for use in numerical models of the tropical atmosphere. *Mon. Weather Rev.*, **112**, 704–716

Harshvardhan, Randall, D. A., Corsetti, T. G. and Dazlich, D. A. 1989 Earth radiation budget and cloudiness simulations with a general circulation model. *J. Atmos. Sci.*, **46**, 1922–1942

Jabouille, P., Redelsperger, J. L. and Lafore, J. P. 1996 Modification of surface fluxes by atmospheric convection in the TOGA COARE region. *Mon. Weather Rev.*, **124**, 816–837

Johnson, R. H. 1976 The role of convective-scale precipitation downdrafts in cumulus and synoptic-scale interaction. *J. Atmos. Sci.*, **33**, 1890–1910

Kao, C.-Y. J. and Ogura, Y. 1987 Response of cumulus clouds to large-scale forcing using the Arakawa–Schubert cumulus parameterization. *J. Atmos. Sci.*, **44**, 2437–2458

Krueger, S. K. 1988 Numerical simulation of tropical cumulus clouds and their interaction with the subcloud layer. *J. Atmos. Sci.*, **45**, 2221–2250

Legates, D. R. and Willmott, C. J. 1990 Mean seasonal and spatial variability in gauge-corrected global precipitation. *Int. J. Climatology*, **10**, 111–127

Lilly, D. K. and Jewett, B. F. 1990 Momentum and kinetic energy budgets of simulated supercell thunderstorms. *J. Atmos. Sci.*, **47**, 707–726

Lin, C.-C. 1994 'Development of an improved cloud model for use in cumulus parameterization'. Ph.D. dissertation. University of California, Los Angeles

Lord, S. J. 1982 Interaction of a cumulus cloud ensemble with the large-scale environment, Part III: Semi-prognostic test of the Arakawa–Schubert cumulus parameterization. *J. Atmos. Sci.*, **39**, 88–103

Lord, S. J. and Arakawa, A. 1980 Interaction of a cumulus cloud ensemble with the large-scale environment, Part II. *J. Atmos. Sci.*, **37**, 2677–2692

Lord, S. J., Chao, W. C. and Arakawa, A. 1982 Interaction of a cumulus cloud ensemble with the large-scale environment, Part IV: The discrete model. *J. Atmos. Sci.*, **39**, 104–113

Manabe, S., Smagorinsky, J. and Strickler, R. F. 1965 Simulated climatology of a general circulation model with a hydrologic cycle. *Mon. Weather Rev.*, **93**, 769–797

Pan, D.-M. 1995 'Development and application of a prognostic cumulus parameterization'. Ph.D. thesis, Colorado State University

Ramanathan, V., Cess, R. D., Harrison, E. F., Minnis, P., Barkstrom, B. R., Ahmad, E. and Hartmann, D. 1989 Cloud-radiative forcing and climate: Results from the Earth Radiation Budget Experiment. *Science*, **243**, 57–63

Randall, D. A. and Pan, D.-M. 1993 Implementation of the Arakawa–Schubert cumulus parameterization with a prognostic closure. Pp. 137–144 in *Cumulus parameterization*, a Meteorological Monograph. Eds. K. Emanuel and D. Raymond. American Meteorological Society

Randall, D. A., Abeles, J. A. and Corsetti, T. G. 1985 Seasonal simulations of the planetary boundary layer and boundary-layer stratocumulus clouds with a general circulation model. *J. Atmos. Sci.*, **42**, 641–676

Randall, D. A., Harshvardhan, Dazlich, D. A. and Corsetti, T. G. 1989 Interactions among radiation, convection, and large scale dynamics in a general circulation model. *J. Atmos. Sci.*, **46**, 1943–1970

Randall, D. A., Harshvardhan and Dazlich, D. A. 1991 Diurnal variability of the hydrologic cycle in a general circulation model. *J. Atmos. Sci.*, **48,** 40–62

Raymond, D. J. and Blyth, A. M. 1986 A stochastic mixing model for non-precipitating cumulus clouds. *J. Atmos Sci.*, **43,** 2708–2718

Rutledge, S. A. and Houze Jr, R. A. 1987 A diagnostic modeling study of the trailing stratiform region of a midlatitude squall line. *J. Atmos. Sci.*, **44,** 2640–2656

Seitter, K. L. and Kuo, H.-L. 1983 The dynamical structure of squall-line type thunderstorms. *J. Atmos. Sci.*, **40,** 2831–2854

Stull, R. B. 1988 *An introduction to boundary layer meteorology*. Kluwer Academic Publ., Dordrecht, The Netherlands

Suarez, M., Arakawa, A. and Randall, D. A. 1983 Parameterization of the planetary boundary layer in the UCLA general circulation model: Formulation and results. *Mon. Weather Rev.*, **111,** 2224–2243

Warner, J. 1970 On steady state one-dimensional models of cumulus convection. *J. Atmos. Sci.*, **27,** 1035–1040

Wu, X. and Moncrieff, M. W. 1996 Collective effects of organized convection and their approximation in general circulation models. *J. Atmos. Sci.*, **53,** 1477–1495

Xie, P. and Arkin, P. A. 1996 Analyses of global monthly precipitation using gauge observations, satellite estimates and numerical model predictions. *J. Climate*, **9,** 840–858

Xu, K.-M. 1991 'The coupling of cumulus convection with large-scale processes'. Ph.D. dissertation, University of California, Los Angeles

Xu, K.-M. and Arakawa, A. 1992 Semiprognostic tests of the Arakawa–Schubert cumulus parameterization using simulated data. *J. Atmos. Sci.*, **49,** 2421–2436

Xu, K.-M., Arakawa, A. and Krueger, S. K. 1992 The macroscopic behavior of cumulus ensembles simulated by a cumulus ensemble model. *J. Atmos. Sci.*, **49,** 2402–2420

Yanai, M., Esbensen, S. and Chu, J.-H. 1973 Determination of bulk properties of tropical cloud clusters from large-scale heat and moisture budgets. *J. Atmos. Sci.*, **30,** 611–627